GENETIC DIVERSITY IN THE MEROZOITE SURFACE PROTEIN 1 GENE OF PLASMODIUM FALCIPARUM IN DIFFERENT MALARIA-ENDEMIC LOCALITIES

DIPAK KUMAR RAJ, BIBHU RANJAN DAS, A. P. DASH, AND PRAKASH C. SUPAKAR
Institute of Life Sciences, Chandrasekharpur, Bhubaneswar, India

Abstract. A number of stage-specific antigens have been characterized for vaccine development in Plasmodium falciparum malaria. The polymorphic merozoite surface protein 1 (MSP-1) of Plasmodium falciparum is a major asexual blood stage malaria vaccine candidate antigen. In the present study, we analyzed the impact of hyperendemic malaria transmission, mesoendemic malaria transmission, and multiple infection on allelic diversity. We have used a simple strategy of polymerase chain reaction amplification and slot-blot hybridization to analyze variable regions of block-2, block-4 and blocks 6–10 of the MSP-1 gene. The allelic types of isolates collected from regions of hyperendemic malaria transmission (RHEMT) and mesoendemic malaria transmission (RMEMT) were compared. In RHEMT, 20 of 24 possible gene types were found among 163 isolates and more than one allelic type was found in 82 (50.3%) of the isolates. Thirteen of 24 possible gene types were found among 125 isolates in RMEMT and 27 (21.6%) of them contained more than one allelic type. Our results suggest for the first time that the allelic distribution or allelic diversity and chances of finding multi-strain parasites in isolates in an area vary with the rate of transmission. Analyses of isolates containing more than one strain of parasite suggest that allelic types are randomly distributed, no specific type of alleles predominately show multi-strain infection, and neither strain of the parasite affect the process of infection and development of another.

INTRODUCTION

Genetic diversity displayed by Plasmodium falciparum field isolates, the occurrence of variant forms of the parasite at different frequencies in different geographic areas, and the complexity of the infection represent major obstacles for the effective control of malaria. The propagation of multi-drug resistant parasites and insecticide-resistant mosquitoes has led to major difficulties in controlling the spread of malaria. To fight against malaria, an effective vaccine is urgently needed. A number of antigens expressed at different stages of the parasite’s life cycle have been characterized with respect to their use in vaccine development against P. falciparum.1 Merozoite surface protein-1 (MSP-1) is one of the most promising vaccine candidates.2 People naturally exposed to P. falciparum develop antibodies against MSP-1.3–5 Furthermore, an association between a naturally acquired immune response to MSP-1 and reduced malaria morbidity has been observed.6 In a number of independent studies, immunization with purified native MSP-1 or a recombinant fragment of the protein has induced at least partial protection against parasite challenge.7 Sequence comparisons showed that the entire MSP-1 gene could be divided into 17 blocks that are variable, conserved, or semi-conserved.8 In blocks 1, 3, 5, 12, and 17, the sequences are conserved. In seven regions (blocks 2, 4, 6, 8, 10, 14, and 16), the sequence shows extensive diversity, while in the remaining blocks (7, 9, 11, 13, and 15), the sequences are semi-conserved. Variation in the sequences of variable regions are dimorphic (K1/Wellcome or MAD20) in nature with the exception of the trimorphic-encoding region in block-2, which has a third version (RO33) found in natural isolates.9 Naturally acquired antibodies react more frequently against variable rather than conserved MSP-1 blocks and are specific for one of the major versions of variable blocks.10,11 In the current study, we have analyzed polymerase chain reaction (PCR)–amplified fragments containing variable blocks 2, 4, and 12–16 of the MSP-1 gene in the P. falciparum natural population and allelic types were scored by sampling allele-specific radio-labeled oligonucleotide probes. The allelic types were compared among the isolates collected from regions of hyperendemic malaria transmission (RHEMT) and mesoendemic malaria transmission (RMEMT). We have also analyzed the allelic diversity in the isolates showing more than one strain of parasites.

