Genetic Diversity of the Malaria Vaccine Candidate *Plasmodium falciparum* Merozoite Surface Protein-3 in a Hypoendemic Transmission Environment

Stephen J. Jordan, OraLee H. Branch, Jean Carlos Castro, Robert A. Oster, and Julian C. Rayner*

**Department of Cell Biology, University of Alabama at Birmingham, Birmingham, Alabama; William C. Gorgas Center for Geographic Medicine, Division of Infectious Diseases, Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama; Laboratorio de Investigaciones de Productos Naturales y Antiparasitarios, Universidad Nacional de la Amazonia Peruana, Iquitos, Peru; Division of Preventive Medicine, Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama**

**Abstract.** The N-terminal domain of *Plasmodium falciparum* merozoite surface protein-3 (PfMSP3) has been excluded from malaria vaccine development largely because of genetic diversity concerns. However, no study to date has followed N-terminal diversity over time. This study describes PfMSP3 variation in a hypoendemic longitudinal cohort in the Peruvian Amazon over the 2003–2006 transmission seasons. Polymerase chain reaction was used to amplify the N-terminal domain in 630 distinct *P. falciparum* infections, which were allele-typed by size and also screened for sequence variation using a new high-throughput technique, denaturing high performance liquid chromatography. PfMSP3 allele frequencies fluctuated significantly over the 4-year period, but sequence variation was very limited, with only 10 mutations being identified of 630 infections screened. The sequence of the PfMSP3 N-terminal domain is relatively stable over time in this setting, and further studies of its status as a vaccine candidate are therefore warranted.

**INTRODUCTION**

*Plasmodium falciparum* is responsible for between 300 and 500 million clinical episodes of malaria and between 1 and 3 million deaths annually, with ~90% of mortalities occurring in young children.1,2 With the continual threat of the emergence of drug-resistant *P. falciparum* strains, the need for a safe and effective malaria vaccine is more urgent than ever. Although a handful of *P. falciparum* vaccine candidates, such as merozoite surface protein 1 (MSP1),3–5 apical membrane antigen 1 (AMA1),6,7 and circumsporozoite protein (CSP),8–10 have been extensively studied for many years, there are many other promising vaccine candidates that have received much less attention. For many of these candidates, a pool that has radically increased with the recent rapid increase in genomic data,11 there are little, if any, supporting data, particularly regarding the candidate’s sequence diversity or role in the generation of protective immunity *in vivo*. Given the problems of sequence diversity and poor antigenicity that can afflict even the most well-studied vaccine candidates, there is critical need for the rapid and early characterization of novel vaccine candidate potential and the adoption of high-throughput techniques to facilitate this type of analysis. It is with such concerns in mind that a collaborative malaria vaccine roadmap was recently released, which identifies key issues that need to be addressed for a vaccine candidate to effectively progress through the development pipeline (www.malaria vaccineroadmap.net). With the increasing cost of vaccine clinical trials, this roadmap proposes that strict go/no-go criteria be enforced on candidate antigens (i.e., studies be conducted early that characterize the candidate’s vaccine potential and either lay the groundwork for subsequent studies or remove it from the pool of viable candidates).

*Plasmodium falciparum* merozoite surface protein-3 (PfMSP3) is a major vaccine candidate that has not yet advanced to Phase II field trials. PfMSP3 is a highly immunogenic non-integrated protein expressed on the surface of merozoites and has been suggested to be involved in erythrocyte binding, although its function remains unknown.12,13 Structurally, it possesses three blocks of heptad repeats that are proposed to form alpha-helical coil/coil domains, a hydrophobic glutamine-rich domain, and a putative leucine zipper domain at the extreme C terminus.14,15 Importantly, it is a major target of antibodies from malaria-immune African individuals, and immunization studies using full-length PfMSP3 elicited significant protection in a non-human primate model.16–18 PfMSP3 exists as two allele classes, termed K1 and 3D7 after the *P. falciparum* strains in which they were first identified. The majority of both intra- and interallele differences are localized to the heptad repeat region, which defines the N-terminal domain. Variation in the PfMSP3 N-terminal domain occurs largely between the blocks of heptad repeats and consists of both indels and SNPs. Conversely, the C-terminal domain, comprised of the glutamine-rich region and the leucine zipper motif, is almost entirely conserved.19 Although the presence of variation in the PfMSP3 N terminus has been established for some time using laboratory-adapted *P. falciparum* isolates,20 only one study has previously looked at natural variation in PfMSP3 sequences, with 48 and 50 PfMSP3 genotypes sequenced from samples in Nigeria and Thailand, respectively.20 PfMSP3 is therefore a strong vaccine candidate with limited epidemiologic data; data that are needed to support its continued development along the proposed malaria vaccine roadmap.

