

# Reduced Variation Around Drug-Resistant *dhfr* Alleles in African *Plasmodium falciparum*

Richard Pearce,\* Allen Malisa,†‡ S. Patrick Kachur,‡§ Karen Barnes,||  
Brian Sharp,¶ and Cally Roper\*

\*London School of Hygiene and Tropical Medicine, Pathogen Molecular Biology Unit, Department of Infectious Tropical Diseases, London, United Kingdom; †University of Morogoro, Department of Biological Sciences, Faculty of Science, SUA, Morogoro, Tanzania; ‡Ifakara Health Research and Development Centre, Ifakara, Tanzania; §Malaria Branch, Division of Parasitic Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia; ||University of Cape Town Division of Clinical Pharmacology, Cape Town, South Africa; and ¶Malaria Research Lead Programme, Medical Research Council, Durban, South Africa

We have measured microsatellite diversity at 26 markers around the *dhfr* gene in pyrimethamine-sensitive and -resistant parasites collected in southeast Africa. Through direct comparison with diversity on sensitive chromosomes we have found significant loss of diversity across a region of 70 kb around the most highly resistant allele which is evidence of a selective sweep attributable to selection through widespread use of pyrimethamine (in combination with sulfadoxine) as treatment for malaria. Retrospective analysis through four years of direct and continuous selection from use of sulfadoxine-pyrimethamine as first-line malaria treatment on a *Plasmodium falciparum* population in KwaZulu Natal, South Africa, has revealed how recombination significantly narrowed the margins of the selective sweep over time. A deterministic model incorporating selection coefficients measured during the same interval indicates that the transition was toward a state of recombination-selection equilibrium. We compared loss of diversity around the same resistance allele in two populations at either extreme of the range of entomological inoculation rates (EIRs), namely, under one infective bite per year in Mpumalanga, South Africa, and more than one per day in southern Tanzania. EIRs determine effective recombination rates and are expected to profoundly influence the dimensions of the selective sweep. Surprisingly, the dimensions were broadly consistent across both populations. We conclude that despite different recombination rates and contrasting drug selection histories in neighboring countries, the region-wide movement of resistant parasites has played a key role in the establishment of resistance in these populations and the dimensions of the selective sweep are dominated by the influence of high initial starting frequencies.

## Introduction

Drug treatment remains the primary means of clearing potentially lethal *Plasmodium falciparum* malaria infections, and drug use has applied strong positive directional selection for resistance mutations. Theory predicts that selection will have a significant impact on genomic diversity (Smith and Haigh 1974). Initially, there is a complete association between the new favored mutation and the genome in which it arose. As the selected allele increases in frequency, more distant associations are quickly broken down by recombination until only associations between the selected allele and sequences immediately flanking the gene remain—this type of association is often referred to as hitchhiking (Smith and Haigh 1974; Kaplan, Hudson, and Langley 1989). Eventually, the signature of selection is a pattern of reduced gene diversity (expected heterozygosity [ $H_e$ ]) at the region of sequence flanking the selected locus (Kaplan, Hudson, and Langley 1989). This loss of diversity is known as a selective sweep. Selective sweeps have been described in maize, humans, and *Drosophila* (Quesada et al. 2003; Bersaglieri et al. 2004; Palaisa et al. 2004), often long after the initial selection events which created them. Over time, the eroding effects of new mutations and recombination obscure the signature of the original selection event, but in the case of *P. falciparum* drug selection is recent and measurable. We have looked at selection of resistance alleles at *dhfr* through use of pyrimethamine for treatment of *P. falciparum* in

the present day and made direct measurement of selection coefficients over a 4-year period of drug use (Roper et al. 2003). Here we examine the impact of selection, recombination, and migration on chromosome diversity around the highly pyrimethamine-resistant triple-mutant allele, which is prevalent throughout southeast (SE) Africa.

The antimalarial drug pyrimethamine is a competitive inhibitor of the folate biosynthesis pathway that targets the active site of the enzyme dihydrofolate reductase (DHFR), and resistance to pyrimethamine is associated in vitro with substitutions within the active site of DHFR (Cowman et al. 1988; Peterson, Walliker, and Wellems 1988; Snewin et al. 1989). In Africa, resistant forms of the enzyme contain substitutions at three different sites and permutations of these substitutions confer a range of sensitivities to the drug, the higher the number of substitutions, the greater the insensitivity to pyrimethamine (Wu, Kirkman, and Wellems 1996). In SE Asia, more resistant parasites are found containing a mutation at fourth position, codon 164.

Microsatellite variation has been described around *dhfr* carrying two to four mutations sampled from SE Asian populations. Diversity was significantly reduced in sites within a 100-kb region around *dhfr* with a region of 12 kb with strongly reduced gene diversity within a valley extending from 58 kb upstream to over 40 kb downstream (Nair et al. 2003). Selection was not concurrent with the study as use of the drug was discontinued some 10 years prior to the study (Nosten et al. 2000). Importantly, all alleles had identical or very similar flanking microsatellites indicating a single ancestral origin. In Africa the situation is different. Although microsatellite flanking markers are tightly associated with *dhfr* resistance alleles containing two or more mutations, three separate origins of parasites

Key words: *Plasmodium falciparum*, selective sweeps, pyrimethamine resistance.

E-mail: richard.pearce@lshtm.ac.uk.

*Mol. Biol. Evol.* 22(9):1834–1844. 2005

doi:10.1093/molbev/msi177

Advance Access publication May 25, 2005

carrying two mutations but just a single origin of parasites carrying the triple-mutant (N51I + C59R + S108N) *dhfr* alleles were identified in two populations in SE Africa (Roper et al. 2003). In both Africa and Asia, all alleles containing a single mutation have different flanking microsatellites suggesting multiple origins for single mutants. Comparing resistance alleles from Africa and SE Asia and using allele sharing at six markers over a 30-kb region around *dhfr*, we showed that the Africa triple-mutant shared ancestry with the SE Asian resistance expansion (Roper et al. 2004), indicating that it was introduced into Africa and introgressed into the population over considerable distances (>4000 km).

To further explore the dynamics underlying the introgression of this allele in Africa and to describe the extent of the selective sweeps around it, we have measured the loss of diversity at microsatellite loci through comparison to a baseline of high diversity described on chromosomes carrying the drug-sensitive allele. Loci are described at increasing distances from the *dhfr* gene in three different population contexts, each with differing recombination rates and selection histories. It is these parameters that determine the size of a selective sweep. Selection acts to maintain high frequencies of the favored allele and thereby the associations with hitchhiking loci (Barton 2000; Kim and Stephan 2002). A high recombination rate will act to dissolve the associations between hitchhiker and selected site, increasing variation at the flanking loci. The intensity of malaria transmission, by multiplying the number of genotypes infecting the same individual and increasing opportunities for outcrossing, has a profound effect on effective rates of recombination.

Analysis of a genetic cross has estimated the recombination rate in *P. falciparum* to be 17 kb/cM (Su et al. 1999), which is considerably higher than that found in other eukaryotes. However, this rate is moderated in the field according to the epidemiology of the parasite. Blood-stage *Plasmodium* is haploid, and recombination only occurs during meiosis in the mosquito vector. The rate of outcrossing and therefore detectable recombination is determined by transmission intensity (Dye and Williams 1997). In regions of higher transmission intensity, as indicated by high entomological inoculation rates (EIRs), there are more multiple infections, increasing the number of different clones ingested by the mosquito and increasing the probability of outcrossing during meiosis. The effective recombination rate is a more accurate measure of the recombination rate of the parasite as it also incorporates the degree of inbreeding ( $F$ ) in the population (Babiker et al. 1994; Dye and Williams 1997; Conway et al. 1999).

