Full Length Research Paper

Quantification of markers of antimalarial drug resistance from an area of high malaria transmission: Comparing frequency with prevalence

Allen Malisa[1,2]*, Richard Pearce[3], Ben Mutayoba[4], Salim Abdullah[2], Hassan Mshinda[2], Patrick Kachur[5], Peter Bloland[5] and Cally Roper[3]

[1]Sokoine University of Agriculture (SUA), Department of Biological Sciences, Faculty of Science, Box 3038, Morogoro, Tanzania.
[2]Ifakara Health Institute (IHI), Ifakara, Kilombero District, Morogoro, Tanzania.
[3]London School of Hygiene and Tropical Medicine, Pathogen Molecular Biology Unit, Department of Infectious Tropical Diseases, Keppel Street, London, WC1E 7HT, UK.
[4]Sokoine University of Agriculture (SUA), Department of Veterinary Physiology, Biochemistry, Pharmacology and Toxicology, Faculty of Veterinary Medicine, PO Box 3017, Morogoro, Tanzania.
[5]Malaria Epidemiology Branch, Division of Parasitic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention (CDC), Atlanta, GA, USA.

Molecular monitoring of markers of antimalarial drug resistance offers an affordable alternative to the *in vivo* method for the detection of resistance, and has the potential to guide public health policy in a timely manner. However, the optimal way of analyzing and reporting these data, particularly those emanating from areas of moderate to high malaria transmission, has never been fully explored or agreed upon, given the potential of being confounded by coinfections. By using large number of real field samples, we quantified the difference between prevalence and frequency when reporting field data on antimalarial drug resistance obtained by direct counting of haplotypes. Polymerase chain reaction (PCR) and sequence specific oligonucleotide probing was used to generate point mutations which were used to construct haplotypes. Results indicate that frequency underestimates haplotypes present at low levels while also amplifying haplotypes present at high levels; prevalence on the other hand behaved in a vice versa manner. Both prevalence and frequency are therefore essential, as each may have relevance in different contexts in high malaria transmission settings. Frequency is essential to gauge the impact of intervention on antimalarial drug resistance while prevalence may be more relevant when the aim is to determine parasite clearance.

Key words: Molecular markers, polymerase chain reaction (PCR) - sequence specific oligonucleotide probing (SSOP), prevalence, frequency.

INTRODUCTION

Malaria is one of the major public health challenges in areas of the world where the disease is endemic. Early diagnosis and prompt treatment of the disease with effective drugs, along with vector control are the main strategies for its control in endemic areas, but the development and spread of drug resistance to the most commonly used anti-malarial drugs, has complicated this strategy (Olliaro, 2005). To provide essential information to the health authorities responsible for ensuring access to effective drugs by justification of alternative anti-malarial drug policy, it is essential to quantify the level of drug resistance through regular monitoring of parasite resistance (Hastings et al., 2007). *In vivo* efficacy

---

*Corresponding author. E-mail: malisa56@yahoo.com or amalisa@suanet.ac.tz. Tel: 255 232603404.

surveillance is a standard method for detection of resistance yet, the technique suffers logistical and financial constrains while its interpretation is confounded by factors such as reinfection, immunity, and pharmacokinetics (Uhleman et al., 2005).

Surveillance designed to detect molecular markers of drug resistance offers rapid and affordable options to monitor parasite resistance in the field, and this approach has subsequently become an integral part for the evaluation of resistance to treatment (WHO, 2005; Plowe et al., 2007). In this approach, large numbers of infected blood samples can be collected and single nucleotide polymorphisms (SNPs)/mutations associated with drug resistance are rapidly screened and used to guide health policy decisions. For long time, there has been a call for optimization of methods for analyzing and reporting molecular markers of drug resistance, particularly in areas of intense malaria transmission characterized by multiple infections.

An increasing number of molecular studies reporting the frequency of known resistance markers presents opportunities to assess the broader impact the study and reporting methodology may be having on the data being produced, notably reporting data with polyclonal infections (Djimde et al., 2001; Kublin et al., 2002). An account of problems associated with analysis of human blood samples from areas of moderate to high malaria transmission is given elsewhere (Hastings et al., 2010). Blood samples from these areas are often infected with multiple genetically distinct malaria clones and the number of co-infecting clones is known as multiplicity of infection (MOI). Samples for which MOI>1 and contains both wild type (sensitive) and mutant (resistant) genotypes, present an exceeding complication to determine the proportions of sensitive and resistant genotypes (Hastings et al., 2010). Furthermore, by using most of the available genotyping methods including polymerase chain reaction and restriction fragment length polymorphism (PCR-RFLP), it is also either difficult or completely impossible to assign specifically each of the various SNPs into specific parasites contained in a single patient blood sample, complicating the whole exercise of true haplotype construction from SNPs.

Haplotypes are combinations of SNP that are in the same gene in the same parasite, as distinct from associations of point mutations that co-occur because there is a mixture of parasites of different genotypes within a single infection (Pearce et al., 2003). Consequently, haplotypes are biologically meaningful, since they determine the resistance properties of parasites that are exposed to drugs at the time of treatment (Pearce et al., 2003). For example, the genetic determinants of pyrimethamine and sulphadoxine resistance are point mutations at codons 16, 50, 51, 59, 108, 164 and 436, 437, 540, 581, 613 of the dihydrofolate reductase-*dhfr* (Cowman et al., 1998) and dihydropteroate synthase-*dhps* (Brooks et al., 1994)

genes, respectively and a mixed infection containing both 51I+108N and 59R+108N double-mutant haplotypes is less resistant to pyrimethamine than an infection containing the triple-mutant 51I +59R+108N haplotype, despite all three mutations being present in either case. It is therefore advisable to measure the frequency of haplotypes rather than the prevalence of each point mutation separately, because haplotypes are the determinants of resistance levels.

In this study, a high throughput sequence-specific oligonucleotide probe-based approach described by Pearce et al. (2003) was used to generate genotypic data. In this method, relative abundance of a given parasite genotype were inferred from intensity of probe signal such that the greater the intensity the greater the parasite population and vice versa, allowing for reliable assignment of SNPs to alternative genotypes and haplotype construction.

As the number of molecular studies, reporting the frequency/prevalence of known resistance markers continues to mount, researchers have started to realize the need to assess the broader impact that reporting conventions and study methodology may be having on the data being produced (Pearce et al., 2003; Hastings et al., 2010; Sridaran et al., 2010). While there is general consensus that, analysis and interpretation of samples from areas of moderate to high malaria transmission present difficulties because of genetic ambiguity resulting from high MOI (MOI>1), there is a varying opinion on the expression of the data into either frequency or prevalence due to overestimation or underestimation of the rare genotypes (Pearce et al., 2003; Hastings et al., 2010; Sridaran et al., 2010).

