Am. J. Trop. Med. Hyg., 76(6), 2007, pp. 1037–1045

HIGH FREQUENCY OF RECOMBINATION-DRIVEN ALLELIC DIVERSITY AND TEMPORAL VARIATION OF *PLASMODIUM FALCIPARUM MSP1* IN TANZANIA

KAZUYUKI TANABE,* NAOKO SAKIHAMA, INGEGERD ROOTH, ANDERS BJÖRKMAN, AND ANNA FÄRNERT

Laboratory of Malariology, International Research Center of Infectious Diseases, Research Institute for Microbial Diseases, Osaka University, Suita, Osaka 565-0871, Japan; Nyamisati Malaria Research Unit, Dar es Salaam, Tanzania; Unit of Infectious Diseases, Department of Medicine, Karolinska Institute, Karolinska University Hospital, Stockholm, Sweden

Abstract. A major mechanism for the generation allelic diversity in the *Plasmodium falciparum msp1* gene is meiotic recombination in the *Anopheles* mosquito. The frequency of recombination events is dependent on the intensity of transmission. Herein we investigate the frequency of recombination-driven allelic diversity and temporal variation of *msp1* in Rufiji, eastern coastal Tanzania, where malaria transmission is intense. We identified 5′ recombinant types, 3′ sequence types, and *msp1* haplotypes (unique associations of 5′ recombinant types and 3′ sequence types) to measure the extent and temporal variation of *msp1* allelic diversity. The results show that *msp1* haplotype diversity is higher in Tanzania as compared with areas with lower transmission rates. The frequencies of individual polymorphic regions/sites remained stable during the study period. However, the frequency distribution of *msp1* haplotypes varied between 1993 and 1998. These results suggest that frequent recombination events between *msp1* alleles intermittently generate novel alleles in high transmission areas.

INTRODUCTION

The 200-kDa merozoite surface protein-1 (MSP-1) of *Plasmodium falciparum* is a leading vaccine candidate antigen.[1,2] MSP-1 contains at least two regions targeted by host immunity: block 2 near the N terminus and block 17 at the C terminus. Human antibodies against block 2 are associated with protection from clinical malaria in highly endemic areas in Africa.[3] Block 17 encodes a C-terminal 19-kDa polypeptide, a product processed from MSP-1,[4] which confers protection after immunization against challenge with live parasites in animals.[5,6] Sera from individuals living in highly endemic areas contain antibodies against the 19-kDa fragment that inhibit merozoite invasion into red blood cells.[7–9]

MSP-1 exhibits extensive polymorphism,[10,11] which is a potential obstacle to the development of effective vaccines. In animal models, MSP-1 has been shown to be the major antigen involved in inducing "strain-specific immunity," in which the host mounts an immune response that is more effective against the immunizing strain than it is against genetically divergent strains.[12,13] As is the case for other *P. falciparum* antigen genes, *msp1* polymorphism is generated via a number of different mechanisms; point mutations result in single-nucleotide polymorphisms (SNPs), insertion/deletion of repeats cause repeat length polymorphisms, and meiotic recombination involving the exchange of gene fragments between parental alleles produces novel alleles in the progeny. SNPs in *msp1* appear to be stable through time[14] and may be of ancient origin.[15] Repeat-length polymorphisms are common in *msp1*[10,11,16] to the extent that size polymorphism between alleles is widely used as a marker for parasite genotyping.[17] Aside from repeat-length polymorphisms, meiotic recombination is the major mechanism for the generation of *msp1* allelic diversity.[10] Potential recombination sites have previously been mapped to restricted regions within *msp1* (see Figure 1).[10,11] The frequency of recombination in *P. falciparum* is dependent, to a large extent, on the rate of transmission, because meiotic recombination occurs only in the mosquito host. Recombination-driven allelic diversity in *msp1* is expected to be high in areas of intense malaria transmission and lower in areas with less intense transmission dynamics. The validity of this assumption remains to be tested, however, as very few studies have directly measured recombination-driven *msp1* diversity in areas of high transmission.

To investigate the nature and frequency of *msp1* allelic diversity in a highly endemic area, we conducted a study of the prevalence of *msp1* haplotypes in isolates collected 1993, 1998, and 2003 in Rufiji, eastern coastal Tanzania, where malaria transmission is intense and perennial.[18] Our results show that the extent of recombination-driven allelic diversity in *msp1* is higher in Tanzania as compared with areas with lower transmission rates. The frequency distribution of *msp1* haplotypes varied through time, but the frequencies of individual polymorphic regions and sites remained stable throughout the 10-year period of study. These results suggest that frequent recombination events in *msp1* intermittently generate novel *msp1* alleles in a high transmission area.

MATERIALS AND METHODS

Study area and sample collection. *P. falciparum* isolates were collected during malaria surveys from individuals living in Nyamisati village in the Rufiji River Delta, 150 km south of Dar es Salaam, in eastern coastal Tanzania in February and March 1993 ($N = 120$), 1998 ($N = 132$), and January 2003 ($N = 104$). Almost all samples were taken from asymptomatic donors of all ages with a mean age of 14.2 years (range, 1–78) and 16.8 years (range, 1–63) in 1993 and 1998, respectively, and from those aged 10–19 years (mean, 13.8 years) in 2003. Malaria in the study area was holoendemic with perennial transmission with some increase during the two rainy seasons, April to June and December.[18] An annual entomological inoculation rate is not available for the study area, but it is known to be in the range of 94 to 667 infective mosquito bites/person/year in eastern Tanzania.[19] Insecticide-impregnated bed nets were distributed to all houses in the village in 1999. Slide-positive parasite rates were recorded for the 1993 sample (46%) but were unavailable for the other sampling dates because of technical reasons. However, para-

* Address correspondence to K. Tanabe, International Research Center of Infectious Diseases, Research Institute for Microbial Diseases, Osaka University, Suita, Osaka, Japan. E-mail: kztanabe@biken.Osaka-u.ac.jp

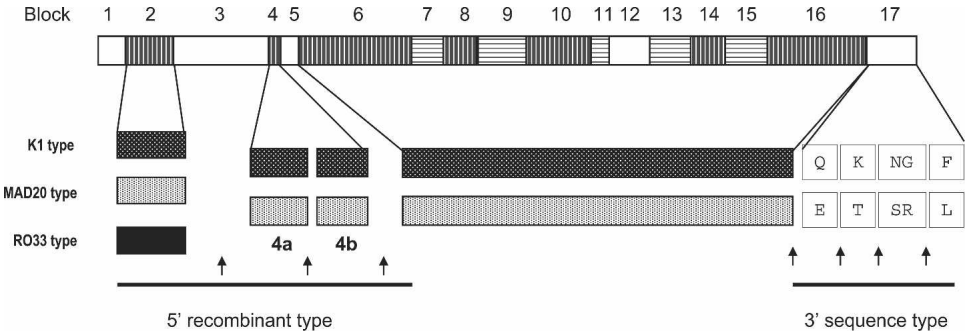

FIGURE 1. Determination of *P. falciparum msp1* haplotype, a unique association of 5′ recombinant types and 3′ sequence types. *msp1* is divided into 17 blocks, in which inter-allele conserved, semi-conserved, and variable blocks are indicated by open, horizontally hatched, and vertically hatched columns, respectively. For K1-type, MAD20-type, and RO33-type variable blocks, sequences are represented by densely toned, half-toned, and black bars, respectively. The 5′ recombinant types were determined by PCR amplification of blocks 2 to 6 using allelic type–specific primers of blocks 2 and 6, followed by nested PCR for blocks 4a and 4b using allelic type-specific primers of blocks 4a and 4b. Five amino acid substitutions in block 17 are indicated by the one-letter codes. The 3′ sequence type is the combination of those residues. Potential recombination sites are shown by arrows.

site positive rates in sampled people as checked by high-sensitivity PCR-based parasite detection (*msp1* typing method used in this study) were 78%, 77%, and 44% in 1993, 1998, and 2003, respectively.