Using a panel of 26 microsatellite loci, we have mapped the gene diversity along drug-sensitive chromosomes, reflecting the ancestral state of diversity on chromosome 4 and triple-mutant chromosomes with *dhfr* as a central point. We have examined a single population where the selection coefficient is known during a 4-year period of drug selection (fig. 1). Having a standardized effective recombination rate across two time points, we can describe the reduction in gene diversity at flanking loci over time due to a known strength of selection. To quantify the effect of recombination on the dimensions of the selective sweep around *dhfr*-

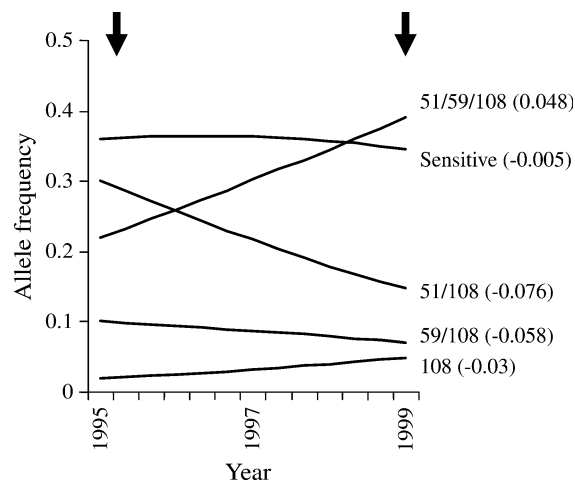


FIG. 1.—The change in allele frequencies of the five most common *dhfr* haplotypes. The haplotype names are derived from the mutated positions in *dhfr* in the codon order 50, 51, 59, and 108. Selection coefficients (in parentheses) were calculated from the change in allele frequencies over 12 generations (1 year being equivalent to three generations). The arrows indicate the sample points. Data are taken from Roper et al. (2003).

resistant chromosomes, we compared loss of diversity around the same resistance allele from two populations on either extreme of a spectrum of transmission intensity in SE Africa. In this comparison, we test the hypothesis that we should see significant differences in the dimensions of selective sweeps for parasite populations with differing effective recombination rates.

## Materials and Methods

### Study Samples

To explore the effects of direct selection through time, parasite populations were sampled over two time points (1996 and 1999) in KwaZulu Natal. Samples were collected from patients presenting to a health care facility. The sites used for contrasting effective recombination rates were in the south of Tanzania and in Mpumalanga, a province in the northeast of South Africa. The Tanzanian samples were collected during household surveys of the three districts of Kilombero, Ulanga, and Rufiji as a part of the Interdisciplinary Monitoring Project for Antimalarial Combination Therapy in Tanzania (IMPACT-T<sub>2</sub>) artesunate combination therapy trial. The samples from Mpumalanga were from symptomatic malaria patients prior to treatment, as a component of the South East African Combination Antimalarial Therapy (SEACAT) evaluation. Table 1 summarizes the drug use and epidemiological context of the populations sampled. All three study sites are endemic for *P. falciparum*.

Finger prick bloodspots were made of blood samples taken at all time points. DNA extraction from bloodspots on filter paper was carried out in a 96-well plate format. A segment of the bloodspot was first soaked in 0.5% saponin in 1 × phosphate-buffered saline (PBS) overnight, and then washed twice in 1 ml 1 × PBS. The segment was then boiled for 8 min in 100 μl polymerase chain reaction (PCR) quality water plus 50 μl 20% chelex suspension in distilled water (pH 9.5). Typing of point mutations at

**Table 1**  
**Drug Use and Epidemiological Context of Populations Sampled**

Population and Sample Date	History of SP Usage (first line unless stated otherwise)	Frequency of the Triple-Mutant <i>dhfr</i> Allele (51/59/108)	EIR (infectious bites per annum)
KwaZulu Natal, 1996–1999	1988–2000	22%–38% <sup>a</sup> (42%–62%) <sup>a,b</sup>	<0.8 <sup>c</sup>
Mpumalanga, 2001	1997 to present day	52% <sup>d,b</sup>	<0.14 <sup>c</sup>
Tanzania, 2001	2001 to present day (after 18 years second line)	40% <sup>d</sup>	584 <sup>c</sup>

<sup>a</sup> Roper et al. (2003).

<sup>b</sup> Clinical samples.

<sup>c</sup> Barnes (unpublished data).

<sup>d</sup> Unpublished data.

<sup>e</sup> Charlwood et al. (1998).

codons 50/51, 59, 108, and 164 was performed according to the method of Pearce et al. (2003), using hybridization of sequence-specific oligonucleotide probes to PCR-amplified products. As blood-stage parasites are haploid, the point mutation haplotypes are immediately evident except where multiple genotypes are coinfecting one person. For the purposes of looking at flanking sequence polymorphism on chromosomes carrying specific allelic forms of *dhfr*, we deliberately selected a subset of samples which were un-mixed at any polymorphic loci at *dhfr*.

#### PCR Amplification and Analysis of Microsatellite Sequences

The microsatellites were amplified in a seminested manner. The primary reaction comprised 1 µl template, 3.0 mmol/Liter Mg<sup>2+</sup>, 0.75 pmol/Liter primer, and 1 unit of *Taq* polymerase. The reaction was cycled as follows: 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.

A third fluorescently labeled primer (Applied Biosystems, Warrington, United Kingdom) was incorporated into a second-round PCR of total volume of 11 µl containing 2.5 mmol/Liter Mg<sup>2+</sup>, 2 pmol/Liter primer, 1 unit of *Taq* polymerase, and 1 µl of outer nest template. Cycling conditions were 2 min at 94°C, then 25 cycles of 20 s at 94°C, 20 s at 45°C, and 30 s at 65°C, and a final step of 2 min at 65°C.

Samples were diluted 1 in 100 and run with LIZ-500 size standard on an ABI 3730 genetic analyzer (Applied Biosystems). Fragments were sized using the GeneMapper software (Applied Biosystems). The samples were preselected, and consequently, multiple alleles in the same isolate were a rare occurrence. In the event of two or more alleles being detected, the majority allele was used if the minority peaks were less than 50% of the height of the majority. If peaks were of equivalent height, data were recorded as missing for that locus in that isolate.

#### Statistics and Software

Gene diversity values were calculated as  $H_e = [n/(n-1)][1 - \sum p_i^2]$ , where  $n$  is the number of samples and  $p_i$  is the frequency of the  $i$ th allele. The variance of

the gene diversity was calculated using the formula of Nei and Roychoudhury (1974):

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

The software PowerMarker (Liu and Muse 2004) was used to calculate population differentiation theta values. The software “bottleneck” (Cornuet and Luikart 1996) was used to determine  $H_e$  levels given the observed number of alleles using a coalescent simulation. The output was used to test the significance of allele distribution shown in the data. While the stepwise mutation model (SMM) is a suitable model for evolution for a majority of microsatellites, the frequency of complex mutations in *P. falciparum* microsatellites suggests that observations based on purely SMM could produce some spurious results (Anderson et al. 2000b). Therefore, both SMM and infinite-allele model (IAM) were used, representing the upper and lower limits of the analysis. Differences between observed and “bottleneck” predicted gene diversity values were tested for significance using Wilcoxon’s tests.  $T$ -tests and nonparametric tests were performed in SPSS 12 (SPSS Inc.)