A number of studies conducted in areas of moderate to high malaria transmission have raised concern about the implication that the use of prevalence and frequency of molecular markers of drug resistance might have when used as a tool for the provision of information to guide policy (Pearce et al., 2003; Hastings et al., 2010).

Hasting et al. (2010) defined prevalence as the proportion of human blood samples where the marker is present, implying that one or more clones in the sample carry the marker. He also defined frequency as the proportion of individual malaria clones that carries the marker. Prevalence requires accounting for all SNPs irrespective of their numbers. In an attempt to achieve consistent and comparable quality data on which rational decision to guide policy can be based, Hastings et al. (2010) wrote and updated a MalHaploFreq software, incorporating an aspect of maximum likelihood (ML) estimation of SNPs and haplotype frequencies, assuming that all clones are detectable and the appropriate genotyping sensitivity limit (GSL) to avoid misclassification in samples with MOI>1 from areas of moderate to high malaria transmission. As this awaits adoption by researchers, it may be important to quantify the difference between reporting field studies in terms of

prevalence and frequency emanating from direct count of SNPs or haplotypes, as until presently, this seems the predominant practice (Pearce et al., 2003; Roper et al., 2003; Pearce et al., 2009; Malisa et al., 2010).

This study reports a comparative analysis of prevalence and frequency of *dhfr* and *dhps* haplotypes in a large scale field samples from high malaria transmission area in south eastern Tanzania.

MATERIALS AND METHODS

Study area, subjects and samples

Cross sectional community surveys were conducted during July, August and September of 2000 and 2006 in three rural districts of south-eastern Tanzania, Rufiji (Population = 170,000), Kilombero (Population = 220,000) and Ulanga (Population = 160,000). The three districts were well matched in terms of predicted intensity and duration of malaria transmission and risk (MARA), relative access and overall utilization of health services (based on surveys), usage of insecticide treated nets (ITNs) and relative proportion of urban peri-urban, rural population.

The surveys were part of large combination therapy pilot implementation programme in Tanzania, the interdisciplinary monitoring programme for antimalarial combination therapy (IMPACT-TZ). Impact-Tanzania is a multiyear implementation research evaluation that rests on a collaborative platform incorporating the United States Center for Disease Control and Prevention (CDC), the Ifakara Health Institute, London School of Hygiene and Tropical Medicine, and the Ministry of Health and Social Welfare including its National Malaria Control Programme, the Tanzania Essential Health Interventions Project and the Council Health Management Teams of Rufiji, Kilombero and Ulanga Districts. Impact-Tanzania is primarily supported by funding from the United States Agency for International Development, CDC and Welcome trust.

For the purpose of the study, Kilombero and Ulanga (KU) Districts were treated as a single district because population movement between these two districts is high and the study population spans the border region. *Plasmodium falciparum* malaria transmission in the study area is intense (with an estimated entomological inoculation rate of 367 infectious bites per person per year) (Hay et al., 2001) and perennial with some seasonal fluctuation.

A total of 14,900 adults and children from the randomly selected households participated in the study. A finger-prick blood sample for blood slide and filter paper (3 MM; Whatman International Ltd., Maidstone, United Kingdom) bloodspot were collected from each individual in the household. The filter paper bloodspots were air-dried and stored at room temperature in self-sealing plastic bags with dessicant. All blood slide samples were screened by light microscopy for *P. falciparum* parasites. Bloodspots from microscopically positive subjects were selected for molecular genotyping.

Ethics

Scientific and ethical clearance was obtained from the Medical Research Council of the National Institute for Medical Research in Tanzania, the Centers for Disease Control and Prevention, USA, and the London School of Hygiene and Tropical Medicine. Written informed consent was obtained from all individuals or their guardians before collection of samples.

DNA extraction

The DNA was extracted from bloodspots dried on filter papers. A section of the dried blood spot filter paper was excised using a sterile blade or scissors, and soaked in a 1 ml, 0.5% saponin-1x phosphate buffered saline (PBS). The section of filter paper was then washed twice in 1 ml of 1x PBS and finally, was boiled for 8 min in 100 µl PCR quality water with 50 µl 20% chelex suspension (pH 9.5).

PCR amplification

Nested PCR was used to amplify a 594 base pair (bp) fragment of *dhfr* and a 711 bp fragment of *dhps* each containing the sequence where mutations are found. Primer sequences and PCR reaction conditions were previously described in the study of Pearce et al. (2003). PCR was performed in 96 well plates with 25 µl PCR reaction volumes containing final concentrations of 0.25 µM each oligonucleotide primers, 2 mM $MgCl_2$, 250 µM each deoxyribonucleotide triphosphate (dNTPs), and 1x Taq polymerase. One microliter (1 µl) of DNA template was used in the outer (primary) PCR reaction mixture for *dhfr* and *dhps* amplifications. For the inner (secondary) *dhps* reactions 1 µl of the outer PCR product was used. The outer *dhfr* PCR products were diluted three fold before 1 µl was introduced into the inner PCR reaction mixtures.

Molecular genotyping of point mutations by sequence specific oligonucleotide probing (SSOP)

The amplified PCR products were screened for *dhfr* and *dhps* sequence variants at 10 loci where SNPs associated with SP drug resistance are known to occur. The sequence changes (and the amino acid substitutions they code for) are summarized in Table 1.

PCR products were spotted in a 12 by 8-grid and cross linked onto nylon membranes and probed for sequence polymorphisms by hybridization to specific oligonucleotide probes described previously in the study of Pearce et al. (2003). For analysis of samples collected in 2000, the visualization of hybridized digoxygenin labeled probes on membranes was performed by the alkaline phosphatase-catalysed breakdown of the CSPD substrate (Roche Boehringer Mannheim, Mannheim, Germany) and visualized by exposure on Hyperfilm-ECL (Amersham Pharmacia Biotech, Little Chalfont, Buckinghamshire, United Kingdom), according to Boehringer Mannheim recommendations and previously described in the study of Conway et al. (1999).

For analysis of samples collected in 2001 and 2002, the probed blots were visualized using ECF substrate and detection using a phosphoimager (STORM®). Inspection of autoradiographic films was carried out by light box illumination, while the phophoimager output was recorded through viewing of digitally-captured images of chemifluorescent signal. The change in the method by which probe hybridization signal was visualized did not affect the results in any way since the probes and hybridization conditions were unchanged.

The stringency and specificity of the hybridization process was confirmed by inspection of a series of 4 controls with a known single genotype variant sequence. All blots with non-specifically bound probes were stripped and re-probed. A SNP was considered to be present in the PCR product when the intensity of signal was higher than that of the background. The blots were scored independently by two people.

Estimation of haplotype frequency

Haplotype frequency is defined as the proportion of individual malaria clones that carries a particular haplotype. In the haplotype

Table 1. The nucleotide and amino acid substitutions at *dhfr* and *dhps* genes screened for by PCR-SSOP.

