

Plasmodium falciparum Genetic Diversity in Western Kenya Highlands

Daibin Zhong,* Yaw Afrane, Andrew Githeko, Zhaoqing Yang, Liwang Cui, David M. Menge, Emmanuel A. Temu, and Guiyun Yan

Program in Public Health, College of Health Sciences, University of California, Irvine, California; Centre for Global Health Research, Kenya Medical Research Institute, Kisumu, Kenya; Department of Parasitology, Kunming Medical College, Kunming, Yunnan, People's Republic of China; Department of Entomology, Pennsylvania State University, University Park, Pennsylvania; Center for Infectious Diseases and Microbiology Translational Research, University of Minnesota, Minneapolis, Minnesota; The Institute of Tropical Medicine, Nagasaki University, Nagasaki-shi, Japan

Abstract. The present study examined the genetic diversity and population structure of *Plasmodium falciparum* in western Kenya by analyzing the polymorphism of 12 microsatellite loci and two antigen loci. Malaria in highland areas is unstable and epidemic whereas malaria in lowland areas is endemic. Transmission intensity and malaria prevalence are substantially lower in the highlands than in the lowlands. Despite that the highland parasite populations exhibited reduced number of alleles, lower expected heterozygosity, and infection complexity in comparison to the surrounding lowland population, genetic diversity of the highland populations remained high in comparison to parasites from other meso-endemic regions. More than 70% of infections from western Kenya highland study sites were mixed genotype infections. Small but statistically significant genetic differentiation between highland and lowland populations was detected. These findings are discussed in the context of human travel and local transmission in the study area.

INTRODUCTION

Epidemic malaria is a major public health problem in the African highlands (highlands are generally referred to areas with elevation more than 1,500 meters above sea level). For instance, since the 1980s a series of *Plasmodium falciparum* epidemic malaria has occurred in the western Kenyan highlands, a region that was previously free of malaria or exhibited low incidences of the disease.^{1–4} The epidemic malaria in the western Kenyan highlands is characterized by expanded geographic areas,⁵ increased frequencies,⁶ and increased case-fatality rates.⁷ Previous studies have examined the risk factors associated with clinical malaria in the highlands,^{8–10} population dynamics of mosquito vectors,¹¹ effects of topography and land use on the spatial distribution of vectors and malaria infections,^{12–14} and the potential role of climate factors^{4,15,16} and drug resistance on malaria epidemics.^{10,17}

Genetic diversity and population structure of malaria parasites have profound impacts on clonal diversity,¹⁸ competitive or positive interactions among the clones,^{19,20} dynamics of drug resistance,²¹ persistence of asexual infection and gametocyte production,²² infectivity in mosquito vectors,^{23,24} and malaria vaccine development.^{25,26} Global-scale studies on *P. falciparum* genetic diversity using specimens with varying transmission intensities from sites in Africa, South America, and Asia found a general pattern of low genetic diversity, strong linkage disequilibrium, and high genetic differentiation in low transmission areas, whereas in areas of high transmission intensity, parasite populations exhibited a panmictic pattern with high genetic diversity and weak linkage disequilibrium.¹⁸ Previous studies established that infection complexity within a host was dependent on exposure (transmission intensity) and age of the host (reflective of level of immunity), and negatively associated with the severity of the disease.^{27–32} A recent study found that allelic diversity of the merozoite surface protein 1 gene (*msp-1*) is affected not only by transmission intensity, but also by other factors such as the number of alleles prevalent and infection complexity.³³ Moreover,

studies of daily parasite population dynamics showed that such complex patterns of multiple infections would undergo rapid changes.³⁴ To date, few studies have examined the genetic diversity and population structure of malaria parasites in highland areas where malaria transmission is unstable and occurrences of epidemic is common.

The objective of this study was to determine genetic diversity of *P. falciparum* in two highland sites (Kakamega and Kisii districts) in western Kenya. We then compared our results with those found in a surrounding, malaria-endemic lowland site (Kisumu district), also in western Kenya. The transmission intensity and malaria prevalence in the highlands was substantially lower than in the lowlands.¹¹ If transmission intensity is an important determinant of the genetic diversity of malaria parasites, one would expect that epidemic malaria in the highlands would lead to few major genotypes in the population, thus low genetic diversity and low infection complexity. To the contrary, high genetic diversity and infection complexity is expected in the malaria-endemic lowlands because of frequent meiotic recombination events in mosquito vectors resulting from high infection rates. Conversely, travel to malaria-endemic areas was found to be significantly associated with increased malaria risk in highland residents.¹⁰ Introduction of new alleles by travelers from malaria-endemic areas to the highlands would increase population genetic diversity. We examined genetic diversity of block 2 of *msp-1* and block 3 of *msp-2* genes that code for major blood-stage surface antigens. In malaria-endemic areas in Africa, allelic types of block 2 of *msp-1* were strongly associated with protection from *P. falciparum* malaria,³⁵ demonstrating that these genes are under strong selection by human immune systems.^{36,37} Thus, interpretation of population structure and diversity results based on a small number of strongly selected markers such as *msp-1* and *msp-2* is problematic because the results are confounded by population history and natural selection.¹⁸ Therefore, we also included 12 microsatellite markers that are presumably selective neutral.^{38,39}

MATERIALS AND METHODS

Parasite sampling. Blood samples were collected from malaria outpatients attending three clinics, Iguhu (1,500–1,580