Wiehe (1998) developed a model to describe the pattern of reduced variation in microsatellites flanking a selected site. This framework was modified by Nair et al. (2003) to describe changes in expected heterozygosity (or gene diversity) rather than variance in repeat number because the often complex nature of *P. falciparum* microsatellites estimates of repeat length are often confounded by significant error. The model uses estimates of mutation rate  $\mu$ , the recombination rate  $r$ , the selection coefficients  $s$ , and the effective population size  $N_e$ .

$$H_e = 1 - e^{-(8\mu + 2r')/s},$$

where  $r'$  is the effective recombination rate and  $e$  is equal to  $m/N_e$  estimating the initial frequency of resistant parasites in the population. In African populations,  $m$  is estimated as the number of migrants from one population to another,

reflecting that the presence of the triple-mutant allele in Africa is due to gene flow from SE Asia rather than through mutational events.

$N_e$  for each population was estimated under IAM using the formula  $N_e\mu = H/4(1 - H)$ , where  $H$  = gene diversity of neutral markers not under the influence of selection and substituting with previous estimates of  $\mu$  of  $1.59 \times 10^{-4}$  mutations/locus/generation (Anderson et al. 2000b).

The recombination rate  $r$  has been estimated by Su et al. (1999) to be  $5.88 \times 10^{-4}$  Morgan/kb/generation from observations of the genetic cross between parasite lines Hb3 and Dd2. Depending on the transmission intensity in a region, a number of *P. falciparum* infections result in self-fertilization (Paul et al. 1995), and this inbreeding affects the apparent rate of recombination. Dye and Williams (1997) established a relationship between the coefficient of inbreeding  $F$  and the recombination rate  $r$  such that  $r' = r(1 - F)$  is the effective recombination rate.  $F$  was estimated from the number of mixed infections detected at *dhfr* in the original survey. As we are most interested in the amount of recombination that occurs between chromosomes carrying *dhfr* alleles, these estimates are not unduly affected by being unable to detect multiple genotypes hidden by shared *dhfr* alleles.

## Results and Discussion

Microsatellites occur in high frequency in *P. falciparum*, on average one per kilobase (Su et al. 1999). *dhfr* occurs at a central point on chromosome 4 using the published sequence of the 3D7 parasite line (Gardner et al. 2002). We were able to identify microsatellite markers at or close to 10, 20, 30, 40, 50, 60, 70, 75, 80, 90, 100, 125, 150, 250, and 350 kb both upstream and downstream from codon 108 of *dhfr* (primer sequences, chromosome location, repeat unit, and allele size range for each locus are summarized in supplementary materials online). The markers spanned 700 kb, 58% of chromosome 4. Of the 30 markers, 5 were discarded for having either an inconsistent repeat size or greater than 35% missing data, suggesting the occurrence of null alleles. The KwaZulu Natal sample consisted of 80 unmixed infections collected from patients in 1996 and 1999, which were selected as representatives of the triple-mutant *dhfr* allele (1996,  $n = 14$ ; 1999,  $n = 28$ ) and the sensitive allele (1996,  $n = 14$ ; 1999,  $n = 14$ ). Ten were removed because they were of a chromosomal haplotype present more than once in the sample set and were therefore presumed to be siblings. A further 11 samples were removed as they had greater than 35% missing data.

### The Pattern of Diversity Around the Sensitive *dhfr* Allele

The pattern of diversity at all microsatellites flanking the sensitive alleles supports the interpretation that the sensitive form of the *dhfr* allele is considered to be the ancestral state of *dhfr*. The mean level of gene diversity in microsatellite loci on sensitive chromosomes was similar in the two time points, being  $0.678 \pm 0.077$  in 1996 ( $n = 9$ ) and  $0.699 \pm 0.077$  in 1999 ( $n = 12$ ). The sample sets of the 2 years were merged as there was no significant differentiation between them

**Table 2**  
The Average Expected Gene Diversity of Microsatellite Markers Grouped by Repeat Unit Size

Repeat Type	$n$ (loci)	Expected Gene Diversity (from sensitive population)	Expected Gene Diversity (from Anderson et al. [2000b])
Di-	15	$0.839 \pm 0.076$	0.781
Tri-	7	$0.593 \pm 0.098$	0.688
Other	3	$0.920 \pm 0.114$	0.636

(Wright's fixation index  $F_{ST} = 0.025$  average of all loci). The mean gene diversity of the merged ancestral population was 0.784, consistent with previous estimates in African populations of between 0.76 and 0.80 based on polymorphism at microsatellite markers dispersed throughout the genome (Anderson et al. 2000a). Table 2 summarizes the average expected heterozygosities of markers when grouped by the repeat unit size. The gene diversity value for the trinucleotide repeats is slightly lower than previously reported, but not significantly so. The gene diversity values of the individual markers together with 95% confidence intervals (CIs) (Nei and Roychoudhury 1974) plotted against distance from codon 108 of *dhfr* are shown in figure 2a.

Gene diversity at marker loci around sensitive alleles should reflect baseline gene diversity in the ancestral state. To test for population events that may potentially confound this assumption, we tested for excess of heterozygosity. The software "bottleneck" (Cornuet and Luikart 1996) carries out a Wilcoxon's signed rank test comparing observed heterozygosity at each locus across the 700-kb region and expected values generated under IAM and SMM. It was found that under IAM there was a significant excess of expected heterozygosity (Wilcoxon's test one-tail  $P = 0.00757$ ), but not so under SMM, indicating that the population had not recently undergone a reduction in size. The high level of diversity at each marker and the absence of a difference in alleles present between the two time points, together with the lack of evidence for a population structure event, mean that the sensitive population can be used as a representation of the ancestral state of each locus.

### The Pattern of Diversity Around the Triple-Mutant *dhfr* Alleles Sampled in KwaZulu Natal, South Africa

The diversity at loci on triple-mutant chromosomes was compared with that on sensitive chromosomes. Figure 2b and c shows plots of gene diversity on the sensitive chromosomes compared with that on triple-mutant chromosomes collected in 1996 ( $n = 13$ ) and 1999 ( $n = 17$ ), respectively. The most striking observation in each case is the valley of significantly reduced gene diversity on the selected chromosomes. In addition to this, there are differences between 1996 and 1999. In 1996, when the frequency of triple mutants in the population was 22% (fig. 1), the region of significantly reduced diversity extended across 70 kb from locus U60 upstream to D10 downstream. In 1999, when the frequency of the triple mutant had increased to 38% (fig. 1), the region was constricted to 50 kb, extending from U40 upstream to D10 downstream (fig. 2c). Interestingly, at both time points a significant