Codon	*dhfr*				
	50	**51**	**59**	**108**	**164**
Wild type	Cys (C), TGT	Asn (N), AAT, AAC	Cys (C), TGT	Ser (S), AGC	Ile (I), ATA
Mutant	Arg (R), CGT	Ile (I), ATT	Arg (R), CGT	Asn (N), AAC, Thr (T), ACC	Leu (L), TTA

	dhps				
Codon	**436**	**437**	**540**	**581**	**613**
Wild type	Ser (S), TCT	Ala (A), GCT	Lys (K), AAA	Ala (A), GCG	Ala (A), GCC
Mutant	Phe (F), TTT, Ala (A), GCT, Cys (C), TGT	Gly (G), GGT	Glu (E), GAA	Gly (G), GGG	Ser (S), TCC, Thr (T), ACC

frequency analysis, the aim was to establish the relative abundance of different point mutation haplotypes at *dhfr* and *dhps*. Since blood stage *P. falciparum* is haploid, this is very straightforward when an infection consists of a single genotype because only one form of sequence at every SNP locus is seen.

When infections are composed of multiple genotypes a mixture of different sequence variants occurs making the inference of point mutation haplotypes within that infection more difficult. The presence, absence, and relative abundance of hybridization signal for every probe were recorded at each locus. A sample was considered to have a single haplotype when only one sequence variant was found at each locus. Blood samples were categorized as having a single, a majority, or a mixture of sequence at each SNP locus.

Majority and mixed genotype infections were differentiated according to the relative intensity of signal. If the hybridization signal of the minority sequence was greater than half the intensity of the majority then an infection was classified as mixed. To determine the relative abundance of different point mutation haplotypes in the parasite population, only one haplotype was counted from each infection and those mixed infections where haplotypes could not be resolved were omitted from the calculation of haplotype frequencies. Hence, frequency data is based upon a subset of isolates which were either unmixed or had a predominating majority haplotype. Therefore, haplotype frequency was calculated as the ratio of allelic haplotypes among those infections where a single or majority haplotype was detected at the resistance locus.

Estimation of haplotype prevalence

Prevalence is defined as the proportion of human blood samples where the haplotype is present, implying that one or more clones in the sample carry the marker. In this analysis, the aim was to account for every point mutation haplotype. As for the frequency estimation, when only one sequence variant was found at each locus, the sample was considered to have a single haplotype and its estimation was straightforward.

However, when infections are composed of multiple genotypes a mixture of different sequence variants occurs necessitating the need to disentangle individual point mutation haplotypes. This was achieved by listing all point mutation haplotypes present in each sample irrespective of whether they were of equal, minority or majority parasite population. Since the higher the number of constituent point mutations in a given haplotype, the more complex is the haplotype estimation, *dhps* haplotype prevalence was relatively easier to estimate than *dhfr*.

Unlinked (selectively neutral) microsatellite loci analysis

To examine the extent of the underlying rate of mixed infections and monitor transmission intensity in the study area, unlinked microsatellite markers (*Poly A, Pfpk2 and TA109)* were genotyped in a subset of samples; 180 samples from each of Rufiji and KU populations, respectively in 2000, and an equal samples from each of the two populations in 2006. A semi-nested PCR was used to amplify microsatellites in 11.0 μl reaction volumes. The 1° reaction composed of 1 μ template DNA, 1x thermo buffer, and 3.0 mmol/L Mg^{2+}, 0.75 pmol/L of each primer and 1 unit of Taq polymerase. The cycling conditions were: 2 min at 94°C; and then 25 repeated cycles of 30 s at 94°C; 30 s at 42°C; 30 s at 40°C; and 40 s at 65°C; followed by 2 min at 65°C.

In the second round of PCR, a third fluorescent labeled primer (Applied Biosystems, Warrington, Cheshire UK) was incorporated and the final PCR reaction volume (11.0 μl) consisted of 1x thermo buffer, 2.5 mmol/L Mg^{2+}, 2 pmol/L primer, 1 unit of Taq polymerase and 1 μl of the outer (1°) PCR products. The cycling conditions were: 2 min at 94°C; and then 25 repeated cycles of 20 s at 94°C; 20 s at 45°C; and 30 s at 65°C; followed by 2 min at 65°C.

Semi nested PCR products were diluted at 1:100 ratio and run with LIZ – 500 size standards on an Abi 3730 genetic analyser (Applied Biosystems, Warrington, and Cheshire, UK). Fragments were sized using the Gene Mapper software (Applied Biosystems, Warrington, and Cheshire, UK). Multiplicity of infection for each sample was recorded by counting the number of peaks, each representing clonally different parasite (haploid nature of blood stages of *P. falciparum*).

Statistical analysis

Statistical comparison of allele prevalences and frequencies at *dhfr* and *dhps* was carried out using chi-squared analysis in STATA version 9.2 (Stata, 2002).

RESULTS

A total of 14,900 asymptomatic persons were sampled and 3,294 found infected with *P. falciparum*. DNA was extracted from all *P. falciparum* positive samples and PCR amplification of *dhfr* and *dhps* performed, giving an average amplification success of 60% for both genes (Table 2). The amplified products were screened for all

Table 2. PCR outcome for the 2000 and 2006 house hold survey samples collected from two rural districts, Rufiji and KU in Tanzania.

Parameter	Rufiji		KU	
Year	2000	2006	2000	2006
Survey population	2844	4267	3289	4500
P. falciparum positive	778	916	955	645
PCR amplified *dhfr*	549	683	404	294
PCR amplified *dhps*	521	703	444	323
Single or majority *dhfr*	455	616	376	275
Single or majority *dhps*	417	588	365	275