This study investigates genetic variation in the PfMSP3 N-terminal domain in samples collected from the Malaria Immunology and Genetics in the Amazon (MIGIA) cohort study, a longitudinal study of *P. falciparum* transmission in Zunagarococha, which is a cluster of four villages located in the Peruvian Amazon. In this community, *P. falciparum* transmission is hypoendemic, with inoculation rates of <0.04 infections per person per month during the 7-month transmission season, with most infections consisting of a single genetic type.21 Although transmission rates are low, there is still a significant level of overall genetic diversity, with at least five different genetic types defined for PfMSP1 Block 2 (P Sutton and OH Branch, unpublished data).

To date, no study has yet looked at the extent of PfMSP3 variation in hypoendemic settings such as in South America. Such data are important because the increased frequency of simple infections in such a setting enables us to look at changes in allele frequency over time, which might provide evidence for...
or against the presence of allele- and variant-specific immune responses. To characterize genetic variation within PIMSP3 in a hypoendemic setting, we initiated a retrospective study using blood samples obtained from *P. falciparum*-infected individuals enrolled in the MIGIA cohort study. To significantly increase our power and enable us to screen hundreds of samples, we adapted the recently developed high throughput denaturing high performance liquid chromatography (dHPLC) genotyping technology to screen for PIMSP3 variation. 22–23 The existence of extensive clinical data generated by the MIGIA study also allowed us to test for correlates of protection with specific PIMSP3 genotypes. This study is therefore a comprehensive analysis of PIMSP3 genotypes across multiple transmission seasons and is an important window into the genetic diversity dynamics of this important vaccine candidate in a hypoendemic setting.

MATERIALS AND METHODS

Study location and malaria transmission. A detailed description of the four villages that comprise the longitudinal cohort study has been previously published. 21 Briefly, the study enrolled residents from four villages of the Zungerococha community: Zungerococha village, Puerto Almendra, Narumaru, and Llanchama. The community is located South of Iquitos in the Peruvian Amazon and is supported by the Nanay River, a tributary of the Amazon River. The village of Narumaru contains a port that brings boats from outside the community. The village residents have homogenous housing construction, income levels, and access to healthcare, provided by the MIGIA cohort physicians. The women generally work in or near their home, and the majority of men work as either community fishermen or in local agriculture. Importantly, travel outside the community is rare, with the most frequent travel being to the city of Iquitos, where malaria transmission is nonexistent.

The Zungerococha community was chosen as the focus of the MIGIA cohort because of the presence of continuing stable hypoendemic transmission of both *P. vivax* and *P. falciparum*, introduced during an epidemic in 1992. The annual malaria season lasts 7 months and closely follows the rainy season, which occurs between January and July. The major vector of malaria transmission is *Anopheles darlingi*, an anthropophilic species that has proliferated in the area since the early 1990s. 26

Sample collection and extraction. All individuals enrolled in the MIGIA study have access to healthcare provided at a local health post located in Zungerococha village. On presentation to the health post, symptomatic individuals were tested for malaria parasites by Giemsa-stained microscopy. If positive, and consent was given, patients submitted a 0.5-mL blood sample by finger venipuncture and underwent a comprehensive medical evaluation. PCR verification of the microscopy result was subsequently performed using species-specific primers. *Plasmodium* infection was immediately treated with a combination of mefloquine (0.5 mg/kg for 7 days) and artemunate (4 mg/kg daily for 3 days).

