Research Articles

Genetic polymorphism in *msp-2*, *ama-1* and *csp* genes in *Plasmodium falciparum* field isolates from north and north-western India

Umar Farooq, N. Malla & M.L. Dubey

Department of Parasitology, Post Graduate Institute of Medical Education and Research, Chandigarh, India

Abstract

Background & objectives: Malaria is a major public health problem in tropical and sub-tropical countries. Malaria vaccine is highly desirable as an adjunct to existing malaria control measures. The polymorphism in vaccine candidate antigens might be a hurdle in developing an effective vaccine. Merozoite surface protein-2, apical membrane antigen-1 and circumsporozoite protein of *Plasmodium falciparum* are vaccine candidate antigens. The aim of this study was to detect extent of genetic polymorphism in potential vaccine candidate antigen genes, i.e. *msp-2*, *ama-1* and *csp* of *P. falciparum* isolates prevalent in northern and north-western parts of India.

Methods: Overall 88 parasite isolates of *P. falciparum* were collected during July 1998–March 2002 from different parts of northern and north-western India. DNA was extracted and analyzed for genetic polymorphism by PCR-RFLP method. For *msp-2* gene, family-specific (FC-27 and 3D7) nested PCR was also performed.

Results: PCR showed size polymorphism in all the target genes. Three alleles were observed in *msp-2* and *ama-1*, while only two in *csp*. RFLP of *ama-1* and *csp* with Dra-1 and Ssp-1 endonucleases respectively, failed to differentiate isolates in sub-allelic types, while *Hinf-I* digestion of *msp-2* amplicons differentiated three alleles into two distinct allelic families, i.e. FC-27 and 3D7. The allelic family-specific PCR generally confirmed the results of PCR-RFLP except in a few isolates, which showed mixed (two) clones of *msp-2* gene.

Interpretation & conclusion: There was extensive polymorphism in *msp-2* gene while *ama-1* and *csp* genes showed low polymorphism which may be due to the functional constraints of these proteins. The low level transmission of malaria in the study area may also be a factor for low polymorphism.

Key words *Ama-1 – csp – genetic polymorphism – msp-2 – nested PCR – PCR-RFLP – Plasmodium falciparum*

Introduction

Malaria is an important tropical disease with an estimated global burden of 300 to 660 million cases every year, of which around 90% occur in sub-Saharan Africa where mortality due to malaria is also reported to be higher than elsewhere. The increase in incidence of *Plasmodium falciparum*, the agent of severe and complicated malaria in India is a matter of immense clinical and public health importance. The situation has become more complicated due to the spread of chloroquine resistant strains of *P. falciparum* throughout the world and in India. The alternative new drugs are cost intensive and potentially toxic, therefore, an effective malaria vaccine development is necessary. It will be a useful adjunct
to existing control measures for malaria. Malaria research is now focused on immunologically relevant proteins, especially those expressed on the parasite surface and which are the first ones to come in contact with the host’s immune system. Several antigens of *P. falciparum* from different stages of the life cycle have been characterized for use as a malaria vaccine. Several malaria vaccines have undergone field trials but these have shown low efficacy during the field trials. One of the reasons for the low efficacy could be the antigenic polymorphism in the vaccine candidate antigens. Recently, Phase 1–2b clinical trial of a vaccine based on *msp*-2, in Papua New Guinea has indicated that antigen of one allelic type included in the vaccine may be more effective against the parasites having same allelic type than those having other alleles. Merozoite surface protein-2 (*MSP*-2), apical membrane antigen-1 (*AMA*-1) and circumsporozoite protein (*CSP*) of *P. falciparum* are considered prime candidates for the development of malaria vaccine. Polymorphism in these malaria vaccine candidate antigens has been reported from several parts of the world. However, limited reports are available on the genetic diversity existing among *P. falciparum* populations of India. Some of these vaccine candidate genes have been reported to be dimorphic, e.g. *msp*-1 and *msp*-2. The analysis of genetic variation among the isolates of *P. falciparum* prevalent in a region is important before the development or field trial of a malaria vaccine in that geographical region.

The present study was aimed to investigate the genetic polymorphism and to identify the presence of different allelic types in *msp*-2, *ama*-1 and *csp* genes in clinical isolates of *P. falciparum* collected from north and north-western parts of India.