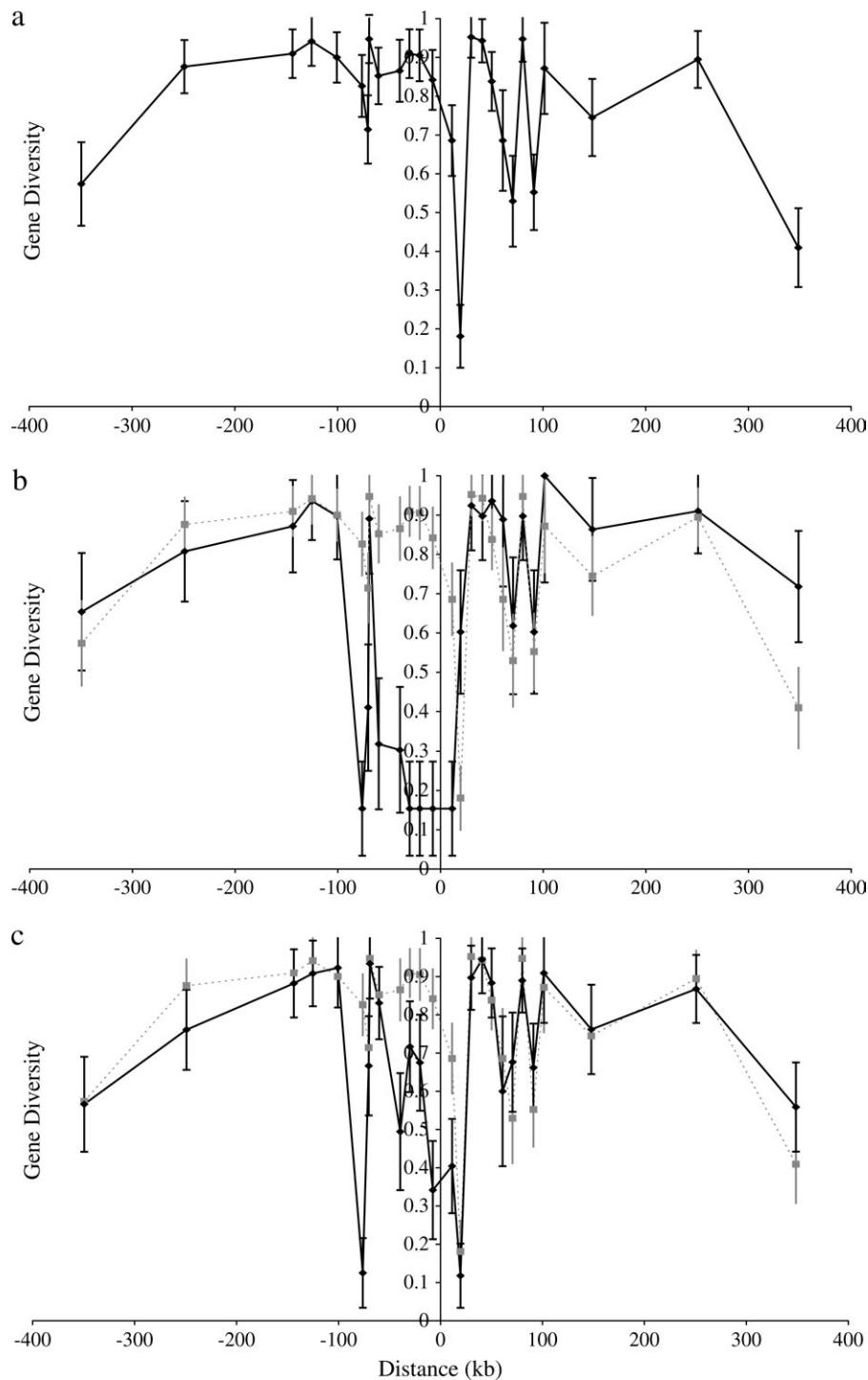


FIG. 2.—The gene diversity at each loci for populations of (a) merged sensitive *dhfr* alleles (dashed line in all subsequent charts), (b) triple-mutant *dhfr* allele from KwaZulu Natal in 1996 (solid line), and (c) triple-mutant *dhfr* allele from KwaZulu Natal in 1999 (solid line). Distances shown are actual physical distances from codon 108 of *dhfr* in base pairs. The 95% CI was calculated from the unbiased variance of gene diversity as given by Nei and Roychoudhury (1974).

asymmetry in the shape of the selective sweep is apparent with the region of reduced gene diversity extending four to six times further upstream than downstream. An additional microsatellite locus was identified between U75 and U70 to ascertain whether the dip in diversity relative to the sensitive chromosome population at U75 was an anomaly or

suggestive of an additional site under selection. Diversity at this intermittent marker was midway between the two flanking markers in 1996 and had returned to baseline levels in 1999. It was concluded that if selection was occurring at a site proximal to locus U75, it was only very weak and more likely that the dip in diversity was an anomaly.

Pop 1	Pop 2	U350	U250	U150	U125	U100	U75	U70d	U70	U60	U40	U30	U20	U10	MA1	MA2	<i>dhfr</i>	D10	D20	D30	D40	D50	D60	D70	D80	D90	D100	D150	D250	D350	
Sensitive	Triple '96						0.67	0.30		0.53	0.56	0.76	0.75	0.69	0.90	0.44		0.53	-0.42												
Sensitive	Triple '99						0.70			0.37	0.19	0.23	0.50	0.90	0.83			0.28													
Triple '96	Triple '99						-0.26			-0.51		-0.56	-0.52	-0.19	Fixed	0.39		-0.25	0.48			0.29									

FIG. 3.—The significance and direction of differences between ancestral and triple-mutant chromosomes at each locus. Light gray =  $P < 0.05$ ; dark gray =  $P < 0.001$ ; black =  $P < 0.0001$ .

The size and direction of difference in gene diversity at microsatellite loci on the triple-mutant chromosome from 1996, the triple-mutant chromosomes from 1999, and the sensitive chromosomes are summarized in figure 3. Paired comparison of sensitive with triple-mutant chromosomes from 1996 found that diversity values were significantly lower than those of sensitive chromosomes at eight loci spanning a 70-kb region. On the triple-mutant chromosomes from 1999, diversity was significantly lower at just six loci (spanning 50 kb) and the magnitude of the difference in diversity was reduced from 0.53–0.76 in 1996 to 0.19–0.50 in 1999.

#### Time and Its Effects on the Pattern of Diversity Around *dhfr*: Longitudinal Data from KwaZulu Natal

The observed recovery in diversity over time could be caused by recombination through outcrossing with sensitive chromosomes or by the generation of diversity through de novo mutation at microsatellite loci. To further examine the contributory factors influencing the dimensions of the selective sweep, we have used the Wiehe model (Wiehe 1998), modified by Nair et al. (2003), which simulates the combined effects of recombination, mutation, selection, effective population size, and inbreeding to predict the dimension of a selective sweep at equilibrium. To estimate the size of the effective population of *P. falciparum* in KwaZulu Natal, we used gene diversity averaged across the loci on the sensitive (ancestral) chromosomes, which gave a value of 4,904 under an assumption of IAM. The selection coefficient is calculated as 0.048 based on the measured changes in frequency of the triple-mutant allele during the interval between 1996 and 1999 (described in fig. 1). An inbreeding coefficient ( $F$ ) of 0.7 was assumed. This is the rate at which selfing occurs and was estimated from the number of infections in which a single *dhfr* haplotype was detected in the original survey material. As it has been shown that the triple-mutant allele did not arise in Africa but was imported (Roper et al. 2003), rather than assuming an initial frequency reflecting the single mutation event, we used an estimate that reflects the number of migrants ( $m$ ) between populations within the region and therefore the spread by gene flow. We estimated the baseline number of migrants to a SE African population from the  $F_{ST}$  estimate of population differentiation of parasites randomly sampled from Mpumalanga province of South Africa and Tanzania and typed at eight microsatellites dispersed throughout the parasite genome (data not shown). The  $F_{ST}$  of 0.011 estimates 24 migrants per generation from one population to the other. This is comparable to an  $F_{ST}$  estimate between two East African populations separated by approximately 2,000 km, Zimbabwe and Uganda, where  $m$  was determined as 20 (Anderson et al. 2000a). Estimating the initial frequency based on the number of migrants in this manner is a coarse approach as it may be an

overestimate, as this simple approach assumes that all migrants are carrying the triple-mutant allele. However, it provides a relative scale of the differences between populations in the initial starting frequencies in the explanation of the observed selective sweep data. For this analysis, we used an estimate of 20 migrants per generation.

The dimensions of the selective sweep in 1999 very closely resemble those of one predicted on the basis of these parameters (fig. 4). The goodness of fit was measured by the modeled data being not significantly different, by 95% CIs, from the observed data (loci marked by an asterisk in fig. 4). This was true at four out of the five loci within the region U40 to D10 already defined as having reduced diversity significantly below that of the sensitive population, and at 13 out of the total 26 loci, with the majority of the similarity in the portions of the selective sweep closest to *dhfr*. There is an apparent transition between 1996 and 1999 as the competing effects of selection and recombination reach equilibrium.