*Plasmodium falciparum* infections were considered asymptomatic if the medical history and physical exam showed no malaria-related signs or symptoms, and the patient had not presented to the clinic. Clinical presentation or symptomatic infections identified in the community were labeled symptomatic infections.

The submitted blood sample was separated into a serum and packed red blood cell (RBC) fraction by centrifugation, and each was cataloged and stored at −80°C. *P. falciparum* DNA was extracted using a commercially available kit (Qiagen) and stored at −80°C until needed.

Only *P. falciparum* infections were included in this study, which spanned 4 years, from 2003 to 2006. Samples were chosen at random, but any infection occurring within 60 days of the initial infection date was excluded to minimize the risk of parasite recrudescence caused by incomplete drug clearance.

Nested PCR and allele-typing. Nested PCR was used to amplify the N-terminal region of PIMSP3 using the external PCR primers 5'-ATAATTGTCGTAGTAAGAAATTG-3' and 5'-AATACATCATCATTTCCTTAG-3' and the internal primers 5'-ATAATCTTAACCTAAAGAAATGC-3' and 5'-CGGCGGGGGCGATAGACATTATTCCGCC-3'. The bold annotation identifies the CG clamp, which was added to the 5' end of the internal reverse primer to facilitate the subsequent analysis by dHPLC. *P. falciparum*-infected patient extracted DNA (1.5 µL) was amplified using ChoiceTaq (Denville, South Plainfield, NJ) and underwent 35 cycles of 95°C for 30 seconds, 50°C for 30 seconds, and 65°C for 1 minute, followed by a final elongation step of 5 minutes at 65°C. The nested PCR reaction involved the same conditions except that 1 µL of the external PCR reaction was used as the template. Classification of the PIMSP3 allele was performed by size migration using 2% agarose gel electrophoresis supplemented with ethidium bromide.

dHPLC and sequencing. dHPLC analysis was performed using a WAVE column (Transgenomics, Omaha, NE) and standard protocols. Before screening, optimal WAVE temperatures for each PIMSP3 allele were determined using Wavemaker software (Transgenomics). The nested PCR mixture (2 µL) was diluted 1:10 in ddH2O and submitted for dHPLC analysis. The submitted sample was mixed with a control sequence from the homologous allele, and heterologous strand formation was performed by denaturation and subsequent reannealing of the mixture. The sample was bound to the WAVE column and eluted using a linear acetonitrile gradient. The presence of a second elution peak on the elution chromatograph was indicative of a mutation, and all detected mutants were immediately sequenced to identify the mutation. As positive controls, 3D7 and Borneo were complexed with the Peruvian alleles HB3 and K1, respectively. As an additional quality control step and to define the sensitivity and specificity of the technique, 10% of all samples that exhibited no mutations by dHPLC were randomly sequenced.

Ethics committee approval. This study was approved by review boards of the University of Alabama at Birmingham, Universidad Peruana Cayetano Heredia, and the Peruvian Ministerio de Salud, Instituto Naccional de Salud. Written consent was obtained from all participants before study enrollment.

Statistical analysis. Descriptive statistics, such as frequencies, percentages, means, and SDs, were used to summarize all study variables of interest. Comparisons between proportions of alleles (K1 and 3D7) for communities, mutant, sex, symptom status, subsequent allele, subsequent *P. falciparum* infections, year of infection, age group (< 15 and ≥ 15 years of age), time group to next *P. falciparum* infection (0–6 and ≥ 6 months), and comparisons between all other proportions of interest were performed using the two-group χ² test or Fisher exact test when the assumptions on the χ² were not tenable. Comparisons between means of actual age and days to
next *P. falciparum* infection were performed using the usual two-group *t* test or the two-group *t* test for unequal variances when needed. All statistical tests were two-sided and were performed using a 5% significance level (i.e., *α* = 0.05). SAS software (version 9.1.3; SAS Institute, Cary, NC) was used to perform all statistical analyses.