All samples were collected after informed consent had been obtained from the donors or their guardians. Venous blood was collected into EDTA-containing tubes and stored at –20°C. Individuals with signs of clinical disease, i.e., fever and parasites, were treated with Fansidar. Parasite genomic DNA was extracted using the QIAamp DNA Blood Kit (Qiagen, Hilden, Germany). The volume of extracted DNA template was adjusted to be equivalent to the original blood volume. Ethical approval was obtained from the Ethical Committee of the National Institute for Medical Research, Tanzania, and the Ethical Committee of the Karolinska Institute, Sweden. Data previously obtained from clinical samples in Mae Sot in northwestern Thailand in 1995 and from survey samples in Guadalcanal Island in the Solomon Islands in 1994–1996 were used for geographical comparison.[20] We used clinical isolates ($N = 111$) from patients who attended a malaria clinic in Mae Sot in northwestern Thailand in 1995.[21] The mean age of the donors in Thailand was 24.6 years. A total of 90 isolates were collected in north Guadalcanal, the Solomon Islands: 40 clinical isolates from outpatients with a mean age of 18.3 years of a hospital in Honiara City and 50 isolates from four villages (Kaotave, Tadhimboko, Nugalitav, and Ruavatu).[20] In these rural villages, samples were collected in most cases from parasite-positive asymptomatic individuals during malariometric surveys, and most of the donors were primary-school children aged 8 to 15 years.

Determination of *msp1* polymorphisms. *P. falciparum msp1* (a 5-kb single-copy gene) consists of 17 distinct sequence blocks, according to the degree of sequence similarity among alleles (Figure 1).[10] Sequence variation in *msp1* is principally dimorphic (either one or the other of two major allelic types: K1 type and MAD20 type) in all variable blocks except block 2, which is trimorphic (represented by K1, MAD20, and RO33 types). To monitor the recombination-driven allelic diversity of *msp1*, we divided the gene into three regions: a 5′ 1.1-kb region (blocks 2 to 6), a central 3.5-kb region (blocks 6 to 16), and a 3′ 0.4-kb region (block 17), in which potential recombination sites have been mapped to the 5′ and 3′ regions (Figure 1). No recombination events occur in blocks 6 to 16.[16,21,22] The *msp1* haplotypes are thus defined as unique associations of 5′ recombinant types and 3′ sequence types in this study.

The 5′ recombinant types are defined as unique associations of allelic types of variable blocks 2, 4a, 4b, and 6. In total, 24 distinctive 5′ recombinant types are distinguishable: i.e., $24 = 3 \times 2 \times 2 \times 2$ (three allelic types designated as K, M, and R in block 2 and two allelic types designated as K and M in blocks 4a, 4b, and 6). The 5′ recombinant types were determined by our methods described previously.[20] In brief, they were determined by the following two steps: (i) first-round PCR to determine allelic types of blocks 2 and 6 using allelic-type-specific primers, and (ii) nested PCR to determine allelic types of blocks 4a and 4b (≈ 100 bp) using the first-round PCR products and allelic-type–specific primers. The PCR method allows us to determine the rate of multiple 5′ recombinant-type infections, here referred to as "polyinfection rate," and the mean number of 5′ recombinant-type infections per isolate (MORT). One microliter of template DNA was used for first-round PCR. 5′ Recombinant types were fully determined in 94 of 120 Tanzanian isolates collected in 1993, in 102 of 132 isolates in 1998 samples, and in 46 of 104 isolates in 2003. Thus, 68% (242/356) were PCR-positive in all samples obtained through malariometric surveys, indicating that our data represent a *P. falciparum* population in the study area.

The nucleotide sequence of block 17, which encodes the C-terminal 19-kDa polypeptide, was determined by direct sequencing after amplification of the full-length *msp1*. To see associations of 3′ sequence types (block 17 sequences) and 5′ recombinant types (blocks 2 to 6), only those isolates having a single 5′ recombinant type (i.e., mono-infection) were selected for further analysis. (We did not use cloning of the full-length *msp1* gene because artificial recombination readily occurs during amplification and cloning when samples with mixed genotypes are used.[22]) Because the number of isolates with mono-infections was limited in our Tanzanian samples, we increased the number of mono-infection samples by diluting genomic DNA templates by 20-fold. 5′ Recombinant types were again determined for the diluted samples, and those with a single 5′ recombinant type were selected. The

numbers of isolates sequenced were 38, 23, and 13 in samples collected in 1993, 1998, and 2003, respectively. No significant difference in the frequency distribution of 5′ recombinant types was found between undiluted original samples and diluted samples, indicating no bias of sampling after dilution (not shown). Amplification of the full-length *msp1* was first done with primers UPF1 (5′-GGCTAATGTAAAATGCAAAAATAAATGT) and DWR1 (5′-ACATGACTAAAATATCACTATTCCTGT) in a 20-μL reaction mixture containing 1 μL of template genomic DNA for 37 cycles using LA-*Taq* (TaKaRa, Tokyo, Japan). Two microliters of 10-fold diluted PCR products were amplified by nested PCR using primers UPF3 (5′-AATAAATGTATACATATTTTTGCTAAGTCA) and DWR3 (5′-TTAAGGTAACATATTTTAACTCCTACA) for 20 cycles. The PCR product was purified using the QIA Quick PCR purification kit (Qiagen) and directly sequenced from both directions using primers C17aFs (5′-CAAG(G/A)TATGTTAAACATTTCACAACA) and DWR3 with the BigDye Terminator Cycle sequencing kit (version 3.1) on an automated multicapillary ABI 3100 sequencer (Applied Biosystems, Foster City, CA). Sequences were verified by re-sequencing the PCR products independently amplified from the same DNA. To date, five major amino acid changes have been identified in block 17 from various geographic areas (E or Q at amino acid residue 1644; T or K at 1691; SR or NG at 1700–1701; and L or F at 1716; the positions are numbered according to Ref. 15) (Figure 1).[21,23] Hereafter, we refer to combinations of these residues as 3′ sequence type.