A selective sweep around *dhfr* has been described in a *P. falciparum* population on the Thailand-Myanmar border. In this population, an additional mutation at codon 164 is present and resistance alleles contain between two and four mutations. Drug selection, which took place between 1976 and 1989, completely displaced the sensitive alleles (Nair et al. 2003). Approximately 90 generations after drug selection ceased, the selective sweep conformed to a shape predicted on the basis of a selection coefficient of 0.1, inbreeding coefficient of 0.8, and a population size of 1,000. The underlying microsatellite mutation rate and basic recombination rate not adjusted for the effects of inbreeding are assumed to be the same between KwaZulu Natal and Thailand. The assumption of the same underlying recombination rate assumes that the genetic cross between parasite lines Hb3 and Dd2 is representative of the recombination rate throughout the total parasite population (Su et al. 1999).

The selective sweeps predicted under these two site-specific sets of circumstances are shown in figure 4. While the model data from parameters estimated for SE Asia appear to also fit the observed data set from 1999, only one out of five loci is not significantly different from the observed data within the region of significantly reduced gene diversity and only 8 out of 26 loci across the whole region (loci marked by a square in fig. 4). We argue that the model using the parameters estimated for KwaZulu Natal is a better fit to the observed data. In KwaZulu Natal, the selective sweep is half as wide and just two-thirds of the depth of that found in SE Asia (Nair et al. 2003; Anderson 2004), and this difference is explained by the lower selection coefficient (fig. 4) and to a lesser degree the five times smaller population size in SE Asia. The effect of reduced selection is that more opportunities for recombination occur between selected and unselected chromosomes. A smaller population size results in a shorter time until fixation of a favorable mutation, thereby also reducing the opportunities for recombination

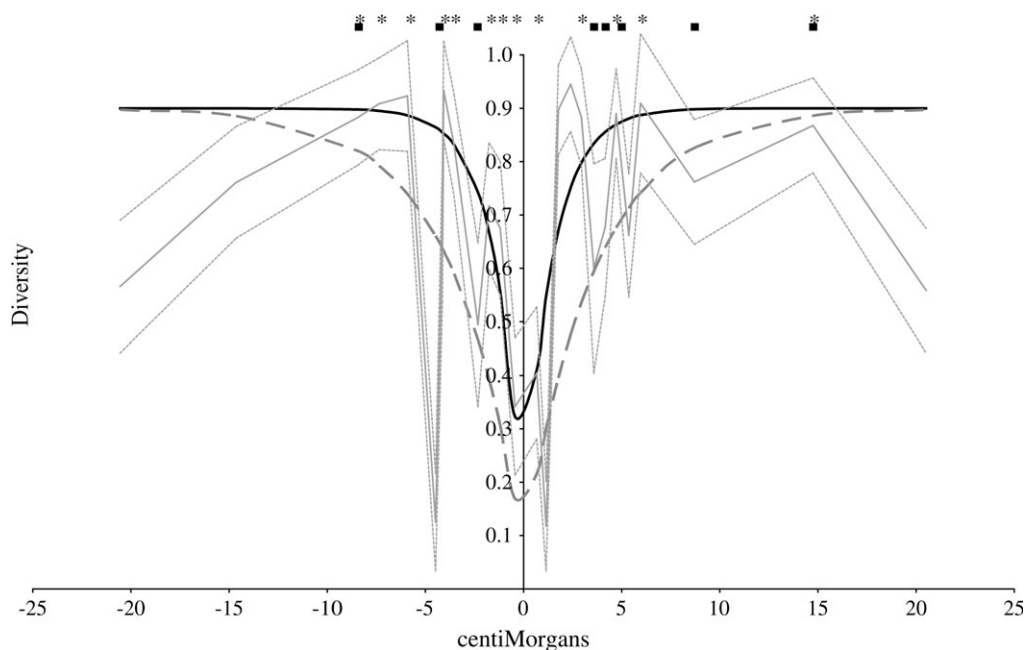


FIG. 4.—Prediction of a line of best fit by a deterministic model of hitchhiking on the chromosome. The gray solid line is the plot of the actual data in KwaZulu Natal in 1999. Dashed lines above this line represent the 95% CI around the actual data. The solid black line and dashed gray line are the model predictions for KwaZulu Natal and Thailand, respectively. Parameters used for KwaZulu Natal were  $N_e = 4,904$ ,  $\mu = 1.59 \times 10^{-4}$ ,  $r = 5.88 \times 10^{-4}$ ,  $s = 0.048$ ,  $m = 20$ ,  $F = 0.7$ . The asterisks above the chart indicate when the modeled data using these parameters fall within the 95% CI of the actual data. Parameters used for the Thai prediction were as reported by Nair et al. (2003). The squares indicate when the modeled data for the Thai population fall within the 95% CI of the KwaZulu Natal actual data.

events (Kim and Stephan 2002). The difference in the magnitude of the selection coefficient between SE Asia and KwaZulu Natal is probably a true reflection of the coverage of treatment and may also reflect the presence of even more highly resistant alleles containing substitutions at codon 164 in SE Asia. There is a spectrum of selection coefficients acting on drug resistance loci in natural populations of *P. falciparum*, and these are reviewed elsewhere (Anderson 2004). Effective recombination rates in Africa are much higher in general, but in the case of KwaZulu Natal, they are very similar to those in Thailand.

#### The Effect of Different Effective Recombination Rates from Low in Mpumalanga, South Africa, to High in Tanzania

Although inbreeding coefficients of KwaZulu Natal and Thailand-Myanmar populations were similar, transmission intensity in the African region spans an enormous range and populations in general have a much higher effective recombination rate there (Babiker et al. 1994; Anderson 2004). To examine the role of effective recombination rates, we compared the selective sweep on triple-mutant chromosomes in KwaZulu Natal with Mpumalanga, where the inbreeding coefficient ( $F$ ) is high, and southern Tanzania, where  $F$  is low (table 1).

The valleys of reduced gene diversity found surrounding the triple-mutant *dhfr* allele present in populations in Tanzania ( $n = 48$ ) and Mpumalanga ( $n = 56$ ) are shown in figure 5a and b. Consistent with previous findings, hitchhiking alleles immediately flanking the triple-mutant alleles in all three populations were identical, indicating that they

are descendants of one ancestral triple mutant (Roper et al. 2003, 2004). In the samples taken from the Tanzanian population, a 50-kb region extending from loci U40 to D10 was found to have significantly lower gene diversity than the sensitive baseline. In Mpumalanga, the region of significantly reduced gene diversity was 70 kb extending from loci U60 to D10.

When the population pairwise comparisons of the gene diversity values for each matched pair of loci were plotted, we found a strong correlation for comparisons between the triple-mutant allele populations of Mpumalanga and KwaZulu Natal 1999 ( $r^2 = 0.934$ ,  $P \ll 0.0001$ ), Tanzania and KwaZulu Natal 1999 ( $r^2 = 0.889$ ,  $P \ll 0.0001$ ), and Mpumalanga and Tanzania ( $r^2 = 0.772$ ,  $P = 0.001$ ). The strong correlation is a good indication of the similarities in the dimensions of the selective sweeps in Tanzania, Mpumalanga, and KwaZulu Natal in 1999. However, correlations between Mpumalanga, Tanzania, and the triple-mutant population from KwaZulu Natal in 1996 were much weaker (Mpumalanga  $r^2 = 0.664$ ,  $P = 0.010$ ; Tanzania  $r^2 = 0.620$ ,  $P = 0.018$ ).