**RESULTS**

**PfMSP3 allele frequency varies significantly between subsequent years and between villages.** The Zungarococha community consists of four villages each ~2–3 km apart and supported by tributaries of the Amazon River: Zungarococha village (ZG, *N* = 805), Puerto Almendra (PA, *N* = 272), Ninariumi (NR, *N* = 590), and Llanchama (LL, *N* = 203) (Figure 1A). *P. falciparum* transmission is hypoendemic at this study site, transmitted largely by the anthropophilic vector *An. darlingi*. *P. falciparum* transmission is seasonal, occurring largely between January and August of each year, with entomologic inoculation rate (EIR) of 0.13 infections/person/mo during the 7-month transmission season. 21

![A](image)

A total of 630 *P. falciparum* infections were randomly chosen for inclusion in this study: 101 from 2003, 174 from 2004, 250 from 2005, and 105 from 2006 (Figure 1B). Samples were chosen to give the widest possible distribution across all four villages, although numbers for each village were not equal because the villages do not carry an equal burden of infections: although Zungarococha village is the largest village in the community, it had the least number of infections included in this study. The highest number of infections sampled was from Ninariumi, the second largest village.

Nested PCR was used to amplify and genotype *PfMSP3* from *P. falciparum*-infected individuals. The primers amplified the *PfMSP3* N-terminal domain, where the majority of genetic diversity has been shown to occur 22 (nucleotides 117–507 in the 3D7 strain, excluding primer sequence) and the allele class identified using agarose gel electrophoresis. *PfMSP3* allele class status. Interestingly, Puerto Almendra showed a significant decrease (> 16-fold, *P* = 0.002) in K1 class infections compared with the other villages. Additionally, Llanchama, the smallest of the villages, was associated with a significant decrease (> 16-fold, *P* = 0.0001) in K1 class infections compared with the other villages, with only a single K1 class infection being detected across all four transmission season.

**dHPLC can detect single nucleotide differences in PfMSP3 sequence.** Because of the large number of samples allelotype in this study, we adopted a high-throughput approach to screen for single nucleotide polymorphisms (SNPs) and indels within each amplicon: dHPLC. To our knowledge, this is the first time that dHPLC has been used to screen for SNPs in *P. falciparum* genes. As described in the Materials and Methods section, the technique detects mutations through the formation of heteroduplex bubbles. Briefly, a test amplicon is denatured and reannealed in the presence of a reference amplicon of known sequence. If the test and reference sequences are identical, only homoduplexes form, but if there are SNPs or indels in the test sequence that are not found in the reference sequence, heteroduplexes will form. The reannealed fragments are bound to a WAVE column (Transgenomics) and

---

**Figure 1.** *P. falciparum* transmission in Zungarococha from 2003 to 2006. **A**, Schema of Zungarococha, a small community of four villages fed by the Nanay and Itaya river systems: Zungarococha village (ZG, *N* = 805), Puerto Almendra (PA, *N* = 272), Ninariumi (NR, *N* = 590), and Llanchama (LL, *N* = 203). The community health post is located in Zungarococha village. **B**, Zungarococha community *P. falciparum* infection demographics from 2003 to 2006. A total of 630 *P. falciparum* infections, spanning all villages, were analyzed in this study.
eluted with a linear acetonitrile gradient. Homoduplexes elute from the column as a single peak. However, if a heteroduplex bubble is present, it will elute at a different acetonitrile concentration and therefore form a second peak distinct from the homoduplex peak. Any samples that elute as two peaks from the WAVE column would be sequenced to confirm the presence and identity of the SNP(s) and/or indels. dHPLC is therefore a rapid method to screen for deviations from a reference sequence, and because the cost is significantly lower (10% or less) than sequencing, dHPLC is of particular use when the expected SNP frequency is low—it allows the user to rapidly screen large numbers of samples and only sequence those that are likely to contain SNPs.

