The origin, genetics and dispersal of drug-resistant *Plasmodium falciparum* in Kenya

Jonathan Maina Mwangi

A thesis submitted to the University of Glasgow for the degree of Doctor of Philosophy

Division of Infection and Immunity

Institute of Biomedical and Life Sciences

University of Glasgow

Glasgow, United Kingdom.

"You will become suddenly, and possibly irrevocably, incredibly boring. People ask, "So, what are you up to these days?" and you say, "Thesis." That's about all you can handle because inexplicably, your facility for the spoken English language has deteriorated. You have stopped seeing people, going places, and doing things. You live entirely in your own head and in the computer lab, if you're as lucky as me. Your life is not your own. You struggle with monumental questions of rhetoric and logic, about minutia. No one likes you anymore. You are paranoid."

'Deconstructing A Nervous Breakdown' *The Declaration*, 29th April, 1999

"The chessboard is the world; the pieces are the phenomena of the universe; the rules of the game are what we call the Laws of Nature. The player on the other side is hidden from us. We know that his play is always fair, and patient. But also we know, to our cost, that he never overlooks a mistake, or makes the smallest allowance for ignorance."

Thomas Henry Huxley (1825-95)

Table of Contents

Table of Figures

List of Tables

Acknowledgements

I would like to thank God for His love and providence in my life.

I would like to express my gratitude to:

The International Atomic Energy Agency (IAEA), University of Glasgow and the Kenya Medical Research Insitute for my scholarships.

My parents Mrs. Mary Mwangi and my late dad Mr. Johnson Mwangi whose enviable skills in my upbringing, love and support contributed to my getting here. Thanks mum for all your prayers.

My co- supervisor and boss at KEMRI, Dr. Sabah Omar for the opportunities, support, friendship and for believing in me.

My immediate Director at KEMRI, Dr. Gerald Mkoji for encouragement and support. The Director KEMRI for the chance to continue my studies.

Dr. 'Sir' Luke Baton for getting me hooked on the 'Metro' crossword, 'Pharaoh's bubbles', for the company and for generally being the young slyboots.

The Mary Calderites especially my three good friends Liz, Fiona and Keith who provided support with all those reagents, lab equipment and also for the good laugh.

My two assessors, Dr. Jonathan Wastling and Dr. Marian Bruce (Thanks for introducing me to Stirling, the friendship and the helpful discussions, Marian).

My very special thanks go to my supervisor at Glasgow University, Dr. Lisa Ranford-Cartwright, for her indefatigable support; her fights for me to get scholarships even when I didn't think I stood a chance; introducing me to South America and tirelessly correcting my 'student waffle'. And of course, the Karaoke with 'Dancing Queen' as the opener!

My brother, David, my sisters Anne and Prisca together and my brother-in-law, Joseph 'John' Muiruri. Thanks for being there (I didn't know what siblings were for till I met you!).

Dr. Joe Mwatha (thanks for the discourse, intellectual and all), Eric 'Jatelo' Lelo (for all the company!) and Ben Mungai. Thanks guys for putting up with me in the office. My uncles John 'JN' and the late Sammy and my aunts Joyce and Priscillah who did all the hard work in my formative years.

Karis 'ZB', Henry, Mike, George, Naomi, Carol and Dr. Nyaga. Thanks for the trips to the airport and Heritage, California and Manyatta. I will 'revenge'.

A special 'thank you' to Lynn. Thanks for everything...

Declaration

I hereby declare that the work presented in this thesis together with the exegetical material herein is my original work. It contains no material previously written or published by another person except where due reference has been acknowledged. This work has not been accepted in substance for any degree or award, and is not being submitted concurrently in candidature for any other degree or award of any university.

Signed

July 2006

Abstract

Three sets of molecular markes were used to investigate the population genetics of three populations of *Plasmodium falciparum* from Kenya; Mwea (low transmission), Tiwi (moderate transmission) and Bondo (high transmission). One set of markers codes for polymorphic antigens while the other two are microsatellite markers; one set located in non coding regions of the genome while the other set is located in the regions flanking two genes whose products are targets of the antimalarial drug sulphadoxine/pyrimethamine (SP).

A comparison of the effectiveness of antigen-coding and the unlinked microsatellite loci in differentiating recrudescence from reinfection revealed that both sets of markers are equally effective. The microsatellite loci however, revealed more alleles per population than the antigen-coding loci possibly due to their different mutation rates.

An analysis of the three populations using the neutral microsatellite loci revealed high levels of diversity, lack of linkage disequilibrium and virtually no population substructuring (F_{ST}<0.008) in the Kenyan *P. falciparum* populations even with the geographical areas being as much as 800 km apart. This indicates a lot of gene flow among these populations a factor that can only be explained by movement of people between the areas studied.

An analysis of the same samples from the three areas at the *dihydrofolate reductase* (*dhfr*) and *dihydropteroate synthase* (*dhps*) gene loci that code for targets of the antimalarial drug SP revealed high prevalence of the multiply substituted alleles associated with SP resistance in the three regions.

An analysis of ~17 kb regions flanking both sides *dhfr* reveal a strong selective sweep of the 108N/51I/59R triple mutation alleles associated with pyrimethamine

resistance. The work presented also demonstrates that alleles of the *dhfr* gene, especially the triple mutant allele, isolated from the three different areas are closely related to one another and probably share a common and very recent ancestor. Most notable is the finding that *dhfr* triple mutants seeem to be imported into the country through immigration from elsewhere. An equivalent region flanking the *dhps* gene also revealed a strong selective sweep of the 437G/540E double mutation allele associated with sulphadoxine resistance in two of the three sites. However, double mutation *dhps* alleles from Mwea revealed no selection at all. While the three populations reveal no geographic substructuring using the results of the unlinked microsatellite loci, they seem to be highly structured in their drug resistance patterns. While it would be expected that these populations would have the same prevalence of drug resistance mutations (due to the apparent panmixia), the Mwea population appears quite different in regard to selection for drug resistance-associted alleles. This is possibly due to the diet, other drug interactions or the hosts' genetics in this area.

A simplistic model on the rate of spread of drug resistance in the three populations reveals that the selection for drug resistance alleles is faster in the lower transmission area of Mwea (selection coefficient, $s = 0.26$) and slowest in Bondo ($s = 0.10$) indicating selection for drug resistant alleles is favoured by low transmission. These observations have implications for malaria drug resistance surveillance programs due to the fact that if treatment failure spreads faster in low transmission areas where almost all the population has low immunity, malaria epidemics are bound to occur resulting in huge morbidity and mortality.

1.0... Introduction and literature review

1.1 Introduction

Malaria is one of the most prevalent parasitic diseases of man affecting over 100 countries with a combined population of 2.4 billion people (Breman 2001). *Plasmodium falciparum, P. ovale, P. malariae* and *P. vivax* are the four species of the parasite that cause malaria in man with the first accounting for most of the lethal infections.

The annual number of clinical cases is estimated at between 300-500 million per year, more than 90% of which occur in sub-Saharan Africa. Malaria mortality is estimated to be over 2 million persons per year with the principal victims being African children below 5 years of age (Phillips 2001; Snow *et al.*, 2005; World Health Organization 2003a). The other high-risk groups for malaria include pregnant women, non-immune travellers, refugees and displaced persons.

The malaria parasite is transmitted by an insect vector of the genus *Anopheles*, the most important of which is *Anopheles gambiae.* The distribution of the disease mirrors that of the vector.

Within the past few decades there has been rapid emergence of *P. falciparum* malaria resistant to almost all of the available antimalarial drugs. This has added a further blow to the war against the disease, making the provision of malaria treatment and control increasingly difficult and costly. This has generally been attributed to the great genetic diversity of the parasite especially in countries with a wide range of transmission intensities (Babiker & Walliker 1997). Infection with genotypically different parasites of the same species, often referred to as multi-clonal infection, is also relatively frequent, and a positive relationship has been reported in three African countries between entomological inoculation rates (EIR) and number of genotypes in infected people (Babiker & Walliker 1997; Walliker, Babiker, & Ranford-Cartwright 1998).

1.2 The Lifecycle of the Malaria parasite, *Plasmodium falciparum*

Plasmodium spp. belong to the phylum Apicomplexa, members of which alternate in one or more vertebrate or invertebrate hosts. *Plasmodium falciparum* undergoes a haploid asexual replication in the human host and a brief diploid sexual phase in the mosquito vector (Good *et al.*, 1988; Walliker *et al.*, 1987; Walliker *et al.*, 1975).

The feeding mosquito injects sporozoites into the skin of the human host, from where they migrate to the liver, invade hepatocytes, and undergo asexual differentiation to form a hepatic schizont. Merozoites burst from the cell and invade erythrocytes. Multiple rounds of asexual replication, schizogony, erythrocyte lysis and merozoite reinvasion then take place, with blood parasite density rising exponentially with each round of replication. The duration of the erythrocytic cycle determines the periodicity of fever and gave rise to the old-fashioned names of tertian and quartian malaria. For *P. falciparum* the cycle takes approximately 48 hours, with fever appearing every third day (reviewed in Good *et al.* 1988).

Gametocytes of *Plasmodium* spp. are formed from a subset of asexual parasites after one or more rounds of intraerythrocytic proliferation. It appears that a combination of innate genetic factors and environmental cues in the blood stream of the host triggers the switch from erythrocytic schizogony to gametocytogenesis (Bruce *et al.*, 1990; Smith *et al.*, 2002). All merozoites within a single schizont of *P. falciparum* are destined either to produce further asexual stages or to develop into gametocytes (Bruce *et al.*, 1990), and all of the merozoites released from one sexually committed schizont become gametocytes of the same sex (Silvestrini F *et al.*, 2000; Smith *et al.*, 2000).

Male and female gametocytes are taken up in the blood meal by the mosquito as it feeds from an infected person. Male and female gametocytes emerge from their red

blood cell hosts; undergo a process of gametogenesis, and the resultant gametes fuse to form the zygote (ookinete). The ookinete migrates through the midgut wall of the mosquito and develops into an oocyst. The first stage of meiosis is known to take place within the ookinete (the second stage is not clear), followed by multiple rounds of mitosis, to produce between 2,000-10,000 (Pringle 1965; Rosenberg & Rungsiwongse 1991) haploid sporozoites per oocyst, which migrate to the salivary glands ready for reinfection into the human host during the next blood meal.

1.3 Ploidy and the Genome of *P. falciparum*

The parasite in the vertebrate host is entirely haploid. Studies on inheritance of isoenzymes and other characters in *P. chabaudi* (Walliker *et al.*, 1975) and in *P. falciparum* (Walliker *et al.*, 1987) have supported haploidy for the erythrocytic stages. Exo-erythrocytic stage haploidy was demonstrated by studies on the inheritance of antigens of these forms (Szarfman *et al.*, 1988).

Studies on DNA content have demonstrated that sporozoites, ring forms, young trophozoites and mature microgametes possess a similar quantity of DNA, assumed to be the haploid amount (Janse *et al.*, 1986). Mature macrogametes were found to have approximately 50% more DNA than the haploid quantity, possibly because of amplification of gametocyte-specific genes.

The only diploid phase in the *Plasmodium* lifecycle is the zygote (ookinete) in the mosquito gut. Approximately 4 times the haploid DNA quantity is synthesised during meiosis that takes place at this time (Janse *et al.*, 1986). This is consistent with the duplication of the diploid chromosome set at the first stage of meiosis.

The nuclear genome of *P. falciparum* is composed of 22.8 megabases (Mb) of DNA distributed among 14 chromosomes. These chromosomes range in size from 0.643 Mb to 3.29 Mb (reviewed in (Gardner *et al.*, 2002)). Overall the genome of *P.*

falciparum is very (A+T)-rich varying from 80.6% in genes to approximately 90% in introns and intergenic regions (Gardner *et al.*, 2002).

Approximately 5,300 protein-encoding genes have been predicted which suggests an average gene density in *P. falciparum* of 1 gene per 4,300bp (Gardner *et al.*, 2002). Introns have been predicted in 54% of *P. falciparum* genes. Excluding introns, the mean length of these genes is 2.3 kilobases (kb), substantially larger than in other organisms in which gene lengths range from 1.3kb to 1.6kb (Louis *et al.*, 1994).

The parasite also has two other genomes: the mitochondrial and the apicoplast genomes (Vaidya *et al.*, 1993; Vaidya *et al.*, 1989). The mitochondrial genome of *P. falciparum* is a 6.0 kb tandemly repeated linear element which is uniparentally inherited (Creasey *et al.*, 1993; Feagin 1992; Joseph *et al.*, 1989). There is 90% conservation of sequence identity among the mitochondrial genomes of *P. falciparum, P. vivax, P. yoelii, and P. gallinaceum* (McIntosh *et al.*, 1998). Detailed processes of replication and expression have been well characterized in the mitochondrial genes of the asexual stage parasite (Feagin & Drew 1995; Preiser *et al.*, 1996).

Based on the malaria genome databases, about 380 proteins are predicted to target the mitochondrion post-translationally. These include the pyruvate dehydrogenase complex, the complete tricarboxylic acid cycle enzymes, many electron transport complexes and ATP synthase (Bender *et al.*, 2003; Gardner *et al.*, 2002). However, functional analyses of the mitochondrion and its products still remain to be elucidated e.g. what controls processes like gene deletion, RNA interference, transcriptomics and proteomics. The mitochondrion is a potential chemotherapeutic target for antimalarial drug development, for instance, the enzyme dihydroorotate dehydrogenase which is essential to parasite function (Baldwin *et al.*, 2002; Krungkrai 1993; Ridley 2002).

There is also evidence that the malaria parasites produce oxidative stress via the mitochondrial superoxide radical generation, a phenomenon that exists in mammalian mitochondria (Dileepan & Kennedy 1985), and is finely balanced against the parasite's antioxidant enzyme activities (Krungkrai 1991). This intracellular redox environment minimizes the generation of reactive oxygen species that damage nucleic acids, proteins, lipids and membranes and still allows essential metabolic functions.

Malaria parasites also harbour a relict plastid, homologous to the chloroplasts of plants and algae, called the apicoplast. This organelle seems to be essential to parasite growth but its exact role is unclear. The apicoplast in malaria parasites has a 35kb genome that encodes only 30 proteins. Among the apicoplast's genes include those that code for a set of tRNAs, *rRNAs, rpoB, rpoC, tufA*, and ribosomal protein (*rp*) genes. The apicoplast genome contains genes typical for plastids (*ycf24 and clpC*) but also several open reading frames of unknown function (Kohler *et al.*, 1997; McFadden & Roos 1999). Translation with prokaryotic features may be present in the apicoplast; however, most of its proteins appear to be nucleus-encoded and imported into the organelle, using an import secretory pathway (Waller *et al.*, 2000).

1.4 Generation of genetic diversity in *Plasmodium falciparum*

The chromosomes of *P. falciparum* vary considerably in size, with most variation occurring in the subtelomeric regions. Chromosomal size variation is also observed in cultures of erythrocytic parasites, but is probably due to chromosome breakage and healing events and not due to recombination (Scherf & Mattei 1992). The subtelomeric regions of the chromosomes show a striking display of conservation within the genome that is probably due to promiscuous inter-chromosome exchange of subtelomeric regions (Scherf & Mattei 1992). Subtelomeric exchanges are known to occur in other eukaryotes (Louis *et al.*, 1994; Rudenko *et al.*, 1996; van Deutekom *et*

al., 1996), but the regions involved are much smaller. Various mechanisms may produce these size polymorphisms in *P. falciparum*. Chromosome breakage and healing during mitotic expansion of parasite populations is one mechanism for loss of subtelomeric sequences and genes from *P. falciparum* chromosomes (Cappai *et al.*, 1989; Pologe & Ravetch 1988; Scherf *et al.*, 1992). Expansion of large tandemly repeated units within individual chromosomes has been shown to be another means of generating large chromosome size polymorphisms (Triglia *et al.*, 1991). Homologous recombination among subtelomeric repetitive elements has been postulated as a third mechanism of polymorphic variation (Corcoran *et al.*, 1988). Homologous chromosomes of different sizes readily undergo meiotic recombination after zygote formation to produce new size variants (Hinterberg K *et al.*, 1994; Sinnis & Wellems 1988; Walliker *et al.*, 1987).

The recognition of parasite polymorphisms and their association with distinct properties has made considerable progress in the last few years. This is largely due to the advent of polymerase chain reaction (PCR)-based assays. Single nucleotide polymorphisms (SNPS) and other types of polymorphisms have been analysed using specific digestion assays with restriction enzymes, DNA hybridisation techniques using specific oligonucleotides, allele specific amplification, or by sequencing of PCR products.

Single nucleotide polymorphisms (SNPs) contribute to most of the polymorphism of the parasite's genes. Some researchers have argued that synonymous mutations are scarce in most of the parasite's genes thus supporting a clonal population structure and a recent bottleneck (Rich *et al.*, 1998; Rich *et al.*, 1997), while others have for the existence of long-lasting polymorphisms and high levels of intraspecific synonymous polymorphism in *P.falciparum* (Hughes & Verra, 1998). Rich and

colleagues' arguments were based on alignment of some sequences from isolates obtained from multiple geographical locations (Rich *et al.*, 2000). However, this work only examined a small proportion of genes such as *Msp1*, which are antigen-coding and infact are vaccine candidates, where there are multiple sequences. Most of the loci examined by these authors (Rich *et al.*, 1998; Rich & Ayala 1998) are housekeeping genes encoding metabolic enzymes, chaperone proteins and surface proteins. The existence of balancing selection at such loci is not expected and most are probably evolving neutrally. It is not expected that the pattern of coalescence at a neutral locus would be the same as that at a locus under balancing selection; rather, polymorphism at the former will be much less long-lasting than at the latter (Takahata & Nei 1990). A closer reexamination of the work by Rich and colleagues (Rich et al., 1998) also reveals that their analysis excluded repeat sequences and the authors erred in calculating the ratios of synonymous to nonsynonymous substitutions which could have led to their conclusions. It is also known that *P. falciparum* has a strong codon bias due to the extraordinary AT-richness of the genome. This nucleotide content bias affects codon usage (Nakamura *et al.*, 1998). The rate of synonymous substitutions in organisms with highly biased genomes such as *P.falciparum* is less than in organisms such as humans, whose codon usage is less biased (Hughes & Verra 1998). The fact that long-lasting polymorphisms have been maintained at certain loci of *P.falciparum* but not at others is also strong evidence that interlocus recombination has been an important factor in the evolution of this species (Hughes & Verra 1998).

Non-synonymous SNPs are found at high frequencies in genes that are under strong selective pressure and an excess of non-synonymous mutations is thought to indicate an area of the genome under diversifying selection (McDonald 1994). Comparison of the frequency of the putatively neutral synonymous and the more

frequent non-synonymous substitutions suggests that natural selection may account for most polymorphisms observed at functional gene loci (Escalante *et al.*, 1998).

SNPs affecting drug resistance in *P. falciparum* have been found in the dihydrofolate reductase-thymidylate synthase (*dhfr-ts*) and the dihydropteroate synthase (*dhps*) genes. Mutations in these genes are associated with resistance to pyrimethamine and sulphadoxine, respectively (Cowman *et al.*, 1988; Peterson *et al.*, 1990; Peterson *et al.*, 1988a; Wang *et al.*, 1997a). SNPs in the chloroquine resistance transporter gene (*Pfcrt*) with a modifying effect from the multidrug resistance gene 1 (*Pfmdr 1*) have been shown to be responsible for chloroquine resistance (Fidock *et al.*, 2000a). Genes where polymorphism has most likely arisen through intragenic recombination and other mechanisms in repetitive segments are characterised by repeat motifs with length variability differing between strains (Jongwutiwes *et al.*, 1993).

1.5 Genetic Analysis of *P. falciparum*: Complexity of Infections

Both clinical and asymptomatic infections are often composed of a mixture of parasite strains, which were assumed to be the result of many superimposed independent infections, hence reflecting transmission intensity (reviewed in (Konate *et al.*, 1999; Smith *et al.*, 1999a). Many studies have been conducted to analyse the genetic diversity, dynamics and population structure of *P. falciparum* using different molecular tools, which are discussed, in later sections.

1.5.1 Multiclonality in malaria infections

P. falciparum infections have been shown to be genetically diverse in all except unusual situations (such as epidemics on islands or in isolated populations), and most infected individuals in moderate to highly endemic transmission areas harbour more than one parasite clone, a condition known as multiclonality (Arnot 1998). PCR-based techniques have shown a high prevalence of multiclonal infections (Sherman 1998).

Diverse multiclonal infections have been found to be almost as common in regions of low endemicity such as eastern Sudan (Babiker *et al.*, 1991a; Babiker *et al.*, 1991b; Bayoumi *et al.*, 1993) as they are in areas of higher transmission such as The Gambia and Tanzania (Babiker *et al.*, 1997; Conway & McBride 1991). In other areas such as South America, where transmission is low, low diversity and population substructuring of parasites are observed (Anderson *et al.*, 2000a).

In humans living in regions endemic for malaria, partial immunity to *Plasmodium falciparum* is acquired as a result of natural exposure to these multiple infections over many years. One study in Senegal demonstrated that while most symptomatic and asymptomatic subjects were infected by multiple *P. falciparum* genotypes, the number of clones present in asymptomatic infections decreased at the age where a more efficient immune response is in place, as indicated by lower parasite densities and reduced clinical attack rates (Ntoumi *et al.*, 1995). This study was conducted in Dielmo, an area with intense transmission and where partial immunity develops with age. Another study in Ndiop, a Senegalese area with low transmission observed different results. This study used the three polymorphic markers *Msp1*, *Msp2* and *Glurp* and showed that the average number of fragments (multiplicity of infection, MOI) per clinical malaria sample was higher than for samples taken from asymptomatic infections: 1.75 for *Msp1*, 2.15 for *Msp2* and 1.4 for *Glurp* compared to 1.5, 1.5, and 1.2 for the asymptomatics, for *Msp1* block 2, *Msp2*, and *Glurp* typing reactions, respectively. They also showed that in both groups, most multiple infections contained 2–3 alleles (Zwetyenga *et al.*, 1998).The later study was conducted in an area of low endemicity where clinical malaria is seen in adults as well as children, and the main finding was that age had no influence on either the complexity of infection or on the distribution of alleles (Zwetyenga *et al.*, 1998). This observation is consistent with other

studies that have shown a reduction with increasing age in the number of genotypes present in infected individuals from high transmission areas (Ntoumi *et al.*, 1995; Smith *et al.*, 1999b) suggesting that this is a reflection of acquired immunity.

Multiple clone infections are important because they provide the possiblility of cross-fertilisation, which itself allows recombination. The extent of recombination between parasites defines population structure.

1.5.2 Population structure, mating and recombination

A clonal structure of *P. falciparum,* where all parasites are related due to inbreeding, was first proposed in 1990 (Tibayrenc *et al.*, 1990). However other authors (Babiker & Walliker 1997) have argued that in highly endemic areas outcrossing (mating between/among different parasite genotypes) occurs frequently enough to generate a great diversity of molecular genotypes.

In the mosquito midgut, fertilisation can occur between a male and a female gamete of the same genotype (self-fertilisation) or between gametes of different genotype (cross-fertilisation). Mating occurs in a random fashion between gametes present in the bloodmeal, in accordance with Hardy-Weinberg equilibrium (Ranford-Cartwright *et al.*, 1993). In naturally infected mosquitoes, both crossing and selfing have been shown to occur (Babiker *et al.*, 1994; Paul *et al.*, 1995), although the ratio varies according to the genetic structure of the population. The population structure of this parasite is very diverse, showing variation in outcrossing rates from 10% in Papua New Guinea (Paul *et al.*, 1995) to 60 % in Tanzania (Babiker *et al.*, 1994).

Strong linkage disequilibrium (LD), low genetic diversity and high levels of geographical variation have been observed in areas of low malaria transmission while random association (linkage equilibrium (LE)), high genetic diversity and minimal geographical differentiation were observed in regions with high transmission especially

in Africa and Papua New Guinea (Anderson *et al.*, 2000a). The observed association between transmission intensity and LD can be explained by the fact that *P. falciparum* has a mixed mating system in which inbreeding predominates in low transmission areas, while higher levels of outbreeding occur in regions with higher transmission. People are rarely superinfected with more than one parasite clone in low transmission areas, and thus unrelated parasites rarely co-occur in the same mosquito blood meal. Mating can only occur between parasites in the same mosquito bloodmeal. While it is theoretically possible that a mosquito could take a partial bloodmeal on two or more people (Soremekun *et al.*, 2004), the second bloodmeal would have to be taken very rapidly after the first one, as gametogenesis and fertilisation occur within 10 minutes of the bloodmeal. Therefore self-fertilisation (and inbreeding) will be much more common in low transmission areas. Conversely, multiple clone infections are frequent in high transmission areas and mosquitoes therefore often ingest unrelated parasites, leading to higher levels of outbreeding (and recombination) (Babiker *et al.*, 1994; Paul *et al.*, 1995). The extent of parasite diversity may vary with transmission intensity, being at least in part generated by recombination during meiosis in the mosquito. Since every inoculation is preceded by sexual reproduction of parasites in the mosquito, the inoculation rate is predicted to reflect the rate of parasite recombination. Sexual reproduction, which generates new genotypes (Babiker *et al.*, 1994; Paul *et al.*, 1995; Rosenberg & Rungsiwongse 1991; Walliker *et al.*, 1987) and novel alleles by intragenic recombination, (Kerr *et al.*, 1994) is likely to play a significant role in the generation of parasite diversity. Recombination during meiosis however, is only able to increase diversity in heterozygotes, formed by crossfertilisation between different gametes. This is mainly found in areas of high transmission where probabilities of multiple clone infections and, hence mating between unlike gametes, will be more frequent. Levels of

LD have important influences on the population biology of *P. falciparum* particularly the rate at which recombination breaks down association between genes, which may in turn influence the persistence of clonal genotypes (Hastings & Wedgewood-Oppenheim 1997; Paul *et al.*, 1995). LD also influences the maintenance of antigenically distinct "strains" (Gupta *et al.*, 1996; Hastings & Wedgewood-Oppenheim 1997), and the spread of drug resistance (Dye & Godfray 1993; Hastings 1997; Hastings & Mackinnon 1998), and may also influence sex ratio (Dye & Godfray 1993; Read *et al.*, 1992).

These studies, though few, demonstrate the high diversity of *P. falciparum* populations especially in Africa where transmission is intense. The strong geographical sub-structuring in South America may be explained by multiple colonization events; malaria is thought to have been introduced into South America ≅ 500 years ago with the arrival of Europeans and that subsequent reintroduction from Africa occurred with the slave trade (Anderson *et al.*, 2000a).

Regardless of the causes, the differences in genetic diversity, population differentiation, and LD in different locations have important consequences for our understanding of *P. falciparum* biology (Anderson *et al.*, 2000b). In parasite populations with low microsatellite diversity, we would expect to see reduced diversity in antigen-encoding loci (Ferreira *et al.*, 1998) and a smaller repertoire of variant surface antigens. Hence, under a model of genotype-specific immunity, we might expect effective immunity to malaria to be generated following a relatively small number of infective mosquito bites in low transmission regions (Gupta & Day 1994). In areas with low recombination, multilocus genotypes will be maintained through multiple generations making it possible to track the spread of multilocus genotypes within communities (Anderson *et al.*, 2000a). Comparison of infection characteristics of multiply represented haplotypes can be used to investigate which aspects of *P.*

falciparum traits are a product of parasite genetics rather than host factors (Anderson *et al.*, 2000a; Walliker *et al.*, 2005).

1.5.3 Choice of molecular markers

Genetic markers are powerful tools for investigating the population structure and epidemiology of parasites. For both these purposes the most effective markers would be those that are neutral, highly polymorphic and stable during the course of infection. However, such markers do not exist. Each genetic marker has a range over which it can detect genetic variation (Sunnucks 2000). For instance, if the marker is too polymorphic, either the required sample size to describe diversity will be too high, or the samples will be too different resulting in homoplasy. If the marker has little polymorphism, then samples will appear the same when infact they are genetically different. Genetic markers should also be chosen based on the question being asked and also their applicability for the particular situation. For example, microsatellite markers might be good for population studies but they are not ideal for situations like differentiating recrudescences from reinfections (Mwangi *et al.*, 2006).

Multiple loci are used since measures of population structure characteristically show high levels of variance among loci (Mcdonald 1994; Nei 1973). Selectively 'neutral' loci are preferred, since the aim is to use these markers to make inferences about population history or transmission of parasites between hosts. These markers are most likely to be found in the non-coding sequences of the DNA.

The molecular markers so far used in population genetic analyses can be classified as antigen-encoding genes (Abderrazak *et al.*, 1999; Babiker *et al.*, 1995; Maitland *et al.*, 2000; Zwetyenga *et al.*, 1998), enzymes with electrophoretic variants (Abderrazak *et al.*, 1999), and microsatellite loci (Anderson *et al.*, 1999). Given that some of the antigen genes encode proteins that are recognised by the host's immune

mechanism, the observed high levels of diversity have been attributed to positive or diversifying natural selection, an outcome of accumulation and frequent switch of suitable mutations, by means of which the parasite escapes the host's immune defences (Anders & Saul 2000; Hughes 1991; Hughes 1992; Hughes & Hughes 1995). Evidence for selection comes from comparison of the ratio of synonymous to non-synonymous mutation rates (Escalante *et al.*, 1998).(Felger *et al.*, 1997; Hughes 1991; Hughes & Hughes 1995; Rich *et al.*, 1997) In most organisms, synonymous nucleotide substitutions are usually less common than non-synonymous substitutions (Lockyer *et al.*, 1989; Shi *et al.*, 1992; Thomas *et al.*, 1990a; Thomas *et al.*, 1990b). Synonymous substitutions are likely to be neutral whereas non-synonymous substitutions may be functionally constrained and thus, subject to purifying natural selection (Kimura 1977; Ohta 1996). However, selection can also act to favour temporal diversity, and therefore non-synonymous mutation. For allozyme markers, natural selection may account for the rapid spread of a favoured genotype throughout populations, particularly when the population is large and/or the selection is strong (Rich *et al.*, 1998). However, this would overtime reach saturation of the favoured genotype and the diversity would be lost.

Some of the markers that have been employed in epidemiological studies of *P. falciparum* are discussed in the following sections.

1.5.3.1 Allozyme typing

Allozymes, in particular variants of parasite enzymes such as glucose phosphate isomerase (GPI) and lactate dehydrogenase (LDH) (Carter & McGregor 1973) were the first plasmodial variant proteins to be analysed. Electrophoretic variants of these enzymes were observed in individual malaria patients, and were used to demonstrate the

existence of mixed infections with differing *P. falciparum* clones in individuals (Carter & McGregor 1973).

Allozymes however, have various disadvantages over other loci; for instance, it is possible that differential selection could be acting to maintain different combinations of Allozyme alleles in a population. This impacts on the neutrality of these markers (Jordan *et al.*, 1997). The number of electrophoretic variants in a given population and therefore the resolution of these markers are quite low. The other disadvantage of Allozymes over other markers is that the likelihood of an amino acid change translating into a mobility shift is very low and this increases the frequency of different alleles being typed as similar.

1.5.3.2 Polymerase chain reaction and DNA polymorphism

The use of polymerase chain reaction (PCR) amplification of deoxyribonucleic acid for the detection of genetic polymorphism in malaria parasites has broadened our knowledge of the parasites' natural genetic diversity. The main polymorphic *Plasmodium* genes studied are those encoding merozoite surface proteins 1 and 2 (MSP1 and MSP2) and the glutamate rich protein (GLURP). The genes have been used to make community-based comparisons between parasites infecting different human populations exposed to malaria, and have also been used to determine clonal multiplicity in individuals.

In an analysis of the genetic characteristics of parasite populations, the PCR typing methodology provides the advantage of a great sensitivity and specificity, without the need to culture parasites, as was necessary for allozyme typing. The single copy genes *Msp1, Msp2, and Glurp* show high allelic polymorphism, allowing easy detection of parasite diversity; alleles with similar or identical serologic properties can be distinguished by size polymorphism of the PCR product.

1.5.3.2.1 Merozoite surface protein-1 (MSP-1)

The merozoite surface protein-1 (MSP-1), also known as merozoite surface antigen-1 (MSA-1), glycoprotein 195 (gp195), and P190, is one of the most studied plasmodial antigens. It is synthesised by liver- and blood-stage schizonts (Holder & Freeman 1982; Szarfman *et al.*, 1988) and varies in size between different parasite lines from 180-220kDa (McBride *et al.*, 1985). The antigen is encoded by a single gene located on chromosome 9 (Kemp *et al.*, 1987). Each haploid parasite therefore produces one allelic variant of this protein (Howard *et al.*, 1986; McBride & Heidrich 1987). The extensive polymorphism of MSP-1 was originally characterized using a panel of monoclonal antibodies, some to polymorphic epitopes of the molecule (Holder 1988; Holder *et al.*, 1988b; Holder *et al.*, 1988a; McBride *et al.*, 1985). Sequencing of the gene from different *P. falciparum* isolates has revealed the extent of the allelic polymorphism (Holder *et al.*, 1985; Kemp *et al.*, 1986; Miller *et al.*, 1993; Ranford-Cartwright *et al.*, 1991; Tanabe *et al.*, 1987; Holder *et al.*, 1985; Kemp *et al.*, 1986; Holder *et al.*, 1985).

By comparing nucleotide sequences, Tanabe and colleagues (Tanabe *et al.*, 1987) identified polymorphic domains, conserved domains, and intermediate or 'semi-conserved domains' based on the percentage identity of conserved nucleotide sequences. On this basis the gene was divided into 17 domains or blocks; 7 of which are variable, interspersed with five conserved and five semi-conserved blocks (Figure 1). Sequences are dimorphic, i.e., substitutions at given positions are either one or the other of any 2 residues, in most parts of the gene with the exception of a polymorphic tripeptide-encoding region in block 2. Sequences from the MAD20 and K1 strains serve as representative types for the two dimorphs (Peterson *et al.*, 1988b; Tanabe *et al.*, 1987).

The block 2 domain contains tripeptide repeats and exhibits considerable polymorphism. This block differs by sequence and/or by numbers of repeats in almost every allele investigated so far. Three basic families of block 2 have been defined largely on the basis of non-repetitive sequences at the ends of the repeat regions and are named after the isolates from which they were originally described; the MAD20-type, the K1-type and the RO33-type (Tanabe *et al.*, 1987). The RO33-type block 2 lacks the typical N-terminal repeat structure found in the other two alleles (Certa 1990) and is non-polymorphic between alleles of this family.

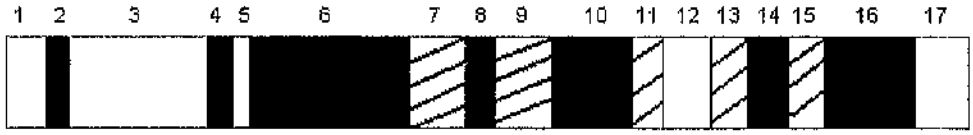

Figure 1: Schematic diagram of the 17 domains of the *msp*1 gene of *P. falciparum*. The open, hatched and completely filled regions are the conserved semi-conserved and variable domains respectively.

Msp1 alleles may be defined as unique combinations of: (a) one of three versions of block 2 (MAD20, K1 or RO33), (b) one of four possible versions of block 4, because recombination within this region generates MAD20/K1 and K1/MAD20 hybrids in addition to the 'pure' allelic types MAD20 and K1 (Conway *et al.*, 1991; Kaneko *et al.*, 1996), and (c) one of two versions (MAD20 or K1) of the segment between the variable blocks 6 and 16, that comprises about 60% of the gene. Recombinant-type alleles have not been described in this 3' portion of the gene covering blocks 6 to 16 (Conway *et al.*, 1991; Jongwutiwes *et al.*, 1993; Kaneko *et al.*, 1996; Tanabe *et al.*, 1989; Tanabe *et al.*, 1987). Minor differences also exist between homologous versions of the same variable block, and nucleotide substitutions (most of

which are dimorphic) occur in semi-conserved and conserved blocks (Tanabe *et al.*, 1987).

Excluding the polymorphism seen in block 2, most allelic diversity is thought to be generated by intragenic recombination, at the diploid stage in the mosquito host, within blocks 3, 4 and 5. Direct evidence of recombination in block 3 has been demonstrated following a genetic cross (Kerr *et al.*, 1994).

1.5.3.2.2 Merozoite surface protein-2 (MSP-2)

This antigen, also known as merozoite surface antigen-2 (MSA-2), is synthesised in the blood-stage schizonts and, like MSP-1, is incorporated on the schizont surface, and on the surface of free merozoites (Miettinen-Bauman *et al.*, 1988; Stanley *et al.*, 1985). MSP-2 was first defined by monoclonal antibodies (Fenton *et al.*, 1989), and the size of the protein (Stanley *et al.*, 1985) was found to vary greatly in size from 35-36kDa on Western blots with protein extracted from different parasite isolates (Ramasamy 1987). MSP-2 is known to be one of the antigens present in immune complexes that form at the surface of merozoites when antibodies in immune serum inhibit merozoite dispersal (Thomas *et al.*, 1990c). The *Msp2* locus has been mapped to a region of approximately 800 bp on chromosome 2 (Kemp *et al.*, 1987).

The MSP2 protein consists of conserved amino- and carboxy-terminal regions flanking a central variable region that is in turn composed of non-repetitive sequences surrounding a repetitive sequence (Figure 2). Comparisons of *Msp2* sequences from different parasite isolates revealed a similar block structure to that seen with *Msp1*, but with four distinct blocks by virtue of local differences in the degree of homology between sequences from different strains (Thomas *et al.*, 1990; Smythe *et al.*, 1988; Smythe *et al.*, 1990).

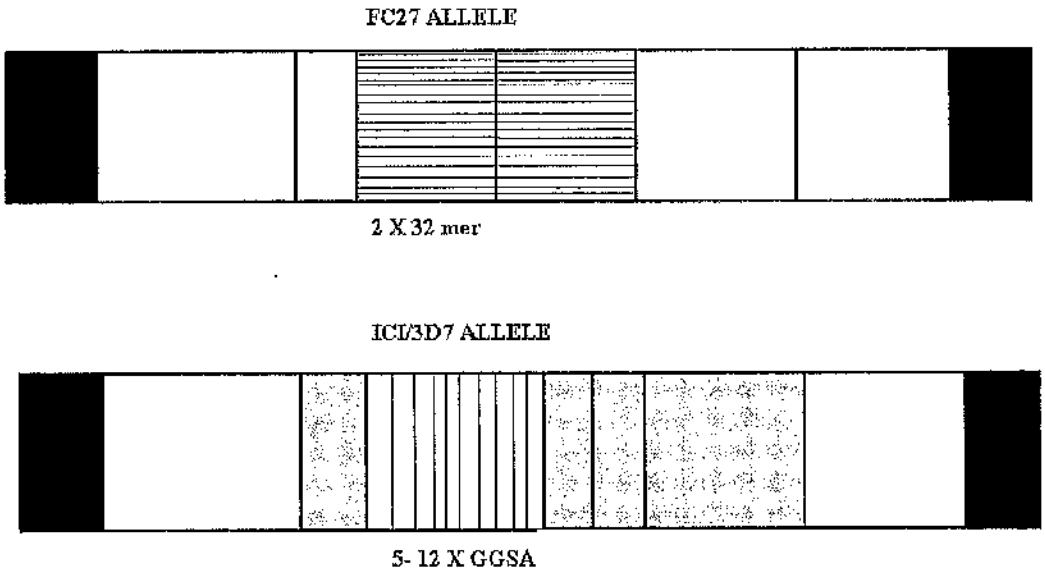

Figure 2: Schematic diagrams of the two main MSP2 allelic families.

Dark regions at the left and right indicate signal and anchor sequences. Open regions internal to these represent conserved regions, whereas grey and striped regions represent variable nonrepetitive and repetitive sequences, respectively.

The repetitive part of the gene exhibits marked polymorphism (Fenton *et al.*, 1991; Smythe *et al.*, 1988; Smythe *et al.*, 1990; Thomas *et al.*, 1990c). All of the alleles of *Msp2* so far examined can be grouped into two families, defined by the non-repetitive region 3 (Thomas *et al.*, 1990c). These are denoted the Indochina 1-like (IC1-like) and FC27-like families (Figure 2).

The 5' conserved region, with 100% conservation of nucleotide sequence (region 1), encompasses the first 129 bp from the initiation codon. The translation of this region includes the predicted signal peptide, which is presumably absent from the mature parasite protein (Smythe *et al.*, 1990).

Region 2 is strain variable, and is characterized by the presence of tandemly arranged repeat sequences. It begins with a short stretch of amino acids, some of which are conservatively substituted within the IC1-type allele. This stretch lies immediately

N-terminal to a glycine, serine and alanine-rich block that is present in all IC1-type strains, but absent from the FC27-type strains. The MSP2 sequence from parasite FC27 contains two identical copies of a 96-bp repeat unit commencing at base 178. In contrast, the alleles from parasites 3D7 and IC1 contain multiple copies (5 and 12, respectively) of a 12-bp repeat unit commencing in both genes at base 169. The repeat in alleles of IC1-family encodes the amino acid sequence Gly Gly Ser Ala and bears no apparent relationship at either the nucleotide or amino acid level to the 32-amino-acid repeat in FC27-type alleles (Thomas *et al.*, 1990b).

Sequencing of a large number of *Msp2* alleles from isolates drawn from various geographical regions has revealed hybrid alleles that do not fall into either the IC1-type or FC27-type families (Marshall et al. 1991). An allele (NIG60) was sequenced with an N-terminal region of the IC1 allelic family and a C-terminal region of the FC27 allelic family. As the *Msp2* gene is found in a single copy in the haploid genome, this allele was suggested to have been formed by intragenic recombination, presumably occurring during meiosis. The apparent crossover event occurred within the repetitive variable regions (Marshall et al. 1991).

1.5.3.2.3 Glutamate-Rich Protein (GLURP)

The glutamate-rich protein (*glurp*) gene has been used as a marker in many malaria studies (Borre *et al.*, 1991). The single copy *glurp* gene is located on chromosome 10, and encodes a polypeptide of 1271 residues with a predicted molecular mass of 145 kDa. Rabbit antiserum against a fusion protein expressing the C-terminal end of the molecule detects a protein with a molecular mass of 220 kDa (Hogh *et al.*, 1993). The 220-kDa GLURP protein is located in the parasitophorous vacuole of liver-stage parasites, in erythrocytic schizonts, and on the surface of newly released merozoites and may facilitate merozoite invasion (Borre *et al.*, 1991). The gene encodes

an antigen with an amino-terminal nonrepeat region, denoted R0 ($GLURP_{94-489}$), as well as a central repeat region ($GLURP_{489-705}$) and a carboxy-terminal ($GLURP_{705-1178}$) repeat region (denoted R1 and R2, respectively). All 3 regions are quite polymorphic among isolates (Borre *et al.*, 1991) . GLURP has not been as widely studied as the merozoite surface proteins but its variability, exhibited by length polymorphisms of the repeat regions, has been described and determined in studies on multiclonal *P. falciparum* infections (Boudin *et al.*, 1993; Dodoo *et al.*, 2000; Hogh *et al.*, 1992; Theisen *et al.*, 1998).

1.5.2.3 Polymorphism of Noncoding Sequences

Intergenic and intronic DNA regions often contain simple sequence repeats of variable lengths known as satellites. These are further classified into satellites, mini- and micro-satellites based on the length of the repeat unit.

Satellite DNA was the first of the tandemly repeating DNA sequences to be discovered, and was so named by its appearance as minor or "satellite" bands that separated from the "bulk" DNA upon buoyant density gradient centrifugation (Britten & Kohne 1968). Human satellite DNA is not transcribed, and is found in heterochromatin (condensed areas of chromatin consistently lacking actively expressed genes), especially centromeric heterochromatin (Bennett 2000). However, there is no evidence to suggest that these restrictions upon chromosomal localization are associated with any functional importance. Furthermore, because of its enormous size and restricted localization, satellite DNA is of no real use for either individual DNA profiling or genetic linkage studies (Bennett 2000).

Minisatellites are blocks of repeated core elements longer than 10 to 15 bp in tandem arrays. It can be subdivided into two types, the first of which is known as telomeric. Telomeric DNA consists of 10–15 kb of hexanucleotide repeats (mainly

TTAGGG), added to the telomeres of all chromosomes by the enzyme telomerase. Such DNA is most definitely functional in that it protects the ends of chromosomes from degradation and provides a means for the complete replication of telomeric sequences (Bennett 2000). It is also thought to play a role in the pairing and orientation of chromosomes during cell division (Bennett 2000). The second type of minisatellite sequences is hypervariable minisatellite DNA, and includes those first discovered in 1985 by Jeffreys and colleagues (Jeffreys *et al.*, 1985). The basic repeat unit may vary in length from six to > 50 nucleotides, with the overall number of repeats at any one locus usually being highly polymorphic between individuals (Bennett 2000).

The molecular processes that generate variability at minisatellites remain obscure. A subset of minisatellites share a common "core" sequence, which, based on sequence similarity to the Chi sequence of *E. coli*, has led to speculation that the core sequence may serve as a recombination signal to promote unequal crossing over, at least in human minisatellites (Jeffreys *et al.*, 1988a; Jeffreys *et al.*, 1985). Direct estimation of minisatellite mutation rates to new length alleles in human pedigrees has shown that paternal and maternal mutations arise with similar frequency, consistent with length change events being restricted to one stage of gametogenesis and possibly meiosis (Jeffreys *et al.*, 1988b). However, new mutants have been detected in clonal tumor cell populations (Armour *et al.*, 1989) and early in mouse development, indicating that mutation events are not restricted to the germline but can also arise in the soma (Jeffreys *et al.*, 1990). It is still not clear the extent to which interallelic unequal exchange through meiotic or mitotic recombination is involved in generating minisatellite variability, or whether other processes, such as unequal exchange between sister chromatids and replication slippage, are the primary source of minisatellite variation (Jeffreys *et al.*, 1990).

Minisatellite-type repeats appear to be ubiquitous in many antigen-encoding genes in *P. falciparum,* such as merozoite surface proteins (MSP-1 and MSP-2), S-antigen and circumsporozoite protein, and may be important in generating an immune response to malaria parasites (Hughes & Hughes 1995). This immune response generated to the repetitive regions, however, is thought to be a diversion by the parasite and is generally thought to be very poor at affecting the parasite. This is referred to as the 'smokescreen' hypothesis (Anders *et al.*, 1986). A few of these minisatellites have also been found in housekeeping genes, sequenced during the *P. falciparum* genome project (Gardner *et al.*, 1998).

Using data from 12 laboratory parasite lines together with field isolates from Papua New Guinea and the Democratic Republic of the Congo to analyse patterns of repeat structure, Anderson and colleagues observed numerous minisatellite/microsatellite combinations suggesting that minisatellites might be just as common as microsatellites in the *Plasmodium* genome (Anderson *et al.*, 2000b).

Microsatellites are defined as simple tandemly repeated DNA sequence elements of 2-6nucleotides as repeat units. They were first isolated by chance by Wyman and White (Wyman & White 1980) from a library of random segments of human DNA. Microsatellites are found in greater or less abundance in the genomes of just about every known organism and organelle (Bennett, 2000), and are sometimes referred to as short tandem repeats (STR) or simple sequence repeats. To denote the type of motif being iterated, terms such as mono-, di-, tri- and tetranucleotide repeats are used. Polymorphism of microsatellites is mainly due to allelic length variation, caused by a difference in the number of repeat units between alleles. Microsatellites therefore belong to a class of sequences termed variable number of tandem repeats (VNTR), which refer to any tandemly repetitive DNA that can show length polymorphism

(reviewed by Ellegren (Ellegren 2000)). PCR amplification and subsequent gel electrophoresis can determine the length of a distinct allele.

At present the function of microsatellite sequences is not clear. It is likely that a large proportion of microsatellites evolves neutrally, but in some cases, microsatellite variation can be associated with an altered phenotype (Chambers & MacAvoy 2000). The most direct evidence for the involvement of microsatellites in gene function comes from studies of human genetic disorders that result from trinucleotide-repeat expansion, in particular from the observed strong correlation between repeat length, age of onset and severity of disease (Rubinsztein *et al.*, 1999). This is exemplified by the trinucleotide repeat associated diseases such as Huntington's disease, myotonic dystrophy, and certain types of spinocerebellar ataxia (Brook *et al.*, 1992; McDonald 1994; Orr *et al.*, 1993; Schols *et al.*, 1995). The primary genetic cause seems to be the expansion of a trinucleotide repeat far outside of its "normal" polymorphic range (Debrauwere *et al.*, 1997). Such repeats are usually inside the disease gene, wherein most encode runs of glutamine residues; others, which are outside, are close enough to disrupt its correct functioning.

Microsatellites offer several advantages as genetic markers. These include the fact that they are relatively easy to isolate in many species (Armour *et al.*, 1994; Schlotterer *et al.*, 1991; Su & Wellems 1996), different loci can be used according to the level of variation, which range from very low to very high (Beaumont 1999), they can easily be amplified by PCR and thus can be used on a wide range of sample material for example blood, hair, saliva, skin and faeces, and their genetic systems are easily automated enabling the analysis of a large number of samples (Heyen *et al.*, 1997; Luikart & England 1999). Their applications range from estimation of the spatial

relationships between chromosome segments to the elucidation of temporal relationships between origins of species and genera (Chambers & MacAvoy 2000).

However microsatellites have several disadvantages. These include reports that for certain groups of organisms they are difficult to isolate (Beaumont & Bruford 1999), the tendency of *Taq* polymerase to add an additional dATP to PCR products, *Taq* polymerase-generated slippage products in some mono- and di-nucleotide microsatellite loci causing base shifts and sizing problems (Gill *et al.*, 1985; Ginot *et al.*, 1996; Schlotterer & Tautz 1992), the considerable technical challenges of microsatellite analysis of some types of samples e.g. saliva, hair or faecal material (Taberlet *et al.*, 1997; Taberlet & Fumagalli 1996) and the reports that data generated in different laboratories using different chemistries have proved difficult to amalgamate (Beaumont & Bruford 1999).

However the popularity of microsatellites as markers remains undiminished, probably because researchers are of the opinion that the advantages offered by the microsatellites, for example their resolving power, outweighs their disadvantages.

1.6 The Evolution of microsatellites

Microsatellites have been estimated to mutate at rates in the order of 10^{-3} to 10^{-4} changes per locus per generation, which is relatively high compared to the rate of point mutations (Hancock 1999). The high rates of change mean that the microsatellites differ from most other types of DNA sequences in their unusual degree of polymorphism, making them attractive as genetic markers (Goldstein *et al.*, 1999; Goldstein & Schlotterer 1999). The mechanisms by which microsatellites mutate are poorly understood but two main mechanisms have been proposed, which may act in concert.

Levinson and Gutman formulated the first model for microsatellite mutation about 15 years ago (Levinson & Gutman 1987). This model is based on the replication

slippage or slipped-strand mispairing. It refers to the out-of-register alignment of the two strands following dissociation at the time when the DNA polymerase traverses the repetitive region. If the 3' repeat unit of the nascent strand rehybridises with a complementary repeat unit downstream along the template strand, a loop will be formed in the nascent strand and the new sequence will become correspondingly longer than the template sequence upon elongation. Conversely, if the incorrect alignment occurs upstream along the template strand, the new strand will become shorter than the template strand (Ellegren 2000).

The other mechanism for microsatellite evolution is interchromosomal exchange (Ellegren 2000). In theory, new-length variants at repetitive DNA sequences can form through interchromosomal exchange, for example, in conjunction with recombination (unequal crossing-over) or gene conversion. However, the accumulated evidence is in favour of replication slippage, and includes:

- Mutant alleles are generally non-recombinant for flanking markers, which would indicate intrahelical mutation rather than recombination (Mahtani & Willard 1993; Morral *et al.*, 1991).
- The rate and pattern of microsatellite mutation does not seem to differ between hemizygote chromosomes (e.g. the Y chromosome) and chromosome pairs (autosomes of diploids) (Kayser *et al.*, 2000), suggesting that mutation events do not require contact between homologous chromosomes (Heyer *et al.*, 1997).
- It has been demonstrated from experiments done *in vitro*, that microsatellite sequences have the intrinsic ability to undergo DNA slippage (Schlotterer & Tautz 1992).

The mutation rates of microsatellites differ not only between species but also within species; that is, between loci. For most species studied, the average repeat number at

loci is directly proportional to the degree of length polymorphism indicating that long loci mutate more often than short loci (Weber 1990).

It has also been speculated that the flanking sequence content might affect the mutation rate e.g. through a simple GC content effect. Brock and colleagues showed that the %GC in the 100 bp region flanking the repeat sequence was positively associated with relative expandability of the microsatellite (Brock *et al.*, 1999). Another factor proposed to affect microsatellite mutation rate is the size difference between the two alleles present within an individual, denoted the 'allele span'. Amos and colleagues (Amos *et al.*, 1996) suggested that microsatellite mutations were more likely to occur in heterozygous individuals with a larger difference between the sizes of their two alleles. However, this idea is not supported by the evidence for mutations representing intrahelical events (Ellegren 2000).

The mutation patterns of microsatellites are assumed to evolve via two models; stepwise mutation model (SMM) or the infinite allele model (IAM). The SMM assumes that length changes result from either losses or gains of one or a few repeat units (Valdes *et al.*, 1993; Zhivotovsky & Feldman 1995). The SMM can be easily modelled and the length of repeat sequences contains information on the relationships between alleles. Evidence in support of this model comes from direct observation of mutations in human pedigrees (Weber & Wong 1993), experimental yeast systems (Weirdl *et al.*, 1997) and in tetranucleotide repeats of swallows (Primmer *et al.*, 1998). However, the problem with SMM is that it has little power to detect rare mutation events resulting in large changes in allele length. For instance SMM has failed to detect indels in flanking regions, which also play a significant role in generating length variation (Angers & Bernatchez 1997; Grimaldi & Crouau Roy 1997; Weber & Wong 1993).

The IAM model was developed to study loci that violate the SMM model (Kimura & Crow 1964). Under this model, a mutation involves any number of tandem repeats and always results in an allelic state not previously encountered in the population.

1.7 Microsatellites in malaria genetic studies

Microsatellites have been shown to be extremely widespread in *P. falciparum,* occurring every 2-3 kb throughout the genome (Su *et al.*, 1999; Su & Wellems 1996). A total of 901 markers have now been isolated and mapped to the genome of clone 3D7(Su *et al.*, 1999) . The most common repeat motifs seen in microsatellites have been those with an AT bias, which is not surprising given the extreme AT- richness of the parasite genome (Gardner *et al.*, 2002). Typical repetitive sequences are (TA)n, (CA)n, (A or T)n, (AAT)n and (CAA)n motifs (Su & Wellems 1996).

Microsatellite markers were first used to map a genetic cross. Su and Wellems (Su & Wellems 1996) described 224 microsatellite markers for *P. falciparum*, and used 188 polymorphic microsatellites to map the segregation patterns in the progeny from a genetic cross. They also investigated the variations in allele size present in 12 laboratory lines, from diverse geographical locations, at all 224 loci. From the prevalence of nonparental alleles in 35 progeny of this genetic cross, Anderson and colleagues (Anderson *et al.*, 2000a) estimated the microsatellite mutation rate for *P. falciparum* to be 1.59×10^{-4}. Anderson and colleagues (Anderson *et al.*, 2000b) used data from the same 12 laboratory parasite lines together with field isolates from Papua New Guinea and the Democratic Republic of the Congo to analyse patterns of repeat structure and expected heterozygosity at 114 of the same microsatellite loci. They also examined the sequence variation at 12 tri-nucleotide repeat loci. They found that there were complex mutation patterns in at least five of the twelve loci in the sequencing study. While alleles for some loci differed in size by one or two repeats, others were found to contain

compound repeats each of which showed length variation. Still others were found to have multiple sources of length variation including sequences of indels. Another finding from this study was the differences in allele sizes between the two populations studied. While alleles from Papua New Guinea were extremely conserved in size, parasite alleles from the Democratic Republic of Congo showed large size variations (Anderson *et al.*, 2000b). This was found to be as a result of various insertions, including minisatellites, in the microsatellite repeats of the parasites from Congo. In fact, some of the sequences observed in some loci could be best described as minisatellites containing microsatellite repeats. However, despite the complex repeat structures in these sequences, point mutations were found to be virtually absent from the flanking regions (Anderson *et al.*, 2000b).

These studies also suggested that the mutation rate of dinucleotide-bearing loci is 1.6-2.1 times faster than trinucleotides. The mean *He* (calculated as the proability of drawing two alleles of different length (for the same locus) from a population sample) over all loci was found to be 0.72 but was higher in di- than trinucleotide repeats. *He* was also found to be lower in seven loci that contained monomers or repeats >4 bp (Anderson *et al.*, 2000b). A positive relationship was also found between the repeat number and *He*, with *He* reaching a plateau around 13 repeats. This suggests that loci with large arrays of repeats have a higher mutation rate than loci containing small arrays (Anderson *et al.*, 2000b) until the maximum allele length is attained.

In perhaps the most comprehensive study on *P. falciparum* population genetics to date, Anderson and colleagues analysed the population structure in 465 infections from 9 locations worldwide using 12 microsatellites (Anderson *et al.*, 2000a). This study revealed an array of parasite population structures in different locations.

A study done in Senegal (Leclerc *et al.*, 2002) also revealed high mean heterozygosities for both microsatellites (0.745) and antigenic (0.702) loci but the values were lower than those obtained by Anderson and colleagues (Anderson *et al.*, 2000a) for the Africa region. This can perhaps be explained by different transmission intensities between the areas from which the samples were collected. Higher transmission rates would lead to higher recombination rates which would be expected to result in more parasites with different alleles (read higher *He*).

Some of the microsatellites used in the two studies that investigated parasite variation in field isolates (Anderson *et al.*, 2000a; Leclerc *et al.*, 2002) are situated in introns within protein coding sequences (loci denoted *2490*, *PfPK2*, *Polyα*, *G377* and *ARAII* (Anderson *et al.*, 2000b). These markers may therefore not be entirely neutral, since they are situated within protein-encoding genes subject to selection. Therefore, the values of *He* obtained may be lower than the true values as a result of the removal of genetic variation in these markers through "hitchhiking".

Hitchhiking refers to the process by which polymorphism of sites within the vicinity of a positive selection (selective sweep) is reduced (Hedrick 1980). This phenomenon has been observed on microsatellite loci around some genes of *P. falciparum*, *Pfcrt*, (Cortese *et al.*, 2002; Wootton *et al.*, 2002) *dhfr* and *dhps* (Nair *et al.*, 2003; Nash *et al.*, 2005; Roper *et al.*, 2003). In one of the studies with Southeast Asian isolates *He* of the microsatellite loci around *dhfr* was observed to reduce from 0.80 at markers situated 58 kb to 0.10 in markers situated 1.2 kb from the gene (Nair *et al.*, 2003). Two studies; one with parasites isolated from worldwide locations (Wootton *et al.*, 2002) and the other on Thai and Laos isolates (Nash *et al.*, 2005) made the same observation with markers around the *Pfcrt* gene. Though it is still quite difficult to unequivocally prove neutrality of a marker (Kreitman 1996; Rand *et al.*, 1994), it is

imperative to use markers that are as neutral as possible for any conclusive arguments on the population structure of *P. falciparum* to be made.

1.8 Drug resistance in malaria

The malaria parasite *P. falciparum* has developed clinically significant resistance to all classes of antimalarial drugs, with the possible exception of the artemisinin derivatives. Drug resistance is defined as the ability of a parasite strain to multiply or to survive in the presence of concentrations of a drug that normally destroy parasites of the same species or prevent their multiplication (World Health Organization 1973). Such resistance may be relative, yielding to increased doses of the drug that can be tolerated by the host (also referred to as drug tolerance); or complete, when the parasite withstands maximum doses tolerated by the host (outright resistance). It is usually demonstrated as a shift to the right of the dose-response curve, thus requiring higher drug concentrations to achieve the same parasite clearance (White 2004). Resistance emerges *de novo* through spontaneous mutations or gene duplications, which are thought to be independent of drug selection pressure. The mutated parasites have a selective advantage in the presence of the drug, and their frequency in the population will increase at a rate proportional to drug use (Mackinnon & Hastings 1998).