A

B
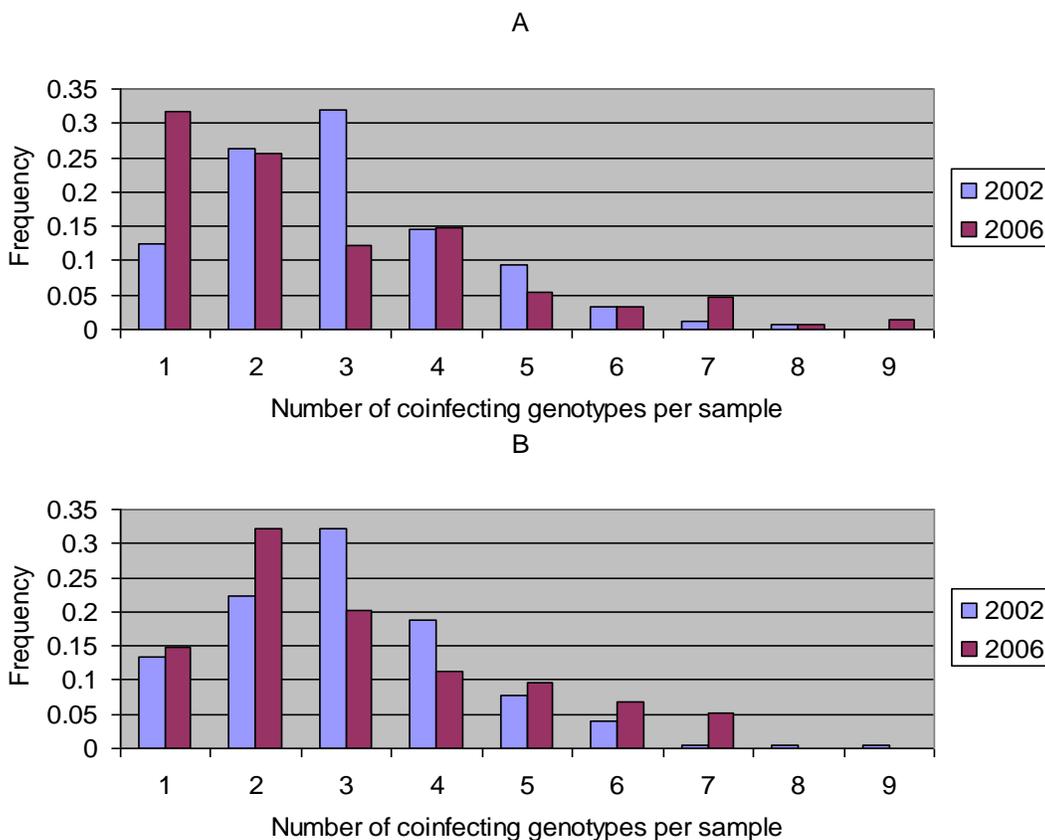
Figure 1. MOI distribution in 2002 and 2006 isolates from Rufiji and KU populations determined by genotyping of three neutral microsatellites markers (*Poly A, Pfpk2 and TA109*). A = KU, B = Rufiji.

the variant sequences described in Table 1. Out of the 1,930 isolates which amplified successfully for *dhfr*, 89% were single or majority genotype infections. Of the 1,991 samples which amplified successfully for *dhps,* 83% were single or majority genotype with known haplotypes. In reporting frequency, allelic haplotypes, the conformations of point mutations in *dhfr* and *dhps* were recorded only from single and majority genotype infections, while in the case of prevalence, all genotype infections were recorded irrespective of whether they represent minority or majority parasite population.

The number of distinct *P. falciparum* genotypes detectable per sample is defined as the 'multiplicity of infection' (MOI). Since analysis of human blood samples from areas of moderate to high malaria transmission is complicated by MOI, it was necessary for this study to determine the underlying MOI in the study area at the time of sampling. Figure 1 shows estimation of MOI from the two sites, Rufiji and KU in years 2000 and 2006.

In the year 2000, isolates from both Rufiji and KU districts had remarkably similar distribution of MOI (Figure 1). The most common MOI was 3 in both districts,
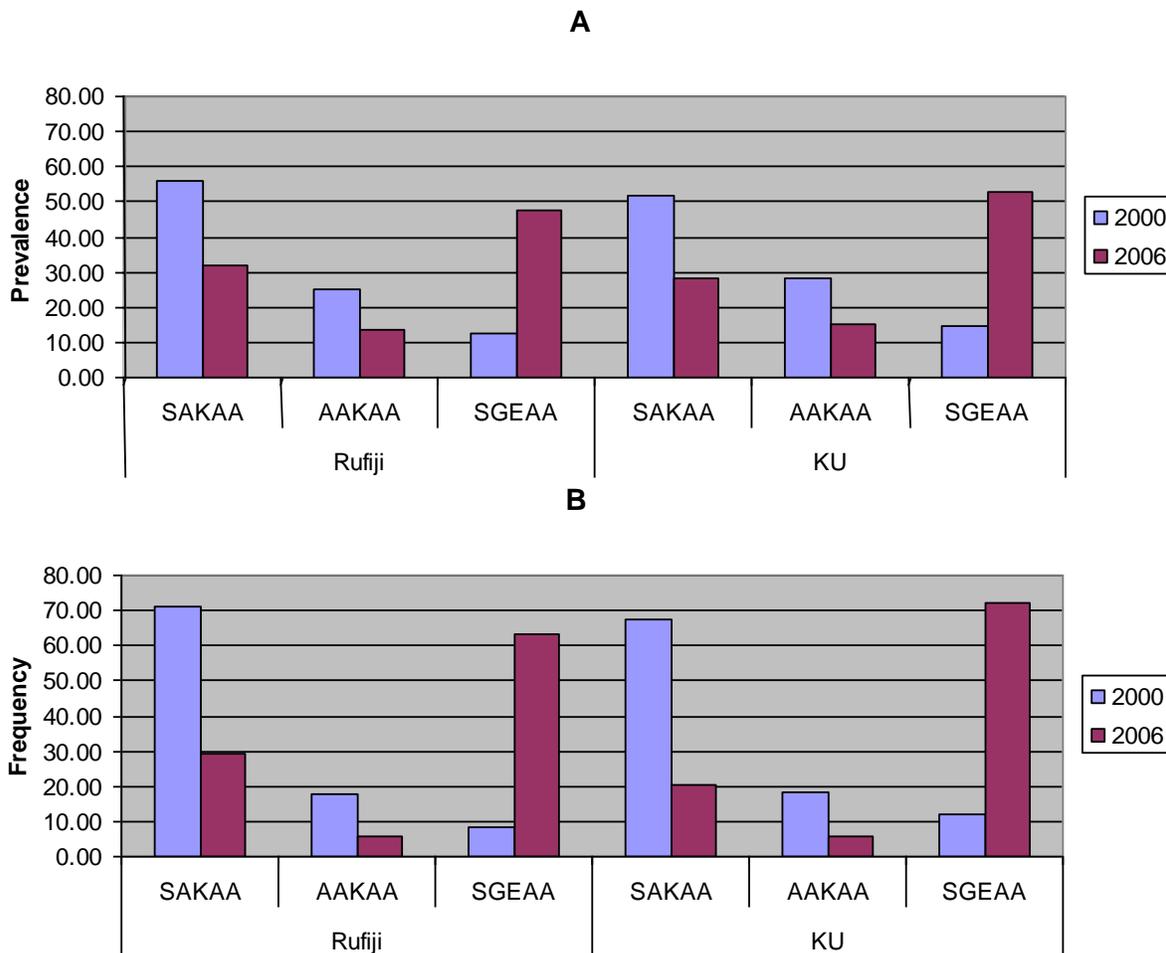
A