* Address correspondence to Daibin Zhong, Program in Public Health, College of Health Sciences, University of California at Irvine, Irvine, CA 92697. E-mail: dzhong@uci.edu

meters above sea level) in Kakamega district, Marani (1,520–1,700 meters above sea level) in Kisii district, and Kombewa (1,170–1,250 meters above sea level) in Kisumu district, western Kenya (Figure 1). The three study sites have been previously described in detail.¹¹ The transmission intensity, measured by annual entomologic inoculation rate (EIR), was estimated at 16.6 infectious bites per person per year in the Kakamega site, 0.4 in Kisii, and 31.1 in Kombewa.¹¹ In 2005, *P. falciparum* prevalence in asymptomatic children (6–14 years of aged) ranged from 20% to 48% in Kakamega, less than 10% in Kisii, and consistently more than 50% in Kombewa (Afrane Y, unpublished data). *Plasmodium falciparum* constituted more than 98% of malaria infections in these sites. On the basis of elevation, the Kakamega and Kisii sites are classified as highland sites, and the Kombewa site as a lowland site. The catchment area of these village clinics is small, typically less than 10,000 residents within a 15-km² area. A blood sample of approximately 200 μ L was obtained from symptomatic volunteers using the finger prick method and then preserved on Whatman (Maidstone, United Kingdom) filter papers. All samples were collected in August 2005. An exclusion criterion was unwillingness to participate in the study. Blood samples from all study participants were tested for *P. falciparum* infection by microscopy, and all positive cases were treated with artemether/lumefantrine (Coartem®; Novartis, Basel, Switzerland). The mean age of donors was 16.8 years (range = 1–86 years) in Kakamega (n = 98), 16.7 years (range = 1–80 years) in Kisii (n = 90), and 15.4 years (range = 1–60 years) in Kombewa (n = 100). The study was reviewed and approved by the ethical review committee of the Kenya Medical Research Institute and the institutional review board of the University of California at Irvine.

Parasite DNA extraction and microsatellite genotyping.

Genomic DNA of the malaria parasite was extracted from 50

microscopic-positive samples per site using the Saponin/Chelex method.⁴⁰ The nested ribosomal DNA–polymerase chain reaction (PCR) method described by Singh and the others⁴¹ was used to confirm *P. falciparum* species identification. Parasites were genotyped using 12 microsatellite markers (Poly α , TA42, TA81, TA1, TA87, TA109, ARA2, 2490, TA40, PfkPK2, Pfg377, and TA60) that have previously been used for population genetic studies.^{18,42–44} The single-clone parasite supplied by the Malaria Research and Reference Reagent Resource Center (Manassas, VA) (MR4) and distilled water were used as positive and negative controls, respectively. We followed the protocol of Anderson and others³⁹ for PCR conditions, primer specificity and reproducibility, with small modifications for automated genotyping in a model 4300 automated DNA analyzer (Li-Cor, Lincoln, NE). For the apparatus to detect PCR products, one primer in every microsatellite marker must be fluorescently labeled (infrared dye, IRD800). To reduce the cost associated with synthesis of fluorescently labeled primers, we used the tailed primer method.⁴⁵ In this method, the forward primer for each marker was synthesized with an additional 19-basepair M13 sequence (5'-CACGACGTTGTAAAACGAC-3') added to the 5' end of the oligonucleotide. A primer with the same 19-basepair sequence was directly labeled with the fluorescent dye (IRD800) and was used as the sole type of labeled primer for the detection of microsatellite alleles for all 12 markers. The tailed primer method dramatically reduced the synthesis cost of fluorescently labeled primers.

The PCR was carried out in a total volume of 25 μ L for each sample and locus and contained 20 μ L of 1.1 \times PCR Master Mix (Abgene, Rochester, NY), 0.25 μ L of 10 μ M each primer, 0.5 μ L of M13 primer, an additional 1 μ L of 25 mM MgCl₂ and 10 ng of template DNA. An initial denaturation period of for 5 minutes at 95°C preceded the 45 amplification cycles (annealing for 30 seconds at 45°C, extension for 45 seconds at 65°C, and denaturation for 30 seconds at 94°C), followed by a final extension of 7 minutes. A Gene ImagIR 4.33 software (Li-Cor) program was used to quantify allele size on the basis of the height pattern of signal peaks. Multiple alleles were scored at a given locus if minor peaks were more than one-third the height of the predominant peak. An infection was considered to contain multiple clones if one or more loci showed more than one allele.^{18,42} Infection complexity provides a surrogate indicator of the level of transmission within populations, as well the opportunity for genetic recombination between different malaria clones.⁴⁶

Merozoite surface protein genotyping. We genotyped three *P. falciparum* populations using polymorphic block 2 of *msp-1* gene on chromosome 9 and block 3 of *msp-2* gene on chromosome 2. A nested PCR were performed individually using previously published primer sequences and conditions that distinguish three major allelic families (K1, MAD20, and RO33) for *msp-1* and two allelic families (FC27 and IC1) for *msp-2*.^{47,48} The PCR products were subjected to electrophoresis on a 2.0% agarose gel. Fragment length was determined using the Kodak Digital Science 1D Image Analysis Software (Eastman Kodak Company, New Haven, CT). Each *P. falciparum* infection was characterized on the basis of the allele size for each locus. Sizes of alleles were allocated into bins of 40-basepair size ranges.⁴⁹

Statistical analysis. For microsatellite loci, the analyses focused on population genetic diversity, infection complexity,

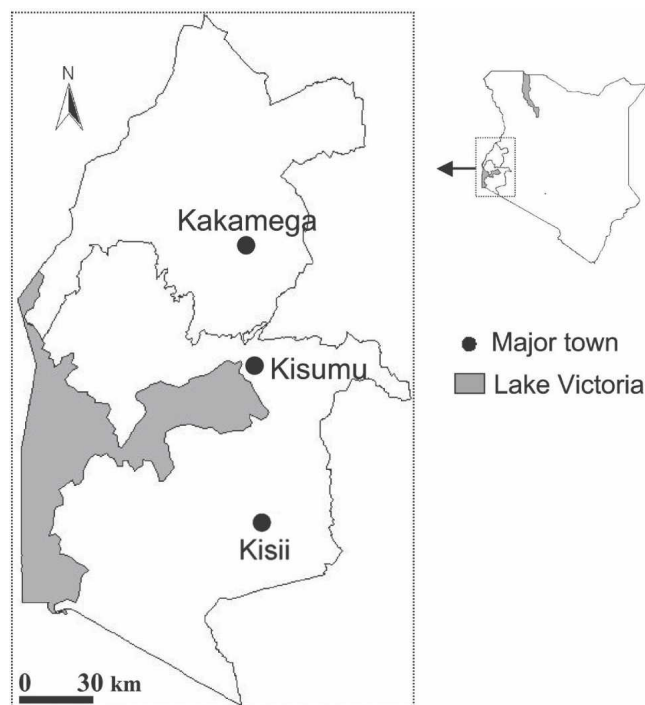


FIGURE 1. Map of Kenya showing the location of the study sites.