A number of different mechanisms are known to give rise to drug resistance including alterations of drug transport and permeability, conversion of the drug to an altered form with lower activity, increased expression of the drug target, and alterations to the drug target that decrease the affinity of binding of the inhibitor (White 1992). Different levels of drug resistance are encountered in malaria chemotherapy. Studies monitoring the spread of drug resistant malaria have tended to use various tools ranging from the standard *in vivo* method involving the examination of a blood film for parasites during the study period (WHO, 1996), to the more recent, parasite genotyping.

1.8.1 Therapeutic response and levels of drug resistance

The *in vivo* drug efficacy test is normally carried out following protocols published by the WHO (World Health Organization 2003b) and the therapeutic response is classified as outlined below. The first day of treatment corresponds to day 0; the days of follow-up are then counted from this date.

Early treatment failure (ETF):

- development of danger signs or severe malaria (see below for definition) on day 1, day 2 or day 3.
- axillary temperature ≥ 37.50C on day 2 with parasitaemia > that of day 0;
- axillary temperature ≥ 37.50C on day 3 in the presence of parasitaemia;
- Parasitaemia on day 3 ≥ 25% of the count on day 0.

Late treatment failure (LTF):

- danger signs or severe malaria in the presence of parasitaemia on any day from day 4 to day 14, [or 21, 28 or 42 depending on the study period] without previously meeting any of the criteria of early treatment failure;
- axillary temperature ≥ 37.50C in the presence of parasitaemia on any day from day 4 to any day within the follow-up period without previously meeting any of the criteria of early treatment failure.

Adequate clinical response (ACR):

- absence of parasitaemia on day 7, 14, 21, 28 or 42 irrespective of axillary temperature, without previously meeting any of the criteria of early or late treatment failure;
- axillary temperature < 37.50C irrespective of the presence of parasitaemia, without previously meeting any of the criteria for early or late treatment failure.

Adequate clinical and parasitological response

- absence of any malaria clinical signs and parasitaemia on any day within the follow-up period after treatment.

Treatment response is also classified according to the level of parasitaemia during the follow-up period using as a reference, the parasitaemia on day 0 (World Health Organization 1996).

RIII resistance: this is defined as day-3 parasite density ≥ 75% of the day 0 parasitaemia (Figure 3).

RII resistance: this is defined as day-3 parasitaemia equal to 25% of the day 0 density with the patient continuing to be parasitaemic on days 4-7 (Figure 3).

RI resistance: this is defined as day-7 parasite densities < 25% of day-0 density following either an initial parasite clearance on day 3 or presence of parasite density on day-3 at < 25% of day 0 (Figure 3).

Sensitive response: this defined as day 3 parasitaemia < 25% of that on day 0 [or absence of parasitaemia], a negative blood smear on day 7 and negative smears for the remainder of the follow-up period (Figure 3).

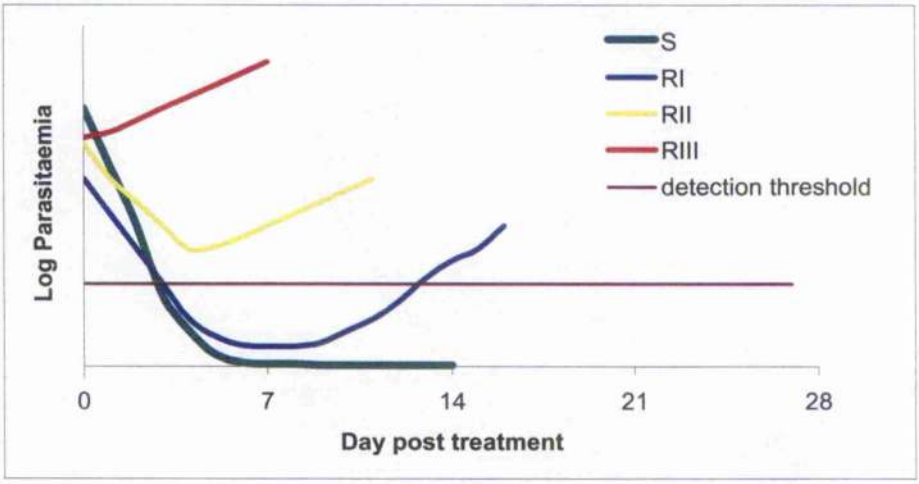

Figure 3: Graph showing parasite resistance after treatment over a 28-day period

One explanation for the different types of resistant parasite observed is that RI parasites may be the result of a low proportion of resistant parasites in a mixed infection. Drug treatment would kill the sensitive parasites, thereby reducing the resistant parasite population to a parasitaemia below that detectable by microscopy and possibly below the threshold for symptomatic malaria. These resistant parasites would continue to grow in the presence of the drug and eventually reach microscopically detectable parasitaemias (RI) or the threshold for symptomatic malaria to return (LTF) (Figure 3). An alternative explanation is that resistance is not an 'all-or-nothing' response. Some parasites could be highly resistant and their growth would be unaffected by the treatment; these would be classified as RIII resistant. Other parasites could show slightly lower resistance, such that their growth would be initially inhibited by the drug, but as concentrations fall they would be able to grow again; these would be RII. Finally, RI parasites could be resistant to the drug at a lower level still, such that growth is inhibited until the drug concentration drops further, although this level would still be able to prevent growth of parasites sensitive to the drug.

There is some correlation between clinical and parasitological response, but generally the rate of resistance/treatment failure is higher when measured by parasitaemia than by clinical response.

1.8.3 History of antimalarial drug resistance

The emergence of drug resistance became particularly evident during the Global Malaria Eradication campaign launched in 1955 by WHO (Bruce-Chwatt 1956). At the beginning, the campaign's strategy was mainly based on DDT indoor spraying; the insecticidal properties of DDT would interrupt malaria transmission by decreasing the survival of potentially infected mosquitoes. However, the first reports of insecticide resistance (to dieldrin) in *Anopheles gambiae* in Nigeria prompted the Second African

Malaria Conference held in Lagos in 1955 to stress the importance of obtaining the complete interruption of local transmission as quickly as possible (Bruce-Chwatt 1956). Therefore, the use of chemotherapeutic methods in association with residual insecticides was recommended whenever the rapid elimination of malaria was thought to be possible.

The introduction of cooking salt medicated with an antimalarial drug (mass drug administration (MDA)) was adopted for coverage of larger populations with antimalarials (Pinotti, 1953; World Health Organization 1961). It soon became clear that these interventions, besides the operational difficulties, could easily select for resistant parasites. In a few Kenyan villages, MDA with monthly pyrimethamine initially decreased the parasite prevalence but was followed by the emergence of parasites showing increased tolerance to the drug: sensitive parasites were rapidly replaced by resistant ones that disappeared when the administration of monthly pyrimethamine was stopped (Avery Jones 1958; Clyde & Shute 1954). A similar observation was reported from Ghana, although in this case it was explained by the irregular ingestion of tablets (Charles 1961; Charles *et al.*, 1962).

Interventions based on the introduction of pyrimethamine-medicated salt were implemented, among others, in The Netherlands, New Guinea, Brazil and Cambodia (Eyles *et al.*, 1963). Usually, despite an initial decrease, parasite rates returned to pre-operational levels within a period of 6 months. In New Guinea, the emergence of pyrimethamine resistance, unsatisfactory distribution of medicated salt and underestimation of the salt ration were identified as major causes for the campaign failure (World Health Organization 1961). Better results were reported with chloroquine (CQ)-medicated salt from Guyana (Giglioli *et al.*, 1967) and Uganda (Hall & Wilks 1967), where parasite rates decreased. In general, the widespread use of medicated salt

produced a wide variation of drug levels in the population because of the great variation in salt intake. This resulted in a very high selection pressure that induced an almost instantaneous *P. falciparum* resistance to pyrimethamine. CQ seemed to induce resistance less easily. However, the first cases of CQ-resistant strains originated from or near areas where CQ-medicated salt had been distributed (Payne 1988).

The first reports of confirmed *P. falciparum* resistance (RI) to CQ came, almost simultaneously, from South America (Colombia, Brazil, Venezuela) in 1960 (Moore & Lanier 1961; Wernsdorfer & Payne 1991) and from South-east Asia (Thailand, Kampuchea) in 1961 (Hartinuta, Migasen& Boonag, 1962). By 1973, CQ resistance had been reported in several countries in South America (Brazil, Colombia, Guyana and Venezuela) and in Asia (Burma, Cambodia, Malaysia, Philippines, Thailand and Vietnam) but not in sub-Saharan Africa (World Health Organization 1973). In Africa, *P. falciparum* CQ resistance was firstly reported from the eastern region, in Kenya (Fogh *et al.*, 1979) and Tanzania (Campbell *et al.*, 1979), in the late 1970s and it spread from east to west. By 1989, the distribution of CQ resistance was almost identical to that of *P. falciparum* (Wernsdorfer & Payne 1991).

Levels of resistance have been shown to vary widely between and within countries. In Kenya, for example, there were major differences in CQ resistance between the North, where malaria transmission is low, and the West and Southeast, where transmission is intense. In the early 1990s, CQ resistance (positive slide on day 14 after treatment) was about 18% around Lake Turkana (Clarke *et al.*, 1996) while in Kisumu, on the shores of Lake Victoria, it was around 70% (RII and RIII) (Bloland *et al.*, 1993).

This kind of situation within one country presents a real dilemma for health managers needing to decide whether and with which drug to change the country's antimalarial drug policy. Such a change constitutes a major undertaking that can take

several years before being fully operational. In optimal conditions the 'reaction time' has been estimated to be at least 2 years (Baudon 1995). Once the decision of changing the first-line drug is taken, there is no way back. However, it has always been difficult to give detailed guidelines or to establish a threshold at which a change should be implemented because this depends on several factors such as the cost of the drug and compliance.

1.8.4 Chloroquine resistance

Chloroquine (CQ) has been the mainstay of malaria treatment for well over half a century, but its exact mechanism of action remains controversial. Much evidence has pointed to the formation of a chloroquine-haem complex in the acid food vacuole of the parasite as being responsible for the antimalarial activity of chloroquine (Bray *et al.*, 1998; Chou *et al.*, 1980). The concentration of the drug at the site of action, the acid food vacuole of the parasite, is thought to be essential for its activity.

It has been known for some time that there is a reduced accumulation of chloroquine in chloroquine-resistant parasites (Geary *et al.*, 1986; Saliba *et al.*, 1998). Most studies have indicated a reduced rate of chloroquine uptake in resistant (CQR) compared to sensitive (CQS) strains. Several studies have suggested that this could be due to an elevated vacuolar pH, leading to reduced accumulation of chloroquine (Bray *et al.*, 1992a; Bray *et al.*, 1992b). Others have suggested that this is a consequence of CQ being actively transported out of the digestive vacuole (DV) away from its presumed site of accumulation in resistant parasites (Krogstad & Schlesinger 1987). The efforts to understand the molecular basis of CQ resistance with the goal of circumventing or suppressing it have identified two genes that have been shown to play a role in chloroquine resistance. These have been identified as *Pfcrt*, coding for the

chloroquine resistant transporter (PfCRT) (Fidock *et al.*, 2000a) and *pfmdr1*, (Foote *et al.*, 1989) coding for the multidrug resistance 1 genes.

Resistance to chloroquine resembles that of multidrug resistance (mdr) seen in mammalian cells: there is a reduced accumulation of the drug, and resistance can be reversed by the presence of verapamil (Geary *et al.*, 1986) (Krogstad *et al.*, 1987; Martin *et al.*, 1987). Multidrug resistance occurs when cells are exposed to one chemotherapeutic agent, but become resistant to a broad range of structurally unrelated drugs (Juliano & Ling 1976). The major protein mediating this phenotype in many mammalian cell-lines is the multidrug-resistance (*mdr*) transporter or P-glycoprotein (Pgh) (Gerlach *et al.*, 1986). Gene amplification, increased mRNA expression and mutation of the *mdr* genes have all been shown to confer resistance to various drugs across different cell lines (Roepe *et al.*, 1996). In many tumour cells the mdr phenotype has been shown to be dependent on the hydrolysis of ATP molecules, suggesting that Pgh may act as an energy-dependent drug efflux pump. Alternatively, it has been suggested that Pgh indirectly mediates substrate accumulation by regulating the plasma pH gradient and/or membrane potential (Roepe *et al.*, 1996).

Chloroquine–resistance in malaria parasites and multidrug resistance in cancer cell-lines share the features of reduced accumulation of the drug, and the ability to reverse this resistance with verapamil (Krogstad *et al.*, 1987). Therefore *P. falciparum* sequences corresponding to homologs of the mdr transporters that mediate multidrug resistance in mammalian cell-lines were identified and investigated for their possible role in chloroquine resistance (Foote *et al.*, 1989; Wilson *et al.*, 1989). Two homologues of *mdr* were found in the *P. falciparum* genome, and these were denoted *Pfmdr1* and *Pfmdr2* (Wilson *et al.*, 1989).

The *pfmdr2* gene, located on chromosome 14 and present as a single-copy gene, encodes a 110-kDa protein with structural similarity to the heavy-metal tolerance gene, *hmt1*, of the yeast *Schizosaccharomyces pombe* (Riffkin *et al.*, 1996). *Pfmdr2* contains only 10 transmembrane domains, demonstrating 26% amino acid homology with mammalian *mdr1*, and a single nucleotide-binding site (Rubio & Cowman 1994; Zalis *et al.*, 1993); a profile that is somewhat atypical for ABC transporters. Although some studies have reported increased transcription of *pfmdr2* in chloroquine-resistant parasites (Ekong *et al.*, 1993), other studies have demonstrated neither amplification, nor overexpression of *pfmdr2* (Rubio & Cowman 1994; Zalis *et al.*, 1993). To date, there appears to be no evidence to suggest that *pfmdr2* is involved in mediating drug resistance in *P. falciparum*.

1.8.4.1 *Pfmdr1* and chloroquine resistance

Pfmdr1 is a typical member of the ABC transporter superfamily, with a conserved structure of two domains consisting of six predicted transmembrane segments, coupled to a nucleotide-binding fold. The two domains are joined together by a linker region. The ~162 kD protein was termed Pgh1 (P-glycoprotein homolog 1) (Cowman *et al.*, 1991; Cowman & Karcz 1991; Higgins 1992). Pgh1 was subsequently localized to the parasite vacuole throughout the asexual cycle of the parasite, where it was postulated to regulate intracellular drug concentrations (Cowman *et al.*, 1991).

Sequence polymorphisms in *pfmdr 1* have been identified which are associated with chloroquine-resistant phenotypes *in vitro*. One polymorphism, exemplified by the K1 allele at codon 86, involves a single amino-acid change, N86Y, (Foote *et al.*, 1990a) while the other, exemplified by the 7G8 allele, involves four amino acid changes, Y184F, S1034C, N1042D, and D1246Y, the latter three being contained in transmembrane domain 11 (Foote *et al.*, 1990a). Molecular epidemiological analysis of

these polymorphisms in a double-blind study correctly predicted the chloroquine resistance profile of 34 of 36 isolates (Foote *et al.*, 1990a). Molecular epidemiological analysis of field isolates using rapid polymerase chain reaction and endonuclease restriction digestion techniques have demonstrated association of some of these polymorphisms with chloroquine resistance in studies from sub-Saharan Africa (Basco *et al.*, 1995), Nigeria (Adagu *et al.*, 1995a), Malaysia (Cox-Singh *et al.*, 1995), Guinea-Bissau (Adagu *et al.*, 1995b), and Thailand (Duraisingh *et al.*, 2000). Studies from the Sudan, (Awad-el-Kariem *et al.*, 1992) South America (Povoa *et al.*, 1998), and Southeast Asia (Basco *et al.*, 1996; Mungthin *et al.*, 1999), however, have failed to identify an association with these intra-allelic variations. In an attempt to establish linkage of chloroquine resistance to *pfmdr 1*, a genetic cross between a chloroquine-sensitive central American clone (HB3), and a multidrug-resistant Southeast Asian clone (Dd2) was performed (Wellems *et al.*, 1990). Chloroquine resistance was mapped as a Mendelian trait to a 36-kb locus on chromosome 7 (Wellems *et al.*, 1990), which segregated independently of the *pfmdr 1* locus found on chromosome 5 (Foote *et al.*, 1989; Foote *et al.*, 1990a; Foote & Kemp 1989). Amplification of *Pfmdr1* has also been suggested as a mechanism of resistance (Triglia *et al.*, 1991). When field isolates are analysed, both mutation and amplification are found to be widespread in numerous geographical areas (Basco *et al.*, 1995; Price *et al.*, 1999). Together these data suggest that mutations in *pfmdr1* can affect *P. falciparum* parasite response to CQ, but how big a role it plays in clinical resistance remains uncertain.

1.8.4.2 *Pfcrt* and chloroquine resistance

Analysis of the progeny arising from the HB3 x Dd2 cross mapped a single genetic locus, localized to a 36 kb segment on chromosome 7, that segregated with the inheritance of verapamil-reversible CQ resistance (CQR) (Su *et al.*, 1997; Wellems *et*

al., 1991). Subsequent studies of this segment identified nine open reading frames, with that denoted *cg2* having the closest linkage to the phenotype (Su *et al.*, 1997). Polymorphism in *cg2* was characterised by a complex series of repeat unit length differences and point mutations. Later studies used DNA transfection and allelic exchange to replace *cg2* polymorphisms in CQR parasites with those sequences from chloroquine sensitive parasites (Fidock *et al.*, 2000b) . Drug assays of the allelically modified lines showed no change in the degree of CQR. Similarly, a parasite (106/1), which contains the exact sequence of *cg2* from CQR parasites, was found to be CQS, and *cg2* haplotypes were found to be present within CQS parasites (Fidock *et al.*, 2000b).

Cg2 was ruled out as the candidate Chloroquine resistance gene. Re-examination of the 36kb segment on chromosome 7 revealed a highly interrupted gene, with 13 exons, some of which were very short. The gene now known as *Plasmodium falciparum chloroquine resistance transporter (Pfcrt)* spans 3.1 kb and encodes a 424 amino acid (48.6 kDa) protein (PfCRT) with 10 predicted transmembrane domains (Fidock *et al.*, 2000a). The protein was localized by immunofluorescence to the digestive vacuole (DV) membrane (Cooper, *et al.* 2002; Fidock, *et al.* 2000).

Transfection studies provided direct evidence of a role for PfCRT in CQ resistance. Coexpression of mutant forms of *Pfcrt* in the presence of a wild-type background conferred a modest degree of Verapamil-reversible CQ resistance to CQ sensitive (CQS) parasites (Fidock *et al.*, 2000a). More recently, a system was developed to replace the entire *Pfcrt* allele in a CQS parasite with *Pfcrt* alleles representative of mutant sequences prevalent in CQR parasites from Asia, Africa, South America and Papua New Guinea (Sidhu *et al.*, 2002). The resulting recombinant clones displayed reduced PfCRT expression as a result of the genetic modifications, but nevertheless

demonstrated all the hallmarks of a *bona fide* CQ resistance phenotype, i.e. increased CQ IC_{50} values (exceeding 100 nM), acquisition of VP-reversibility, and decreased CQ accumulation.

Mutations in the *Pfcrt* gene have been very closely associated with *in vitro* CQ resistance in culture-adapted isolates from around the world (Fidock *et al.*, 2000a; Wootton *et al.*, 2002). The critical mutation in CQR isolates results in a change of Lysine to Threonine at codon 76 (K76T). Other amino acid changes exist in flanking regions (72-75) with the allele present dependent on the geographical origin of the isolate. For CQR African and most South-east Asian strains, the resistance-linked allele encodes CVIET (amino acids 72-76), whereas CQS parasites have *Pfcrt* encoding CVMNK (Mehlotra *et al.*, 2001; Wellems & Plowe 2001). By contrast, most South American (and Papua New Guinea and Philippines) CQR strains encode SVMNT, while CQS strains have a CVMNK or related haplotype (Chen *et al.*, 2003; Durand *et al.*, 2001; Durrand *et al.*, 2004; Mehlotra *et al.*, 2001; Nagesha *et al.*, 2003; Wellems & Plowe 2001). A number of additional mutations are predicted to lie within the transmembrane segments. The recent discovery of additional *Pfcrt* mutations in Cambodian isolates (Lim *et al.*, 2003) brings the total number of variant residues to 15, making this quite a polymorphic gene for a non-surface molecule. It is presumed that many of these mutations serve to compensate for a loss of the endogenous function associated with the K76T variation, although some may serve to confer resistance to related antimalarial agents (Bray *et al.*, 2005).

Mutations in *Pfcrt* are believed to confer CQ resistance by reducing the amount of CQ accumulated by the parasite. The bulk of the intraparasitic CQ is believed to be concentrated within the digestive vacuole (DV). This, combined with the DV membrane location of PfCRT is consistent with the view that mutations in *Pfcrt* cause CQ

resistance by reducing the drug concentration in the DV (Saliba *et al.*, 1998). There are three main hypotheses on how PfCRT might exert this effect on DV CQ concentration. The first is that the reduced partitioning of chloroquine into the vacuole of CQR parasites is a consequence of alterations in the pH of the digestive vacuole (Bennett *et al.*, 2004). The second is that it is a consequence of the transport of CQ out of the DV (Saliba *et al.*, 1998) while the third is a consequence of conformational changes in the PfCRT as a result of mutations in the *Pfcrt* gene (Sanchez *et al.*, 2003). These hypotheses are discussed in the following sections.

1.8.4.2.1 Chloroquine resistance and pH of the digestive vacuole

One of the early hypothesis proposed that CQR parasites maintained an elevated pH in the digestive vacuole or DV, leading to reduced accumulation of CQ at steady state (Ginsburg & Stein 1991; Yayon 1985). Each molecule of CQ can associate with two protons, thus in theory leading to a concentration gradient across the DV membrane (de Duve *et al.*, 1974; Yayon *et al.*, 1984), it was thought that relatively small changes in pH should have a dramatic effect on the concentration of CQ in this organelle (Geary *et al.*, 1986).

The first reports of a comparison of pH in digestive vacuoles of CQS and CQR parasites suggested that CQR parasites have a more acidic DV than CQS parasites (Dzekunov *et al.*, 2000; Ursos *et al.*, 2000). However, these results were thought to be erroneous as they relied on measurements of fluorescence arising from acridine orange monomer, which was later shown to be undetectable in the DV and is, in fact, localized to the parasite cytoplasm (Bray *et al.*, 2002; Wissing *et al.*, 2002).

Roepe and colleagues (Bennett *et al.*, 2004) repeated their studies using the ratiometric dextran-tagged pH probe 5-(and 6-) carboxy-2′, 7′-dimethyl-3′-hydroxy-6′-*N*-ethylaminospiro [isobenzofuran-1 (3*H*), 9′-(9*H*) xanthen]-3-one (DM-NERF), which

again showed that CQR parasite lines exhibited a more acidic pH in the DV than CQS parasites. CQS parasites were reported to have a DV pH of around 5.7 and this was found to be decreased to around 5.2 in CQR lines, including lines in which the endogenous wild *Pfcrt* allele had been replaced by alleles from CQR lines (Bennett *et al.*, 2004). This study provided the first evidence of a direct link between CQ resistance-associated *Pfcrt* alleles and altered DV physiology. However, the reported differences in pH are difficult to reconcile with the CQ accumulation data. According to weak base theory, an increase of 0.5 in the pH gradient across the DV membrane in CQR parasites has the potential to produce up to a 10-fold increase in the uptake of CQ into the DV of CQR parasites compared to CQS lines. However, what is observed when the uptake is measured directly is a six fold decrease in the uptake of CQ into CQR isolates, relative to that in CQS isolates (Bray *et al.*, 1998; Sanchez *et al.*, 2003; Sanchez *et al.*, 1997).

1.8.4.2.2 Chloroquine resistance and efflux from the digestive vacuole

An alternative hypothesis for the reduced uptake of chloroquine by CQR parasites is that the drug is transported out of the DV, away from its presumed site of accumulation (Krogstad *et al.*, 1987). Krogstad and colleagues showed that addition of glucose to the medium reduced the steady-state accumulation of CQ by CQR parasites. By contrast, adding glucose to suspensions of CQS parasites markedly stimulated CQ accumulation (Krogstad *et al.*, 1992). These data suggest that CQS parasites have an energy-dependent CQ uptake mechanism (with energy being required both to maintain the DV proton gradient and to traffic and digest haemoglobin) and that CQR parasites have an energy-dependent CQ efflux mechanism.

However, these compelling data are inconsistent with some of the available literature. It has been shown previously that replacing glucose with its non-metabolizable analogue 2- deoxy-D-glucose, a procedure known to deplete the parasite

of ATP (Kirk *et al.*, 1996), has no effect on the steady-state accumulation of CQ by CQR parasites (Bray *et al.*, 1996). In the absence of ATP the DV H^+-ATPase fails and the H+ gradient across the DV membrane decreases markedly (Saliba *et al.*, 2003). Under these conditions there is very little possibility of active CQ efflux out of the DV. If CQR parasites were reliant on an active efflux process to reduce the CQ within the DV, then replacement of glucose with 2-deoxy-D-glucose would have been expected to result in a substantial increase in CQ accumulation. Other studies have shown that lowering the incubation temperature reduces the CQ accumulation of CQR parasites(Sanchez *et al.*, 1997). This is in contrast to the increased drug accumulation that might be expected at low temperature if an active efflux process were operating in the resistant parasites (Gottesman & Pastan 1993). These data suggest that the reduced CQ uptake in CQR lines that is observed upon addition of glucose (Sanchez *et al.*, 2003) might be unrelated to the concentration of ATP and, by implication, might not be caused by the active efflux of CQ from the DV.

1.8.4.2.3 The 'Charged drug leak' hypothesis

The mutations in *Pfcrt* that cause CQ resistance are associated with a loss of basic and hydrophobic residues (Warhurst *et al.*, 2002). These mutations (including the one for K76T) are predicted to be in the barrel of the transporter on the DV side of the membrane, in a region of PfCRT implicated in substrate selectivity (Martin & Kirk 2004). The presence of the positively charged lysine residue at position 76 in wild- type PfCRT is predicted to repulse the doubly protonated form of CQ (CQ^{2+}), and prevent its interaction with the transporter. The CQ resistance-conferring mutation at codon 76 resulting in threonine (or isoleucine or asparagine; (Cooper, *et al.*, 2002; Fidock, *et al.*, 2000) removes the positive charge, possibly allowing CQ^{2+} to interact with PfCRT and exit the vacuole down the steep outward CQ^{2+} concentration gradient. This would result

in a decrease in the overall CQ concentration at its site of action within the DV, and hence a decreased CQ sensitivity of the parasite. This 'charged drug leak' hypothesis provides a plausible mechanism by which PfCRT could directly mediate CQ resistance.

This hypothesis is supported by some data obtained from earlier studies on CQ resistance. It was shown that CQR culture-adapted field isolates are often hypersensitive to the anti-influenza drug amantadine (Evans & Havlik 1993) and to the phenanthrene methanol, halofantrine (Nateghpour *et al.*, 1993). Using a stepwise selection protocol, CQR lines were selected for increasing levels of resistance to amantadine or halofantrine. Surprisingly, as the parasites became resistant to these drugs, they became sensitive to CQ (Nateghpour *et al.*, 1993; Ritchie *et al.*, 1996; Ward 2002). Detailed examination of these parasite lines revealed that CQ accumulation and access to haematin were restored and the resistance-reversing effect of verapamil was lost (Johnson *et al.*, 2004). Both the amantadine-resistant line and the halofantrine-resistant line exhibited novel mutations of *Pfcrt* including one mutation (coding for S163R) that was common to both drug-selected lines. Surprisingly, both lines retained the *Pfcrt* K76T-coding mutation, making these the first fully characterised examples of CQS parasite lines with this mutation. The novel mutation common to both lines (resulting in S163R) is proposed to replace a positive charge in the transporter barrel, thereby compensating for the loss of positive charge associated with the K76T variation. This single amino acid change in PfCRT is proposed to block the leak of diprotonated CQ and thereby return the parasites to full CQS status, restoring access of CQ to ferriprotopophyrin IX and ablating verapamil synergism (Johnson *et al.*, 2004).

The debate on the mechanism of chloroquine and how its resistance arises is still ongoing. Nevertheless, a combination of *Pfcrt* and *pfmdr 1* polymorphisms together have been shown to result in higher levels of chloroquine resistance (Babiker *et al.*,

2001). This observation is also supported by a whole genome survey which, using microsatellite markers to detect linkage disequilibrium in a range of parasites, has shown that *Pfcrt* is the major locus (Djimde *et al.*, 2001; Fidock *et al.*, 2000a), but the *pfmdr 1* locus on chromosome 5 is also an important secondary mutation (Wootton *et al.*, 2002) for chloroquine resistance.

1.8.5 *Antifolates*

Antifolates are generally classified as type I (sulfonamides and sulfones) and type II (pyrimethamine, Trimethoprim and cycloguanil, the active metabolite of proguanil). They mimic the essential metabolites of the malaria parasite in the folate pathway, and are active against all the growing stages in the liver (liver schizonts), erythrocytic stages in the blood (blood schizonts) and growing stages in the mosquito (sporogonic stages) (Sherman 1979). In *Plasmodium falciparum*, two genes involved in the pyrimidine biosynthetic pathway, *dhfr* and *ts* are contiguous and are expressed as a bifunctional enzyme DHFR-TS (Bzik *et al.*, 1987). DHFR is the biochemical target of pyrimethamine, and also of cycloguanil, the biologically active metabolite of proguanil. Pyrimethamine binds to the parasite DHFR with a high affinity, inhibiting enzyme activity, resulting in the eventual depletion of deoxythymidine monophosphate (dTMP) and the disruption of DNA synthesis (Hyde 1990). The DHFR inhibition results in a reduction in the intracellular pool of tetrahydrofolate cofactors, which are used in most cells for the *de novo* synthesis of pyrimidines, methionine and thymidylate and for the interconversion of glycine with serine. Serine is the only identified source of methyl groups of methionine and thymidylate. Serine hydroxymethyltransferase converts serine to glycine with the formation of N5-N10-methylene tetrahydrofolate, which is then reduced to N-5-methyl tetrahydrofolate, the cofactor used to provide the methyl group that converts deoxyuridine monophosphate to the deoxythymidine monophosphate

required for DNA synthesis. This activity is catalyzed by the enzyme thymidylate synthase (TS) and results in the regeneration of dihydrofolate (DHF) (Figure 4).

Differences in folate metabolism between malaria parasites and their mammalian and avian hosts permit the use of antifolates as antimalarial agents. The selective activity of pyrimethamine has usually been attributed to higher affinity of the drug for *Plasmodium* DHFR-TS than for human DHFR (Foote *et al.*, 1990b; Hitchings 1969; Peterson *et al.*, 1990). Recent work has shed light on the differential specificity of the drug for parasite and human DHFR. Pyrimethamine treatment removes the translational repression of the human enzyme, and more DHFR is produced to partially overcome the drug barrier. However, in *Plasmodium* DHFR-TS, mRNA binding is not coupled to enzyme active sites. Therefore, antifolate treatment does not relieve translational inhibition and the parasites do not have the ability to make fresh enzyme (Zhang & Rathod 2002). Other authors (Nirmalan *et al.*, 2004) have however, reported lack of discernible effect of antifolate drugs on the level of mRNA expression from the *dhfr-ts* gene. Nirmalan and colleagues (2004) showed that expression levels of DHFR-TS are elevated three-fold by pyrimethamine, and up to seven-fold upon challenge by TS inhibitors, but not by a drug whose mechanism of action is unrelated to the folate pathway. They therefore concluded that the increases reflect enhanced synthesis of the DHFR-TS enzyme in the parasite as compared to human cells (Nirmalan *et al.*, 2004).

Studies on the sulfonamides and 4-aminobenzoic acid (PABA) in malaria date back to when the azo dye sulfachrysoidine and sulfanilamide were first used to treat human malarial infections and PABA was found to reverse sulfonamide inhibition in *Plasmodium gallinaceum* (Maier & Riley 1942).

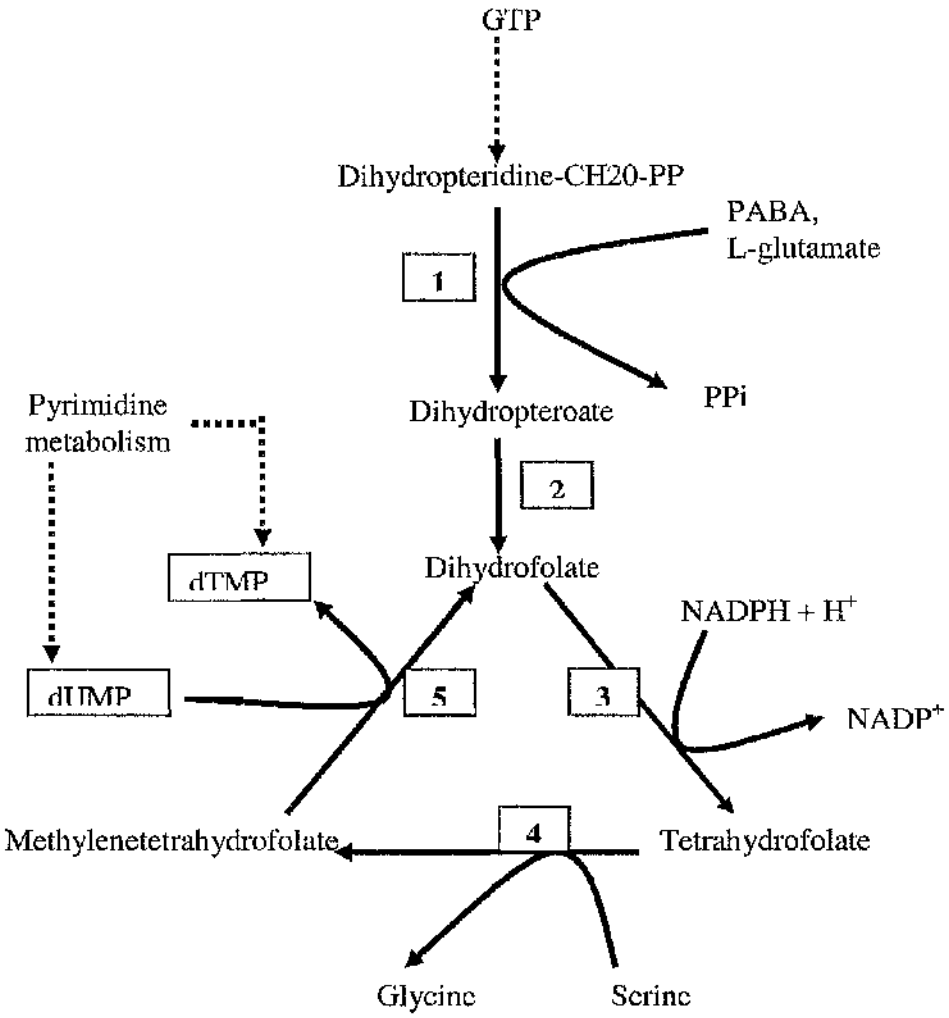

Figure 4: Simplified diagram of the Folate pathway indicating some the enzymes. The numbered boxes indicate the relevant enzymes 1. Dihydropteroate synthase, 2. Dihydrofolate synthase, 3. Dihydrofolate reductase, 4. Serine hydroxymethyltransferase, 5. Thymidylate synthase. Dashed arrows indicate multiple steps

The success of proguanil as an antimalarial agent in humans (Curd *et al.*, 1945) stimulated further study of pyrimidine derivatives, which resulted in the development of pyrimethamine (Falco *et al.*, 1951). The enzyme DHPS (dihydropteroate synthase) (Figure 4), which is not found in mammals, is the primary target of type I antifolates. DHPS has been shown to be a bifunctional enzyme with 6-hydroxymethyl-7, 8-

dihydropterin pyrophosphate (PPPK) at the N-terminus of the protein (Brooks *et al.*, 1994; Triglia & Cowman 1994). Expression and purification of the enzyme from *P. falciparum* in *E. coli* have shown that the activity purifies as an 83-kDa protein (Triglia *et al.*, 1997). DHPS is responsible for the synthesis of 7,8-dihydropteroate from 2-amino-4-hydroxy-6-hydroxymethyl-7,8- dihydropteridine pyrophosphate and (Para aminobenzoic acid) PABA (Figure 4). The drugs of the sulfa group are analogues of PABA and have been shown to be competitive inhibitors of the DHPS enzyme from the murine malarias *P. berghei* (Ferone 1973; Ferone 1977; McCullough & Maren 1974) and *P. chabaudi* (Ferone 1977; Walter & Konigk 1980), as well as in *P. falciparum* (Ferone *et al.*, 1970; Sirawaraporn *et al.*, 1990; Sirawaraporn *et al.*, 1997; Sirawaraporn & Yuthavong 1984; Triglia *et al.*, 1997; Zhang & Mcshnick 1991; Zolg *et al.*, 1989)

Sulfones and sulfonamides (sulfa drugs) were also used as antimalarials during World War II, but their use declined with the advent of more effective drugs such as chloroquine and pyrimethamine. The ability of sulfa drugs to show potentiating effect with proguanil or pyrimethamine (Chulay *et al.*, 1984; Greenberg *et al.*, 1948; Rollo 1955; Sirawaraporn & Yuthavong 1986) suggested that combinations of these two groups of drugs would not only be more effective antimalarial agents but would also delay, if not avoid, the development of resistance. The first combination used was the sulfone, dapsone, which was combined with pyrimethamine (Maloprim), but clinical trials were not encouraging and parasite resistance developed rapidly (Peters 1985). The most commonly used combination is that of a sulfonamide, sulfadoxine, and pyrimethamine (Fansidar), which has been particularly effective in areas such as Africa where chloroquine resistance is now widespread (Landgraf *et al.*, 1994; Peters 1985; Wernsdorfer & Kouznetsov 1980).

1.8.5.2 Mechanisms of resistance to DHFR inhibitors

Early studies done with *P. berghei* showed that it was possible to select resistance rapidly to pyrimethamine, suggesting that a single point mutation was involved (Bishop 1963; Diggens 1970; Diggens *et al.*, 1970). Genetic crossing experiments in both *P. chabaudi* and *P. yoelii* using pyrimethamine-sensitive and pyrimethamine-resistant strains showed that the resistance phenotype segregated as a single gene in both cases (Knowles *et al.*, 1981; Walliker *et al.*, 1976; Walliker *et al.*, 1975). This gene was confirmed to be *dhfr*, when it was shown that in *P. berghei* strains increased expression of the DHFR protein (Ferone 1970) and decreased binding affinity of the drug to the enzyme had occurred (Diggens 1970). In the case of *P. chabaudi*, no large differences in the enzyme amounts or turnover numbers were observed between resistant and sensitive clones. However, a large decrease in affinity for binding of pyrimethamine with the enzyme from the resistant clone, together with changes in kinetic and other properties, indicated that the resistance is due to genetic change leading to a structurally different enzyme (Sirawaraporn & Yuthavong 1984).

The major mechanism of resistance to pyrimethamine appears to be altered drug binding to DHFR. Studies of the enzyme kinetics of DHFR in the rodent malaria parasites, *P. berghei, P. vinckei,* and *P. chabaudi*, as well as studies in *P. falciparum* demonstrated that the inhibition constant (*Ki*) for pyrimethamine of the DHFR enzyme had increased in pyrimethamine-resistant strains and isolates (Chen *et al.*, 1987; Ferone 1970; Sirawaraporn & Yuthavong 1984; Walter 1986; Zolg *et al.*, 1989). One exception of this work was the observation that the DHFR activity from a pyrimethamine-resistant *P. falciparum* isolate had the same *Ki* as the DHFR enzyme from a sensitive isolate; however it was suggested that the enzyme was increased in the level of expression 30 – 80-fold (Kan & Siddiqui 1979). Other analysis of equivalent isolates has not shown any

altered levels of enzyme, and suggested that the DHFR enzyme was structurally altered (Cowman *et al.*, 1988; Peterson *et al.*, 1988a). However, these studies used crude extracts, making it hard to justify the assumptions of the Michaelis-Menten (MM) kinetics. The kinetic properties of DHFR enzymes purified from highly resistant, moderately -resistant and pyrimethamine-sensitive isolates of *P. falciparum* have been compared (Zolg *et al.*, 1989). DHFR enzymes extracted from the moderately resistant and highly resistant parasites were found to bind to pyrimethamine with 15 X and 500 X lower affinity respectively, when compared to the enzyme of the sensitive isolate.

The cloning of the gene encoding the DHFR enzyme from *P. falciparum* (Bzik *et al.*, 1987; Cowman *et al.*, 1988; Cowman & Lew 1989) and *P. chabaudi* (Cowman & Lew 1989) allowed a detailed analysis of alterations in this enzyme and their role in the mechanism of resistance. The DHFR enzyme had earlier been shown to co-purify with thymidylate synthase (TS) in *Crithidia fasciculata* (Ferone & Roland 1980) and the association of the DHFR and TS activities was confirmed in *P. falciparum* when it was shown that a single open reading frame encoded both activities (Bzik *et al.*, 1987; Cowman *et al.*, 1988; Sirawaraporn *et al.*, 1990), a property that is shared by most apicomplexan organisms (Garrett *et al.*, 1984; Grumont *et al.*, 1986).

Analysis of the alterations in the *dhfr* gene of experimentally-induced pyrimethamine-resistant *P. chabaudi* and *P. falciparum* strains has shown that both amplification and mutation of the gene can occur (Cowman & Lew 1989; Inselburg *et al.*, 1987; Tanaka *et al.*, 1990a; Thaithong *et al.*, 2001; Watanabe & Inselburg 1994; Cowman & Lew 1989; Inselburg *et al.*, 1987; Cowman & Lew 1989). Pyrimethamine resistance was selected for in both *P. chabaudi* (Cowman & Lew 1989; Cowman & Lew 1990) and in two independent experiments with *P. falciparum* (Inselburg *et al.*, 1987; Tanaka *et al.*, 1990b; Tanaka *et al.*, 1990a; Thaithong *et al.*, 2001) by slowly

increasing the level of drug. In one selected line of pyrimethamine-resistant *P. falciparum*, parasites exhibited increased expression of DHFR-TS, which correlated with a duplication of part of the chromosome containing the gene, which resulted in an increase in karyotype from 14 chromosomes to 15 chromosomes (Inselburg *et al.*, 1987; Tanaka *et al.*, 1990b). The selected lines of *Plasmodium chabaudi*, resistant to high levels of the pyrimethamine, showed rearrangement and duplication of a portion of chromosome 7 (Cowman & Lew 1989). This chromosomal duplication resulted in an increase in the chromosome number from 14 to 15; two derived chromosomes (450 kilobases and 1.1 megabases) were smaller than the original chromosome 7 (1.3 megabases), so that essentially only a 200-kilobase region was duplicated. This region contained the *dhfr-ts* gene and the closely linked *Hsp70* gene. The duplication of the DHFR-TS gene presumably conferred the observed drug resistance (Cowman & Lew 1989). Cowman and Lew extended this analysis by selecting a highly pyrimethamine resistant line of *P. chabaudi* from the line that had duplicated and rearranged chromosome 7. This resulted in reversion back to the original parental type chromosome 7 but a point mutation was found to have occurred in the *dhfr* gene (Cowman & Lew 1990). Increased selection for high levels of pyrimethamine resistance resulted in the selection of parasites carrying a mutated *dhfr* gene, encoding a protein where the amino acid at position 108 was replaced with asparagine (Cowman & Lew 1989). The second *P. falciparum* selected line was found to have amplification of the DHFR gene up to a copy number of 44, correlating with a 1000X increase in minimum inhibitory concentration (MIC) of pyrimethamine (Thaithong *et al.*, 2001). The resistance was unstable, and following withdrawal of pyrimethamine, MIC levels dropped over 16 months almost to that of the unselected parental line, accompanied by a decrease in DHFR copy number to 4. No mutations in the *dhfr* gene were observed.

The mutation at codon 108 was subsequently also found in naturally occurring pyrimethamine-resistant *P. falciparum* parasites collected from the naturally infected individuals. Peterson and colleagues examined the molecular basis of pyrimethamine resistance in isolates and clones of *Plasmodium falciparum* from geographically distant sources (Peterson *et al.*, 1988a). A genetic cross between pyrimethamine-resistant clone HB3 and and pyrimethamine-sensitive clone 3D7 showed absolute linkage of resistance to the point mutation (Asn-108) in *dhfr* (Peterson *et al.*, 1988a).

Peterson and colleagues examined the *in vitro* pyrimethamine susceptibilities and DHFR sequences of eight additional isolates and clones from geographically distant locations (Peterson *et al.*, 1988a). All of the pyrimethamine sensitive parasites had Thr-108 (ACC) or Ser-108 (AGC), whereas all of the resistant parasites exhibited the point mutation to Asn-108. An additional mutation to Ile-51 (AAT to ATT) was associated with a 4- to 8-fold increase in resistance in some lines, and a third mutation from Cys-59 to Arg-59 (TGT to CGT) in one clone was associated with even higher levels of resistance (Peterson *et al.*, 1988a).

Polymorphic sequences of *P. falciparum dhfr* were first described in the late 1980s (Cowman *et al.*, 1988; Peterson *et al.*, 1988a; Snewin *et al.*, 1989), and it is now well established that high-level Pyrimethamine resistance results from the accumulation of mutations in the *dhfr* gene, principally at codons 108, 59 and 51, where allelic polymorphism gives rise to S108N, N51I and C59R (Table 1). Evidence also exists that parasites with additional mutations of 164 to Leu (I164L) have increased levels of resistance to Pyrimethamine and Cycloguanil (Cowman *et al.*, 1988; Peterson *et al.*, 1988a). Variation at residue 16 (Ala-16 to Val A16V) confers resistance only to cycloguanil and the variation is always found to be associated with S108T variation (Foote *et al.*, 1990b; Peterson *et al.*, 1990). Analysis of the single A16V and S108T

variations revealed that the former is highly resistant to cycloguanil without cross resistance to pyrimethamine, while the latter is susceptible to both Pyrimethamine and cycloguanil, indicating that the cycloguanil resistance in the A16+SI08T must be attributed almost completely to the A16V mutation (Sirawaraporn *et al.*, 1997). Therefore, it is conceivable that the A16V+SI08T mutant would not have survived pyrimethamine pressure, but instead must have arisen in response to cycloguanil selection.

Kinetic studies of the different alleic variants of DHFR have revealed that the A16V protein has a severely-impaired enzyme kinetic (V_{max}), which could have significant effects on the enzyme catalytic efficiency to support DNA synthesis, rendering the mutant unable to survive in nature under selective drug pressure (Sirawaraporn *et al.*, 1997). The S108T enzyme shows almost unperturbed V_{max} kinetics (Sirawaraporn *et al.*, 1997). The enzyme is as inhibited by both pyrimethamine and cycloguanil as the wild-type enzyme, so it is very likely that the A16V+S108T mutant must have derived from the S108T, a variant which is rare (Reeder *et al.*, 1996), reviewed in (Hyde 2002; Hyde 2005).

The two double mutants found in nature (N51I/S108N and C59R/S108N) could have arisen from point mutation of the S108N as they both provide a small increase in resistance toward pyrimethamine, which would have led to their selection (Siriwaraporn *et al.*, 1997). The N51I/S108N mutant enzyme had almost similar catalytic constant, k_{cat}, to the wild-type DHFR but was about 2 to 3-fold more resistant to pyrimethamine than S108N. C59R/S108N had a significantly reduced k_{cat} compared to wild-type and S108N DHFR, but was 5-fold more resistant to pyrimethamine than S108N.

The triple mutant N108S/N51I/C59R had a significantly reduced k_{cat} compared to the wild-type DHFR, but its inhibition constant, K_i, by pyrimethamine was

approximately 80-fold that of the wild-type DHFR (Sirawaraporn *et al.*, 1997). This would offer it a huge selective advantage over the wild-type in a pyrimethamine environment. The same was observed for the quadriple mutant S108N/N51I/C59R/I164L whose k_{cat} was almost six-fold lower than that of wild-type DHFR but had a K_i by pyrimethamine about 600-fold higher than that of the wild-type DHFR (Sirawaraporn *et al.*, 1997).

The mutations appear sequentially in treated populations, with the serine-asparagine mutation at codon 108 appearing first, followed by asparagine- isoleucine (codon 51) or cysteine-arginine (codon 59), and finally isoleucine-leucine (codon 164) (Plowe *et al.*, 1998). In all parasite isolates studied to date, only the S108N mutation has been found to occur singly in nature. The mutations at codon 51 and 59 are always found to occur together with S108N (as N51I + S108N, C59R + S108N or triple mutant) but never singly. In fact, the enzymes from genes containing these single mutations were found to remain highly sensitive to both pyrimethamine and cycloguanil, with one (N51I) also showing significantly reduced catalytic properties (Sirawaraporn *et al.*, 1997). This has led to the suggestion that mutations either or both at codons 51 and 59 act predominantly by restoring the enzymatic defects that occur as a consequence of the original mutation at position 108 (Hastings & Donnelly 2005).

Infections with parasites carrying single (Asn108) or double mutations in *dhfr* (encoding Asn 108 plus either Ile 51 or Arg 59) show increased parasite clearance times following treatment with SP than wild-type parasites (Mendez *et al.*, 2002; Su & Wellems 1996), and parasites with three or four resistance mutations are refractory to treatment (Plowe *et al.*, 1998). Some of the isolates with various combinations of mutations at *dhfr* and their levels of pyrimethamine resistance are shown in Table 1.

Codons / Isolates	16	51	59	108	164	Level of Pyrimethamine resistance [a]
3D7, SLD6, T9/96	A	N	C	S	I	0.025
HB3, PA2	A	N	C	**N**	I	1.5
FAC3, FAC8, ItG2F6, FCB	**V**	N	C	**T**	I	0.006
7G8, ItD12	A	**I**	C	**N**	I	12
K1, V1	A	N	**R**	**N**	I	20
W2, Dd2, NT1	A	**I**	**R**	**N**	I	25
CsI 2	A	N	**R**	**N**	**L**	40
V1/S, VP35	A	**I**	**R**	**N**	**L**	40

Table 1: Common laboratory isolates with their variant amino acids and IC50 to Pyrimethamine *in vitro*.

The variant conferring resistance is indicated in bold.
[a]The figures indicated for levels of resistance are approximate values for the concentration of pyrimethamine (μM) required to reduce parasitaemias by more than 50% relative to untreated controls (Hyde 1989).

Another polymorphism has also been observed in *dhfr* sequences of samples obtained from Bolivia, which is an insertion coding for a five-amino acid repeat after codon 30, termed the Bolivia repeat (Plowe *et al.*, 1997). However this polymorphism has only been observed in Bolivia and its role in pyrimethamine resistance is not yet known.

1.8.5.3 Mechanisms of action and resistance to DHPS inhibitors

Comparisons of *dhps* alleles from lines of *P. falciparum* with different IC50 values to sulfadoxine *in vitro* pinpointed amino acid differences suggested to be important in resistance (Brooks *et al.*, 1994; Triglia & Cowman 1994). There was no apparent amplification of the gene or increased expression of the PPPK-DHPS enzyme in sulfadoxine-resistant isolates.

Polymorphisms in codons 436, 437, 581 and 613 in the *dhps* domain of cultured parasite lines have been broadly correlated with estimated levels of sulfadoxine resistance *in vitro* (e.g. (Brooks *et al.*, 1994; Triglia & Cowman 1994). Later studies revealed that polymorphism in codon 540 was also commonly observed in field samples (e.g. (Plowe *et al.*, 1997; Triglia *et al.*, 1997; Wang *et al.*, 1997a)), with some association to SP treatment failure. Trials in East Africa and elsewhere attested to the enhanced effectiveness of the SP combination (Nguyen-Dinh *et al.*, 1982) compared with monotherapy with either agent, even where pyrimethamine resistance was already prevalent. Sulphadoxine is never used as a single antimalarial as it is quite ineffective and therefore all field data is from SP treatment failures. To date, only changes in these five codons have been observed in a large number of field samples of diverse geographical origins, (e.g. (Plowe *et al.*, 1997; Reeder *et al.*, 1996). Codon 436 appears to be tetramorphic, with alternative alleles encoding alanine, phenylalanine, cysteine, or serine. S436 is usually associated with sensitivity to SP and for the purposes of this thesis it is considered as the wild type allele. Three alternative amino acids are encoded by different alleles at codon 613: serine, threonine and alanine. Codons 437, 540 and 581 appear to be dimorphic, with codons 437 and 581 both encoding either alanine or glycine, while codon 540 encodes either lysine or glutamine (Table 2).

The most common polymorphism in *dhps* is the one coding for 437G. Of the 13 fully characterized resistance-associated alleles described to date, 10 involve this polymorphism, although parasites with DHPS 437G in the absence of other changes appear to be rare. Alleles defined by 437G/540E and 437G/581G are common in SP resistant parasites from East African patients (Nzila *et al.*, 2000a; Plowe *et al.*, 1997; Wang *et al.*, 1997a). The allele encoding 437G/581G is also common overall in South-East Asian populations (Wang *et al.*, 1997a), while in areas of South America, parasites

from patients with SP treatment failure are frequently found to have all three of these mutations (encoding 437G/540E/581G) (Kublin *et al.*, 1998; Plowe *et al.*, 1997).

While the mutations affecting codons 436, 437 and 540 can occur singly within an allele, the data so far suggest that the 581G variation is always found with 437G, and that the 613S/T alterations must also be coupled to changes in either residue 436 or 437, presumably reflecting steric constraints of the enzyme (Plowe *et al.*, 1998). However, this has not been confirmed by any kinetic data due to the difficulties of working with sulfadoxine *in vitro*.

The role of *dhps* mutations in sulfadoxine resistance is not certain, partly because *in vitro* assays for sulfadoxine have been done under varying folate conditions (Wang *et al.*, 1997b), and also the use of host folate by some parasite isolates antagonises sulfa drugs irrespective of their *dhps* genotype (Milhous *et al.*, 1985; Watkins *et al.*, 1985). Analysis of a genetic cross between two parasites differeing in IC50 to sulfadoxine showed direct correlation of inheritance of *dhps* allele and IC50 (Wang *et al.*, 1997b). Consistently reproducible results showing large differences between the most sensitive and most resistant strains were obtained as a result of using an improved drug assay in the absence of folate (Triglia *et al.*, 1997; Wang *et al.*, 1997b).

Codons / Isolates/origin	436	437	540
3D7 (Netherlands)	Serine	Glycine	Lysine
Dd2 (IndoChina)	Phenylalanine	Glycine	Lysine
T9-96 (Thailand)	Alanine	Glycine	Lysine
SL/D6 (Sierra Leone)	Phenylalanine	Alanine	Lysine
HB3 (Honduras)	Serine	Alanine	Lysine
IEC513/86 (Brazil)	n.d.	n.d.	Glutamine

Table 2: DHPS variants in common laboratory isolates.
n.d.= not determined

Another mechanism associated with folate uptake and utilisation is thought to have an effect on sulfadoxineresistance levels (Wang *et al.*, 1999). A reduction in the susceptibility of parasites to sulfadoxine inhibition as a result of low levels of folate (folate effect) was displayed in the Dd2 (mutant) clone, producing marked changes in IC_{50} values. A very small change or no change in the IC_{50} values was observed for the HB3 clone under varying folate concentrations. This difference is thought to be due to another gene that strongly influences resistance of a specific clone to sulfadoxine depending on the usage of exogenous folate.

1.8.5.4 Correlations of *dhfr* and *dhps* mutations with SP treatment response

Wang and colleagues made the first attempt to define a set of *dhfr* and *dhps* mutations predictive of SP treatment (Wang *et al.*, 1997a). Using samples for SP drug efficacy trials, they found that all infections that cleared totally had wild type *dhfr* and *dhps* sequences. Those that cleared initially but recurred, or did not clear at all carried a

variety of *dhfr* and/or *dhps* mutations, with no clear set of mutations accounting for initial or no clearance (Wang *et al.*, 1997a).

Several studies have shown that treatment with antifolate drugs results in the selection of parasites with *dhfr* and *dhps* mutations. The study by Wang (Wang *et al.*, 1997a) described 18 previously examined Tanzanian samples taken on the day of SP treatment and 13 that were again positive during the 28 days following treatment (Wang *et al.*, 1997a). Even with this relatively small sample size, analysis of the data shows that *dhfr* coding for 108N, and *dhps* coding for 437G and 581G were significantly more prevalent in the post-treatment infections (Plowe *et al.*, 1998), and *dhfr* encoding 51I and 59R were also more common post-treatment. A second study in Tanzania found that SP selected for *dhfr* encoding 51I, 59R and 108N as well as *dhps* encoding 436S (thought to be wild-type), 437G and 540K (Curtis *et al.*, 1998).

A study in Peru provided the first evidence of a distinct set of *dhfr* and *dhps* mutations that is strongly associated with *in vivo* SP resistance (Kublin *et al.*, 1998). The genotype *dhfr* coding for 51I, 108N, 164L and *dhps* coding for 437G, 540K, 581G was highly correlated with the level of *in vivo* resistance. This resistant genotype was present in seven out of eight cases (87.5%) of RIII resistance, nine out of 13 cases (69.2%) of RII resistance, 5 out of 13 cases (38.5%) of RI resistance, and 0/11 cases (0.0%) sensitive to SP treatment. The Bolivia repeat was less frequent at higher levels of resistance, consistent with *in vitro* evidence that it does not play a direct role in resistance to pyrimethamine (Cortese & Plowe 1998). *Dhfr* coding for 108N was present in all sensitive and resistant cases, demonstrating clearly that this mutation is not a suitable marker for SP resistance, despite having been used for this purpose in several studies (Curtis *et al.*, 1996; Edoh *et al.*, 1997; Peterson *et al.*, 1991; Plowe *et al.*, 1996). A drug-sensitive genotype consisting of DHFR Asn-108 and no other *dhfr* or

dhps mutations was present in all sensitive infections, 61.5% of RI infections, 23.1% of RII infections and in only a single RIII infection (Cortese & Plowe 1998).

It is important to emphasize that the resistance of parasites to drugs *in vitro* is not synonymous with clinical failure of pyrimethamine–sulfadoxine treatment. This is especially clear in endemic areas in which semi-immune individuals compose a majority of the population at risk. For example, on the coast of Kenya, the triple-mutant allele of *dhfr* is observed in ~90% of the parasites but SP treatment failure rates are about 40%. Double-mutant alleles of *dhps* are also common (Nzila *et al.*, 2000a). Semi-immune patients who carry parasites with resistant genotypes very often do resolve any fever or other clinical manifestations. These patients are classified as having an adequate clinical response to the therapy (no visible parasites or clinical manifestations at day 7), even though some resistant parasites might escape from the drug action, remain at a low level for a time and then cause a new malaria episode (Jelinek *et al.*, 1998; Plowe *et al.*, 1998; White 1998).

The relationship of the parasite genotype to the outcome of pyrimethamine–sulfadoxine treatment in an individual patient is influenced by many factors; the genotype of the parasite is only one determinant. The complexity of the parasite infection and the genotype, the nutritional status and the immune response of the host, as well as the rates of drug metabolism, are all critical in determining whether a patient will clear the infection (Plowe *et al.*, 1998; White 1998). In the following section, a discussion is offered on the determination of 'true' drug resistance.