Unique associations of 5′ recombinant types and 3′ sequence types are referred to as *msp1* haplotypes. Partial sequencing of blocks 2 to 6 of the PCR amplicons (full-length *msp1*) confirmed 5′ recombinant types determined by PCR-based typing (Tanabe, unpublished). This indicates that our analysis of linkage between polymorphisms in the 5′ region and 3′ region is not affected by artificial recombination.

Statistical analyses. Frequency distributions of *msp1* 5′ recombinant types, 3′ sequence types, and *msp1* haplotypes were compared using the χ^2 test with Yates correction and Fisher's exact test for data sets fewer than 5. Differences in mean number of 5′ recombinant types per isolate (MORT) were tested for significance using a two-tailed Mann-Whitney *U* test. The diversity level of *msp1* haplotypes was expressed in two ways: (i) relative frequency of the number of unique *msp1* haplotypes per total number of *msp1* haplotypes, and (ii) expected heterozygosity (*h*). *h* and its variance were calculated as previously described.[20] Differences in the relative frequency were tested by *t* test. The frequency of recombination events in *msp1* was inferred from analysis of linkage disequilibrium within and between polymorphic blocks 2 to 6 and polymorphic sites in block 17. To assess linkage disequilibrium within *msp1*, pairs of polymorphic blocks 2, 4a, 4b, and 6 and polymorphic sites in block 17 were subjected to an R^2 test as described elsewhere.[24] Non-informative pairs (frequency < 10% in a polymorphic block or nucleotide site) were excluded from the R^2 test. Significance of linkage disequilibrium was assessed using the χ^2 test with Yates correction or two-tailed Fisher's exact probability test. A $P < 0.05$ was considered statistically significant.

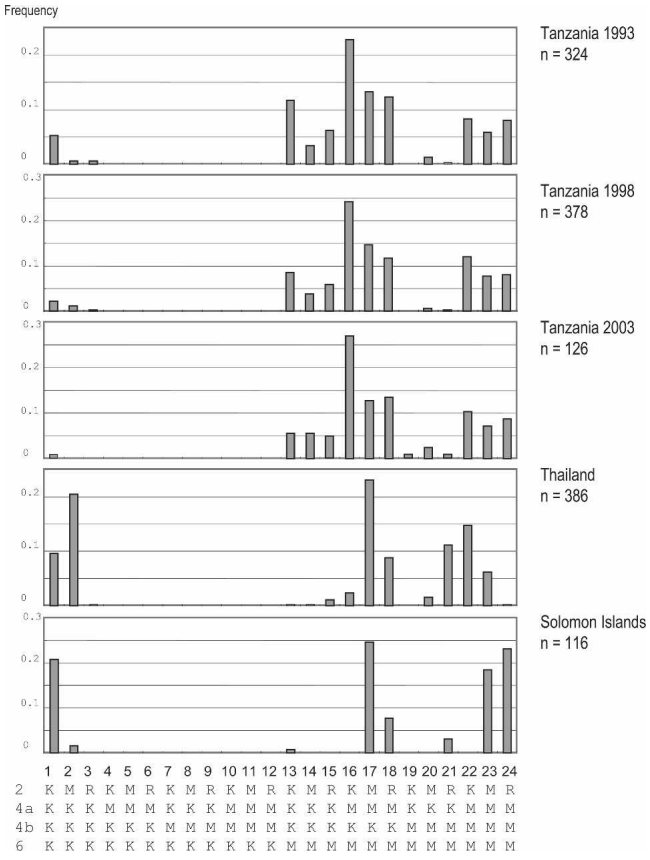

FIGURE 2. Frequency distribution of *P. falciparum msp1* 5′ recombinant types in Tanzania. Twenty-four distinct types—unique associations of allelic types in variable blocks 2, 4a, 4b, and 6—are shown at the bottom of the figure. Data from Thailand and Solomon Islands are from Sakihama et al.[20]

RESULTS

***msp1* 5′ recombinant types (blocks 2 to 6).** The frequency distributions of *msp1* 5′ recombinant types are shown in Figure 2. Types #1 to #12 are those with K1 allelic type in block 6, and types #13 to #24 are those with MAD20 allelic types. Most of the Tanzanian isolates were MAD20 allelic type in block 6 in 1993, 1998, and 2003. The overall pattern of fre-

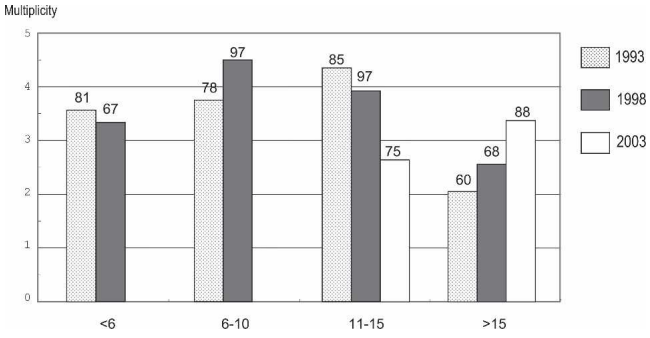

FIGURE 3. Age distribution of mean number of 5′ recombinant-type infections per isolates (MORT) in Tanzania. Ages are categorized into four classes: < 6, 6–10, 11–15, and > 15 yrs. Percentage of multiple 5′ recombinant infections is shown above each bar. Total numbers of isolates are 87, 95 and 44 in 1993, 1998 and 2003, respectively.

quency distribution of 5′ recombinant types was very similar from 1993 to 2003. The frequency distribution of 5′ recombinant types in Rufiji did not differ significantly from that reported previously in Tanga, northeastern Tanzania.[25] Tanzania was, however, significantly different from other geographic areas: Thailand and Solomon Islands ($P < 10^{-10}$), where frequencies of those types having K1 type in block 6 were substantially higher (19% in Solomon Islands and 30% in Thailand) as compared with Rufiji (< 7%).

Rates of multiple infections of the 5′ recombinant types (polyinfection rate) were 76.6%, 87.4%, and 78.3% in 1993, 1998, and 2003, respectively, and the mean number of 5′ recombinant types per isolate (MORT) was 3.48, 3.76, and 2.74, respectively. The reduction of MORT from 1998 to 2003 was significant ($P = 0.008$, Mann-Whitney U test). Both polyinfection rate and MORT are considerably higher in Tanzania than in the Solomon Islands (35.4–60.7% for polyinfection rate and 1.41–1.73 for MORT in Solomon Islands[20]). In Thailand, the polyinfection rate was 96.3% and MORT was 3.61, a level comparable to that observed in Tanzania. Thai isolates, however, were obtained from symptomatic patients, whereas Tanzanian isolates were from asymptomatic carriers, thus making direct comparison somewhat difficult. (There was no significant difference in polyinfection rate and MORT between individuals with clinical malaria and those with asymptomatic malaria in the Solomon Islands.[20])