We initially tested the sensitivity of dHPLC using PfMSP3-specific primers containing a 3’-CG clamp by mixing the HB3 allele either with itself or the same fragment amplified from genomic DNA from the 3D7 strain. Both HB3 and 3D7 are of the same allele class, but 3D7 has previously been shown to contain a single SNP relative to HB3. dHPLC analysis showed a single peak when HB3 was reannealed only to itself but two elution peaks when 3D7 and HB3 alleles were reannealed (Figure 3A). Sequencing the 3D7 amplicon confirmed the single SNP, showing that dHPLC has the sensitivity to detect single nucleotide changes in PfMSP3.

Mixing K1 sequences with those from a different K1-class allele, genomic DNA from the Borneo P. falciparum strain (a kind gift from John Barnwell) also showed the presence of a heteroduplex. Sequencing the Borneo fragment showed that K1 and Borneo differ by 37 nucleotides, establishing a range of detection for the PfMSP3-specific dHPLC assay of 1–37 nucleotide differences. The sensitivity of dHPLC withstood wide changes in either reference or unknown amplicon concentrations (data not shown) as has
been previously noted, increasing its utility with field \textit{P. falciparum} samples where DNA concentrations can vary.

**PfMSP3 coil/coil domain sequence variation is limited.**

To characterize the level of sequence variation within the N-terminal domain of PfMSP3, all 630 samples that had been allele-typed were subjected to dHPLC analysis. All 3D7 allele class samples were mixed with HB3 sequences as a reference, again because our initial sequencing had identified this as the predominant K1 class allele. Of the 630 samples screened, dHPLC detected heteroduplexes in 14 samples, all in the 3D7 allele class. All 14 samples were sequenced: 3 were false positives with no detectable mutation, whereas 11 samples had mutations that were different to the HB3 reference sequence (10 true mutations and 1 mutation that had double overlapping peaks at one position on the sequencing chromatogram and therefore presumably resulted from Taq polymerase error). Eighty-six samples that were mutation negative by dHPLC were also submitted to direct sequencing to test for sensitivity, and all 86 samples were found to be identical to the reference sequence, as predicted by dHPLC. We therefore estimate that using dHPLC to screen for PfMSP3 variation has a sensitivity of 100% and a specificity of 96.6%. Sensitivity is calculated by the equation \( A/(A + C) \), where \( A \) is the number of true positives in the dHPLC-mutant negative group (11) and \( C \) is the number of true positives in the dHPLC-mutant negative group (0). Specificity is calculated by the equation \( D/(B + D) \), where \( D \) is the number of true negatives in the dHPLC-mutant negative group (86) and \( B \) is the number of true negatives in the dHPLC-mutant positive group (3).

In 630 samples screened, we therefore found SNPs in only 10 samples (Figure 3B), all of which were 3D7-class alleles. Direct sequencing of these 10 mutants established that all were identical and contained the same SNP: a single C→T nucleotide substitution at position 203, which resulted in a non-synonymous mutation L68S. Interestingly, this is the SNP that differentiates the HB3 and 3D7 PfMSP3 sequences. All 10 samples were collected during the 2003 transmission season, but they were distributed widely across the community: 3 occurred in Zungarococha village, 5 in Ninarumi, and 2 in Puerto Almendra. No significant correlation was detected between these samples and age, sex, symptom status, or village (data not shown). To rule out the possibility that the detected mutations were caused by the introduction of a completely new strain into the Zungarococha community, another polymorphic loci, \textit{PfMSP2}, was amplified and sequenced from 6 of these 10 samples. Three of the \textit{PfMSP2} genes were identical to the HB3 \textit{PfMSP2}, but three \textit{PfMSP2} genes were significantly different (data not shown), suggesting that not all \textit{PfMSP3} mutant strains shared the same origin.