B
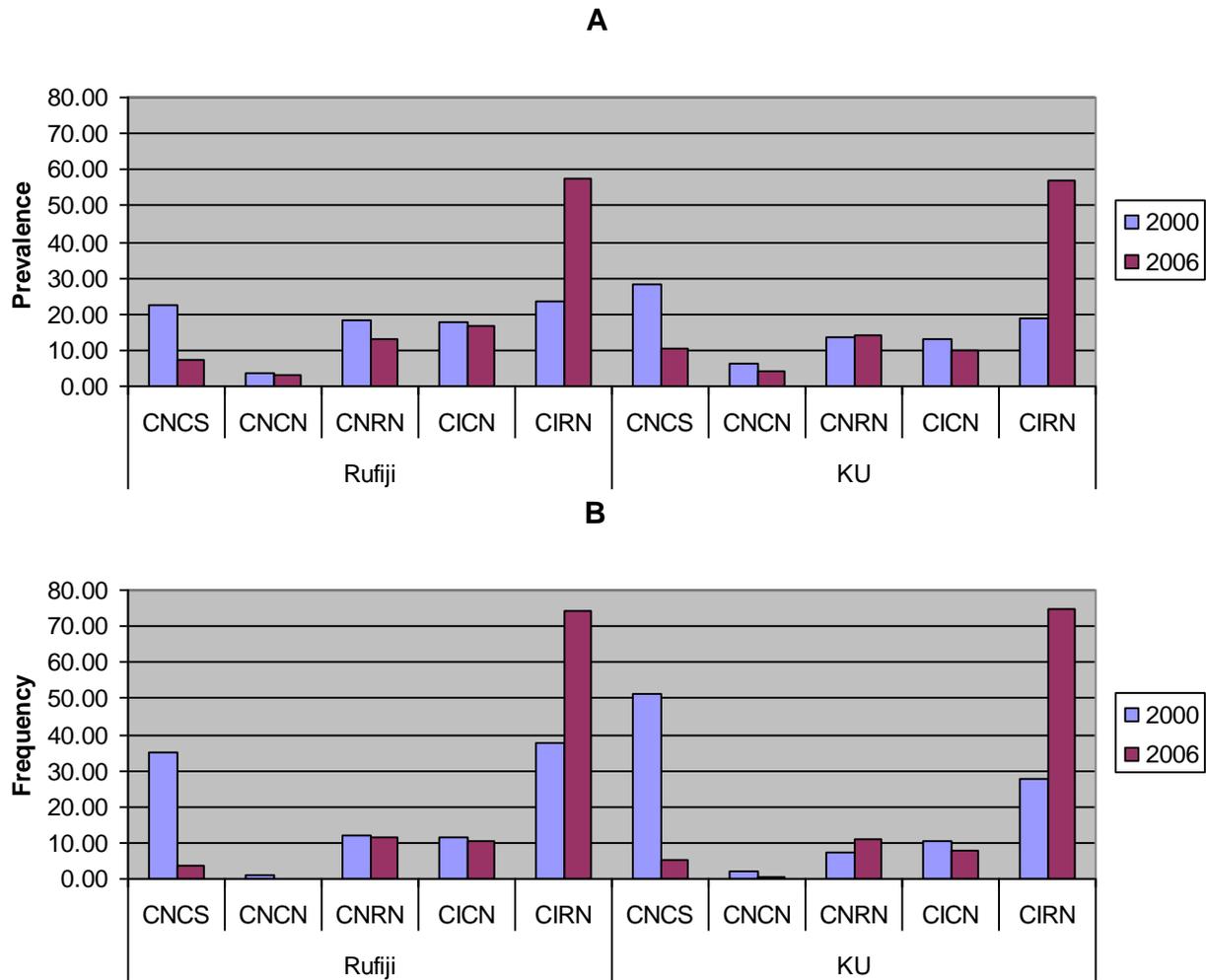
Figure 2. Comparison of *dhfr* haplotype prevalence and frequency in two populations of Rufiji and KU. A = prevalence, B = frequency.

and an estimated 60 (33%) of 180 patients screened from each of Rufiji and KU were found to be coinfected with 3 genetically distinct *P. falciparum* parasites. However, in 2006, a trend towards reduction of malaria transmission was observed which was associated with the decrease of the most common MOI from 3 in both populations to 2 in Rufiji and 1 in KU. Generally, MOI analysis confirmed high malaria transmission in the populations with MOI ranging from 1 to 9 in both surveys.

The point mutations found in the study area were N51I, C59R and S108N for the *dhfr* and S436A, S436F, S436C, A437G, and K540E for the *dhps* gene. While all the three *dhfr* point mutations were common and widespread, two of the *dhps* mutations, S436F and S436C, were very rare and the rest three mutations, S436A, A437G, and K540E, were common and widespread. The three *dhfr* mutations identified in this study were present in seven haplotypic conformation; CNCS, CNRN, CICN and CIRN, which were common and widespread and CNCN, CNRS and CICS, which

were extremely rare and occurred in a very low proportion. Likewise, for the *dhps* gene, the observed five point mutations were present in four common and widespread haplotypic conformations; SAKAA, AAKAA, SGEAA, SGKAA and five rare and low frequency haplotypes; SAEAA, CAKAA, FAKAA, AAEAA, and FAEAA. Detailed account of haplotype changes with time are described elsewhere (Malisa et al., 2010).

Comparison between prevalence and frequency of *dhfr* and *dhps* haplotypes in the two study sites revealed a remarkably similar progression; indicating that the highly resistant haplotypes (CIRN and SGEAA) were increasing with time displacing the wild type haplotypes (CNCS, SAKAA) while the moderately resistant haplotypes (CNCN, CNRN, CICN and AAKAA) displayed almost a constant frequency or prevalence (Figures 2 and 3). The observed progression of *dhfr* and *dhps* haplotype prevalence and frequency is consistent with drug use history in Tanzania. Sampling in 2000 was done one year prior to policy change from CQ first line to SP first

A

B

Figure 3. Comparison of *dhps* haplotype prevalence and frequency in two populations of Rufiji and KU. A = prevalence, B = frequency.

line in 2001 (17 years of SP use as second line treatment) while sampling in 2006 took place after 5 years of SP use as first line treatment, and thus it is plausible that the increased SP drug pressure resulting from its widespread use as national first line antimalarial drug could be attributed to the observed rise of resistant *dhfr* and *dhps* haplotypes in 2006.

Our analysis has found a striking contrasting behaviour when comparing frequency with prevalence among the wild type (sensitive), the moderately resistant and the highly resistant *dhfr* and *dhps* haplotypes. While the frequencies were greater than prevalencies in the wild type and the highly resistant haplotypes (except the SGEAA in 2000 which behaved like the moderately resistant haplotypes), the moderately resistant haplotypes displayed an opposite relationship, showing prevalences which were greater than frequencies (Tables 1 and 2).