MATERIALS AND METHODS

Study area. The study area was in Orissa, a state of India located at 30°20′N and 85°54′E. The population of Orissa is more than 40 million. Orissa is a region endemic for malaria transmission. For our study, we divided Orissa into two regions depending upon the degree of infection with P. falciparum (Figure 1). The first was a region of hyperendemic malaria transmission located in forested and hilly areas that covered approximately 25% of the area (Keonjhar, Phulbani, Koraput, and Malkanagiri). Transmission persists throughout the year, with the peak season from August to October. Plasmodium falciparum accounts for more than 80% of the infections in this region; the rest are caused by P. vivax and other Plasmodium species. The second region was one of mesoendemic malaria transmission located in coastal districts (Jajpur, Cuttack, Khurda, and Berahmpur). Only seasonal malaria infections are observed (except for some cases of malaria detected throughout the year in migratory populations from hyperendemic areas). In this region, approximately 40% of the infections are caused by P. falciparum; the rest are caused by P. vivax. The study was reviewed and approved by the Ministry of Health of the Government of Orissa.

Isolates of P. falciparum. Sample collection was carried out in 2001 and 2003. Blood samples (n = 350) were obtained from P. falciparum-infected individuals who attended the malaria clinics at district hospitals, primary health centers, or directly from the field. Blood (1 mL) was collected by venipuncture from consenting volunteers into heparinized tubes or as 20-μL samples that were spotted and dried on Whatman (Clifton, NJ) filter paper. Precautions were taken to avoid collection of samples from the migratory population. DNA was prepared from these samples using a rapid DNA isolation method essentially as described by Foley and others.12 Briefly, 500 μL of ice-cold 5 mM sodium phosphate (pH 8)
was added to 20 \mu L of venipuncture blood or an equivalent amount of blood spotted on filter paper and vortexed. After centrifugation for 10 minutes in a microcentrifuge tube at 4°C, the supernatant was discarded. The pellet was suspended in 100 \mu L of phosphate-buffered saline containing 0.01% saponin and washed twice with same buffer by repeating the above steps of vortexing and centrifugation. After washing, the pellet was suspended in 50 \mu L of water, vortexed and then boiled for 20 minutes at 100°C. The samples were centrifuged and the supernatant was collected and used in a PCR.

**Polymerase chain reaction.** The detection of malaria parasites was done by a microscopy method and multiplex PCR as described elsewhere.\textsuperscript{13} Three pairs of oligonucleotide primers were used in separate PCRs to amplify four variable regions of the MSP-1 gene from genomic DNA of each of the natural isolates. The codon numbers given here refer to the sequence alignment previously described.\textsuperscript{14} The three regions and primer pairs were used as previously described.\textsuperscript{15} The regions and primers used were 1) codons 34-234 (block 1-3, primers BK1F: 5'-gtc gta gcc gct aca-3' and BK3R: 5'-gta act ttc cat ttt gcc gac att a-3'); 2) codons 198-359 (block 3-5, primers BK3F: 5'-ttc aat ctt aaa att cgt gca-3' and BK5R: 5'-aaa ttt aat agt ttt ggc aat ttc ttt-3'); and 3) codons 1235-1638: (blocks 12-16, primers BK12F: 5'-aaa aat tat aca ggt aat tct cca ag-3' and BK16R: 5'-tac gca ttg gtg ttg tga aat gtt-3'). The PCR amplification was performed in 20-\mu L volumes in 48-well plates with initial denaturation for five minutes at 94°C and 40 cycles for one minute at 94°C, one minute at 55°C, and 90 seconds at 72°C. A final extension was performed at 72°C for 10 minutes. Each reaction contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl\textsubscript{2}, 0.001% gelatin, 2 mM of each dNTP, and two units of Taq DNA polymerase. A 15-\mu L aliquot of the PCR product was separated by electrophoresis on a 1.2% agarose gel to check amplified products and the rest was kept for further analysis.