**Assessing for evidence of protection for PfMSP3 alleles.**

The longitudinal MIGIA cohort study uses active case detection in addition to passive case detection. This allows for the identification of not only symptomatic infections but also asymptomatic infections that remain undiagnosed in the absence of active detection. Because premunition is characterized by the absence of malaria symptoms in combination with an inability to clear \textit{P. falciparum} parasites, asymptomatic infections in Peruvian individuals might indicate protection against \textit{P. falciparum}. We therefore compared our \textit{PfMSP3} dHPLC genotyping data, which identified all infections as either HB3, 3D7, or K1 alleles, with the available epidemiologic data to analyze whether there was any propensity for a specific allele type to be associated with a specific clinical outcome.

Because of the hypoendemic transmission and low rate of mixed infections at this study site, it is possible to follow subsequent infections in single individuals that are separated by time. We therefore assessed whether the \textit{PfMSP3} allele frequency in these subsequent infections in the same individual was any different to the population level allele frequency. If there is strong allele-specific immunity to the PfMSP3 N-terminal domain, one would expect that allele frequencies would change in subsequent infections—that is, an HB3 infection would be more likely to be followed by a K1 infection or vice versa. There were 106 successive infection pairs (individuals with two \textit{P. falciparum} infections in a < 500-day period) in the 630 infections for which we had genotyped \textit{PfMSP3}. Comparing the genotypes for the first and second infections in these individuals, we found no statistically significant deviation from whole population allele frequencies—that is, there was no trend for an infection of one allele to be followed by an infection of another allele, as would be expected if protection was largely allele specific (Table 1A; \( P = 0.93 \)). Additionally, we tested for the possibility that infection with either allele class would lead to a decrease in subsequent \textit{P. falciparum} infections, but this relationship was not significant (\( P = 0.17 \)).

![Image](image.png)
TABLE 1
Assessing for allele-specific protection in the Zungarococha community: (A) assessing for an association between PfMSP3 allele and subsequent allele, the potential to experience a subsequent P. falciparum infection, and the time-to-next-infection failed to identify a significant association; and (B) comparison of PfMSP3 allele-typed infection and subsequent infection symptom status

<table>
<thead>
<tr>
<th>Subsequent allele</th>
<th>HB3</th>
<th>K1</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>HB3</td>
<td>86</td>
<td>13</td>
<td>0.9309</td>
</tr>
<tr>
<td>K1</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Subsequent Pf infection</td>
<td>0.1651</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>412</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>161</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Next infection (days)</td>
<td>0.5179</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>263</td>
<td>237</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>73</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

**A**

<table>
<thead>
<tr>
<th></th>
<th>Asym</th>
<th>Sym</th>
<th>%Asym</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterologous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HB3:HB3</td>
<td>26</td>
<td>42</td>
<td>38.2%</td>
<td>0.4335</td>
</tr>
<tr>
<td>K1:K1</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
<td></td>
</tr>
<tr>
<td>Homologous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HB3:K1</td>
<td>1</td>
<td>2</td>
<td>33.3%</td>
<td>0.9126</td>
</tr>
<tr>
<td>K1:K1</td>
<td>3</td>
<td>7</td>
<td>30.0%</td>
<td></td>
</tr>
<tr>
<td>Total Homologous</td>
<td>26</td>
<td>42</td>
<td>38.2%</td>
<td>0.6095</td>
</tr>
<tr>
<td>Total Heterologous</td>
<td>4</td>
<td>9</td>
<td>30.8%</td>
<td></td>
</tr>
</tbody>
</table>

No significant association was found that suggested any shift in the asymptomatic rate.

The propensity of individuals to experience another P. falciparum infection within a 6-month period was also used to assess for possible protection, as well as assessing for an increase in the “time to next infection.” Because the infection rate has remained relatively steady at 0.13 infections/person/mo during the 7-month transmission season, any significant decrease in infection rates in individuals infected with either allele would be suggestive of immune-mediated protection. As shown in Table 1A, no significant delay in time-to-next-infection was detected for either homologous or heterologous allele infections. There was, however, a non-significant trend that K1-infected individuals had a decreased “time to next infection” value (on average 26 days earlier, $P > 0.05$). However, because we showed above that the K1 allele class frequency is significantly associated with community, any difference is most likely caused by transmission differences between the villages.