There was a noticeable difference in age distribution of MORT from Tanzania (Figure 3). In 1993, MORT increased from age group < 6 years to age group 11–15 years and thereafter declined. MORT was significantly lower in age group > 15 years than other age groups ($P = 0.035$ against < 6 years, $P = 0.001$ against 6–10 years, and $P = 0.003$ against 11–15 years). In 1998, a peak MORT was observed in age group 6–10 years, followed by a significant reduction in age group > 15 years ($P = 0.003$). In contrast to the reduction in age group > 15 years in 1993 and 1998, MORT increased from age group

TABLE 1
Frequency distribution of *P. falciparum msp1* haplotypes in Tanzania

5′ Recombinant type	3′ Sequence type					Total	No. of *msp1* haplotypes
	QKNGL	QKNGF	EKNGL	EKNGF	ETSRL		
Tanzania 1993							
KKKK	1	0	0	0	0	1	
KKKM	2	0	1	0	0	3	
KMKM	2*	3	7	0	1	13	
KMMM	0	1 (QKSGF)	1	0	0	2	
MMKM	0	0	5	0	0	5	
(M/R)MMM	1	0	0	0	0	1	
RKKM	0	0	1	0	0	1	
RMKM	1	3	2*	2	0	8	
RMMM	3	0	0	0	1	4	
Total	10	7	17	2	2	38	20
Tanzania 1998							
MKKK	1	0	0	0	0	1	
KMKM	2	1	8	0	0	11	
KMMM	1 (QTSRL)	1	1	0	0	3	
MKKM	0	0	1	0	0	1	
MMKM	1	1	1	0	1+1 (EKSRL)	5	
RMKM	0	1	0	0	1	2	
Total	5	4	11	0	3	23	15
Tanzania 2003							
KKKK	1	0	0	0	0	1	
KKKM	1	0	0	0	0	1	
KMKM	3	2	2	0	0	7	
MMKM	0	0	0	0	1	1	
RMKM	2*	1	0	0	0	3	
Total	7	3	2	0	1	13	9
Thailand							
KKKK	2	0	0	0	0	2	
MKKK	9	0	4	0	0	13	
KMKM	0	0	0	0	1	1	
MMKM	0	2	12	0	2	16	
RMKM	0	0	3	0	0	3	
RKMM	0	0	3	0	1 (ETSGL)	4	
KMMM	1	1	2	0	3	7	
MMMM	0	0	1	0	1	2	
Total	12	3	25	0	8	48	16
Solomon Islands							
KKKK	9	0	0	0	0	9	
MMKM	1	0	0	0	13	14	
RMKM	0	0	0	0	1	1	
MMMM	0	1	5	0	3	9	
RMMM	0	0	0	0	14	14	
Total	10	1	5	0	31	47	8

Identical *msp1* haplotypes shared between 1993 and 1998 are boxed.
* A variant having a substitution from S to N at 1699 is included.

TABLE 2
Diversity of *P. falciparum msp1* haplotype in Tanzania

	No. of samples	No. of *msp1* haplotypes	Relative frequency	*P* value vs. Thailand vs. *Solomon*	$h \pm SE$*	*P* value vs. Thailand vs. *Solomon*
Tanzania						
1993	38	20	0.53	0.040	0.94 ± 0.02	0.212
				0.002		*0.0002*
1998	23	15	0.65	0.007	0.89 ± 0.06	0.84
				0.006		*0.21*
2003	13	9	0.69	0.015	0.94 ± 0.05	0.515
				0.002		*0.022*
Thailand†	52	16	0.31		0.89 ± 0.03	
Solomon Islands	47	8	0.17		0.80 ± 0.03	

* *h*, expected heterozygosity as an index of haplotype diversity.[20]
† Data from Sakihama et al.[20]

11–15 years to age group > 15 years in 2003, but this trend was not statistically significant. (In 2003, sampling was limited to those of age > 10 years for technical reasons in the survey, and therefore MORT in age groups < 6 years and 6–10 years was not shown.) Polyinfection rates also showed similar patterns of age dependency. A sharp fall was noted from 11–15 years to > 15 years: 85% to 60% in 1993 and 97% to 68% in 1998.

3′ Sequence polymorphism (block 17). Five major nucleotide polymorphisms in block 17, all resulting in amino acid replacements, were observed in Tanzanian isolates ($N = 74$). We obtained 10 unique 3′ sequence types: Q-K-NG-L, Q-T-SR-L, Q-K-NG-F, Q-K-SG-F, E-K-NG-L, E-K-NG-F, E-T-SR-L, and E-K-SR-L (Table 1). In addition, minor variants showing Q-K-NNG-L ($N = 2$) and E-K-NNG-L ($N = 1$) were also observed, where the underlined "N" are substitutions for S at 1699, as detected earlier.[21,23] The number of 3′ sequence types was 5 and 4 in Thailand ($N = 48$) and Solomon Islands ($N = 47$), respectively.

Distribution and diversity of *msp1* haplotypes. The numbers of distinct *msp1* haplotypes were 20 in 38 isolates in 1993, 15 in 23 isolates in 1998, and 9 in 13 isolates in 2003 (Table 2). The *msp1* haplotype diversity, as expressed by relative frequency of the number of unique *msp1* haplotypes per total number of samples, was high in Tanzania in 1993 to 2003 (0.53–0.69) (Table 2). These levels were significantly higher than the level observed in Thailand (0.31; $P < 0.04$) and Solomon Islands (0.17; $P < 0.006$). Rare *msp1* haplotypes with a frequency of < 5% were abundant in Tanzania as compared with Thailand and Solomon Islands: 21/26 haplotypes (81%) in Tanzania in 1993 and 1998, 10/16 haplotypes (63%) in Thailand, and 3/8 haplotypes (38%) in Solomon Islands. Expected heterozygosity (h) was also high from 1993 to 2003 (Table 2). The difference in h reached statistical significance in 1993 and 2003 between Tanzania and Solomon Islands but not between Tanzania and Thailand.

Temporal variation in *msp1* polymorphisms. The frequencies of polymorphisms in polymorphic blocks 2, 4a, 4b, and 6 and five major polymorphic nucleotide sites in block 17 were compared from 1993 to 2003 (Table 3). A frequency variation was only observed in block 4a. Pairwise comparisons were also made between 1993 and 1998 ($P = 0.71$) and between 1998 and 2003 ($P = 0.06$). In contrast to the stable frequencies of individual polymorphisms, the frequency distribution of *msp1* haplotypes was clearly different between 1993 and 1998 (Figure 4) (χ^2 test, $P = 0.001$), indicating temporal variation of *msp1* haplotypes during this 5-year interval. (Rare *msp1* haplotypes were excluded from analysis: $N = 4$ in 1993 and $N = 2$ in 1998; see Table 1.) Among 26 distinct haplotypes found in 1993 and 1998 in a total of 61 isolates, only six haplotypes ($N = 35$) were shared between 1993 and 1998. Because of limited numbers of samples, a comparison with samples collected in 2003 was not made. The frequency distribution of *msp1* haplotypes in Tanzania was considerably different from that of Thailand and Solomon Islands (χ^2 test, $P < 10^{-10}$).