**Slot-blot hybridization.** Five microliters of each PCR product was heated in 0.5 mL of 0.4 M NaOH, 10 mM EDTA at 100°C for 10 minutes. The denatured DNA fragments were transferred to a nylon N + membrane (Amersham, Arlington).
POLYMORPHISM IN THE MSP-1 GENE OF *P. FALCIPARUM*

287

RESULTS

Scoring of MSP-1 alleles in the isolates. Of 350 samples collected, only 288 samples showed positive typing for the MSP-1 gene (Figure 2). One hundred sixty-three isolates were typed from RHEMT and 125 isolates were typed from RMEMT. In case of RHEMT, 82 (50.3%) isolates showed more than one gene type parasite, while in RMEMT, 27 (21.6%) of 125 isolates showed more than one allele. Most of the multi-strain isolates contain two different gene types and only nine isolates show three different gene types. Twenty of the 24 possible MSP-1 gene types were identified in RHEMT, and 13 of the 24 possible gene types were identified in RMEMT. When all the isolates containing more than one strain of parasites were analyzed separately, we found 16 of the 24 possible gene types.

Variation observed in the distribution of the MSP-1 allele. We compared the MSP-1 gene type frequencies in isolates collected from RHEMT and RMEMT (Figure 3A). No statistical significant difference was detected in the comparison of the overall distribution of this gene in RHEMT and RMEMT. A block-by-block comparison showed statistically significant differences in the frequencies of allelic types in block 6–16 (χ² = 7.49, degrees of freedom = 1, P < 0.006). The largest variation was found in the frequencies of gene types 17, 20, 23, and 24 in RHEMT and 16, 17, 23, and 24 in RMEMT. In RHEMT, type 17 and 20 differ by block 4a and 4b allelic type (K1/Wellcome and MAD20/MAD20 and K1/Wellcome, respectively). In RMEMT, type 16 and 17 differ by block 2 allelic type (either K1/Wellcome or MAD20) and type 23 and 24 differ by the same block 2 allelic type (either

The MSP-1 gene typing strategy and data analysis. The typing procedure was designed to identify the 24 major MSP-1 gene types shown in Table 1, which are defined as unique combinations of 1) one of the three variations (K1/Wellcome, MAD20, or RO33) of block-2; 2) one of the four versions of block-4; and 3) one of the two versions of (K1/Wellcome or MAD20) in the segment between block 6 and 16 where intragenic recombination does not occur. Isolates with multi-strain parasites from both RHEMT and RMEMT were combined to identify specific gene types more frequently found in multi-strain infections. The frequency of each MSP-1 gene type was analyzed as the proportion of total typed parasite populations among the isolates that included more than one sub-population per isolate in cases of genetically mixed infections. The proportions expected from the null hypothesis of random association of variable block allelic types were derived from a simple probability model analogous to those used in population genetics to estimate the expected frequency of multi-locus genotypes. For instance, multiplying the observed proportion of the parasites with K1/Wellcome type sequence in blocks 2, 4a, 4b, and 6–16 gave the expected proportion of gene type-1, which has K1/Wellcome type sequences in all variable blocks. To test the null hypothesis of random assortment of allelic types in this parasitic population, expected and observed frequencies were compared by using chi-square tests for goodness of fit. We also used chi-square tests for independent samples to analyze variations in MSP-1 gene type distributions and diversity.

Figure 2. Slot-blot hybridization showing binding of probes B4 MD (A) and B4 K1/Wellcome (B) to the polymerase chain reaction–amplified product of Block-4a for the merozoite surface protein 1 of *Plasmodium falciparum*. Wells 3, 12, 28, 32 and 33 showing isolates with more than one strain of parasites. Wells 14, 17, and 30 show the lack of hybridization.
MAD20 or RO33) as seen in isolates from RHEMT. The eight most frequent gene types (2, 15−18, 20, 23, and 24) account for 76% in RHEMT and 90% in RMEMT. Gene types 4−9 were present in RHEMT, but absent in RMEMT.