1.8.6 Molecular assays for determination of 'true' drug resistance

Until recently, our knowledge of the epidemiology of drug-resistant malaria was based on the collection of *in vivo* data from symptomatic patients to whom different antimalarial drugs were administered and, to a lesser extent, on *in vitro* drug sensitivity

assays. The limitations of these methods for studying drug-resistant malaria and elucidating molecular mechanisms of resistance to some antimalarial drugs have stimulated the use of a third approach based on molecular markers (Greenwood, 2002). The standard *in vivo* method of testing for drug resistance involves the examination of a blood film for parasites 14 or 28 days after an infection is diagnosed and treated, the detection of parasites at this time being considered an indication of treatment failure or recrudescence of drug resistant parasites (WHO, 1996). In high transmission areas such as tropical Africa, most people are bitten by malaria-infected mosquitoes almost every night (Greenwood 2002). Parasites found in the peripheral blood during the follow-up period could therefore, represent new and possibly drug-sensitive infections, and not recrudescence of resistant parasites. Estimates of the frequency of drug resistance based on uncontrolled tests *in vivo* in highly endemic areas may therefore be exaggerated. In order to determine true drug resistance levels, it is important to differentiate between true recrudescence (treatment failure) and reinfection.

Molecular genotyping is increasingly being used to help distinguish recrudescence from reinfection in antimalarial drug efficacy studies. Infecting malaria parasites can be "fingerprinted" through polymerase chain reaction (PCR) amplification of polymorphic genes. The "fingerprint" patterns of isolates causing successive episodes of malaria can then be compared to distinguish recrudescent from newly infecting parasites (Babiker *et al.* 1994; Ohrt *et al.* 1997; Ranford-Cartwright *et al.* 1997). The technique involves amplification by the polymerase chain reaction (PCR) of regions of highly polymorphic parasite genes. The likelihood of a patient being reinfected with a parasite containing exactly the same 'alleles at each of 2-3 loci is very small. If the parasites present on follow-up are genetically identical at all loci to the parasites found on presentation then it is assumed that the infection has recrudesced and that persisting

parasites are resistant to the drug. If the parasites are genetically different, it is assumed that the original parasites were cleared by the drug treatment, and the parasites present during the follow-up period are the result of a re-infection. The loci commonly employed in this work include the *Msp1*, *Msp2* and *glurp* (Babiker *et al.*, 1994; Babiker & Walliker 1997; Brockman *et al.*, 1999; Ranford-Cartwright *et al.*, 1997).

There is abundant evidence that these loci are under rather strong immune selection. A comparison of ratios of synonymous and non-synonymous mutations in various genes reveals that strong selective constraints against silent variation in *P. falciparum*, a fact that has been attributed to immune selection (Escalante *et al.*, 1998; Hughes 1991; Hughes & Hughes 1995). Additional evidence that these markers are under strong immune selection comes from experimental immunology studies. For *Msp1*, vaccination with the recombinant protein was shown to be protective in monkeys (Etlinger *et al.*, 1991), while antibodies against it protect against clinical disease (Egan et al. 1996). Antibodies to MSP2 have been shown to inhibit parasite invasion of red blood cells (Epping *et al.*, 1988), and are associated with protection from clinical disease (Al Yaman *et al.*, 1994; Al Yaman *et al.*, 1995; Taylor *et al.*, 1998). This protection appears to be dependent on MSP2 genotype (Al Yaman *et al.*, 1994; Ranford-Cartwright *et al.*, 1996).

Recently, after the discovery of microsatellites, some malariologists have started to question epidemiological data generated by the use of antigenic loci under strong immune selection. The argument is that selected loci might be giving a distorted view of population structure and transmission dynamics, since selection rather than population history may determine the patterns of allele distribution within populations for these loci (Anderson *et al.*, 1999; Hastings 1996; Mcdonald 1994).

In this thesis the antigen-coding loci; *Msp1, Msp2* and *glurp* have been compared with five putatively neutral microsatellite loci in their ability to differentiate recrudescences from reinfections in *P. falciparum* infections from Kenya.

1.9 Drug resistance and selective sweeps

When a beneficial mutation spreads through a population, flanking neutral mutations 'hitchhike', resulting in removal of genetic variation from the chromosomal regions surrounding the selected site (Barton 2000; Maynard Smith & Haigh 1974). The extent to which this happens depends on when in the history of the sample population the selected substitutions happen, the strength of positive selection and the amount of crossing over between the flanking locus and the locus at which the selected substitutions occur (Kaplan *et al.*, 1989). The spread of the favoured allele also results in increased linkage disequilibrium (LD) with flanking markers (Sabeti *et al.*, 2002; Tishkoff *et al.*, 2001) and skews in the allele frequencies observed at loci nearby on the chromosome (Payseur *et al.*, 2002). The size of genomic regions affected is influenced by the strength of the selection, as well as the rates of recombination and mutation. There is currently great interest in using such characteristic patterns of variation to identify regions of the genome that are under selection (Harr *et al.*, 2002; Kim & Stephan 2002; Kohn *et al.*, 2000; Schlotterer 2003; Vigouroux *et al.*, 2002; Vigouroux *et al.*, 2005).

Antimalarial drug resistance provides an excellent system for investigating the genomic effects of selection events in a recombining eukaryote, and the scars left in the *Plasmodium falciparum* genome by drug selection may serve as a useful tool for locating drug resistance genes (Anderson *et al.*, 2000a; Wootton *et al.*, 2002). Fortunately the date of introduction of new antimalarials is usually known with some accuracy, allowing the rate of spread of resistance to be calculated.

Recombination and mutation rates have been estimated from a genetic cross (Mu *et al.*, 2002; Su *et al.*, 1999), and haplotypes can easily be constructed using haploid blood stage parasites. This aids in the estimation of the rates of spread of resistance through calculations of selection intensities driving resistance to the drugs (Wiehe, 1998).

The repeated appearance throughout global malaria endemic regions of drug-resistant phenotypes, determined by nonsynonymous substitutions at the *Dhfr*, *Dhps*, *Pfcrt* and other loci, is most likely due to natural selection (Roper *et al.*, 2004; Wootton *et al.*, 2002). Selection sweeps are known in other organisms, such as *Drosophila melanogaster*, where a single nucleotide sequence at the *Sod* locus is present in about 50% of all haplotypes throughout the world, without any silent substitutions along the 1,500-bp sequence (Hudson *et al.*, 1997; Hudson 1994).

The worldwide geographical prevalence of the alleles of *dhfr* and *dhps* associated with SP treatment failure reflects both the duration of antifolate use and the level of therapeutic resistance. However, at the time the research for this thesis was begun, the modes by which antifolate resistance develops had not been elucidated. For example it was not clear whether resistance mutations could repeatedly arise *de novo* or spread by dissemination after rare or infrequent mutation events (Cortese & Plowe 1998; Nair *et al.*, 2003; Nash *et al.*, 2005; Pearce *et al.*, 2005; Roper *et al.*, 2003; Roper *et al.*, 2004). Molecular analysis of the chloroquine-resistance transporter (*Pfcrt*) gene have suggested that all extant CQ resistant parasites originate from only four separate mutational events, two of which are believed to have occurred in South America, one in Southeast Asia and one in Papua New Guinea (Wellems & Plowe 2001). These mutant alleles are thought to have subsequently spread over large geographical distances(Wootton *et al.*, 2002). Interestingly, it appears that CQ resistance did not arise

de novo in Africa, where most *P. falciparum* mortality occurs, despite wide-scale deployment of the drug (Wellems & Plowe 2001). Analysis of the *Pfcrt* gene and flanking markers suggests that CQ resistant parasites in Africa are most similar to those in Southeast Asia, suggesting the spread of resistance from countries in this area. The usual explanation for resistance occurring so rarely is that CQ resistance is a complex trait, requiring several sequential mutations in the *Pfcrt* gene (Wellems & Plowe 2001) and, arguably, requiring mutations in other genes (Hastings 2003), perhaps such as *pfmdr1* (Babiker *et al.*, 2001). By contrast, the genetic basis of resistance to SP is known in detail, and simple logic leads to the *a priori* expectation that SP resistance originates from numerous mutational events. Resistance arises extremely rapidly after the introduction of pyrimethamine treatment (Clyde & Shute 1957; Doumbo *et al.*, 2000; Molineaux & Gramiccia 1980) and the mutations involved can be selected readily in the laboratory (Paget-McNicol & Saul 2001). Thus it was generally assumed that *dhfr* mutations underlying resistance evolved multiple times in nature. Since infected people contain 10^{10} to 10^{12} parasites, and key point mutations in *dhfr* conferring resistance to pyrimethamine occur at frequencies as high as 2.5×10^{-9} per parasite replication in the laboratory (Paget-McNicol & Saul 2001), we might expect such mutations to arise independently in every treated malaria patient. In this case, we would expect resistant *dhfr* alleles to be associated with different alleles at flanking microsatellite loci and to see little evidence for diminished variation around *dhfr* (Doumbo *et al.*, 2000). However, three recent publications have disproved this conjecture (Cortese *et al.*, 2002; Nair *et al.*, 2003; Roper *et al.*, 2003). Molecular analyses of the *dhfr* gene and flanking regions have shown that high level resistance to SP in South America appears to have spread from a single mutational event in *dhfr* (Cortese *et al.*, 2002) a pattern now known

to have also occurred in Southeast Asia (Nair *et al.*, 2003) and southern Africa (Roper *et al.*, 2003).

In a study conducted using samples from Thailand/Myanmar border, Nair and colleagues found minimal variation in microsatellite markers over a 12 kb region (0.7 cM) immediately surrounding *dhfr*, and variation was reduced in a region approximately 100 kb (6 cM) around this locus (Nair *et al.*, 2003). In comparison expected heterozygosity (H_e) at markers situated more than 58 kb from the 5' end and more than 40 kb from the 3' end of *dhfr* was high (H_e = 0.81), and these were not significantly different from H_e at 56 unlinked dinucleotide microsatellites (H_e= 0.80) sampled from chromosomes 1, 2, 3 and 12 and genotyped from the same parasite collection (Nair *et al.*, 2003).

In a study of *P. falciparum* populations from 5 locations in S. America (Brazil, Colombia, Haiti, Peru and Venezuela), Cortese and colleagues showed that the mutations *dhfr* C50R, *dhfr* I164L, *dhps* K540E, and *dhps* A581G, (denoted RLEG) conferring mid- and high-level SP resistance have a common origin (Cortese *et al.*, 2002). First, sequences in coding and noncoding regions within or flanking these genes show identical or nearly identical polymorphic patterns among South American parasites harboring these alleles. Second, alleles of 2 genes involved in chloroquine resistance that are not linked to *dhfr* and *dhps* tracked with the RLEG alleles. Either of 2 *Pfcrt* alleles first identified in Brazilian parasite lines from the 1980s and subsequently detected in all or most Brazilian and Bolivian isolates from the 1990s were found in all of the *P. falciparum* isolates carrying these alleles but only in a minority without the mutations. *Pfmdr1* D1246Y also was predominant among RLEG malaria parasites. The association of the RLEGs with the mutant *pfmdr1* allele, as well as with 2 resistant *Pfcrt* alleles carrying a unique C72S consequent to mutually exclusive nucleotide

substitutions, suggests that the RLEGs may have spread by the selective pressures of both the antifolates and the 4-aminoquinolines in S. America (Cortese *et al.*, 2002).

In another study done in South Africa, Roper and colleagues (Roper *et al.*, 2003) showed that only a single haplotype of microsatellite loci flanking a region of 8 kb upstream of *dhfr* gene, in 43 samples, was associated with the triple *dhfr* mutant coding for 51I/59R/108N. The double mutants encoding 51I/108N and 59R/108N were found to be associated with one and two haplotypes respectively. This is in contrast with 8 and 31 haplotypes associated with the single mutant coding for 108N and sensitive alleles respectively in the same region (Roper et al. 2003). Likewise, analysis of an 8 kb region flanking the *dhps* gene (downstream) showed that isolates with double mutations coding for 437G/540E were associated with just a single flanking haplotype. By contrast, the single mutant (coding for 436A) isolates were associated with 3 unrelated flanking haplotypes (Roper *et al.*, 2003).

The alleles with several resistance-conferring mutations in both of these genes showed a high degree of conservation in the flanking sequences. In fact, a comparison of the flanking sequences with isolates from northern Tanzania revealed that these isolates had the same flanking haplotype as those found in South African isolates, 4000 km apart. This suggests that rather than arising many times, a few mutant alleles have been subject to positive drug selection and have introgressed through large geographic regions, a fact that can only be attributed to the use of sulphadoxine/pyrimethamine in these areas (Roper *et al.*, 2003). Although single mutants seem to have arisen independently on multiple occasions, the fitness benefits they confer are perhaps too weak to overcome the associated fitness costs, and thus do not seem to be strongly selected for. On the rare occasions where a multiple mutant arises, it has high relative fitness and can spread rapidly through large geographic regions (Roper *et al.*, 2004).

In parasite populations such as those in South America (Anderson *et al.*, 2000a; Cortese *et al.*, 2002), Southeast Asia (Anderson *et al.*, 2000a; Nair *et al.*, 2003) and even some parts of Africa where transmission is low (Roper *et al.*, 2003), self-fertilization events predominate, and levels of recombination are consequently much lower. Most of the studies that have looked at the evolution of drug resistance have used isolates from these areas and thus an understanding of the evolution and spread of drug resistance in high transmission areas, where recombination is high, is still lacking.

1.10 Rationale I for this thesis

The molecular genotyping analyses have largely used the antigen-coding loci encoding merozoite surface proteins 1 and 2 (MSP1 and MSP2 respectively) and the glutamate-rich protein (GLURP). The fact that these genetic markers encode antigens that are under immune selection has led some malaria researchers to question the data generated by their use (Hughes and Hughes 1995; Anthony *et al.* 2000). Microsatellites, which have recently been discovered to be abundant in the genome of *P. falciparum*, are usually in non-coding regions of DNA, and are therefore assumed to be selectively neutral (Goldstein & Schlotterer 1999).

In this study we set out to compare the effectiveness of the antigenic loci markers *msp1*, *msp2* and *glurp* with five microsatellite markers, in differentiating recrudescence from re-infection in field isolates from Kenya.

1.11 Rationale II

Malaria is for the most part endemic in developing countries with limited budgets, and there is a need to extend the useful life of cheaper antimalarial drugs. An understanding of the evolution and spread of resistance will enable the effective monitoring of resistance patterns and inform policy changes and better interventions.

The studies on the evolution and spread of antimalarial drug resistance published to date have been conducted in areas with low malaria transmission. A more extensive study involving both low and high transmission areas needs to be conducted to examine the effects of antimalarial drugs in areas with varying transmission indices, especially in high transmission areas, where recombination is high. Therefore, we set out to compare the rates of evolution and spread of SP resistance in malaria endemic areas of Kenya.

SP was officially introduced as the first-line antimalarial treatment in Kenya in 1997 though it was being widely used before then by most people, especially those visiting private health facilities, due to the ineffectiveness of chloroquine. By the time it will be replaced as the official first-line antimalarial by Artemisinin/lumefantrine (Coartem® [Norvatis]) in 2006, treatment failure rates (from *in vivo* data) with SP in most endemic areas might be well over 60% (EANMAT data at http://www.eanmat.org).

1.12 Research objectives

1.12.1 General objective

To investigate the evolution and spread of antimalarial drug resistance in Kenya and the effect of drug resistance mutations on the general population of *P. falciparum* in Kenya.

1.12.2 Specific objectives

i) To study the evolution of SP-resistant parasites and their spread in Kenya as a consequence of SP use.

ii) To compare the effectiveness of unlinked microsatellite markers with antigenic markers of *Msp1*, *Msp2* and *glurp* to differentiate between recrudescence and reinfection in malaria field isolates.

iii) To compare the genotypes of malaria parasites, and their population structure, collected from three different locations (with varying geographical settings and transmission intensities) in Kenya.

iv) To study the relationship between the spread of drug-resistance and the different transmission intensities.

2.0 CHAPTER TWO...Materials and methods

2.1 Comparing unlinked microsatellites with antigenic loci

2.1.1 Study sites

The samples used in this study were collected from regions designated as sentinel sites for monitoring drug resistance established by the Kenyan Ministry of Health and the East African Network for Monitoring Antimalarial Treatment (EANMAT). The three areas represent varying geographical features and degrees of malaria endemicity. The studies were part of *in vivo* efficacy tests for antimalarials and were performed according to standard guidelines (World Health Organization 1996). Ethical clearance for the studies was obtained from the Kenya Medical Research Institute and the University of Glasgow Ethical Review Committee.

The study in Mwea (n= 68 samples) was carried out between October and November 2001. Mwea area (1S 37W) lies at an altitude of 1200m in central Kenya (Figure 5), in a region that is mostly reliant on agriculture by irrigation. Malaria in this area is classified as 'irrigation, seasonal endemic' (http://www.eanmat.org), meaning that the vector environment is maintained by the irrigation agriculture. Malaria transmission is seasonal and the infections prevalence is between 30-40%, (meso-endemic). The maximum temperature can be as high as 33°C. In this region, *P. falciparum* is highly resistant to SP, with early treatment failure rate (ETF) of more than 35%, (EANMAT, 2003) that indirectly or directly impacts on the increased number of malaria illnesses and malaria-related deaths in this region.

Anopheles arabiensis (73%), *An. pharoensis* (7%) and *An. funestus* (3%) have been identified as the principal vectors of malaria in this area (Rapuoda 1995). Studies conducted between 1989 & 1990 showed the parasite rates among children aged 1-9 years to be between 0.7% and 5.4% (Ijumba *et al.*, 1990). Malaria in this area appears to affect both adults and children (52% of admissions were below 15 years between 1996

and 1997) (Division of Malaria Control 1998). At the time of this study the SP treatment failure rate was 24%.

The studies in Tiwi (n= 60 samples) and Bondo (n= 60 samples) were carried out between October and December 2000. Tiwi area is located in the coastal region of Kenya, 4S and 39W (Figure 5), where malaria is holoendemic. It lies at an altitude of just 2m above sea level. Tiwi has an annual entomologic inoculation rate (EIR) of four infective bites per year (range 0-18) (Mbogo *et al.*, 1995). An annual EIR of 120 has been recorded for a site with the highest transmission intensity in this region (Mbogo *et al.*, 2003). This area is characterized by continuous malaria transmission with pronounced seasonal variation and relatively high incidence of severe disease occurring in time-space clusters, and patients frequently carry polyclonal isolates (Kyes *et al.*, 1997). At the time of this study SP *in vivo* failure rate was about 30% in Tiwi. Bondo is located in the western zone of Kenya (Figure 5) in an area where malaria is hyperendemic with very high transmission rates. It lies at 0N and 34W at an altitude of about 1230m. At the time of this study the SP *in vivo* failure rate in Bondo was about 40%. Parasite prevalence rates in these two areas are >90% especially in children less than five years (Division of Malaria Control 1998).

2.1.2 Patient recruitment and management

All studies were carried out using the same inclusion criteria. Patients who presented at the respective hospitals with symptoms suggestive of malaria had a blood smear undertaken and were later considered for recruitment if they fulfilled the following inclusion criteria: 6–59 (6-120 for Mwea) months of age; positive blood smears with *P. falciparum* monoinfection and parasite density of between 2000-100,000 asexual parasites per microliter of blood; uncomplicated malaria with absence of symptoms of severe malaria or danger signs e.g. convulsions, excessive vomiting,

drowsiness or inability to feed; axillary temperature of 37.5 °C and above or history of fever in the last 48 h; absence of history of allergy to sulfa drugs; hemoglobin level of 5g/dl or more; residence within an accessible address and agreement to return to the clinic for follow-up and provision of informed consent.

Clinical history was taken from the recruited patients including anti-malarial therapy in the last 72 h. A physical examination was carried out and the axillary (armpit) temperature was measured using a clinical thermometer. Fever was defined as axillary temperature of 37.5 °C and above.

Patients who fulfilled the inclusion criteria and agreed to participate in the study were treated with the standard paediatric dose of Sulfadoxine 25mg/kg and Pyrimethamine 1.25mg/kg (SP) as a fixed-dose combination. Patients were asked to return for follow-up on days 1, 3, 7 and 14 for the Bondo and Tiwi studies and on days 1, 2, 3, 7, 14, 21 and 28 for the Mwea study. The Mwea study was extended to 28 days followup because of lower transmission intensity in this region (World Health Organization 1996).

On each day of follow-up, thin and thick blood smears were prepared and examined, hemoglobin concentration in the blood was measured on days 0 and 14 and also day 28 for Mwea using a haemoglobinometer. On each appointed day of the follow-up period, and on any other day if the patient returned to the clinic with fever during this period, the patient's history was taken and they were given a physical examination. Patients who failed to return on the designated days were found and requested to return to the health unit by the assigned social worker/health visitor. A thick blood smear from a fingerprick sample was prepared and examined to assess parasitaemia. Thick smears were stained with 10% Giemsa's stain for 10 minutes, and the number of asexual parasite forms seen against every 200 white blood cells (WBCs) was evaluated.

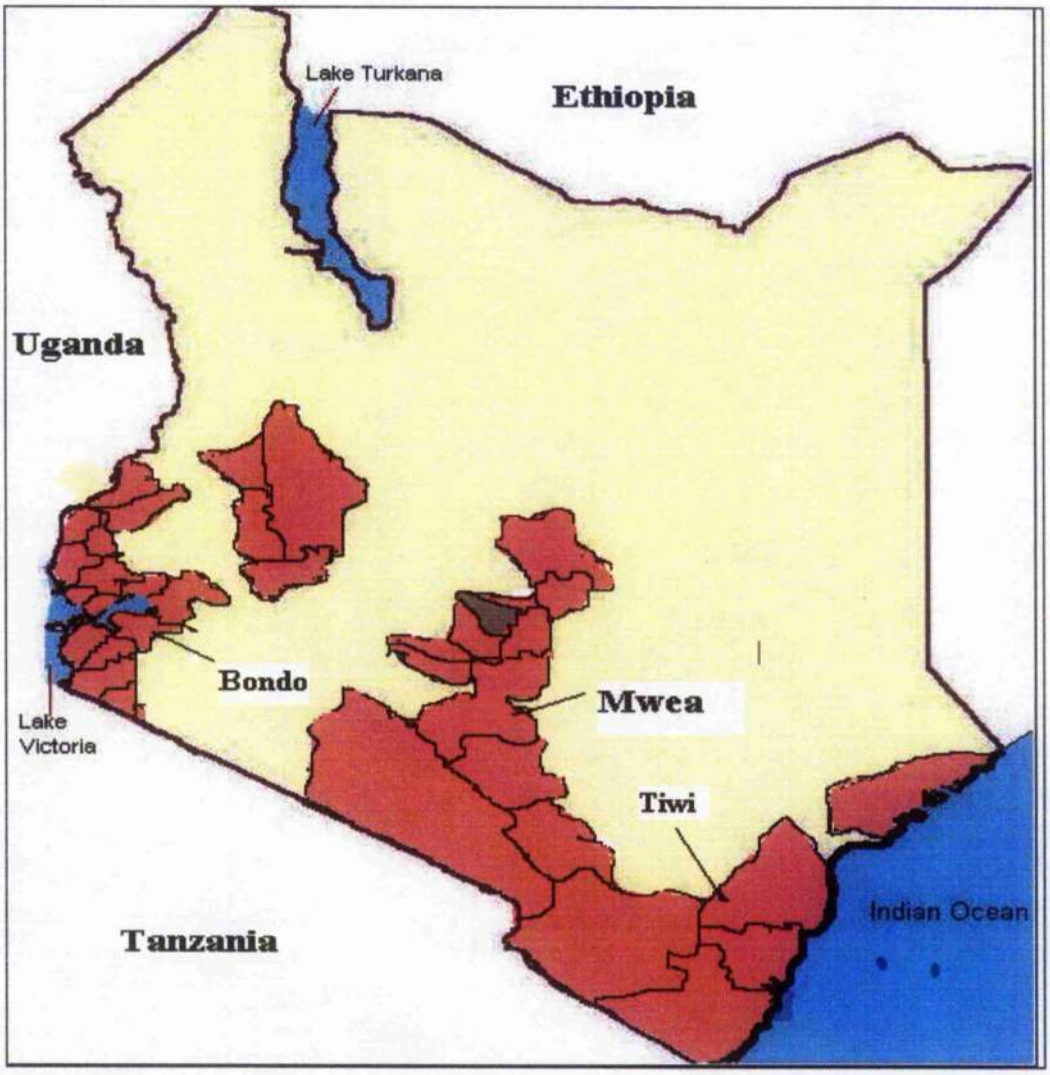

Figure 5: Map of Kenya showing the areas endemic for malaria (shaded brown) and the study sites

The parasite count obtained was then multiplied by a factor of 40 in order to obtain number of circulating parasites per microliter of blood, assuming a WBC count of 8000/μL of blood (World Health Organization 1991a). In the case of a positive smear, alternative drug treatment was provided according to the treating physician, and severe cases were referred to hospitals.

Approximately 2-3 drops of the fingerprick blood collected on day 0 and on the follow-up days was spotted on Whatman 3M filter paper, allowed to dry, and then

packaged individually in sealed plastic bags for molecular analysis as part of this PhD. Project. These samples were marked with the patient identification number and the date of sample, and were stored with desiccant at 4°C.

All outcomes were registered on a case record form labeled with the patient's study identification number.

2.1.3 DNA extraction

2.1.3.1 DNA from cultured material

DNA was extracted from *in vitro* cultures of laboratory lines originating from various geographic origins and this was used to optimize and select the microsatellite markers for the study. The parasite lines chosen were K1 (Thailand), RO33 (Ghana), 7G8 (Brazil), HB3 (Honduras), 3D7 (The Netherlands), SL-D6 (Sierra Leone), ITG2F6 (Brazil), Dd2 (Indo-China), ZIM V160 (Zimbabwe), V1S (Vietnam), FCB1 (Colombia), IEC 513/86 (Brazil), F32 (Tanzania), K28 (Thailand), and SUD106/11, SUD124/8, SUD126/1, and SUD128/5 (all from Sudan). The isolates had been grown *in vitro* according to the standard methods (Trager & Jensen 1976). Culture material was centrifuged at 1500 x g, the supernatant removed and the red blood cell pellet used for DNA extraction. Each culture was at around 6% parasitaemia at the time of extraction. 3 µL of the red cell material was used for DNA extraction. The Instagene® (Bio-Rad, UK) extraction kit was used to extract DNA from the pellet following the manufacturer's instructions. The DNA was not quantified using a spectrophotometer since it is normally too impure to get an accurate reading.

2.1.3.2 DNA extraction from filter papers

DNA from the filter papers was extracted using the Chelex-100 method (Plowe *et al.*, 1995). An approximate 4 mm^2 of the blood-soaked filter paper was excised using

a sterile scalpel blade. This was used for extraction of DNA according to the protocol (Plowe *et al.*, 1995). The extracted DNA (approx. volume 200ul) was divided into 20 µl aliquots to avoid repeated freeze-thaw cycles and degradation. These aliquots were stored at –80°C and the aliquot in immediate use was stored at –20°C.

2.1.4 Microsatellite markers

Over 900 microsatellite loci have been mapped in the genome of *P. falciparum* with at least one every 2-3 kb (Su *et al.*, 1999). The sizes of PCR products for all of these loci are known for the 3D7 clone. However, there is little information on the population size diversity for most microsatellite markers except for twelve that have been used for the population studies that have been carried out to date (Anderson *et al.*, 1999). Of these 12 markers used previously, some have been shown to be located in introns of genes (Anderson *et al.*, 1999), or very close to coding regions, and were therefore not considered suitable for this work.

2.1.4.1 Selection of microsatellite markers

Initially, ten microsatellite loci were selected on the basis of (i) not being in coding sequence or very close to an open reading frame (ORF), (ii) having a PCR product of at least 200 bp with clone 3D7 (www.ncbi.nlm.nih.gov/Malaria/ index.html), and (iii) consisting of trinucleotide repeats. Microsatellites located in chromosomes 4 and 7 were also avoided since these chromosomes are known to code for drug targets which might interfere with normal Mendelian inheritance of alleles through selective sweeps or 'hitchhiking'. The chosen markers and their chromosomal locations are given in Table 3. Five of these markers, C2M3, C3M85, TA48, TA17 and C14M108, have not been used in previous studies.

2.1.4.2 Optimisation of amplification conditions for microsatellite loci

The initial optimisation using DNA extracted from *in vitro* cultures was done using a single round of PCR using the primers designed by Su and colleagues (Su *et al.*, 1999). This is because, with the high parasitaemia present in *in vitro* cultures, one round of PCR is sufficient to generate enough PCR product to visualise on agarose gels. The ten microsatellites were amplified from DNA extracted from 18 laboratory-cultured isolates, with a wide range of geographical locations, 50% coming from East and Southern Africa, as described in section 2.1.3. One round of amplification (30 cycles) was performed for each DNA sample, using 1 μl of DNA extracted previously.

Approximately 5 μl of the PCR product were mixed with 1-2 μl of 6X loading dye (Promega; 0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll® 400, 10mM Tris-HCl (pH 7.5) and 50mM EDTA (pH 8.0)) and analysed by agarose gel electrophoresis on a 3% Metaphor® agarose gel (supplier) in 1X Tris-borate-EDTA (TBE) buffer (prepared as one litre 10X stock with 108gm Tris base, 55gm boric acid and 40 ml 0.5 M EDTA). Metaphor® is a high-resolution agarose that enables better resolution than standard agarose, allowing discrimination of PCR products with small size differences (4–10 bp). Although polyacrylamide gels or automatic sequencing machines give excellent resolution of very small differences in PCR product size, this benefit is outweighed by their high cost, toxicity of reagents and relatively low throughput. High-resolution agarose gels provide almost the same level of resolution for a lower cost without the need to purchase expensive equipment and reagents and may therefore be more suitable for use in laboratories in malaria-endemic countries. Stutter bands are also not seen with Metaphor agarose, whereas these can cause problems of interpretation on polyacrylamide gels.

After electrophoresis, the gel was stained with 1μg/ml solution of ethidium bromide for 45 minutes. The DNA was visualised by ultraviolet transillumination and band sizes determined by comparison with a standard 100-base pair DNA ladder (Promega) run at each side of the gel. The sizes were calculated using a gel documentation package, Labworks 3.0 Image Acquisition and Analysis Software (UVP Inc., Upland California, USA).

DNA from any parasite line that failed to amplify at any locus after repeating the amplification reaction a second time was scored as a null allele at the specific locus.

Five of the markers were chosen for the field analysis on the basis of having the widest range of product sizes, highest degree of polymorphism, and lowest number of null alleles when tested with the 18 laboratory isolates. The loci chosen, with their repeat units and size of the respective 3D7 PCR product, are shown in Table 4. A microsatellite locus with dinucleotide repeats, C2M3 (chr2), was included for comparison with the trinucleotide repeat loci.

2.1.4.2.1 Basic PCR conditions

The cycling parameters for all loci were optimised to work with the same concentrations of PCR reagents. These were final concentrations of 1X buffer, 200 uM dNTPs, 100 nM of each primer and 1 unit of Taq polymerase. A PCR machine (MJ Research) fitted with a gradient block was used to define the optimum annealing temperatures. The full primer sequences and PCR conditions for each of the selected loci are listed in (Table 4 and Table 5).

2.1.4.2.2 Design of primers and optimisation for nested PCR

Field isolates generally have lower parasite densities (normally 0.1- 3%) compared to those found in cultures (about 6%) and thus require two rounds of PCR to get adequate DNA for visualisation. Therefore a nested PCR reaction was used with the

extra primers being designed outside the first pair using the 3D7 sequence available at www.PlasmoDB.org. The primers were designed such that the primary PCR product was bigger than the secondary product by at least 150 bp. This was to aid in better resolution in case of carry-over PCR material from the primary PCR during template addition.

These PCR conditions for the reactions were optimised using a few of the culture isolates and are shown in Table 5.

Microsatellite loci	Chromosome	3D7 PCR product
C2M3	2	340
C3M85	3	270
POLYα	4	224
TA42	5	183
TA1	6	198
TA17	8	199
TA40	10	210
TA48	12	253
TA60	13	202
C14M108	14	241

Table 3: List of microsatellite loci, chromosomal location and size of the 3D7 PCR product

Marker	Repeat	Outer Primers (shown 5' to 3')	Nested Primers (shown 5' to 3')
C2M3	TA	F-CACTTATGTTATGACAAGAAC R-GTGAAAAGGATATGCTTCC	F-AAAGTGGGATTCATCCAG R-TCGGGGTATTATTAACATG
C3M85	TTA	F-CTTAACCGTTCAGGAGAT R-GAGTTGATAACTTGTTGGT	F-AAGGGATTGCTGCAAGGT R-CATCAATAAAATCACTACTA
TA48	TAA	F-AATCTATCGGCGTTGGTAGA R-TGGGTTGTGTTACATGAACG	F-TTTTGATATCTCTCAATCAT R-CTTCACGACAGAGGTGTC
TA17	TTA	F-GTATCTCTTAGATGTTAG R-CTTATGGATGTTAATGAC	F-CTGTACCTTTAGGTATCATA R-GAAAATAAAACATTAAAACTATG
TA40	TAA	F-TGGAAAATACAGCAACGGAG R-AGCCAAATCAAAAAGGAATC	F-GGTTAATATGATCACAAAATG R-ATTGTTGATTCATGAAATGCA

Table 4: Microsatellite markers used in the study, their repeat unit, and the sequences of the outer and nested primers

Marker	Cycling Parameters
C2M3	94°C for 2 min, then 30 cycles of 94°C for 40 sec, 45°C for 40 sec and 60°C for 1 min.
C3M85	94°C for 2 min, then 30 cycles each of 94°C for 40C sec, 50°C for 1 min and 60°C for 1 min
TA48	94°C for 2 min, then 30 cycles each of 94°C for 40 sec, 49°C for 40 sec and 60°C for 1 min.
TA17	94°C for 2 min then 30 cycles of 94°C for 40 sec, 50°C for 40 sec and 60°C for 1 min.
TA40	94°C for 2 min, then 30 cycles each of 94°C for 40 sec, 47°C for 40 sec and 60°C for 1 min.

Table 5: Microsatellite markers chosen, and the optimised cycling conditions for each locus.

The same conditions were used for both primary and secondary (nested) reactions

2.1.5 Distinguishing recrudescence from new reinfections

To distinguish genuine recrudescence of resistant parasites from new reinfections arising during the follow-up period, day 0 (D0, pre-treatment) and the positive follow-up (post treatment) isolates (R) were genotyped at three highly polymorphic gene loci, merozoite surface protein-1 (*Msp1*), merozoite surface protein-2 (*Msp2*) and the glutamate-rich protein (*glurp*), as described (Ranford-Cartwright *et al.* 1997).

The isolates were also genotyped with the five polymorphic microsatellite loci, C2M3, C3M85, TA40, TA17 and TA48, as selected in section 2.1.4 above, to compare with the results obtained with the gene loci.

2.1.5.1 Genotyping at the *Msp1, Msp2* and *glurp* loci

DNA extracted from patient samples at the time of admission to the study (day 0) and at the time of reappearance of parasites and/or symptoms (R) was amplified at the *Msp1*, *Msp2* and *glurp* loci using previously published conditions (Ranford-Cartwright *et al.* 1997). Two microlitres of parasite DNA was used as template in the primary reaction, and 3 µl of the primary PCR product was used as template in the nested reaction. Amplifications were performed using a PTC-100^{TM} thermocycler (MJ Research, Inc., Waltham, MA, USA). Primer sequences and reaction conditions are given in Table 6.

Locus	Primers	Outer PCR parameters	Nested PCR parameters
Msp1	01-CACATGAAAGTTATCAAGAACTTGTC 02-GTACGTCTAATTCATTTGCACG N1-GCAGTATTGACAGGTTATGG N2-GATTGAAAGGTATTTGAC	94°C for 3 min, 30 cycles each of 94°C for 25S, 50oC for 35S and 68°C for 2.5 min, then 68°C for 8 min	-As for primary reaction
Msp2	S2-GAGGGATGTTGCTGCTCCACAG S3-GAAGGTAATTAAAACATTGTC S1-GAGTATAAGGAGAAG TATG S4-CTAGAACCATGCATATGT CC	30 cycles each of 94°C for 25S, 42°C for 60S; 65°C for 2 min, then a final extension at 65°C 8min	30 cycles each of 94°C for 25S, 50°C for 60S, 70°C 60S then final extension at 70°C 8 min
Glurp	G4-ACATGCAAGTGTTGATCC G5-GATGGTTTGGGAGTAACG G3-TGTAGGTACCACGGGTTCTTGTGG G1-TGAATTCGAAGATGTTCACACTGAAC	94 °C, 3 min, 30 cycles each of 94 °C, 25S, 45°C, 60S, 68°C, 2min; then final extension at 72°C, 3 min	30 cycles each of 94°C, 60S 55°C, 2 min 70 °C, 2min,then final extension 72 °C, 3 min

Table 6: Primers and PCR cycling conditions for the antigen-coding loci

Ten microlitres of the nested PCR product were mixed with 1-2 ul of 6X loading dye (0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll® 400, 10mM Tris-HCl (pH 7.5) and 50mM EDTA (pH 8.0), Promega) and analysed by agarose gel electrophoresis on a 1.5% agarose gel stained with 0.5 µg/ml ethidium bromide in 1X Tris-borate-EDTA (TBE) buffer. Each A or day 0 (D0) isolate was run alongside its post treatment (R) isolate for ease of comparison. DNA was visualised by ultraviolet transillumination and band sizes determined by comparison with a standard 100-base pair DNA ladder (Promega) run at each side of the gel. The sizes were calculated using a gel documentation package, Labworks 3.0 Image Acquisition and Analysis Software (UVP Inc., Upland California, USA).

If there was no amplification for any sample, the PCR was repeated with up to 2.5 times the quantity of template DNA in the primary reaction. If no amplification was detected after this second reaction, genotyping was classified as unsuccessful.

2.1.5.2 Genotyping at microsatellite loci

Each pair of patient samples (A or D0 and R) was subjected to PCR amplification with each of the five microsatellite loci described in section 2.1.4. As for the antigenic loci *Msp1, Msp2* and *glurp*, two microlitres of parasite DNA was used as template in the primary reaction, and 3 µl of the primary PCR product was used as template in the nested reaction. After amplification the PCR products were subjected to electrophoresis on 3% Metaphor® agarose gels (FMC Bioproducts) in 1 x TBE buffer. Metaphor® enables better resolution than standard agarose, allowing discrimination of products with small size differences (5-10bp.). The DNA was stained with 0.5 µg/ml ethidium bromide and visualised using UV illumination. PCR product sizes were calculated using Labworks 3.0 Image Acquisition and Analysis Software (UVP Inc., Upland California, USA) by comparison with molecular weight markers (100bp ladder, Promega) run on each gel.

2.1.6 Analysis of results

2.1.6.1 Definition of recrudescence or reinfection

It was assumed that after a patient was initially treated for malaria, a subsequent episode was caused by either the same parasite genotype(s) present before treatment (recrudescence) or by parasites acquired after treatment (reinfection) or by a mixture of the two (recrudescence plus reinfection). An outcome was defined as recrudescence if a follow-up sample contained identical alleles or a subset of the alleles present in the admission sample. An outcome was defined as reinfection if a follow-up sample contained only new alleles. If a follow-up sample contained alleles present in the first

sample in addition to new alleles, the outcome was considered as both recrudescence and reinfection.

2.1.6.2 Comparison of the two classes of marker

The comparison of the two sets of the results of the two sets of markers was done using the Cohen's *Kappa* (κ) test of agreement in the SAS/STAT software program version 8.2 (SAS Institute Inc. Cary, North Carolina, USA 1999- 2001). Possible values of κ are constrained to the interval 0 to 1; where κ=0 means no agreement above that expected by chance, and κ=1 means perfect agreement. A value of κ=0.5 and above was considered indicative of significant agreement (Cohen J. 1960). The likelihood of the null hypothesis that κ=0, i.e. no agreement above that expected by chance, was also calculated. Mathematically, kappa is defined as the improvement upon chance agreement, divided by the maximum possible improvement upon chance agreement: K = P(A) – P(E)/1 - P(E) where P(A) is observed agreement and P(E) is agreement expected by chance. The method used to find P(E) is the same as that used in the chi-squared test (Fleiss *et al.*, 1979).

2.1.6.3 Multiple clone infections

Plasmodium falciparum isolates were characterised according to the size of the fragment at each locus. DNA fragments were categorised into different size classes (bins) for the statistical analyses. DNA fragments were allocated into bins of 10 bp for the antigen gene loci, and 6 bp for the microsatellite loci, as described by Brockman (Brockman *et al.*, 1999) and by Paul (Paul *et al.*, 1998). The reason different bin sizes were used for the two sets of markers is because of the use of gels with different resolving powers. The multiplicity of infection (MOI) was estimated by taking the highest number of alleles (bands) per individual. For example, if an isolate had one allele at each of the loci, the MOI was taken to be one, and if the isolate had two alleles

at two loci but three at a third, then the MOI was three. Mean MOI was calculating by dividing the total MOI by the number of individuals analysed.

Since malaria infections in humans are caused by the haploid stage of the parasite, the presence of two or more differently sized PCR products in the same sample using single copy markers indicates a multiple clone infection. Multiple alleles were scored if two or more alleles of the same locus could be distinctly observed on the gel.

2.1.7 Genetic diversity

An analysis of the genetic structure of the studied populations was conducted using only the alleles of the microsatellite loci. While a lot of software programs are freely available for population genetics data, there are very few that can be used for the analysis of population genetics of *P. falciparum* and even these require a lot of reorganization of data and, most times, the exclusion of some. This is mostly due to the genetics of the parasite especially its haploid nature (one allele per locus per parasite), while most software was designed for diploid organisms, and also its multiple clone infections (difficult to tell which allele goes with which in multilocus genotyping). This analysis used some of the software that can analyse haploid data although some of the data had to be excluded or reorganised to fit the programs' requirements.

2.1.7.1 Allele numbers and Heterozygosity

The mean number of alleles (MNA) detected in each population and the expected heterozygosity (*He*) are good indicators of the genetic polymorphism within the populations under study. The MNA is the average number of alleles, at all loci studied, observed in a population (Nei 1987a).

Generally the MNA is highly dependent on the sample size because of the presence of unique alleles that occur in low frequencies, which are less likely to be detected in small samples. In addition, the number of observed alleles tends to increase

as sample size increases, in populations with a large effective population (*Ne*) size (Nei 1987b). Therefore the comparison of the MNA between populations may not be meaningful unless sample sizes are more or less the same (Hartl & Clark 1989; Nei 1987b).

Since the malaria parasites studied in this case were haploid, expected heterozygosity (H_e) for each locus is defined as the probability of drawing 2 alleles of different length from a population sample and is calculated as follows (Anderson *et al.*, 1999):

$$H_e=[n/(n-1)][1-\Sigma^{n}_{i=1} P^2_i]$$

where P is the frequency of the ' i ' th allele and n is the number of alleles in the sample.The MNA and expected heterozygosity were calculated using the program Microsatellite Toolkit (http://oscar.gen.tcd.ie/~sdepark/ms-toolkit/index.html).

2.1.7.2 Linkage disequilibrium

In the study described here, allele frequencies were calculated using only the predominant allele present at each locus within each infection (Anderson *et al.*, 1999). The predominant allele at each locus was defined as the brightest band on the gel under UV illumination. While this procedure results in unbiased estimation of allele frequencies within a population if we assume the composition of PCR products is representative of the composition of templates, it is not the most ideal since the most common parasite in an infection might not be reflected in the gametocyte numbers and therefore in the next generation.

The predominant allele detected at each locus was used to construct ''infection haplotypes''. Where blood samples contain a single parasite clone, this results in recovery of true parasite haplotypes. Where two or more clones are present, the infection haplotypes may be a composite of alleles from two or more clones. This may

impose additional recombination on the data and bias the data against detection of LD although the reverse may be true where detection of LD is missed due to exclusion of minor alleles.

Genotypic linkage disequilibrium was tested by the exact probability test performed using GENEPOP software (Raymond & Rousset 1995). The null hypothesis is that genotypes at one locus are independent of genotypes at the other locus. This test computed unbiased estimates by randomization (2,000,000 iterations) and by the Markov-chain method for the exact probabilities of random association for all contingency tables corresponding to all possible loci pairs within the combined populations.

The program LIAN, version 3 (Haubold & Hudson 2000), was used to compute the number of alleles shared between all pairwise comparisons of complete 5-locus haplotypes and to calculate the variance, V_D, of these pairwise differences. The sample variance was then compared with the variance expected under linkage equilibrium, V_E. To investigate if the observed data differed from random expectations, the observed V_D was compared with the distribution of V_D values in 10,000 simulated data sets in which alleles at each locus were randomly reshuffled among genotypes. The index of association (I_A) was used to measure the strength of LD. The "classical" I_A was defined as $I_A = (V_D/V_e - 1)$, where V_e is the mean variance of the reshuffled data sets (Brown *et al.*, 1980). However, since this statistic scales with $r - 1$, where r is the number of loci analyzed (Hudson 1994), a "standardized" *IA* statistic (I_{AS}), calculated as $I_{AS} = (V_D/V_E - 1)/(r - 1)$ was used. A measure of haplotype-wide linkage and 95% confidence limits, L_{MC}, was determined by Monte Carlo simulations.

2.1.7.3 Genetic differentiation of the populations

The understanding of genetic structuring or differentiation within and between populations is of interest to population geneticists because it reflects the number of alleles exchanged between populations, which influences the genetic composition of individuals within these populations (Balloux & Lugon-Moulin 2002; Souza *et al.*, 1992). Gene flow between populations determines the effects of selection and genetic drift, generates new polymorphisms and increases the local effective population size (Balloux & Lugon-Moulin 2002; Souza *et al.*, 1992). F_{ST} and G_{ST} are very commonly used to describe population differentiation (Nei 1973).

2.1.7.3.1 The fixation index, F_{ST}

F_{ST} is a fixation index that was developed by Wright (Wright 1921) to account for inbreeding within samples. Wright (Wright 1951; Wright 1965) defined F_{ST} as the relatedness between two alleles chosen at random within sub-populations relative to alleles sampled at random from the total population. F_{ST} therefore measures inbreeding due to the relatedness among alleles because they are found in the same sub-population. F_{ST} can be defined as:

$F_{ST} = (H_T - H_S)/ H_S$

Where H_T is the expected total heterozygosity and H_S is the observed within-population heterozygosity.

F_{ST} estimates were calculated using the programs FSTAT 2.9.3 (Goudet 2000) and GENEPOP (Raymond & Rousset 1995).

2.1.7.3.2 The coefficient of gene differentiation, G_{ST}

The coefficient of gene differentiation (G_{ST}) developed by Nei, (Nei 1973) is an extension of his genetic theory to apply to a pair of populations applied to the structured populations (Nei & Feldman 1972). G_{ST} can be computed directly from allele

frequencies in terms of expected heterozygosities within and between populations. Unlike F_{ST}, the estimation of G_{ST} relies only on allele frequencies (Nei 1987a). This method offers several advantages because the number of alleles at the locus does not affect the value and neither do evolutionary forces such as mutation, selection and migration.

G_{ST} can be defined as:

$G_{ST} = D_{ST}/H_T$

Where D_{ST} is the average gene diversity between and within sub-populations and H_T is the mean heterozygosity among the populations studied. G_{ST} estimates in this study were calculated using the program FSTAT (Ota 1993).

2.1.7.4 Population subdivision

The isolation by distance model of population structure was used to test for population subdivision. This model assumes that mobile 'propagules' with continuous geographical distribution in an undifferentiated environment distribute themselves at various distances from their origin (Rousset 1997). There are no discrete boundaries between different parts of the range and allele frequencies change gradually without any discontinuities (Nadler 1995). The extent of population subdivision as calculated by isolation by distance was performed using the ISOLDE program incorporated in the GENEPOP software (Raymond & Rousset 1995). The subdivision analysis was done per population/site.

2.2 Analysis of loci linked to drug resistance

The samples taken at admission were subjected to PCR amplification of two genes linked to resistance of SP, which is the drug the patients were treated with. This section describes the analysis of the genes *dhfr* and *dhps* and the loci flanking approximately 8 kb on either side of each gene. The genes were analysed using dot blot

hybridization with oligonucleotide probes specific to different mutations previously linked to drug resistance (Abdel-Muhsin *et al.*, 2002).

2.2.1 Analysis at the **dhfr** *locus*

A 700 bp fragment of *dhfr* was amplified from each admission sample using a fully nested PCR, as previously described (Plowe *et al.*, 1995). DNA from parasites 3D7, Dd2 and T994 were included as controls for each of the alleles at positions 51, 59 and 108 of *dhfr* (Table 9). The outer PCR was performed with primers AMP1 and AMP2 with the reagents and cycling parameters shown in Table 7. 2-5 ul of DNA was added as template to each outer PCR.

The nested PCR was prepared using the SP1 and SP2 primers as shown in Table 7, in a final reaction volume of 30ul. 2 µl of the outer PCR product was used as the template. The reaction parameters of the nested PCR are outlined in Table 7.

5 µl of the PCR product was mixed with1 µl gel loading buffer (Promega) and subjected to electrophoresis through a 1.5% agarose gel in 1X TBE buffer with 0.5µg/ml ethidium bromide. A 100 bp molecular weight marker was included in the gel to help with product sizing. PCR products were visualised by UV transillumination. The remaining 25µl of PCR product was stored at –20°C prior to preparation of the dot blot.

	PRIMERS	REAGENTS CONC.	PCR PARAMETERS
Outer	AMP1- TTTATATTTTCTCCTTT TTA AMP2 - CATTTTATTATTCGTTTTCT	100nM of primers, 200 µM dNTPs, 1 unit Taq, 1.5mM $MgCl_2$* and 1X PCR buffer	95°C for 3 min [92°C for 30 sec, 45°C for 45 sec and 72°C for 45 sec] 45 then 72°C for 3 min
Nested	SP1- ATGATGGAACAAGTCTG CGAC SP2 -ACATTTTATTATTCGTTTTC	-As above-	95°C for 3 min [92°C for 30 sec, 45°C for 30 sec, 72°C for 30 sec] 30 then 72°C for 3 min

Table 7: PCR reagents and cycling parameters for *dhfr*.

The primers are shown in the 5' to 3' orientation. * The $MgCl_2$ was included in the PCR buffer.

2.2.2 Analysis at the dhps *locus*

An 1152 bp fragment of *dhps* was amplified for each isolate using a fully nested PCR as previously described (Abdel-Muhsin *et al.*, 2002). In addition to the field samples, control samples to be used in each blot were included. The parasite lines 3D7, Dd2, T996, SLD6, HB3 and IEC513/86 were used as controls as they have various combinations of mutations at the *dhps* locus Table 9. The outer PCR was performed using primers 186 and M3717 with the other PCR reagents as shown in Table5. 2 µl of the sample DNA were added to the premix before amplification using the cycling parameters shown in Table 8.

The nested PCR premix was prepared using primers 185 and 218 and the PCR reaction performed with the cycling parameters shown in Table 8, in a final volume of 30ul. 2-5 ul of the outer PCR product were used as template.

Analysis of the PCR products was performed as described earlier (section 2.2.1).

Reaction	PRIMERS	REAGENTS CONC.	PCR PARAMETERS
Primary	186 GTTTAATCACATGTTTGCACTTTC M3717 CCATTCCTCATGTGTATACACAC	100nM each primer, 1X PCR buffer, 1.5mM $MgCl_2$*, 200µM dNTPs, 1 unit of Taq	95°C for 3 min [92°C for 30 sec, 50°C for 45 sec and 72°C for 60 sec] X30 then 72°C for 3 min.
Secondary	185 TGATACCCGAATATAAGCATAATG 218 ATAATAGCTGTAGGAAGCAAT TG	-As above-	95°C, 3 min [92°C for 30 sec, 48°C for 30 sec 72°C for 30 sec] X30 then 72°C for 3 min

Table 8: PCR reagents and cycling parameters for *dhfr*.

The primers are shown in the 5' to 3' orientation. * The $MgCl_2$ was included in the PCR buffer.

Parasite	Dhfr51	Dhfr59	Dhfr108	Dhps436/37 combination	Dhps540
3D7	Asn	Cys	Ser	Ser/Gly	Lys
Dd2	**Ile**	**Arg**	**Asn**	Phe/Gly	Lys
T9-94	Asn	Cys	**Thr**		
T9-96				Ala/Gly	Lys
SL/D6				Phe/Ala	Lys
HB3				Ser/Ala	Lys
IEC513/86				n.d.	Glu

Table 9: The parasite isolates used as controls, with the amino acids present at specific codons in dhfr and dhps.

2.2.3 Dot Blot Hybridization

2.2.3.1 Preparation of samples for dot blot

For each sample that gave a positive result by nested PCR, 20 µl of PCR product was denatured with EDTA and NaOH to final concentrations of 10mM EDTA and 0.4M NaOH. To ensure complete denaturation, the mixture was boiled at 100°C for 10 minutes in a boiling waterbath. The tubes were then spun briefly in a microfuge.

The mixture was neutralised using an equal volume of 2M ammonium acetate, pH 7 and blotted onto Genescreen nylon membrane (New England Nuclear, Houndslow, United Kingdom) using a dot-blotting apparatus (Bio-Rad, Hemel Hempstead, United Kingdom) according to the manufacturer's instructions. Each PCR product was loaded into two different wells of the dotblotter, with half of the remaining PCR product (30ul) being placed into each well, and the duplicates being placed in different parts of the blot. After DNA transfer, the membrane was rinsed in 2X SSC (0.15M sodium chloride, 15mM Trisodium citrate) for 1 minute, denatured in 0.4M NaOH for 1 minute and finally neutralized in a mixture of 1M Tris-HCl and 1.5 M NaCl, pH 8, for 30 seconds. To fix the PCR products on to the membrane, the blot was then exposed to ultraviolet light in a UV crosslinker (Stratalinker, Stratagene, Amsterdam, The Netherlands) according to the manufacturer's instructions.

2.2.3.2 Oligonucleotide probe labelling and hybridisation

In order to detect the *dhfr* and *dhps* alleles associated with changes at amino acid positions 51, 59 and 108 by the dot-blot hybridisation technique, specific probes were designed and labelled with [γ-^{32}P]-ATP (Amersham Pharmacia Biotech Inc, Little Chalfont, United Kingdom) (Abdel-Muhsin *et al.*, 2002). The sequences of the probes and the specific hybridisation conditions are given in Table 10 and Table 11.

Ten picomoles (pmol) of each probe were labelled using a single Ready-To-Go polynucleotide kinase reaction mix (Amersham Pharmacia Biotech Inc, Little Chalfont, United Kingdom) containing 10 units of T4 polynucleotide kinase, 50 mM Tris-HCl, (pH 7.6), 10mM $MgCl_2$, 5mM DTT, 0.1mM spermidine, 0.1mM EDTA (pH 8.0), 0.2μM ATP and stabilizers in a total volume of 50 μl, to which 0.37 MBq of [γ-^{32}P]-ATP was added. Each probe was incubated with the labelling mixture at 37°C for 30 minutes and the unincorporated [γ-^{32}P]-ATP was removed using Microspin G-25 columns (Amersham Pharmacia Biotech Inc, Little Chalfont, United Kingdom) according to the manufacturer's instructions. The labelled probes were stored at –20°C until required.

The membrane was prehybridised with hybridisation buffer [5 X SSPE (0.15M NaCl, 10mM sodium phosphate, 1mM EDTA), 5 X Denhardt's reagent, 0.5% sodium dodecyl sulphate (SDS), 0.02 mg/ml of sonicated salmon sperm DNA] using 0.25-0.125 ml of hybridisation solution per cm^2 of membrane. The membrane and the buffer were incubated in a hybridisation oven with rotisserie (Grant Boekel HTR 12, Cambridge, England) for at least 30 minutes at a temperature specific for the respective probe Table 10 and Table 11. The labelled and purified oligonucleotide probe was added to the hybridisation buffer (1 μl probe for every 1 ml of the hybridisation buffer) and the blot hybridised overnight at a temperature specific for the probe used. The blot was washed once with 2 X SSC for 20 minutes and twice for 10 minutes in 1 X SSC/ 0.1% SDS or 0.5 X SSC/ 0.1% SDS, at the temperature specific for the respective probe (Table 10 and Table 11).

The blot was wrapped in cling film and taped into an autoradiography cassette complete with intensifying screens. Location markers (Glogos II markers, Stratagene, Amsterdam, The Netherlands) were used to allow accurate positioning of the

autoradiograph and blot. Each blot was exposed to Kodak MXB X-ray film overnight at –80°C after which the X-ray film was developed using an automatic developer (X-OGRAPH Imaging System Compact X4). The probe was stripped from the blot with two washes in 0.1M NaOH for 15 minutes each at room temperature. The blot was then rinsed briefly in 5 X SSC, air-dried and then re- hybridised with other probes or kept at room temperature until required. Each blot was hybridised with each of the probes in Table 10 or Table 11 according to whether the PCR product was from the *dhfr* or *dhps* locus.

Probe	Sequence (5' to 3')	Hybridisation temperature	Stringent washes
DHFR 108			
Ser-specific (wild-type)	AACAAGCTGCGAAAGCATTCCAA	50°C	[(1XSSC/0.1%SDS) 10min] X2
Asn-specific (mutant)	AACAAACTGGGAAAACATTCCAA	54.5°C	[(1XSSC/0.1%SDS) 10min] X2
DHFR-51			
Asn-specific (wild-type)	ATGGAAATGTAATTCCCTAGAT	50°C	[(0.5XSSC/0.1% SDS) 10min] X2
Ile-specific (mutant)	ATGGAAATGTATTTCCCTAGAT	50°C	[(0.5XSSC/0.1% SDS) 10 min] X2
DHFR-59			
Cys-specific (wild-type)	GAAATATTTTTGTGCAGTTAC	52°C	[(0.5XSSC/0.1% SDS) 10 min] X2
Arg-specific (mutant)	GAAATATTTTCGTGCAGTTAC	50°C	[0.5XSSC/0.1%SDS] 10 min X2

Table 10: *Dhfr* variant codons with their specific oligonucleotide probes and hybridisation and washing conditions

Probe	Sequence (5' to 3')	Hybridisation temperature	Stringent washes
DHPS 436/437			
Ser/Gly-specific (mutant)	GAATCTTCTGGTCCTTTT	42.5°C	[(1XSSC/0.1%SDS) 10min] X2
Ala/Gly-specific (mutant)	GAATCCGCTGGTCCTTTT	52°C	[(1XSSC/0.1%SDS) 10min] X2
Ser/Ala-specific (wild-type)	GAATCCTCTGCTCCTTTT	50°C	[(1XSSC/0.1%SDS) 10min] X2
DHPS-540			
Glu-specific (mutant)	CAATGGATGAACTAACAA	35°C	[(1XSSC/0.1%SDS) 10min] X2
Lys-specific (wild-type)	CAATGGATAAACTAACAA	35°C	[(1XSSC/0.1%SDS) 10min] X2

Table 11: *Dhps* variant codons with their specific oligonucleotide probes and hybridisation and washing conditions.

The blots were scored based on the specificity of the controls. Experimental samples and controls were dot-blotted in duplicate to reinforce correct scoring. Blots that had autoradiograph results with controls showing non-specific hybridisation were given an extra stringent wash to ensure specificity. If this was not successful in removing the non-specific hybridisation, then the hybridisation and washing steps were re-optimised.

2.2.4 Selective sweep due to SP use

2.2.4.1 Analysis of microsatellite loci around *Dhfr*

The DNA samples were typed at variable microsatellite loci flanking both sides of the *dhfr* gene. The microsatellite sequences analysed are located at 0.3 kb, 4.7 kb, 7.0 kb and 8.0 kb and 0.5 kb, 4.0 kb, 6.5 kb and 8.0 kb, downstream and upstream of the

dhfr gene, respectively (Table 12 and Table 13). A semi-nested PCR protocol was used to amplify each locus using primers designed from the sequence obtained from parasite clone 3D7 on the PlasmoDb database (Bahl *et al.*, 2002) except DHFR.3F and DHFR.3R which were described by Su and Wellems (1996), and DHFR.3R* which was described by Roper and colleagues (Roper *et al.*, 2003).