Linkage disequilibrium in *msp1*. To determine the frequency of recombination events in *msp1*, we performed linkage disequilibrium (LD) analysis, in which pairs of four polymorphic blocks (blocks 2, 4a, 4b, and 6) and four polymorphic sites were analyzed. Two sites at 1700 and 1701 in block 17 were always linked, and so they were combined for LD analysis. LD was undetectable in most pairs in Tanzania (Figure 5). Only one pair of 10 informative pairs in 1993 ($N = 37$) and one pair of 15 informative pairs in 1998 ($N = 23$) were significant. These pairs were within block 17. LD analysis was not carried out on samples from 2003, due to limited numbers ($N = 13$). These results indicate that the frequency of recombination events in *msp1* is high in the Tanzanian populations. In contrast, in Thailand and Solomon Islands 12 of 21 pairs

TABLE 3
Stable frequency of polymorphism in *Plasmodium falciparum msp1* in Tanzania

Block	Polymorphic type	*n* (frequency) 1993	1998	2003	*P* value
2	K1	156 (0.481)	176 (0.466)	56 (0.444)	0.869
	MAD20	79 (0.244)	104 (0.275)	35 (0.278)	
	RO33	89 (0.275)	98 (0.259)	35 (0.278)	
4a	K1	95 (0.293)	84 (0.222)	26 (0.206)	0.048
	MAD20	229 (0.707)	294 (0.778)	100 (0.794)	
4b	K1	247 (0.762)	271 (0.717)	88 (0.698)	0.261
	MAD20	77 (0.238)	107 (0.283)	38 (0.302)	
6	K1	21 (0.065)	13 (0.034)	1 (0.008)	0.06*
	MAD20	303 (0.935)	365 (0.966)	125 (0.992)	
17	1644:Q†	16 (0.432)	9 (0.391)	9 (0.692)	0.339
	1644:E	21 (0.568)	14 (0.609)	4 (0.308)	
	1691:T	2 (0.054)	3 (130)	1 (0.077)	0.775
	1691:K	35 (0.946)	20 (0.870)	12 (0.923)	
	1700-01:SR	3 (0.081)	4 (0.174)	1 (0.077)	0.772
	1700-01:NG	34 (0.919)	19 (0.826)	12 (0.923)	
	1716:L	28 (0.757)	19 (0.826)	10 (0.769)	0.926
	1716:F	9 (0.243)	4 (0.174)	3 (0.231)	

* Comparison between 1993 and 1998. Frequency in 2003 was not informative.
† Positions are after Miller et al.[11]

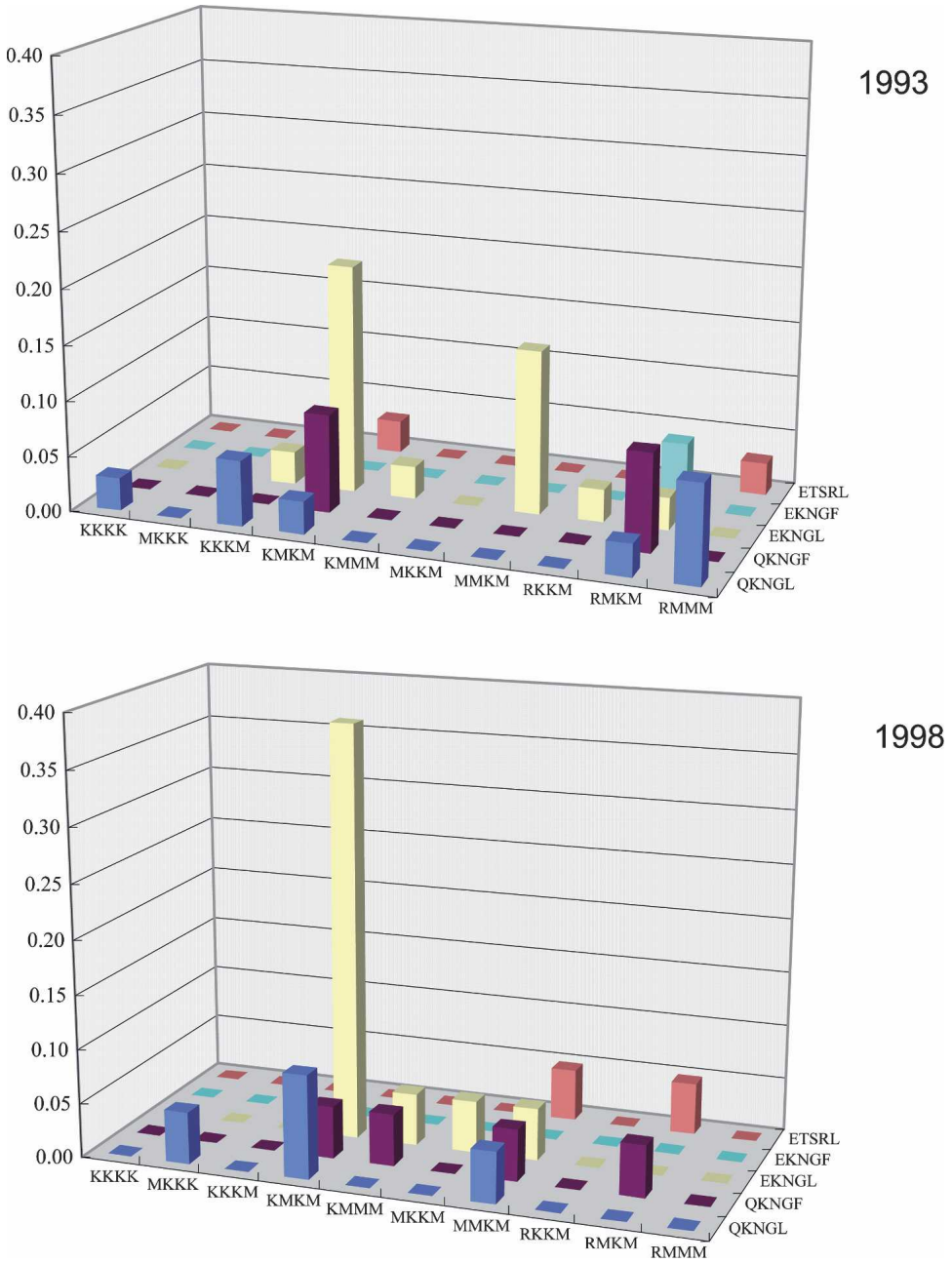

FIGURE 4. Temporal variation in frequency distribution of *P. falciparum msp1* haplotypes between 1993 and 1998. *msp1* haplotypes are unique associations of 5′ recombinant types (*x* axis) and 3′ sequence types (*y* axis). Frequencies are shown on the vertical axis.

and 19 of 21 pairs showed LD, indicating limited or little recombination in those areas.[20,24]

