SHORT REPORT: RARE \textit{PLASMODIUM FALCIPARUM} MEROZOITE SURFACE PROTEIN 1 19-KDA (MSP-1\textsubscript{19}) HAPLOTYPES IDENTIFIED IN MALI USING HIGH-THROUGHPUT GENOTYPING METHODS

SHANNON L. TAKALA, DAVID L. SMITH, MAHAMADOU A. THERA, DRISSA COULIBALY, OGOBARA K. DOUMBO, AND CHRISTOPHER V. PLOWE*

Center for Vaccine Development, University of Maryland School of Medicine, Baltimore, Maryland; Fogarty International Center, National Institutes of Health, Bethesda, Maryland; Malaria Research and Training Center, University of Bamako, Bamako, Mali

Abstract. Genetic diversity in malaria vaccine antigens may compromise malaria vaccine efficacy, so it is important to understand this diversity and the processes that generate it. By applying new high-throughput genotyping methods to a large sample of infections from Mali \((N = 1369)\), seven new 19-kDa merozoite surface protein \(1\) (MSP-1\textsubscript{19}) haplotypes were identified. Herein we report the sequences of these new haplotypes and discuss their possible origins. Although they are present in < 1% of the samples examined, the existence of these rare haplotypes reveals a greater degree of diversity at this locus than previously reported and highlights the potential for \textit{Plasmodium} to evolve under selective pressure from the immune system and from such interventions as vaccines and drugs.

INTRODUCTION

Progress toward a malaria vaccine has been slow due in part to the extensive genetic variability in \textit{Plasmodium}. Such genetic variability is generated through mutation under selective pressure from the human immune system and sexual recombination in the mosquito vector, and is particularly prevalent in surface antigens being targeted for malaria vaccines.\(^1\) Merozoite surface protein \(1\) is a candidate antigen for a blood-stage malaria vaccine. The 195-kDa precursor of this protein undergoes two rounds of proteolytic cleavage, leaving only the C-terminal 19 kDa on the surface of the merozoite as it invades the erythrocyte.\(^2\) MSP-1\textsubscript{19} contains two epidermal growth factor (EGF)-like motifs that are thought to play a role in erythrocyte invasion.\(^3\) Antibodies to this region can block erythrocyte invasion in vitro\(^2\) and are associated with protection from clinical malaria in field studies.\(^4-9\) The sequence of MSP-1\textsubscript{19} is highly conserved;\(^10\) however, six non-synonymous single nucleotide polymorphisms (SNPs) have been documented at amino-acid positions 1644, 1691, 1699, 1700, 1701, and 1716,\(^10-14\) and it is unclear how these polymorphisms affect immunity. Intragenic recombination has been proposed as an important mechanism for generating novel genetic variants in MSP-1\textsubscript{19}; however, new variants can also be derived from single-nucleotide mutations that are maintained by positive natural selection.\(^1\)

If vaccine efficacy is allele-specific, then vaccination against polymorphic antigens could lead to selection for nontarget alleles in the parasite population, compromising vaccine efficacy. Such vaccine-induced selection has been suggested by theoretical studies\(^15-19\) and has been observed in a clinical trial of a blood-stage vaccine.\(^20\) We report the sequence of seven rare MSP-1\textsubscript{19} haplotypes identified at a malaria vaccine-testing site in Mali and discuss the possible origins of these new haplotypes and the potential implications of genetic diversity for the efficacy of MSP-1-based vaccines.

New MSP-1\textsubscript{19} haplotypes were identified from samples collected in a cohort study conducted in Bandiagara, Mali, during the years 1999–2001.\(^21\) From July to January of each year, individuals were visited weekly and contributed a filter paper blood sample at least monthly and at every clinical malaria episode. Among 629 study participants, 100 who had at least 2 years of follow up were randomly selected within three age strata: 30 children of age ≤ 5 years, 32 children of age 6–10 years, and 38 children of age ≥ 11 years.\(^21\) Samples were collected under protocols reviewed and approved by Institutional Review Boards of the University of Maryland School of Medicine and the University of Bamako Faculty of Medicine. Informed consent was obtained from all study participants or their guardians.