The optimised primary reaction contained 2 µl template, 1X PCR buffer, 200 µM dNTPs, 3.0 nM $MgCl_2$, 100 nM of each primer and 0.5 units Taq polymerase. The secondary reaction was performed in a total volume of 20 µl made up of 3 µl of primary PCR product, 1X PCR buffer, 200 µM dNTPs, 3.0 mM $MgCl_2$, 100 nM of each primer and 1 unit Taq polymerase. The cycling conditions are as shown in Table 14.

The PCR products were analysed using Spreadex SEL 500® gels (Elchrom scientific, Switzerland) by electrophoresis in 1X Tris-Acetate EDTA (TAE) buffer supplied with the gels by the manufacturer. The precast Spreadex® gels are made of new types of matrices with unique properties that make it possible to get better results by submarine electrophoresis than by vertical or flat-bed electrophoresis using mini gels. These gels are recommended for applications that require the highest resolution in a narrow DNA size range (Kapitanovic *et al.*, 2001). The gels were stained in freshly prepared 0.4 µg/ml ethidium bromide solution in distilled water for 35 minutes. The staining was carried out in a shaker away from direct light according to the

Upstream region of *Dhfr*					Downstream Region of *Dhfr*			
DhfrD4	DhfrD3	DhfrD2	DhfrD1		DhfrU1	DhfrU2	DhfrU3	DhfrU4
8.0 kb	7.0 kb	4.7 kb	0.3 kb	***Dhfr* Gene**	0.5 kb	4.0 kb	6.5 kb	8.0 kb

Table 12: A diagram showing the location of microsatellite markers selected and their location from the ends of the *dhfr* gene.

Locus	DHFR primers (shown 5' to 3')
DOWNSTREAM	
DHFRU1	F-ATTCCAACATTTTCAAGA R-GGCATAAATATCGAAAAC R*-TCCATCATAAAAAGGAGA
DHFRU2	R-TATACAGGACGACGTTCT F- CAATCTCATGTAGACAAA F*-TGGTTTCTGCGATGAAACG
DHFRU3	F*-GCCTTCTTATTTTAAAGGG R-CTTCAAATATATGATGACAT F-GAAATATGTTTACAAGGAGG
DHFRU4	R-TTCCATGCTACAGATAAAACG F-GTTCCTGTTATTTGTTTG F*-GACATGTCTTCACTTTTAG
UPSTREAM	
DHFRD1	F-ATTTTACAATTTCGGATTTTAC R-CATTGAGATAAATAAGTGTTCA F*-TAAAGAAGGCATAATTTTCA
DHFRD2	F-GTATTATATACATGGATCAC R-CTATACATTTCTTTTTCA R*-CACCTTATTTTATTTGAAGGC
DHFRD3	F-GTGATGATATACCAAGCAG R-ATACTTATATCATCAACCT R*-CCGTGTTATTATCTATTC
DHFRD4	F-CATACGATATATGAAGCT R-TCATCCATACAATTCATAT F*-AATAAGAATAAATCAGGAGG

Table 13: Microsatellite loci flanking *dhfr* and their primers.

The study used a seminested PCR strategy. The F and R primers were used in the primary PCR while the* Primer with its corresponding primer (also used in the primary PCR) were used in the secondary reaction.

manufacturer's instructions. They were then destained in distilled water for 30 minutes before visualisation by UV transillumination. The bands were analysed for size variation using Labworks 3.0 Image Acquisition and Analysis Software (UVP Inc., Upland California, USA) by comparison with M3 DNA ladder (Elchrom scientific). This marker contains over 50 DNA fragments in the size range from 50 to 622 bp specially designed for precise estimation of the size of unknown DNA fragments on Spreadex® gels. These are fragments produced from the digestion of pBR322 by *HaeIII, MspI* and *HhaI* enzymes.

Markers	Primary reaction parameters	Nested reaction parameters
DHFRU1	94°C for 2 mins, 35X [94°C for 30S, 47°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 47°C and 65°C for 1 min], then 65°C for 2 mins.
DHFRU2	94°C for 2 mins, 30X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 50°C and 65°C for 1 min], then 65°C for 2 mins.
DHFRU3	94°C for 2 mins, 35X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 49°C and 65°C for 1 min], then 65°C for 2 mins.
DHFRU4	94°C for 2 mins, 30X [94°C for 30S, 47°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 50°C and 65°C for 1 min], then 65°C for 2 mins.
DHFRD1	94°C for 2 mins, 35X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 50°C and 65°C for 1 min], then 65°C for 2 mins.
DHFRD2	94°C for 2 mins, 35X [94°C for 30S, 45°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 50°C and 65°C for 1 min], then 65°C for 2 mins.
DHFRD3	94°C for 2 mins, 30X [94°C for 30S, 46°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.
DHFRD4	94°C for 2 mins, 35X [94°C for 30S, 47°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.

Table 14: PCR parameters for the microsatellite loci flanking *dhfr*

2.2.4.2 Analysis of microsatellite loci around *dhps*

The analysed microsatellites sequences for *dhps* are located at 0.4 kb, 3.7 kb, 7.3 kb and 8.5 kb, and 0.1 kb, 3.8 kb, 6.5 kb and 8.7, upstream and downstream of the *dhps* gene, respectively (Table 15 and Table 16). The PCR mixtures were all set up as above described for *dhfr* loci and the cycling parameters are listed in Table 17.

Upstream region of *Dhps*					Downstream Region of *Dhps*			
DhfrD4	DhfrD3	DhfrD2	DhfrD1		DhfrU1	DhfrU2	DhfrU3	DhfrU4
8.5 kb	7.3 kb	3.7 kb	0.4 kb	***Dhps*** **Gene**	0.1 kb	3.8 kb	6.5 kb	8.5 kb

Table 15: A diagram showing the location of microsatellite markers selected and their location from the ends of the *dhfr* gene.

Locus	**PRIMERS**
DOWNSTREAM	
DHPSU1	F-CTTGACATATAATGAGCATG R-ATTGTGGACAAATCACAC R*-GGAAAGTGCAAACATGTG
DHPSU2	F-TCTATAGTATACATGGAT R- ATTTCAAATTGTTCGTCC F*-CAATGTCCATTGTGCATCA
DHPSU3	F-ATGTACATATTGATAACC R-ATTGTTAATCTTCCTTAG F*-CAATAACCTGAAAAGTGA
DHPSU4	F-GAATAAATTAATTACACACGG R-GTAATACACATAAAACAACAG R*-TGTACATTTAAAGATAGATG
UPSTREAM	
DHPSD1	F-AGTTCTTGTATAGTTTCC R-TCTATAATCGATACCAAG F*-CACTTTTATAGTTTAAGTTG
DHPSD2	F- CAACTTATATCTGAATGG R-CATACAGCAAGTGCAAGAGC F*-GATTGTAATTACTAAATGG
DHPSD3	F-CACTTAATGTAAATGGAG R- ATATTAAGCTTGTACATG F*-CACATATATATGTACGTAGT
DHPSD4	F-CATACAAAACAAAATGCG R-AATATACAAATAGCTAAG R*-CATATGTTTTTAAGATATAAGC

Table 16: Microsatellite loci flanking *dhps* and their primer sequences.

The study used a seminested PCR strategy. The F and R primers were used in the primary PCR while the* Primer with its corresponding primer (also used in the primary PCR) were used in the secondary reaction.

Marker	Primary reaction parameters	Nested reaction parameters
DHPSU1	94°C for 2 mins, 35X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 51°C and 65°C for 1 min], then 65°C for 2 mins.
DHPSU2	94°C for 2 mins, 35X [94°C for 30S, 46°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 51°C and 65°C for 1 min], then 65°C for 2 mins.
DHPSU3	94°C for 2 mins, 35X [94°C for 30S, 40°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 45°C and 65°C for 1 min], then 65°C for 2 mins.
DHPSU4	94°C for 2 mins, 35X [94°C for 30S, 50°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.
DHPSD1	94°C for 2 mins, 35X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 47°C and 65°C for 1 min], then 65°C for 2 mins.
DHPSD2	94°C for 2 mins, 35X [94°C for 30S, 47°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.
DHPSD3	94°C for 2 mins, 35X [94°C for 30S, 46°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 47°C and 65°C for 1 min], then 65°C for 2 mins.
DHPSD4	94°C for 2 mins, 35X [94°C for 30S, 43°C and 65°C for 1 min], then 65°C for 2 mins.	94°C for 2 mins, 30X [94°C for 30S, 48°C and 65°C for 1 min], then 65°C for 2 mins.

Table 17: PCR cycling parameters for the *dhps* flanking markers

2.2.5 Drug resistance selective sweep analysis

2.2.5.1 Drug resistant mutants per site

The frequencies of the drug resistant alleles at *dhfr* and *dhps* were compared among the three sites using Fisher's Exact statistical test. This was to determine whether there were any statistically significant differences in the distribution of drug resistant

alleles among the three sites. The R X C package (downloadable free software [http://www.marksgeneticsoft ware.net]) for the analysis of contingency tables (Miller 1997) was used for Fisher's Exact statistical test since some of the expected values were less than 5.

2.2.5.2 Expected Heterozygosity

Expected heterozygosity (*He*) at each microsatellite locus was measured as described above (section 2.1.7.1). *He* was used in preference to variance in repeat number because *P. falciparum* microsatellites frequently contain indels in the flanking regions or have complex repeat structure (Anderson *et al.*, 2000b). Hence, inference from number of repeats from PCR product length results in frequent errors. These errors are due to indels, which might not be the same size as the repeat unit or, where the repeat structure is made up of mixed repeat units e.g. di- and trinucleotides. This means that use of repeat number is difficult to ascertain in these cases.

2.2.5.3 Testing for population bottlenecks

It has been noted that allelic diversity is reduced faster than heterozygosity during a bottleneck. If a population has gone through a bottleneck, it is found to retain the excess heterozygosity (Nei *et al.* 1975).

To test whether the population had undergone a recent bottleneck, the presence of excess of heterozygosity was tested for using the software '*bottleneck*' (Cornuet & Luikart 1996). This software carries out a Wilcoxon signed rank test comparing observed heterozygosity at each locus across the ~16 kb region of both triple mutant *dhfr* and double mutant *dhps* loci. Expected values are generated under infinite allele (IAM) and step-wise allele models (SMM).

2.2.5.4 Linkage disequilibrium for measuring selective sweeps

Linkage disequilibrium (LD) at a distance x from the core region was measured by calculating the extended haplotype homozygosity (EHH) (Sabeti *et al.*, 2002). EHH is defined as the probability that two randomly chosen chromosomes carrying the core haplotype of interest are identical by descent (as assayed by homozygosity at all loci) for the entire interval from the core region to the point x. Homozygosity was used as a multilocus measure of linkage disequilibrium (Sabatti & Risch 2002). Haplotype homozygosity (HH) measures the variation at linked sites and this is exactly what is affected by selection on one locus within a chromosome, an observation termed 'hitchhiking'. HH was evaluated for all haplotypes for wild type, single, double and triple mutants for *dhfr* and wild type, single and double mutants for *dhps*. This was done across the whole gene sequence inclusive of the flanking loci (referred to as EHH). HH is evaluated as:

$$HH = \left(\sum p_i^2 - 1/n\right) / \left(1 - 1/n\right)$$

with p_i being the relative haplotype frequency and n the sample size. In this way, the calculation corrects for sampling effects (Sabatti & Risch 2002). The variance of each HH [Var(HH)] was estimated according to Nei (1975).

$$Var(HH) = \frac{2(n-1)}{n^3}\left\{2(n-2)\left[\sum p_i^3 - \left(\sum p_i^2\right)^2\right] + \sum p_i^2 - \left(\sum p_i^2\right)^2\right\}$$

EHH thus detects the transmission of an extended haplotype without recombination. This test for positive selection involves finding a core haplotype with a combination of high frequency and high EHH, as compared with other core haplotypes at the locus. We used a web-tool to calculate EHH (Mueller & Andreoli 2004) available at http://ihg.gsf.de/cgi-bin/mueller/webehh.pl

2.2.5.5 Strength of selective sweep in Kenyan parasite populations

Selection coefficients (s), which measure the stability of the selected allele in a population, were estimated from SP efficacy studies done by the East African Network for Monitoring Antimalarial Treatment (EANMAT). EANMAT has been carrying out antimalarial efficacy testing studies in the three East African countries of Kenya, Uganda and Tanzania since 1999 (EANMAT 2001) and the data has been compiled and made available on their website (www.eanmat.org). This data is grouped according to the recent WHO classification guidelines on treatment response (World Health Organization 2003b). For this analysis it was assumed that all patients who did not achieve an adequate clinical and parasitological response (ACPR) (negative blood smear with no clinical signs on day 14 or 28) was a treatment failure. This is probably an overestimation of treatment failures and so the calculated selection coefficients might be overestimates. The frequencies of resistant (p) and sensitive (q) alleles were inferred by assuming that the frequency of treatment failures is proportional to p and that $q = 1-p$. We plotted $\ln(p/q)$ against time in generations using an estimate of six generations per year (Joy *et al.*, 2003). The gradient of the slope, which is the rate at which drug sensitive alleles are replaced by drug-resistant ones, is the selection coefficient, s (Hartl & Dykhuizen 1981; Nair *et al.*, 2003).

3.0..... Results

3.1 Introduction

The results section has been divided into three parts. The first part deals with comparison of the results of the antigen-coding loci with those of the microsatellite loci in differentiating recrudescence versus reinfection. This work has already been published ((Mwangi *et al.*, 2006); Publication annexed). The second part deals with the genetic diversity of *P. falciparum* in Kenya, and the relationships within and between the Kenyan parasite populations. The third part deals with SP resistance in Kenya and the selective sweep driven by both pyrimethamine and sulphadoxine treatment in the Kenyan malaria parasite populations.

3.2 Recrudescence versus Reinfection

Most clinical studies of malaria drug resistance, which apply the WHO guidelines of assessing clinical or parasitological outcome (WHO, 2001), do not distinguish recrudescence of resistant parasites from re-infection by new parasite strains. The use of molecular genotyping to distinguish recrudescence from new infections has become increasingly common in antimalarial drug efficacy studies, but questions still remain regarding the most suitable markers to be used.

In this study, we evaluated molecular genotyping results using two types of markers using samples collected as part of a large national drug efficacy study. One set of markers was the antigenic loci (*Msp1*, *Msp2* and *glurp*), and the other markers were microsatellite loci.

The key questions addressed here are:

i. *Since the antigenic gene loci are under selection, are they effective in differentiating recrudescence from reinfections?* This was determined by comparing their results with those of five putatively neutral microsatellite loci.

ii. *How do the two sets of markers compare in their ability to pick out multiclonal infections?* This was determined by looking at the multiplicity of infection (MOI) values generated using the two sets of markers.

The study used 42, 44 and 47 samples of *P. falciparum* from Tiwi, Bondo and Mwea respectively. The samples were taken from patients both before treatment began (denoted admission, D0 or A sample), and after treatment failure (denoted R sample). For each locus, the PCR products from the two samples, collected before and after treatment, were loaded next to each other on electrophoresis gels. The presence of the

same sized PCR products in the admission sample and its follow-up R sample was scored as recrudescence. If all of the PCR products in the A sample were of a different size to those in the R sample, this was scored as reinfection. In most cases there was a mixture of the two, meaning that the individual didn't clear the initial infection even after treatment, and in addition acquired infections with novel parasites (Figure 6). A comparison of the results obtained using both sets of markers is provided in Table 18.

3.2.1 Msp1, Msp2 and glurp

Using the antigen-coding loci 40, 40 and 45 samples from Bondo, Tiwi and Mwea respectively were successfully typed and classified into the three groups, namely recrudescence, recrudescence and reinfection, or pure reinfection (Table 18).

Most of the samples were typed as containing both recrudescent and reinfecting genotypes (42%, 46% and 70% in Tiwi, Mwea and Bondo respectively) (Table 18).

In all three areas, there were more samples typed as pure recrudescence than there were pure reinfections (Fisher's Exact, P=0.00058).

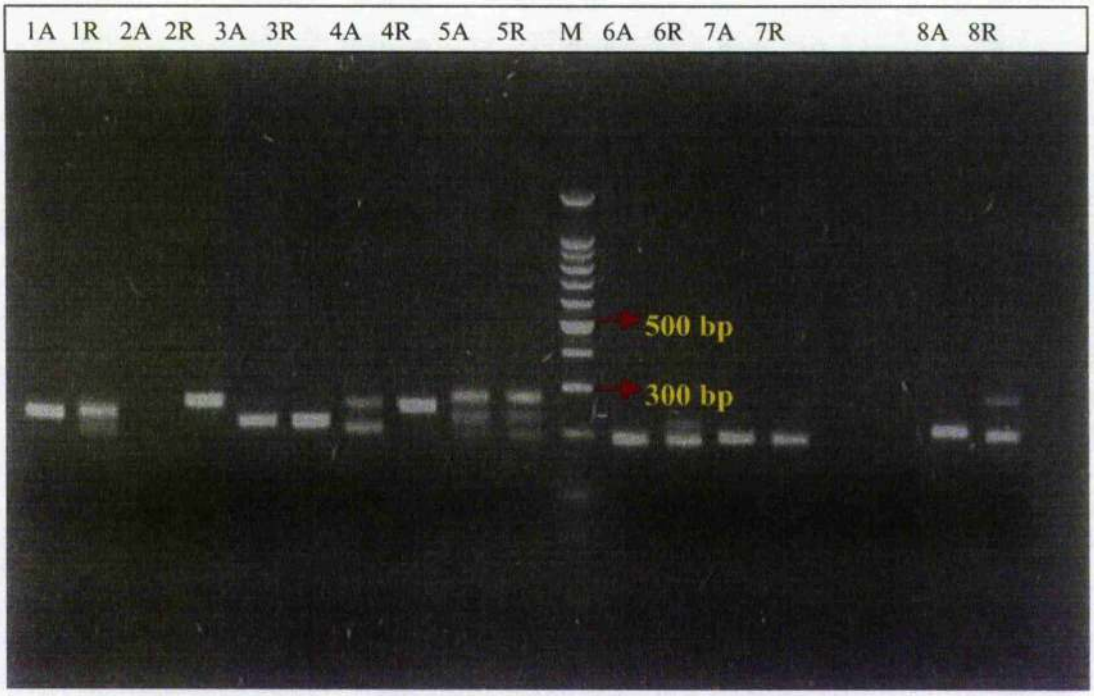

Figure 6: An example of an agarose gel of TA 17 locus for various samples with both recrudescent and reinfecting parasites.

Adjacent lanes are loaded with the PCR product obtained from the Admission (A or D0) or R samples. The lane labelled M contains a 100bp ladder molecular weight marker (Promega). Sizes are indicated in base pairs. Sample pairs displaying patterns interpreted as recrudescent are 3A/3R, 5A/5R and 7A/7R. Sample pairs displaying patterns interpreted as reinfection are 4A/4R and 8A/8R. Sample pairs showing a mixture of recrudescence and reinfection are 1A/1R and 6A/6R. The 2A sample did not amplify in this case.

3.2.2 Microsatellites

41, 42 and 46 samples from Tiwi, Bondo and Mwea, respectively, were successfully typed with the fivemicrosatellite loci. Only a few samples could not be fully resolved with the microsatellite markers (Table 18). As observed with the antigen-coding loci most samples were typed as containing both recrudescences and reinfections (RE+RI) (Range= 47.62% in Tiwi to 79.55% in Bondo).

More samples were also typed as being pure recrudescence in all the three areas than those typed as pure reinfections. Recrudescences ranged from 11.26% in Bondo to 33.33% in Tiwi. Reinfections ranged from 4.54% in Bondo to 16.67% in Tiwi.

SITES / OUTCOME	MWEA (n=47)		TIWI (n=42)		BONDO (n=44)	
	Ags	Microsats	Ags	Microsats	Ags	Microsats
Re	13 (27.65%)	10 (21.28%)	13 (30.95%)	14 (33.33%)	5 (11.36%)	5 (11.36%)
Re + RI	22 (46.81%)	32 (68.09%)	18 (42.86%)	20 (47.62%)	31 (70.45%)	35 (79.55%)
RI	10 (21.28%)	4 (8.51%)	9 (21.43%)	7 (16.67%)	4 (9.09%)	2 (4.54%)
Indeterminate	2 (4.26%)	1 (2.13%)	2 (4.76%)	1 (2.38%)	4 (9.09%)	2 (4.54%)

Table 18: A summary of the results comparing microsatellites and antigen-loci typing for the samples of the patients who did not clear their parasitaemia following SP treatment.

Each pair of samples was classified as recrudescence (RE), reinfection (RI) or both recrudescence plus reinfection (Re + RI). Indeterminate means samples that could not be resolved as either RE or RI or both. Values in parentheses indicate percentages, n= sample sizes, Ags= antigen-coding loci, microsats= microsatellites.

3.2.3 Microsatellite versus antigen-coding loci

We assessed the extent of the agreement between the two sets of markers using the Cohen's *Kappa* (κ) test of agreement in the SAS/STAT software program version 8.2 (SAS Institute Inc. Cary, North Carolina, USA 1999- 2001). The κ values for all the three populations studied were higher than 0.5 (range 0.559-0.606) (Table 19). We obtained significant agreement between the two sets of markers for all three sites than what might be expected from chance alone (*P<0.001 - P=0.0056*) (Table 19).

TESTS OF AGREEMENT	
Antigen-coding vs. microsatellite loci	
Mwea	
kappa coeff.	0.606
95% CI	0.391 - 0.821
P	< 0.0001
Bondo	
kappa coeff.	0.559
95% CI	0.233- 0.884
P	0.0056
Tiwi	
kappa coeff.	0.602
95% CI	0.388 - 0.815
P	<0.0001

Table 19: Results for the tests of agreement between the two marker subsets using the Kappa test for each collection site.

The *P* value given represents the probability of obtaining the *k* value by chance alone

3.3 Multiplicity of Infection (MOI)

All the three sites showed high levels of multiple infections, which were measured by counting the number of samples showing two or more distinctive alleles per locus (Section 2.1.6.3). The mean MOI values were calculated by dividing the total number of highest alleles per individual by the number of samples analysed .

The MOI values ranged from 1.56 (C2M3, Tiwi) to 6.12 (C3M85, Bondo). The average MOI values for the eight markers combined were 2.87, 3.61 and 3.99 for Mwea, Tiwi and Bondo respectively. The marker with the lowest MOI values was

C2M3 with 1.56, 1.88 and 2.01 in Tiwi, Mwea and Bondo respectively. The marker that gave the highest MOI values was C3M85 with 4.42, 5.7 and 6.12 for Mwea, Tiwi and Bondo respectively (Figure 7).

The five microsatellite markers combined had a MOI for all sites of 3.84 compared to that of antigen coding loci of 2.91. The average MOI for the three most polymorphic of the microsatellite loci was 4.69.

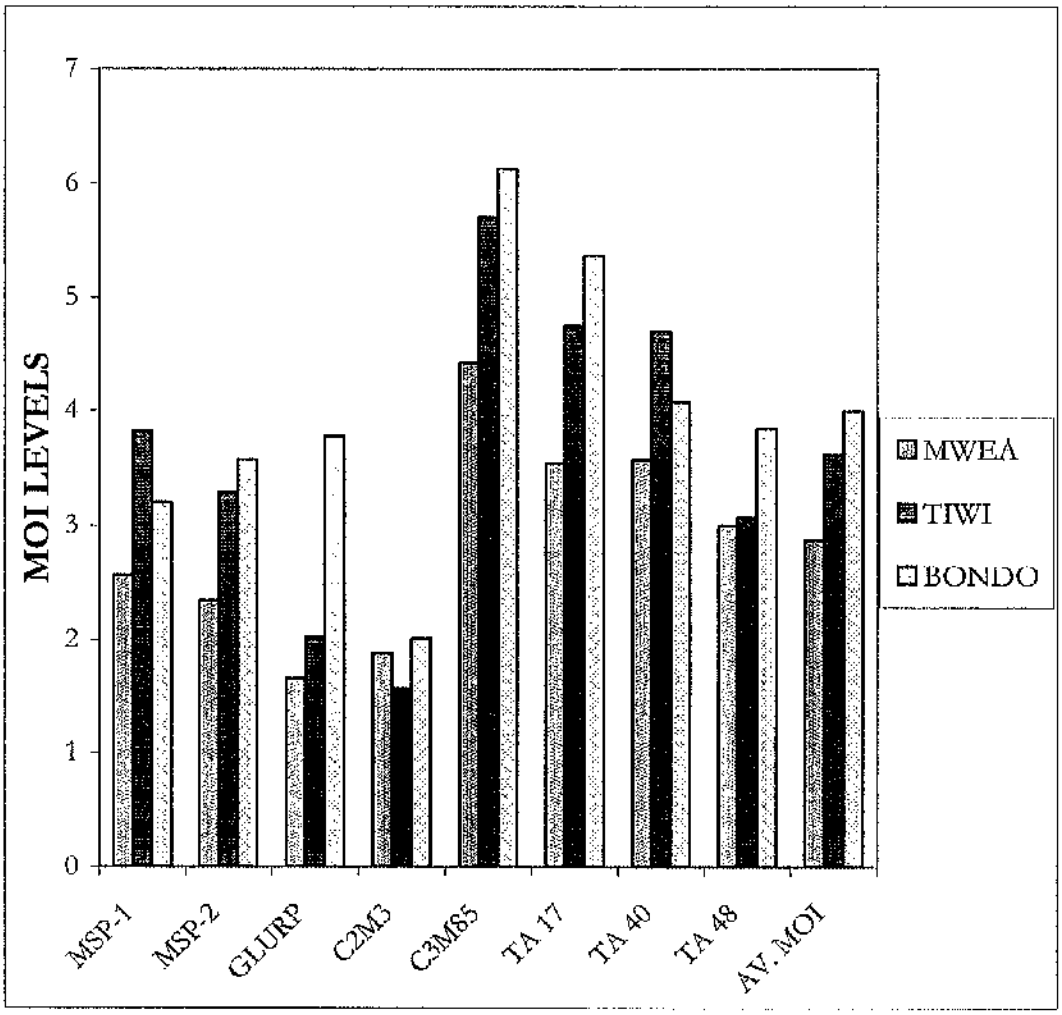

Figure 7: A chart showing the average multiplicities of infection (MOI) for the three areas studied with the two sets of markers.

Also included is the average MOI for the three areas with all the eight markers.

3.4 Genetic diversity study

3.4.1 Introduction

The key questions to be addressed in this section are:

i. *What is the extent of the genetic diversity of the Kenyan P. falciparum populations?* This was determined by calculating the heterozygosities and the mean number of alleles of the three Kenyan *P. falciparum* populations studied.

ii. *What is the level of association among loci in Kenyan parasite populations?* This was determined through the calculation of linkage disequilibrium between all possible pairs of loci.

iii. *How genetically differentiated are the Kenyan malaria populations from each other?* This was determined through the calculation of F_{ST} and G_{ST}.

The results of the microsatellite typing were used to analyse the population structures in the three areas studied. Three measures of diversity of the populations were analysed: the number of alleles per locus, expected heterozygosity and variance (range) in allele length.

3.4.2 Genetic diversity

3.4.2.1 Mean Number of Alleles (MNA)

MNA is a measure of the average number of alleles found in a population (see section 2.1.7.1). The MNA ranged from 12.94 in Tiwi to 15.12 in Bondo when all the individuals within the population were considered (Figure 8). The most polymorphic marker of all eight studied was the microsatellite C3M85 with a mean number of alleles per population of 23.1 in both Tiwi and Bondo, while the least polymorphic marker was C2M3 with 5.1 alleles per population.

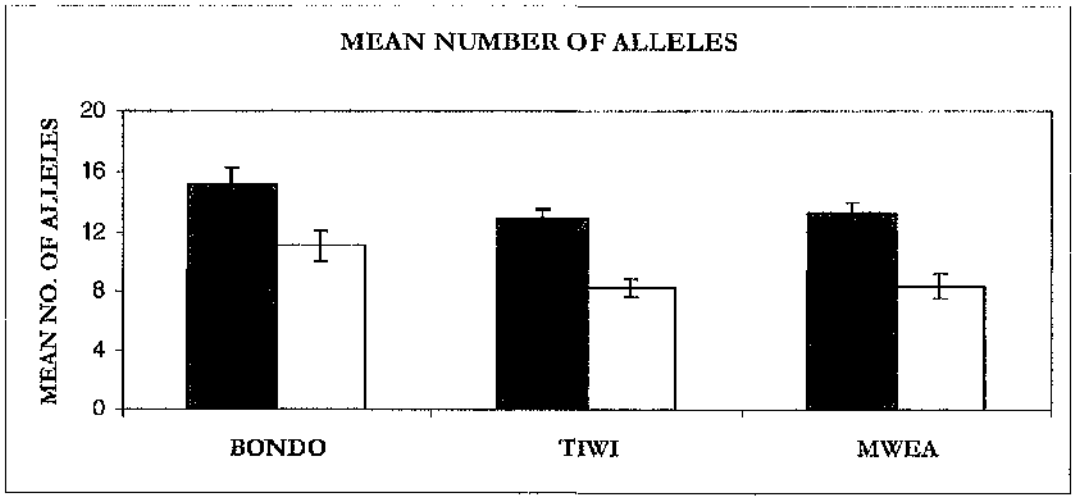

Figure 8: Average mean number of alleles per site for the two sets of markers. These were calculated using the dominant allele per isolate (section 2.1.7.1). Grey bars represent the MNA for the five microsatellite loci while white bars represent the three antigen-coding loci. The y bars show the standard error of the means.

3.4.2.2 Expected Heterozygosities (H_E)

The expected heterozygosities (H_E) were calculated for each locus for each site. The values of mean H_E (± S.d.) obtained ranged from 0.63 ± 0.006 in Mwea to 0.82± 0.005 in Bondo. The locus with the highest H_E was C3M85 with 0.960 ± 0.006 in Tiwi while the lowest H_E was at locus C2M3 with 0.5 ± 0.003 in Mwea (Figure 9).

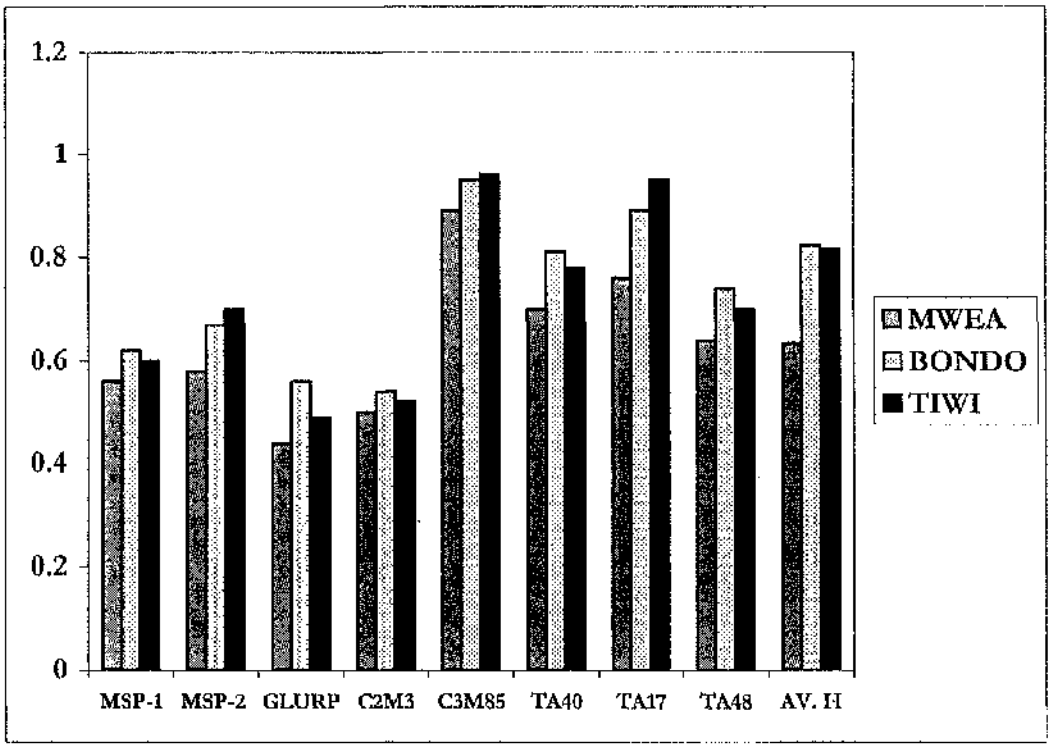

Figure 9: Values of expected heterozygosity (H_E) for the three areas using microsatellite loci and antigen-coding markers.

Av. H represents the average *He* across all loci per site

3.4.3 Multilocus Linkage Disequilibrium

The program GENEPOP was used to test for linkage disequilibrium where the null hypothesis is that genotypes at one locus are independent from genotypes at the other locus (see section 2.1.7.2 for methodology). Results of GENEPOP for all possible pairs of loci and their *P* values are presented in Table 20.

A Monte Carlo simulation method was also used to test the significance of linkage disequilibrium for each locus pair for the complete data set, using the program LIAN version 3 (Haubold & Hudson 2000). The observed mismatch variance of the distance between all pairwise comparisons of complete 5-locus comparisons, V_D, was 0.2970. This was similar to the expected mean variance of the reshuffled data (10,000 times), V_E, which was 0.2925, indicating lack of LD at these loci. The "standardized"

index of association, I^S_A, was 0.0039 (P=0.2427). The I^S_A measures the strength of the linkage disequilibrium; in this case the value obtained is not significant suggesting there is no LD. Therefore for this data set, no significant linkage disequilibrium was observed between any of the markers used by both methods using GENEPOP and LIAN programs. This was expected for these populations where the high recombination driven by the high transmission breaks down any linkage between unlinked loci.

Locus pair	P-value
C2M3 & C3M85	0.124
CM3 & TA40	0.728
C2M3 & TA17	0.961
C2M3 & TA48	0.547
C3M85 & TA40	0.914
C3M85 & TA17	0.292
C3M85 & TA48	0.211
TA40 & TA17	0.720
TA40 & TA48	0.468
TA17 & TA48	0.242

Table 20: P-values for linkage disequilibrium between various locus pairs using GENEPOP.

3.4.4 Genetic differentiation

3.4.4.1 F_{ST} and G_{ST}

F_{ST} and G_{ST} measures inbreeding due to the relatedness among alleles because they are found in the same sub-population. These were calculated according to section 2.1.7.2.

The F_{ST} and G_{ST} estimates for all five loci are presented in Table 21.

The F_{ST} values ranged from 0.002 (TA40) to 0.211 (C2M3) with an overall F_{ST} of 0.048. These F_{ST} are very low and indicate that there is little if any substructuring of the population between the three sites. The G_{ST} values are also very low which means that while there is a lot of diversity among individuals within the same population, there is less diversity among the individual populations.

LOCI	H_E	Ho	D_{ST}	G_{ST}	F_{ST}
C2M3	0.520	0.630	0.141	0.141	0.211
C3M85	0.933	0.988	0.002	0.002	0.003
TA40	0.763	0.765	0.002	0.002	0.002
TA17	0.867	0.872	0.004	0.004	0.007
TA48	0.693	0.710	0.013	0.014	0.020
OVERALL	**0.757**	**0.793**	**0.0332**	**0.0332**	**0.0486±0.0093**

Table 21: Population genetics parameters for the combined populations per locus and also the overall parameters.

H_E = Expected Heterozygosity; H_O = Observed Heterozygosity; D_{ST} = Average gene diversity between subpopulations; G_{ST} = coefficient of gene differentiation; F_{ST} = fixation index. Values were obtained using the program FSTAT.

3.4.5 Isolation by distance

A graph showing the results of the test for population sub-structuring and its relationship to the geographic distance is presented in Figure 10.

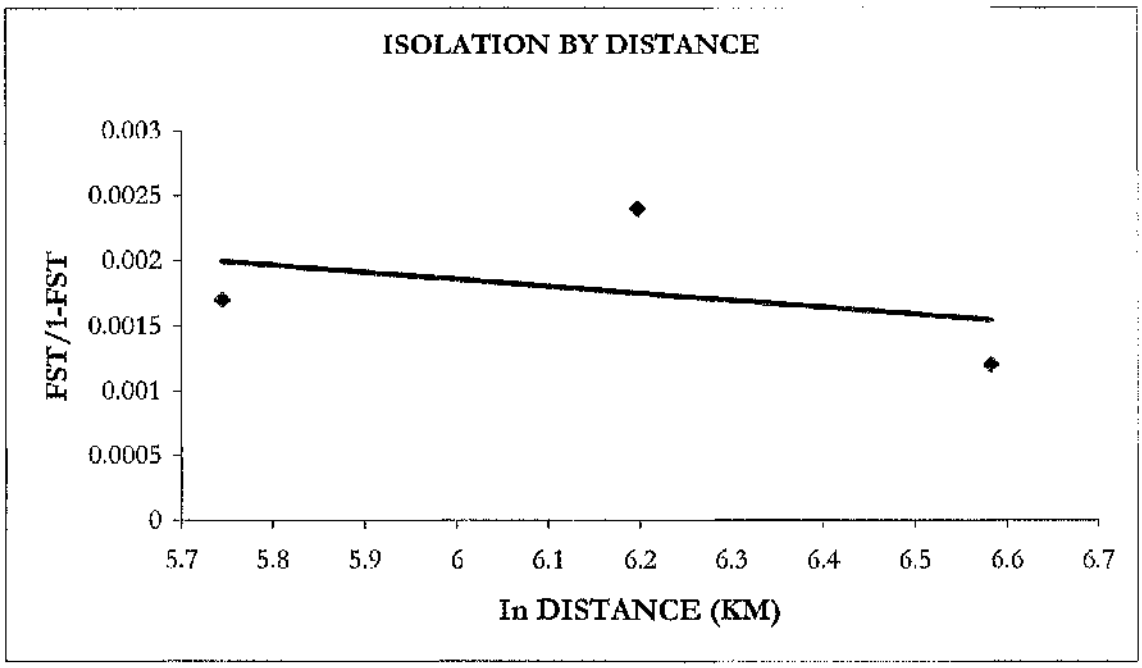

Figure 10: Graph showing isolation by distance between the Kenyan malaria populations studied.

The distance has been transformed using natural logarithm.

The correlation coefficient, $R^2 = 0.1387$ (P=0.8354). The R^2, which is the fit of points to the line, indicates that the line is not a good relationship between the points. Although there are only three sites to compare it emphasizes the finding that there is no relationship between geographic and genetic distance.

3.4.8 Summary and conclusion

The main results and of this section are summarized here.

The high MNA and expected heterozygosities observed with the three Kenyan *P. falciparum* populations, is indicative of the high levels of genetic diversity of malaria parasites in Kenya.

The results obtained using the GENEPOP and LIAN programs confirm the observation of no linkage disequilibrium of the 5-locus haplotypes. A 'standardized' (see section 2.1.7.2) measure of the strength of LD, the index of association I^S_A gave a value of 0.0039, with no significance (P= 0.2427). The simulated variance of distance measure (V_E) was similar to the observed one (V_D). This reveals a lack of LD in these Kenyan populations, an observation that is consistent with population data from Uganda and Congo (Anderson *et al.*, 2000a), that in regions of high transmission, there is random association among loci (low LD) due to frequent recombination which breaks down any association between unlinked loci.

The results of the genetic differentiation (F_{ST} and G_{ST}) show that the populations studied are not sub-structured (Figure 10 and Table 21) and can be said to be undifferentiated. This is confirmed by the test for isolation by distance that reveals a negligible 13% of the variation in these three parasite populations can be attributed to their geographical areas (Figure 10).

The results suggest that there is a lot of out crossing among these three populations. Since the populations are separated by up to 800 km and the range of the mosquito vector is very limited (up to 6km) it seems most likely that the human host, through migration, facilitates outcrossing between parasites from different geographical origins across the studied areas.

3.5 Drug resistance and selective sweeps

3.5.1 Introduction

The key questions to be addressed in this chapter are the following:

1. *What is the prevalence of the mutations in* Pfdhfr *and* Pfdhps *that confer pyrimethamine resistance?* This was determined by typing samples from patients who did not clear the parasites after treatment with SP, using dot blot hybridization.
2. *Do all resistant parasites from Kenya have the same alleles at* Pfdhfr *and* Pfdhps *flanking microsatellite loci?*
3. *Do all Kenyan* Pfdhfr *and* Pfdhps *mutants have the same evolutionary origin?* This was determined by examining the haplotypes of resistant isolates compared to those of sensitive ones for relatedness.
4. *Does the triple mutant* Pfdhfr *and the double mutant* Pfdhps *have an ancestral origin from among the extant wild type* Pfdhfr *in Kenya?*

These four questions were addressed by analysing each of the samples (with mutations at codons 51, 59 or 108 for *dhfr* and 437/540 for *dhps*) at microsatellite loci up and downstream of the resistance-associated locus. Haplotypes were then compared between samples with different *dhfr* and *dhps* alleles. Samples with multiple genotypes present were excluded from the analysis, as it is difficult to determine which allele to associate with which at the respective codons.

3.5.2 Dhfr dot blot hybridization

43, 57 and 60 isolates from Tiwi, Bondo and Mwea respectively were successfully typed for the three codons in *dhfr* associated with SP drug resistance in *P. falciparum*. The samples were classified as either mutant or wild-type based on their respectively genotypes at codons 51, 59 and 108 based on their signals, following

probe-specific radiolabelling, of the dot blot on an autoradiograph (Figure 11). The samples were further classified as having single, double or triple mutations. The results for the various *dhfr* genotypes are summarised in Table 22.

3.5.2.1 Prevalence of *dhfr* mutations

All the samples successfully typed from the two most endemic areas, Bondo and Tiwi, had only the allele encoding Asn at codon 108. 20 % of the samples from Mwea were typed as wild type (encoding serine) at this codon (Table 22 and Figure 12). There were 7%, 8% and 10% samples with mutations at codons 51 and 108, while 14%, 11% and 18% of the samples had mutations at codons 59 and 108 in Bondo, Tiwi and Mwea respectively. Samples with the triple mutation (108, 51 and 59) were found at a prevalence of 33%, 63% and 70% in Mwea, Tiwi and Bondo respectively (Table 22 and Figure 12).

Sites/ genotypes	Mwea (n=60)	Tiwi (n=43)	Bondo (n=57)
Dhfr Wild-type	12 (20%)	0	0
Dhfr 108N (Single mutant)	11 (18%)	5 (11%)	9 (16%)
Dhfr 108N/51I (Double mutant 1)	6 (10%)	3 (8%)	4 (7%)
Dhfr 108N/59R (Double mutant 2)	11 (18%)	5 (11%)	8 (14%)
Dhfr 108N/51I/59R (Triple mutant)	20 (33%)	30 (70%)	36 (63%)

Table 22: Prevalence (and percentage) of the *dhfr* alleles in the three study sites of Kenya.

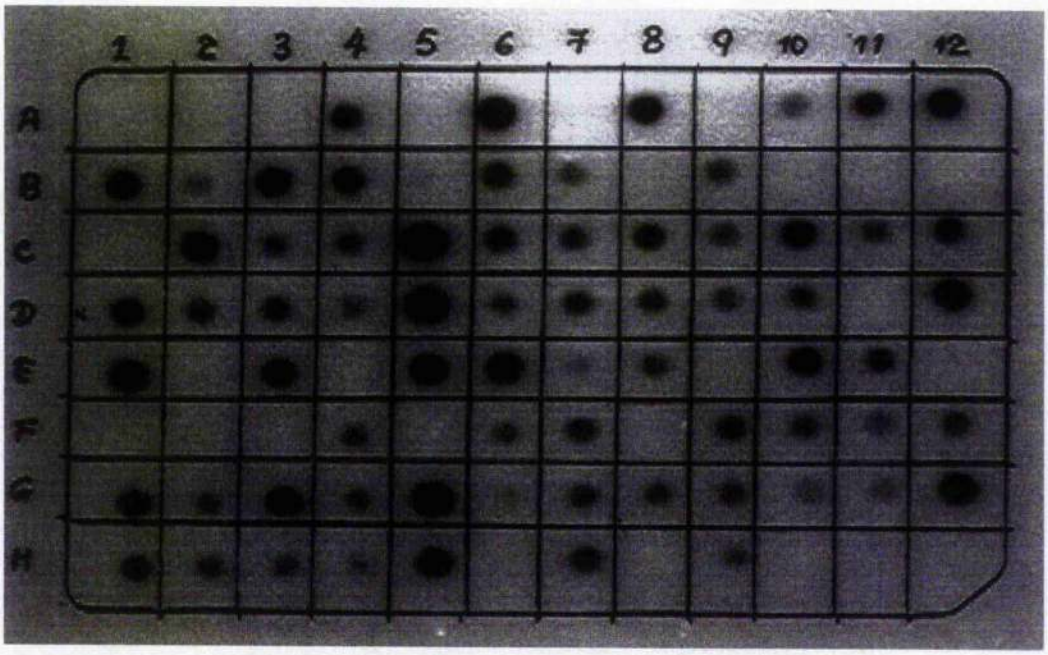

Figure 11: An example of an autoradiograph following hybridization of a dotblot with a radiolabelled probe for 59R DHFR.

Amplification reactions were loaded in duplicate with the negative control and the positive controls 3D7, T996 and Dd2 loaded in wells A1, A2, A3 and A4 respectively and then in H12, H11, H10 and H9 respectively. Note the positive hybridisation with Dd2 control in wells A4 and H9.

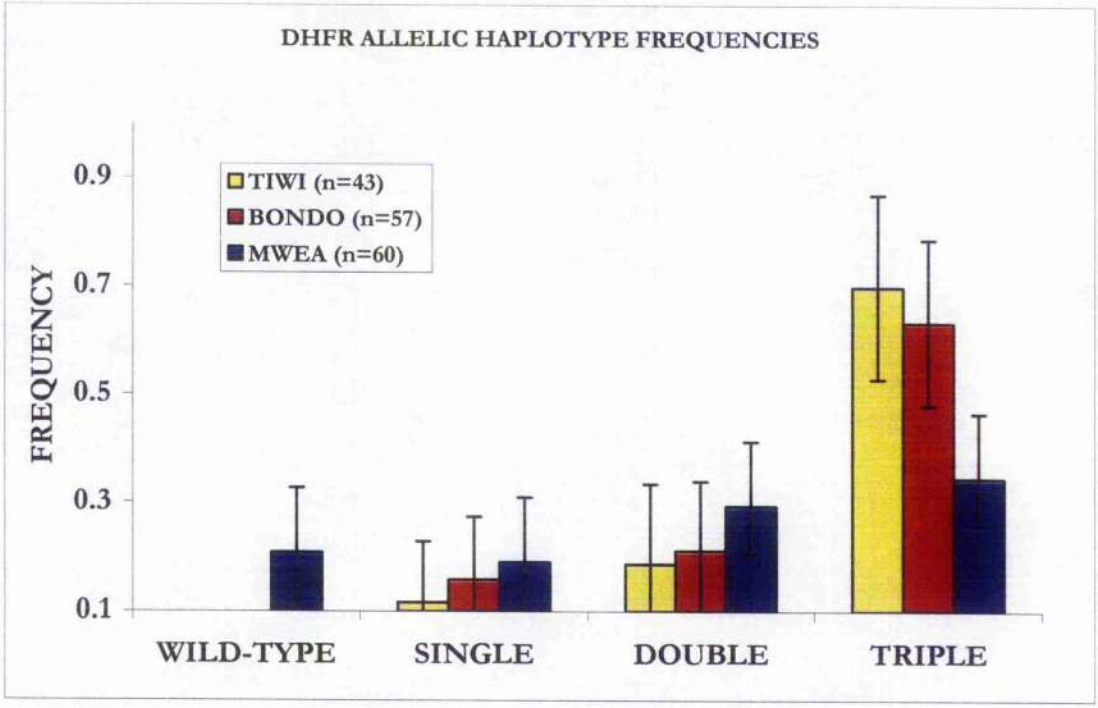

Figure 12: Frequencies of the *dhfr* allelic haplotypes for the three study areas.

Frequency values are shown with 95% CI (bars) calculated by multinomial (3 cases) exact method.

3.5.3 Dhps dot blot hybridization

37, 39 and 56 samples from Bondo, Tiwi and Mwea respectively were included in the analysis for *dhps* alleles. The two codons examined, A437G and K540E, have been associated with resistance to sulphadoxine in *P. falciparum* (Curtis, *et al.* 1998).

13%, 27% and 39% of amplified samples from Tiwi, Bondo and Mwea respectively had wild-type alleles at both codons typed (Table 23 and Figure 13).

	Study Sites		
Genotypes	Bondo (n=37)	Tiwi (n=39)	Mwea (n=56)
Dhps 437A/540K (Wild-type)	10 (27%)	5 (13%)	22 (39%)
Dhps**437G**/540K (Single mutant)	19 (51%)	14 (36%)	10 (18%)
Dhps**437G/540E** (Double mutant)	8 (22%)	20 (51%)	24 (43%)

Table 23: Results of *dhps* genotyping of samples from the three study sites.

The A437G mutation was found as a single mutant (i.e. wild-type at codon 540) in all the three areas, at a frequency of 18%, 36% and 51% in Mwea, Tiwi and Bondo respectively. No samples had the wild-type codon at 437 and a mutated codon 540.

43%, 51% and 22% of samples in Mwea, Tiwi and Bondo respectively had mutant codons at both position 437 and 540 (double mutants) (Figure 13).

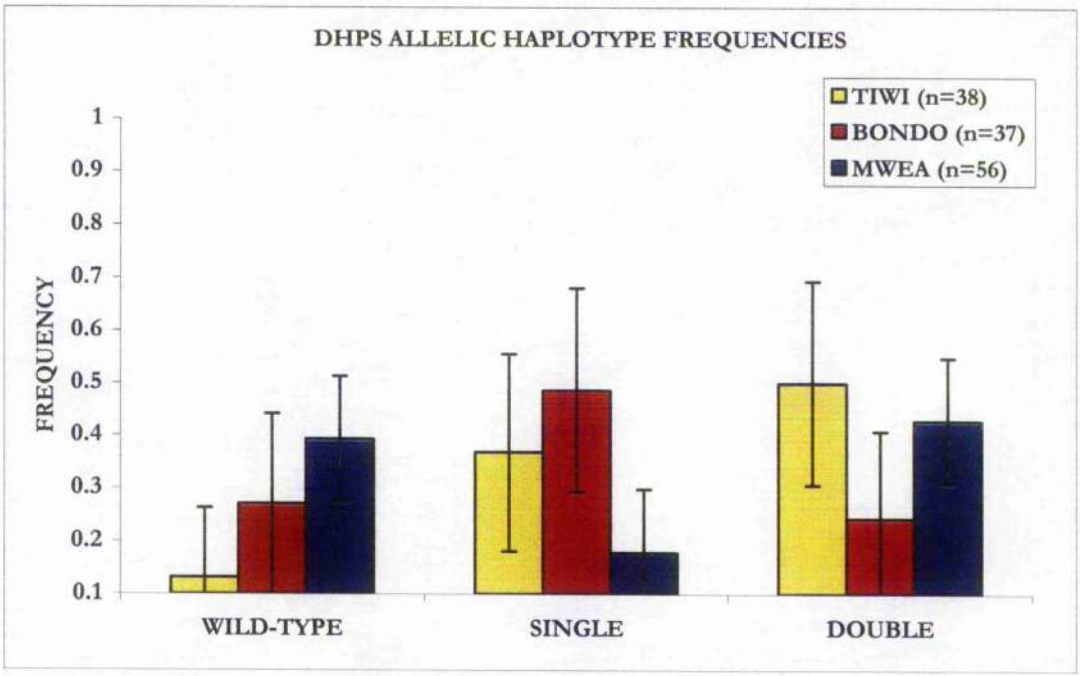

Figure 13: Frequencies of the *dhps* alleles in the three study sites.
Frequency values are shown with 95% CI (bars) calculated by multinomial (3 cases) exact method.

3.5.4 Comparison of the frequencies of dhfr *and* dhps *alleles per site*

The distribution of alleles was analysed using the Fisher's Exact test. A comparison of the distribution of the *dhfr* alleles revealed highly significant differences ($P<0.001$) among the three sites (Table 22). The distribution of the *dhps* alleles was also significantly different between sites (Fisher's exact, $P= 0.0023$) (Table 23). When the analysis was repeated with the exclusion of Mwea (since this where most of the differences were) there were no significant differences between Bondo and Tiwi in the distribution of the genotypes (Fisher's Exact, $P= 0.733$ for *dhfr* and $P= 0.0567$ for *dhps*).

	Study sites		
Genotypes	Tiwi (n=39)	Bondo (n=37)	Mwea (n=56)
DhfrW/DhpsW	0	0	3
DhfrM/DhpsW	5	10	22
DhfrW/DhpsM	0	0	12
DhfrM/DhpsM	34	27	19

Table 24: Multilocus genotypes per study site.

Samples were classified into four categories; those that had no mutations in the two genes (W/W), those with at least one mutation in *dhfr* with none in *dhps* (DhfrM/DhpsW), those with no mutation in *dhfr* but at least one in *dhps* (DhfrW/DhpsM) and those that had at least one mutation in each gene (DhfrM/DhpsM)

The combined distribution of the alleles defined by all codons examined at both genes (multilocus genotype) was analysed by scoring whether samples had at least one mutation in *dhfr* and none in *dhps* (dhfrM/dhpsW), at least one mutation in *dhps* but none in *dhfr* (dhfrW/dhpsM), at least one mutation in both genes (dhfrM/dhpsM) or wild-type in both genes (dhfrW/dhpsW) (Table 24). An analysis using this classification revealed highly significant differences (Fishers Exact, $P<0.0001$) between the three sites. A comparison between Tiwi and Bondo found no significant differences (Fisher's Exact, $P = 0.1557$) in the distribution of substituted alleles.

A comparison between Mwea and either Tiwi or Bondo found highly significant differences (Fisher's Exact, $P< 0.0001$; Table 24) in genotype frequencies in Mwea and in either of the other two sites.

3.5.5 Microsatellite loci around Dhfr

Four microsatellite loci were identified within an 8kb region downstream, and four within an 8kb region downstream of *dhfr* on chromosome 4. Microsatellite DHFRU1 is situated 0.3kb downstream of *dhfr* in the 3D7 genome. Analysis of this locus in the Kenyan samples revealed 24 distinguishable alleles ranging in size from 107 – 152 bp. Analysis of microsatellite locus DHFRU2, located 4.7kb downstream of *dhfr*, revealed 22 alleles of 158 – 192 bp. Similar analysis of microsatellite DHFRU3, located 7.0kb downstream of *dhfr*, revealed 21 alleles of 170- 195 bp, and of DHFRU4, located 8.0kb downstream of *dhfr*, revealed 24 distinguishable alleles of 187-238 bp.

Upstream of *dhfr*, analysis of microsatellite DHFRD1 (located 0.5kb) revealed 24 alleles of 88-152 bp, DHFRD2 (located 4.0kb) had 26 alleles of 230-269 bp, DHFRD3 (located 6.5kb) had 28 alleles of 175-229 bp, and DHFRD4 (located 8.0kb) had 17 alleles of 263-300 bp. The number of alleles could be an underestimation as all the sizes were determined following electrophoresis on Spreadex gels, where the minimum size difference distinguishable is 4bp.

3.5.5.1 Microsatellite haplotypes and dhfr mutations

Samples with multiple alleles at more than one of the microsatellite loci preclude the construction of unambiguous haplotypes. Therefore only samples with single alleles at the microsatellite loci were used to construct microsatellite haplotypes. Haplotypes were ranked according to allele size at DHFRU1, then at DHFRU2, DHFRU3, etc. until finally at DHFRD4 loci, and numbered 1-59 when all the three sites (n=160 samples) were combined. Full haplotype descriptions are provided in Appendix 2.

	Haplotype codes			
DHFR allele	Overall (n= 131)	Mwea (n=60)	Tiwi (n=43)	Bondo (n=57)
108S (Wild-Type)	3, 15, 19, 20, 22, 30, 48, 49, 53, 55, 58, 59	Mwea had each of these haplotypes each with n=1	None	None
108N/51N/59C (single mutant)	1, 4, 6, 7, 8, 9, 15, 16, 17, 18, 21, 23, 24, 25, 31, 46, 47, 51, 52, 54, 56, 57	6 (n=1) 17 (n=1) 18 (n=1) 23 (n=1) 31 (n=1) 47 (n=1) 51 (n=1) 52 (n=1) 54 (n=1) 57 (n=1)	2 (n=1) 4 (n=1) 7 (n=1) 15 (n=1) 16 (n=1)	1 (n=1) 4 (n=1) 8 (n=1) 9 (n=1) 21 (n=1) 24 (n=1) 25 (n=1) 46 (n=1) 56 (n=1)
108N/51I/59C (Double mutant)	10, 12, 13, 14	12 (n=2) **13 (n=3)** 14 (n=1)	**10 (n=3)**	12 (n=1) 13 (n=1) **14 (n=3)**
108N/51N/**59R** (Double mutant)	11, 26, 27, 29, 50	26 (n=1) 27 (n=2) **29 (n=8)**	**26 (n=3)** 27 (n=2)	11 (n=1) 27 (n=1) **29 (n=3)** **50 (n=3)**
108N/51I/59R (Triple mutant)	5, 28, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45	**33 (n=19)** 43 (n=1)	**35 (n=25)** 37 (n=1) 38 (n=1) 43 (n=1) 44 (n=1) 45 (n=1)	5 (n=1) 28 (n=1) **33 (n=16)** 34 (n=2) 35 (n=5) 36 (n=4) 39 (n=4) 40 (n=1) 41 (n=2)

Table 25: Microsatellite haplotypes associated with different *dhfr* alleles per study site.

The major allele (with at least 50% of samples with a particular *dhfr* allele per site) is indicated in bold.

The results are shown in Table 25 and the haplotype frequencies for each allele of *dhfr* (as defined by polymorphism at codons 51, 59 and 108) are presented in figures 14 – 17. Overall 59 haplotypes were identified in the 160 samples from Tiwi, Mwea and

Bondo. There were 13 distinguishable haplotypes in the samples from Tiwi, 20 from Bondo and 30 from Mwea.

The *dhfr* triple mutant encoding N51I, C59R and S108N was associated with 15 microsatellite haplotypes (Table 25). Out of the 86 samples typed as triple mutants, 35 samples (43.75%) had haplotype 33 (composed of alleles 133 bp/174 bp/188 bp/212 bp/100 bp/244 bp/183 bp/283 bp at the DHFRU1, DHFRU2, DHFRU3, DHFRU4, DHFRD1, DHFRD2, DHFRD3, DHFRD4 respectively) while 30 samples (35.71%) had haplotype 35 (133 bp/174 bp/188 bp/212 bp/100 bp/244 bp/183 bp/291 bp). Haplotypes 37, 38, 43, 44 and 45 were detected only in Tiwi at quite low frequencies (Table 25). Haplotypes 36, 39, 40 and 41 were detected only in Bondo in 4, 4, 1 and 2 samples respectively. 19 out of 20 and 16 out of 34 triple mutant samples from Mwea and Bondo respectively contained parasites with haplotype 33. 25 out of 30 (95%) samples from Tiwi contained parasites with haplotype 35 (Table 25).