DISCUSSION

Intragenic meiotic recombination in the mosquito is a major mechanism of generation of allelic variation in *P. falciparum msp1*. The frequency of recombination in *P. falciparum* generally depends on the intensity of malaria transmission, which varies greatly in different endemic areas.[26,27] In areas of Africa experiencing high perennial transmission, the entomological inoculation rate (the number of infective mosquito bites per person per year) can reach several hundred,[19] whereas it is at least 2 orders of magnitude lower in areas of low and seasonal transmission such as Southeast Asia. Thus, the recombination-driven allelic diversity of *msp1* may be assumed to be higher in an intense transmission area than in a low transmission area. The present study is the first to measure the recombination-driven allelic diversity of *P. falciparum msp1* in Africa. The results demonstrate that the diversity of *msp1* haplotypes in Tanzania is high compared with areas of lower transmission such as Southeast Asia and Melanesia.[20] In the present study, geographic comparisons of *msp1* diversity were performed using a PCR-based typing method, which may lead to underestimation of the frequency of recombination events. Nevertheless, we observed a sub-

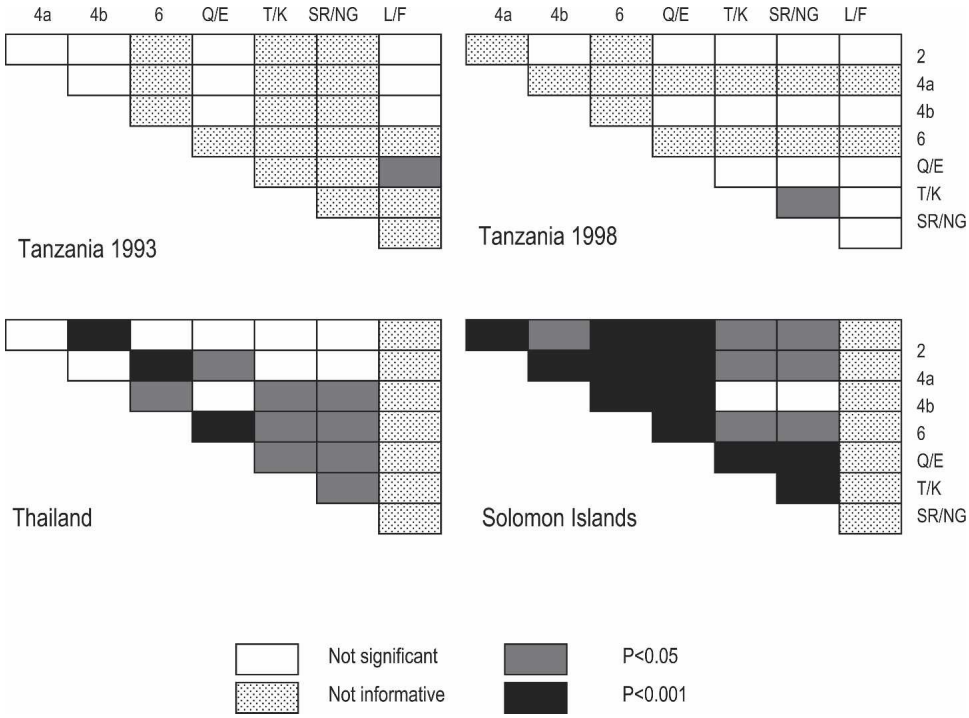

FIGURE 5. Linkage disequilibrium in *P. falciparum msp1* in populations from Tanzania. Pairs of polymorphic blocks 2, 4a, 4b, and 6 and four polymorphic sites (Q/E, T/K, SR/NG, and L/F) in block 17 were subjected to the R^2 test. Non-informative pairs (frequency < 10% in a polymorphic block or nucleotide site) were excluded from the R^2 test. Data from Thailand and Solomon Islands are from Sakihama et al.[20]

stantially high frequency of recombination-driven allelic diversity of *msp1,* suggesting that the extent of recombination-driven allelic diversity of *P. falciparum msp1* is much higher in Africa than we observed.

Although the intensity of transmission is a major factor determining *msp1* allelic diversity, other factors may also be important. These factors include, but are not necessarily limited to, the rate of multiple-genotype infections (polyinfection rate), the mean number of 5′ recombinant type infections per isolate (MORT), and the prevalence of *msp1* haplotypes as well as the parasite-positive rate in a given area. In the Solomon Islands, where the transmission rate is comparable to that of Africa, *msp1* allelic diversity is considerably lower than in Tanzania (Table 3). The polyinfection rate, MORT, and *msp1* haplotype prevalence are relatively limited in the Solomon Islands compared with Tanzania, and therefore the frequencies of out-crossing may be relatively low, resulting in the limited allelic diversity of *msp1* observed in this area.

The frequency distribution of *msp1* haplotypes varied in Nyamisati village between 1993 and 1998. During the same period, frequencies of individual polymorphisms in four polymorphic blocks (blocks 2 to 6) and 4 polymorphic sites (in block 17) remained stable. These two findings appear to contradict each other. However, they are readily reconciled when frequent recombination events are taken into consideration. We observed little linkage disequilibrium in *msp1* in 1993 and 1998, suggesting frequent recombination events in the study area. Therefore, we consider it highly probable that frequent recombination events generate novel *msp1* haplotypes (while simultaneously breaking down previously existing haplotypes), resulting in a temporal variation in their frequency distribution. This explanation is supported by a previous study that showed a rapid decline of linkage disequilibrium along a map distance in *msp1* in highly endemic areas of Africa.[28] Temporal variations in *msp1* polymorphisms in relatively short periods have been reported in Brazil.[29] Epidemic propagations of parasite populations bearing discrete *msp1* alleles along with human movements have been suggested as a likely reason for such temporal variations. Recombination events may play a minor role, if any, in the temporal variation of *msp1* allelic diversity in low transmission areas.

Variation of the frequency distribution of *msp1* haplotypes through time has important implications regarding the parasite's ability to evade the host's immune response. In highly endemic areas, children gradually gain protective immunity to malaria after repeated infections. Although the mechanisms that generate this protective immunity are little understood, it is believed that protective immunity is acquired by cumulative immune responses to multiple antigenic variants after repeated infections.[30–32] Therefore, the extent and prevalence of antigen diversity in a local area is important for the acquisition of protective immunity. MSP-1 is highly immunogenic and induces antibody responses to the entire MSP-1 molecule.[33] Antibodies specific to different regions of MSP-1 inhibit, when combined, parasite growth in an additive manner.[33] Individuals living in endemic areas raise serum antibodies against MSP-1 in an age-dependent manner.[34] The intermittent appearance of novel *msp1* alleles generated by meiotic recombination would produce a number of novel tertiary structure-associated combinational epitopes, and would therefore be likely to induce "epitope"-specific immunity even when frequencies of individual polymorphic blocks and sites are stable. Human antibodies that inhibit merozoite invasion into red cells are known to recognize conformational epitopes.[9] We consider, therefore, that frequent recombination-driven generation of novel *msp1* alleles may affect the

efficiency of acquiring "strain"-specific immunity in highly endemic areas.