MSP-1\textsubscript{19} was amplified from samples collected at monthly surveys and clinical episodes occurring during the transmission season in the 3 years of the incidence study. A single PCR was used to amplify MSP-1\textsubscript{19} from samples with parasitemia > 1,000 parasites/µL, and a nested PCR was used to amplify MSP-1\textsubscript{19} from samples with parasitemia < 1,000 parasites/µL and microscopy-negative samples. Of the 2309 samples that underwent PCR (including microscopy-negative samples), 1375 were parasite-positive (by PCR).\(^21\)

All PCR-positive samples underwent Pyrosequencing to determine allele frequencies at each of the six SNPs in MSP-1\textsubscript{19}.\(^22\) Pyrosequencing (Biotage, Charlottesville, VA) is a high-throughput method that allows quantification of the proportions of alternative nucleotides at each SNP. Of the 1,375 PCR-positive samples, 1,369 gave successful MSP-1\textsubscript{19} genotyping results.\(^21\)

A mathematical model was used to estimate the frequency of 14 confirmed MSP-1\textsubscript{19} haplotypes in each genotyped sample.\(^21,22\) The haplotype-estimating algorithm uses maximum likelihood methods to determine the most probable combination of haplotypes given the allele frequencies for an infection, the haplotypes known to be circulating in the population, and a probability distribution of the measurement errors. Three of the 14 haplotypes included in the haplotype-estimating algorithm had not been observed previously and were confirmed by reamplification of MSP-1\textsubscript{19} followed by PCR cloning.\(^22\) When applied to the 1,369 genotyped samples, the algorithm was able to resolve haplotype frequencies for all but six samples given the list of 14 haplotypes.\(^21\)

The six samples that were unable to be resolved by the algorithm had allele frequencies consistent with the presence of additional new MSP-1\textsubscript{19} haplotypes. The PCR and Pyrose-
quencing were repeated to rule out genotyping error as an explanation for the observed allele frequencies. Upon obtaining the same results, MSP-119 was reamplified from these six samples, using non-biotinylated MSP-119 primers, and cloned. Twelve clones were picked from each transformation. Pyrosequencing and direct sequencing of MSP-119 clones from the six samples revealed four additional new haplotypes: QKSNRF, QKSSGF, EKSNRL, and EKNNGF. Nucleotide and amino-acid alignments of the seven new haplotypes (three from the previous study22 and four from the current study) are shown in Figure 1, as well as haplotypes with substitutions at positions other than those at the six known SNPs (sequences available in GenBank, accession numbers DQ677569–DQ677579). Sequences for the 3D7 and FVO P. falciparum strains are also included for reference. As indicated in the figure, the QKSSGF haplotype also has non-synonymous substitutions at positions 1674 (N → D) and 1690 (A → V). Additional non-synonymous substitutions were observed in other clones at residues 1647 (G → R, clones 0818c1-6 and 1-8), 1673 (E → G, clone 0818c1-6), and 1677 (P → L, clone 0818c1-7). The QKSSRL haplotype has a synonymous substitution at codon 1667. Synonymous substitutions were also observed in other clones at codons 1659 (0818c1-6), 1666 (0818c4-1), 1698 (0818c1-7, 4-1), and 1699 (0818c1-6, 1-8). It is possible that other polymorphic sites exist in the parasites infecting the cohort; however, because Pyrosequencing™ genotypes the short regions surrounding known SNPs, only those samples giving unusual Pyrosequencing™ results were flagged for cloning and direct sequencing.

Using the haplotype-estimating algorithm, QKSSRL, QKSSGF, and QTSSRF had prevalences of 0.07%, 1.3%, and 0.51%, respectively, in the cohort. QKSNRF, QKSSGF, EKSNRL, and EKNNGF were not included in the haplotype-estimating algorithm and were found in one sample each. Table 1 contains a comprehensive list of MSP-119 haplotypes reported in the literature. Including the seven new haplotypes identified in Mali, 22 haplotypes have been documented, including one isolate from India that contained a Y allele at position 1700.23 All but four of the reported haplotypes (EKSSGL, QTSSRF, ETSSRF, and EKSYGF) have been observed in the samples from Mali.

The role of recombination in the generation of genetic diversity in Plasmodium has been debated24,25, but intragenic recombination has been implicated as a factor in generating diversity.
diversity in MSP-1. Three of the haplotypes observed in this study (QKSSRL, QKSSGL, and EKSNRL) were predicted to exist based on single and double crossover events between previously identified alleles; however, until now they had not been identified in field isolates. Table 2 shows how the seven new haplotypes identified in this study could have arisen via recombination events. As indicated in the table, all but one of the seven new haplotypes (QKSNRF) could have arisen from single crossover events between haplotypes observed in Mali. If all known haplotypes are considered (including those not observed at the site), then all seven haplotypes could have been generated via single crossovers.

Although reshuffling of known polymorphisms via recombination could be responsible for the observed haplotypes, these new haplotypes could also be the result of convergence (i.e., selection for mutations that are identical but that do not share common ancestry). For example, the haplotype QKSSRL could have arisen from a single nucleotide change from GAA to CAA in codon 1644 of the EKSSRL haplotype. However, without additional sequence information from adjacent regions (e.g., neighboring microsatellites), it is difficult to distinguish recombination from convergent point mutations in this context.