**Association of allelic type in variable blocks.** We compared the distribution of MSP-1 gene types with that of expected types from the hypothesis of random assortment allelic types in variable blocks. Significant deviation from the expected distribution was detected in both samples of RHEMT and RMEMT. Block-by-block analysis showed a number of linkage groups in variable regions of the MSP-1 gene. For example, most of the parasites with MAD20 type sequences in blocks 6−16 have concordant MAD20 type sequence in blocks 4a and 4b.

**DISCUSSION**

Of the 350 isolates, MSP-1 typing was completed in 288. The remainder showed either partial typing (hybridization with one or two variable blocks) or no hybridization. The reason for the failure of hybridization in some of the isolates may be polymorphism in repeats of variable blocks and variations such as point mutations with in the sequence homologous to one or more of the probes.

In contrast to previous studies done in Vietnam,18,19 relatively restricted MSP-1 gene types were found in an area of malaria holoendemicity (Tanzania).15 Our finding of 50.3% of the isolates with more than one-gene types in RHEMT showed higher degree of multi-strain infections in comparison to isolates from Vietnam (44%) and Brazil (38%).20 However, RMEMT showed a lower percentage of (21.6%) of multi-strain parasites in the isolates in comparison to other mesoendemic localities studied.21 The presence of more possible allelic strains (20 of 24) in RHEMT in comparison to RMEMT (13 of 24) indicates that the rate of transmission may be a key factor in the allelic distribution i.e., the higher the rate of infection, the higher the number of allelic types in an endemic area. The presence of more than one parasitic gene type in a single human host may lead to cross-fertilization and meiotic recombination during the develop-
ment stages in mosquito vectors. Moreover, analysis of isolates with multi-strain parasites showed that the allelic types are randomly distributed and no significant differences were observed in comparison to RHEMT (Figure 3B). The present study suggests that one allelic type has no effect on others in process of infection and development in the human host, so chances of new allele types emerging increase. All the allelic types present in the RMEMT and in the RHEMT indicate that the 13 types of allelic present in RMEMT are the basic types in these regions in comparison to the remaining seven allelic types, which are present only in RHEMT and may have been generated due to natural selection, followed by high rate of cross-fertilization and recombination. Finding of all 24 allele types may be possible if more isolates in RHEMT are analyzed. The result suggests that the allelic distribution and diversity depends upon the degree of transmission, and not only on geographic distributions and other biologic factors.

In conclusion, an understanding of genetic diversity of *P. falciparum* in natural population is a pre-requisite for efficacious control measures and may be helpful in designing multivariate vaccines for malaria caused by this parasite. It is also clear that parasite diversity is implicated in various phenomena of pathology, immunity, drug resistance, and transmission potential. Results from this study suggest that diverse genetic variations and multi-strain infections may be responsible for the failure of recombinant vaccines generated from variable regions of MSP-1 gene, thus making the parasite more successful in avoiding the immune system of human host.

Received September 16, 2003. Accepted for publication February 24, 2004.

Acknowledgments: Administrative and initial technical help in sample collection from Dr. R. C. Sethi and M. M. Rout are gratefully acknowledged. We sincerely thank Dr. B. Ravindran for critically reading the manuscript.

Financial support: This work was supported by the Department of Biotechnology, Government of India. Dipak Kumar Raj was supported by a Research Fellowship from the University Grant Commission, Government of India.

Authors' addresses: Dipak Kumar Raj, A. P. Dash, and Prakash C. Supakar, Institute of Life Sciences Nalco Square, Chandrasekharpur, Bhubaneswar 751 023, India, Telephone: 91- 674-230-0137 or 230-1476, Fax: 91- 674-230-0728, E-mail: psupakar@hotmail.com. Bibhuan Ranjan Das, Research and Development Division, Sisco Research Laboratories-Ranbaxy, Limited, Clinical Reference Laboratories, Mumbai 400 093, India.

REFERENCES