In the context of strain-specific immunity, our observation of a significant reduction of MORT from 1993/1998 to 2003 deserves attention. During this period, the age group displaying the highest MORT shifted from those of ages 6–10 years (highest MORT in 1998) to those > 15 years (highest MORT in 2003). This trend was also seen in the polyinfection rates. The reason for this shift is unknown, but it is possibly related to the introduction of insecticide-treated bed nets (ITNs) to the study village in 1999. ITNs have previously been shown to reduce malaria infections substantially in Tanzania.[35] It is also possible that the establishment of a health clinic with continuous monitoring of malaria infections and provision of early treatment of patients contributed to an overall reduction of the mean number of multiple *msp2* genotype infections.[36] The shift of the peak of MORT toward older age groups may be explained in terms of the acquisition of strain-specific immunity. Measures such as ITNs and better health-care facilities will effectively reduce transmission in the areas in which they are deployed. Reduced transmission could lead to an increase in the time it takes an individual to contract, and therefore to develop immunity to, all the different strains present in the area. This would lead to a shift in the peak of MORT to older individuals, as observed in this study. Similarly, the overall reduction of multiple infections may also be a function of reduced transmission.

MSP-1 induces protective antibody responses in individuals living in highly endemic areas.[3,8,9] It may be argued that *msp1* polymorphism is maintained by immune selection, and hence rare polymorphisms increase in frequency over predominant polymorphisms because of the low rates of acquired immunity against them. However, the present study revealed a very stable frequency distribution of *msp1* polymorphisms throughout the period of study (10 years) in Tanzania. Polymorphism in *msp1* has previously been shown to remain stable over a study period of 7 years in the Gambia as determined by typing using monoclonal antibodies.[37] We propose, therefore, that *msp1* polymorphism is not subject to frequency-dependent immune selection.

In conclusion, the present study demonstrates that allelic diversity of *msp1* is higher in Tanzania than in Thailand and the Solomon Islands and suggests that intragenic recombination contributes to the allelic diversity of *P. falciparum msp1* to a greater extent. In Tanzania, frequent recombination events appear to generate novel *msp1* haplotypes intermittently and cause a temporal variation in the frequency distribution of *msp1* haplotypes, whereas the frequencies of individual polymorphisms are stable.

Received October 31, 2006. Accepted for publication December 11, 2006.

Acknowledgments: The authors thank Richard Culleton for reading this manuscript and his comments. We are grateful to the villagers and the research team in Nyamisati who participated in this study.

Financial support: This study was supported by a Grant-in-Aid for Scientific Research on Priority Areas from The Japanese Ministry of Education, Culture, Sports, Science and Technology (18073013), Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (18390131, 18GS03140013), the Japanese Ministry of Health, Labor and Welfare (H17-Sinkou-ippan-019), and the Swedish International Development Agency.

Authors' addresses: Kazuyuki Tanabe and Naoko Sakihama, Laboratory of Malariology, International Research Center of Infectious Diseases, Research Institute for Microbial Diseases Osaka University, 3-1, Yamada-Oka, Suita, 565-0871, Japan, Telephone: +81-6-6879-4260, Fax: +81-6-6879-4262, E-mail: kztanabe@biken.osaka-u.ac.jp. Ingegerd Rooth, Nyamisati Malaria Research Unit, P.O. Box, 663, Dar es Salaam, Tanzania. Anders Björkman and Anna Färnert, Unit of Infectious Diseases, Department of Medicine, Karolinska Institute, Karolinska University Hospital, Solna, SE-17176 Stockholm, Sweden.

Reprint requests: Kazuyuki Tanabe, Laboratory of Malariology, International Research Center of Infectious Diseases, Research Institute for Microbial Diseases Osaka University, 3-1, Yamada-Oka, Suita, 565-0871, Japan. Telephone: +81-6-6879-4260, Fax: +81-6-6879-4262, E-mail: kztanabe@biken.osaka-u.ac.jp.

REFERENCES

1. Holder AA, Guevara Patino JA, Uthaipibull C, Syed SE, Ling IT, Scott-Finnigan T, Blackman MJ, 1999. Merozoite surface protein 1, immune evasion, and vaccine against asexual blood stage malaria. *Parassitologia 41:* 409–414.
2. Mahanty S, Saul A, Miller LH, 2003. Progress in the development of recombinant and synthetic blood-stage malaria vaccines. *J Exp Biol 206:* 3781–3788.
3. Conway DJ, Cavanagh DR, Tanabe K, Roper C, Mikes ZS, Sakihama N, Bojang KA, Oduola AMJ, Kremsner PG, Arnot DE, Greenwood BM, McBride JS, 2000. A principal target of human immunity to malaria identified by molecular population genetic and immunological analyses. *Nat Med 6:* 689–692.
4. Blackman MJ, Ling IT, Nicholls SC, Holder AA, 1991. Proteolytic processing of the *Plasmodium falciparum* merozoite surface protein-1 produces a membrane-bound fragment containing two epidermal growth factor-like domains. *Mol Biochem Parasitol 49:* 29–34.
5. Chang SP, Gibson HL, Lee-Ng CT, Barr PJ, Hui GS, 1992. A carboxyl-terminal fragment of *Plasmodium falciparum* gp195 expressed by a recombinant baculovirus induces antibodies that completely inhibit parasite growth. *J Immunol 149:* 548–555.
6. Kumar S, Yadava A, Keister DB, Tian JH, Ohl M, Perdue-Greenfield KA, Miller LH, Kaslow DC, 1995. Immunogenicity and in vivo efficacy of recombinant *Plasmodium falciparum* merozoite surface protein-1 in *Aotus* monkeys. *Mol Med 1:* 325–333.
7. Egan AF, Morris J, Barnish G, Allen S, Greenwood BM, Kaslow DC, Holder AA, Riley EM, 1996. Clinical immunity to *Plasmodium falciparum* malaria is associated with serum antibodies to the 19-kDa C-terminal fragment of the merozoite surface antigen, PfMSP-1. *J Inf Dis 173:* 765–769.
8. Nwuba R, Sodeinde O, Anumudu C, Omosum Y, Odaibo A, Holder A, Nwagwu M, 2002. The human immune response to *Plasmodium falciparum* includes both antibodies that inhibit merozoite surface protein 1 secondary processing and blocking antibodies. *Infect Immun 70:* 5328–5331.
9. O'Donnel RA, de Koning-Ward TF, Burr RA, Bockarie M, Reeder JC, Cowman A, Crabb B, 2001. Antibodies against merozoite surface protein (MSP)-1_{19} are a major component of the invasion-inhibitory response in individuals immune to malaria. *J Exp Med 193:* 1403–1412.
10. Tanabe K, Mackay M, Goman M, Scaife JG, 1987. Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum. J Mol Biol 195:* 273–287.
11. Miller LH, Roberts T, Shahabuddin M, McCutchan TF, 1993. Analysis of genetic diversity in the *Plasmodium falciparum* merozoite surface protein-1 (*MSP1*). *Mol Biochem Parasitol 59:* 1–14.
12. Martinelli A, Cheesman S, Hunt P, Culleton R, Raza A, Mackinnon M, Carter R, 2005. A genetic approach to the de novo identification of targets of strain specific immunity in malaria parasites. *Proc Natl Acad Sci USA 102:* 814–819.
13. Cheesman S, Raza A, Carter R, 2006. Mixed strain infections and strain-specific protective immunity in the rodent malaria parasite *Plasmodium chabaudi chabaudi* in mice. *Infect Immun 74:* 2996–3001.