In conclusion, by applying new high-throughput genotyping methods to a large sample of infections from Mali, seven new MSP-1 haplotypes have been identified, representing a 50% increase in the number of haplotypes previously reported in the literature. Few studies have examined MSP-1 genetic diversity in Africa, even though Africa carries the heaviest malaria burden and most MSP-1–based vaccines currently being developed and tested target this region of the protein.

Additional MSP-1 diversity may continue to be discovered as high-throughput methods are used to conduct large molecular epidemiology studies of this locus in other malaria endemic areas of Sub-Saharan Africa.

Although the impact of genetic diversity in MSP-1 on immunity and efficacy of MSP-1–based vaccines is not clear, the Y allele at position 1700 was recently reported in one isolate from India. Haplotype diversity in Africa, using developed and tested target this region of the protein.
it is possible that vaccination with one of the common MSP-1α haplotypes could give a competitive advantage to rare haplotypes such as those observed in this study, allowing them to increase in frequency in the parasite population. Understanding the mechanisms by which diverse MSP-1α haplotypes arise may improve our ability to predict how Plasmodium will evolve in response to interventions such as vaccines and drugs.

Received October 9, 2006. Accepted for publication January 8, 2007.

Acknowledgments: The authors thank the population of Bandiagara, Mali, for their continued participation in our studies, as well as the regional and district health authorities of Bandiagara, Mali for their continued support. We also thank Dr. Alan Shuldiner and the Division of Endocrinology, Diabetes, and Nutrition, University of Maryland School of Medicine, for use of the Pyrosequencer. In addition, we acknowledge Dr. Ananias Escalante for helpful comments on the manuscript.

Financial support: This study was funded by NIAID (N01AI85346) and the USAID Malaria Vaccine Program.

Authors’ addresses: Shannon L. Takala and Christopher V. Plowe, Center for Vaccine Development, University of Maryland School of Medicine, Baltimore, MD, Telephone: +1 (410) 706-3082, Fax: +1 (410) 706-6205, E-mail: cplowe@medicine.umaryland.edu. David L. Smith, Fogarty International Center, National Institutes of Health, Bethesda, MD. Mahamadou A. Thera, Drissa Coulibaly, and Ogbbara K. Doumbo, Malaria Research and Training Center, University of Bamako, Bamako, Mali.

REFERENCES


types of sequence polymorphism in the circumsporozoite gene
27. Chang SP, Case SE, Gosnell WL, Hashimoto A, Kramer KJ, Tam
LQ, Hashiro CQ, Nikiado CM, Gibson HL, Lee-Ng CT, Barr
PJ, Yokota BT, Hut GS, 1996. A recombinant baculovirus
42-kilodalton C-terminal fragment of Plasmodium falciparum
merozoite surface protein 1 protects Aotus monkeys against
28. Stoute JA, Gombe J, Withers MR, Siangla J, McKinney D, On-
yango M, Cummings JF, Milman J, Tucker K, Soisson L, Stew-
art VA, Lyon JA, Angov E, Leach A, Cohen J, Kester KE,
Ockenhouse CF, Holland CA, Diggis CL, Wittes J, Gray HD
Jr, 2007. Phase 1 randomized double-blind safety and immu-
nogenicity trial of Plasmodium falciparum malaria merozoite
surface protein FMP1 vaccine, adjuvanted with AS02A, in
Dec 7.
29. Holder AA, Lockyer MJ, Odink KG, Sandhu JS, Riveros-
Moreno V, Nicholls SC, Hillman Y, Davey LS, Tizard ML,
Schwarz RT, 1985. Primary structure of the precursor to the
three major surface antigens of Plasmodium falciparum mero-
phism in a surface antigen gene of the malaria parasite Plas-
WA, 1988. Plasmodium falciparum: gene structure and hy-
dropathy profile of the major merozoite surface antigen (gp195)
of the Uganda-Palo Alto isolate. Exp Parasitol 67: 1–11.
ation in the C-terminal part of the precursor to the major
merozoite proteins (MSP-1) of Plasmodium falciparum from
33. Kang Y, Long CA, 1995. Sequence heterogeneity of the C-
terminal, Cys-rich region of the merozoite surface protein-1
(MSP-1) in field samples of Plasmodium falciparum. Mol Bio-
chem Parasitol 73: 103–110.
Plasmodium falciparum: variations in the C-terminal cysteine-
rich region of the merozoite surface protein-1 in field samples
among Indian isolates. Exp Parasitol 92: 12–18.