A closer comparison between haplotype prevalence and frequency revealed interesting differences between the two mutation quantification methods (Table 3). Most

striking differences were observed in wild type haplotypes and the highly resistant haplotypes in both the *dhfr* and *dhps* genes. The prevalence of the *dhps* wild type haplotype, the SAKAA in 2000 was 56 and 52% in Rufiji and KU while its frequency was 71 and 68% in Rufiji and KU, respectively; this difference was statistically significant (p≤0.0001). By contrast, in 2006, the prevalence and frequency of the wild type *dhps* haplotype (SAKAA) were remarkably closely similar, with prevalence of 32 and 28% in Rufiji and KU and frequency of 29 and 20% in Rufiji and KU, respectively this difference was not statistically significant (p≥0.05).

For the highly resistant *dhps* haplotype (SGEAA), its prevalence in 2000 was 12 and 15% in Rufiji and KU, respectively while its frequency was 8 and 12% in Rufiji and KU, respectively this difference was statistically insignificant (p≥0.05). In 2006 its prevalence was 47.8% and 53.0% in Rufiji and KU, respectively while its frequency was 63.3 and 72.1% in Rufiji and KU, respectively, this difference was statistically significant (p≤0.0001).

Table 3. Comparison of *dhfr* and *dhps* haplotype prevalence and frequency in Rufiji and KU for the 2000 and 2006 isolates.

Population	Haplotype	Year					
		2000			**2006**		
		Prevalence	**Frequency**	**P value (95% CI)**	**Prevalence**	**Frequency**	**P value (95% CI)**
Rufiji	SAKAA*	55.8	71.1	0.0001	31.9	29.1	0.435
	SGEAA	12.3	8.1	0.231	47.8	63.3	0.0001
	CNCS*	22.6	35.2	0.0001	7.1	3.6	0.189
	CIRN	23.4	37.4	0.0001	57.8	74.3	0.0001
KU	SAKAA*	51.8	67.7	0.0001	28.4	20.7	0.05
	SGEAA	14.6	12.1	0.402	53.0	72.1	0.0001
	CNCS*	28.5	51.2	0.0001	10.3	5.5	0.265
	CIRN	18.7	28.0	0.01	57.1	74.9	0.0001

*Wild type haplotype.
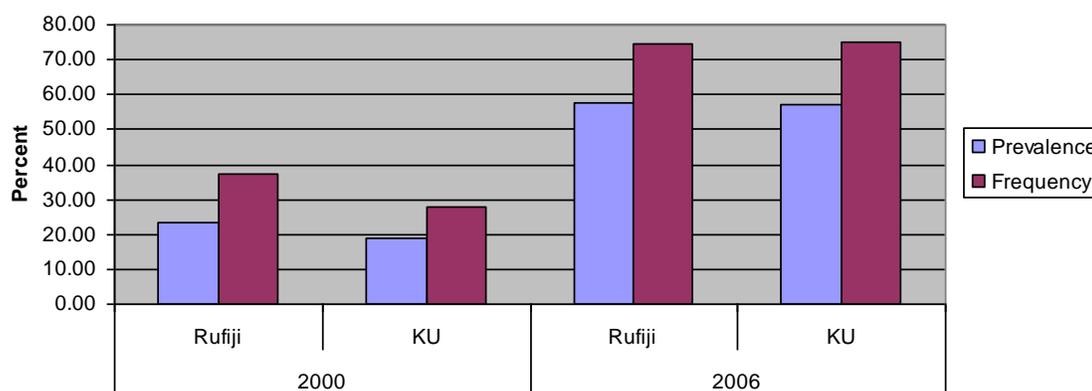
Figure 4. Comparison of the prevalence and frequency of triple mutant *dhfr* (CIRN) parasites in the two districts, Rufiji and KU.
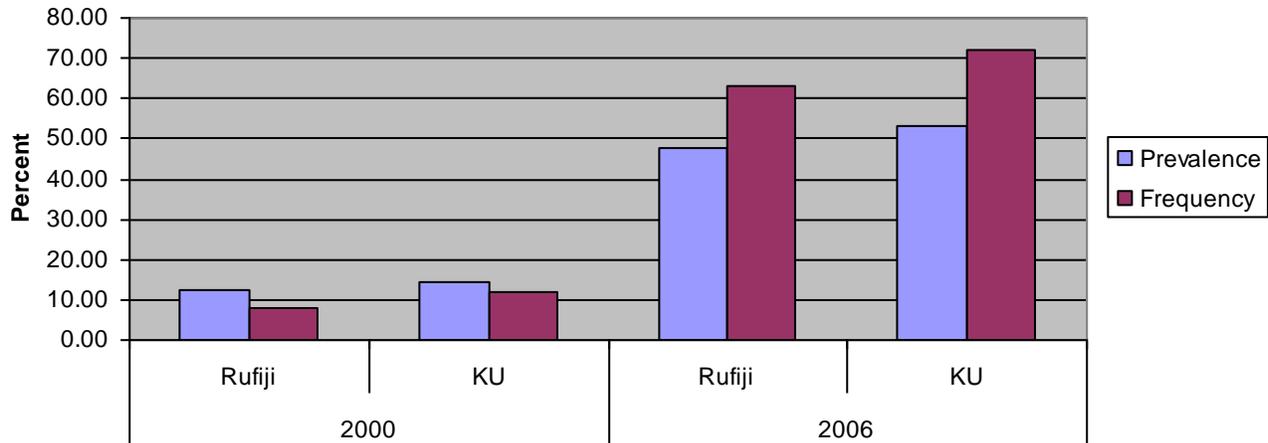
Comparison of prevalence and frequency of the wild type *dhfr* haplotype (CNCS) revealed significant differences in 2000 but insignificant differences in 2006 isolates in the two study sites. In 2000, its prevalence was 22.6 and 28.5% in Rufiji and KU, respectively while its frequency was 35.2 and 51.2% in Rufiji and KU, respectively; this difference was statistically significant (p≤0.0001). However, in 2006, its prevalence was 7.1 and 10.3% in Rufiji and KU, respectively while its frequency was 3.6 and 5.5% in Rufiji and KU, respectively, this difference was not statistically significant (p≥0.05).

Comparison of prevalence and frequency of the highly resistant *dhfr* haplotype (CIRN) revealed significant differences in both 2000 and 2006 isolates in the two study sites (Figure 4). In 2000, its prevalence was 23.4 and 18.7% in Rufiji and KU, respectively while its frequency was 37.4 and 28.0% in Rufiji and KU, respectively, this difference was statistically significant (p≤0.001). Similarly, in 2006, its prevalence was 57.8 and 57.1% in Rufiji and KU, respectively while its frequency was 74.3 and 74.9% in Rufiji and KU, respectively; this difference was statistically significant (p≤0.0001).