linkage disequilibrium, and genetic differentiation. Genetic diversity was measured by the number of alleles per locus and expected heterozygosity on the basis of allele frequency data in each population. In the case of the presence of multiple alleles at a locus within an infection, only the predominant allele defined by the allele with the highest peak in pherograms was used for calculation of allele frequency.¹⁸ This procedure is appropriate for estimation of population allele frequency if the composition of PCR products is representative of the composition of templates. The unbiased expected heterozygosity, H_e , was calculated as $H_e = [n/(n-1)][1 - \sum p_i^2]$, where n is the number of infections sampled and p_i is the frequency of the i th allele.¹⁸ Observed infection complexity is defined as the largest number of alleles at any *m*sp locus detected in the sample; it is often referred to as minimum number of clones.^{46,50} For example, an infection with two *m*sp-1 alleles and one *m*sp-2 allele is scored as two clones. This measure is conservative and it likely underestimates the number of clones likely to be present.⁴⁶ Therefore, we used the maximum likelihood methods described by Hill and Babiker⁴⁶ to estimate the mean number of clones, the support limit and *m*sp allele frequencies, assuming a Poisson distribution of number of clones per individual. Because detection of multiple alleles per locus depends critically on the polymorphism of the markers and sensitivity of the methods used, the data provides relative assessment rather than absolute measures of infection complexity. The non-parametric Wilcoxon test was used to determine the statistical differences in genetic diversity among populations.

Linkage disequilibrium (LD), which characterizes statistical independence of alleles at pairs of loci, was determined for all pairs of microsatellite loci, using the LIAN 3.0 software developed for multilocus haploid data.⁵¹ This software tests the null hypothesis of no linkage by a Monte Carlo simulation (10,000 permutations) on the variance of genetic distances between isolates (V_D). The sample variance is then compared with the variance expected under linkage equilibrium V_E . A distribution of V_E is generated by Monte Carlo simulations, and its percentiles provide 95% confidence intervals. The output file gives V_D and V_E values, as well as a standardized index of association ($I_{AS} = (V_D/V_E - 1)/(1 - r)$), where r is the number of loci, a measure of haplotype-wide linkage and the 95% confidence limits determined by Monte Carlo simulations (L_{MC}).⁵¹ Because *m*sp markers are not selectively neutral,^{35–37} population genetic structure was determined using microsatellite markers alone. The analysis of molecular variance (AMOVA) was conducted using the Arlequin genetic analysis software.⁵² The AMOVA partitions the molecular variance (microsatellite allele size) into three categories: among groups (highland and lowland), among populations within groups, and among individuals within populations. A Permutation test ($n = 1,000$) was used to test the null hypothesis that the variance component is zero. Wright's F -statistics between pairs of populations is also computed using the FSTAT computer software.⁵³

Structure 2.0 computer program was used to assign individuals from all populations to a predetermined number of clusters (K) based on multilocus microsatellite data.⁵⁴ For each run, a burn-in period of 50,000 steps was followed by 1,000,000 iterations under the admixture model and the assumption of correlated allele frequencies. Various values of K (2 to 4) were tested. Individuals were assigned to the cluster

only if the proportion ancestry was ≥ 0.8 .⁵⁵ The individual was unassigned when the proportion ancestry from any single cluster was < 0.8 .

RESULTS

Microsatellite polymorphism. High levels of polymorphism were shown among parasites because all 12 microsatellite markers used were polymorphic for the three populations studied. The number of microsatellite alleles per locus ranged from 7 (TA42) to 16 (Poly α). The lowland population showed the highest level of diversity, and the highland samples showed a significantly lower level of diversity (Table 1). For example, an average of 8.6 alleles was detected in the lowland Kombewa population (range = 3–14), whereas the two highland populations, Kakamega and Kisii, showed an average of 6.7 alleles (range = 4–9) and 5.8 alleles (range = 2–9), respectively ($\chi^2 = 7.44$, $df = 2$, $P = 0.02$). The levels of unbiased expected heterozygosity of the lowland Kombewa population were marginally higher than those observed in the highland population ($\chi^2 = 4.86$, degrees of freedom [df] = 2, $P = 0.09$). The estimated expected heterozygosity (H_e) was 0.75 (range = 0.55–0.89) in Kombewa, 0.69 (range = 0.53–0.78) in Kakamega, and 0.66 (range = 0.38–0.81) in Kisii (Table 1).

The overall infection complexity (minimum number of clones in an infection detected by all 12 microsatellite markers) was marginally higher in the lowland Kombewa population than in the two highland populations ($\chi^2 = 5.54$, $df = 2$, $P = 0.06$; Table 2). The proportion of infections with more than one clone was 87.0%, 82.4%, and 80.9% in Kombewa, Kakamega and Kisii, respectively ($\chi^2 = 1.44$, $df = 2$, $P = 0.49$).