**Material & Methods**

The blood samples were collected from symptomatic malaria patients found positive for *P. falciparum* by microscopy. The patients included were those attending Nehru Hospital of Postgraduate Institute of Medical Education and Research, Chandigarh, India or malaria clinics and hospitals in northern and north-western India. Since Nehru Hospital is a referral, tertiary care centre, patients from various states in north and north-western India are referred to this institute. Though the majority of patients (72) were from north and north-western India (27 from Punjab and Haryana; 20 from Uttar Pradesh; 13 from Rajasthan; and 12 from Delhi), a few patients from other regions of India (Assam, Chhattisgarh, Madhya Pradesh, West Bengal and Tripura) were also included. Overall, 88 parasite isolates were collected during July 1998–March 2002. The patients with mixed infection (with *P. falciparum* and *P. vivax*) were excluded. Two to three millilitres of blood was obtained from each patient by venepuncture in citrated anticoagulant.

The DNA was isolated by a rapid method as described by Foley et al with slight modification. Briefly, 50 μl of parasitized blood was washed thrice with 1 ml of ice-cold sodium phosphate (5 mM, pH 8.0). Finally, the suspension was centrifuged at 10,000 rpm for 10 min and 50 μl of sterile distilled water was added to the pellet. After mixing thoroughly, it was boiled for 10 min in a water bath. The suspension was centrifuged at 10,000 rpm and 10 μl of supernatant was used as DNA template in 50 μl polymerase chain reaction (PCR) mixture.

The PCR was performed for variable regions of *msp*-2, *ama*-1 and *csp* genes. In addition, for *msp*-2 gene, a nested PCR was also performed using allelic family-specific primers for FC-27 and 3D7 families. All PCR reactions were carried out in a 50 μl reaction volume in a thermocycler (Eppendorf master cycler gradient, Germany). Amplification was performed in 50 mM KCl, 10 mM Tris-Cl (pH 9.0), 1.5 mM MgCl₂, 0.01% gelatin, 0.2 mM each of dNTPs, 1.25 unit AmpliTaq DNA polymerase (GIBCO-BRL, USA), 250 nM each of primers and 10 μl of parasite DNA as template. The central repeat and non-repeat regions of *msp*-2, hyper-variable region (HVR) of *ama*-1 and the central polymorphic region of *csp* gene were amplified by PCR using specific primers (Table 1).
The PCR products were electrophoresed on 1.5% Agarose gel and the fragment size of PCR products was determined by comparison with appropriate molecular weight marker. The amplified products of \( msp-2 \), \( csp \) and \( ama-1 \) genes were digested with \( Hinf-1 \), \( Ssp-1 \) and \( Dra-1 \) restriction enzymes respectively. Digestion was performed in a 20 μl reaction mixture utilizing 16 μl of PCR product, 1.5 μl sterile double distilled water, 2 μl digestion buffer (1 X) and 0.5 to 1 μl (5 units) of restriction enzyme. The digestion was carried out at 37°C overnight. Finally, 10 μl of each digested PCR product was electrophoresed on 2% agarose gel containing ethedium bromide and digitized in gel documentation system (UVI Pro, UK). No PCR product was obtained in control samples which included the DNA extracted from \( P. vivax \) and blood samples from healthy individuals. The study was approved by the Institute’s Ethical Committee. The blood samples from patients were collected after obtaining informed written consent from the subjects.

### Results

The central repeat region of \( msp-2 \), hyper variable region of \( ama-1 \) and T-cell epitopic region of \( csp \) genes of \( P. falciparum \) were analyzed by PCR-RFLP. Out of 88 clinical isolates, 50 (56.8%), 67 (76.13%) and 84 (95.45%) were amplified for \( msp-2 \), \( ama-1 \) and \( csp \) respectively. Size polymorphism was seen in all the target gene regions. The molecular weight of PCR products of \( ama-1 \) ranged from 800 to 1000 bp, \( csp \) from 280 to 300 bp (Table 2) and \( msp-2 \) gene from 760 bp to 870 bp (Table 3). On the basis of molecular weight of amplicons, the isolates were categorized into three allelic types each for \( ama-1 \) and \( msp-2 \) and in two allelic types for \( csp \) (Tables 2 and 3). The RFLP analysis of \( ama-1 \) and \( csp \) did not show sub-allelic types (Table 2) as the enzymatic digestion could not reveal any difference among the isolates. The dimorphic region of \( msp-2 \) gene showed family specific \( Hinf-1 \) restriction sites. The \( Hinf-1 \) digestion of repetitive and non-repetitive dimorphic regions of \( msp-2 \) gene produced characteristic fragments (FC-27 family–96 and 115 bp; for 3D7/IC-1 family–70 and 108 bp), which are specific for allelic families, i.e. FC-27 and 3D7\textsuperscript{27}. Two distinct families, i.e. FC-27 and 3D7 were observed by RFLP and also by family-specific nested PCR.