The significant differences between loci and their direction and magnitude are summarized in figure 6. The differences between triple-mutant chromosomes from the three sites are limited. Two significant observations were that the selective sweep from Mpumalanga was larger and, secondly, the KwaZulu Natal 1999 selective sweep had reduced depth at the markers closest to *dhfr* relative to that in Tanzania. The lack of major differences in the dimensions of the three selective sweeps from Mpumalanga and Tanzania in 2001 as compared with that from KwaZulu Natal in 1999 is in contrast to the reduction in size that

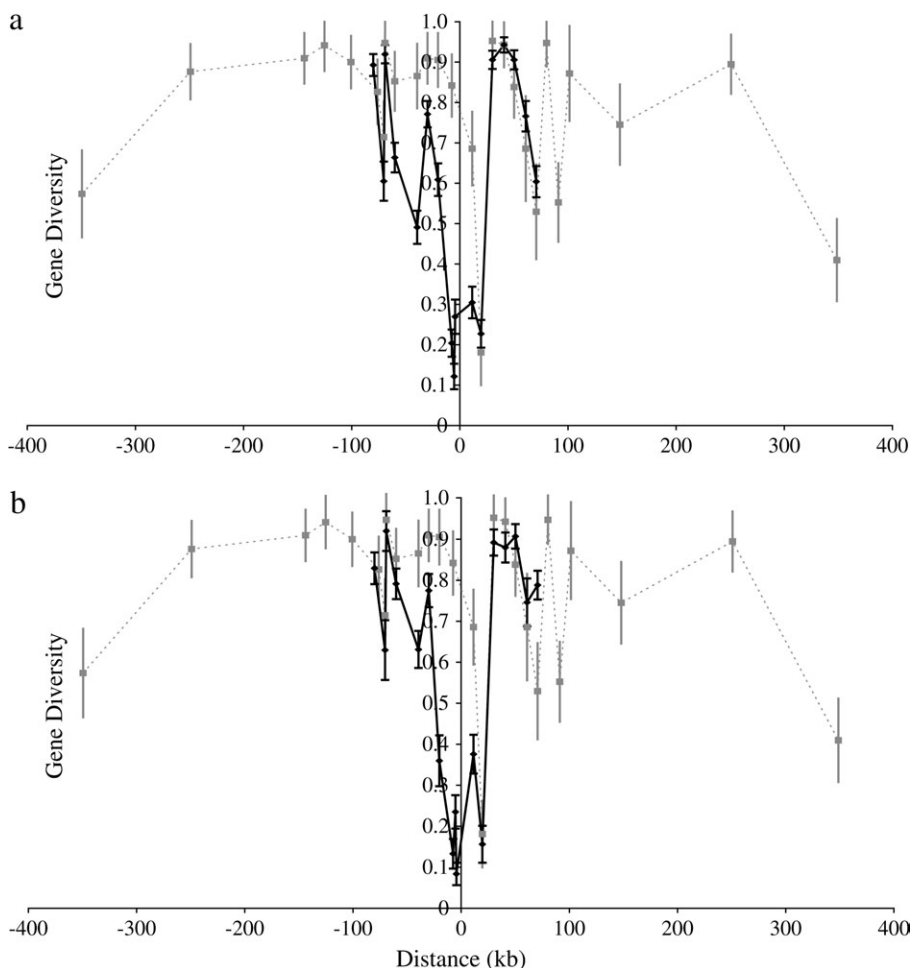


FIG. 5.—The gene diversity at each locus for populations of (a) triple-mutant *dhfr* allele from Mpumalanga and (b) triple-mutant *dhfr* allele from Tanzania. Distances shown are actual physical distances from codon 108 of *dhfr* in base pairs. The 95% CIs were calculated from the unbiased variance of gene diversity as given by Nei and Roychoudhury (1974).

occurred over 3 years from 1996 to 1999 in KwaZulu Natal. We have suggested that a relatively weak selection coefficient versus a high rate of recombination explains the differences in the size of the selective sweeps in KwaZulu Natal between the two time points. Extrapolating the rates of breakdown of the selective sweep between 1996 and 1999, one may have expected the Mpumalangan and Tanzanian derived chromosomes in 2001 to be more than 10 kb smaller than that found in KwaZulu Natal in 1999. By contrast, we find that the Mpumalangan selective sweep is actually significantly wider and that the Tanzanian selective sweep while being the same width is significantly deeper than in 1999. Differences in both width and depth imply that the recombination has not occurred to the same extent in these populations as it has in KwaZulu Natal. The differ-

ences between the populations expected from the differences in EIR and effective recombination rate were not apparent. In addition to EIR, assortative mating through drug pressure killing sensitive parasites in mixed infections could also cause an increase in the inbreeding coefficient.

### The Importance of Gene Flow and Initial Starting Frequencies on Patterns of Diversity

Using the model of expected diversity along a chromosome at equilibrium, we determined a line of best fit to the valleys of reduced gene diversity in Tanzania and Mpumalanga, using population-specific parameters calculated as follows. Inbreeding coefficients of 0.7 and 0.4 were estimated for Mpumalanga and Tanzania, respectively,

Pop 1	Pop 2	U70d	U70	U60	U40	U30	U20	U10	MA1	MA2	D10	D20	D30	D40	D50	D60	D70
Sensitive	Triple Mpu	0.11		0.19	0.37	0.14	0.30	0.64	0.78	0.65	0.38				-0.07		
Sensitive	Triple Tanz			0.23	0.13	0.55	0.71	0.67	0.83	0.83	0.31		0.06				-0.26
Sensitive	Triple '99				0.37	0.19	0.23	0.50	0.90	0.83	0.28						
Triple '99	Triple Mpu			0.17				0.14	-0.12	-0.18	0.10	-0.11				-0.17	
Triple '99	Triple Tanz				-0.14		0.32	0.21	-0.23								-0.11
Triple Mpu	Triple Tanz			-0.13	-0.14		0.25	0.07	-0.11	0.19	-0.07	0.07		0.06			-0.18

FIG. 6.—The significance and direction of differences between ancestral and triple-mutant chromosomes from Mpumalanga, Tanzania, and KwaZulu Natal 1999 at each locus. Light gray =  $P < 0.05$ ; dark gray =  $P < 0.001$ ; black =  $P < 0.0001$ .



from the number of single infections detected at *dhfr* across all samples in the initial surveys. The effective population sizes ( $N_e$ ) were estimated from the average gene diversity of eight microsatellites (Anderson et al. 1999) dispersed throughout the parasite genome (data not shown). The  $N_e$  was determined to be 5,987 in Tanzania and 4,642 in Mpumalanga. These estimates are comparable to KwaZulu Natal and to other SE African populations (Anderson et al. 2000a) and larger than estimated for SE Asia. The selection coefficient of 0.048 is directly measured from frequency changes in KwaZulu Natal (fig. 1) and is the only direct measurement available for Africa at present.

In SE Asia, the initial frequency of the favored allele will reflect the underlying mutation rate of *dhfr*, whereas in SE Africa, it is a reflection of the amount of gene flow between populations. Using the model, we estimated that the numbers of migrants ( $m$ ) required for best fit to the actual data were 1,000 in Tanzania and 200 in Mpumalanga, 50- and 10-fold larger than the estimate based on an  $F_{ST}$  in the region. These may be overestimates because, unlike KwaZulu Natal, the selective sweeps in populations of Tanzania and Mpumalanga are not yet likely to be at equilibrium. Alternatively the selection coefficient used in our simulation may be underestimated. In the case of Tanzania, an increase in the selection coefficient to approximately 0.2 is equivalent to an increase of  $m$  from 20 to 1,000. However, it is unlikely that the coverage of pyrimethamine usage in Tanzania was higher than that in KwaZulu Natal, in Mpumalanga, or on the Thailand-Myanmar border because in Tanzania the first-line treatment at the time of sampling was chloroquine; the EIR is high resulting in acquired malaria immunity and more asymptomatic infections (Snow et al. 1997; Kleinschmidt and Sharp 2001).