Prevalence and frequency comparison of the triple mutant *dhfr* CIRN and the double mutant *dhps* SGEAA haplotypes, which have the greatest significance for SP efficacy, are shown in Figures 4 and 5, respectively. As described previously, the CIRN haplotype maintained similar relationship between prevalence and frequency, with frequencies being greater than prevalences at the two sampling points of 2000 and 2006 (Figure 4). However, the relationship was different for the SGEAA haplotype, whereby prevalences were greater than frequencies in 2000 isolates but a reciprocal relationship was observed in the 2006 isolates where frequencies were greater than prevalences (Figure 5).

DISCUSSION

This study has quantified the difference between prevalence and frequency when reporting field data on antimalarial drug resistance obtained by direct counting of SNPs or haplotypes. Until recently, frequency method is still a common reporting method (Pearce et al., 2003,

Figure 5. Comparison of the prevalence and frequency of double dhps mutant parasites in the two districts, Rufiji and KU. A = prevalence, B = frequency.

2009; Roper et al., 2003; Malisa et al., 2010, 2011).

We found greater frequencies than prevalences for the sensitive haplotypes and resistant haplotypes conferring high resistance to pyrimethamine (CIRN) and sulfadoxine (SGEAA), but greater prevalences than frequencies in haplotypes conferring mild resistance to pyrimethamine (CNCN, CNRN, and CICN) and to sulfadoxine (AAKAA) in 2000 and 2006 isolates from the two study sites. This finding is consistent with others (Hastings and Smith, 2008) who assert that frequency reporting method has a weakness of underestimation of SNPs/haplotypes present at low frequencies. Taken together, the findings of this study revealed that in reporting field data by direct counting method, frequency underestimates haplotypes present at low levels while also amplifying haplotypes present at high levels; prevalence on the other hand behaved in a vice versa manner.

The two reporting methods may have relevance in different contexts in high malaria transmission setting; if the intention is to monitor progression of resistance to guide policy makers in a form of early warning signal of a deteriorating drug treatment, it is important to use frequency to avoid underestimation of the highly resistant haplotypes/SNPs. When the aim is to determine parasite clearance, information about all forms of the parasite becomes useful and hence prevalence may be more informative. It is difficult to explain why the frequencies of the highly resistant SNPs/haplotypes were greater than their prevalences but we speculated that, for the successful spread of a resistance, the mutants must be transmitted at a faster rate than the sensitive forms. This occurs in the presence of drug because of the differential survival and reproduction rates conferred by these alleles (Anderson et al., 1989).

The highly resistant mutant parasites survive treatment while sensitive and mildly resistant genotypes do not. In addition to that, the drug treatment itself, by purging the co-infecting sensitive genotypes, promotes assortative mating among resistant survivors of treatment. Consequently, by virtue of their higher fitness the proportion of the unmixed highly resistant SNPs/haplotypes by far outweighs the mixed form and mildly resistant SNPs/haplotypes, taken together. In our previous reports (Malisa et al., 2010), we found that the reproduction rates of the mildly resistant *dhfr* and *dhps* haplotypes were very slow even at a peak SP drug pressure keeping an almost constant frequency.

Observation of higher SGEAA prevalences than frequencies in year 2000 was an exception of the previous description and requires some explanation. To better understand this, one need to first understand the malaria treatment policy in Tanzania. Treatment of malaria in Tanzania is typically guided by official recommendations from the Ministry of Health and Social Welfare (MOHSW) regarding drugs of choice for various situations. "First-line" treatment refers to the drug officially recommended as the drug of first choice for the treatment of uncomplicated malaria. "Second-line" treatment refers to the drug officially recommended as an alternative primarily to be used for treatment of patients in whom the first-line treatment failed to clear the infection and other select patients (such as those who are hypersensitive to the first-line treatment). "Third-line" treatment typically refers to the drug recommended for severely ill patients (a rescue drug).

In practice, few treatment failures are recognized and patients are often moved directly from first to third-line treatment, consequently, little second-line drug is used compared to the first-line drug. At the time of sample collection in 2000, Sulfadoxine/pryrimethamine (SP) was still a second line treatment and it only replaced chloroquine (CQ) as the recommended first-line antimalarial treatment on the Tanzanian mainland in August 2001, after 18 years of its use as a second-line

treatment since 1983. Clearly, the SP drug selection pressure was very low and a result of monitoring of resistant *dhfr* and *dhps* allelic haplotypes in the same area found that the frequency of the most resistant allele, the double *dhps*-triple *dhfr* (SGAA-CIRN) mutant genotype, increased by only 1% during 17 years of SP second line use, but there was a dramatic increase by a 45% during five (2001 to 2006) years of SP first line use (Malisa et al., 2011). Consequently, as a result of low selection the proportion of mildly resistant alleles also remained low leading to the observed higher prevalences than frequencies.

It was important to determine the malaria transmission intensity in the two study districts, Rufiji and KU, as the complication associated with genotyping of human blood infected with malaria parasite occurs in moderate to high malaria transmission areas. This is because infections in these areas tend to be composed of multiple genetically distinct parasite lines (Babiker et al., 1999) and this may occur in two ways; through super infection of a host following receiving multiple infective bites or through the bites of a mosquito carrying a mixture of genotypes within its saliva inoculum.

Analysis of three selectively neutral microsatellites markers (*Poly A, Pfpk2 and TA109*) confirmed high malaria transmission in both study districts, in 2000 and 2006 surveys with MOI ranging from 1 to 9 in both populations. Moreover, the study found a remarkably similar distribution of MOI in Rufiji and KU districts suggesting existence of homogenizing gene flow between the two populations. Furthermore, comparison between 2000 and 2006 MOI distribution revealed a trend toward decreasing malaria transmission which was more marked in KU than Rufiji. In 2000, the most common MOI was 3 in both populations yet by 2006, the most common MOI in Rufiji was 2 and KU was 1.

In recent years, several studies have reported a decrease in malaria transmission (Killeen et al., 2007; Okiro et al., 2007) attributing the reduction to the effect of transmission blocking interventions such as integrated vector control and use of ITNs.

In conclusion, as researchers continue to report data on molecular genotyping of antimalarial drug resistance through direct counting of haplotype/SNP, information on the implication of reporting the data either in frequency or prevalence is important. Our findings have provided insight on that area while also confirming previous findings (Hasings et al., 2010) which asserts that both prevalence and frequency are therefore essential, as each may have relevance in different contexts in high malaria transmission settings.

Frequency is essential to gauge the impact of intervention on antimalarial drug resistance while prevalence may be more relevant when the aim is to determine parasite clearance. Researchers are also urged to measure haplotypes rather than SNPs because the latter rarely act independently and their combination in haplotypes often determines the resistance properties of a given parasite.

ACKNOWLEDGEMENT

REFERENCES

Anderson RM, May RM, Gupta S (1989). Non-linear phenomena in host-parasite interactions. Parasitology 99:S59-S79.