Genetic diversity of *m*sp-1 and *m*sp-2 genes. The *P. falciparum* parasite populations exhibited a high degree of diversity on the basis of length polymorphism of the two antigen loci. Using a conservative bin size range of 40 basepairs for *m*sp-1 and *m*sp-2, a total of 13 alleles were identified for *m*sp-1 and 15 alleles for *m*sp-2 for the three locations (Table 3). At the *m*sp-1 locus, a total of six alleles of the K1 family (size range = 140–370 basepairs), five MAD20 alleles (range = 120–320 basepairs) and two alleles of the RO33 family (range = 180–260 basepairs) were detected in the three

TABLE 1
Genetic diversity of *Plasmodium falciparum* from western Kenya at 12 microsatellite loci*

Locus	Chromosome	Lowland		Highland		Allele size (basepairs)
		Kombewa		Kakamega	Kisii	
		<i>Na</i>	<i>He</i>	<i>Na</i>	<i>He</i>	
Poly α	4	14	0.89	9	0.73	179–248
TA42	5	5	0.66	4	0.53	204–277
TA81	5	10	0.80	5	0.63	129–160
TA1	6	10	0.58	9	0.53	152–221
TA87	6	10	0.86	7	0.78	109–144
TA109	6	10	0.82	8	0.69	170–237
ARA2	11	8	0.81	7	0.82	79–108
2490	10	5	0.64	4	0.38	90–115
TA40	10	10	0.75	9	0.61	195–252
PfPK2	12	9	0.85	8	0.77	167–217
Pfg377	12	3	0.55	4	0.57	110–130
TA60	13	9	0.84	6	0.81	188–237
Mean		8.6	0.75	6.7	0.66	

* *Na* = number of alleles for microsatellite loci; *He* = expected heterozygosity.

TABLE 2

Mean infection complexity and percentage of mixed infections detected by microsatellite markers in three populations of *Plasmodium falciparum* from highland and lowland areas in western Kenya

Locus	Lowland		Highland			
	Kombewa		Kakamega		Kisii	
	Mean complexity \pm SE*	% Multi-infection†	Mean complexity \pm SE*	% Multi-infection†	Mean complexity \pm SE*	% Multi-infection†
Poly α	2.38 \pm 0.22	33.9	2.00 \pm 0.24	36.4	1.71 \pm 0.24	23.5
TA42	1.14 \pm 0.10	20.3	1.10 \pm 0.11	20.5	1.10 \pm 0.17	17.7
TA81	3.33 \pm 0.25	72.9	3.00 \pm 0.25	62.7	2.70 \pm 0.39	60.6
TA1	2.43 \pm 0.21	59.3	1.90 \pm 0.19	50.0	1.73 \pm 0.26	47.1
TA87	2.65 \pm 0.23	49.2	2.30 \pm 0.26	47.7	2.33 \pm 0.37	52.9
TA109	1.83 \pm 0.16	28.8	1.56 \pm 0.15	29.6	1.69 \pm 0.25	35.3
ARA2	2.82 \pm 0.24	55.9	2.10 \pm 0.20	61.4	2.00 \pm 0.26	58.8
2490	1.92 \pm 0.16	66.1	1.70 \pm 0.17	62.7	1.60 \pm 0.26	64.7
TA40	1.71 \pm 0.13	44.1	1.20 \pm 0.12	31.8	1.50 \pm 0.26	35.3
PfPK2	2.06 \pm 0.16	50.9	1.29 \pm 0.16	29.6	1.25 \pm 0.17	17.7
Pfg377	2.43 \pm 0.16	71.2	2.20 \pm 0.18	68.2	2.00 \pm 0.31	60.6
TA60	2.31 \pm 0.20	44.1	2.10 \pm 0.24	38.6	1.80 \pm 0.25	41.2
Overall‡	4.76 \pm 0.22	87.0	4.02 \pm 0.26	82.4	3.76 \pm 0.32	80.9

* Complexity is the minimum number of parasite genotypes per infected person.

† Percentage of infections that exhibited more than one genotype.

‡ The overall mean infection complexity is the average of the highest number of clones detected by any microsatellite markers tested. The overall proportion of multi-infection refers to the percentage of infections with more than one clone detected by any 12 microsatellite markers.

populations. Alleles of the RO33 family in the *msp-1* gene were not found in the Kisii population. At the *msp-2* locus, seven IC1 alleles (range = 440–760 basepairs) and eight FC27 alleles (range = 300–720 basepairs) were observed. Overall, the Kombewa population showed the highest number of alleles, and the Kisii population showed the lowest number of alleles (Table 3). The K1 and IC1 allele families were dominant alleles in western Kenya study sites. The minimum number of clones was 1.98, 1.92, and 1.58 in Kombewa, Kakamega, and Kisii, respectively ($P > 0.05$). The number of clones estimated by the maximum likelihood method was substantially higher than the observed number of clones (Table 3), but the trend on the difference between highland and lowland sites was similar to those detected by the minimum number of clones. The proportion of mixed infections based on *msp-1* and *msp-2* genes was 81.7%, 76.7%, and 73.3% in Kombewa, Kakamega, and Kisii, respectively ($\chi^2 = 1.21$, $df = 2$, $P = 0.55$). This result indicated that multiple infections in western Kenya highlands were as frequent as in the lowland populations.

Linkage disequilibrium. For microsatellite analysis, 9 and 17 of the 66 possible tests for LD showed significant results at $P \leq 0.05$ in the lowland and highland populations, respectively. None of the significant pairwise associations involved loci located on the same chromosome. The LD estimated by the LIAN program was significant in the lowland population (observed mismatch variance $V_D = 3.66$, expected mismatch variance $V_E = 1.93$, standardized index of association $I_{AS} = 0.082$, simulated 5% critical value $L_{MC} = 2.09$, $P = 10^{-4}$) and in the highland populations ($V_D = 3.92$, $V_E = 1.94$, $I_{AS} = 0.093$, $L_{MC} = 2.16$, $P = 10^{-4}$).