In attempting to identify key parameters that explain the dimensions of the selective sweeps in the different sites using the model, it becomes clear that it is not possible to distinguish a clear effect of recombination. However, over and above selection and recombination is an unquantifiable but important influence, namely, the homogenizing effect of gene flow across the entire SE African region. Drug treatment use in the various nation states has historically varied widely, and local frequencies of resistance alleles will reflect this. The number of migrant triple-mutant alleles entering a country will be determined by the movement of people from neighboring states and the frequency of those alleles in the region or country they come from.

Although we have quantified recombination rates, the strength of pyrimethamine selection and its duration cannot be inferred from national treatment policy histories. In Tanzania, sulfadoxine-pyrimethamine (SP) was the recommended second-line treatment at the time of sampling and had been for 18 years. In addition, it is available through private suppliers and is used in self-treatment (Goodman et al. 2004). In northern Tanzania, the frequency of the triple-mutant allele can be as high as 84% (Pearce et al. 2003), reflecting a higher local level of use and a history of resistance (Ronn et al. 1996; Trigg et al. 1997). Gene flow between northern and southern Tanzania is likely to be high (Clyde 1967). A possible explanation for the larger than expected dimensions of the selective sweep is that its initial frequency in the population was generated

by high levels of gene flow from northern Tanzania, effectively reducing the population size and thus opportunities for outcrossing.

The parameters used to fit the line of equilibrium to the selective sweep on the triple-mutant chromosomes in Mpumalanga are feasibly close to those of the actual equilibrium values. The estimate of  $s$  equivalent to the  $m$  of 200 is only slightly higher at 0.06 than the measured  $s$  in KwaZulu Natal (0.048). The drug selection history in Mpumalanga is shorter than that in KwaZulu Natal as SP only became first-line antimalarial in 1997, 9 years after its implementation in KwaZulu Natal.

#### Asymmetric Pattern of Diversity Around *dhfr*

A major characteristic of the selective sweeps around the triple-mutant *dhfr* allele in KwaZulu Natal, Tanzania, Mpumalanga, and SE Asia is a pronounced asymmetry, and interestingly, in every case, the region of reduced gene diversity extends further on the upstream side of *dhfr*. There are a number of competing explanations for this phenomenon, and the most obvious is that rates of recombination differ on either side of the gene. Considerable variation in the distribution of crossover events, differing greatly from their genome average, can reflect recombination hotspots or cold spots as well as other genome features such as proximity to centromeres and telomeres (Barnes et al. 1995; Lichten and Goldman 1995). However, the *dhfr* gene is in the center of chromosome 4, and the distribution of meiotic crossover events on *P. falciparum* chromosomes has been shown to be relatively uniform in the progeny of a laboratory cross (Su et al. 1999).

Alternatively, the asymmetry may be due to the stochastic nature of recombination events during short phases of intense selection. In a model of the pattern of genetic variation along a recombining chromosome, Kim and Stephan (2002) showed that in a population where the time to fixation was short, such as when effective population size ( $N_e$ ) was small, the selective sweep would be asymmetrical around the selected site. The short lineage of the rapidly expanding selected allele decreases the amount of time for recombination events to occur, and the pattern is more of stochastic noise. In SE Asia, pyrimethamine resistance swept to fixation in only 6 years (White 1992). In Africa, the triple mutant is yet to reach fixation, so it is likely that the lineage is long and therefore the asymmetry cannot be easily explained. Intense selection pressure can result in rapid epidemic expansions of resistant parasites in local populations in Africa, and as such, this would facilitate the formation of the asymmetry. The caveat to this hypothesis is that it assumes that the asymmetry is random, and therefore, by chance, it has fallen to the same side of *dhfr* in East Africa as it has in SE Asia.

Recent analysis of the genetic cross between parasite lines Hb3 and Dd2 has identified a 48.6-kb region of chromosome 4 as in complete linkage with the folate salvage phenotype, thought to abrogate the killing effect of the sulfadoxine component of SP (Wang et al. 2004). On this fragment were seven open reading frames including *dhfr*. If there was a resistance-enhancing or a resistance-compensating adaptation in the region upstream of the *dhfr* triple-mutant

allele, such as one favorable for folate salvage, then there would be reason to expect consistent asymmetry in the same direction. However, the changes we observed over time do not support this as the area of lowest diversity became more condensed and mapped onto the *dhfr* gene itself. We observed a gradient in the depth of the selective sweep, which becomes progressively deeper as you get nearer to the gene. We observed no further dips in the levels of diversity along the chromosome that would indicate a putative second site under selection.

### Concluding Remarks

In African populations, there was a consistently narrower selective sweep than that found on the Thailand-Myanmar border, and this was expected because effective recombination rates are in general much higher in Africa due to the higher transmission intensity. We observed that over time in KwaZulu Natal the selective sweep reduced in size toward equilibrium between selection and recombination.

When we compared the three African populations of KwaZulu Natal, Mpumalanga, and Tanzania, we did not observe a clear relationship between the width of the area of reduced diversity and transmission intensity/effective recombination rates. The key factors appear to be the strength of selection, length of time over which drug selection has acted, and the starting frequency of resistance mutants. Although the extent of gene flow is high between populations in this region, resistance levels are heterogeneous because of selection imposed in neighboring countries implementing different treatment policies. Importantly, the founder events in Asia were determined by mutation rates, whereas in Africa, these were dictated by resistance levels in neighboring states, and the contribution of diffusion of resistance alleles in determining starting frequencies has been significant.

The size of the observed selective sweep in Tanzania is contrary to expectation, given the population genetic parameters of the region, but may be explained by large numbers of migrants into that population. Our data underline the importance of gene flow in the spread of resistance between African countries.

### Supplementary Material

Details of the microsatellite primer sequences are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

### Acknowledgments

R.J.P. and C.R. are supported by a Wellcome Trust career development fellowship awarded to C.R. We would like to thank all the people involved in carrying out prevalence surveys in 2000 and collecting the samples in southern Tanzania, which is part of the Interdisciplinary Monitoring Project for Antimalarial Combination Therapy in Tanzania, which is funded by United States Agency for International Development, Centers for Disease Control, and Wellcome Trust, particularly Co-Principal Investigators Salim Abdulla and Peter Bloland. We thank A. Mabuza and his Mpumalanga Malaria Control Programme team,

who collected the samples in 2001 from Steenbok, Mangweni, and Komatipoort clinics as part of the South East African Combination Antimalarial Therapy evaluation, which received partial financial support from the United Nations Development Programme/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases. We would also like to thank A. Keyser for assistance with typing point mutation present at *dhfr* in the Mpumalanga data set.