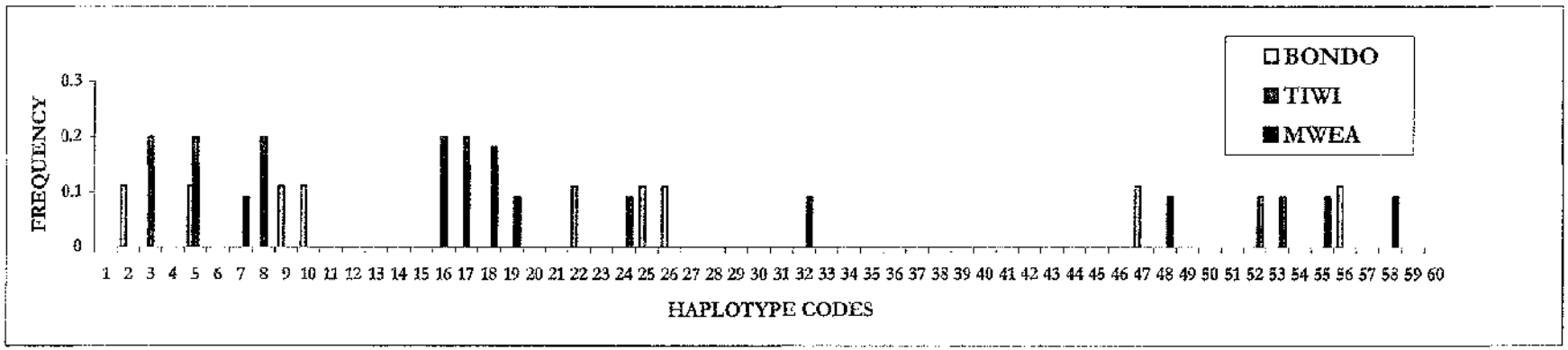

Figure 14: Frequencies of microsatellite haplotypes associated with the 108N *dhfr* single mutant allele in Bondo, Mwea and Tiwi.

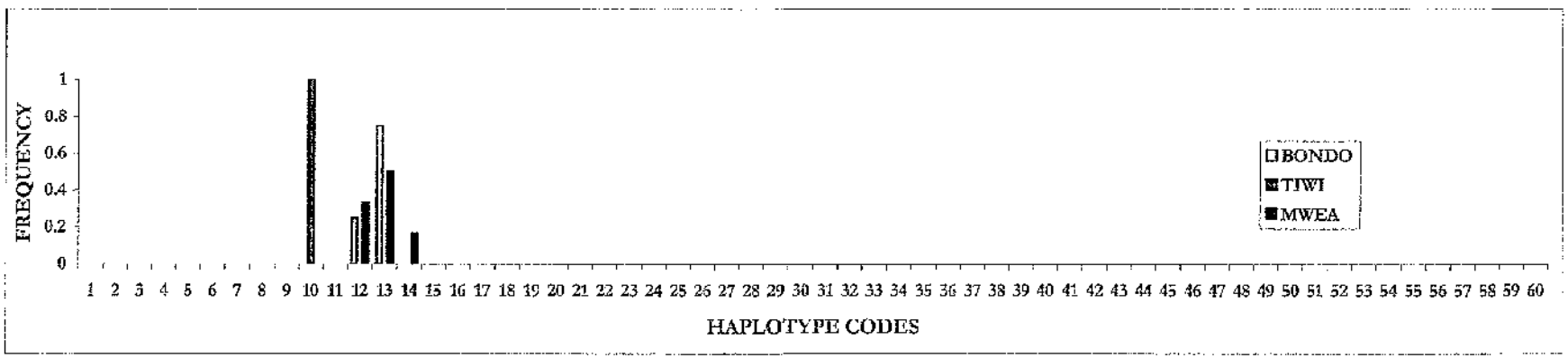

Figure 15: Frequencies of microsatellite haplotypes associated with 51I/108N *dhfr* double mutant in Bondo, Mwea and Tiwi

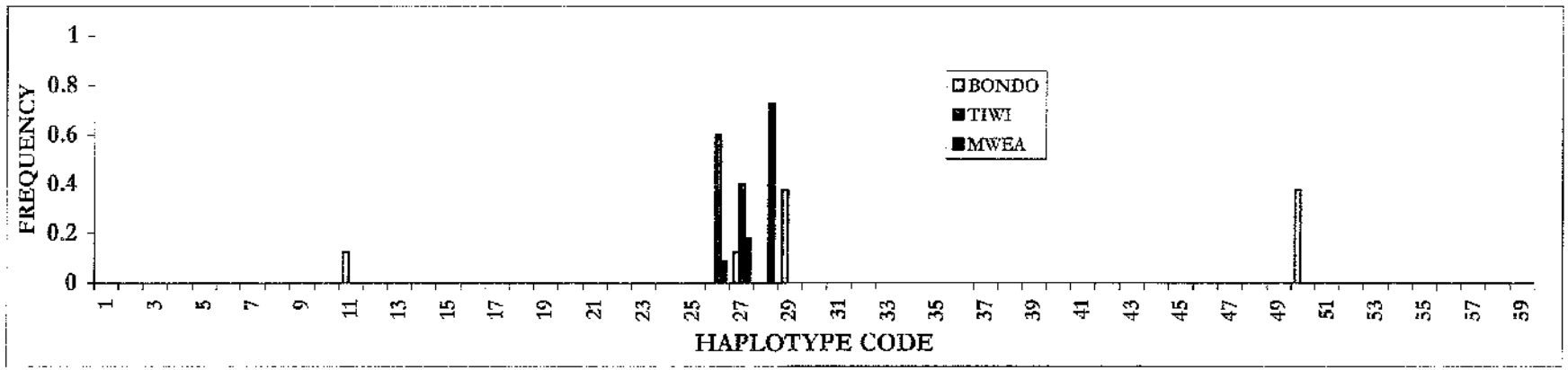

Figure 16: Frequencies of microsatellite haplotypes associated with 59R/108N *dhfr* double mutant in Bondo, Mwea and Tiwi.

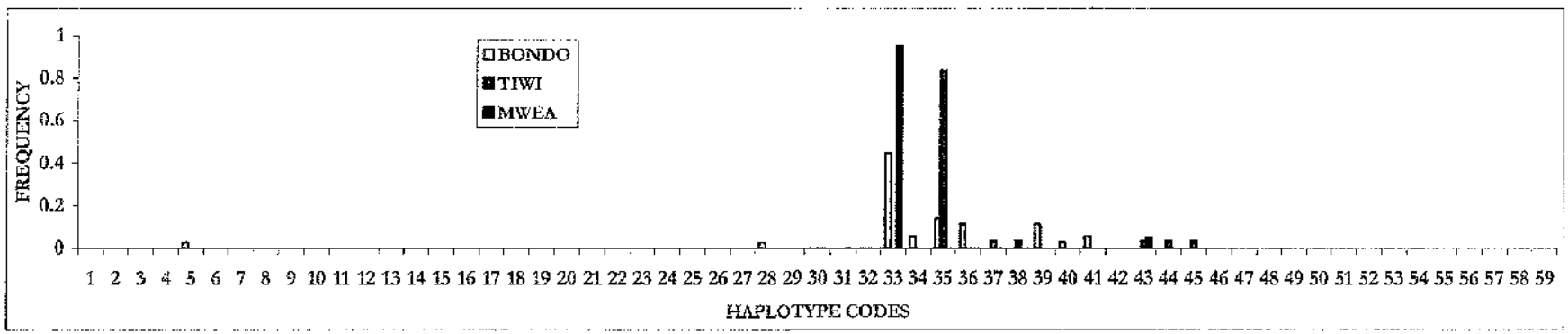

Figure 17: Frequencies of microsatellite haplotypes associated with the *dhfr* triple mutant in Bondo, Mwea and Tiwi.

3.5.5.2 Haplotype distribution among sites

Some haplotypes were found in more than one of the three sites while others were unique to one of the sites ('private haplotypes'). Eleven haplotypes were shared between the sites but only one haplotype was found to be common in the three sites studied: haplotype 27, which was associated with the 108N/59R double mutant (Figure 18). The site with the highest number of private haplotypes was Mwea with 22, while the site with the lowest number of private haplotypes was Tiwi, which had 8. Bondo had 17 private haplotypes (Figure 18).

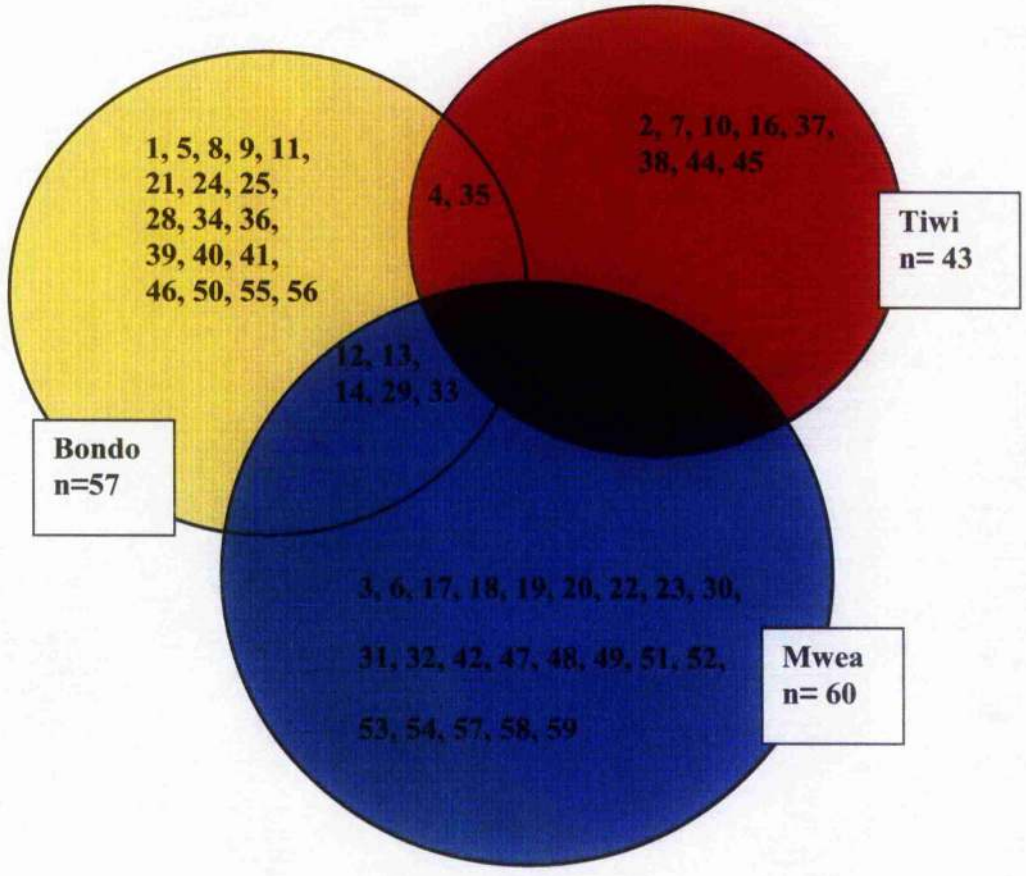

Figure 18: Distribution of haplotypes in the three study sites

Only 11 out of the total 59 haplotypes observed are shared between sites out of which only one (no. 27) is common among all three sites. n = number of haplotypes seen.

The two sites with the highest number of shared haplotypes were Bondo and Mwea with 5 shared haplotypes. Tiwi shared two and three haplotypes with Bondo and Mwea respectively (Figure 18).

3.5.6 Microsatellite loci around dhps

Several microsatellite loci were identified around *dhps* and eight (four on either side of the gene) were selected for this study. Downstream of the gene, 23 alleles with PCR products ranging in size from 120-165 bp were found in the 131 samples analysed at DHPSU1 (located 0.1kb downstream of *dhps*). Similarly, 19 alleles of 192- 232 bp were found at DHPSU2 (located 3.8kb downstream), 20 alleles of 113-170 bp at DHPSU3 (located 6.5kb downstream), and 22 alleles of 185-240 bp at DHPSU4 (located 8.7kb downstream).

Upstream of the *dhps* gene, 20 alleles of 104-124 bp were observed in the samples at the DHPSD1 locus (at 0.4kb), 24 alleles of 205-280 bp at DHPSD2 (at 3.7kb), 17 alleles of 190-246 bp at DHPSD3 (at 7.3kb), and 18 alleles of 189-230 bp at DHPSD4 (at 8.5kb).

Haplotypes observed from all the three sites (n=131 samples) were combined for numbering purposes. Haplotypes were ranked according to allele size at DHPSU1, then at DHPSU2, DHPSU3, and finally at DHPSD4 loci. The haplotypes were then numbered 1-82. Most of the haplotypes (49) were associated with the samples from the Mwea site (Table 26).

Full haplotype descriptions are provided in Appendix 3.

	Haplotype codings			
Dhps alleles	Overall (n=131)	Bondo (n=37)	Tiwi (n= 38)	Mwea (n=56)
Wild-Type	1, 2, 4, 5, 6, 7, 8, 12, 17, 18, 22, 23, 26, 27, 39, 46, 49, 50, 51, 52, 53, 55, 56, 58, 59, 61, 67, 70, 71, 73, 74, 75, 77, 78, 82	1 (n=1), 2 (n=1) 5 (n=1) 7 (n=2) 8 (n=1) 22 (n=1) 23 (n=1) 49 (n=1) 75 (n=1)	4 (n=1) 6 (n=1) 17 (n=1) 59 (n=1) 61 (n=1)	6, 12, 18, 26, 27, 39, 46, 50, 51, 52, 53, 55, 56, 58, 67, 70, 71, 73, 74, 77, 78, 82 (n=1 for each)
437G/540K Single mutant	3, 9, 10, 14, 16, 19, 20, 21, 28, 30, 31, 32, 33, 35, 36, 45, 50, 51, 62, 64, 65, 66, 76, 79, 80	3 (n=1) **9 (n=3)** 16 (n=1) **28 (n=4)** 30 (n=1) 31 (n=1) **35 (n=4)** 36 (n=2) 66 (n=1)	3 (n=1) 10 (n=2) 16 (n=1) **19 (n=4)** 21 (n=2) 32 (n=1) 33 (n=1) 36 (n=1) 66 (n=1)	14, 20, 50, 51, 62, 64, 65, 76, 79, 80 (n=1 for each)
437G/540E Double Mutant	11, 13, 15, 24, 25, 29, 34, 37, 38, 40, 41, 42, 43, 44, 47, 48, 51, 53, 54, 57, 60, 63, 68, 69, 72, 81	41 (n=1) **42 (n=5)** 43 (n=2) 48 (n=1)	40 (n=1) **41 (n=3)** **42 (n=14)** 43 (n=1) 48 (n=1)	11, 13, 15, 24 (n=2), 25, 29, 34, 37, 38, 44, 47, 51 (n=2), 53, 54, 57, 60 63, 68 (n=2), 69, 72, 81 (n=1 for each except for those indicated in brackets)

Table 26: Codes for haplotypes and their prevalence in the 3 study sites

The wild-type *dhps* alleles (n=37) from all three sites were not found to be strongly associated with any particular microsatellite haplotype; a wide range of haplotypes was observed and no one haplotype predominated (Table 26 and Figure 21).

The allele of *dhps* encoding the 437G change (n=42) was associated with 29 haplotypes. All the alleles with the single *dhps* mutation from Mwea had different haplotypes (Table 26 and Figure 22). In Tiwi and Bondo some haplotypes were observed to be more prevalent in the single mutant *dhps* than others (Figure 22). The most common haplotype found in samples from the Tiwi site with the 437G allele was haplotype 19 (42.86%). The two most common haplotypes associated with the 437G

dhps mutant in Bondo were haplotype 28 (22.22%) and haplotype 35 (22.22%). These two haplotypes were not found in Tiwi while haplotype 13 was not found with the Bondo samples. Four haplotypes (3, 16, 36 and 66) were found in single mutant *dhps* alleles from both Tiwi and Bondo but in quite low frequencies (Figure 22).

The 437G/540E *dhps* double substituted allele in Tiwi and Bondo (n= 28) was associated with haplotypes 40, 41, 42, 43 and 48 (Figure 19 and Figure 23). The most common haplotype associated with the double mutant *dhps* allele was haplotype 42 with a frequency of 56% and 79% in Bondo and Tiwi respectively (Figure 23). There was no dominant haplotype associated with the double substitute *dhps* allele in Mwea (Figure 23).

While there seemed to be apparent reduction in the number of haplotypes associated with the single and double mutant *dhps* samples from Bondo and Tiwi, those from Mwea seemed to have multiple haplotypes, all found at very low frequencies. For example, the 437G/540E allele found in Mwea was associated with 21 haplotypes in 24 samples, all with frequencies <0.1. (Table 26, Figure 20, Figure 22 and Figure 23).

3.5.6.1 *Dhps*-linked microsatellite haplotype sharing among the three sites

There was less sharing of haplotypes of the *dhps*-linked microsatellites than was observed with the *dhfr*-linked ones. Mwea did not share any haplotype with Bondo and was only found to share one haplotype with Tiwi (Figure 19). Tiwi and Bondo shared 8 of the total haplotypes of the microsatellites flanking the *dhps* gene. 50% of the shared haplotypes between Bondo and Tiwi were those associated with the double-substituted *dhps* allele (Table 26 and Figure 19) indicating a close relationship of the double mutant *dhps* allele between the two sites.

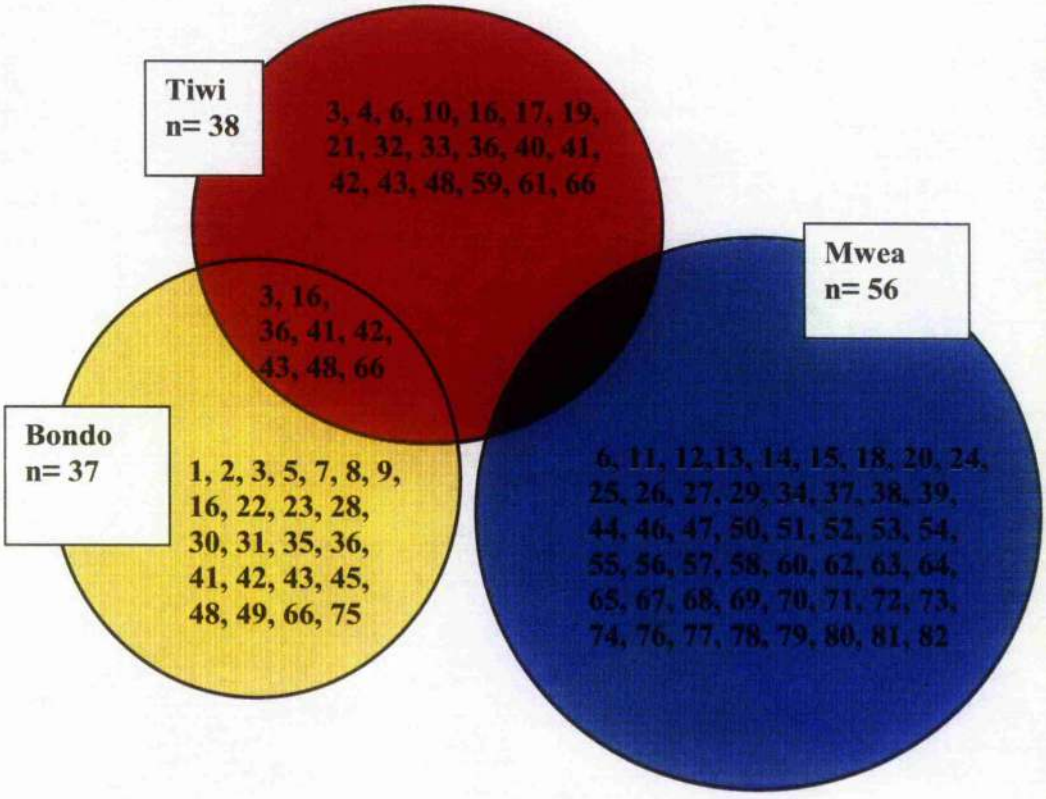

Figure 19: *Dhps*-linked microsatellite haplotype distribution in the 3 sites.

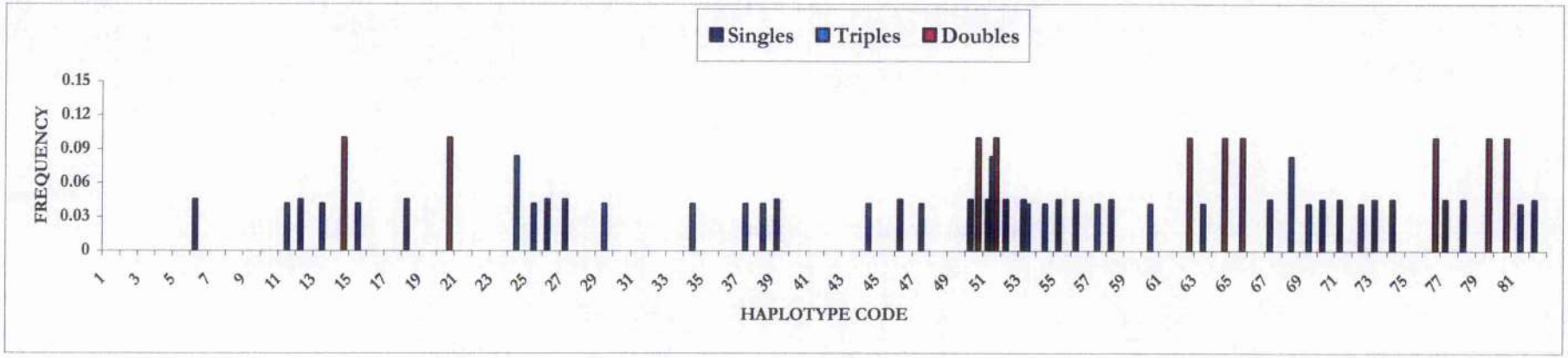

Figure 20: Haplotype frequencies emphasizing lack of selection for Mwea *dhps* mutants

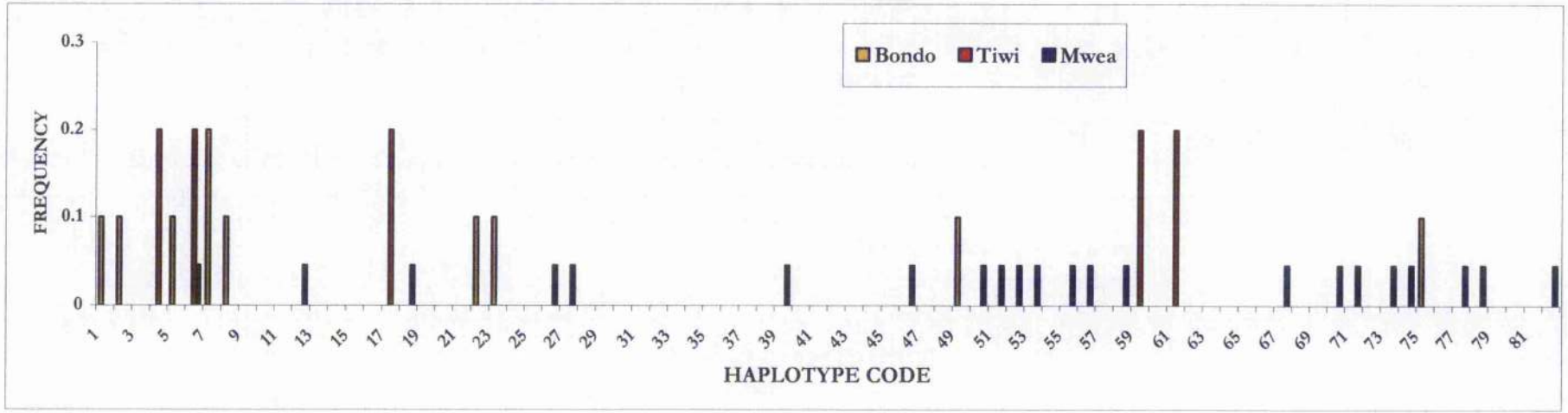

Figure 21: Haplotype alleles associated with the wild-type *dhps* in the three sites

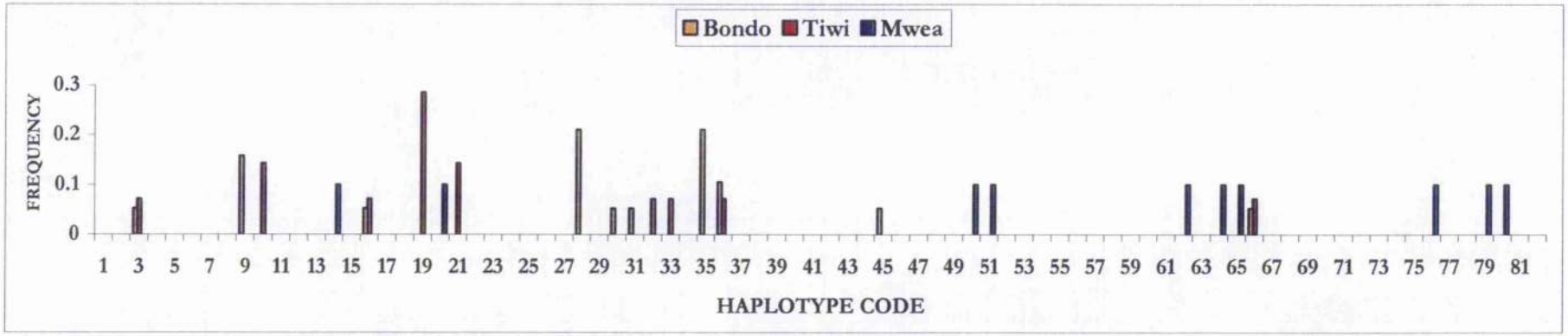

Figure 22: Haplotype frequencies associated with the single mutant *dhps* allele.

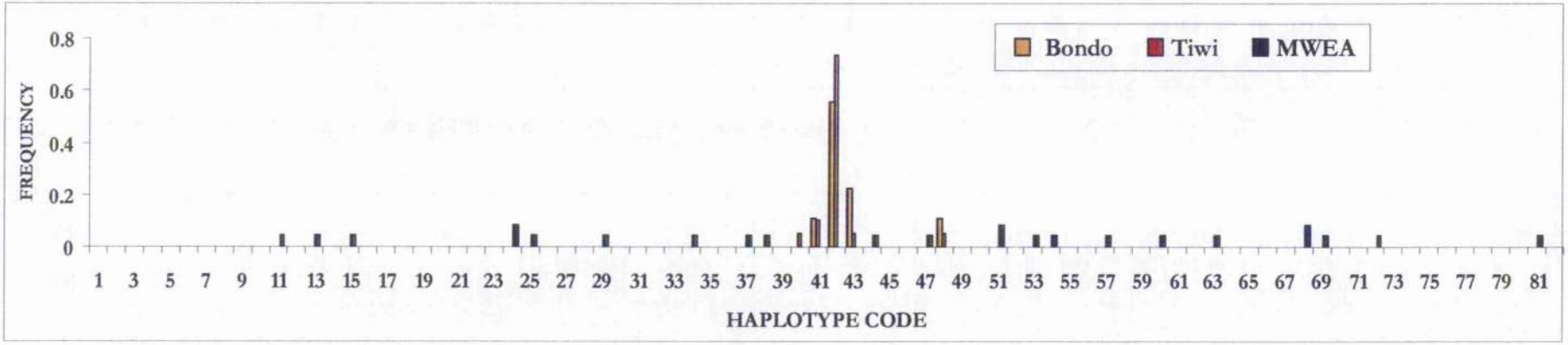

Figure 23: Comparison of haplotype frequencies associated with double mutants *dhps* for the 3 sites.

3.5.7 Expected Heterozygosity in the flanking microsatellites

Expected heterozygosity (*He*) at each microsatellite locus was calculated as explained in section 2.1.7.1. Alleles of microsatellites were categorised according to the variation at all three of the codons examined for *dhfr* and two codons examined for *dhps*. Thus, for *dhfr* there were four categories of alleles; wild-type, single mutants (**108N**/51N/59C), double mutants (**108N/51I**/59C and/or **108N**/51N/**59R**) and triple mutants (**108N/51N/59R**). For *dhps* there were three categories of alleles; wild-type, single mutants (**437G**/540K) and double mutants (**437G/540E**). Based on these categories, *He* at the genes (both *dhfr* and *dhps*) was set at zero, as it is presumed there is no variation within the gene when the same alleles are considered. A comparison of the *He* values of the microsatellites between the wild-type and triple mutant alleles revealed markedly reduced variation around the *dhfr* alleles carrying the triple mutations. The reduction in variation was more pronounced upstream of the gene compared to the downstream region. *He* values were lower nearer the gene than further from it (Figure 24). This is in agreement with the expectations of a selective sweep where the regions carried over depend largely on the strength of the selective force (i.e. antimalarial use) and the strength of recombination. In contrast the wild-type alleles show almost equal variation (as measured by *He*) with the unlinked microsatellites. This is discussed further in Section 4.5. A similar scenario was observed with the *dhps* alleles where those carrying the double mutations have reduced *He* at markers flanking the gene compared to those carrying the wild-type gene Figure 25.

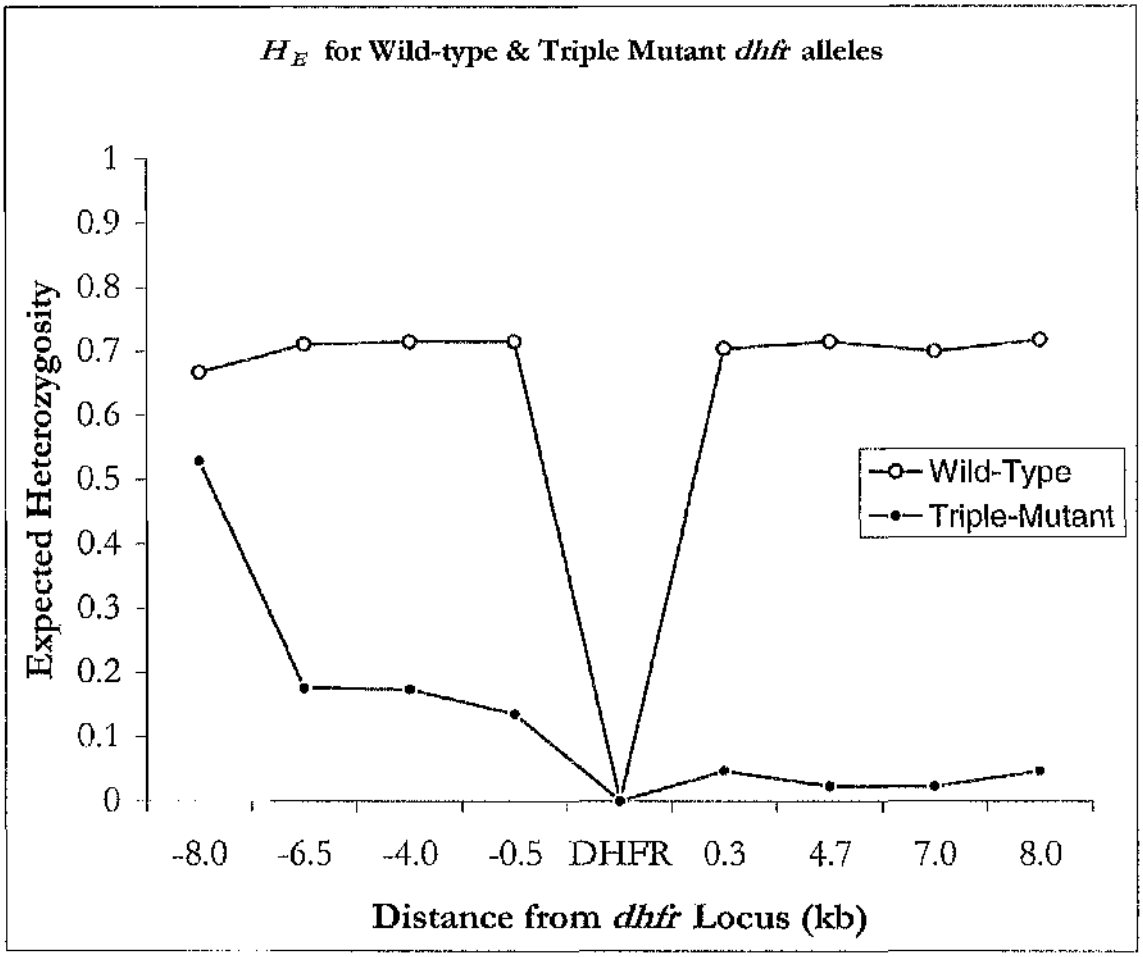

Figure 24: Microsatellites flanking *dhfr* resistant alleles (triple mutant) show reduced *He* compared with those flanking wild-type *dhfr* alleles.

All sites have been combined for this analysis. The X-axis refers to the location (in Kb) of the microsatellite loci used with respect to the *dhfr* gene. Filled dots represent H_e of microsatellite markers flanking *dhfr* triple mutant alleles and open dots represent H_e of microsatellites flanking wild-type *dhfr* alleles from the three study areas. Note the pronounced asymmetry of the mutant allele curve.

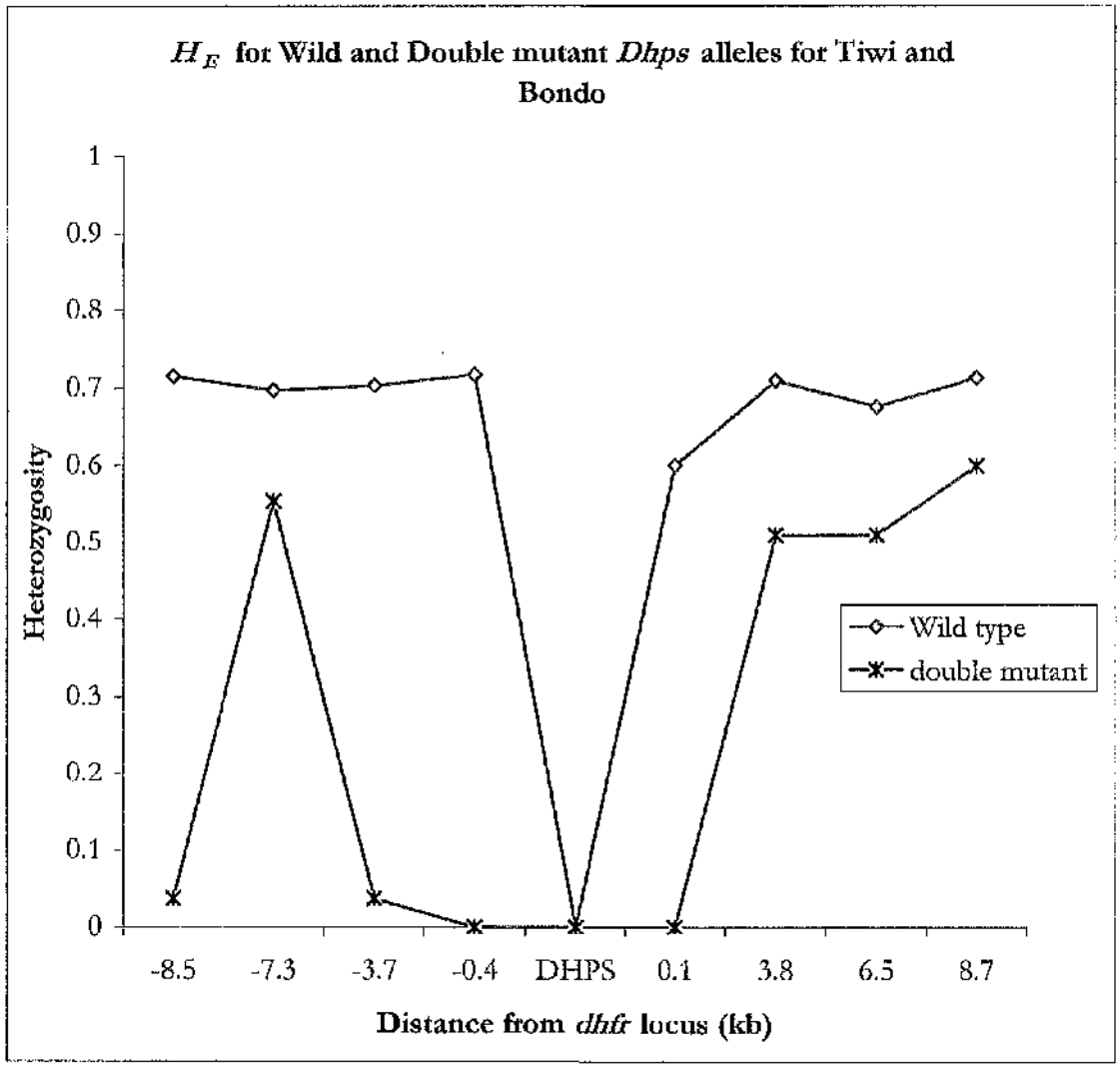

Figure 25: Variation (*He*) at microsatellite loci flanking wild-type and double mutant *dhps* alleles in Tiwi and Bondo

The x-axis refers to the location (in kb) of the microsatellite loci used with respect to the *dhps* gene. Note the variation at the gene is set at 0.

3.5.8 Extended Haplotype Homozygosity

To explore how linkage disequilibrium breaks down with increasing distance to a specified core region, haplotype homozygosity (HH) was calculated in a stepwise manner for each haplotype length (extended HH, EHH) according to published methods (Sabeti *et al.*, 2002). HH is calculated between a distance x and the specified core region for a chromosome population carrying a single core haplotype. Distance x increases stepwise to the most outlying marker. The procedure was repeated for each core haplotype. If the candidate region is not at the margin of the genotyped range, linkage disequilibrium patterns can be estimated on both sides of each core haplotype. In other words, the extended HH estimates the level of haplotype splitting due to recombination and mutation at each distance to the core haplotype. An attractive aspect of this approach is that the various core haplotypes at a locus serve as internal controls for one another at the same chromosomal region. This is important given the variability of local recombination rates across genomic regions. The values of EHH range from 0 to 1 with a high EHH (approaching 1) indicating complete linkage disequilibrium (LD) between the marker and the gene and vice versa. The results of EHH are given in Figure 26and Figure 27.

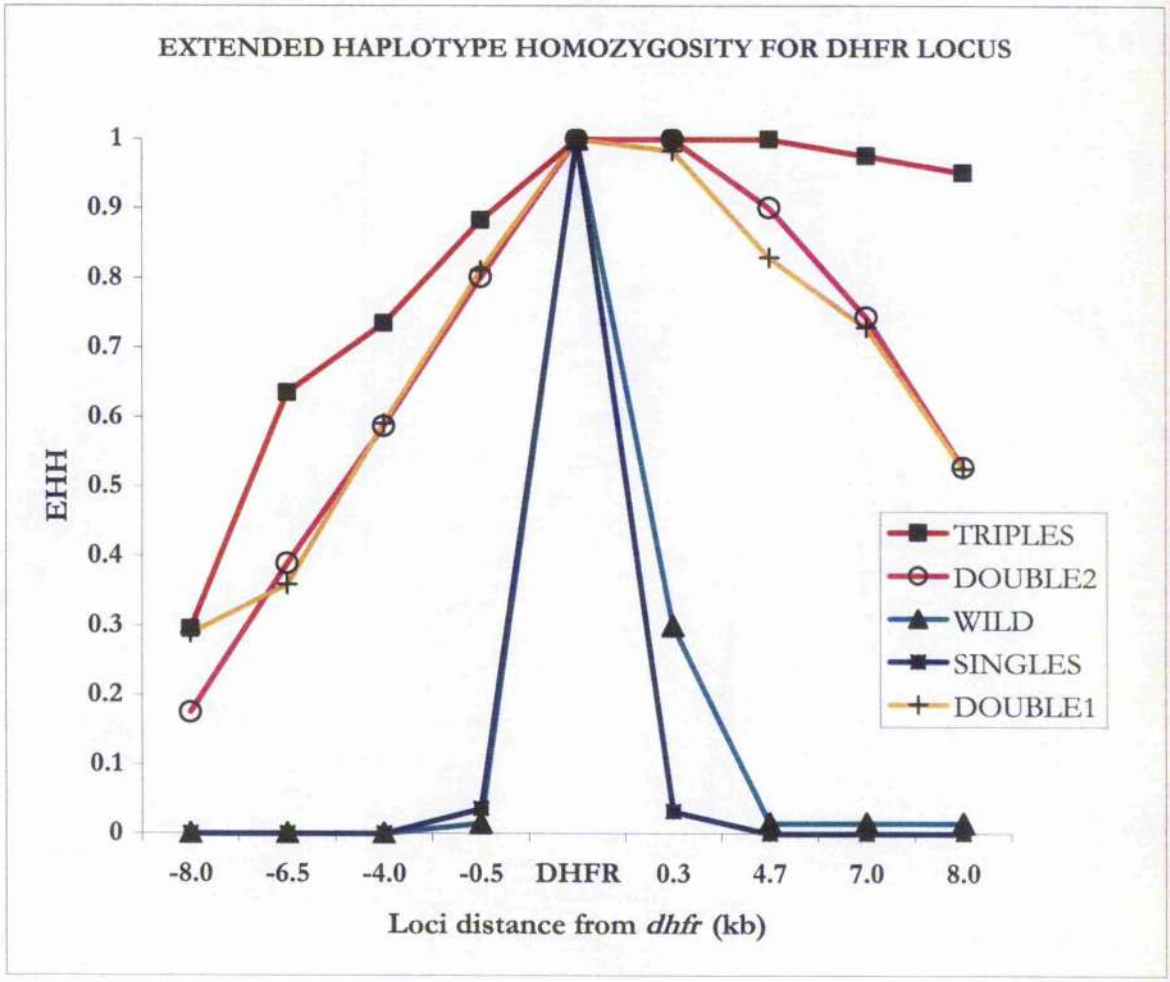

Figure 26: LD, measured by EHH, in *dhfr* wild-type and mutant parasites from Tiwi, Bondo and Mwea.

EHH for *dhfr* was set at 1. 'Double1' and 'Double2' refer to the mutants having 51I/108N and 59R/108N mutations respectively. 'Singles' refer to parasites with the 108N mutation only. Triples refer to parasites with the *dhfr* allele 51I/59R/108N.

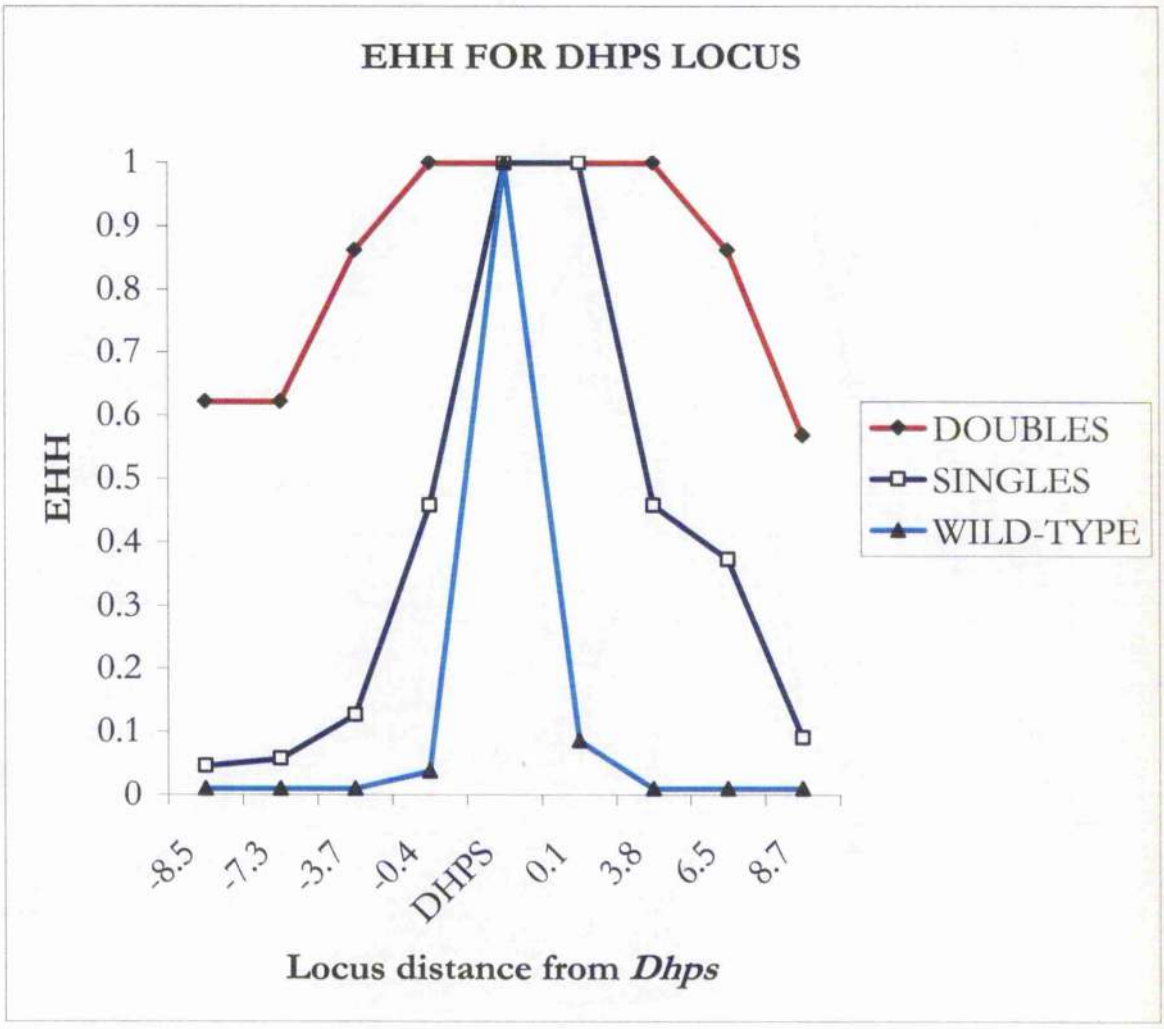

Figure 27: LD, measured by EHH, for *dhps* wild-type and mutant parasites from Tiwi, Bondo.

Wild-type refer to parasites with no mutations in *dhps*, singles are those parasites with the 437G while doubles are the ones with the allele 437G/540E.
EHH for *dhps* was set at 1.

3.5.9 Testing for population bottlenecks

It has been noted that allelic diversity is reduced faster than heterozygosity during a bottleneck (Nei *et al.* 1975). If a population has gone through a bottleneck, it is found to retain the excess heterozygosity.

To test whether the population of substituted *dhfr* alleles had undergone a recent bottleneck, the presence of excess of heterozygosity was determined using the software '*bottleneck*' (Cornuet and Luikart 1996). This software carries out a Wilcoxon signed rank test comparing observed heterozygosity at each locus across the ~17 kb region flanking the triple mutant *dhfr*. The same analysis was repeated for double mutant *dhps* loci. These analyses were carried out using the both the site-specific and pooled data for *dhfr*, and with the exclusion of Mwea for the pooled *dhps* data due to the observation that Mwea's double mutant *dhps* mutants had not been selected for. Expected values are generated under infinite allele model (IAM) and step-wise allele model (SMM).

It was found that under both IAM and SMM there was no significant excess of expected heterozygosity for the triple mutant *dhfr* isolates from all three areas (Wilcoxon test one tail P=0.99 [IAM] and P=1.000 [SMM]) indicating that the population had not recently undergone a reduction in size (Table 27).

The same result was observed when analysing samples with the double mutant *dhps* allele from Bondo and Tiwi, excluding Mwea (Wilcoxon test one tail P=0.22 [IAM] and P=0.34 [SMM]) (Table 28).

Locus	DhfrU1	DhfrU2	DhfrU3	DhfrU4	DhfrD1	DhfrD2	DhfrD3	DhfrD4
Empirical data								
Sample size (n)	86	86	86	86	86	86	86	86
Observed heterozygosity (H_O)	0.046	0.023	0.023	0.046	0.135	0.173	0.175	0.528
No. of alleles	3	2	2	3	5	3	4	4
IAM								
Average heterozygosity (H_E)	0.337	0.195	0.196	0.338	0.509	0.337	0.437	0.230
Standard deviation	0.180	0.168	0.166	0.179	0.169	0.181	0.167	0.175
P Heterozygote excess	0.050	0.218	0.212	0.039	0.018	0.254	0.093	0.384
SMM								
Average heterozygosity (H_E)	0.463	0.231	0.233	0.461	0.674	0.465	0.585	0.587
Standard deviation (S.D.)	0.137	0.165	0.171	0.140	0.083	0.139	0.111	0.108
P Heterozygote excess	0.003	0.119	0.138	0.005	0.000	0.460	0.003	0.230

Table 27: Results obtained using the Bottleneck software for samples carrying the triple mutant *dhfr* allele from Tiwi, Bondo and Mwea.

H_E and SD are the average heterozygosity and standard deviation of the distribution of the heterozygosity *(H)* obtained in simulated samples (1000 replicates) with a size and number of alleles identical to the empirical data, assuming mutation-drift equilibrium (null hypothesis). The overall *P* for heterozygote excess is 0.998 (IAM) and 1.000 (SMM).

Locus	DhpsU1	DhpsU2	DhpsU3	DhpsU4	DhpsD1	DhpsD2	DhpsD3	DhpsD4
Empirical data								
Sample size (n)	28	28	28	28	28	28	28	28
Observed heterozygosity (H_O)	0.000	0.037	0.553	0.037	0.000	0.509	0.509	0.599
No. of alleles	1	2	3	2	1	2	2	4
IAM								
Average heterozygosity (H_E)	-	0.211	0.367	0.215	-	0.218	0.209	0.479
Standard deviation	-	0.163	0.172	0.162	-	0.167	0.162	0.165
P Heterozygote excess	-	0.234	0.152	0.234	-	0.025	0.019	0.271
SMM								
Average heterozygosity (H_E)	-	0.256	0.473	0.257	-	0.259	0.252	0.609
Standard deviation (S.D.)	-	0.166	0.136	0.168	-	0.165	0.166	0.099
P Heterozygote excess	-	0.156	0.314	0.159	-	0.037	0.020	0.386

Table 28: Results of the Bottleneck software for isolates carrying the double mutant *dhps* alleles from Bondo and Tiwi sites combined.

H_E and SD are the average heterozygosity and standard deviation of the distribution of the heterozygosity *(H)* obtained in simulated samples (1000 replicates) with a size and number of alleles identical to the empirical data, assuming mutation-drift equilibrium (null hypothesis). The overall *P* for heterozygote excess (one tail) is 0.219 (IAM) and 0.344 (SMM).

3.5.10 Modeling for selective sweep in Kenyan parasite populations

A very basic model was employed to investigate the rate at which drug-resistance alleles replace drug-sensitive ones in areas of Kenya with different endemicity of *P. falciparum,* under drug pressure with SP. This model was also used to determine rate of selection using SP data collected over the years in the three areas. This model is described in section 2.2.5.5. For the coefficient to be significant the fit of the points (R^2) has to be >0.80 (Hartl & Dykhuizen 1981). The results are presented in Figure 28, Figure 29 and Figure 30.

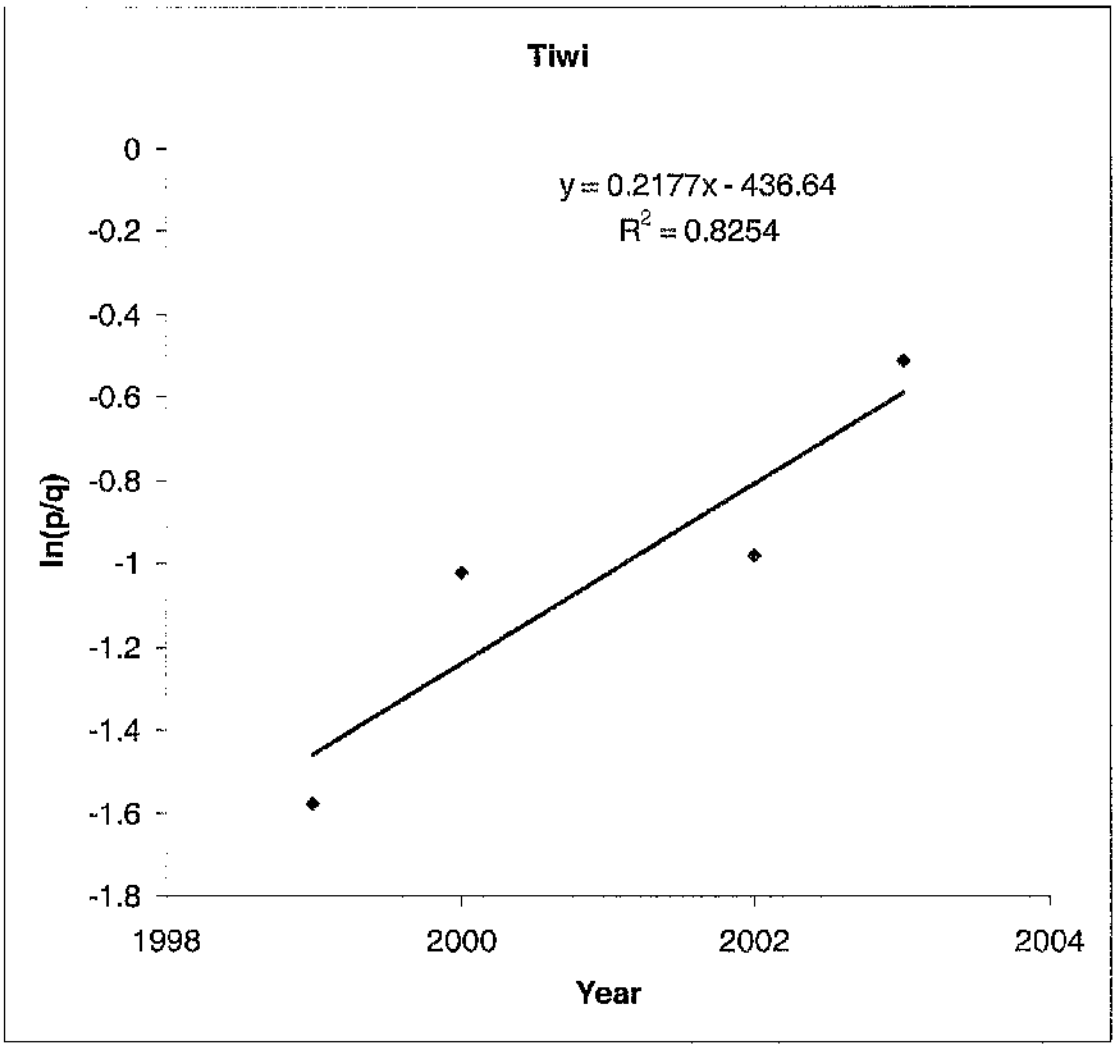

Figure 28: Estimation of selection coefficient (*s*) driving SP resistance in Tiwi.
p is the frequency of resistant alleles inferred from treatment failure rate data, and *q* (=1-*p*) is the inferred frequency of sensitive alleles. The *x*-axis shows the generation number (assuming six generations per year (Joy *et al.*, 2003). The slope gives an estimate of *s* = 0.2177.

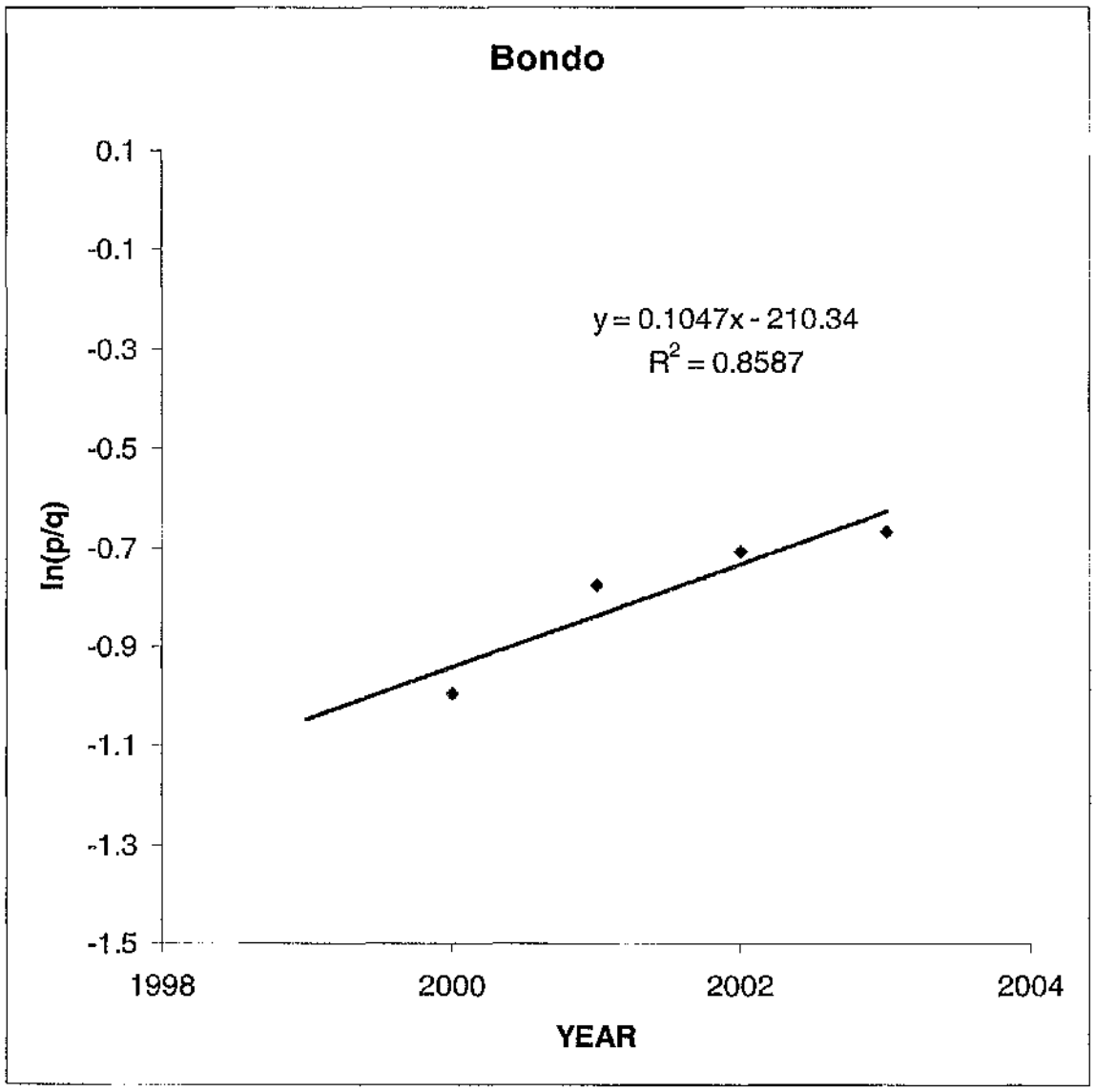

Figure 29: Estimation of selection coefficient (*s*) driving SP resistance in Bondo. *p* is the frequency of resistant alleles inferred from treatment failure rate data, and *q* (=1-*p*) is the inferred frequency of sensitive alleles. The *x*-axis shows the generation number (assuming six generations per year). The slope gives an estimate of *s* = 0.1047.