### Table 1. The primers, annealing temperatures and number of cycles used for PCR amplifications of polymorphic segments of genes

<table>
<thead>
<tr>
<th>Name of genes</th>
<th>Sequence of primers</th>
<th>Annealing temperature (°C)</th>
<th>No. of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ama-1 )</td>
<td>5’ CCT TTG AGT TTA CAT ATA TG -3’</td>
<td>47</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>5’ TAT CTT CAC AAT TTC CAT CG -3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( csp )</td>
<td>5’ GAC CCA AAC CCG AAA TGT AGA TG -3’</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>5’ CGA CAT TAA ACA CAC TGG A -3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( msp-2 )</td>
<td>5’ GAA GGT AAT TAA AAC ATT GT -3’</td>
<td>48</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>5’ GAG GGA TGT TGC TGC TGC AC -3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FC-27</td>
<td>5’ GCAATGAAGTATATCTAATACTAATAG -3’</td>
<td>58</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>5’ GCTTTGGGTCTCCTTCTCAGTTGATTC-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D7</td>
<td>5’GCAGAAATGAAGCCTTCTACTGTTGCT-3’</td>
<td>63</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>5’GATTTGGTTCGGCATTATTAGA-3’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Though, in the present study, few isolates showed both the fragments characteristic of the two families, most isolates in the study showed only one character-istic fragment, i.e. 96 bp for FC-27 and 70 bp for...
3D7/IC-1 family on RFLP, while the other fragment characteristic of these families 115 bp and 108 bp respectively for FC-27 and 3D7/IC-1 were not obtained exactly as described by Felger et al. By the criteria of Felger et al. majority of our isolates (35/50) fell in FC-27 allelic family while 15/50 isolates belonged to 3D7 type (Table 3). Most of the isolates msp-2 type-1 allele (88.8%) and type-2 allele (87.5%) (on the basis of size of amplicon as shown in Table 3) belonged to allelic family FC-27, while the majority of isolates in allelic type-3 were found to be of 3D7 type (Table 3). The analysis of msp-2 gene performed by nested PCR by using allelic family-specific primers generally confirmed the results of PCR-RFLP with few differences. Overall, 32/50 isolates belonged to FC-27 type and 13 isolates fell into 3D7 type while 5 isolates showed mixed clones of both the allelic families indicating the prevalence of more than one type of clones among the field isolates (Table 3).

| Allelic types | Mol. weight of PCR products | Mol. weight of fragments on RFLP with Dra-1 | Frequency (%) | PCR types (%)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>950</td>
<td>500, 450</td>
<td>33 (49.25)</td>
<td>1 (20)</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>550, 450</td>
<td>31 (46.26)</td>
<td>2 (40)</td>
</tr>
<tr>
<td>3</td>
<td>800</td>
<td>350, 450</td>
<td>3 (4.47)</td>
<td>5 (12)</td>
</tr>
</tbody>
</table>

**Discussion**

Strategies to prevent the rapid spread of the parasite resistance to novel drugs or efficacy trials of potential vaccines require an understanding of population structure of the parasite. The availability of polymorphic genetic markers combined with the relative ease of their characterization in the field isolates by methods like PCR, have made such investigations possible. We have observed three allelic types of ama-1 in 67 isolates of *P. falciparum* and two genotypes of csp gene in 84 isolates on the basis of size polymorphism, and RFLP analysis did not reveal any sub-allelic types. Marshall et al. identified four allelic types of ama-1 gene based on the presence or absence of particular restriction sites for Dra-1, Ssp-1, Mae-III and Sau3A restriction enzymes. Similarly, Eisen et al. also identified four distinct allelic types in Papua New Guinea by using same restriction enzymes. However, in the present study, ama-1 PCR

| Allelic type of *msp-2* gene of *P. falciparum* according to molecular weight of amplicons and typing of isolates in allelic families assessed from PCR-RFLP and family-specific nested PCR |
|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| Size of PCR product (bp) | Allelic types (Mol. weight) | Frequency (%) | RFLP (n = 50) | Nested PCR (n = 50) |
|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| 870                                           | 1                                            | 18/50 (36) | FC-27 family (%) | 14/18 | 2 (10) |
| 800                                           | 2                                            | 16/50 (32) | 14/16 (88.88) | 13/16 | 2/16 |
| 760                                           | 3                                            | 16/50 (32) | 5/16 (31.25) | 5/16  | 9/16 |
products were digested by Dra-1 in all the isolates. According to criteria proposed by Marshall et al, all our isolates of *P. falciparum* belong to group-2. Marshall et al have observed four allelic types because they used more number of restriction enzymes which increased the possibilities to detect higher degree of polymorphism while in the present study, we used only two restriction enzymes. Marshall et al and Eisen et al also found the evidence of intragenic recombination in ama-1 gene. Esclante et al suggested that the polymorphism observed in the ama-1 gene may be due to positive natural selection and not due to intragenic recombination. The sequencing analysis of pfama-1 gene from Amazon basin of Peru has reported three distinct allelic types.