We also tested for any association that might exist between the presence of either PfMSP3 allele and the rate of asymptomatic infections in subsequent infections. If the asymptomatic rate is higher in subsequent infections with the homologous antigen (i.e., subsequent HB3 infection after a primary HB3 infection), immune selection might be a significant factor in mediating this phenomenon. The rate of asymptomatic infections in subsequent heterologous infection is also immensely informative because, if rates are similar to homologous infections, it suggests some degree of cross-protection. Conversely, a significant increase in asymptomatic infections found in homologous infections but absent in asymptomatic heterologous infections would suggest a significant strain-specific immune response. Compared with primary infections, no statistical significant difference in HB3 or K1 asymptomatic rates was detected for either homologous or heterologous subsequent infections (Table 1B). Even after pooling the infections into homologous/heterologous subsequent infections, we were unable to detect a significant increase in asymptomatic infections for either group.

**DISCUSSION**

In this study, we described the use of dHPLC to screen for frequency within the malaria vaccine candidate PfMSP3 in a hypoendemic setting. Because dHPLC has been shown to exhibit a sensitivity and specificity of up to 100% in some tests, possesses a high-throughput capacity, and is considerably more cost-efficient compared with direct sequencing, we decided to adapt it to screen for PfMSP3 N-terminal domain variation. Even with our resequencing of both mutation positive and negative samples for specificity and sensitivity analysis, our overall costs were only 25% of the cost of direct sequencing all 630 isolates. We therefore believe that dHPLC is an effective high-throughput approach for P. falciparum genotyping, particularly in hypoendemic transmission regions where genetic variation is suspected to be low. Although we applied dHPLC here to study a vaccine candidate antigen, another obvious application would be for drug resistance screening, where SNPs associated with drug resistance in PfCRT, PfDHFR, and PfDHPS are well established.

Using nested PCR, we allele typed 630 individual P. falciparum infections and found that both 3D7 and K1 alleles classes of PfMSP3 were present in the Zungarococha community. Direct sequencing of those samples identified 3D7 class alleles as identical to the previously published sequences for HB3 and K1 class alleles as identical to the K1 sequence itself. Interestingly, 3D7 class alleles were dominant throughout the community and, over time, became increasingly more dominant. Between 2003 and 2006, the allele frequency shifted significantly, with 3D7 class percent increasing from 80.9% to 99% of all infections. Although the possibility exists that the observed changes in allelic frequency are caused by extraneous factors such as genetic drift, which is a major contributor to allele variation in small populations, we believe our study population is sufficiently large that the effects of genetic drift are minimal. Such large-scale and directional changes in allele frequencies over such a short time frame are unlikely to be caused by genetic drift, but this remains a formal possibility. Thus, it seems most likely that an as of yet unidentified selection pressure is responsible for driving the observed allele frequency variation. Importantly, if this selection pressure is immunologically mediated, the target of that response is most likely against the N-terminal domain, because the C-terminal domain is almost entirely conserved between both alleles.

Screening for sequence variation identified only 10 mutations in PfMSP3 in 630 infections. All mutations occurred within the 2003 transmission season and were identical, a C>T non-synonymous substitution at position 203. It is interesting to note that this is the same SNP that was most commonly observed in other global studies, being present in 48% (20 of 42 total) of 3D7 allele class variants in Nigeria and Thailand. The frequency and global distribution of this SNP suggests a globally conserved function, perhaps in immune evasion. Although the possibility exists that the lack of PfMSP3 variation that we observed in the Zungarococha community is caused by an absence of genetic diversity in this setting, work by Sutton and others (unpublished data) has shown that MSP1-block 2 variation in the Zungarococha community is
strains difficult to distinguish. In contrast, the hypoendemic
tion of new polymorphisms from infections with heterologous
vironments, the high inoculation rate makes the differentia-
should be considered hypervariable. In such hyperendemic
sity and does not necessarily mean that the N-terminal domain
resent a much larger pool of population-level genetic diver-
shown more extensive PfMSP3 N-terminal domain variation
of PfMSP3 in our study is significant. Previous reports have
ments, the observed stable nature of the N-terminal domain
a much larger pool of population-level genetic diver-