Population structure. The AMOVA found that most (92.2%) of the variance in allele frequencies was among individuals within populations; only 6.9% of the variation was explained by differences between highland and lowland populations ($P < 0.001$) (Table 4). Wright's F-statistics indicated that 9 of 12 microsatellite markers detected significant population structure (Table 5). The overall genetic differentiation index among the three populations (F_{ST}) was 0.027 ($P <$

TABLE 3

Polymorphism of merozoite surface protein (*msp*) genes and infection complexity in *Plasmodium falciparum* from highland and lowland areas in western Kenya

Alleles	Lowland		Highland			
	Kombewa		Kakamega		Kisii	
	No. alleles	Frequency	No. alleles	Frequency	No. alleles	Frequency
<i>msp-1</i>						
K1	6	0.602	4	0.622	5	0.815
MAD20	5	0.186	5	0.180	2	0.185
RO33	2	0.212	2	0.197	0	0.000
Total	13		11		7	
<i>msp-2</i>						
IC1	7	0.615	6	0.644	3	0.554
FC27	8	0.385	7	0.356	3	0.446
Total	15		13		6	
Average minimum no. of clones (95% confidence interval)	1.98 (1.77–2.19)		1.92 (1.69–2.15)		1.58 (1.34–1.82)	
No. of clones estimated by the maximum likelihood method (range of support limit)	3.86 (3.02–4.53)		2.83 (2.15–3.79)		2.19 (1.53–2.87)	
Percentage of infections with more than one clone	81.7		76.7		70.3	

TABLE 4

Results of analysis of molecular variance for genetic variation of microsatellite markers in *Plasmodium falciparum* populations from highland and lowland areas in western Kenya

Source of variation	Degrees of freedom	Sum of squares	Expected mean squares	Variance component	Percentage of variation	P
Among groups	1	32.4	32.4	0.65	6.9	< 0.001
Among populations within groups	1	11.6	11.6	0.08	0.9	0.113
Among individuals within populations	147	1,231.6	8.4	8.63	92.2	< 0.001
Total	149	1,275.6	52.4	9.36		

0.0001), whereas differences between highland and lowland populations were pronounced ($F_{ST} = 0.036$, $P < 0.0001$). Seven of 12 microsatellite markers detected significant structure between Kombewa and Kisii, which had the lowest transmission intensity ($F_{ST} = 0.032$, $P < 0.0001$). Between the two highland populations, four of 12 markers detected significant structure, giving an average differentiation index of 0.017 ($P < 0.05$).

Assignment test. Two clusters, each representing parasites with similar microsatellite genotypes, were identified by the structure analysis. Assignment test found that 91.1–94.0% individuals could be assigned to particular locations (Table 6). From the location perspective, most (76%) of the malaria parasites from the lowland site (Kombewa) were assigned to one cluster, whereas parasites from the two highland sites represented a mixture of both clusters in similar frequencies (Table 6). These results indicate a high rate of parasite infiltration from lowland to highland areas.

DISCUSSION

In this study, microsatellite markers detected a significantly lower number of alleles, marginally lower expected heterozygosity, and reduced infection complexity in the highland populations in comparison to the lowland population. Similar trends in genetic diversity measures were found in *msp-1* and *msp-2*; however, the difference between highland and lowland populations was not statistically significant. Malaria in the lowland site used in this study is hyperendemic according to its transmissions intensity (annual EIR = 31.1) and prevalence in asymptomatic children (< 50%). The lack of statistical significance in population genetic diversity by *msp* genes is likely due to the fact that *msp* genes are under strong selec-

tion pressure,^{35–37} the limited sample size analyzed in this study,¹⁸ and limited difference in malaria transmission intensities among the study sites.

The expected heterozygosity of microsatellite markers in our highland samples (He range = 0.66–0.69) was substantially higher than that observed in South America (Colombia, Bolivia, and Brazil; He range = 0.30–0.40) and in Thailand ($He = 0.51$),¹⁸ but was comparable to the samples from cosmopolitan Africa (Benin, Cameroon, The Gambia, Ghana and Uganda; He range = 0.69).⁴³ In comparison to other malaria-endemic African regions, genetic diversity of malaria parasites in western Kenya highlands was slightly lower. Studies with microsatellite markers showed expected heterozygosity of 0.76–0.80 for *P. falciparum* from malaria-endemic sites in Uganda, Congo, and Zimbabwe,¹⁸ and 0.79 for those from the Republic of the Congo.⁴³ These estimates were comparable with our lowland Kombewa sample.

Interestingly, although highland parasite populations exhibited reduced infection complexity in comparison with the lowland population, the proportion of mixed infections was similar between the highland and lowland populations. Both microsatellite markers and *msp* gene detected a high proportion of mixed infections in the highland populations. The high proportion of mixed infections paralleled high genetic diversity in the highland. This observation raises a question of how high genetic diversity and high prevalence of mixed infections are generated and maintained in the highlands where malaria transmission intensity is generally low (e.g., EIR was estimated as merely 0.4 infectious bites per person in Kisii).¹¹ One possible mechanism is the introduction of parasite clones from malaria-endemic lowland areas through human travel. Human travel is an important risk factor for malaria in the highlands.¹⁰ During the epidemic season, malaria transmission intensity is increased substantially over that in the non-epidemic season. Diversity in microsatellite alleles or *msp* haplotypes may be generated through meiotic recombination during the epidemic, a period with enhanced parasite infection rates in vector populations. If low-transmission during non-epidemic seasons did not lead to a substantial genetic

TABLE 5

F_{ST} indices of the 12 microsatellite loci for *Plasmodium falciparum* populations from highland and lowland areas in western Kenya

Marker	Among the three populations	Highland versus lowland	Within highland (Kakamega versus Kisii)
TA42	0.002	0.005	0.003
Pfg377	0.047*	0.059*	0.042*
ARA2	0.054†	0.056†	0.004
TA87	0.005	0.028	0.012
2490	0.027‡	0.041*	0.035†
TA40	0.029‡	0.038†	0.032†
PfPK2	0.02	0.015	0.006
TA109	0.038†	0.036†	0.015
Poly α	0.023*	0.043*	0.029
TA60	0.036†	0.048*	0.007
TA81	0.023‡	0.030†	0.011
TA1	0.022‡	0.031†	0.013
Mean	0.027*	0.036*	0.017‡

* $P < 0.0001$.

† $P < 0.001$.

‡ $P < 0.01$.