We have observed low degree of polymorphism in T-cell region of csp gene in *P. falciparum* collected from north and north-western parts of India. Similarly, Sidhu & Madhubala have also shown little variation in this region of csp amongst the Indian isolates. The findings of the present study are also in agreement with Doolan et al who have also reported limited polymorphism in csp gene of *P. falciparum* prevalent in certain geographical regions of Papua New Guinea. Contrarily, Chenet et al was reported high degree of variation in csp gene from Peru. In the present study, the presence of low polymorphism in ama-1 and csp may also be due to functional constraints of these genes. Bhattacharya et al, noted low polymorphism in csp gene T-cell epitopic region of several isolates of *P. falciparum* collected from single area (Shankargarh, India). Limited polymorphism in csp T-cell regions has been found in isolates collected from a particular geographical region. However, the comparison of *P. falciparum* isolates from different geographical regions of the world revealed extensive polymorphism in T-cell region. Doolan et al have classified the csp gene into three groups, i.e. African, Asian and one other group.

In the present study, three allelic types of msp-2 gene were observed on PCR amongst the 50 isolates from north and north-western parts of India suggesting high degree of variation. Similarly, Joshi et al from north-eastern and eastern part of India and Ranjit & Sharma from south-east, central and western part of India have reported high degree of polymorphism in *P. falciparum* isolates. Similarly, Dolmazone et al also reported high degree of polymorphism in msp-2 gene from central African Republic. The larger parasite diversity in these regions may be due to higher malaria transmission intensity, as extent of polymorphism may vary with transmission intensity in the regions. The intensity of transmission affects the frequency of inter-clonal parasite mating and intragenic recombinations, which might be playing a role in generation of genetic diversity. High degree of size polymorphism in the target regions of genes studied indicates that the *P. falciparum* population in the study area is genetically a mixture of different clones. Contrarily, low degree of variation has been reported from French Guyana and Colombia in msp-2 gene.

In the present study, RFLP analysis as well as allele specific PCR of msp-2 gene have shown the predominance of FC-27 allele. This is in accordance with earlier studies, which also reported higher prevalence of FC-27 type of allele than 3D7 from different parts of the world, including, Senegal, The Gambia, Colombia and Honduras. In earlier studies, in small number of Indian isolates of *P. falciparum*, Bhattacharya et al identified only one allelic family (FC-27). Bhattacharya et al have stated that the apparent presence of just one allelic family in India might be due to clonal propagation of *P. falciparum*. The presence of size variation may result from the addition or deletion of the nucleotides in the msp-2 gene among the isolates, which may occur during multiplication of parasite or during intragenic recombination at the time of crossing over during meiosis within mosquito vectors. In the present study, 5 (10%) of the isolates showed mixed clones, suggesting the presence of more than one allelic type in the isolates from patients.

PCR–RFLP and nested PCR techniques are rapid,
convenient and allow handling of large number of samples at one time. These techniques are more useful tools for the study of genetic polymorphism in these genes in which the events of intragenic recombination is frequent such as *msp-1*\(^{16}\) and *msp-2*\(^{27}\) genes. While RFLP is not a suitable tool to study the genes in which only point mutations are reported i.e., *csp* and *ama-1* genes in which the polymorphism might appear due to the natural selection and not by intragenic recombination\(^{32}\). This is because RFLP has a limitation as it can detect limited point mutations unless they are in restriction sites of enzyme used. On the other hand the sequencing is able to detect all types of mutations and is a strong tool to study the genetic polymorphism but it is time consuming, costly and may not be available in all the laboratories, and thus, it may not be possible to apply this technique for large epidemiological studies.

**Acknowledgement**

We gratefully acknowledge the help of National Institute of Malaria Research, Delhi for providing few isolates of *P. falciparum*. The research was supported by Indian Council of Medical Research (ICMR) grant to Dr Umar Farooq.

**References**


Corresponding author: Dr M.L. Dubey, Professor, Department of Parasitology, Post Graduate Institute of Medical Education and Research, Chandigarh–160 012, India.
E-mail: mldubey@gmail.com

Received: 8 September 2008  
Accepted in revised form: 28 January 2009