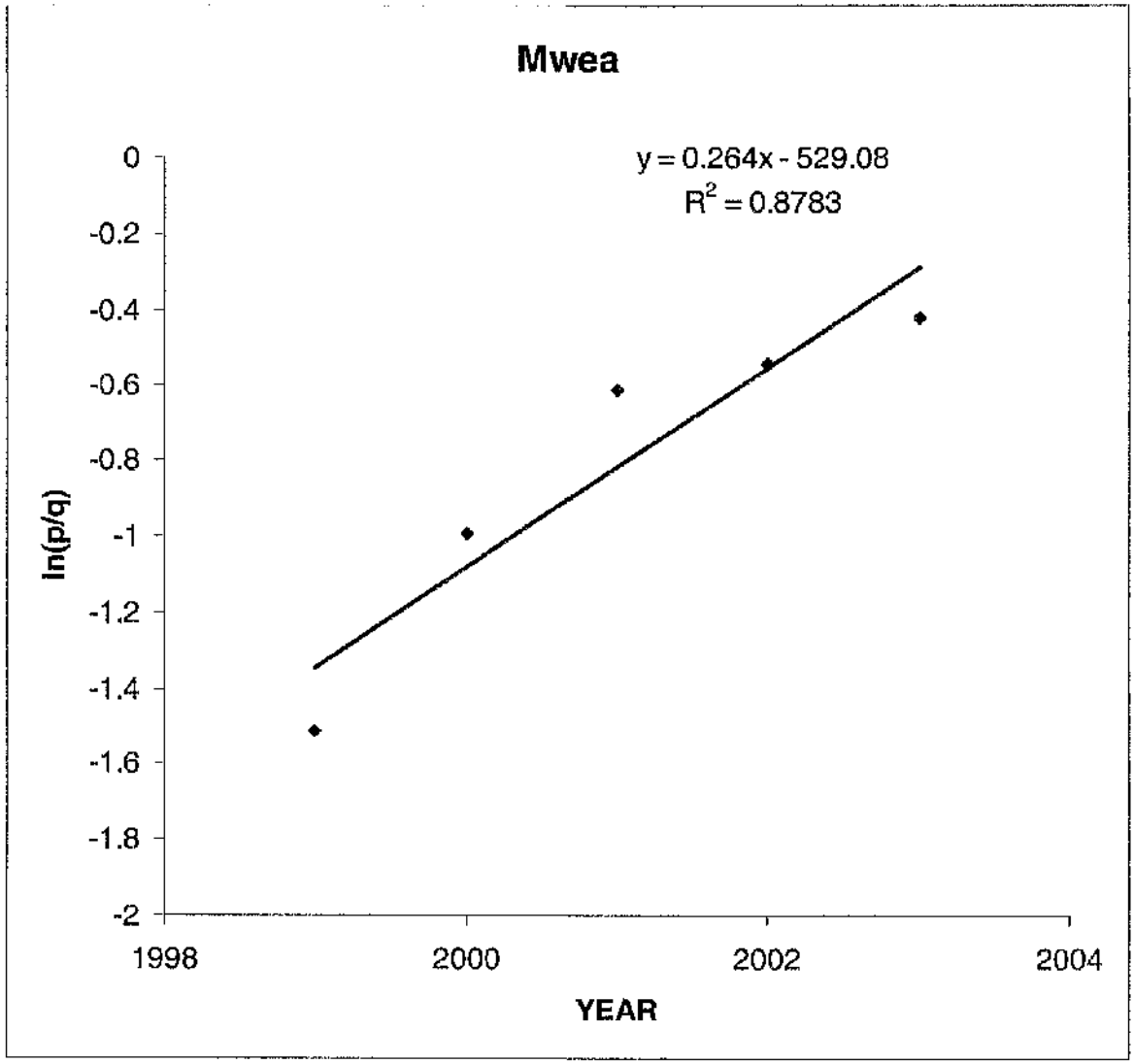

Figure 30: Estimation of selection coefficient (*s*) driving SP resistance in Mwea.

***P* is the frequency of resistant alleles inferred from treatment failure rate data, and *q* (=1-*p*) is the inferred frequency of sensitive alleles. The *x*-axis shows the generation number (assuming six generations per year). The slope gives an estimate of *s* = 0.264.**

4.1 Recrudescence versus reinfection

In this study two different sets of genetic markers used to distinguish recrudescences from reinfections were compared: five putatively neutral microsatellites and three antigen coding loci that are thought to be under strong natural selection (Hughes 1992;Hughes & Hughes 1995;Conway 1997;Escalante *et al.* 1998). The antigen-coding loci were found to be as effective as microsatellites in differentiating recrudescence from reinfection, despite the fact that antigen-coding loci are likely to be under strong selection pressure while the microsatellite loci are presumably not (Table 19).

More alleles were observed at the microsatellite loci than at the antigen-coding loci in the populations studied (Figure 8 and Figure 9). There are three possible nonexclusive hypotheses for this; a) that the microsatellites are more polymorphic than the antigen loci; b) that selection within partially immune hosts removes certain parasites from infections on the basis of their antigens, whereas there is no selection based on microsatellite alleles (since they are non-coding); c) that microsatellite loci are unstable.

An assumption of the second hypothesis is that these parasite populations are recombining frequently. Therefore, there is a mixture of all microsatellite/antigen allele combinations, and when immunity removes a subset of parasites based on a particular antigen, all the microsatellites are still represented in combination with the non-selected antigen alleles.

Microsatellite loci could behave in an unstable manner for two reasons. Recombination may occur *in vitro* (during PCR); recombination, or chimera formation, has been shown to occur between related template sequences present in a single PCR amplification (Judo *et al.* 1998). This phenomenon has also been observed in *P.*

4.0..... DISCUSSION AND SUMMARY

falciparum especially for the antigen gene *Msp1* (Tanabe *et al.* 2002), although this was over a much longer distance with long stretches of homologous DNA. The PCR used in this study for the antigen-coding loci are much shorter and do not contain long stretches of homologous DNA thus there are reduced chances of recombination. Loci may also be unstable *in vivo* within an asexual infection or during meiosis in the mosquito. The number of repeats within a microsatellite locus may change during an infection, possibly during mitotic replication, giving rise to parasites genetically identical at all loci except that particular microsatellite, although this does not appear to be the case in this study, as there were no parasites observed with multiple bands at only one microsatellite locus.

Genetic markers should be chosen based on the research question being asked (Sunnucks 2000). Each genetic marker has a range over which it can usefully detect genetic variation (Thompson *et al.* 1998). If the marker is too polymorphic, either the required sample size to describe diversity will be too high, or the samples will be too different resulting in homoplasy. If the marker has little polymorphism, as we have seen with C2M3 (Figure 7 and Figure 9), it might misrepresent some differences between the samples, and more so in a field setting.

We still do not know much about evolution of microsatellites and even their neutrality and hence data from their use should be treated with caution especially when that data is meant to inform policy change. From the results of this work it is clear that microsatellites might be the better markers for population analysis but more work remains to be done to justify their application as markers for monitoring drug efficacy in malaria.

Another potential problem with microsatellites is the issue of null alleles, which was noted with some of the microsatellites used in this study, where they failed to

amplify DNA from some samples, including laboratory isolates, despite repeated attempts. This is often caused by polymorphism in the primer recognition sequences in non-coding DNA. Since non-coding DNA is usually under less constraint because it does not encode proteins, mutations that occur have a greater chance of remaining within the population. This is in contrast with antigen loci where primers are in coding regions within blocks of sequences with 90% conservation of nucleotides in well studied genes where many allele sequences are available (Miller *et al.*, 1993).

4.2 MOI and transmission intensities

Is there a relationship between the numbers of clones per infection (MOI) and the transmission rate of malaria? In the study presented in this thesis, the MOI values seem to increase with increased levels of EIR (Figure 7), although it is hard to be certain of this trend, as there are only three sites to compare.

One fact that emerges in this study is that the average MOI values with both sets of markers do not reflect the intensity of transmission. This is because the transmission intensities of, for instance Bondo are over 1000 infective bites per person annually compared to Mwea where the rate is about 100 but the difference in mean MOI levels in the two areas is just one (Figure 7). This finding is in broad agreement with previous work, which indicated that increasing entomological inoculation rates (EIRs) were associated with progressively smaller increases in the average number of clones per host (MOI) (Arnot 1998).

The transmission rate of malaria is measured in entomological inoculation rate (EIR), which is a product of the proportion of mosquitoes containing sporozoites (the sporozoite rate) and the human biting rate. The human biting rate is usually estimated from the landing rate on collectors (Burkot & Graves 1995). EIR is also a fairly crude estimate, and infection rates will vary in time, especially in areas with seasonal

rains/changes in mosquito density. However, our estimates of strain diversity may have underestimated the true probability of two infections having the same genotype due to clustering of parasite strains in time and space, as well as the frequency of infections containing multiple strains.

Another inherent limitation of genotyping is the assumption that parasite strains detected in pretreatment samples reflect the complete population of infecting parasites. It is possible that strains circulating at low parasite density, or sequestered at the time of sampling, are not detected in pretreatment samples but could emerge to cause treatment failure, leading to the misclassification of a recrudescence as a new infection. Even if this was the case, the average number of clones detected in the patients under study in the two higher transmission areas of Bondo and Tiwi do not approach the potential number of input clones implied by the high EIRs of the two areas. The same observation was made in a study of infection complexity in Thailand (Paul *et al.*, 1998), but can be explained by some nonexclusive hypotheses.

Measurements of clone multiplicities are affected by the sensitivity and kinetics of PCR. The rapid DNA extraction methods used for the large numbers of small blood samples generated by field studies are known to limit the sensitivity of the method (Bottius *et al.*, 1996; Roper *et al.*, 1996). Ordinary PCR methods might fail to detect parasite clones with fewer than 5 million circulating parasites (genomes) per individual infection (equivalent to 1 parasite/ µl of blood).

PCR-derived estimates of the number of clones per individual can also be biased when alleles originating from different genomes are present in a sample in unequal concentrations (Arnot 1998). If one sequence dominates in the reaction mixture (>90%), it will tend to obscure minority sequences, which will fail to be amplified due to the stochastic nature of sampling in the early reaction cycles (Ranford-Cartwright *et al.*,

2002). This problem will be compounded if the same primers are amplifying two or more different sequences of unequal length, since the PCR is known to amplify preferentially smaller, more rapidly synthesized sequences (Lukyanov *et al.*, 1997). Underestimation of the true MOI values through such biases has been noted as a possible cause of the tendency for PCR to detect fewer genetically complex infections during clinical episodes than during asymptomatic infections (Roper *et al.*, 1998). This observation could also be an artefact due to fast-growing 'dominant' populations of parasites overgrowing 'resident' clones during the clinical attack, and not necessarily an indication that clinical infections actually contain fewer clones than asymptomatic infections (Contamin *et al.*, 1996; Mercereau-Puijalon 1996).

Parasite sequestration is also a factor that contributes to the low MOI values observed. The parasite population in an individual has been shown to vary on a day-to-day basis in asymptomatically infected children in Tanzania (Farnert *et al.*, 1997), presumably due to synchronicity of parasite populations of a particular genotype.

The origin of the genetic diversity itself could also play a part in the observations recorded. It is generally considered that the presence of genetically distinct parasite clones in patient infections results from one of two routes. The first is where the patient is inoculated with mixed genotype sporozoites derived from mosquitoes containing at least one hybrid oocyst. In hybrid oocysts, sporogony mitotically amplifies the meiotic products of a transiently diploid zygote/ookinete derived from the fusion of genetically different gametes (Ranford-Cartwright *et al.*, 1993). The second is simply superinfection with novel clones from inoculations from different infectious mosquitoes. This assumption that all genetic diversity originates from either meiotic recombination or superinfection has been queried by other authors (Druilhe *et al.*, 1998), who suggest that mitotic recombination may also play a role. The results of the

study presented in this thesis, that there is wide diversity of parasites even within populations with low transmission, suggest that mitotic recombination, or other mechanisms leading to a change in length of repetitive regions of DNA, during the various multiplication cycles in the vertebrate host (Arnot & Gull 1998) might contribute to the observed parasite diversity.

On the other hand, the observations may have nothing to do with experimental biases and may be a true picture of the MOI. If so then the lack of correlation between MOI and EIR could be explained by:(i) a high proportion of sporozoite inoculations either fail to initiate, or to complete successfully, intrahepatocytic development and thus do not mature into erythrocytic infections; (ii) most sporozoite inoculations successfully invade hepatocytes and complete exo-erythrocytic schizogony but result in short-lived, and usually low level, erythrocytic infections (de Roode *et al.*, 2004).

The possibility of sporozoites not completing the intrahepatocytic schizogony would be of interest for vaccine development against exo-erythrocytic stages (Frevert *et al.*, 1998), because it suggests some form of pre-erythrocytic stage immunity as a major component of natural immunity capable of preventing a high proportion of the genotypes inoculated as sporozoites from developing into bloodstream infections (Arnot 1998).

However, if the average number of clones is controlled by the immune response to erythrocytic stage antigens (e.g. MSP1, MSP2 and GLURP, and especially PfEMP1), the establishment of new clones would require them to have novel variants of these protective antigens not recognised by the existing host immunity. If the parasite population 'mix' of such polymorphic antigens were finite, increasing the infection pressure would expose the host to more antigenic variants and ultimately make it more difficult for new parasite clones to establish (Arnot 1998). This translates into a scenario

where older individuals, particularly in high transmission areas, harbour lower average MOI than children. Interestingly such a scenario has been observed in a holoendemic area in Senegal (Ntoumi *et al.*, 1995), but not in a hypoendemic area in Sudan (Roper *et al.*, 1998).

Natural immunity to malaria remains poorly understood. Various models involving eventual acquisition of long-term immunity have been advanced. The most popular seems to be the premunition hypothesis. This proposes that protection from clinical disease operates largely via a blockade against superinfection maintained by established chronic infections composed of numerous clones. Alternative models involving the acquisition of immunity to diverse variable antigens (Gupta & Day 1994; Hviid 1998), immunity to relatively conserved parasite antigens (Druilhe & Perignon 1997), or mechanisms dependent on age and immunological maturity (Baird *et al.*, 1991; Ntoumi *et al.*, 1995) have also been proposed. This study seems to reinforce the findings of others in Africa that showed most individuals harbour more than one parasite clone. It also seems to reinforce the fact that in African malaria endemic situations, the average MOIs are quite high (<2 clones per individual per infection) and this reflects the rates of transmission.

4.3 Genetic Diversity

In this study, microsatellite markers have been used to study the population structure of *P. falciparum* from geographical areas with varying transmission dynamics in Kenya. The data obtained seem to agree with others (Anderson *et al.*, 2000) that have suggested very high diversity levels of *P. falciparum* in African populations (Figure 9). Lack of LD was also observed in physically unlinked markers suggesting high levels of recombination in these areas (Table 20). The study areas are as far apart as 800 km but there seems to be no population genetic subdivision between the sites (F_{ST} = 0.008

(excluding C2M3) Table 21), and no correlation with distance between the sites (Figure 10) suggesting that the Kenyan *P. falciparum* populations studied are in effective panmixia

With a maximum flying range of approximately 6 km, it is unlikely that the mosquito vector is responsible for this observed panmixia, and it can only be attributed to the human host through migration and inter-regional travel. A closer look at the areas under the study reveals a lot of human movement especially between Bondo and Tiwi. Tiwi is located on the South Coast of Kenya, an area that is quite popular with tourism. This means there is frequent movement of people from upcountry to this area in search of jobs and business opportunities. The immigrants normally leave their families upcountry thus necessitating frequent travel back home. This perhaps explains the parasite gene flow and the almost indistinguishable parasite populations among these three areas.

A striking observation from the microsatellite analyses is that while the trinucleotide loci showed high levels of variation and low levels of diversity between subpopulations (G_{ST} = 0.006) (Figure 9 and Table 21), the dinucleotide marker C2M3 revealed low levels of variation with higher levels of G_{ST} (G_{ST} = 0.141) (Table 21). Studies of *P. falciparum* suggest that dinucleotide repeat microsatellites evolve up to twice as fast as trinucleotide repeats (Anderson *et al.* 2000b), which would be expected to translate into more variation for the dinucleotides. However, the dinucleotide repeat locus C2M3 revealed lower levels of variation (heterozygosity) and diversity (as measured by G_{ST}) than the trinucleotide repeats (Figure 9 and Table 21). This contradictory observation could be explained by the fact that the locus C2M3 is not a perfect microsatellite. In the 3D7 clone, C2M3 has two repeat runs; $(TA)_{18}$ and $(TA)_{24}$ with an insert of about 90bp in between. This insert contains various combinations of

bases including mono- and trinucleotide repeats, but is not a pure repetitive sequence. Although no C2M3 alleles were sequenced, it is possible that the non-repetitive middle sequence has affected the mutation/slippage rate and hence the variation at this locus.

4.3.1 Genetic differentiation and drug resistance

The F_{ST} /G_{ST} values are indicators of population differentiation. The lower the values, the more closely related are the parasite populations under consideration. The results of the study presented in this thesis reveal that the three Kenyan parasite populations are almost indistinguishable with very low levels of differentiation (F_{ST} < 0.008). In fact, the Kenyan parasite populations act as a single large gene pool. This finding is very important when considering drug resistance.

It would be expected that with the low differentiation of the Kenyan malaria parasite populations, drug resistance would spread quite fast once it has evolved or been introduced, since there seems to be a high gene flow between the geographically separated populations. This is consistent with the proposal that in populations with low F_{ST}, newly arising resistance alleles will spread more rapidly than in populations with higher F_{ST}. However, other factors will also determine the spread of these alleles within a population and these include the relative fitness of the alleles in the drug environment, transmission intensity, and recombination.

4.4 Drug resistance

Mutations in the parasite genes *Pfdhfr* and *Pfdhps* genes are associated with SP treatment failure, although their relative contribution to SP clinical treatment failure is not completely understood (Nash *et al.*, 2005; Roper *et al.*, 2003; Roper *et al.*, 2004). In the three Kenyan sites studied, SP treatment failure rates varied between 30% and 40 %, and this was reflected in the absence of wild-type alleles of *Pfdhfr* in two of the sites studied, Tiwi and Bondo. Wild-type alleles of both genes were seen in the third site,

Mwea, where SP treatment failure rates are lower (between 18 and 25%). The variation in the rates of treatment failure was reflected in the differences between the prevalence and combinations of mutations in the two genes studied. Mwea had the lowest percentage of triple mutant *dhfr* alleles (33%) thus reflecting the low treatment failure rates (TFR) *in vivo* while Tiwi, with the higher TFR, had 70% of the analysed samples having the triple mutant *dhfr* allele (Table 22). Bondo with the highest TFR had 63% prevalence of the triple mutant *dhfr* (Table 22), which is slightly lower than Tiwi (which has a slightly lower TFR than Bondo). Again the same pattern was observed with *dhps* alleles. Mwea had a prevalence of the double mutant *dhps* at 43%, Bondo had 22% while Tiwi had the highest with 51%. This difference is probably because the analysis excluded the multiple clone infections and, as can be expected most of these were from Bondo thus biasing the sample.

All of the parasites from Bondo and Tiwi with single and double mutant *dhps* alleles had the triple mutant *dhfr* allele, suggesting that in these two areas, *dhps* mutants occur against a background of *dhfr* mutations. However, in the Mwea site 12 isolates with the *dhps* A437G single mutation were found to have a wild-type *dhfr* allele (Table 24). It is commonly thought that *dhps* mutations arise against a background of *dhfr* mutant alleles (Plowe *et al.*, 1996), implying that resistance to SP requires the presence of mutations in *dhfr*. Field studies in Africa have clearly shown association of the *dhfr* triple mutant with increased likelihood of treatment failure following SP treatment (Nzila *et al.*, 2000b; Omar *et al.*, 2001a) and the odds ratio (OR) is usually slightly higher when mutations at *dhps* are included. However it has not been possible to obtain OR for *dhps* mutations alone because they are usually found only in association with substituted *dhfr* alleles.

A few recent studies have reported parasites with *dhps* mutant alleles with no corresponding mutations in their *dhfr* genes. In a recent study in Uganda, which is very close to our Bondo site, Sendagire and colleagues observed very high levels of the *dhps* mutants at codons 437 and 540 (89% and 74.6% respectively), compared to the most prevalent (63.6%) *dhfr* mutation at codon 108 (Sendagire *et al.*, 2005). One parasite sample out of 113 described was identified with wildtype *Dhfr* coupled with a single mutant (437G) *dhps*, and two samples had wildtype *dhfr* with double mutant *dhps*. One hypothesis for this observation is that SP is not the only drug driving the selection of *dhps* mutant alleles (see section 4.9).

While the F_{ST} levels are quite low among the three sites (Table 21), signifying adequate gene flow among them, the prevalence of *dhfr* alleles and also the TFR do not support this. For instance, the observation of wild-type *dhfr* in Mwea but none in the other two sites and also the unequal distribution of the triple substituted *dhfr* and double substituted *dhps* alleles (Table 22 and Table 23). While the prevalence of alleles in Bondo and Tiwi show no significant differences (Fisher's Exact, P= 0.733 for *dhfr* and P= 0.0567 for *dhps*), there are highly significant differences in allele prevalence between Mwea and the other two sites (Fisher's exact, P<0.001 for *dhfr* and P<0.002 for *dhps*). In an apparently panmictic population it is expected that all the alleles would be equally distributed which is not the case in this data. A hypothesis for this is that panmixia was established long before the introduction of the selection pressure in all the three populations. However sometime later the population in Mwea might have drifted apart from the other two with no apparent loss of the diversity in the noncoding DNA regions leading to the apparent panmixia being observed. However, the introduction of SP in these now separate populations selects for drug resistant parasites at different rates leading to the differences being observed.

In Kenya, SP was introduced as a first-line treatment in 1997. This means the useful therapeutic life for SP in Kenya has only been approximately four years although it was still available in private healthcare before its introduction as the first-line therapy for uncomplicated malaria in Kenya. At the outset, the long elimination half-life of SP was considered to be advantageous, since a single-dose cleared parasites, and prevented re-infection for a period of approximately 50 days. This relatively long period allowed patients to recover from the sequelae of malaria infection, particularly anaemia (Winstanley *et al.*, 1992).

However, the long half-life of SP also appears to have had a detrimental effect: the rapid selection of resistance. SP is a long acting drug and will persist at sub-therapeutic concentrations for more than a month (Nzila *et al.*, 2000b; Winstanley *et al.*, 1992). Sub-therapeutic levels of drug may or may not clear fully sensitive parasites [depending on whether the level is higher than the minimum inhibitory concentration (MIC) (most therapeutic drug levels are higher than the MIC)], but not highly resistant ones, therefore exerting pressure for resistance. This selective pressure eventually eliminates the alleles conveying drug sensitivity in parasite populations, and leads to the fixation of drug resistant alleles. For SP in Kenya the selection appears to have taken about 4 years, as demonstrated by this study. In Asia and Southern Africa, the time taken for resistance alleles to spread to fixation was shown to be approximately 6 years (Nair *et al.*, 2003; White 1992), which is approximately 36 *Plasmodium* generations assuming an estimate of six generations per year (Joy *et al.*, 2003).

As drug resistance spreads, it leaves genomic signatures that can be used to trace its evolution and introgression through a population. These signatures appear in the form of hitchhiking of genomic sequences flanking the drug resistance-conferring loci.

For SP resistance, the important loci are thought to be the *Pfdhfr* and *Pfdhps* genes (Nair *et al.*, 2003).

4.5 Evidence of selective sweep of resistant alleles

Microsatellite loci spanning a region approximately 17kb (both up and downstream) around the *dhfr* and *dhps* loci were used to construct haplotypes of parasites present in patients in three Kenyan sites.

The haplotype analysis provides evidence suggestive of selection around both *dhfr* and *dhps* loci in two of the sites, Bondo and Tiwi (Figure 17, Figure 23, Figure 24, Figure 25, Figure 26 and Figure 27). For the Mwea site evidence for selection was only observed on *dhfr* but not in *dhps* (Figure 17, Figure 20 and Figure 23). Although parasites bearing the wild-type *dhfr* were absent in Bondo and Tiwi, both wild-type and mutant *dhfr* and *dhps* alleles were found amongst the samples collected from Mwea. Hence, it was possible to compare patterns of microsatellite variation flanking both classes of alleles (Figure 24, Figure 25 and Figure 26). These data provide strong evidence of selection for the mutant *dhfr* allele, especially the triple mutant, in the three areas studied (Figure 17 and Figure 24). The results also suggest a strong selective pressure on the double mutant *dhps* in Bondo and Tiwi.

Markers flanking the triple mutant *dhfr* and the double mutant *dhps* alleles show an almost complete lack of diversity in Bondo and Tiwi (Figure 26 and Figure 27). The linkage disequilibrium (LD), as measured by EHH, across this ~17kb haplotype is also very elevated with skewed allele frequency distributions relative to the markers flanking the wild-type and single mutant alleles (Figure 26 and Figure 27). While these parasite isolates show a complete lack of LD with the unlinked markers in chromosomes 2, 3, 5,

8 and 10, they show elevated LD with the markers flanking the two genes. This is suggestive of very strong selection of the *dhfr* and *dhps* loci.

Allele distributions also provide additional evidence for a recent selection of the resistant alleles. Microsatellites flanking the triple mutant *dhfr* showed three predominant allelic haplotypes varying by a single locus in the outermost region of the sequence and occasional rare variants, indicative of a recent bottleneck (Cornuet & Luikart 1996). This is in contrast with haplotypes flanking wild-type *dhfr* alleles, which showed a wider variation of allele frequency distribution (Figure 24). Looking at the number of alleles observed at each locus, microsatellite *He* of haplotypes carrying mutant *dhfr* alleles was significantly lower than expectation, assuming both infinite alleles (IAM) and stepwise mutation models (SMM) of microsatellite mutation (Wilcoxon test, P <0.0039 [IAM]; P< 0.0020 [SMM]). In comparison, chromosomes carrying wild-type *dhfr* alleles showed no significant deviation from equilibrium expectations (Wilcoxon test, P < 0.056 [IAM]; P< 0.74 [SMM]). These findings suggest that triple mutant *dhfr* alleles have arisen on relatively few occasions, as evidenced by little variation in flanking markers. Different haplotypes are predominant in the three populations studied, which might indicate independent derivation of these mutants, or could indicate evolution within the parasites in a particular area.

4.6 Origin of the mutant *dhfr* alleles

Analysis of the eight-locus microsatellite haplotype flanking triple mutant *dhfr* alleles provides strong evidence for a single or very limited origin of these alleles. Of the 86 isolates bearing the mutant *dhfr* alleles, 65 had either one of two eight-locus allelic haplotypes that differed at only one of the eight markers. Exclusion of the DHFRD4 outer locus (8.0 Kb upstream of *dhfr*) from the analysis showed that the 65/84 isolates had a single seven-locus haplotype. Of the remaining 21, 13 differed at only

two of the eight markers while the other 8 differed at a single locus but with different allelic combinations (Table 25). None of the isolates bearing the triple mutant *dhfr* allele differed at more than three of the eight markers. This variation was also only observed in the downstream region flanking the *dhfr* gene. Most of the 86 isolates bearing the triple *dhfr* mutant examined, showed very similar four-locus haplotype at the markers in the downstream region of the gene (see discussion on asymmetry on page 174 and Figure 24).

Comparisons of microsatellite haplotype frequency distributions (Figure 14, Figure 15, Figure 16 and Figure 17) reveals that neither the haplotypes associated with the double mutant *dhfr* or those of the triple mutant *dhfr* alleles could be found associated with either the wild-type or single mutant haplotypes. Although the single mutants theoretically preceded the double and definitely the triple mutants, the extant single mutant allele haplotypes do not seem to be ancestral, on the basis of flanking markers, to either of the more substituted haplotypes. This finding is in agreement with the results of previous studies that have suggested that African *dhfr* double and triple mutants are possibly imported from elsewhere (Roper *et al.*, 2003). It would also be in agreement with theoretical predictions that new resistance mutations are rare and even constrain the rate at which the parasite population can adapt to drug (Anderson & Roper 2005; Hastings 2004; Hastings & Watkins 2005). However, a number of other triple mutant haplotypes were found to be present in the three areas but at very low frequencies, which would suggest that they are quite rare in the population. These rarer triple mutant haplotypes could have arisen through background mutation and have been selected by drug pressure. Single mutant alleles (coding for 108N) had unrelated flanking sequences, suggesting that they did arise independently and on multiple genetic backgrounds, although it is possible that the selection on them is too weak for the

flanking sequences to be carried through. Although single mutations seem to have arisen independently on multiple occasions, the fitness benefits they confer are perhaps too weak to overcome the associated fitness costs, and they do not seem to be strongly selected for at the population level. If, as has been postulated, there are additional compensatory mutations (Wang *et al.*, 2004) elsewhere in the genome (Schrag *et al.*, 1997), required to limit the deleterious effects of mutations in *dhfr*, then we might expect that the S108N single mutant might not have developed these compensatory mutations to enable it overcome the fitness cost associated with the *dhfr* mutations even with the drug pressure working to its advantage. Using this argument, the single mutant alleles seem unlikely to spread throughout the population or to persist long enough to accumulate more mutations. In contrast, when a multiple mutant allele arises, through successive selection and recombination of drug tolerant alleles, it has high relative fitness and can spread rapidly throughout large geographic regions (Roper *et al.*, 2003).

Due to time limitations, it was not possible to type samples from other areas such as Southeast Asia and Southern Africa to determine the possible geographical origin of the multiple mutant alleles observed in this study (assuming that the alleles spread into Kenya from other areas, rather than arising there). However, it is possible that their origin is in Southeast Asia, since another study has shown that Tanzanian (which borders Kenya to the south and is very close to one of the sites of this study) multiple mutants have a shared origin in Asia (Roper *et al.*, 2004).

4.7 The selective sweep 'valley'

The valley of reduced variation around the triple mutant *dhfr* alleles appears to be asymmetrical (Figure 24). This is in contrast with the symmetrical valley of reduced variation around the double mutant *dhps* alleles in Tiwi and Bondo (Figure 27) and also with what has been suggested by simulation studies (Kim & Stephan 2002). The valley

of reduced variation around *dhfr* in other populations has been shown to extend to as much as 100 kb away from the gene (Nair *et al.*, 2003) but the slopes of the valley on both sides of the gene do not appear to be symmetrical. The region of reduced diversity extends much more downstream than upstream with the dip being narrower on the upstream side of the gene (Figure 24). Pearce and colleagues (Pearce *et al.*, 2005) have hypothesized three explanations for this asymmetry;

i) Differing rates of recombination on either side of *dhfr*,

ii) Stochastic noise,

iii) Existence of a resistance enhancing or resistance-compensating gene downstream of *dhfr*.

It has been shown that varying distribution of crossover events, differing substantially from their genome average can reflect recombination hotspots as well as other genome features e.g. proximity to centromeres and telomeres (Barnes *et al.*, 1995; Symington *et al.*, 1991; Szankasi *et al.*, 1988). Even certain nucleotide substitutions have been shown to create recombination hotspots (Ponticelli *et al.*, 1988). The downstream region of dhfr does not appear to be refractory to recombination (for example, due to proximity to a centromere), since the samples with the wild-type *dhfr* still show high diversity (Figure 14 and Figure 24).

Asymmetry may also be caused by the stochastic nature of recombination events during short phases of intense selection. Kim and Stephan used a model of genetic variation along a recombining chromosome to show that in a population where the time to fixation is short, e.g. when effective population (*Ne*) is small; the selective sweep could be asymmetrical around the selected site (Kim & Stephan 2002). The short period in which fixation is reached reduces the amount of time for recombination events to occur resulting in stochastic noise (Pearce *et al.*, 2005). This fits with the scenario

observed in Southeast Asia where fixation of resistance took only six years. In Africa where fixation of the triple mutant is yet to be achieved, it is not easy to explain. However, intense selection pressure may result in rapid epidemic expansions (Maynard Smith *et al.*, 1993) of resistant parasites through introgression in local populations and as such this would lead to asymmetry (Pearce *et al.*, 2005). This hypothesis however assumes that asymmetry is random and therefore has appeared on the same side of *dhfr* in East and Southern Africa as it has in SE Asia by chance alone (Nair *et al.*, 2003; Pearce *et al.*, 2005).

There is some experimental evidence to support the third explanation. Wang and colleagues analysed a genetic cross between parasite lines HB3 and Dd2; in this cross the two parents differ in sensitivity to sulfadoxine (Wang *et al.*, 2004). They identified a segregating modifier of sulfadoxine resistance, which they termed the 'folate effect'. A 48.6 kb region of chromosome 4 was found to be in complete linkage with the folate effect phenotype. It is tempting to speculate that a resistance-enhancing or fitness-compensating gene, such as the one proposed for folate salvage in the upstream region of the resistant *dhfr* allele, could explain the reduced variation observed in flanking microsatellite markers upstream *dhfr*. The existence of this 'folate gene' is yet to be proved (Pearce *et al.*, 2005). However, if such an adaptation gene exists, we would expect it to be very close to the *dhfr* gene to prevent their linkage from being wiped out by recombination.

Nash and colleagues have observed the same phenomenon of asymmetry in an analysis of selective sweep of *pfcrt* on SE Asian isolates (Nash *et al.*, 2005). Given these observations with two genes with different sources of selective pressure, the most probable cause of the asymmetry is heterogeneity in recombination rates on the two sides of *dhfr*.

4.8 Drug resistance and selection on *Pfdhps*

In Tiwi and Bondo, the results of the analysis of microsatellites flanking the *dhps* gene reveal a similar pattern to that of the markers flanking *dhfr*, with a limited number of haplotypes flanking the double mutant (*dhps* A437G/K540E). A very different picture emerges with the isolates from Mwea. While the isolates bearing the double mutant *dhps* allele (A437G, K540E) from Bondo and Tiwi reveal signatures of selection (Figure 23, Figure 25 and Figure 27), the isolates with double mutant alleles from Mwea do not exhibit these signatures of selection (Figure 20 and Figure 23). When isolates (n =132) from all sites are analysed together for their allelic haplotypes, a total of 82 haplotypes are generated. When the Mwea isolates are excluded only 31 haplotypes are generated for the isolates from Bondo and Tiwi (n = 75). *Dhps* double mutants from Mwea are therefore present on a much larger genetic background, with no clear predominance of any one haplotype (Figure 20 and Figure 23). The distribution of microsatellite haplotypes around the single single (A437G) and double mutant *dhps* alleles from Mwea reveals a wide distribution of these alleles across all haplotypes. There are a large number of different haplotypes for each allele of *dhps* (wild-type, single, double), with no evidence of a reduction in haplotype diversity associated with increasing substitution at the *dhps* locus.

4.9 Lack of selective sweep around *dhps* in Mwea

Our observations in Mwea with regard to the distribution of drug resistance alleles require some explanation (Figure 20 and Figure 23). The uniform distribution of haplotypes associated with the wild-type, single and double mutant *dhps* alleles are significantly different findings to those observed in the other two sites in this study, and those reported elsewhere in the literature.

The lack of a selective sweep on *dhps* in Mwea is also puzzling given the fact that the same isolates show a strong selective sweep around *dhfr* (Figure 17). One possible explanation for this observation is that SP is not the selective force driving the evolution of *dhps* mutation in *P. falciparum*, at least not in Mwea, and possibly not at all. Two other possible factors may also be responsible: antibiotic drug use and diet.

Various antibiotics are known to affect the folic acid synthesis and some such as cotrimoxazole have even shown potential as antimalarials (Omar *et al.*, 2001b; Saliba & Kirk 1998). These antibiotics, coupled with the widespread practice for self-medication, might have pushed the evolution of *dhps* in *P. falciparum* in Mwea, perhaps even before the introduction of SP as an antimalarial. Cotrimoxazole is commonly used to treat acute respiratory tract infections (ARIs) especially pneumonia and its use is widespread in East Africa because of its lower cost (World Health Organization 1991b).

It has been shown that even in the presence of a biosynthetic synthesis for PABA, malaria parasites require exogenous dietary PABA for survival (Greene 1999; Kicska *et al.*, 2003; van Doorne *et al.*, 1998). As has been discussed previously, PABA and folate are competitive inhibitors of sulfadoxine (Peters 1997), so it is possible that in patients with high dietary folate/PABA, sulfadoxine will have no inhibitory effect on the parasite. Investigations of the dietary patterns of the inhabitants of Mwea reveal that most of their diet is composed of rice and/or maize taken with vegetables, as the farming in this area is rice irrigation. Leafy greens such as turnip greens (diet folate equivalent = 65 μg), orange juice (70 μg), peas (105 μg), beans (90 μg), and many other types of fruits and vegetables are rich sources of folate. The diet in Mwea is very low in animal proteins (Bwibo & Neumann 2003), suggesting that the source of protein is likely to be legumes, which are high in folates. This is in contrast to the diet of the populations of Tiwi and Bondo, whose diet is mainly fish and maize meal as they live

very close to fishing areas. The vegetable intake in these other areas is also poor as the soil is not suitable for horticulture.

It is possible therefore that parasites infecting the Mwea population encounter higher levels of physiological folate/PABA than those of Tiwi and Bondo, and that this removes the effectiveness of sulfadoxine in the combination drug SP. Mutations in *dhps* would therefore confer no selective advantage to parasites, and the gene would therefore not undergo a selective sweep. Further studies comparing PABA and folate levels between the three populations would be required to investigate this hypothesis further.

4.10 Transmission intensity and drug resistance

The areas in which this study was carried out have relative higher transmission indices than the other areas of Southeast Asia and Southern Africa in which similar studies have been carried out. What is the link between transmission and antimalarial drug resistance?

There are currently two schools of thought concerning the link between transmission intensity and evolution of drug resistance both of which are discussed below.

4.10.1 Low transmission increases drug resistance

In the mid-1990s it was proposed that parasite inbreeding could elevate multigenic drug resistance by preventing the breakdown and dilution of drug-resistant haplotypes that occur through genetic recombination (Schmidt 1995). Inbreeding is thought to be more frequent when transmission is low, as this is predicted to lessen the likelihood of mixed infections comprising organisms carrying resistant and sensitive alleles (Curtis & Otoo 1986). The proponents of this hypothesis also argue that most infections in low-transmission areas are in nonimmune people who become symptomatic and are then treated with the drug. The increased proportion of infections

treated results in the selective advantage of the mutation being larger in areas of low transmission than that in areas of high transmission, so resistance mutations are more likely to survive and spread in such areas (Hastings 2004). Perusal of the survival probabilities suggest that, even if this was not the case, most origins of drug resistance might occur in areas of low transmission because the differences in survival probabilities caused by their differing selective advantages will probably greatly outweigh any differences in their rate of input. For example, Nair and colleagues estimated the selective advantage of *dhfr* resistance mutations to be 10% in Southeast Asia, giving them an 18% chance of survival assuming a Poisson distribution, falling to 2% under a more realistic assumption of highly heterogeneous malaria transmission (Nair *et al.*, 2003). Similarly, Roper and colleagues calculated the selective advantage of the *dhfr* triple mutation in Southern Africa to be 5%, giving it a chance of survival of between 9% and 10%, depending on degree of transmission heterogeneity (Roper *et al.*, 2003). Both studies (Roper *et al.*, 2003) were carried out in areas of low transmission where most infections are symptomatic, hence inducing the patient to seek treatment (Nair *et al.*, 2003).

By contrast, many infections in areas of higher transmission in Africa are asymptomatic and more likely to remain untreated (Ntoumi *et al.*, 1995). This could substantially reduce selective advantage and hence the probability of a new mutation surviving (Hastings 2004). The ability of mutations to survive can therefore vary substantially depending on local epidemiology, which determines the selective advantage of the mutation and the degree of heterogeneity in the transmission of the mutation.

4.10.2 Low transmission decreases drug resistance

It has been proposed by theoretical argument (Hastings & Mackinnon 1998) and in a review paper (Molyneux *et al.*, 1999) that low transmission can decrease multigenic drug resistance. More direct field evidence has recently emerged in Zimbabwe, during a comprehensive vector control programme. In Zimbabwe, which lies on the southern fringes of malaria endemicity in Africa, transmission is markedly seasonal and epidemic, and asymptomatic carriage is rare (Taylor & Mutambu 1986). Annual house spraying was selectively instituted (during 1995–1999) in one of two mesoendemic sites that both used chloroquine as first line treatment and had similar starting levels of chloroquine resistance. In the sprayed region during 4 years when annual spraying was used, the chloroquine failure rate decreased such that the odds of drug failure became four-fold lower than before spraying was instituted (Mharakurwa 2004). The odds of drug failure did not significantly change in the unsprayed site. After the selective spraying was stopped in 1999, drug failure odds for the sprayed area had increased four-fold by 2003, back to the original level (Mharakurwa 2004). During the 1998–1999 transmission seasons, a cross-sectional assessment of parasite genotypes in the *Pfmdr1* and *Pfcrt* showed higher likelihood for mixed infections, containing both mutated and wild-type variants at one or more loci, in the sprayed area. This was in spite of lower transmission in the sprayed area (Mharakurwa 2004). Thus reduced transmission, due to vector control, or environmental factors such as high altitude, appears to delay the escalation of drug resistance.

4.10.3 Transmission and rate of selection for resistance in Kenya

The three areas under study in this thesis have varying transmission intensities. Bondo is holoendemic and transmission is perennial, Tiwi is hyperendemic and transmission is perennial with seasonal variations, while Mwea is hypoendemic and the

selective advantage of 0.26 (meaning theoretically resistant parasites can replace the entire population in only 4 years). The resistant parasite population in Bondo has only a 0.10 selective advantage (theoretically it would take 10 years to replace the entire parasite population in Bondo) (Figure 28 and Figure 29). Selection in Mwea would appear to be the strongest, followed by that in Tiwi, while it would be slowest in Bondo (Figure 28, Figure 29 and Figure 30).

Taken together, these findings suggest that the high transmission rates in Bondo, resulting into high recombination rates, break down the triple mutant haplotype thus taking a long time for it to introgress through the population; this explanation is supported by the higher variation seen in the loci flanking regions of triple mutant *dhfr* alleles.

An upshot of this finding is that during the initial stages of evolution and spread of drug resistance, the human population at highest risk would be that living in low transmission areas. This is because even with their higher clinical cases with sensitive malaria parasites (due to their low immunity) they have the added risk of getting drug resistant parasites which would lead to epidemics and eventually higher morbidities and mortalities. This emphasizes the need for surveillance programs especially after the introduction of new antimalaria treatment regimens.

transmission is seasonal. From the arguments in section 4.10.1, drug resistance would be expected to evolve and spread fastest in Mwea while it would be slowest in Bondo. The mutant forms of *dhfr* associated with SP treatment failure were higher in Tiwi and Bondo compared to Mwea. For example, the triple mutant allele was found in 63% and 70% of samples in Tiwi and Bondo respectively, but in only 33% of samples from Mwea (Table 22), and Mwea was the only site where wild-type *dhfr* alleles were detected. For *dhps*, wildtype alleles were seen in 13% and 27% of samples from Tiwi and Bondo respectively, and 39% of those from Mwea. Currently SP treatment failure rates in the three sites have been estimated to be between 20-60%. Data from previous studies reveal that SP treatment failure rate in Kilifi (an area including the Tiwi site) was 20% in 1998. Most of these persistent parasites carried the triple-mutant allele of *dhfr* with or without mutations in *dhps* (Nzila *et al.*, 2000). These failure rates were shown to have increased to about 30% in 1998 (Nzila *et al.*, 1998). This trend has now been confirmed by the results of this study, which reveal that in Tiwi SP failure rates were as high as 40% in 2000 when this study was carried out. This trend is expected to be the same in the other two sites with Mwea having lower rates and Bondo with higher rates consistent with the transmission intensities.

Examination of the eight-locus haplotypes reveals that most of the variation in the triple mutant *dhfr* alleles was found in the isolates from Bondo (Figure 17 and Table 25), with only 15 out of 36 isolates bearing the triple *dhfr* allele having the same identical haplotype. This contrasts with Mwea, where 19 out of the twenty isolates bearing the triple *dhfr* mutant had a single eight-locus haplotype. In Tiwi, 25 out of the 30 isolates bearing the triple *dhfr* allele had the same eight-locus haplotype.

Finally the estimated selection coefficients for SP resistant mutants (section 3.5.9) suggest that the resistant parasite population in Mwea (lowest transmission) has a

4.11 SUMMARY

The work presented in this thesis has demonstrated that microsatellite markers perform no better than antigenic markers in distinguishing genuine recrudescence of resistant parasites from reinfections during the follow-up period of *in vivo* drug efficacy trials.

Data have been presented showing that populations of parasites from three areas of Kenya, separated by more than 800 km, are genetically closely related such that there is apparent panmixia. This has implications for the spread of drug resistance within the country; resistant parasites arising spontaneously, or moving into an area from outside of the country, have a high likelihood of spreading rapidly to all areas of Kenya.

The work presented also demonstrates that alleles of the *dhfr* gene, especially the triple mutant allele, isolated from the three different areas are closely related to one another and probably share a common and very recent ancestor. There is strong evidence for a selective sweep around *dhfr* in the Kenya *P. falciparum* population. The same situation has occurred for *dhps* alleles (both the single and more especially the double mutant) in two of the sites studied. However at a third study site there was no evidence for selection at the *dhps* locus. Markers flanking the gene showed high and similar diversity around wild-type and mutant alleles with multiple haplotypes present. This finding suggests that in this region, sulfadoxine is not driving selection of mutant *dhps* alleles, and, that factors such as diet or sulfa-antibiotic use may be affecting the efficacy of the sulfadoxine component of SP.

These findings have implications for the spread of antimalarial drug resistance in Kenya. SP resistant parasites appear to arise by immigration from other areas rather than by spontaneous emergence, as has been reported for other areas of Africa with lower

intensities of transmission. The strength of selection appears to be lower in areas of high transmission, suggesting a more rapid increase in treatment failure in areas of lower transmission. Since individuals living in regions of low transmission develop partial immunity more slowly, they are even more at risk because drug resistant levels will rise more quickly, and treatment failure rates will be higher due to higher parasite resistance as well as lower immunity. Malaria control programmes therefore should take great care to assess treatment failure rates/resistance in low transmission areas, since there is the potential for epidemics with high morbidity and mortality.

APPENDIX 1

Allele sizes (in base pairs) for all individuals studied across the five unlinked loci (0= nonamplification)

Sample ID	*C2M3*	*C3M85*	*TA40*	*TA17*	*TA48*
BONDO.1	265	292	222	276	288
BONDO.2	300	336	308	256	279
BONDO.3	306	289	233	254	275
BONDO.4	325	350	233	289	288
BONDO.5	319	296	239	327	275
BONDO.6	313	336	265	313	288
BONDO.7	299	353	289	302	296
BONDO.8	325	309	300	274	283
BONDO.9	313	335	323	252	283
BONDO.10	275	329	221	244	292
BONDO.11	280	332	308	294	318
BONDO.12	260	306	229	323	335
BONDO.13	295	269	221	228	271
BONDO.14	290	215	207	300	335
BONDO.15	275	290	243	250	276
BONDO.16	252	254	207	211	252
BONDO.17	274	233	207	167	281
BONDO.18	260	296	196	331	295
BONDO.19	279	243	226	256	262
BONDO.20	274	354	189	256	267
BONDO.21	307	289	237	233	271
BONDO.22	300	299	222	272	306
BONDO.23	320	234	235	228	295
BONDO.24	264	263	222	324	285
BONDO.25	305	293	198	338	280
BONDO.26	332	343	210	260	290
BONDO.27	342	322	190	265	312
BONDO.28	300	215	226	331	289
BONDO.29	268	278	190	245	285
BONDO.30	276	332	214	300	225
BONDO.31	305	363	204	313	265
BONDO.32	256	222	215	183	322
BONDO.33	0	273	0	200	0
BONDO.34	288	298	242	228	301
BONDO.35	311	269	188	217	272
BONDO.36	260	231	227	244	268
BONDO.37	279	289	212	233	312
BONDO.38	232	299	220	336	332

Sample ID	C2M3	C3M85	TA40	TA17	TA48
BONDO.39	248	342	206	244	265
BONDO.40	232	292	216	228	232
BONDO.41	228	285	208	244	280
BONDO.42	248	275	186	213	285
BONDO.43	300	290	185	200	268
BONDO.44	265	345	196	198	323
BONDO.45	248	218	195	207	288
BONDO.46	268	322	200	213	300
BONDO.47	232	318	204	183	265
BONDO.48	311	335	215	192	285
BONDO.49	284	222	250	184	289
BONDO.50	256	299	235	192	295
BONDO.51	305	318	213	210	322
BONDO.52	260	314	225	220	272
BONDO.53	264	298	190	194	232
BONDO.54	244	325	185	207	290
BONDO.55	264	275	210	213	300
BONDO.56	279	0	225	213	322
BONDO.57	288	311	234	227	265
BONDO.58	248	243	213	227	290
BONDO.59	232	265	189	200	306
BONDO.60	272	292	236	220	344
TIWI.1	321	335	206	217	293
TIWI.2	311	330	213	230	303
TIWI.3	286	297	188	215	340
TIWI.4	257	289	213	270	330
TIWI.5	282	312	229	259	310
TIWI.6	276	262	213	233	280
TIWI.7	300	330	210	226	293
TIWI.8	300	267	185	230	247
TIWI.9	286	287	178	225	310
TIWI.10	306	344	226	189	290
TIWI.11	313	290	200	200	323
TIWI.12	300	285	163	248	320
TIWI.13	258	293	173	256	0
TIWI.14	283	315	190	220	308
TIWI.15	296	285	200	252	331
TIWI.16	296	320	213	244	280
TIWI.17	265	332	237	252	287
TIWI.18	293	266	216	199	310
TIWI.19	252	267	220	197	340

Sample ID	*C2M3*	*C3M85*	*TA40*	*TA17*	*TA48*
TIWI.20	312	307	200	197	320
TIWI.21	296	258	211	212	287
TIWI.22	254	289	227	204	260
TIWI.23	311	319	188	190	300
TIWI.24	306	296	230	193	267
TIWI.25	252	285	230	190	273
TIWI.26	248	357	170	211	304
TIWI.27	291	349	181	233	260
TIWI.28	252	321	236	207	280
TIWI.29	291	345	218	215	297
TIWI.30	248	322	178	189	280
TIWI.31	287	298	202	195	287
TIWI.32	289	324	191	230	293
TIWI.33	311	294	221	254	297
TIWI.34	305	255	214	190	280
TIWI.35	300	300	178	220	290
TIWI.36	286	268	212	187	279
TIWI.37	294	306	233	199	267
TIWI.38	312	292	190	195	293
TIWI.39	333	288	239	239	280
TIWI.40	321	338	228	258	254
TIWI.41	252	273	182	265	287
TIWI.42	265	332	258	235	260
TIWI.43	257	271	189	217	343
TIWI.44	265	252	240	256	290
TIWI.45	321	328	223	212	267
TIWI.46	326	277	203	248	300
TIWI.47	289	299	217	239	289
TIWI.48	312	305	234	223	290
TIWI.49	325	289	211	202	254
TIWI.50	291	269	219	234	310
TIWI.51	316	327	225	226	293
TIWI.52	353	270	231	230	340
TIWI.53	296	290	226	218	293
TIWI.54	274	313	217	248	289
TIWI.55	280	305	213	225	267
TIWI.56	300	311	209	220	290
TIWI.57	235	288	224	199	254
TIWI.58	288	268	203	203	310
TIWI.59	252	273	210	200	343
TIWI.60	296	275	192	241	287

Sample ID	*C2M3*	*C3M85*	*TA40*	*TA17*	*TA48*
MWEA.1	262	267	207	254	308
MWEA.2	296	311	186	195	293
MWEA.3	265	300	207	224	342
MWEA.4	305	267	194	198	325
MWEA.5	237	255	241	238	286
MWEA.6	296	324	230	215	293
MWEA.7	254	311	214	200	250
MWEA.8	260	290	207	272	286
MWEA.9	329	300	197	261	248
MWEA.10	273	289	192	229	267
MWEA.11	265	353	175	214	273
MWEA.12	262	324	200	187	267
MWEA.13	327	295	216	187	273
MWEA.14	305	339	212	224	267
MWEA.15	257	281	212	200	273
MWEA.16	318	300	208	203	270
MWEA.18	323	329	224	207	325
MWEA.19	282	318	190	188	286
MWEA.20	309	311	207	220	293
MWEA.21	308	282	216	210	252
MWEA.22	298	329	212	248	270
MWEA.23	330	277	207	271	320
MWEA.24	282	319	192	248	273
MWEA.25	254	284	197	187	286
MWEA.26	309	342	197	192	270
MWEA.27	264	294	223	187	262
MWEA.28	296	300	229	207	243
MWEA.29	309	364	163	230	293
MWEA.30	296	329	207	248	271
MWEA.31	265	290	212	212	300
MWEA.32	231	306	176	232	276
MWEA.33	309	355	187	188	286
MWEA.34	235	322	191	283	250
MWEA.35	303	266	172	219	284
MWEA.36	318	277	188	192	270
MWEA.37	327	256	190	212	349
MWEA.38	283	286	218	226	281
MWEA.39	234	281	207	191	340
MWEA.40	303	302	212	245	323
MWEA.41	300	286	228	212	278
MWEA.42	299	288	192	218	247

Sample ID	*C2M3*	*C3M85*	*TA40*	*TA17*	*TA48*
MWEA.43	266	265	176	248	287
MWEA.44	296	280	206	188	293
MWEA.45	279	311	189	215	296
MWEA.46	264	308	213	272	310
MWEA.47	285	306	209	225	350
MWEA.48	316	324	220	244	320
MWEA.49	320	356	200	230	254
MWEA.50	266	334	213	245	296
MWEA.51	274	289	254	262	286
MWEA.52	310	306	202	204	309
MWEA.53	282	323	229	256	250
MWEA.54	262	318	252	238	267
MWEA.55	242	342	176	197	276
MWEA.56	265	322	210	207	287
MWEA.57	242	271	198	238	343
MWEA.58	237	300	229	188	330
MWEA.59	260	290	193	220	310
MWEA.60	258	346	234	197	271
MWEA.61	332	295	193	215	343
MWEA.62	324	283	218	232	296
MWEA.63	262	319	213	257	270
MWEA.64	273	281	239	200	262
MWEA.65	318	342	183	212	276
MWEA.66	310	312	232	228	282
MWEA.67	265	290	175	254	308
MWEA.68	309	282	194	232	250

APPENDIX 2

Dhfr-linked microsatellite haplotypes arranged by sorting isolates according to their DhfrU1 allele sizes (in base pairs). Values in parentheses indicate whether the double mutant has the secondary mutation at either the 51 or 59 codons.

Sample ID	DhfrU1	DhfrU2	DhfrU3	DhfrU4	DhfrD1	DhfrD2	DhfrD3	DhfrD4	Mutations	Haplotyp
BONDO.54	107	158	174	200	120	235	180	263	1	1
TIWI.20	107	167	178	202	116	279	200	300	1	2
MWEA.61	110	181	210	219	92	276	218	294	0	3
BONDO.49	111	162	178	235	122	240	185	266	1	4
BONDO.17	111	174	188	212	100	244	183	283	3	5
MWEA.2	111	178	170	212	100	247	178	283	1	6
TIWI.38	111	180	191	238	108	251	200	298	1	7
BONDO.20	111	189	170	224	116	238	190	296	1	8
BONDO.46	115	158	210	187	118	260	176	300	1	9
TIWI.1	115	169	183	206	97	269	205	288	2 (51)	10
TIWI.17	115	169	183	206	97	269	205	288	2 (51)	10
TIWI.60	115	169	183	206	97	269	205	288	2 (51)	10
BONDO.38	115	169	183	224	97	252	205	279	2 (59)	11
BONDO.53	115	169	183	224	97	269	192	275	2 (51)	12
MWEA.29	115	169	183	224	97	269	192	294	2 (51)	12
MWEA.34	115	169	183	224	97	269	192	294	2 (51)	12
BONDO.44	115	169	183	229	97	252	192	279	2 (51)	13
BONDO.47	115	169	183	229	97	252	192	279	2 (51)	13
MWEA.22	115	169	183	229	97	252	192	279	2 (51)	13
MWEA.11	115	169	183	229	97	252	192	279	2 (51)	13
BONDO.45	115	169	183	229	97	252	192	279	2 (51)	13
MWEA.23	115	169	183	229	97	252	192	279	2 (51)	13
MWEA.26	115	169	183	229	97	252	192	294	2 (51)	14
TIWI.45	117	176	182	200	124	260	196	285	1	15
MWEA.47	117	176	182	200	124	260	196	285	0	15
TIWI.55	118	177	175	210	127	267	200	286	1	16
MWEA.24	118	192	167	171	109	266	211	289	1	17
MWEA.66	119	162	170	216	122	255	229	289	1	18
MWEA.33	119	190	199	212	122	245	178	283	0	19
MWEA.48	121	160	168	219	119	234	209	283	0	20
BONDO.40	122	192	183	187	115	256	179	290	1	21
MWEA.15	122	200	175	212	152	248	215	289	0	22
MWEA.35	125	166	211	198	119	252	202	278	1	23
BONDO.35	126	160	174	218	90	231	176	275	1	24
BONDO.41	126	167	183	224	120	248	213	295	1	25
TIWI.43	126	169	183	206	97	252	192	288	2 (59)	26
TIWI.9	126	169	183	206	97	252	192	288	2 (59)	26
TIWI.8	126	169	183	206	97	252	192	288	2 (59)	26
MWEA.62	126	169	183	229	97	252	192	288	2 (59)	26
TIWI.49	126	184	183	229	97	252	218	296	2 (59)	27
TIWI.35	126	184	183	229	97	252	218	296	2 (59)	27
BONDO.50	126	184	183	229	97	252	218	296	2 (59)	27
MWEA.6	126	184	183	229	97	252	218	296	2 (59)	27
MWEA.28	126	184	183	229	97	252	218	296	2 (59)	27
BONDO.58	126	184	188	224	100	244	205	283	3	28
MWEA.1	126	184	195	224	97	252	205	279	2 (59)	29
BONDO.48	126	184	195	224	97	252	205	279	2 (59)	29
BONDO.11	126	184	195	224	97	252	205	279	2 (59)	29

Sample ID	DhfrU1	DhfrU2	DhfrU3	DhfrU4	DhfrD1	DhfrD2	DhfrD3	DhfrD4	Mutations	Haplotyj
BONDO.51	126	184	195	224	97	252	205	279	2 (59)	29
MWEA.14	126	184	195	224	97	252	205	279	2 (59)	29
MWEA.49	126	184	195	224	97	252	205	279	2 (59)	29
MWEA.50	126	184	195	224	97	252	205	279	2 (59)	29
MWEA.51	126	184	195	224	97	252	205	279	2 (59)	29
MWEA.55	126	184	195	224	97	252	205	279	2 (59)	29
MWEA.57	126	184	195	224	97	252	205	279	2 (59)	29
MWEA.68	126	184	195	229	97	252	205	279	2 (59)	29
MWEA.40	127	181	170	180	96	272	216	294	0	30
MWEA.39	129	179	178	178	122	244	180	289	1	31
TIWI.24	133	174	183	212	100	244	183	291	3	35
BONDO.1	133	174	188	212	100	244	183	283	3	33
BONDO.2	133	174	188	212	100	244	183	283	3	33
BONDO.3	133	174	188	212	100	244	183	291	3	35
BONDO.7	133	174	188	212	100	244	183	291	3	35
BONDO.8	133	174	188	212	100	244	197	291	3	36
BONDO.10	133	174	188	212	100	244	197	291	3	36
BONDO.13	133	174	188	212	100	244	183	283	3	33
BONDO.14	133	174	188	212	100	244	183	283	3	33
BONDO.15	133	174	188	212	100	244	183	283	3	33
BONDO.18	133	174	188	212	100	244	183	291	3	35
BONDO.22	133	174	188	212	100	244	197	291	3	36
BONDO.24	133	174	188	212	100	244	183	283	3	33
BONDO.26	133	174	188	212	100	244	183	283	3	33
BONDO.33	133	174	188	212	100	244	183	283	3	33
BONDO.34	133	174	188	212	100	244	183	283	3	33
TIWI.3	133	174	188	212	100	244	183	291	3	35
TIWI.4	133	174	188	212	100	244	183	291	3	35
TIWI.7	133	174	188	212	100	244	183	291	3	35
TIWI.10	133	174	188	212	100	244	183	291	3	35
TIWI.15	133	174	188	212	100	244	183	291	3	35
TIWI.18	133	174	188	212	100	244	183	291	3	35
TIWI.19	133	174	188	212	100	244	183	291	3	35
TIWI.21	133	174	188	212	100	244	183	291	3	35
TIWI.22	133	174	188	212	100	244	183	291	3	35
TIWI.23	133	174	188	212	100	244	183	291	3	35
TIWI.25	133	174	188	212	100	244	183	291	3	35
TIWI.28	133	174	188	212	100	244	183	291	3	35
TIWI.30	133	174	188	212	100	244	183	291	3	35
TIWI.31	133	174	188	212	100	244	183	291	3	35
TIWI.34	133	174	188	212	100	244	183	291	3	35
TIWI.36	133	174	188	212	100	244	200	291	3	37
TIWI.41	133	174	188	212	100	244	183	291	3	35
TIWI.46	133	174	188	212	100	244	183	291	3	35
TIWI.47	133	174	188	212	100	244	183	291	3	35
TIWI.48	133	174	188	212	100	244	183	291	3	35
TIWI.52	133	174	188	212	100	244	183	291	3	35
TIWI.53	133	174	188	212	100	244	183	291	3	35
TIWI.58	133	174	188	212	100	244	183	291	3	35
TIWI.59	133	174	188	212	100	244	183	291	3	35
MWEA.8	133	174	188	212	100	244	183	278	3	33
MWEA.10	133	174	188	212	100	244	183	289	3	33
MWEA.12	133	174	188	212	100	244	183	283	3	33

Sample ID	DhfrU1	DhfrU2	DhfrU3	DhfrU4	DhfrD1	DhfrD2	DhfrD3	DhfrD4	Mutations	Haplotyj
MWEA.27	133	174	188	212	100	244	183	283	3	33
MWEA.31	133	174	188	212	100	244	183	283	3	33
MWEA.37	133	174	188	212	100	244	183	283	3	33
MWEA.41	133	174	188	212	100	244	183	283	3	33
MWEA.42	133	174	188	212	100	244	183	283	3	33
MWEA.43	133	174	188	212	100	244	183	283	3	33
MWEA.44	133	174	188	212	100	244	183	283	3	33
MWEA.45	133	174	188	212	100	244	183	283	3	33
MWEA.54	133	174	188	212	100	244	183	283	3	33
MWEA.59	133	174	188	212	100	244	183	283	3	33
MWEA.60	133	174	188	212	100	244	183	283	3	33
MWEA.63	133	174	188	212	100	244	183	283	3	33
MWEA.64	133	174	188	212	100	244	183	283	3	33
MWEA.65	133	174	188	212	100	244	183	283	3	33
MWEA.67	133	174	188	212	100	244	183	283	3	33
TIWI.33	133	174	188	212	100	247	183	291	3	38
BONDO.4	133	174	188	212	100	269	205	283	3	40
BONDO.5	133	174	188	212	100	269	183	291	3	39
BONDO.12	133	174	188	212	100	269	183	291	3	39
BONDO.16	133	174	188	212	100	269	197	291	3	41
BONDO.21	133	174	188	212	100	269	183	291	3	39
BONDO.43	133	174	188	212	100	269	183	291	3	39
BONDO.39	133	174	188	212	106	244	183	283	3	34
BONDO.37	133	174	188	212	106	269	183	283	3	41
MWEA.5	133	174	188	212	113	244	183	283	3	42
TIWI.57	133	174	188	212	115	244	183	291	3	43
TIWI.27	133	174	188	212	121	244	183	291	3	44
TIWI.54	133	174	188	215	100	244	183	291	3	45
BONDO.25	133	174	188	212	100	244	183	291	3	35
BONDO.27	133	174	188	212	100	244	197	291	3	36
BONDO.28	133	174	188	212	100	244	183	283	3	33
BONDO.56	133	174	188	212	100	244	183	291	3	35
TIWI.56	133	174	188	212	100	244	183	291	3	35
BONDO.29	133	174	188	212	100	244	183	283	3	33
BONDO.30	133	174	188	212	100	244	183	283	3	33
BONDO.31	133	174	188	212	100	244	183	283	3	33
BONDO.32	133	174	188	212	100	244	183	283	3	33
BONDO.57	133	174	188	212	100	244	183	283	3	33
BONDO.60	133	174	188	212	100	244	183	283	3	33
BONDO.59	133	174	188	212	106	244	183	283	3	34
MWEA.25	133	174	188	212	100	244	183	283	3	33
BONDO.52	136	177	171	220	116	250	178	271	1	46
MWEA.36	137	158	198	187	126	245	218	283	1	47
MWEA.38	138	166	162	210	118	266	186	283	0	48
MWEA.56	139	189	176	237	117	248	189	294	0	49
BONDO.19	141	169	183	224	111	269	192	275	2 (59)	50
BONDO.9	141	169	183	224	111	269	192	275	2 (59)	50
BONDO.36	141	164	183	224	111	269	192	275	2 (59)	50
MWEA.46	141	188	182	175	88	230	223	289	1	51
MWEA.17	146	183	201	178	134	230	229	283	1	52
MWEA.53	147	166	174	243	109	249	187	289	0	53
MWEA.21	148	179	204	196	127	257	226	278	1	54
MWEA.13	151	189	190	212	113	250	209	294	0	55

Sample ID	DhfrU1	DhfrU2	DhfrU3	DhfrU4	DhfrD1	DhfrD2	DhfrD3	DhfrD4	Mutations	Haplotyj
BONDO.55	152	162	170	220	90	238	175	271	1	56
MWEA.58	154	179	170	253	116	266	195	300	1	57
MWEA.52	158	181	170	216	111	233	198	278	0	58
MWEA.9	160	185	180	220	113	253	187	294	0	59

APPENDIX 3

Dhps-linked microsatellite haplotypes arranged in their allele sizes based on the DhpsU1 allele.