TABLE 6

Results of Bayesian cluster analysis and assignment test for *Plasmodium falciparum* populations from highland and lowland areas in western Kenya

Population	Sample size	Cluster 1*	Cluster 2*	Not assigned†	Percent assigned
Kombewa	50	38	8	4	92.0
Kakamega	50	26	21	3	94.0
Kisii	45	21	20	4	91.1
Total	145	85	49	11	

* Numbered clusters represent distinct groups identified by Bayesian cluster analysis using Structure computer software.⁵⁴

† Not assigned refers to the number of individuals that could not be assigned to any cluster statistically.

bottleneck effect, population diversity and mixed infection could be maintained in the population. The question of whether low transmission intensity in the long non-epidemic season leads to parasite selfing and thus dramatic reduction in genetic diversity has not been determined. Evidence from the lack of significant seasonal fluctuations in the allele frequencies of *msp-1*, *msp-2*, and the gene coding for glutamate-rich protein in Sudan, after a six-month dry season during which malaria transmission was not even detectable, suggests that genetic bottleneck in *P. falciparum* during low transmission season should not be severe.⁵⁶ It would be interesting to determine the changes in genetic diversity and infection complexity throughout an epidemic season in the highland areas.

We observed small, but statistically significant population structure between the highland and lowland malaria populations ($F_{ST} = 0.036$, $P < 0.0001$). The lack of strong genetic differentiation between highland and lowland parasite populations is consistent with the notion of parasite introduction by human travel from the malaria-endemic lowland to the epidemic highland region. Shanks and others¹⁰ found that travel to malaria-endemic areas was significantly associated with increased malaria risk in highland residents. In addition to human travel, genetic structure of malaria parasite populations may be affected by parasite life history, site-specific epidemiology, and malaria control measures such as insecticide-treated bed nets and indoor residual spray. Selection by antimalarial drugs would also result in unique genetic imprints on parasite populations.

We found significant linkage disequilibrium in both highland and lowland populations in the present study. None of the significant linkage disequilibrium involved loci located on the same chromosome, which suggested that these loci are only statistically linked and not physically linked. Linkage disequilibrium in our highland populations was consistent with the expectation of low transmission. However, linkage disequilibrium in our lowland population present with a substantially stronger index of association ($I_{AS} = 0.082$) than those reported in other African sites ($I_{AS} < 0.02$) with comparable malaria prevalence.¹⁸ Sakihama and others³³ also reported strong linkage disequilibrium in the *msp-1* gene in *P. falciparum* populations from the Solomon Islands where the EIR exceeds several hundred infectious bites per person per year. These investigators suggested that transmission intensity is not the sole determinant of the strength of linkage disequilibrium. Other variables such as the number of prevalent alleles, the prevalence of mixed infections, and infection complexity may also be involved. In our study, more than 70% of infection cases were mixed genotype infections in the highlands and lowlands on the basis of microsatellite or *msp* markers. Multiple repeated multilocus genotypes may have contributed to the observed linkage disequilibrium. No significant linkage disequilibrium was found in *P. falciparum* populations from hypo-endemic areas in Africa when only genotypes of the mono-infected isolates were used.⁴⁴

In summary, the present study examined the genetic diversity of *P. falciparum* from symptomatic infections in western Kenya highland and lowland areas with contrasting transmission intensities. Despite reduced genetic diversity in the highland populations in comparison to populations in the surrounding endemic lowland, population genetic diversity in the highland area remained high. More than 70% of the infections were mixed genotype infections, even in areas of with an

EIR < 1 infectious bite per person per year. Small geographic differentiation was detected between highland and lowland populations. Together with the results of risk factor analysis¹⁰ and malaria infection dynamics^{11,12} in the highlands, these results suggest that both human travel and local transmission may be important in shaping the genetic structure of malaria parasites in the highlands.

Received May 2, 2007. Accepted for publication August 23, 2007.

Acknowledgments: This study is published with the permission of the Director of the Kenya Medical Research Institute. We thank Claire Garro, Hong Chen, and two anonymous reviewers for helpful suggestions on the manuscript.

Financial support: This work was supported by grants R01 A150243, D43 TW01505, and R03 TW007360 from the National Institutes of Health.

Author's addresses: Daibin Zhong and Guiyun Yan, Program in Public Health, College of Health Sciences, University of California at Irvine, Irvine, CA 92697. Yaw Afrane and Andrew Githeko, Centre for Global Health Research, Kenya Medical Research Institute, Kisumu, Kenya. Zhaoqing Yang, Department of Parasitology, Kunming Medical College, Kunming 650031, Yunnan, People's Republic of China. Liwang Cui, Department of Entomology, Pennsylvania State University, University Park, PA 16802. David M. Minge, Center for Infectious Diseases and Microbiology Translational Research, University of Minnesota, Minneapolis, MN 55455. Emmanuel A. Temu, The Institute of Tropical Medicine, Nagasaki University, Nagasaki-shi 852-8523, Japan.

Reprint requests: Daibin Zhong, Program in Public Health, College of Health Sciences, University of California at Irvine, Irvine, CA 92697, E-mail: dzhong@uci.edu.

REFERENCES

- Garnham PC, 1945. Malaria epidemics at exceptionally high altitudes in Kenya. *BMJ* 2: 45-47.
- Roberts JM, 1964. The control of epidemic malaria in the highlands of western Kenya. Part I. Before the campaign. *J Trop Med Hyg* 61: 161-168.
- Malakooti MA, Biomndo K, Shanks GD, 1998. Reemergence of epidemic malaria in the highlands of western Kenya. *Emerg Infect Dis* 4: 671-676.
- Hay SI, Simba M, Busolo M, Noor AM, Guyatt HL, Ochola SA, Snow RW, 2002. Defining and detecting malaria epidemics in the highlands of western Kenya. *Emerg Infect Dis* 8: 555-562.
- World Health Organization, 2003. *A Practical Guideline for Prevention and Control of Malaria Epidemics in Africa*. Geneva: World Health Organization.
- Snow RW, Craig MH, Deichmann U, le Sueur D, 1999. A preliminary continental risk map for malaria mortality among African children. *Parasitol Today* 15: 99-104.
- Shanks GD, Biomndo K, Hay SI, Snow RW, 2000. Changing patterns of clinical malaria since 1965 among a tea estate population located in the Kenyan highlands. *Trans R Soc Trop Med Hyg* 94: 253-255.
- Brooker S, Clarke S, Njagi JK, Polack S, Mugo B, Estambale B, Muchiri E, Magnussen P, Cox J, 2004. Spatial clustering of malaria and associated risk factors during an epidemic in a highland area of western Kenya. *Trop Med Int Health* 9: 757-766.
- John CC, Koech DK, Sumba PO, Ouma JH, 2004. Risk of *Plasmodium falciparum* infection during a malaria epidemic in highland Kenya, 1997. *Acta Trop* 92: 55-61.
- Shanks GD, Biomndo K, Guyatt HL, Snow RW, 2005. Travel as a risk factor for uncomplicated *Plasmodium falciparum* malaria in the highlands of western Kenya. *Trans R Soc Trop Med Hyg* 99: 71-74.
- Ndenga B, Githeko AK, Mushinzimana E, Omukunda E, Otsyula G, Minakawa N, Zhou G, Yan G, 2006. Population dynamics of malaria vectors in highlands and lowlands of western Kenya. *J Med Entomol* 43: 200-206.