A significant strength of the MIGIA cohort study is the use of
active case detection in addition to passive detection of P. falciparum
fections. Such infections, marked by a lack of malaria-
related symptoms in the presence of active parasitemia, mimic
tection in malaria-immune individuals, a phenomenon
 termed premunition, and might predict malaria-immune status
these individuals. Thus, the use of asymptomatic individuals,
detected through active case detection, allows us to assess for
correlates of protection for P. falciparum alleles.

To identify correlates of protection for PfMSP3 alleles, we
combined our PfMSP3 allele data with the available epide-
mologic data. Evidence of protection was assessed by testing
whether the presence of either PfMSP3 allele correlated with an
increased frequency of subsequent allele infections, the rate of
subsequent infections, or an increase in the rate of subsequent
asymptomatic infections. Although we were unable to find sig-
nificant correlations of protection for either PfMSP3 allele,
this outcome does not rule out the potential for allele-specific
protection. Such correlates of protection for PfMSP3 alleles,
although important to assess for, are indirect outcomes of host
protection and infrequent P. falciparum infections, which define
the hypoendemic environment, make the identification of sig-
nificant correlations unfavorable because of a lack of statistical
power. More direct tests, such as directly testing antibody levels,
will be critical to clearly identifying the role that PfMSP3 plays
in mediating immunity in the hypoendemic setting.

In terms of PfMSP3 vaccine development, our central find-
ing of genetic stability in the PfMSP3 N-terminal domain
supports a reassessment of the inclusion of this domain in a
PfMSP3-based P. falciparum vaccine, especially because
genetic variation has been the major factor excluding
N-terminal domain-based vaccine development. Because a
large population of the world lives in hypoendemic environ-
ments, the observed stable nature of the N-terminal domain
of PfMSP3 in our study is significant. Previous reports have
shown more extensive PfMSP3 N-terminal domain variation
in the hyperendemic setting, but this variation may simply rep-
resent a much larger pool of population-level genetic diver-
sity and does not necessarily mean that the N-terminal domain
should be considered hypervariable. In such hyperendemic
environments, the high inoculation rate makes the differentia-
tion of new polymorphisms from infections with heterologous
strains difficult to distinguish. In contrast, the hypoendemic

significant and is similar in spectrum, although not in scale,
to diversity seen in a hyperendemic environment (P Sutton
and OH Branch, unpublished data). Additionally, microsatel-
lite studies by Branch and others suggested extensive diversity
within the community (OH Branch, personal communication).
This implies that the lack of variation in PfMSP3 is likely not
caused by a lack of genome-wide diversity in the community
but by a combination of the relative stability of the N-terminal
domain antigen and the relatively recent introduction of
P. falciparum in the Zungarococha community. Subsequent
studies of PfMSP3 mutants by sequencing the highly poly-
morphic PfMSP2 showed that three of those infections con-
tained PfMSP2 sequences that were significantly different
than HB3, suggesting that the introduction of new 3D7 class
strains into the community can explain some, but perhaps not
all, PfMSP3 mutations. It is likely that the introduction of new
P. falciparum strains occurred through the naval port located
in Ninarumi. In support of this, all three PfMSP2-variant infec-
tions occurred in Ninarumi (data not shown).

REFERENCES
1. Breman JG, 2001. The ears of the hippopotamus: manifestations,
determinants, and estimates of the malaria burden. Am J Trop
Med Hyg 64: 1–11.