Sample ID	DhpsU1	DhpsU2	DhpsU3	DhpsU4	DhpsD1	DhpsD2	DhpsD3	DhpsD4	Mutations	Haplotype
BONDO.46	120	232	154	193	106	271	236	205	0	1
BONDO.49	122	196	136	199	109	280	246	200	0	2
BONDO.39	124	221	150	194	112	253	218	229	1	3
TIWI.6	124	221	150	194	112	253	218	229	1	3
TIWI.46	124	222	154	189	114	279	226	208	0	4
BONDO.42	127	196	154	197	119	279	227	204	0	5
MWEA.4	129	194	149	200	117	243	210	196	0	6
TIWI.56	129	194	149	200	106	243	210	196	0	6
BONDO.12	130	205	136	189	106	268	231	208	0	7
BONDO.41	130	205	136	189	106	268	231	208	0	7
BONDO.40	132	206	136	192	112	260	246	230	0	8
BONDO.24	133	204	140	194	112	253	218	212	1	9
BONDO.48	133	204	140	194	112	253	218	212	1	9
BONDO.9	133	204	140	194	112	253	218	212	1	9
TIWI.24	133	212	163	188	112	265	218	225	1	10
TIWI.57	133	212	163	188	112	265	218	225	1	10
MWEA.46	136	198	125	218	119	268	218	208	2	11
MWEA.36	136	205	113	188	115	268	215	213	0	12
MWEA.32	136	210	125	213	111	246	210	208	2	13
MWEA.64	136	212	125	220	114	224	200	189	1	14
MWEA.9	139	200	130	233	108	219	210	200	2	15
BONDO.21	139	200	160	188	112	253	235	208	1	16
TIWI.38	139	200	160	188	112	253	235	208	1	16
TIWI.21	139	210	141	196	114	271	220	218	0	17
MWEA.43	139	210	160	207	112	238	215	224	0	18
TIWI.35	139	212	163	188	112	253	218	208	1	19
TIWI.41	139	212	163	188	112	253	218	208	1	19
TIWI.55	139	212	163	188	112	253	218	208	1	19
TIWI.58	139	212	163	188	112	253	218	208	1	19
MWEA.54	139	224	125	213	117	238	210	218	1	20
TIWI.60	139	224	150	194	112	271	224	229	1	21
TIWI.9	139	224	163	188	112	253	235	213	1	21
BONDO.37	140	196	133	195	110	250	226	225	0	22
BONDO.56	140	200	154	198	116	266	210	210	0	23
MWEA.28	140	210	125	220	121	265	215	217	2	24
MWEA.48	140	210	125	220	121	265	215	217	2	24
MWEA.44	141	210	125	220	104	265	215	217	2	25
MWEA.52	143	192	147	220	111	253	190	229	0	26
MWEA.58	143	196	125	213	119	253	200	229	0	27
BONDO.1	143	204	130	194	112	265	218	212	1	28
BONDO.11	143	204	130	194	112	265	218	212	1	28
BONDO.14	143	204	130	194	112	265	218	212	1	28
BONDO.52	143	204	130	194	112	265	218	212	1	28
MWEA.40	143	208	141	227	109	246	238	217	2	29
BONDO.60	143	208	147	198	112	253	235	216	1	30
BONDO.44	143	210	135	188	112	265	218	224	1	31
TIWI.1	143	212	140	194	112	265	218	204	1	32

Sample ID	DhpsU1	DhpsU2	DhpsU3	DhpsU4	DhpsD1	DhpsD2	DhpsD3	DhpsD4	Mutations	Haplotype
TIWI.13	143	212	150	188	112	271	224	213	1	33
MWEA.66	143	216	146	207	115	257	215	200	2	34
BONDO.19	143	224	140	194	112	253	218	229	1	35
BONDO.2	143	224	140	194	112	253	218	229	1	35
BONDO.38	143	224	140	194	112	253	218	229	1	35
BONDO.59	143	224	140	194	112	253	218	229	1	35
BONDO.27	143	224	150	188	112	265	218	213	1	36
BONDO.43	143	224	150	188	112	265	218	213	1	36
TIWI.52	143	224	150	188	112	265	218	213	1	36
MWEA.68	145	208	119	213	116	253	224	229	2	37
MWEA.33	148	200	125	213	124	224	210	217	2	38
MWEA.42	148	200	138	198	116	210	210	217	0	39
TIWI.54	148	200	138	200	119	276	212	213	2	40
BONDO.26	148	200	140	200	119	276	212	224	2	41
TIWI.23	148	200	140	200	119	276	212	224	2	41
TIWI.27	148	200	140	200	119	276	212	224	2	41
BONDO.22	148	200	160	200	119	276	212	213	2	42
BONDO.28	148	200	160	200	119	276	212	208	2	43
BONDO.30	148	200	160	200	119	276	212	208	2	43
BONDO.31	148	200	160	200	119	276	212	213	2	42
BONDO.45	148	200	160	200	119	276	212	213	2	42
BONDO.50	148	200	160	200	119	276	212	213	2	42
BONDO.57	148	200	160	200	119	276	212	213	2	42
TIWI.11	148	200	160	200	119	276	212	213	2	42
TIWI.18	148	200	160	200	119	276	212	213	2	42
TIWI.19	148	200	160	200	119	276	212	213	2	42
TIWI.22	148	200	160	200	119	276	212	213	2	42
TIWI.25	148	200	160	200	119	276	212	213	2	42
TIWI.28	148	200	160	200	119	276	212	213	2	42
TIWI.29	148	200	160	200	119	276	212	213	2	42
TIWI.3	148	200	160	200	119	276	212	208	2	43
TIWI.30	148	200	160	200	119	276	212	213	2	42
TIWI.4	148	200	160	200	119	276	212	213	2	42
TIWI.40	148	200	160	200	119	276	212	213	2	42
TIWI.48	148	200	160	200	119	276	212	213	2	42
TIWI.5	148	200	160	200	119	276	212	213	2	42
TIWI.53	148	200	160	200	119	276	212	213	2	42
TIWI.59	148	200	160	200	119	276	212	213	2	42
MWEA.50	148	204	138	213	119	261	200	196	2	44
BONDO.15	148	206	130	194	112	265	218	229	1	45
MWEA.2	148	210	130	213	119	219	210	197	0	46
MWEA.3	148	216	113	220	115	238	215	200	2	47
BONDO.58	148	216	138	200	108	276	218	213	2	48
TIWI.34	148	216	138	200	108	276	218	213	2	48
BONDO.51	150	193	150	185	117	268	220	224	0	49
MWEA.55	150	196	125	198	104	229	210	200	0	50
MWEA.67	150	196	125	198	119	229	210	200	1	50
MWEA.11	152	200	125	213	113	224	210	200	0	51
MWEA.29	152	200	125	213	105	224	210	200	2	51
MWEA.45	152	200	125	213	112	224	210	200	1	51

Sample ID	DhpsU1	DhpsU2	DhpsU3	DhpsU4	DhpsD1	DhpsD2	DhpsD3	DhpsD4	Mutations	Haplotype
MWEA.65	152	200	125	213	118	224	210	200	2	51
MWEA.27	152	200	143	220	120	243	210	205	0	52
MWEA.56	152	204	119	233	113	238	224	196	0	53
MWEA.62	152	204	119	233	109	238	224	196	2	53
MWEA.19	152	210	141	207	109	267	218	200	2	54
MWEA.47	152	216	125	224	112	229	210	229	0	55
MWEA.41	152	216	135	233	116	219	215	200	0	56
MWEA.22	152	232	125	207	114	253	215	200	2	57
MWEA.34	155	192	125	220	117	243	210	200	0	58
TIWI.15	155	196	156	194	117	272	224	225	0	59
MWEA.17	155	200	125	213	106	205	210	200	2	60
TIWI.10	155	200	170	198	112	250	224	229	0	61
MWEA.1	155	210	133	213	108	229	210	197	1	62
MWEA.37	155	224	146	207	108	253	206	229	2	63
MWEA.31	157	206	125	240	116	253	190	229	1	64
MWEA.20	157	210	125	213	106	210	224	208	1	65
BONDO.33	157	220	160	194	112	265	218	224	1	66
TIWI.49	157	220	160	194	112	265	218	224	1	66
MWEA.13	158	200	113	240	118	243	210	200	0	67
MWEA.23	158	200	119	227	114	229	190	196	2	68
MWEA.26	158	200	119	227	114	229	190	196	2	68
MWEA.8	158	200	130	220	107	243	215	197	2	69
MWEA.25	158	200	141	213	120	261	210	200	0	70
MWEA.14	158	210	133	213	120	210	210	205	0	71
MWEA.5	158	216	130	220	120	219	210	193	2	72
MWEA.51	158	216	160	207	122	246	210	218	0	73
MWEA.7	160	210	130	220	122	219	210	200	0	74
BONDO.55	160	222	138	187	118	258	218	218	0	75
MWEA.57	160	224	141	220	118	205	215	217	1	76
MWEA.24	163	200	125	233	108	246	210	189	0	77
MWEA.60	163	210	125	207	106	246	200	224	0	78
MWEA.30	165	192	125	240	110	266	215	210	1	79
MWEA.12	165	210	125	213	113	219	215	197	1	80
MWEA.35	165	216	125	209	109	219	200	196	2	81
MWEA.38	165	220	113	220	117	205	200	217	0	82

Bibliography

Abdel-Muhsin, A. M., Ranford-Cartwright, L. C., Medani, A. R., Ahmed, S., Suleiman, S., Khan, B. *et al.* (2002). Detection of mutations in the *Plasmodium falciparum* dihydrofolate reductase (dhfr) gene by dot-blot hybridization: a potential method for epidemiological surveys of drug resistance genes. *American Journal Of Tropical Medicine & Hygiene*. **67:** 24-27.

Abderrazak, S. B., Oury, B., Lal, A. A., Bosseno, M. F., Force-Barge, P., Dujardin, J. P. *et al.* (1999). *Plasmodium falciparum*: population genetic analysis by multilocus enzyme electrophoresis and other molecular markers. *Experimental Parasitology*. **92:** 232-238.

Adagu, I. S., Dias, F., M Pinheiro, L., Rombo, L., do Rosario, V. E., & Warhurst, D. C. (1995b). Guinea Bissau: association of chloroquine resistance of *Plasmodium falciparum* with the Tyr86 allele of the multiple drug-resistant gene *Pfmdr1*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **90:** 90-91.

Adagu, I. S., Warhurst, D. C., Carucci, D. J., & Duraisingh, M. T. (1995a). *Pfmdr1* mutations and chloroquine-resistance in *Plasmodium falciparum* isolates from Zaire, Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **89:** 132

Al Yaman, F., Genton, B., Anders, R., Taraika, J., Ginny, M., Mellor, S. *et al.* (1995). Assessment of the role of humoral response to *Plasmodium falciparum* MSP2 compared to RESA and SPf66 in protecting Papua New Guinea children from clinical malaria. *Parasite Immunology*. **17:** 493-501.

Al Yaman, F., Genton, B., Anders, R. F., Falk, M., Triglia, T., Lewis, D. *et al.* (1994). Relationship between humoral response to *Plasmodium falciparum* merozoite surface antigen-2 and malaria morbidity in a highly endemic area of Papua New Guinea. *American Journal Of Tropical Medicine & Hygiene*. **51:** 593-602.

Amos, W., Sawcer, S. J., Feakes, R. W., & Rubinsztein, D. C. (1996). Microsatellites show mutational bias and heterozygote instability. *Nature Genetics*. **13:** 390-391.

Anders, R. F., Brown, G. V., Coppel, R. L., & Kemp, D. J. (1986). Repeat structures in malaria antigens. *Papua New Guinea Medical Journal*. **29:** 87-93.

Anders, R. F. & Saul, A. (2000). Malaria vaccines. *Parasitol Today*. **16:** 444-447.

Anderson, T. J. C., Haubold, B., Williams, J. T., Estrada-Franco, J. G., Richardson, L., Mollinedo, R. *et al.* (2000a). 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. & Roper, C. (2005). The origins and spread of antimalarial drug resistance: Lessons for policy makers. *Acta Tropica*. **94:** 269-280.

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

Anderson, T. J. C., Su, X. z., Roddam, A., & Day, K. P. (2000b). Complex mutations in a high proportion of microsatellite loci from the protozoan parasite *Plasmodium falciparum*. *Molecular Ecology*. **9:** 1599-1608.

Angers, B. & Bernatchez, L. (1997). Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Molecular Biology and Evolution*. **14:** 230-238.

Armour, J. A., Neumann, R., Gobert, S., & Jeffreys, A. J. (1994). Isolation of human simple repeat loci by hybridization selection. *Human Molecular Genetics*. **3:** 599-565.

Armour, J. A., Wong, Z., Wilson, V., Royle, N. J., & Jeffreys, A. J. (1989). Sequences flanking the repeat arrays of human minisatellites: association with tandem and dispersed repeat elements. *Nucleic Acids Research*. **17:** 4925-4935.

Arnot, D. E. (1998). Unstable malaria in Sudan: the influence of the dry season. Clone multiplicity of *Plasmodium falciparum* infections in individuals exposed to variable levels of disease transmission. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **92:** 580-585.

Arnot, D. E. & Gull, K. (1998). The *Plasmodium* cell-cycle: facts and questions. *Annals of Tropical Medicine and Parasitology*. **92:** 361-365.

Avery Jones, S. (1958). Mass treatment with pyrimethamine A study of resistance and cross resistance resulting from a field trial in the hyperendemic malarious area of Makueni, Kenya. September 1952 -- September 1953. *Transactions of the Royal Society of Tropical Medicine and Hygiene* . **52:** 547-561.

Awad-el-Kariem, F. M., Miles, M. A., & Warhurst, D. C. (1992). Chloroquine-resistant *Plasmodium falciparum* isolates from the Sudan lack two mutations in the *pfmdr1* gene thought to be associated with chloroquine resistance. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **86:** 587-589.

Babiker, H. A., Charlwood, J. D., Smith, T., & Walliker, D. (1995). Gene flow and cross-mating in *Plasmodium falciparum* in households in a Tanzanian village. *Parasitology*. **111:** 433-442.

Babiker, H. A., Creasey, A. M., Bayoumi, R. A., Walliker, D., & Arnot, D. E. (1991a). Genetic diversity of *Plasmodium falciparum* in a village in eastern Sudan. 2. Drug resistance, molecular karyotypes and the mdr1 genotype of recent isolates. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **85:** 578-583.

Babiker, H. A., Creasey, A. M., Fenton, B., Bayoumi, R. A., Arnot, D. E., & Walliker, D. (1991b). Genetic diversity of *Plasmodium falciparum* in a village in eastern Sudan. 1. Diversity of enzymes, 2D-PAGE proteins and antigens. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **85:** 572-577.

Babiker, H. A., Lines, J. D., Hill, W. G., & Walliker, D. (1997). Population structure of *Plasmodium falciparum* in villages with different malaria endemicity in east Africa. *American Journal Of Tropical Medicine & Hygiene*. **56:** 141-147.

Babiker, H. A., Pringle, S. J., Abdel-Muhsin, A. A., Mackinnon, M. J., Hunt, P., & Walliker, D. (2001). High-level chloroquine resistance in Sudanese isolates of *Plasmodium falciparum* is associated with mutations in the chloroquine resistance transporter gene *pfcrt* and the multidrug resistance Gene *pfmdr1*. *Journal of Infectious Diseases*. **183:** 1535-1538.

Babiker, H. A., Ranford-Cartwright, L. C., Currie, D., Charlwood, J. D., Billingsley, P., Teuscher, T. *et al.* (1994). Random mating in a natural population of the malaria parasite *Plasmodium falciparum*. *Parasitology*. **109:** 413-421.

Babiker, H. A. & Walliker, D. (1997). Current views on the population structure of *Plasmodium falciparum*: Implications for control. *Parasitology Today*. **13:** 262-267.

Bahl, A., Brunk, B., Coppel, R. L., Crabtree, J., Diskin, S. J., Fraunholz, M. J. *et al.* (2002). PlasmoDB: the *Plasmodium* genome resource. An integrated database providing tools for accessing, analyzing and mapping expression and sequence data (both finished and unfinished). *Nucleic Acids Research*. **30:** 87-90.

Baird, J. K., Jones, T. R., Danudirgo, E. W., Annis, B. A., Bangs, M. J., Basri, H. *et al.* (1991). Age-dependent acquired protection against Plasmodium falciparum in people having two years exposure to hyperendemic malaria. *American Journal Of Tropical Medicine & Hygiene*. **45:** 65-76.

Baldwin, J., Farajallah, A. M., Malmquist, N. A., Rathod, P. K., & Phillips, M. A. (2002). Malarial dihydroorotate dehydrogenase: substrate and inhibitor specificity. *Journal of Biological Chemistry*. **277:** 41827-41834.

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

Barnes, T. M., Kohara, Y., Coulson, A., & Hekimi, S. (1995). Meiotic Recombination, Noncoding DNA and Genomic Organization in Caenorhabditis elegans. *Genetics*. **141:** 159-179.

Barton, N. H. (2000). Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. **355:** 1553-1562.

Basco, L. K., de Pecoulas, P. E., Le Bras, J., & Wilson, C. M. (1996). *Plasmodium falciparum*: molecular characterization of multidrug-resistant Cambodian isolates. *Experimental Parasitology*. **82:** 97-103.

Basco, L. K., Le Bras, J., Rhoades, Z., & Wilson, C. M. (1995). Analysis of *pfmdr1* and drug susceptibility in fresh isolates of *Plasmodium falciparum* from subsaharan Africa. *Molecular and Biochemical Parasitology*. **74:** 157-166.

Baudon, D. (1995). La prise de decision dans la politique d'utilisation des antipaludiques en reponse a une modification de l'efficacite de la chloroquine. Reflexions appliquees a l'Afrique. *Medecine Tropicale*. **55:** S37-S45.

Bayoumi, R. A., Creasey, A. M., Babiker, H. A., Carlton, J. M., Sultan, A. A., Satti, G. *et al.* (1993). Drug response and genetic characterization of *Plasmodium falciparum* clones recently isolated from a Sudanese village. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **87:** 454-458.

Beaumont, M. A. (1999). Detecting population expansion and decline using microsatellites. *Genetics*. **153:** 2013-2029.

Beaumont, M. A. & Bruford, M. W.(1999). Microsatellites in conservation genetics: In *Microsatellites. Evolution and Applications*. (Ed.Goldstein, D. B. and Schlotterer, C.) 165-182. Oxford University Press, Oxford.

Bender, A., van Dooren, G. G., Ralph, S. A., McFadden, G. I., & Schneider, G. (2003). Properties and prediction of mitochondrial transit peptides from Plasmodium falciparum. *Molecular and Biochemical Parasitology*. **132:** 59-66.

Bennett, P. (2000). Demystified... Microsatellites. *Journal of Clinical Pathology: Molecular Pathology*. **53:** 177-183.

Bennett, T. N., Kosar, A. D., Ursos, L. M. B., Dzekunov, S., Sidhu, A. B. S., Fidock, D. A. *et al.* (2004). Drug resistance-associated pfCRT mutations confer decreased *Plasmodium falciparum* digestive vacuolar pH. *Molecular and Biochemical Parasitology*. **133:** 99-114.

Bishop, A. (1963). Some recent developments in the problem of drug resistance in malaria. *Parasitology*. **53:** 10.

Bloland, P. B., Lackritz, E. M., Kazembe, P. N., Were, J. B. O., Steketee, R. W., & Campbell, C. C. (1993). Beyond chloroquine: implications of drug resistance for evaluating malaria therapy efficacy and treatment policy in Africa. *Journal of Infectious Diseases*. **167:** 932-937.

Borre, M. B., Dziegiel, M., Hogh, B., Petersen, E., Rieneck, K., Riley, E. *et al.* (1991). Primary structure and localization of a conserved immunogenic Plasmodium falciparum glutamate rich protein (GLURP) expressed in both the preerythrocytic and erythrocytic stages of the vertebrate life cycle. *Molecular and Biochemical Parasitology*. **49:** 119-131.

Bottius, E., Guanzirolli, A., Trape, J. F., Rogier, C., Konate, L., & Druilhe, P. (1996). Malaria: even more clonic in nature than previously thought; evidence for subpatent parasitaemia detectable by the polymerase chain reaction. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **90:** 15-19.

Boudin, C., Chumpitazi, B., Dziegiel, M., Peyron, F., Picot, S., Hogh, B. *et al.* (1993). Possible role of specific immunoglobulin M antibodies to *Plasmodium falciparum*

antigens in immunoprotection of humans living in a hyperendemic area, Burkina Faso. *Journal of Clinical Microbiology*. **31:** 636-641.

Bray, P. G., Hawley, S. R., Mungthin, M., & Ward, S. A. (1996). Physicochemical properties correlated with drug resistance and the reversal of chloroquine resistance in *Plasmodium flaciparum*. *Molecular Pharmacology*. **50:** 1559-1566.

Bray, P. G., Howells, R. E., Ritchie, G. Y., & Ward, S. A. (1992a). Rapid chloroquine efflux phenotype in both chloroquine-sensitive and chloroquine-resistant *Plasmodium falciparum*. A correlation of chloroquine sensitivity with energy-dependent drug accumulation. *Biochemical Pharmacology*. **44:** 1317-1324.

Bray, P. G., Howells, R. E., & Ward, S. A. (1992b). Vacuolar acidification and chloroquine sensitivity in *Plasmodium falciparum*. *Biochemical Pharmacology*. **43:** 1219-1227.

Bray, P. G., Martin, R. E., Tilley, L., Ward, S. A., Kirk, K., & Fidock, D. A. (2005). Defining the role of PfCRT in *Plasmodium falciparum* chloroquine resistance. *Molecular Microbiology*. **56:** 323-333.

Bray, P. G., Mungthin, M., Ridley, R., & Ward, S. A. (1998). Access to hematin: the basis of chloroquine resistance. *Molecular Pharmacology*. **54:** 170-179.

Bray, P. G., Saliba, K. J., Davies, J. D., Spiller, D. G., White, M. R., Kirk, K. *et al.* (2002). Distribution of acridine orange fluorescence in *Plasmodium falciparum*-infected erythrocytes and its implications for the evaluation of digestive vacuole pH. *Molecular and Biochemical Parasitology*. **119:** 301-304.

Breman, J. G. (2001). The ears of the hippopotamus: manifestations, determinants, and estimates of the malaria burden. *American Journal Of Tropical Medicine & Hygiene*. **64:** 1-11.

Britten, R. J. & Kohne, D. E. (1968). Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science*. **161:** 529-540.

Brock, G. J., Anderson, N. H., & Monckton, D. G. (1999). Cis-acting modifiers of expanded CAG/CTG triplet repeat expandability: associations with flanking GC content and proximity to CpG islands. *Human Molecular Genetics*. **8:** 1061-1067.

Brockman, A., Paul, R. E., Anderson, T. J. C., Hackford, I., Phaiphun, L., Looareesuwan, S. *et al.* (1999). Application of genetic markers to the identification of recrudescent *Plasmodium falciparum* infections on the northwestern border of Thailand. *American Journal Of Tropical Medicine & Hygiene*. **60:** 14-21.

Brook, J., McCurrach, M. E., & Harley, H. G. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell*. **68:** 799-808.

Brooks, D. R., Wang, P., Read, M., Watkins, W. M., Sims, P. F., & Hyde, J. E. (1994). Sequence variation of the hydroxymethyldihydropterin pyrophosphokinase: dihydropteroate synthase gene in lines of the human malaria parasite, *Plasmodium falciparum*, with differing resistance to sulfadoxine. *European Journal of Biochemistry*. **224:** 397-405.

Brown, A. H. D., Feldman, M. W., & Nevo, E. (1980). Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics*. **96:** 523-536.

Bruce-Chwatt, L. J. (1956). Chemotherapy in relation to possibilities of malaria eradication in tropical Africa. *Bulletin of the World Health Organization*. **15:** 852-862.

Bruce, M. C., Alano, P., Duthie, S., & Carter, R. (1990). Commitment of the malaria parasite *Plasmodium falciparum* to sexual and asexual development. *Parasitology*. **100:** 191-200.

Burkot, T. R. & Graves, P. M. (1995). The value of vector-based estimates of malaria transmission. *Annals of Tropical Medicine and Parasitology*. **89:** 125-134.

Bwibo, N. O. & Neumann, C. G. (2003). The Need for Animal Source Foods by Kenyan Children. *Journal of Nutrition*. **133:** 3936S-3940.

Bzik, D. J., Li, W. B., Horii, T., & Inselburg, J. W. (1987). Molecular cloning and sequence analysis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. *Proceedings of the National Academy of sciences*. **84:** 8360-8364.

Campbell, C., Collins, W. E., Chin, W., Teutsch, S., & Moss, D. (1979). Chloroquine-resistant *Plasmodium falciparum* from East Africa: Cultivation and Drug Sensitivity of the Tanzanian I/CDC Strain from an American Tourist. *The Lancet*. **314:** 1151-1154.

Cappai, R., van Schravendijk, M. R., Anders, R. F., Peterson, M. G., Thomas, L. M., Cowman, A. F. *et al.* (1989). Expression of the RESA gene in *Plasmodium falciparum* isolate FCR3 is prevented by a subtelomeric deletion. *Molecular and Cellular Biology*. **9:** 3584-3587.

Carter, R. & McGregor, I. A. (1973). Enzyme variation in Plasmodium falciparum in the Gambia. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **67:** 830-837.

Certa, U. (1990). *Plasmodium falciparum* surface antigen p 1 90: prevalence of a nonrepetitive allele in humans. *Parasitology Today*. **6:** 373.

Chambers, G. K. & MacAvoy, E. S. (2000). Microsatellites: consensus and controversy. *Comparative Biochemical Physiology B: Biochemical Molecular Biology*. **126:** 455-476.

Charles, L. J. (1961). Field trials with chlorproguanil in the prophylaxis of malaria in Ghana. *Bulletin of the World Health Organization*. **24:** 457-463.

Charles, L. J., van der Kaay, H. J., Vincke, I. H., & Brady, J. (1962). The appearance of pyrimethamine resistance in *Plasmodium falciparum* following self-medication by a rural community in Ghana. *Bulletin of the World Health Organization*. **26:** 108-

Chen, G. X., Mueller, C., Wendlinger, M., & Zolg, J. W. (1987). Kinetic and molecular properties of the dihydrofolate reductase from pyrimethamine-sensitive and pyrimethamine-resistant clones of the human malaria parasite *Plasmodium falciparum*. *Molecular Pharmacology*. **31:** 430-437.

Chen, N., Kyle, D. E., Pasay, C., Fowler, E. V., Baker, J., Peters, J. M. *et al.* (2003). *pfcrt* Allelic types with two novel amino acid mutations in chloroquine-resistant *Plasmodium falciparum* isolates from the Philippines. *Antimicrobial Agents and Chemotherapy*. **47:** 3500-3505.

Chou, A. C., Chevli, R., & Fitch, C. D. (1980). Ferriprotoporphyrin IX fulfills the criteria for identification as the chloroquine receptor of malaria parasites. *Biochemistry*. **19:** 1543-1549.

Chulay, J. D., Watkins, W. M., & Sixsmith, D. G. (1984). Synergistic antimalarial activity of pyrimethamine and sulfadoxine against *Plasmodium falciparum in vitro*. *American Journal Of Tropical Medicine & Hygiene*. **33:** 325-330.

Clarke, R., Odialla, H., Ouma, J., Kenny, V., MacCabe, R., Rapuoda, B. *et al.* (1996). A malariametric survey in Turkana District Kenya: chemo-sensitivity of *Plasmodium falciparum* infections and the identity of the vector. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **90:** 320-324.

Clyde, D. F. & Shute, G. T. (1954). Resistance of East African varieties of *Plasmodium falciparum* to pyrimethamine. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **48:** 495-500.

Clyde, D. F. & Shute, G. T. (1957). Resistance of *Plasmodium falciparum* in Tanganyika to pyrimethamine administered at weekly intervals. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **51:** 505-513.

Cohen J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological measurement*. **20:** 37-46.

Contamin, H., Fandeur, T., Rogier, C., Bonnefoy, S., Konate, L., Trape, J. F. *et al.* (1996). Different genetic characteristics of *Plasmodium falciparum* isolates collected during successive clinical malaria episodes in Senegalese children. *American Journal Of Tropical Medicine & Hygiene*. **54:** 632-643.

Conway, D. J. & McBride, J. S. (1991). Population genetics of *Plasmodium falciparum* within a malaria hyperendemic area. *Parasitology*. **103:** 7-16.

Conway, D. J., Rosario, V., Oduola, A. M. J., Salako, L. A., Greenwood, B. M., & McBride, J. S. (1991). *Plasmodium falciparum*: intragenic recombination and nonrandom associations between polymorphic domains of the precursor to the major merozoite surface antigens. *Experimental Parasitology*. **73:** 469-480.

Corcoran, L. M., Thompson, J. K., Walliker, D., & Kemp, D. J. (1988). Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in P. falciparum. *Cell*. **53:** 807-813.

Cornuet, J. M. & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*. **144:** 2001-2014.

Cortese, J. F., Caraballo, A., Contreras, C. E., & Plowe, C. V. (2002). Origin and dissemination of *Plasmodium falciparum* drug-resistance mutations in South America. *Journal of Infectious Diseases*. **186:** 999-1006.

Cortese, J. F. & Plowe, C. V. (1998). Antifolate resistance due to new and known *Plasmodium falciparum* dihydrofolate reductase mutations expressed in yeast. *Molecular and Biochemical Parasitology*. **94:** 205-214.

Cowman, A. F., Karcz, S., Galatis, D., & Culvenor, J. G. (1991). A P-glycoprotein homologue of *Plasmodium falciparum* is localized on the digestive vacuole. *Journal of Cell Biology*. **113:** 1033-1042.

Cowman, A. F. & Karcz, S. R. (1991). The *pfmdr* gene homologues of *Plasmodium falciparum*. *Acta Leiden*. **60:** 121-129.

Cowman, A. F. & Lew, A. M. (1989). Antifolate drug selection results in duplication and rearrangement of chromosome 7 in *Plasmodium chabaudi*. *Molecular and Cellular Biology*. **9:** 5182-5188.

Cowman, A. F. & Lew, A. M. (1990). Chromosomal rearrangements and point mutations in the DHFR-TS gene of *Plasmodium chabaudi* under antifolate selection. *Molecular and Biochemical Parasitology*. **42:** 21-29.

Cowman, A. F., Morry, M. J., Biggs, B. A., Cross, G. A. M., & Foote, S. J. (1988). Amino Acid Changes Linked to Pyrimethamine Resistance in the Dihydrofolate Reductase-Thymidylate Synthase Gene of *Plasmodium falciparum*. *Proceedings of the National Academy of sciences*. **85:** 9109-9113.

Cox-Singh, J., Singh, B., Alias, A., & Abdullah, M. S. (1995). Assessment of the association between three pfmdr1 point mutations and chloroquine resistance in vitro of Malaysian *Plasmodium falciparum* isolates. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **89:** 436-437.

Creasey, A. M., Ranford-Cartwright, L. C., Moore, D. J., Williamson, D. H., WILSON, R. J. M., Walliker, D. *et al.* (1993). Uniparental inheritance of the mitochondrial gene cytochrome b in *Plasmodium falciparum*. *Current Genetics*. **23:** 360-364.

Curd, F. H. S., Davey, D. G., & Rose, F. L. (1945). Studies on synthetic antimalarial drugs. X. Some biguanide derivatives as new types of antimalarial substances with both therapeutic and causal prophylactic activity. *Annals of Tropical Medicine and Parasitology*. **39:** 208-216.

Curtis, C. F. & Otoo, L. N. (1986). A simple model of the build-up of resistance to mixtures of anti-malarial drugs. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **80:** 889-892.

Curtis, J., Duraisingh, M. T., Trigg, J. K., Mbwana, H., Warhurst, D. C., & Curtis, C. F. (1996). Direct evidence that asparagine at position 108 of the *Plasmodium falciparum* dihydrofolate reductase is involved in resistance to antifolate drugs in Tanzania. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **90:** 678-680.

Curtis, J., Duraisingh, M. T., & Warhurst, D. C. (1998). *In vivo* selection for a specific genotype of dihydropteroate synthetase of *Plasmodium falciparum* by pyrimethamine-sulfadoxine but not chlorproguanil-dapsone treatment. *Journal of Infectious Diseases*. **177:** 1429-1433.

de Duve, C., De Barsy, T., Poole, B., Trouet, A., Tulkens, P., & Van Hoof, F. (1974). Lysosomotrophic agents. *Biochemical Pharmacology*. **23:** 2495-2531.

de Roode, J. C., Culleton, R., Cheesman, S. J., Carter, R., & Read, A. F. (2004). Host heterogeneity is a determinant of competitive exclusion or coexistence in genetically diverse malaria infections. *Proceedings: Biological Sciences*. **271:** 1073-1080.

Debrauwere, H., Gendrel, C. G., & Lechat, S. (1997). Differences and similarities between various tandem repeat sequences: minisatellites and microsatellites. *Biochimie*. **79:** 577-586.

Diggens, S. M. (1970). Single step production of pyrimethamine-resistant *P. berghei*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **64:** 9.

Diggens, S. M., Gutteridge, W. E., & Trigg, P. I. (1970). Altered dihydrofolate reductase associated with a pyrimethamine-resistant *Plasmodium berghei berghei* produced in a single step. *Nature*. **228:** 579-580.

Dileepan, K. N. & Kennedy, J. (1985). Complete inhibition of dihydroorotate and superoxide production by 1,1,1-trifluoro-3-thenoylacetone in rat liver mitochondria. *Biochemical Journal*. **225:** 194.

Division of Malaria Control 1998. Malaria: A situation analysis for Kenya. Ministry of Health, Kenya. Nairobi.

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

Dodoo, D., Theisen, M., Kurtzhals, J. A., Akanmori, B. D., Koram, K. A., Jepsen, S. *et al.* (2000). Naturally acquired antibodies to the glutamate-rich protein are associated with protection against *Plasmodium falciparum* malaria. *Journal of Infectious Diseases*. **181:** 1202-1205.

Doumbo, O. K., Kayentao, K., Djimde, A. A., Cortese, J. F., Diourte, Y., Konare, A. *et al.* (2000). Rapid selection of *Plasmodium falciparum* dihydrofolate reductase mutants by pyrimethamine prophylaxis. *Journal of Infectious Diseases*. **182:** 993-996.

Druilhe, P., Daubersies, P., Patarapotikul, J., Gentil, C., Chene, L., Chongsuphajaisiddhi, T. *et al.* (1998). A primary malarial infection is composed of a very wide range of genetically diverse but related parasites. *Journal of Clinical Investigation*. **101:** 2008-1016.

Druilhe, P. & Perignon, J. L. (1997). A hypothesis about the chronicity of malaria infection. *Parasitology Today*. **13:** 353-357.

Duraisingh, M. T., Jones, P., Sambou, I., von Seidlein, L., Pinder, M., & Warhurst, D. C. (2000). The tyrosine-86 allele of the *pfmdr1* gene of *Plasmodium falciparum* is associated with increased sensitivity to the anti-malarials mefloquine and artemisinin. *Molecular and Biochemical Parasitology*. **108:** 13-23.

Durand, R., Jafari, S., Vauzelle, J., Delabre, J. F., Jesic, Z., & Le Bras, J. (2001). Analysis of *pfcrt* point mutations and chloroquine susceptibility in isolates of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **114:** 95-102.

Durrand, V., Berry, A., Sem, R., Glaziou, P., Beaudou, J., & Fandeur, T. (2004). Variations in the sequence and expression of the *Plasmodium falciparum* chloroquine resistance transporter (Pfcrt) and their relationship to chloroquine resistance *in vitro*. *Molecular and Biochemical Parasitology*. **136:** 273-285.

Dye, C. & Godfray, H. C. F. (1993). On Sex Ratio and Inbreeding in Malaria Parasite Populations. *Journal of Theoretical Biology*. **161:** 131-134.

Dzekunov, S., Ursos, L. M. B., & Roepe, P. D. (2000). Digestive vacuolar pH of intact intraerythrocytic *P. falciparum* either sensitive or resistant to chloroquine. *Molecular and Biochemical Parasitology*. **110:** 107-124.

EANMAT (2001). Monitoring Antimalarial Drug Resistance within National Malaria Control Programmes (NMCPs): the EANMAT experience. *Tropical Medicine and International Health*. **6:** 1-8.

Edoh, D., Mshinda, H., Jenkins, J., & Burger, M. (1997). Pyrimethamine-resistant *Plasmodium falciparum* parasites among Tanzanian children: a facility-based study using the polymerase chain reaction. *American Journal Of Tropical Medicine & Hygiene*. **57:** 342-347.

Ekong, R., Robson, K., Baker, D., & Warhurst, D. (1993). Transcripts of the multidrug resistance genes in chloroquine-sensitive and chloroquine-resistant *Plasmodium falciparum*. *Parasitology*. **106:** 107-115.

Ellegren, H. (2000). Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*. **16:** 551-558.

Epping, R. J., Goldstone, S. D., Ingram, L. T., Upcroft, J. A., Ramasamy, R., Cooper, J. A. *et al.* (1988). An epitope recognised by inhibitory monoclonal antibodies that react with a 51 kilodalton merozoite surface antigen in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **28:** 1-10.

Escalante, A. A., Lal, A. A., & Ayala, F. J. (1998). Genetic Polymorphism and Natural Selection in the Malaria Parasite *Plasmodium falciparum*. *Genetics*. **149:** 189-202.

Etlinger, H. M., Caspers, P., Matile, H., Schoenfeld, H. J., Stueber, D., & Takacs, B. (1991). Ability of recombinant or native proteins to protect monkeys against heterologous challenge with *Plasmodium falciparum*. *Infection and Immunity*. **59:** 3498-3503.

Evans, R. E. & Havlik, I. (1993). *Plasmodium falciparum*: effects of amantadine, an antiviral, on chloroquine-resistant and -sensitive parasites in vitro and its influence on chloroquine activity. *Biochemical Pharmacology*. **45:** 1168-1170.

Eyles, D., Moo, C., Warren, M., & Sandosham, A. (1963). *Plasmodium falciparum* resistant to chloroquine in Cambodia. *American Journal of Tropical Medicine and Hygiene*. **12:** 840-843.

Falco, E. A., Goodwin, L. G., Hitchings, G. H., Rollo, I. M., & Russell, P. B. (1951). 2:4-diaminopyrimidines- a new series of antimalarials. *British Journal of Pharmacology*. **6:** 185-200.

Farnert, A., Snounou, G., Rooth, I., & Bjorkman, A. (1997). Daily dynamics of Plasmodium falciparum subpopulations in asymptomatic children in a holoendemic area. *American Journal Of Tropical Medicine & Hygiene*. **56:** 538-547.

Feagin, J. E. (1992). The 6-kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. *Molecular and Biochemical Parasitology*. **52:** 145-148.

Feagin, J. E. & Drew, M. F. (1995). *Plasmodium falciparum*: alterations in organelle transcript abundance during the erythrocytic cycle. *Experimental Parasitology*. **80:** 430-440.

Felger, I., Marshal, V. M., Reeder, J. C., Hunt, J. A., Mgone, C. S., & 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.

Fenton, B., Clark, J. T., Khan, C. M., Robinson, J. V., Walliker, D., Ridley, R. *et al.* (1991). Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface antigen (MSA-2) of the malaria parasite *Plasmodium falciparum*. *Molecular and Cellular Biology*. **11:** 963-971.

Fenton, B., Clark, J. T., Wilson, C. F., McBride, J. S., & Walliker, D. (1989). Polymorphism of a 35-48 kDa *Plasmodium falciparum* merozoite surface antigen. *Molecular and Biochemical Parasitology*. **34:** 79-86.

Ferone, R. (1970). Dihydrofolate reductase from pyrimethamine-resistant *Plasmodium berghei*. *Journal of Biological Chemistry*. **245:** 850-854.

Ferone, R. (1973). The enzymic synthesis of dihydropteroate and dihydrofolate by *Plasmodium berghei*. *Journal of Protozoology*. **20:** 459-464.

Ferone, R. (1977). Folate metabolism in malaria. *Bulletin of the World Health Organization*. **55:** 291-298.

Ferone, R., O'Shea, M., & Yoeli, M. (1970). Altered dihydrofolate reductase associated with drug-resistance transfer between rodent plasmodia. *Science*. **167:** 1263-1264.

Ferone, R. & Roland, S. (1980). Dihydrofolate reductase: thymidylate synthase, a bifunctional polypeptide from *Crithidia fasciculata*. *Proceedings of the National Academy of sciences*. **77:** 5802-5806.

Ferreira, M. U., Liu, Q., Kaneko, O., Kimura, M., Tanabe, K., Kimura, E. A. *et al.* (1998). Allelic diversity at the merozoite surface protein-1 locus of *Plasmodium falciparum* in clinical isolates from the southwestern Brazilian Amazon. *American Journal of Tropical Medicine and Hygiene*. **59:** 474-480.

Fidock, D. A., Nomura, T., Cooper, R. A., Su, X. z., Talley, A. K., & Wellems, T. E. (2000b). Allelic modifications of the cg2 and cg1 genes do not alter the chloroquine response of drug-resistant *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* . **110:** 1-10.

Fidock, D. A., Nomura, T., Talley, A. K., Cooper, R. A., Dzekunov, S. M., Ferdig, M. T. *et al.* (2000a). Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Molecular Cell*. **6:** 861-871.

Fleiss, J. L., Nee, C. M., & Landis, J. R. (1979). Large sample variance of kappa in the case of different sets of raters. *Psychological Bulletin*. **86:** 974-977.

Fogh, S., Jepsen, S., & Effersoe, P. (1979). Chloroquine-resistant *Plasmodium falciparum* malaria in Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **73:** 228-229.

Foote, S. J., Galatis, D., & Cowman, A. F. (1990b). Amino Acids in the Dihydrofolate Reductase-Thymidylate Synthase Gene of *Plasmodium falciparum* Involved in Cycloguanil Resistance Differ from those Involved in Pyrimethamine Resistance. *Proceedings of the National Academy of sciences*. **87:** 3014-3017.

Foote, S. J. & Kemp, D. J. (1989). Chromosomes of malaria parasites. *Trends in Genetics*. **5:** 337-342.

Foote, S. J., Kyle, D. E., Martin, R. K., Oduola, A. M. J., Forsyth, K., Kemp, D. J. *et al.* (1990a). Several alleles of the multidrug-resistance gene are closely linked to chloroquine resistance in *Plasmodium falciparum*. *Nature*. **345:** 255-258.

Foote, S. J., Thompson, J. K., Cowman, A. F., & Kemp, D. J. (1989). Amplification of the multidrug resistance gene in some chloroquine-resistant isolates of *P. falciparum*. *Cell*. **57:** 921-930.

Frevert, U., Galinski, M. R., Hugel, F. U., Allon, N., Schreier, H., Smulevitch, S. *et al.* (1998). Malaria circumsporozoite protein inhibits protein synthesis in mammalian cells. *EMBP Journal*. **17:** 3816-3826.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W. *et al.* (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*. **419:** 498-511.

Gardner, M. J., Tettelin, H., Carucci, D. J., Cummings, L. M., Aravind, L., Koonin, E. V. *et al.* (1998). Chromosome 2 Sequence of the Human Malaria Parasite *Plasmodium falciparum* . *Science*. **282:** 1126-1132.

Garrett, C. E., Coderre, J. A., Meek, T. D., Garvey, E. P., Claman, D. M., Beverley, S. M. *et al.* (1984). A bifunctional thymidylate synthetase-dihydrofolate reductase in protozoa. *Molecular and Biochemical Parasitology*. **11:** 257-265.

Geary, T. G., Jensen, J. B., & Ginsburg, H. (1986). Uptake of [3H]chloroquine by drug-sensitive and -resistant strains of the human malaria parasite *Plasmodium falciparum*. *Biochemical Pharmacology*. **35:** 3805-3812.

Gerlach, J. H., Endicott, J. A., Juranka, P. F., Henderson, G., Sarangi, F., Deuchars, K. L. *et al.* (1986). Homology between P-glycoprotein and a bacterial haemolysin transport protein suggests a model for multidrug resistance. *Nature*. **324:** 485-489.

Giglioli, G., Rutten, F. J., & Ramjattan, S. (1967). Interruption of malaria transmission by chloroquinized salt in Guyana. *Bulletinof the World Health Organization*. **36:** 283-301.

Gill, P., Jeffreys, A. J., & Werrett, D. J. (1985). Forensic application of DNA 'fingerprints'. *Nature*. **318:** 577-579.

Ginot, F., Bordelais, I., Nguyen, S., & Gyapay, G. (1996). Correction of some genotyping errors in automated fluorescent microsatellite analysis by enzymatic removal of one base overhangs. *Nucleic Acids Research*. **24:** 540-541.

Ginsburg, H. & Stein, W. D. (1991). Kinetic modelling of chloroquine uptake by malaria-infected erythrocytes: assessment of the factors that may determine drug resistance. *Biochem Pharmacol*. **41:** 1463-1470.

Goldstein, D. B., Roemer, G. W., Smith, D. A., Reich, D. E., Bergman, A., & Wayne, R. K. (1999). The use of microsatellite variation to infer population structure and demographic history in a natural model system. *Genetics*. **151:** 797-801.

Goldstein, D. B. & Schlotterer, C. (1999). Microsatellites: Evolution and applications. Oxford University Press,

Good, M. F., Pombo, D., Maloy, W. L., de, l. C., V, Miller, L. H., & Berzofsky, J. A. (1988). Parasite polymorphism present within minimal T cell epitopes of *Plasmodium falciparum* circumsporozoite protein. *Journal of Immunology*. **140:** 1645-1650.

Gottesman, M. M. & Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annual Reviews in Biochemistry*. **62:** 385-427.

Goudet, J. (2000). FSTAT, a program to estimate and test gene diversities and fixation indices (Version 2.9.3). *http://www.unil.ch/izea/software/fstat.html*.

Greenberg, J., Boyd, B. L., & Josephson, E. S. (1948). Synergistic effect of chlorguanide and sulfadiazine against *Plasmodium gallinecium* in the chick. *Journal of Pharmacology and Experimental Therapeutics*. **94:** 60-64.

Greene, L. S. (1999). Genetic and dietary adaptation to malaria in human populations. *Parassitologia*. **44:** 185-192.

Grimaldi, M. C. & Crouau Roy, B. (1997). Microsatellite allelic homoplasy due to variable flanking sequences. *Journal of Molecular Evolution*. **44:** 336-340.

Grumont, R., Washtien, W., Caput, D., & Santi, D. (1986). Bifunctional thymidylate synthase-dihydrofolate reductase from *Leishmania tropica*: sequence homology with the corresponding monofunctional proteins. *Proceedings of the National Academy of sciences*. **83:** 5387-5391.

Gupta, S. & Day, K. P. (1994). A theoretical framework for the immunoepidemiology of *Plasmodium falciparum* malaria. *Parasite Immunology*. **16:** 361-370.

Gupta, S., Maiden, M. C., Feavers, I. M., Nee, S., May, M., & Anderson, R. M. (1996). The maintenance of strain structure in populations of recombining infectious agents. *Nature Medicine*. **2:** 437-442.

Hall, S. A. & Wilks, N. E. (1967). A trial of chloroquine-medicated salt for malaria suppression in Uganda. *American Journal of Tropical Medicine and Hygiene*. **16:** 429-442.

Hancock, J. M.(1999). Microsatellites and other simple sequences: Genomic context and mutational mechanisms: In *Microsatellites: Evolution and Applications*. (Ed.Goldstein, D. B. and Schlotterer, C.) 1-9. Oxford University Press, New York.

Harr, B., Kauer, M., & Schlotterer, C. (2002). Hitchhiking mapping: A population-based fine-mapping strategy for adaptive mutations in *Drosophila melanogaster*. *Proceedings of the National Academy of sciences*. **99:** 12949-12954.

Hartinuta, T., Migasen, S., Boonag, D. (1962). Chloroquine resistance in Thailand. *UNESCO 1st Regional Symposium on Science Knowledge of Tropical Parasites, 5-9 November, 1962*.143-153.

Hartl, D. L. & Clark, A. G. (1989). Darwinian evolution in mendelian populations. Principles of Population Genetics. Sinauer Associates Inc., Sunderland, Massachusets.

Hartl, D. L. & Dykhuizen, D. E. (1981). Potential for selection among nearly neutral allozymes of 6-phosphogluconate dehydrogenase in *Escherichia coli*. *Proceedings of the National Academy of sciences*. **78:** 6344-6348.

Hastings, I. M. (1996). Population genetics and the detection of immunogenic and drug resistant loci in *Plasmodium*. *Parasitology*. **112:** 155-164.

Hastings, I. M. (1997). A model for the origins and spread of drug-resistant malaria. *Parasitology*. **115:** 133-141.

Hastings, I. M. (2003). Malaria control and the evolution of drug resistance: an intriguing link. *Trends in Parasitology*. **19:** 70-73.

Hastings, I. M. (2004). The origins of antimalarial drug resistance. *Trends in Parasitology*. **20:** 512-518.

Hastings, I. M. & Donnelly, M. J. (2005). The impact of antimalarial drug resistance mutations on parasite fitness, and its implications for the evolution of resistance. *Drug Resistance Updates*. **8:** 43-50.

Hastings, I. M. & Mackinnon, M. (1998). The emergence of drug-resistant malaria. *Parasitology*. **117:** 411-417.

Hastings, I. M. & Watkins, W. M. (2005). Intensity of malaria transmission and the evolution of drug resistance. *Acta Tropica*. **94:** 218-229.

Hastings, I. M. & Wedgewood-Oppenheim, B. (1997). Sex, strains and virulence. *Parasitology Today*. **13:** 375-383.

Haubold, B. & Hudson, R. R. (2000). LIAN version 3: a program for detecting linkage disequilibrium in multilocus data. *Bioinformatics*. **16:** 847-848.

Hedrick, P. W. (1980). Hitchhiking: a comparison of linkage and partial selfing. *Genetics*. **94:** 791-808.

Heyen, D. W., Beever, J. E., Da, Y., Evert, R. E., Green, C., Bates, S. R. *et al.* (1997). Exclusion probabilities of 22 bovine microsatellite markers in fluorescent multiplexes for semiautomated parentage testing. *Animal Genetics*. **28:** 21-27.

Heyer, E., Puymirat, J., Dieltjes, P., Bakker, E., & de Knijff, P. (1997). Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Human Molecular Genetics*. **6:** 799-803.

Higgins, C. F. (1992). ABC transporters: from microorganisms to man. *Annual Reviews in Cell Biology*. **8:** 67-113.

Hinterberg K, Mattei D, Wellems TE, & Scherf A (1994). Interchromosomal exchange of a large subtelomeric segment in a Plasmodium falciparum cross. *EMBP Journal.* **13:** 4174-4180.

Hitchings, G. H. (1969). Species differences among dihydrofolate reductases as a basis for chemotherapy. *Postgraduate Medical Journal.* **45:** S7-S10.

Hogh, B., Petersen, E., Dziegiel, M., David, K., Hanson, A., Borre, M. B. *et al.* (1992). Antibodies to a recombinant glutamate-rich *Plasmodium falciparum* protein: evidence for protection of individuals living in a holoendemic area of Liberia. *American Journal Of Tropical Medicine & Hygiene.* **46:** 307-313.

Hogh, B., Thompson, R., Zakiuddin, I. S., Boudin, C., & Borre, M. B. (1993). Glutamate rich *Plasmodium falciparum* antigen (GLURP). *Parassitologia.* **35:** S47-S50.

Holder, A. A. (1988). The precursor to major merozoite surface antigens: structure and role in immunity. *Progress in Allergy* . **41:** 72-97.

Holder, A. A. & Freeman, R. R. (1982). Biosynthesis and processing of a *Plasmodium falciparum* schizont antigen recognized by immune serum and a monoclonal antibody. *Journal of Experimental Medicine.* **156:** 1528-1538.

Holder, A. A., Freeman, R. R., & Nicholls, S. C. (1988b). Immunization against *Plasmodium falciparum* with recombinant polypeptides produced in *Escherichia coli*. *Parasite Immunology.* **10:** 607-617.

Holder, A. A., Lockyer, M. J., & Hardy, G. W. (1988a). A hybrid gene to express protein epitopes from both sporozoite and merozoite surface antigens of *Plasmodium falciparum*. *Parasitology*. **97:** 373-382.

Holder, A. A., Lockyer, M. J., Odink, K. G., Sandhu, J. S., Riveros-Moreno, V., Nicholls, S. C. *et al.* (1985). Primary structure of the precursor to the three major surface antigens of *Plasmodium falciparum* merozoites. *Nature*. **317:** 270-273.

Howard, R. F., Ardeshir, F., & Reese, R. T. (1986). Conservation and antigenicity of N-terminal sequences of GP185 from different *Plasmodium falciparum* isolates. *Gene*. **46:** 197-205.

Hudson, R. R., Saez, G. A., & Ayala, F. J. (1997). DNA variation at the *Sod* locus of *Drosophila melanogaster*: An unfolding story of natural selection. *Proceedings of the National Academy of sciences*. **94:** 7725-7729.

Hudson, R. R. (1994). How can the low levels of DNA sequence variation in regions of the drosophila genome with low recombination rates be explained? *Proceedings of the National Academy of sciences*. **91:** 6815-6818.

Hughes, A. L. (1991). Circumsporozoite protein genes of malaria parasites (*Plasmodium spp.*): evidence for positive selection on immunogenic regions. *Genetics*. **127:** 345-353.

Hughes, A. L. (1992). Positive selection and interallelic recombination at the merozoite surface antigen-1 (MSA-1) locus of *Plasmodium falciparum*. *Molecular Biology and Evolution* . **9:** 381-393.

Hughes, A. L. & Verra, F. (1998). Ancient Polymorphism and the Hypothesis of a Recent Bottleneck in the Malaria Parasite *Plasmodium falciparum*. *Genetics*. **150:** 511-513.

Hughes, M. K. & Hughes, A. L. (1995). Natural selection on *Plasmodium falciparum* surface proteins. *Molecular and Biochemical Parasitology*. **71:** 99-113.

Hviid, L. 1998. Clinical disease, immunity and protection against *Plasmodium falciparum* malaria in populations living in endemic areas. *Expert Reviews in Molecular Medicine*. [http://www-ermm.cbcu.cam.ac.uk]

Hyde, J. E. (1989). Point mutations and pyrimethamine resistance in *Plasmodium falciparum*. *Parasitol Today*. **5:** 252-255.

Hyde, J. E. (1990). The dihydrofolate reductase-thymidylate synthetase gene in the drug resistance of malaria parasites. *Pharmacology & Therapeutics*. **48:** 45-59.

Hyde, J. E. (2002). Mechanisms of resistance of *Plasmodium falciparum* to antimalarial drugs. *Microbes and Infection*. **4:** 165-174.

Hyde, J. E. (2005). Exploring the folate pathway in *Plasmodium falciparum*. *Acta Tropica*. **94:** 191-206.