12. Munyekenye GO, Githeko AK, Zhou G, Mushinzimana E, Minakawa N, Yan G, 2005. Spatial analysis of *Plasmodium* infections in western Kenya highlands. *Emerg Infect Dis* 10: 1571–1577.
13. Minakawa N, Munga S, Atieli F, Mushinzimana E, Zhou G, Githeko AK, Yan G, 2005. Spatial distribution of anopheline larval habitats in western Kenyan highlands: effects of land cover types and topography. *Am J Trop Med Hyg* 73: 157–165.
14. Githeko A, Ayisi J, Odada P, Atieli F, Ndenga B, Githure J, Yan G, 2006. Topography and malaria transmission heterogeneity in western Kenya highlands: prospects for focal vector control. *Malaria J* 5: 107.
15. Shanks GD, Hay SI, Stern DI, Biomndo K, Snow RW, 2002. Meteorologic influences on *Plasmodium falciparum* malaria in the highland tea estates of Kericho, Western Kenya. *Emerg Infect Dis* 8: 1404–1408.
16. Zhou G, Minakawa N, Githeko AK, Yan G, 2004. Association between climate variability and malaria epidemics in the east African highlands. *Proc Natl Acad Sci U S A* 101: 2375–2380.
17. Bodker R, Kisinza W, Malima R, Msangeni H, Lindsay SW, 2000. Resurgence of malaria in the Usambara mountains, Tanzania, an epidemic of drug-resistant parasites. *Glob Change Hum Health* 1: 134–153.
18. Anderson TJC, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurwa S, French N, Whitworth J, Velez ID, Brockman AH, Nosten F, Ferreira MU, Day KP, 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Mol Biol Evol* 17: 1467–1482.
19. de Roode JC, Pansini R, Cheesman SJ, Helinski ME, Huijben S, Wargo AR, Bell AS, Chan BH, Walliker D, Read AF, 2005. Virulence and competitive ability in genetically diverse malaria infections. *Proc Natl Acad Sci U S A* 102: 7624–7628.
20. Gilbert SC, Plebanski M, Gupta S, Morris J, Cox M, Aidoo M, Kwiatkowski D, Greenwood BM, Whittle HC, Hill AV, 1998. Association of malaria parasite population structure, HLA, and immunological antagonism. *Science* 279: 1173–1177.
21. Awadalla P, Walliker D, Babiker H, Mackinnon M, 2001. The question of *Plasmodium falciparum* population structure. *Trends Parasitol* 17: 351–353.
22. Nassir E, Abdel-Muhsin A-MA, Suliaman S, Kenyon F, Kheir A, Geha H, Ferguson HM, Walliker D, Babiker HA, 2005. Impact of genetic complexity on longevity and gametocytogenesis of *Plasmodium falciparum* during the dry and transmission-free season of eastern Sudan. *Int J Parasitol* 35: 49–55.
23. Collins WE, Nguyen-Dinh P, Sullivan JS, Morris CL, Galland GG, Richardson BB, Nesby S, 1998. Adaptation of a strain of *Plasmodium vivax* from Mauritania to New World monkeys and anopheline mosquitoes. *J Parasitol* 84: 619–621.
24. Babiker HA, Abdel-Wahab A, Ahmed S, Suleiman S, Ranford-Cartwright L, Carter R, Walliker D, 1999. Detection of low level *Plasmodium falciparum* gametocytes using reverse transcriptase polymerase chain reaction. *Mol Biochem Parasitol* 99: 143–148.
25. Miller LH, Roberts T, Shahabuddin M, McCutchan TF, 1993. Analysis of sequence diversity in the *Plasmodium falciparum* merozoite surface protein-1 (MSP-1). *Mol Biochem Parasitol* 59: 1–14.
26. Healer J, Murphy V, Hodder AN, Masciantonio R, Gemmill AW, Anders RF, Cowman AF, Batchelor A, 2004. Allelic polymorphisms in apical membrane antigen-1 are responsible for evasion of antibody-mediated inhibition in *Plasmodium falciparum*. *Mol Microbiol* 52: 159–168.
27. Babiker HA, Ranford-Cartwright LC, Currie D, Charlwood JD, Billingsley P, Teuscher T, Walliker D, 1994. Random mating in a natural population of the malaria parasite *Plasmodium falciparum*. *Parasitology* 109: 413–421.
28. Paul RE, Packer MJ, Walmsley M, Lagog M, Ranford-Cartwright LC, Paru R, Day KP, 1995. Mating patterns in malaria parasite populations of Papua New Guinea. *Science* 269: 1709–1711.
29. Ntoumi F, Contamin H, Rogier C, Bonnefoy S, Trape JF, Mercereau-Puijalon O, 1995. Age-dependent carriage of multiple *Plasmodium falciparum* merozoite surface antigen-2 alleles in asymptomatic malaria infections. *Am J Trop Med Hyg* 52: 81–88.
30. Contamin H, Fandeur T, Rogier C, Bonnefoy S, Konate L, 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.
31. Al-Yaman F, Genton B, Reeder JC, Anders RF, Smith T, Alpers MP, 1997. Reduced risk of clinical malaria in children infected with multiple clones of *Plasmodium falciparum* in a highly endemic area: a prospective community study. *Trans R Soc Trop Med Hyg* 91: 602–605.
32. Beck HP, Felger I, Huber W, Steiger S, Smith T, Weiss N, Alonso PL, Tanner M, 1997. Analysis of multiple *Plasmodium falciparum* infections in Tanzanian children during the phase III trial of the malaria vaccine SPf66. *J Infect Dis* 175: 921–926.
33. Sakihama N, Ohmae H, Bakote EB, Kawabata M, Hirayama K, Tanabe K, 2006. Limited allelic diversity of *Plasmodium falciparum* merozoite surface protein 1 gene from populations in the Solomon Islands. *Am J Trop Med Hyg* 74: 31–40.
34. Farnert A, Snounou G, Rooth I, Bjorkman A, 1997. Daily dynamics of *Plasmodium falciparum* subpopulations in asymptomatic children in a holoendemic area. *Am J Trop Med Hyg* 56: 538–547.
35. Conway DJ, Cavanagh DR, Tanabe K, Roper C, Mikes ZS, Sakihama N, Bojang KA, Oduola AMJ, Krensner 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.
36. Hughes AL, 1992. Positive selection and interallelic recombination at the merozoite surface antigen-1 (MSA-1) locus of *Plasmodium falciparum*. *Mol Biol Evol* 9: 381–393.
37. Hughes MK, Hughes AL, 1995. Natural selection on *Plasmodium* surface proteins. *Mol Biochem Parasitol* 71: 99–113.
38. Su XZ, Wellems TE, 1996. Toward a high resolution *Plasmodium falciparum* linkage map: polymorphic markers from hundreds of simple sequence repeats. *Genomics* 3: 430–444.
39. Anderson TJ, Su XZ, Bockarie M, Lagog M, Day KP, 1999. Twelve microsatellite markers for characterisation of *Plasmodium falciparum* from fingerprick blood samples. *Parasitology* 119: 113–126.
40. Wooden J, Kyes S, Sibley CH, 1993. PCR and strain identification in *Plasmodium falciparum*. *Parasitol Today* 9: 303–305.
41. Singh B, Bobogare A, Cox-Singh J, Snounou G, Shukri M, Rahman HA, 1999. A genus- and species specific nested polymerase chain reaction malaria detection assay for epidemiologic studies. *Am J Trop Med Hyg* 60: 686–692.
42. Ferreira MU, Nair S, Hyunh TV, Kawamoto F, Anderson TJC, 2002. Microsatellite characterization of *Plasmodium falciparum* from cerebral and uncomplicated malaria patients in southern Vietnam. *J Clin Microbiol* 40: 1854–1857.
43. Durand P, Michalakis Y, Cestier S, Oury B, Leclerc MC, Tibayrenc M, Renaud F, 2003. Significant linkage disequilibrium and high genetic diversity in a population of *Plasmodium falciparum* from an area (Republic of the Congo) highly endemic for malaria. *Am J Trop Med Hyg* 68: 345–349.
44. Bogreau H, Renaud F, Bouchiba H, Durand P, Assi S-B, Henry M-C, Garnotel E, Pradines B, Fusai T, Wade B, Adehossi E, Parola P, Kamil MA, Puijalon O, Rogier C, 2006. Genetic diversity and structure of African *Plasmodium falciparum* populations in urban and rural areas. *Am J Trop Med Hyg* 74: 953–959.
45. Oetting WS, Armstrong CM, Ronan SM, Young TL, Sellers TA, King RA, 1998. Multiplexed short tandem repeat polymorphisms of the Weber 8A set of markers using tailed primers and infrared fluorescence detection. *Electrophoresis* 19: 3079–3083.
46. Hill WG, Babiker HA, 1995. Estimation of numbers of malaria clones in blood samples. *Proc R Soc Lond B Biol Sci* 262: 249–257.
47. Kaneko O, Kimura M, Kawamoto F, Ferreira MU, Tanabe K, 1997. *Plasmodium falciparum*: allelic variation in the merozoite surface protein 1 in wild isolates from southern Vietnam. *Exp Parasitol* 86: 45–57.
48. Snounou G, Zhu X, Siripoon N, Jarra W, Thaitong S, Brown KN, Viriyakosol S, 1999. Biased distribution of *msp-1* and *msp-2* allelic variants in *Plasmodium falciparum* populations in Thailand. *Trans R Soc Trop Med Hyg* 93: 369–374.