### Literature Cited

- Anderson, T. J. 2004. Mapping drug resistance genes in *Plasmodium falciparum* by genome-wide association. *Curr. Drug Targets Infect. Disord.* **4**:65–78.
- Anderson, T. J., B. Haubold, J. T. Williams et al. (16 co-authors). 2000a. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Mol. Biol. Evol.* **17**:1467–1482.
- Anderson, T. J., X. Z. Su, M. Bockarie, M. Lagog, and K. P. Day. 1999. Twelve microsatellite markers for characterization of *Plasmodium falciparum* from finger-prick blood samples. *Parasitology* **119**(Pt. 2):113–125.
- Anderson, T. J., X. Z. Su, A. Roddam, and K. P. Day. 2000b. Complex mutations in a high proportion of microsatellite loci from the protozoan parasite *Plasmodium falciparum*. *Mol. Ecol.* **9**:1599–1608.
- Babiker, H. A., L. C. Ranford-Cartwright, D. Currie, J. D. Charlwood, P. Billingsley, T. Teuscher, and D. Walliker. 1994. Random mating in a natural population of the malaria parasite *Plasmodium falciparum*. *Parasitology* **109**(Pt. 4):413–421.
- Barnes, T. M., Y. Kohara, A. Coulson, and S. Hekimi. 1995. Meiotic recombination, noncoding DNA and genomic organization in *Caenorhabditis elegans*. *Genetics* **141**:159–179.
- Barton, N. H. 2000. Genetic hitchhiking. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **355**:1553–1562.
- Bersaglieri, T., P. C. Sabeti, N. Patterson, T. Vanderploeg, S. F. Schaffner, J. A. Drake, M. Rhodes, D. E. Reich, and J. N. Hirschhorn. 2004. Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**:1111–1120.
- Charlwood, J. D., T. Smith, E. Lyimo, A. Y. Kitua, H. Masanja, M. Booth, P. L. Alonso, and M. Tanner. 1998. Incidence of *Plasmodium falciparum* infection in infants in relation to exposure to sporozoite-infected anophelines. *Am. J. Trop. Med. Hyg.* **59**:243–251.
- Clyde, D. F. 1967. *Malaria in Tanzania*. Oxford University Press, Oxford.
- Conway, D. J., C. Roper, A. M. Oduola, D. E. Arnot, P. G. Kremsner, M. P. Grobusch, C. F. Curtis, and B. M. Greenwood. 1999. High recombination rate in natural populations of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **96**:4506–4511.
- Cornuet, J. M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**:2001–2014.
- Cowman, A. F., M. J. Morry, B. A. Biggs, G. A. M. Cross, and S. J. Foote. 1988. 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.
- Dye, C., and B. G. Williams. 1997. Multigenic drug resistance among inbred malaria parasites. *Proc. R. Soc. Lond. B Biol. Sci.* **264**:61–67.
- Gardner, M. J., N. Hall, E. Fung et al. (45 co-authors). 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**:498–511.

- Goodman, C., S. P. Kachur, S. Abdulla, E. Mwageni, J. Nyoni, J. A. Schellenberg, A. Mills, and P. Bloland. 2004. Retail supply of malaria-related drugs in rural Tanzania: risks and opportunities. *Trop. Med. Int. Health* **9**:655–663.
- Kaplan, N. L., R. R. Hudson, and C. H. Langley. 1989. The “hitchhiking effect” revisited. *Genetics* **123**:887–899.
- Kim, Y., and W. Stephan. 2002. Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics* **160**:765–777.
- Kleinschmidt, I., and B. Sharp. 2001. Patterns in age-specific malaria incidence in a population exposed to low levels of malaria transmission intensity. *Trop. Med. Int. Health* **6**:986–991.
- Lichten, M., and A. S. Goldman. 1995. Meiotic recombination hotspots. *Annu. Rev. Genet.* **29**:423–444.
- Liu, K., and S. Muse. 2004. PowerMarker: new genetic data analysis software. Distributed by the author at <http://www.powermarker.net>.
- Nair, S., J. T. Williams, A. Brockman et al. (12 co-authors). 2003. A selective sweep driven by pyrimethamine treatment in south-east asian malaria parasites. *Mol. Biol. Evol.* **20**:1526–1536.
- Nei, M., and A. K. Roychoudhury. 1974. Sampling variances of heterozygosity and genetic distance. *Genetics* **76**:379–390.
- Nosten, F., M. van Vugt, R. Price, C. Luxemburger, K. L. Thway, A. Brockman, R. McGready, F. ter Kuile, S. Looareesuwan, and N. J. White. 2000. Effects of artesunate-mefloquine combination on incidence of *Plasmodium falciparum* malaria and mefloquine resistance in western Thailand: a prospective study. *Lancet* **356**:297–302.
- Palaisa, K., M. Morgante, S. Tingey, and A. Rafalski. 2004. Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep. *Proc. Natl. Acad. Sci. USA* **101**:9885–9890.
- Paul, R. E., M. J. Packer, M. Walmsley, M. Lagog, L. C. Ranford-Cartwright, R. Paru, and K. P. Day. 1995. Mating patterns in malaria parasite populations of Papua New Guinea. *Science* **269**:1709–1711.
- Pearce, R. J., C. Drakeley, D. Chandramohan, F. Mosha, and C. Roper. 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.
- Peterson, D. S., D. Walliker, and T. E. Wellems. 1988. Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in falciparum malaria. *Proc. Natl. Acad. Sci. USA* **85**:9114–9128.
- Quesada, H., U. E. Ramirez, J. Rozas, and M. Aguade. 2003. Large-scale adaptive hitchhiking upon high recombination in *Drosophila simulans*. *Genetics* **165**:895–900.
- Ronn, A. M., H. A. Msangeni, J. Mhina, W. H. Wernsdorfer, and I. C. Bygbjerg. 1996. High level of resistance of *Plasmodium falciparum* to sulfadoxine-pyrimethamine in children in Tanzania. *Trans. R. Soc. Trop. Med. Hyg.* **90**:179–181.
- Roper, C., R. Pearce, B. Bredenkamp, J. Gumede, C. Drakeley, F. Mosha, D. Chandramohan, and B. Sharp. 2003. Antifolate antimalarial resistance in southeast Africa: a population-based analysis. *Lancet* **361**:1174–1181.
- Roper, C., R. Pearce, S. Nair, B. Sharp, F. Nosten, and T. Anderson. 2004. Intercontinental spread of pyrimethamine-resistant malaria. *Science* **305**:1124.
- Smith, J. M., and J. Haigh. 1974. The hitch-hiking effect of a favourable gene. *Genet. Res.* **23**:23–35.
- Snewin, V. A., S. M. England, P. F. Sims, and J. E. Hyde. 1989. Characterisation of the dihydrofolate reductase-thymidylate synthetase gene from human malaria parasites highly resistant to pyrimethamine. *Gene* **76**:41–52.
- Snow, R. W., J. A. Omumbo, B. Lowe et al. (13 co-authors) 1997. Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. *Lancet* **349**:1650–1654.
- Su, X., M. T. Ferdig, Y. Huang, C. Q. Huynh, A. Liu, J. You, J. C. Wootton, and T. E. Wellems. 1999. A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. *Science* **286**:1351–1353.
- Trigg, J. K., H. Mbwana, O. Chambo, E. Hills, W. Watkins, and C. F. Curtis. 1997. Resistance to pyrimethamine/sulfadoxine in *Plasmodium falciparum* in 12 villages in north east Tanzania and a test of chlorproguanil/dapsone. *Acta Trop.* **63**:185–189.
- Wang, P., N. Nirmalan, Q. Wang, P. F. Sims, and J. E. Hyde. 2004. Genetic and metabolic analysis of folate salvage in the human malaria parasite *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **135**:77–87.
- White, N. J. 1992. Antimalarial drug resistance: the pace quickens. *J. Antimicrob. Chemother.* **30**:571–585.
- Wiehe, T. 1998. The effect of selective sweeps on the variance of the allele distribution of a linked multiallele locus: hitchhiking of microsatellites. *Theor. Popul. Biol.* **53**:272–283.
- Wu, Y., L. A. Kirkman, and T. E. Wellems. 1996. Transformation of *Plasmodium falciparum* malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine. *Proc. Natl. Acad. Sci. USA* **93**:1130–1134.

Geoffrey McFadden, Associate Editor

Accepted May 23, 2005