Ijumba, J. N., Mwangi, R. W., & Beier, J. C. (1990). Malaria transmission potential of *Anopheles* mosquitoes in the Mwea-Tebere irrigation scheme, Kenya. *Medical & Veterinary Entomology*. **4:** 425-432.

Inselburg, J. W., Bzik, D. J., & Horii, T. (1987). Pyrimethamine resistant *Plasmodium falciparum*: overproduction of dihydrofolate reductase by a gene duplication. *Molecular and Biochemical Parasitology*. **26:** 121-134.

Janse, C. J., van der Klooster, P. F., van der Kaay, H. J., van der, P. M., & Overdulve, J. P. (1986). Rapid repeated DNA replication during microgametogenesis and DNA synthesis in young zygotes of *Plasmodium berghei*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **80:** 154-157.

Jeffreys, A. J., Neumann, R., & Wilson, V. (1990). Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell*. **60:** 473-485.

Jeffreys, A. J., Royle, N. J., Wilson, V., & Wong, Z. (1988b). Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature*. **332:** 278-281.

Jeffreys, A. J., Wilson, V., Neumann, R., & Keyte, J. (1988a). Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Research*. **16:** 10953-10971.

Jeffreys, A. J., Wilson, V., & Thein, S. L. (1985). Individual-specific 'fingerprints' of human DNA. *Nature*. **316:** 76-79.

Jelinek, T., Ronn, A. M., Lemnge, M. M., Curtis, J., Mhina, J., Duraisingh, M. T. *et al.* (1998). Polymorphisms in the dihydrofolate reductase (DHFR) and dihydropteroate synthetase (DHPS) genes of *Plasmodium falciparum* and in vivo resistance to sulphadoxine/pyrimethamine in isolates from Tanzania. *Tropical Medicine and International Health*. **3:** 605-609.

Johnson, D. J., Fidock, D. A., Mungthin, M., Lakshmanan, V., Sidhu, A. B. S., Bray, P. G. *et al.* (2004). Evidence for a Central Role for PfCRT in Conferring *Plasmodium*

falciparum Resistance to Diverse Antimalarial Agents. *Molecular Cell.* **15:** 867-877.

Jongwutiwes, S., Tanabe, K., & Kanbara, H. (1993). Sequence conservation in the C-terminal part of the precursor to the major merozoite surface proteins (MSP1) of Plasmodium falciparum from field isolates. *Molecular and Biochemical Parasitology.* **59:** 95-100.

Jordan, W. C., Verspoor, E., & Youngson, A. F. (1997). The effect of natural selection on estimates of genetic divergence among populations of the atlantic salmon. *Journal of Fish Biology.* **51:** 546-560.

Joseph, J. T., Aldritt, S. M., Unnasch, T., Puijalon, O., & Wirth, D. F. (1989). Characterization of a conserved extrachromosomal element isolated from the avian malarial parasite *Plasmodium gallinaceum. Molecular and Cellular Biology.* **9:** 3621-3629.

Joy, D. A., Feng, X., Mu, J., Furuya, T., Chotivanich, K., Krettli, A. U. *et al.* (2003). Early origin and recent expansion of *Plasmodium falciparum. Science.* **300:** 318-321.

Juliano, R. L. & Ling, V. (1976). A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochimia Biosphysica Acta.* **455:** 152-162.

Kan, S. C. & Siddiqui, W. A. (1979). Comparative studies on dihydrofolate reductases from *Plasmodium falciparum* and *Aotus trivirgatus. Journal of Protozoology.* **26:** 660-664.

Kaneko, O., Jongwutiwes, S., Kimura, M., Kanbara, H., Ishii, A., & Tanabe, K. (1996). *Plasmodium falciparum*:Variation in Block 4 of the Precursor to the Major Merozoite Surface Proteins in Natural Populations. *Experimental Parasitology*. **84:** 92-95.

Kapitanovic, S., Cacev, T., Spaventi, R., & Pavelic, K. (2001). Submerged gel electrophoresis on Spreadex gels-a new method for APC gene mutation detection. *Journal of Molecular Medicine*. **79:** 333-337.

Kaplan, N. L., Hudson, R. R., & Langley, C. H. (1989). The ``Hitchhiking Effect" Revisited. *Genetics*. **123:** 887-899.

Kayser, M., Roewer, L., Hedman, M., Henke, L., Henke, J., Brauer, S. *et al.* (2000). Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *American Journal of Human Genetics*. **66:** 1580-1588.

Kemp, D. J., Coppel, R. L., Stahl, H. D., Bianco, A. E., Corcoran, L. M., McIntyre, P. *et al.* (1986). The Wellcome Trust lecture. Genes for antigens of *Plasmodium falciparum*. *Parasitology*. **92:** S83-108.

Kemp, D. J., Thompson, J. K., Walliker, D., & Corcoran, L. M. (1987). Molecular karyotype of *Plasmodium falciparum*: conserved linkage groups and expendable histidine-rich protein genes. *Proceedings of the National Academy of sciences*. **84:** 7672-7676.

Kerr, P. J., Ranford-Cartwright, L. C., & Walliker, D. (1994). Proof of intragenic recombination in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **66:** 241-248.

Kicska, G. A., Ting, L. M., Schramm, V. L., & Kim, K. (2003). Effect of dietary p-aminobenzoic acid on murine *Plasmodium yoelii* infection. *Journal of Infectious Diseases*. **188:** 1776-1781.

Kim, Y. & Stephan, W. (2002). Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics*. **160:** 765-777.

Kimura, M. (1977). Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature*. **267:** 275-276.

Kimura, M. & Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics*. **49:** 725-738.

Kirk, K., Horner, H. A., & Kirk, J. (1996). Glucose uptake in *Plasmodium falciparum*-infected erythrocytes is an equilibriative not an active process. *Molecular and Biochemical Parasitology*. **82:** 195-205.

Knowles, G., Sanderson, A., & Walliker, D. (1981). *Plasmodium yoelii*: genetic analysis of crosses between two rodent malaria subspecies. *Experimental Parasitology*. **52:** 243-247.

Kohler, S., Delwiche, C. F., Denny, P. W., Tilney, L. G., Webster, P., & WILSON, R. J. M. (1997). A plastid of probable green algal origin in apicomplexan parasites. *Science*. **275:** 1485-1489.

Kohn, M. H., Pelz, H. J., & Wayne, R. K. (2000). Natural selection mapping of the warfarin-resistance gene. *Proceedings of the National Academy of sciences*. **97:** 7911-7915.

Konate, L., Zwetyenga, J., Rogier, C., Bischoff, E., Fontenille, D., Tall, A. *et al.* (1999). Variation of *Plasmodium falciparum* msp1 block 2 and msp2 allele prevalence and of infection complexity in two neighbouring Senegalese villages with different transmission conditions. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **93:** S21-S28.

Kreitman, M. (1996). The neutral theory is dead. Long live the neutral theory. *Bioessays*. **18:** 678-683.

Krogstad, D. J., Gluzman, I. Y., Herwaldt, B. L., Schlesinger, P. H., & Wellems, T. E. (1992). Energy dependence of chloroquine accumulation and chloroquine efflux in *Plasmodium falciparum*. *Biochemical Pharmacology*. **43:** 57-62.

Krogstad, D. J., Gluzman, I. Y., Kyle, D. E., Oduola, A. M. J., Martin, S. K., Milhous, W. K. *et al.* (1987). Efflux of chloroquine from *Plasmodium falciparum*: mechanism of chloroquine resistance. *Science*. **238:** 1283-1285.

Krogstad, D. J. & Schlesinger, P. H. (1987). The basis of antimalarial action: non-weak base effects of chloroquine on acid vesicle pH. *American Journal Of Tropical Medicine & Hygiene*. **36:** 213-220.

Krungkrai, J. (1991). Malarial dihydroorotate dehydrogenase mediates superoxide radical production. *Biochemistry International*. **24:** 833-839.

Krungkrai, J. (1993). Dihydroorotase and dihydroorotate dehydrogenase as a target for antimalarial drugs. *Drugs of the Future*. **18:** 441-450.

Kublin, J. G., Witzig, R. S., Shankar, A. H., Zurita, J. Q., Gilman, R. H., Guarda, J. A. *et al.* (1998). Molecular assays for surveillance of antifolate-resistant malaria. *Lancet* . **351:** 1629-1630.

Kyes, S., Harding, R., Black, G., Craig, A., Peshu, N., Newbold, C. *et al.* (1997). Limited spatial clustering of individual *Plasmodium falciparum* alleles in field isolates from coastal Kenya. *American Journal Of Tropical Medicine & Hygiene*. **57:** 205-215.

Landgraf, B., Kollaritsch, H., Wiedermann, G., & Wernsdorfer, W. H. (1994). *Plasmodium falciparum*: susceptibility in vitro and in vivo to chloroquine and sulfadoxine-pyrimethamine in Ghanaian schoolchildren. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **88:** 440-442.

Leclerc, M. C., Durand, P., de Meeus, T., Robert, V., & Renaud, F. (2002). Genetic diversity and population structure of *Plasmodium falciparum* isolates from Dakar, Senegal, investigated from microsatellite and antigen determinant loci. *Microbes and Infection* . **4:** 685-692.

Levinson, G. & Gutman, G. A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*. **4:** 203-221.

Lim, P., Chy, S., Ariey, F., Incardona, S., Chim, P., Sem, R. *et al.* (2003). *pfcrt* Polymorphism and Chloroquine Resistance in *Plasmodium falciparum* Strains Isolated in Cambodia. *Antimicrobial Agents and Chemotherapy*. **47:** 87-94.

Lockyer, M. J., Marsh, K., & Newbold, C. I. (1989). Wild isolates of *Plasmodium falciparum* show extensive polymorphism in T cell epitopes of the circumsporozoite protein. *Molecular and Biochemical Parasitology*. **37:** 275-280.

Louis, E. J., Naumova, E. S., Lee, A., Naumov, G., & Haber, J. E. (1994). The chromosome end in yeast: its mosaic nature and influence on recombinational dynamics. *Genetics*. **136:** 789-802.

Luikart, G. & England, P. R. (1999). Statistical analysis of microsatellite DNA data. *Trends in Ecology & Evolution*. **14:** 253-256.

Lukyanov, K., Diatchenko, L., Chenchik, A., Nanisetti, A., Siebert, P., Usman, N. *et al.* (1997). Construction of cDNA libraries from small amounts of RNA using the suppression PCR effect. *Biochemical and Biophysical Research Communications*. **230:** 285-288.

Mackinnon, M. & Hastings, I. M. (1998). The evolution of multiple drug resistance in malaria parasites. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **92:** 188-195.

Mahtani, M. M. & Willard, H. F. (1993). A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: implications for mechanisms of mutation at short tandem repeat loci. *Human Molecular Genetics*. **2:** 431-437.

Maier, J. & Riley, E. M. (1942). Inhibition of antimalarial action of sulfonamides by *p*-aminobenzoic acid. *Proceedings of the Society of Experimental Biology and Medicine*. **50:** 152-154.

Maitland, K., Kyes, S., Williams, T. N., & Newbold, C. I. (2000). Genetic restriction of *Plasmodium falciparum* in an area of stable transmission: an example of island evolution? *Parasitology*. **120:** 335-343.

Martin, R. E. & Kirk, K. (2004). The malaria parasite's chloroquine resistance transporter is a member of the drug/metabolite transporter superfamily. *Molecular Biology and Evolution*. **21:** 1938-1949.

Martin, S. K., Oduola, A. M. J., & Milhous, W. K. (1987). Reversal of chloroquine resistance in *Plasmodium falciparum* by verapamil. *Science*. **235:** 899-901.

Maynard Smith, J. & Haigh, J. (1974). The hitchhiking effect of a favourable gene. *Genetic Research*. **23:** 23-35.

Maynard Smith, J., Smith, N. H., O'Rourke, M., & Spratt, B. G. (1993). How clonal are bacteria? *Proceedings of the National Academy of sciences*. **90:** 4384-4388.

Mbogo, C. N., Mwangangi, J. M., Nzovu, J. S. E., Gu, W. E., Yan, G. U., Gunter, J. T. *et al.* (2003). Spatial and temporal heterogeneity of Anopheles mosquitoes and Plasmodium falciparum transmission along the Kenyan coast. *American Journal of Tropical Medicine and Hygiene*. **68:** 734-742.

Mbogo, C. N., Snow, R. W., Khamala, C. P., Kabiru, E. W., Ouma, J. H., Githure, J. I. *et al.* (1995). Relationships between Plasmodium falciparum transmission by vector populations and the incidence of severe disease at nine sites on the Kenyan coast. *American Journal Of Tropical Medicine & Hygiene*. **52:** 201-206.

McBride, J. S. & Heidrich, H. G. (1987). Fragments of the polymorphic Mr 185,000 glycoprotein from the surface of isolated *Plasmodium falciparum* merozoites form an antigenic complex. *Molecular and Biochemical Parasitology*. **23:** 71-84.

McBride, J. S., Newbold, C. I., & Anand, R. (1985). Polymorphism of a high molecular weight schizont antigen of the human malaria parasite *Plasmodium falciparum*. *Journal of Experimental Medicine*. **161:** 160-180.

McCullough, J. L. & Maren, T. H. (1974). Dihydropteroate synthetase from *Plasmodium berghei*: isolation, properties, and inhibition by dapsone and sulfadiazine. *Molecular Pharmacology*. **10:** 140-145.

Mcdonald, J. H.(1994). Detecting natural selection by comparing geographical variation in protein and DNA sequences.: In *Non Neutral Evolution: Theories and Molecular Data*. (Ed.Golding, B.) 88-100. Chapman and Hall, New York.

McFadden, G. I. & Roos, D. S. (1999). Apicomplexan plastids as drug targets. *Trends in Microbiology*. **7:** 328-333.

McIntosh, M. T., Srivastava, R., & Vaidya, A. B. (1998). Divergent evolutionary constraints on mitochondrial and nuclear genomes of malaria parasites. *Molecular and Biochemical Parasitology*. **95:** 69-80.

Mehlotra, R. K., Fujioka, H., Roepe, P. D., Janneh, O., Ursos, L. M. B., Jacobs-Lorena, V. *et al.* (2001). Evolution of a unique *Plasmodium falciparum* chloroquine-resistance phenotype in association with pfcrt polymorphism in Papua New Guinea and South America. *Proceedings of the National Academy of Sciences of the United States of America*. **98:** 12689-12694.

Mendez, F., Munoz, A., Carrasquilla, G., Jurado, D., Arevalo-Herrera, M., Cortese, J. F. *et al.* (2002). Determinants of Treatment Response to Sulfadoxine-Pyrimethamine and Subsequent Transmission Potential in *falciparum* Malaria. *American Journal of Epidemiology*. **156:** 230-238.

Mercereau-Puijalon, O. (1996). Revisiting host/parasite interactions: molecular analysis of parasites collected during longitudinal and cross-sectional surveys in humans. *Parasite Immunology*. **18:** 173-180.

Mharakurwa, S. (2004). *Plasmodium falciparum* transmission rate and selection for drug resistance: a vexed association or a key to successful control? *International Journal for Parasitology*. **34:** 1483-1487.

Miettinen-Bauman, A., Strych, W., McBride, J. S., & Heidrich, H. G. (1988). A 46000 kDa *Plasmodium falciparum* merozoite surface antigen not related to the 185000-195000 kDa schizont precursor molecule: isolation and characterisation. *Parasitology Research*. **74:** 317-323.

Milhous, W. K., Weatherly, N. F., Bowdre, J. H., & Desjardins, R. E. (1985). In vitro activities of and mechanisms of resistance to antifol antimalarial drugs. *Antimicrobial Agents and Chemotherapy*. **27:** 525-530.

Miller, L. H., Roberts, T., Shahabuddin, M., & McCutchan, T. F. (1993). Analysis of sequence diversity in the Plasmodium falciparum merozoite surface protein-1 (MSP-1). *Molecular and Biochemical Parasitology*. **59:** 1-14.

Miller, M. P. (1997). The R X C Package: A computer program for the analysis of contingency tables. *http://www.marksgeneticsoftware.net*.

Molineaux, L. & Gramiccia, G. 1980. *The Garki Project*. World Health Organization. Geneva.

Molyneux, D. H., Floyd, K., Barnish, G., & Fevre, E. M. (1999). Transmission Control and Drug Resistance in Malaria: A Crucial Interaction. *Parasitology Today*. **15:** 238-240.

Moore, D. V. & Lanier, J. E. (1961). Observations on two *Plasmodium falciparum* infections with abnormal response to chloroquine. *American Journal of Tropical Medicine and Hygiene*. **10:** 5-9.

Morral, N., Nunes, V., Casals, T., & Estivill, X. (1991). CA/GT microsatellite alleles within the cystic fibrosis transmembrane conductance regulator (CFTR) gene are not generated by unequal crossingover. *Genomics*. **10:** 692-698.

Mu, J., Duan, J., Makova, K. D., Joy, D. A., Huynh, C. Q., Branch, O. H. *et al.* (2002). Chromosome-wide SNPs reveal an ancient origin for *Plasmodium falciparum*. *Nature*. **418:** 323-324.

Mueller, J. C. & Andreoli, C. (2004). Plotting haplotype-specific linkage disequilibrium patterns by extended haplotype homozygosity. *Bioinformatics*. **20:** 786-787.

Mungthin, M., Bray, P. G., & Ward, S. A. (1999). Phenotypic and genotypic characteristics of recently adapted isolates of *Plasmodium falciparum* from Thailand. *American Journal of Tropical Medicine and Hygiene*. **60:** 469-474.

Mwangi, J. M., Omar, S. A., & Ranford-Cartwright, L. C. (2006). Comparison of microsatellite and antigen-coding loci for differentiating recrudescing *Plasmodium falciparum* infections from reinfections in Kenya. *International Journal for Parasitology*. **36:** 329-336.

Nadler, S. A. (1995). Microevolution and the genetic structure of parasite populations. *Parasitology*. **81:** 395-403.

Nagesha, H. S., Casey, G. J., Rieckmann, K., Fryauff, D. J., Laksana, B. S., Reeder, J. C. *et al.* (2003). New haplotypes of the *Plasmodium falciparum* chloroquine resistance transporter (*pfcrt*) gene among chloroquine-resistant parasite isolates. *American Journal Of Tropical Medicine & Hygiene*. **68:** 398-402.

Nair, S., Williams, J. T., Brockman, A., Paiphun, L., Mayxay, M., Newton, P. N. *et al.* (2003). A selective sweep driven by pyrimethamine treatment in Southeast Asian malaria parasites. *Molecular Biology and Evolution*. **20:** 1526-1536.

Nakamura, Y., Gojobori, T., & Ikemura, T. (1998). Codon usage tabulated from the international DNA sequence databases. *Nucleic Acids Research*. **26:** 334.

Nash, D., Nair, S., Mayxay, M., Newton, P. N., Guthmann, J. P., Nosten, F. *et al.* (2005). Selection strength and hitchhiking around two anti-malarial resistance genes. *Proceedings: Biological Sciences*. **272:** 1153-1161.

Nateghpour, M., Ward, S. A., & Howells, R. E. (1993). Development of halofantrine resistance and determination of cross-resistance patterns in *Plasmodium falciparum*. *Microbial Agents Chemotherapy*. **37:** 2337-2343.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of sciences*. **70:** 3321-3323.

Nei, M. (1987b). Genetic distance and molecular phylogeny. In: *Population Genetics and Fishery Management*. Ryman, N. and Utter, F. University of Washington Press, Seattle.

Nei, M. (1987a). Molecular Evolutionary Genetics. Columbia University Press, New York.

Nei, M. & Feldman, M. W. (1972). Identity of genes by descent within and between populations under mutation and migration pressures. *Theoretical Population Biology*. **3:** 460-465.

Nguyen-Dinh, P., Spencer, H. C., Chemangey-Masaba, S., & Churchill, F. C. (1982). Susceptibility of *Plasmodium falciparum* to pyrimethamine and sulfadoxine/pyrimethamine in Kisumu, Kenya. *Lancet*. **1:** 823-825.

Nirmalan, N., Sims, P. F. G., & Hyde, J. E. (2004). Translational up-regulation of antifolate drug targets in the human malaria parasite *Plasmodium falciparum* upon challenge with inhibitors. *Molecular and Biochemical Parasitology*. **136:** 63-70.

Ntoumi, F., Contamin, H., Rogier, C., Bonnefoy, S., Trape, J. F., & Mercereau-Puijalon, O. (1995). Age-dependent carriage of multiple *Plasmodium falciparum* merozoite surface antigen-2 alleles in asymptomatic malaria infections. *American Journal Of Tropical Medicine & Hygiene*. **52:** 81-88.

Nzila, A., Mberu, E., Sulo, J., Dayo, H., Winstanley, P. A., Sibley, C. H. *et al.* (2000a). Towards an understanding of the mechanism of pyrimethamine-sulfadoxine resistance in *Plasmodium falciparum*: genotyping of dihydrofolate reductase and dihydropteroate synthase of Kenyan parasites. *Antimicrobial Agents and Chemotherapy*. **44:** 991-996.

Nzila, A., Nduati, E., Mberu, E., Hopkins, S. C., Monks, S. A., Winstanley, P. A. *et al.* (2000b). Molecular evidence of greater selective pressure for drug resistance exerted by the long-acting antifolate Pyrimethamine/Sulfadoxine compared with

the shorter-acting chlorproguanil/dapsone on Kenyan *Plasmodium falciparum*. *Journal of Infectious Diseases*. **181:** 2023-2028.

Ohta, T. (1996). The current significance and standing of neutral and neutral theories. *Bioessays*. **18:** 673-677.

Omar, S. A., Adagu, I. S., & Warhurst, D. C. (2001a). Can pretreatment screening for dhps and dhfr point mutations in *Plasmodium falciparum* infections be used to predict sulfadoxine-pyrimethamine treatment failure? *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **95:** 315-319.

Omar, S. A., Bakari, A., Owiti, A., Adagu, I. S., & Warhurst, D. C. (2001b). Co-trimoxazole compared with sulfadoxine-pyrimethamine in the treatment of uncomplicated malaria in Kenyan children. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **95:** 657-660.

Orr, H. T., Chung, M. Y., & Banfi, S. (1993). Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nature Genetics*. **4:** 221.

Ota, P. (1993). Dispan: Genetic Distance and Phylogenetic Analysis. Pennsylvania State University, Pennsylvania.

Paget-McNicol, S. & Saul, A. (2001). Mutation rates in the dihydrofolate reductase gene of *Plasmodium falciparum*. *Parasitology*. **122:** 497-505.

Paul, R. E., Hackford, I., Brockman, A., Muller-Graf, R., Price, R., Luxemburger, C. *et al.* (1998). Transmission intensity and *Plasmodium falciparum* diversity on the northwestern border of Thailand. *American Journal Of Tropical Medicine & Hygiene*. **58:** 195-203.

Paul, R. E., Packer, M. J., Walmsley, M., Lagog, M., Ranford-Cartwright, L. C., Paru, R. *et al.* (1995). Mating patterns in malaria parasite populations of Papua New Guinea. *Science*. **269:** 1709-1711.

Payne, D. (1988). Did medicated salt hasten the spread of chloroquine resistance in *P. falciparum*? *Parasitology Today*. **4:** 112-115.

Payseur, B. A., Cutter, A. D., & Nachman, M. W. (2002). Searching for Evidence of Positive Selection in the Human Genome Using Patterns of Microsatellite Variability. *Molecular Biology and Evolution*. **19:** 1143-1153.

Pearce, R., Malisa, A., Kachur, S. P., Barnes, K., Sharp, B., & Roper, C. (2005). Reduced variation around drug-resistant *dhfr* Alleles in african *Plasmodium falciparum*. *Molecular Biology and Evolution*. **22:** 1834-1844.

Peters, W. (1997). Chemotherapy and Drug resistance in Malaria. Academic Press, London.

Peters, W. (1985). The problem of drug resistance in malaria. *Parasitology*. **90:** 705-715.

Peterson, D. S., Di Santi, S. M., Povoa, M., Calvosa, V. S., do Rosario, V. E., & Wellems, T. E. (1991). Prevalence of the dihydrofolate reductase Asn-108 mutation as the basis for pyrimethamine-resistant falciparum malaria in the Brazilian Amazon. *American Journal Of Tropical Medicine & Hygiene*. **45:** 492-497.

Peterson, D. S., Milhous, W. K., & Wellems, T. E. (1990). Molecular basis of differential resistance to cycloguanil and pyrimethamine in *Plasmodium falciparum* malaria. *Proceedings of the National Academy of sciences*. **87:** 3018-3022.

Peterson, D. S., Walliker, D., & Wellems, T. E. (1988a). Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in *falciparum* malaria. *Proceedings of the National Academy of sciences*. **85:** 9114-9118.

Peterson, M. G., Coppel, R. L., McIntyre, P., Langford, C. J., Woodrow, G., Brown, G. V. *et al.* (1988b). Variation in the precursor to the major merozoite surface antigens of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **27:** 291-301.

Phillips, R. S. (2001). Current status of malaria and potential for control. *Clinical Microbiology Reviews*. **14:** 208-226.

Pinotti, M. (1953). Chemoprophylaxis of malaria by the association of an antimalarial drug to the sodium chloride used daily in the preparation of meals. In: *V^e Congres internalionaux de Medecine tropicale et du Pallidisme*.**2:** 248-258.

Plowe, C. V., Cortese, J. F., Djimde, A. A., Nwanyanwu, O. C., Watkins, W. M., Winstanley, P. A. *et al.* (1997). Mutations in *Plasmodium falciparum* dihydrofolate reductase and dihydropteroate synthase and epidemiologic patterns of pyrimethamine-sulfadoxine use and resistance. *Journal of Infectious Diseases*. **176:** 1590-1596.

Plowe, C. V., Djimde, A. A., Bouare, M., Doumbo, O. K., & Wellems, T. E. (1995). Pyrimethamine and proguanil resistance-conferring mutations in *Plasmodium falciparum* dihydrofolate reductase: polymerase chain reaction methods for surveillance in Africa. *American Journal Of Tropical Medicine & Hygiene*. **52:** 565-568.

Plowe, C. V., Djimde, A. A., Wellems, T. E., Diop, S., Kouriba, B., & Doumbo, O. K. (1996). Community pyrimethamine-sulfadoxine use and prevalence of resistant *Plasmodium falciparum* genotypes in Mali: a model for deterring resistance. *American Journal Of Tropical Medicine & Hygiene*. **55:** 467-471.

Plowe, C. V., Kublin, J. G., & Doumbo, O. K. (1998). *P. falciparum* dihydrofolate reductase and dihydropteroate synthase mutations: epidemiology and role in clinical resistance to antifolates. *Drug Resistance Updates*. **1:** 389-396.

Pologe, L. G. & Ravetch, J. V. (1988). Large deletions result from breakage and healing of P. falciparum chromosomes. *Cell*. **55:** 869-874.

Ponticelli, A. S., Sena, E. P., & Smith, G. R. (1988). Genetic and physical analysis of the M26 recombination hotspot of *Schizosaccharomyces pombe*. *Genetics*. **119:** 491-497.

Povoa, M. M., Adagu, I. S., Oliveira, S. G., Machado, R. L. D., Miles, M. A., & Warhurst, D. C. (1998). *Pfmdr1* Asn1042Asp and Asp1246Try polymorphisms thought to be associated with chloroquine resistance, are present in chloroquine-resistant and sensitive Brazilian field isolates of *Plasmodium falciparum*. *Exp Parasitol*. **88:** 64-68.

Preiser, P. R., Wilson, R. J., Moore, P. W., McCready, S., Hajibagheri, M. A., Blight, K. J. *et al.* (1996). Recombination associated with replication of malarial mitochondrial DNA. *EMBP Journal*. **15:** 684 -693.

Price, R. N., Cassar, C., Brockman, A., Duraisingh, M. T., van Vugt, M., White, N. J. *et al.* (1999). The *pfmdr1* gene is associated with a multidrug-resistant phenotype in

Plasmodium falciparum from the western border of Thailand. *Antimicrobial Agents and Chemotherapy*. **43:** 2943-2949.

Primmer, C. R., Saino, N., Moller, A. P., & Ellegren, H. (1998). Unravelling the processes of microsatellite evolution through analysis of germ line mutations in barn swallows *Hirundo rustica*. *Molecular Biology and Evolution*. **15:** 1047-1054.

Pringle, G. (1965). A count of the sporozoites in an ooecyst of *Plasmodium falciparum*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **59:** 289-290.

Ramasamy, R. (1987). Studies on glycoproteins in the human malaria parasite *Plasmodium falciparum*. Identification of a myristilated 45kDa merozoite membrane glycoprotein. *Immunology and Cell Biology*. **65:** 419-424.

Rand, D. M., Dorfsman, M., & Kann, L. M. (1994). Neutral and non-neutral evolution of Drosophila mitochondrial DNA. *Genetics* . **138:** 741-756.

Ranford-Cartwright, L. C., Johnston, K. L., Abdel-Muhsin, A. M., Khan, B. K., & Babiker, H. A. (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.

Ranford-Cartwright, L. C., Balfe, P., Carter, R., & Walliker, D. (1993). Frequency of cross-fertilization in the human malaria parasite *Plasmodium falciparum*. *Parasitology*. **107:** 11-18.

Ranford-Cartwright, L. C., Balfe, P., Carter, R., & Walliker, D. (1991). Direct sequencing of enzymatically amplified DNA of alleles of the merozoite surface

antigen MSA-1 gene from the malaria parasite *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **46:** 185-187.

Ranford-Cartwright, L. C., Taylor, J., Umasunthar, T., Taylor, L. H., Babiker, H. A., Lell, B. *et al.* (1997). Molecular analysis of recrudescent parasites in a *Plasmodium falciparum* drug efficacy trial in Gabon. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **91:** 719-724.

Ranford-Cartwright, L. C., Taylor, R. R., Asgari-Jirhandeh, N., Smith, D. B., Roberts, P. E., Robinson, V. J. *et al.* (1996). Differential antibody recognition of FC27-like Plasmodium falciparum merozoite surface protein MSP2 antigens which lack 12 amino acid repeats. *Parasite Immunology*. **18:** 411-420.

Rapuoda, B. (1995). Ecological and behavioural studies of mosquitoes in Mwea Tebere irrigation scheme, Kirinyaga District, Kenya with special reference to *Anopheles arabiensis* (Diptera: cullicidae). Department of Zoology, Kenyatta University.

Raymond, M. & Rousset, F. (1995). GENEPOP (version 3.4): Population Genetics software for exact tests and ecumenism. *Journal of Heredity*. **86:** 248-249.

Read, A. F., Narara, A., Keymer, A. E., & Day, K. P. (1992). Gametocyte sex ratios as indirect measures of outcrossing rates in malaria. *Parasitology*. **104:** 387-395.

Reeder, J. C., Rieckmann, K., Genton, B., Lorry, K., Wines, B., & Cowman, A. F. (1996). Point mutations in the dihydrofolate reductase and dihydropteroate synthetase genes and in vitro susceptibility to pyrimethamine and cycloguanil of *Plasmodium falciparum* isolates from Papua New Guinea. *American Journal Of Tropical Medicine & Hygiene*. **55:** 209-213.

Rich, S. M. & Ayala, F. J. (1998). The recent origin of allelic variation in antigenic determinants of *Plasmodium falciparum* . *Genetics*. **150:** 515-517.

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

Rich, S. M., Hudson, R. R., & Ayala, F. J. (1997). *Plasmodium falciparum* antigenic diversity: Evidence of clonal population structure. *Proceedings of the National Academy of sciences*. **94:** 13040-13045.

Rich, S. M., Licht, M. C., Hudson, R. R., & Ayala, F. J. (1998). Malaria's Eve: Evidence of a recent population bottleneck throughout the world populations of *Plasmodium falciparum*. *Proceedings of the Association of American Physicians*. **95:** 4425-4430.

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

Riffkin, C. D., Chung, R., & Wall, D. M. (1996). Modulation of the function of human MDR1 P-glycoprotein by the antimalarial drug mefloquine. *Biochem Pharmacol*. **52:** 1545-1552.

Ritchie, G. Y., Mungthin, M., Green, J. E., Bray, P. G., Hawley, S. R., & Ward, S. A. (1996). *In vitro* selection of halofantrine resistance in *Plasmodium falciparum* is not associated with increased expression of Pgh1. *Molecular and Biochemical Parasitology*. **85:** 35-46.

Roepe, P. D., Wei, L. Y., Hoffman, M. M., & Fritz, F. (1996). Altered drug translocation mediated by the MDR protein: direct, indirect, or both? *Journal of Bioenergetics and Biomembranes*. **28:** 541-555.

Rollo, I. M. (1955). The mode of action of sulphonamides, proguanil and pyrimethamine on *Plasmodium gallinaceum*. *British Journal of Pharmacology*. **10:** 208-214.

Roper, C., Elhassan, I. M., Hviid, L., Giha, H., Richardson, W., Babiker, H. A. *et al.* (1996). Detection of very low level *Plasmodium falciparum* infections using the nested polymerase chain reaction and a reassessment of the epidemiology of unstable malaria in Sudan. *American Journal Of Tropical Medicine & Hygiene*. **54:** 325-331.

Roper, C., Pearce, R., Bredenkamp, B., Gumede, J., Drakeley, C., Mosha, F. *et al.* (2003). Antifolate antimalarial resistance in southeast Africa: a population-based analysis. *Lancet*. **361:** 1174-1181.

Roper, C., Richardson, W., Elhassan, I. M., Giha, H., Hviid, L., Satti, G. *et al.* (1998). Seasonal changes in the *Plasmodium falciparum* population in individuals and their relationship to clinical malaria: a longitudinal study in a Sudanese village. *Parasitology*. **116:** 501-510.

Roper, C., Pearce, R., Nair, S., Sharp, B., Nosten, F., & Anderson, T. J. C. (2004). Intercontinental spread of pyrimethamine-resistant malaria. *Science*. **305:** 1124-

Rosenberg, R. & Rungsiwongse, J. (1991). The number of sporozoites produced by individual malaria oocysts. *American Journal Of Tropical Medicine & Hygiene*. **45:** 574-577.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*. **145:** 1219-1228.

Rubinsztein, D. C., Amos, B., & Cooper, G. (1999). Microsatellite and trinucleotide-repeat evolution: evidence for mutational bias and different rates of evolution in different lineages. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. **354:** 1095-1099.

Rubio, J. & Cowman, A. F. (1994). *Plasmodium falciparum. the pfmdr 2* protein is not overexpressed in chloroquine-resistant isolates of the malaria parasite. *Exp Parasitol*. **79:** 137-147.

Rudenko, G., McCulloch, R., Dirks-Mulder, A., & Borst, P. (1996). Telomere exchange can be an important mechanism of variant surface glycoprotein gene switching in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*. **80:** 65-75.

Sabatti, C. & Risch, N. (2002). Homozygosity and linkage disequilibrium. *Genetics*. **160:** 1707-1719.

Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z. P., Richter, D. J., Schaffner, S. F. *et al.* (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*. **419:** 832-837.

Saliba, K. J., Allen, R. J., Zissis, S., Bray, P. G., Ward, S. A., & Kirk, K. (2003). Acidification of the malaria parasite's digestive vacuole by a H^+-ATPase and a H^+-pyrophosphatase. *Journal of Biological Chemistry*. **278:** 5605-5612.

Saliba, K. J., Folb, P. I., & Smith, P. J. (1998). Role for the *plasmodium falciparum* digestive vacuole in chloroquine resistance. *Biochemical Pharmacology*. **56:** 313-320.

Saliba, K. J. & Kirk, K. (1998). Clotrimazole inhibits the growth of *Plasmodium falciparum in vitro*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **92:** 666-667.

Sanchez, C. P., Stein, W., & Lanzer, M. (2003). Trans stimulation provides evidence for a drug efflux carrier as the mechanism of chloroquine resistance in *Plasmodium falciparum*. *Biochemistry*. **42:** 9383-9394.

Sanchez, C. P., Wunsch, S., & Lanzer, M. (1997). Identification of a chloroquine importer in *Plasmodium falciparum*. Differences in import kinetics are genetically linked with the chloroquine-resistant phenotype. *Journal of Biological Chemistry*. **272:** 2652-2658.

Scherf, A. & Mattei, D. (1992). Cloning and characterization of chromosome breakpoints of *Plasmodium falciparum*: breakage and new telomere formation occurs frequently and randomly in subtelomeric genes. *Nucleic Acids Research*. **20:** 1491-1496.

Scherf, A., Petersen, C., Carter, R., Alano, P., Nelson, R., Aikawa, M. *et al.* (1992). Characterization of a *Plasmodium falciparum* mutant that has deleted the majority of the gametocyte-specific Pf11-1 locus. *Memerias do Instituto Oswaldo Cruz*. **87:** S91-S94.

Schlotterer, C. (2003). Hitchhiking mapping--functional genomics from the population genetics perspective. *Trends in Genetics*. **19:** 32-38.

Schlotterer, C., Amos, B., & Tautz, D. (1991). Conservation of polymorphic simple sequence loci in cetacean species. *Nature*. **354:** 63-65.

Schlotterer, C. & Tautz, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*. **20:** 211-215.

Schmidt, K. F. (1995). Malaria research. Inbred parasites may spur resistance. *Science*. **269:** 1670.

Schols, L., Vieira-Saecker, A. M. M., & Schols, S. (1995). Trinucleotide expansion within the MJD1 gene presents clinically as spinocerebellar ataxia (SCA) and occurs most frequently in German SCA patients. *Human Molecular Genetics*. **4:** 1001-1005.

Schrag, S. J., Perrot, V., & Levin, B. R. (1997). Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. *Proceedings: Biological Sciences*. **264:** 1287-1291.

Sendagire, H., Kyabayinze, D., Swedberg, G., & Kironde, F. (2005). *Plasmodium falciparum*: higher incidence of molecular resistance markers for sulphadoxine than for pyrimethamine in Kasangati,Uganda. *Tropical Medicine and International Health*. **10:** 537-543.

Sherman, I. W. (1979). Biochemistry of *Plasmodium* (malarial parasites). *Microbiology Reviews*. **43:** 453-495.

Sherman, I. W. (1998). Malaria: Parasite Biology, Pathogenesis, and Protection. American Society for Microbiology, Washington.

Shi, Y. P., Alpers, M. P., Povoa, M. M., & Lal, A. A. (1992). Diversity in the immunodominant determinants of the circumsporozoite protein of *Plasmodium falciparum* parasites from malaria-endemic regions of Papua New Guinea and Brazil. *American Journal Of Tropical Medicine & Hygiene*. **47:** 844-851.

Sidhu, A. B. S., Verdier-Pinard, D., & Fidock, D. A. (2002). Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by *pfcrt* mutations. *Science*. **298:** 210-213.

Silvestrini F, Alano, P., & Williams JL (2000). Commitment to the production of male and female gametocytes in the human malaria parasite *Plasmodium falciparum*. *Parasitology*. **121:** 465-471.

Sinnis, P. & Wellems, T. E. (1988). Long-range restriction maps of Plasmodium falciparum chromosomes: crossingover and size variation among geographically distant isolates. *Genomics*. **3:** 287-295.

Sirawaraporn, W., Sathitkul, T., Sirawaraporn, R., Yuthavong, Y., & Santi, D. (1997). Antifolate-resistant mutants of *Plasmodium falciparum* dihydrofolate reductase. *Proceedings of the National Academy of sciences*. **94:** 1124-1129.

Sirawaraporn, W., Sirawaraporn, R., Cowman, A. F., Yuthavong, Y., & Santi, D. V. (1990). Heterologous expression of active thymidylate synthase-dihydrofolate reductase from *Plasmodium falciparum*. *Biochemistry*. **29:** 10779-10785.

Sirawaraporn, W. & Yuthavong, Y. (1984). Kinetic and molecular properties of dihydrofolate reductase from pyrimethamine-sensitive and pyrimethamine-resistant Plasmodium chabaudi. *Molecular and Biochemical Parasitology*. **10:** 355-367.

Sirawaraporn, W. & Yuthavong, Y. (1986). Potentiating effect of pyrimethamine and sulfadoxine against dihydrofolate reductase from pyrimethamine-sensitive and pyrimethamine-resistant *Plasmodium chabaudi*. *Antimicrobial Agents and Chemotherapy*. **29:** 899-905.

Smith, T., Beck, H. P., Kitua, A. Y., Mwankusye, S., Felger, I., Fraser-Hurt, N. *et al.* (1999b). Age dependence of the multiplicity of *Plasmodium falciparum* infections and of other malariological indices in an area of high endemicity. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **93 Suppl 1:** 15-20.

Smith, T., Felger, I., Kitua, A. Y., Tanner, M., & Beck, H. P. (1999a). Dynamics of multiple *Plasmodium falciparum* infections in infants in a highly endemic area of Tanzania. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **93 Suppl 1:** 35-39.

Smith, T. G., Lourenco, P., Carter, R., Walliker, D., & Ranford-Cartwright, L. C. (2000). Commitment to sexual differentiation in the human malaria parasite, *Plasmodium falciparum*. *Parasitology*. **121:** 127-133.

Smith, T. G., Walliker, D., & Ranford-Cartwright, L. C. (2002). Sexual differentiation and sex determination in the Apicomplexa. *Trends in Parasitology*. **18:** 315-323.

Smythe, J. A., Coppel, R. L., Brown, G. V., Ramasamy, R., Kemp, D. J., & Anders, R. F. (1988). Identification of two integral membrane proteins of *Plasmodium falciparum*. *Proceedings of the National Academy of sciences*. **85:** 5195-5199.

Smythe, J. A., Peterson, M. G., Coppel, R. L., Saul, A. J., Kemp, D. J., & Anders, R. F. (1990). Structural diversity in the 45-kilodalton merozoite surface antigen of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **39:** 227-234.

Snewin, V. A., England, S. M., Sims, P. F., & Hyde, J. E. (1989). Characterisation of the dihydrofolate reductase-thymidylate synthetase gene from human malaria parasites highly resistant to pyrimethamine. *Gene*. **76:** 41-52.

Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y., & Hay, S. I. (2005). The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature*. **434:** 214-217.

Soremekun, S., Maxwell, C., Zuwakuu, M., Chen, C., Michael, E., & Curtis, C. (2004). Measuring the efficacy of insecticide treated bednets: the use of DNA fingerprinting to increase the accuracy of personal protection estimates in Tanzania. *Tropical Medicine and International Health*. **9:** 664-672.

Souza, V., Nguyen, T. T., Hudson, R. R., Pinero, D., & Lenski, R. E. (1992). Hierarchical analysis of linkage disequilibrium in Rhizobium populations: evidence for sex? *Proceedings of the National Academy of sciences*. **89:** 8389-8393.

Stanley, H. A., Howard, R. F., & Reese, R. T. (1985). Recognition of a Mr 56K glycoprotein on the surface of Plasmodium falciparum merozoites by mouse monoclonal antibodies. *Journal of Immunology*. **134:** 3439-3444.

Su, X. z., Ferdig, M. T., Huang, Y., Huynh, C. Q., Liu, A., You, J. *et al.* (1999). A Genetic Map and Recombination Parameters of the Human Malaria Parasite *Plasmodium falciparum*. *Science*. **286:** 1351-1353.

Su, X. z., Kirkman, L. A., Fujioka, H., & 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. z. & Wellems, T. E. (1996). Toward a high-resolution *Plasmodium falciparum* linkage map: polymorphic markers from hundreds of simple sequence repeats. *Genomics*. **33:** 430-444.

Sunnucks, P. (2000). Efficient genetic markers for population biology. *Trends in Ecology & Evolution*. **15:** 199-203.

Symington, L. S., Brown, A., Oliver, S. G., Greenwell, P., & Petes, T. D. (1991). Genetic analysis of a meiotic recombination hotspot on chromosome III of *Saccharomyces cerevisiae*. *Genetics*. **128:** 717-727.

Szankasi, P., Heyer, W. D., Schuchert, P., & Kohli, J. (1988). DNA sequence analysis of the ade6 gene of *Schizosaccharomyces pombe*. Wild-type and mutant alleles including the recombination hot spot allele ade6-M26. *Journal of Molecular Biology*. **204:** 917-925.

Szarfman, A., Walliker, D., McBride, J. S., Lyon, J. A., Quakyi, I. A., & Carter, R. (1988). Allelic forms of gp195, a major blood-stage antigen of *Plasmodium falciparum*, are expressed in liver stages. *Journal of Experimental Medicine*. **167:** 231-236.

Taberlet, P., Camarra, J. J., Griffin, S., Uhres, E., Hanotte, O., Waits, L. P. *et al.* (1997). Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular Ecology*. **6:** 869-876.

Taberlet, P. & Fumagalli, L. (1996). Owl pellets as a source of DNA for genetic studies of small mammals. *Molecular Ecology*. **5:** 301-305.

Takahata, N. & Nei, M. (1990). Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics*. **124:** 967-978.

Tanabe, K., Murakami, K., & Doi, S. (1989). *Plasmodium falciparum*: dimorphism of the p190 alleles. *Experimental Parasitology*. **68:** 470-473.

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

Tanaka, M., Gu, H. M., Bzik, D. J., Li, W. B., & Inselburg, J. W. (1990b). Dihydrofolate reductase mutations and chromosomal changes associated with pyrimethamine resistance of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **39:** 127-134.

Tanaka, M., Gu, H. M., Bzik, D. J., Li, W. B., & Inselburg, J. W. (1990a). Mutant dihydrofolate reductase-thymidylate synthase genes in pyrimethamine-resistant *Plasmodium falciparum* with polymorphic chromosome duplications. *Molecular and Biochemical Parasitology*. **42:** 83-91.

Taylor, L. H., Mackinnon, M. J., & Read, A. F. (1998). Virulence of mixed-clone and single-clone infections of the rodent malaria *Plasmodium chabaudi*. *Evolution*. **52:** 583-591.

Taylor, P. & Mutambu, S. L. (1986). A review of the malaria situation in Zimbabwe with special reference to the period 1972-1981. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **80:** 12-19.

Thaithong, S., Ranford-Cartwright, L. C., Siripoon, N., Harnyuttanakorn, P., Kanchanakhan, N. S., Seugorn, A. *et al.* (2001). *Plasmodium falciparum*: Gene Mutations and Amplification of Dihydrofolate Reductase Genes in Parasites Grown in Vitro in Presence of Pyrimethamine. *Experimental Parasitology*. **98:** 59-70.

Theisen, M., Soe, S., Oeuvray, C., Thomas, A. W., Vuust, J., Danielsen, S. *et al.* (1998). The Glutamate-Rich Protein (GLURP) of *Plasmodium falciparum* Is a Target for Antibody-Dependent Monocyte-Mediated Inhibition of Parasite Growth *In Vitro*. *Infection and Immunity*. **66:** 11-17.

Thomas, A. W., Carr, D. A., Carter, J. M., & Lyon, J. A. (1990c). Sequence comparison of allelic forms of the *Plasmodium falciparum* merozoite surface antigen MSA2. *Molecular and Biochemical Parasitology*. **43:** 211-220.

Thomas, A. W., Waters, A. P., & Carr, D. (1990a). Analysis of variation in PF83, an erythrocytic merozoite vaccine candidate antigen of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **42:** 285-287.

Thomas, A. W., Carr, D. A., Carter, J. M., & Lyon, J. A. (1990b). Sequence comparison of allelic forms of the Plasmodium falciparum merozoite surface antigen MSA2. *Molecular and Biochemical Parasitology*. **43:** 211-220.

Tibayrenc, M., Kjellberg, F., & Ayala, F. J. (1990). A clonal theory of parasitic protozoa: the population structures of *Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas,* and *Trypanosoma* and their medical and taxonomical consequences. *Proceedings of the National Academy of sciences*. **87:** 2414-2418.

Tishkoff, S. A., Varkonyi, R., Cahinhinan, N., Abbes, S., Argyropoulos, G., Destro-Bisol, G. *et al.* (2001). Haplotype diversity and linkage disequilibrium at human

G6PD: recent origin of alleles that confer malarial resistance. *Science*. **293:** 455-462.

Trager, W. & Jensen, J. B. (1976). Human malaria parasites in continuous culture. *Science*. **193:** 673-675.

Triglia, T. & Cowman, A. F. (1994). Primary structure and expression of the dihydropteroate synthetase gene of *Plasmodium falciparum*. *Proceedings of the National Academy of sciences*. **91:** 7149-7153.

Triglia, T., Foote, S. J., Kemp, D. J., & Cowman, A. F. (1991). Amplification of the multidrug resistance gene *pfmdr1* in *Plasmodium falciparum* has arisen as multiple independent events. *Molecular and Cellular Biology*. **11:** 5244-5250.

Triglia, T., Menting, J. G., Wilson, C., & Cowman, A. F. (1997). Mutations in dihydropteroate synthase are responsible for sulfone and sulfonamide resistance in *falciparum*. *Proceedings of the National Academy of sciences*. **94:** 13944-13949.

Ursos, L. M. B., Dzekunov, S., & Roepe, P. D. (2000). The effects of chloroquine and verapamil on digestive vacuolar pH of *P. falciparum* either sensitive or resistant to chloroquine. *Molecular and Biochemical Parasitology*. **110:** 125-134.

Vaidya, A. B., Akella, R., & Suplick, K. (1989). Sequences similar to genes for two mitochondrial proteins and portions of ribosomal RNA in tandemly arrayed 6-kilobase-pair DNA of a malarial parasite. *Molecular and Biochemical Parasitology*. **35:** 97-107.

Vaidya, A. B., Morrisey, J., Plowe, C. V., Kaslow, D. C., & Wellems, T. E. (1993). Unidirectional dominance of cytoplasmic inheritance in two genetic crosses of *Plasmodium falciparum*. *Molecular and Cellular Biology*. **13:** 7349-7357.

Valdes, A. M., Slatkin, M., & Freimer, N. B. (1993). Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics*. **133:** 737-749.

van Deutekom, J. C., Bakker, E., Lemmers, R. J., van der Wielen, M. J., Bik, E., Hofker, M. H. *et al.* (1996). Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. *Human Molecular Genetics*. **5:** 1997-2003.

van Doorne, C. W., Eling, W. M., & Luyken, R. (1998). Rodent malaria in rats exacerbated by milk protein, attenuated by low-protein vegetable diet. *Tropical Medicine and International Health*. **3:** 596-600.

Vigouroux, Y., McMullen, M., Hittinger, C. T., Houchins, K., Schulz, L., Kresovich, S. *et al.* (2002). Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proceedings of the National Academy of sciences*. **99:** 9650-9655.

Vigouroux, Y., Mitchell, S., Matsuoka, Y., Hamblin, M., Kresovich, S., Smith, J. S. *et al.* (2005). An analysis of genetic diversity across the maize genome using microsatellites. *Genetics*. **169:** 1617-1630.

Waller, R. F., Reed, M. B., Cowman, A. F., & McFadden, G. I. (2000). Protein trafficking to the plastid of *Plasmodium falciparum* is via the secretory pathway. *EMBP Journal*. **19:** 1794-1802.

Walliker, D., Babiker, H. A., & Ranford-Cartwright, L. C. (1998). In: *Malaria: parasite biology, Pathogenesis and Protection* . ASM Press, Washington.

Walliker, D., Carter, R., & Sanderson, A. (1975). Genetic studies on *Plasmodium chabaudi*: recombination between enzyme markers. *Parasitology*. **70:** 19-24.

Walliker, D., Hunt, P., & Babiker, H. A. (2005). Fitness of drug-resistant malaria parasites. *Acta Tropica*. **94:** 251-259.

Walliker, D., Quakyi, I. A., Wellems, T. E., McCutchan, T. F., Szarfman, A., London, W. T. *et al.* (1987). Genetic analysis of the human malaria parasite *Plasmodium falciparum*. *Science*. **236:** 1661-1666.

Walliker, D., Sanderson, A., Yoeli, M., & Hargreaves, B. J. (1976). A genetic investigation of virulence in a rodent malaria parasite. *Parasitology*. **72:** 183-194.

Walter, R. D. (1986). Altered dihydrofolate reductase in pyrimethamine-resistant *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **19:** 61-66.

Walter, R. D. & Konigk, E. (1980). 7,8-Dihydropteroate-synthesizing enzyme from *Plasmodium chabaudi*. *Methods in Enzymology*. **66:** 564-570.

Wang, P., Lee, C. S., Bayoumi, R. A., Djimde, A. A., Doumbo, O. K., Swedberg, G. *et al.* (1997a). Resistance to antifolates in *Plasmodium falciparum* monitored by sequence analysis of dihydropteroate synthetase and dihydrofolate reductase alleles in a large number of field samples of diverse origins. *Molecular and Biochemical Parasitology*. **89:** 161-177.

Wang, P., Nirmalan, N., Wang, Q., Sims, P. F., & Hyde, J. E. (2004). Genetic and metabolic analysis of folate salvage in the human malaria parasite *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **135:** 77-87.

Wang, P., Read, M., Sims, P. F., & Hyde, J. E. (1997b). Sulfadoxine resistance in the human malaria parasite *Plasmodium falciparum* is determined by mutations in dihydropteroate synthetase and an additional factor associated with folate utilization. *Molecular Microbiology*. **23:** 979-986.

Wang, P., Brobey, R. K. B., Horii, T., Sims, P. F. G., & Hyde, J. E. (1999). Utilization of exogenous folate in the human malaria parasite Plasmodium falciparum and its critical role in antifolate drug synergy. *Molecular Microbiology*. **32:** 1254-1262.

Ward, S. A. (2002). Drug resistance mechanisms in malaria. *Pharmacologist*. **44:** A76-

Warhurst, D. C., Craig, J. C., & Adagu, I. S. (2002). Lysosomes and drug resistance in malaria. *Lancet*. **360:** 1527-1529.

Watanabe, J. & Inselburg, J. (1994). Establishing a physical map of chromosome No. 4 of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. **65:** 189-199.

Watkins, W. M., Sixsmith, D. G., Chulay, J. D., & Spencer, H. C. (1985). Antagonism of sulfadoxine and pyrimethamine antimalarial activity in vitro by p-aminobenzoic acid, p-aminobenzoylglutamic acid and folic acid. *Molecular and Biochemical Parasitology*. **14:** 55-61.

Weber, J. L. (1990). Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. *Genomics*. **7:** 524-530.

Weber, J. L. & Wong, C. (1993). Mutation of human short tandem repeats. *Human Molecular Genetics*. **2:** 1123-1128.

Weirdl, M., Dominska, M., & Petes, T. D. (1997). Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics*. **146:** 769-779.

Wellems, T. E., Panton, L. J., Gluzman, I. Y., do Rosario, V. E., Gwadz, R. W., Walker-Jonah, A. *et al.* (1990). Chloroquine resistance not linked to mdr-like genes in a *Plasmodium falciparum* cross. *Nature*. **345:** 253-255.

Wellems, T. E. & Plowe, C. V. (2001). Chloroquine-resistant malaria. *Journal of Infectious Diseases*. **184:** 770-776.

Wellems, T. E., Walker-Jonah, A., & Panton, L. J. (1991). Genetic mapping of the chloroquine-resistance locus on *Plasmodium falciparum* chromosome 7. *Proceedings of the National Academy of sciences*. **88:** 3382-3386.

Wernsdorfer, W. H. & Kouznetsov, R. L. (1980). Drug-resistant malaria--occurrence, control, and surveillance. *Bulletin of the World Health Organization*. **58:** 341-352.

Wernsdorfer, W. H. & Payne, D. (1991). The dynamics of drug resistance in *Plasmodium falciparum*. *Pharmacology and Therapeutics*. **50:** 95-121.

White, N. J. (1992). Antimalarial drug resistance: the pace quickens. *Journal of Antimicrobial Chemotherapy*. **30:** 571-585.

White, N. J. (1998). Drug resistance in malaria. *British Medical Bulletin*. **54:** 703-715.

White, N. J. (2004). Antimalarial drug resistance. *Journal of Clinical Investigation*. **113:** 1084-1092.

Wilson, C. M., Serrano, A. E., Wasley, A., Bogenschutz, M. P., Shankar, A. H., & Wirth, D. F. (1989). Amplification of a gene related to mammalian *mdr* genes in drug-resistant *Plasmodium falciparum*. *Science*. **244:** 1184-1186.

Winstanley, P. A., Watkins, W. M., Newton, C. R., Nevill, C., Mberu, E., Warn, P. A. *et al.* (1992). The disposition of oral and intramuscular pyrimethamine/sulphadoxine in Kenyan children with high parasitaemia but clinically non-severe falciparum malaria. *British Journal Clinical Pharmacology*. **33:** 143-148.

Wissing, F., Sanchez, C. P., Rohrbach, P., Ricken, S., & Lanzer, M. (2002). Illumination of the malaria parasite *Plasmodium falciparum* alters intracellular pH. Implications for live cell imaging. *Journal of Biological Chemistry*. **277:** 37747-37755.

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

World Health Organization 1961. Chemotherapy of Malaria. WHO Technical Report Series. WHO. Geneva.

World Health Organization 1973. Chemotherapy of Malaria and Resistance to Antimalarials. WHO. Geneva.

World Health Organization 1991b. The overlap in the clinical presentation and treatment of malaria and pneumonia in children: report of a meeting. World Health Organization. Geneva.

World Health Organization (1991a). *Basic malaria microscopy, Part I. Leaner's Guide.* WHO, Geneva.

World Health Organization 1996. *Assessment of Therapeutic Efficacy of Antimalarial Drugs For Uncomplicated Falciparum Malaria in Areas With Intense Transmission.* World Health Organization. Geneva: Division of Control of Tropical Diseases.

World Health Organization 2003b. *Assessment and Monitoring of Antimalarial Drug Efficacy for the Treatment of Uncomplicated Falciparum Malaria.* WHO. Geneva.

World Health Organization 2003a. Human Development Report 2003. Millennium Development Goals: a Compact among Nations to End Human Poverty. World Health Organization. Geneva.

Wright, S. (1921). Systems of mating. *Genetics.* **6:** 111-178.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics.* **15:** 354-

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to the system of mating. *Evolution.* **19:** 420.

Wyman, A. R. & White, R. (1980). A highly polymorphic locus in human DNA. *Proceedings of the National Academy of sciences.* **77:** 6754-6758.

Yayon, A. (1985). The antimalarial mode of action of chloroquine. *Reviews in Clinical and Basic Pharmacology.* **5:** 99-139.

Yayon, A., Cabantchik, Z. I., & Ginsburg, H. (1984). Identification of the acidic compartment of *Plasmodium falciparum*-infected human erythrocytes as the target of the antimalarial drug chloroquine. *EMBP Journal.* **3:** 2695-2700.

Zalis, M. G., Wilson, C. M., Zhang, Y., & Wirth, D. F. (1993). Characterization of the *pfmdr2* gene for *Plasmodium falciparum. Molecular& Biochemical Parasitology.* **62:** 83-92.

Zhang, K. & Rathod, P. K. (2002). Divergent regulation of dihydrofolate reductase between malaria parasite and human host. *Science.* **296:** 545-547.

Zhang, Y. & Meshnick, S. R. (1991). Inhibition of *Plasmodium falciparum* dihydropteroate synthetase and growth *in vitro* by sulfa drugs. *Antimicrobial Agents and Chemotherapy.* **35:** 267-271.

Zhivotovsky, L. A. & Feldman, M. W. (1995). Microsatellite variability and genetic distances. *Proceedings of the National Academy of sciences.* **92:** 11549-11552.

Zolg, J. W., Plitt, J. R., Chen, G. X., & Palmer, S. (1989). Point mutations in the dihydrofolate reductase-thymidylate synthase gene as the molecular basis for pyrimethamine resistance in *Plasmodium falciparum. Molecular and Biochemical Parasitology.* **36:** 253-262.

Zwetyenga, J., Rogier, C., Tall, A., Fontenille, D., Snounou, G., Trape, J. F. *et al.* (1998). No influence of age on infection complexity and allelic distribution in *Plasmodium falciparum* infections in Ndiop, a Senegalese village with seasonal, mesoendemic malaria. *American Journal Of Tropical Medicine & Hygiene.* **59:** 726-735.