14. Tanabe K, Sakihama N, Kaneko A, 2004. Stable SNPs in malaria antigen genes in isolated populations. *Science 303:* 493.
15. Polley SD, Weedal GD, Thomas AW, Golightly LM, Conway DJ, 2005. Orthologous gene sequences of merozoite surface protein 1 (MSP1) from *Plasmodium reichenowi* and *P. gallinaceum* confirm an ancient divergence of *P. falciparum* alleles. *Mol Biochem Parasitol 142:* 25–31.
16. Ferreira MU, Ribeiro WL, Tonon AP, Kawamoto F, Rich SM, 2003. Sequence diversity and evolution of the malaria vaccine candidate merozoite surface protein-1 (*MSP1*) of *Plasmodium falciparum. Gene 304:* 65–75.
17. Snounou G, Zhu X, Siripoon N, Jarra W, Thaithong S, Brown KN, Viriyakosol S, 1999. Biased distribution of *msp1* and *msp2* allelic variants in *Plasmodium falciparum* populations in Thailand. *Trans R Soc Trop Med Hyg 93:* 369–374.
18. Rooth I, Björkman A, 1992. Fever episodes in a holoendemic area of Tanzania: parasitological and clinical findings and diagnostic aspects related to malaria. *Trans R Soc Trop Med Hyg 86:* 479–482.
19. Hay SI, Rogers DJ, Toomer JF, Snow RW, 2000. Annual *Plasmodium falciparum* entomological inoculation rates (EIR) across Africa: literature survey, internet access and review. *Trans R Soc Trop Med Hyg 94:* 113–127.
20. Sakihama N, Ohmae H, Bakote B, Kawabata M, Hirayama K, Tanabe K, 2006. Limited allelic diversity of *Plasmodium falciparum msp1* from populations in The Solomon Islands, a highly endemic area. *Am J Trop Med Hyg 74:* 31–40.
21. Sakihama N, Kimura M, Hirayama K, Kanda T, Na-Bangchang K, Jongwutiwes S, Conway D, Tanabe K, 1999. Allelic recombination and linkage disequilibrium within *Msp1* of *Plasmodium falciparum*, the malignant human malaria parasite. *Gene 230:* 47–54.
22. Tanabe K, Sakihama N, Färnert A, Rooth I, Björkman A, Walliker D, Ranford-Cartwright L, 2002. In vitro recombination during PCR of *Plasmodium falciparum* DNA: a potential pitfall in molecular population genetic analysis. *Mol Biochem Parasitol 122:* 211–216.
23. Qari SH, Shi YP, Goldman IF, Nahlen BL, Tibayrenc M, Lal AA, 1998. Predicted and observed alleles of *Plasmodium falciparum* merozoite surface protein-1 (MSP-1), a potential malaria vaccine antigen. *Mol Biochem Parasitol 92:* 241–252.
24. Tanabe K, Sakihama N, Nakamura Y, Kaneko O, Kimura M, Ferreira MU, Hirayama K, 2000. Selection and genetic drift of polymorphisms within the merozoite surface protein-1 gene of *Plasmodium falciparum. Gene 241:* 325–331.
25. Ferreira MU, Liu Q, Kimura M, Tanabe K, Kawamoto F, 1998. Allelic diversity in the merozoite surface protein-1 and epidemiology of multiple-clone *Plasmodium falciparum* infections in northern Tanzania. *J Parasitol 84:* 1286–1289.
26. Babiker H, Walliker D, 1997. Current views on the population structure of *Plasmodium falciparum*: implications for control. *Parasitol Today 13:* 262–267.
27. Mu J, Awadalla P, Duan J, McGee KM, Joy DA, McVean GAT, Su AZ, 2005. Recombination hot spots and population structure in *Plasmodium falciparum. PLoS Biol 3:* 1734–1741.
28. 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.
29. Silva NS, Silveira LA, Machado RL, Povoa MM, Ferreira MU, 2000. Temporal and spatial distribution of the variants of merozoite surface protein-1 (MSP-1) in *Plasmodium falciparum* populations in Brazil. *Ann Trop Med Parasitol 94:* 675–688.
30. Beck HP, Felger I, Huber W, Steiger S, Smith T, Weiss N, Alonso P, Tanner M, 1997. Analysis of multiple *Plasmodium falciparum* infections in Tanzanian children during the phase III trials of the malaria vaccine SPf66. *J Inf Dis 175:* 921–926.
31. Contamin H, Fandeur T, Rogier C, Bonnefoy S, Trape JF, Mercereau-Puijalon O, 1996. Different genetic characteristics of *Plasmodium falciparum* isolates collected during successive clinical malaria episodes in Senegalese children. *Am J Trop Med Hyg 54:* 632–643.
32. Färnert A, Rooth I, Svensson A, Snounou G, Bjorkman A, 1999. Complexity of *Plasmodium falciparum* infections is consistent over time and protects against clinical disease in Tanzanian children. *J Infect Dis 179:* 989–995.
33. Woehlbier U, Epp C, Kauth CW, Lutz R, Long CA, Coulibaly B, Kouyaté B, Arevalo-Herrera M, Herrera S, Bujard H, 2006. Analysis of antibodies directed against merozoite surface protein 1 of the human malaria parasite *Plasmodium falciparum. Infect Immun 74:* 1313–1322.
34. Ekaka MT, Jouin H, Lekoulou F, Issifou S, Mercereau-Puijalon O, Ntoumi F, 2002. *Plasmodium falciparum* merozoite surface protein 1 (MSP1) genotyping and humoral responses to allele-specific variants. *Acta Trop 81:* 33–46.
35. Fraser-Hurt N, Felger I, Edoh D, Steiger S, Mashaka M, Masanja H, Smith T, Mbena F, Beck HP, 1999. Effect of insecticide-treated bed nets on haemoglobin values, prevalence and multiplicity of infection with *Plasmodium falciparum* in a randomized controlled trial in Tanzania. *Trans R Soc Trop Med Hyg 93 (Suppl 1):* 47–51.
36. Bereczky S, Liljander A, Rooth I, Faraja L, Montgomery SM, Färnert A, 2007. Multiclonal asymptomatic *Plasmodium falciparum* infections predict a reduced risk of malaria disease in a Tanzanian population. *Microbes Infect 9:* 103–110.
37. Conway DJ, Greenwood BM, McBride JS, 1992. Longitudinal study of *Plasmodium falciparum* polymorphic antigens in a malaria-endemic population. *Infect Immun 60:* 1122–1127.