49. Brockman A, Paul RE, Anderson TJ, Hackford I, Phaiphun L, Looareesuwan S, Nosten F, Day KP, 1999. Application of genetic markers to the identification of recrudescing *Plasmodium falciparum* infections on the northwestern border of Thailand. *Am J Trop Med Hyg* 60: 14–21.
50. Roper C, Richardson W, Elhassan IM, Giha H, Hviid L, Satti GM, Theander TG, Arnot DE, 1998. Seasonal changes in the *Plasmodium falciparum* population in individuals and their relationship to clinical malaria: a longitudinal study in a Sudanese village. *Parasitology* 116: 501–510.
51. Haubold B, Hudson RR, 2000. LIAN 3.0, detecting linkage disequilibrium in multilocus data. *Bioinformatics* 16: 847–848.
52. Schneider S, Roessli D, Excoffier L, 2000. *ARLEQUIN: A Software for Population Genetics Data Analysis*. User manual. Geneva: Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva.
53. Goudet J, 1995. FSTAT version 1.2: a computer program to calculate F-statistics. *J Hered* 86: 485–486.
54. Pritchard JK, Stephens M, Donnelly P, 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
55. Randi E, Tabarroni C, Rimondi S, Lucchini V, Sfougaris A, 2003. Phylogeography of the rock partridge (*Alectoris graeca*). *Mol Ecol* 12: 2201–2214.
56. Hamad AA, El Hassan IM, El Khalifa AA, Ahmed GI, Abdelrahim SA, Theander TG, Arnot DE, 2000. Chronic *Plasmodium falciparum* infections in an area of low intensity malaria transmission in the Sudan. *Parasitology* 120: 447–456.