Babiker HA, Ranford-Cartwright LC, Walliker D (1999). The epidemiology of multiple *P. falciparum* infections: Genetic structure and dynamics of *P. falciparum* infections in the Kilombero region of Tanzania. Trans. R Soc. Trop. Med. Hyg. 93(1):11–14.

Brooks DR, Wang P, Read M, Watkins WM, Sims PF, Hyde JE (1994). Sequence variation of the hydroxymethyldihydropterin pyrophosphokinase: dihydropteroate synthase gene in lines of the human malaria parasite, *Plasmodium falciparum*, with differing resistance to sulfadoxine. Eur. J. Biochem. 224:397-405.

Conway DJ, Roper C, Oduola AM, Arnot DE, Kremsner PG, Grobusch MP, Curtis CF, Greenwood BM (1999). High recombination rate in natural populations of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. USA 96:4506-4511.

Cowman AF, Morry MJ, Biggs BA, Cross GA, Foote SJ (1998). Amino acid changes linked to pyrimethamine resistance in the dihydrofolate reductase thymidylate synthase gene of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. USA. 85:9109–9113.

Djimde A, Doumbo OK, Cortese JF, Kayentao K, Doumbo S, Diourte Y, Dicko A, Su XZ, Nomura T, Fidock DA, Wellems TE, Plowe CV, Coulibaly D (2001). A molecular marker for chloroquine-resistant falciparum malaria. N. Engl. J. Med. 344:257-263.

Hastings IM, Korenromp EL, Bloland PB (2007). The anatomy of a malaria disaster: drug policy choice and mortality in African children. Lancet Infect. Dis. 7:739-748.

Hastings IM, Nsanzabana C, Smith TA (2010). A Comparison of Methods to Detect and Quantify the Markers of Antimalarial Drug Resistance. Am. J. Trop. Med. Hyg. 83(3):489-495.

Hastings IM, Smith TA (2008). MalHaploFreq: a computer programme for estimating malaria haplotype frequencies from blood samples. Malar. J. 7:130

Hay SI, Rogers DJ, Toomer JF, Snow RW (2001). Annual *Plasmodium falciparum* entomological inoculation rates (EIR) across Africa: literature survey, Internet access and review. Trans. R Soc. Trop. Med. Hyg. 94:113–127.

Killeen GF, Tami A, Kihonda J, Okumu FO, Kotas ME, Grundmann H, Kasigudi N, Ngonyani H, Mayagaya V, Nathan R, Abdulla S, Charlwood JD, Smith TA, Lengeler C (2007). Cost-sharing strategies combining targeted public subsidies with private-sector delivery achieve high bed net coverage and reduced malaria transmission in Kilombero valley, southern Tanzania. BMC Infect. Dis. 7:121.

Kublin JG, Dzinjalamala FK, Kamwendo DD, Malkin EM, Cortese JF, Martino LM, Mukadam RA, Rogerson SJ, Lescano AG, Molyneux ME, Winstanley PA, Chimpeni P, Taylor TE, Plowe CV (2002). Molecular markers for failure of sulfadoxine-pyrimethamine and chlorproguanil-dapsone treatment of *Plasmodium falciparum* malaria. J. Infect. Dis. 185:380-388.

Malisa A, Pearce R, Abdulla S, Mutayoba B, Mshinda H, Kachur P, Bloland P, Roper C (2011). The evolution of pyrimethamine resistant *dhfr* in *Plasmodium falciparum* of south eastern Tanzania: comparing

selection under SP alone vs SP+Artesunate combination. Malar. J. 10:317.

Malisa A, Pearce RJ, Abdulla S, Mshinda H, Kachur P, Bloland P, Roper C (2010). Drug coverage in treatment of malaria and the consequences for resistance evolution-evidence from the use of sulfadoxine/pyrimethamine. Malar. J. 9:190.

Okiro EA, Hay SI, Gikandi PW, Sharif SK, Noor AM, Peshu N, Marsh K, Snow RW (2007). The decline in paediatric malaria admissions on the coast of Kenya. Malar. J. 6:151.

Olliaro P (2005). Drug resistance hampers our capacity to roll back malaria. Clin. Infect. Dis. 41(4):S247-S257.

Pearce R, Drakeley C, Chandramohan D, Mosha F, Roper C (2003). Molecular Determination of Point Mutation Haplotypes in the Dihydrofolate Reductase and Dihydropteroate Synthase of *Plasmodium falciparum* in Three Districts of Northern Tanzania. Antimicrob. Agents Chemother. 47:1347-1354

Pearce R, Pota H, Evehe MB, El Hadj B, Mombo-Ngoma G, Malisa A, Ord R, Inojosa W, Matondo A, Diallo DA, Mbacham W, van den Broek IV, Swarthout TD, Assefa A, Dejene S, Grobusch MP, Njie F, Dunyo S, Kweku M, Owusu-Ageyi S, Chandramohan D, Bonnet M, Guthmann J-P, Clarke S, Barnes K, Streat E, Katokele ST, Uusiku P, Agboghoroma CO, Elegba OY, Cissé B, A-Elbasit IE, Giha HA, Kachur P, Lynch C, Rwakimari J, Chanda P, Hawela M, Sharp B, Naidoo I, Roper C (2009). Dispersal of drug resistant *dhps* reveals regional migration patterns among African *P. falciparum*. Plos Biology 6(4):e1000055.

Plowe CV, Roper C, Barnwell JW, Happi CT, Joshi HH, Mbacham W, Meshnick SR, Mugittu K, Naidoo I, Price RN, Shafer RW, Sibley CH, Sutherland CJ, Zimmerman PA, Rosenthal PJ (2007). World Antimalarial Resistance Network (WARN) III: molecular markers for drug resistant malaria. Malar. J. 6:121.

Roper C, Pearce R, Bredenkamp B, Gumede J, Drakeley C, Mosha F, Chandramohan D, Sharp B (2003). Antifolate antimalarial resistance in southeast Africa. Lancet 361:1174-1181.

Sridaran S, McClintock SK, Syphard LM, Herman KM, Barnwell JW, Udhayakumar V (2010). Anti-folate drug resistance in Africa: meta-analysis of reported dihydrofolate reductase (dhfr) and dihydropteroate synthase (dhps) mutant genotype frequencies in African Plasmodium falciparum parasite populations. Malar. J. 9:247.

Stata Corp (2002). Stata Statistical Software: Release 7.0. College Station, TX: Stata Corporation.

Uhleman AC, Yuthavong Y, Fidock DA (2005). Mechanisms of antimalarial drug action and resistance. In Molecular Approaches to Malaria Edited by: Sherman IW. Washington: ASM Press, pp. 429-461.

World Health Organization (WHO) (2005). Susceptibility of *Plasmodium falciparum* MAL/2005.1103.