POPULATION STRUCTURE ANALYSIS OF PLASMODIUM VIVAX IN AREAS OF IRAN WITH DIFFERENT MALARIA ENDEMICITY

SEDIGHEH ZAKERI,* AKRAM ABOUIE MEHRIZI, SHADI MAMAGHANI, SOMAYYEH NOORIZADEH, GEORGES SNOUNOU, AND NAVID DINPARAST DJADID

Malaria Research Group, Biotechnology Department, Pasteur Institute of Iran, Tehran, Iran; Parasitologie Comparée et Modèles Expérimentaux, Unité Scientifique Museum 307, Centre de la Recherche Scientifique, Institut Federatif de Recherche 101, Muséum National d’Histoire Naturelle, Paris, France

Abstract. To obtain the genetic structure of Plasmodium vivax populations in the northern and southern malaria-endemic areas in Iran, which differ in endemicity, sequence diversity in the variable block 5 and the C-terminal part of P. vivax merozoite surface protein 1 (Pvmsp 1) was analyzed. The variable block 5 fragment from 52 northern and 94 southern isolates was amplified and sequenced. Type 1, type 2, and recombinant type 3 allelic variants were found in both northern and southern isolates, with type 1 predominant in parasites from the north and type 2 in those from the south. A total of 7 and 27 distinct variants were detected among northern and southern isolates, respectively. A single variant predominated (71%) in the northern isolates, whereas variants were evenly distributed among southern isolates, with only two exceeding 10%. Thus, parasites from the southern malaria-endemic area were more polymorphic than those circulating in the northern area, where malaria is a re-emerging disease. Sequence alignments showed that although some variants were found only in northern or southern isolates, some were common to both and had also been observed in parasites from Azerbaijan, Turkey, Thailand, Bangladesh, and China. The Pvmsp 1 fragment corresponding to the C-terminal region was also amplified and the sequences derived from 20 northern and 50 southern isolates were identical. This high degree of conservation reinforces the potential of this polypeptide fragment for inclusion in synthetic vaccines being developed against P. vivax.

INTRODUCTION

Plasmodium vivax, one of the malaria parasites of humans, is globally distributed and predominates in the Middle East, central and southeast Asia, Latin and South America, and Oceania, where it is a major contributor to morbidity. The geographic distribution of P. vivax and the emergence of chloroquine-resistant P. vivax parasites emphasize the importance of developing control strategies that would effectively prevent or substantially delay the spread of these parasites. The study of the population structure of a parasite by analysis of genetic diversity provides baseline data essential for monitoring drug resistance and for trials and eventual deployment of a malaria vaccine.

In Iran, national malaria control programs initiated in the 1950s resulted in the eradication of the disease from the Caspian Sea region in the north and a substantial reduction in transmission in the coastal plains along the Persian Gulf in the south by 1977. Malaria was then restricted to the southeastern provinces of Hormozga, Sistan and Baluchistan, and the tropical region of Kerman. However, malaria reappeared in northern Iran in 1994 after a large displacement of people from the Republic of Azerbaijan and to some extent from Armenia. Furthermore, substantial human migration mainly from Pakistan and Afghanistan has become characteristic in the provinces bordering these countries. Malaria can thus be considered a re-emerging disease in Iran. In 2004, approximately 23,000 malaria cases were reported in Iran, 80% of which were microscopically diagnosed as P. vivax, and the remainder as P. falciparum. Most (85%) of the cases were recorded in the southeastern provinces and accounted for all the microscopically detected P. falciparum infections, although mixed species infections were detected by sensitive molecular methods in both northern and southern provinces. Recent unpublished reports from the Ministry of Health suggested an increasing trend of malaria introduction into hitherto non-endemic areas.

To date, the single copy genes coding for the polymorphic merozoite surface protein 1 (MSP 1) and circumsporozoite protein (CSP) of P. vivax, two leading vaccine candidate antigens, remain the most studied molecular markers of P. vivax. The P. vivax msp 1 (Pvmsp 1) gene encodes a protein of approximately 1,720 amino acids, and sequence comparison showed that the gene is dimorphic with two known allelic families, Belem (type 1) and Sal 1 (type 2). Interallelic recombination between these two types was also observed and these are denoted type 3. Given the large size of the Pvmsp 1 gene, population genetic studies mainly focused on the diversity of the one of the variable blocks, block 5, and this polymorphic region has been considered suitable for molecular typing of P. vivax. The C-terminus part of the homologous antigen in P. falciparum is considered the most suitable region for inclusion in a vaccine formulation. In this parasite, this domain was shown to be necessary for the invasion of red blood cells, and recent studies support a similar role for the C-terminal region of Pvmsp 1. The relative sequence conservation of the P. falciparum MSP 1 C-terminus, are consistent with its functional role and reinforce its inclusion in experimental vaccines. Comparatively little is known about the diversity of the Pvmsp 1 C-terminal domain.

This study was undertaken to determine the extent of genetic diversity and the population structure of P. vivax parasites found in Iran. A further aim was to compare the parasite populations circulating in the main malaria-endemic areas in the tropical southern provinces with those found in the northern province where P. vivax has been recently re-introduced. The molecular analyses were based on the block 5 and the C-terminal region of Pvmsp 1.

MATERIALS AND METHODS

Study areas and samples collection. Plasmodium vivax isolates were collected from two different malaria-endemic ar-
Figure 1. Map of Iran showing the sites where the *Plasmodium vivax* isolates were collected. Northern endemic area: PA = Pars Abad, Southeastern endemic area: K = Kahnouj in Kerman Province; M = Minab in Hormozgan Province; Ch = Chabahar in Sistan and Baluchistan.

The city of Pars Abad in Ardebil province is in northern Iran, northwest of the Caspian Sea (Figure 1), was selected as a region where malaria was recently introduced. Malaria transmission occurs only from June to October when the average temperature is sufficiently high to allow parasite development in the anopheline vectors. The most prevalent species in this area is *P. vivax*, and the main vectors are *Anopheles maculipennis* and *An. sachaorovi*, and the secondary vectors are *An. superpictus* and *An. hyrcanus*. A total of 174 isolates were collected from symptomatic *P. vivax*-infected patients who sought treatment at the malaria clinic in Pars Abad during the transmission seasons of 2000 to 2003. The demographic information of the patients is shown in Table 1.

The tropical southeastern malaria-endemic area encompasses three provinces, Sistan and Baluchistan, Hormozgan, and the tropical part of Kerman (Figure 1). Transmission is year round with two peaks, the first from May to August with *P. vivax* as the predominant species and the second peak from October to November where both *P. falciparum* and *P. vivax* infections are generally recorded in equal numbers. The main mosquito vectors are *An. stephensi*, *An. culicifacies*, *An. fluviatilis*, and *An. pulcherrimus*. A total of 155 isolates were collected from *P. vivax*-infected patients from the Chabahar District in Sistan and Baluchistan, 17 isolates from Minab in Hormozgan, and 28 from Kahnouj in Kerman Province from 2000 to 2003.

All 374 *P. vivax* clinical isolates were diagnosed by light microscopic examination of Giemsa-stained blood smears. A 1-mL blood sample was then collected on admission after informed consent was obtained from adults or from the parents or legal guardians of children. The study was reviewed and approved by the Ethical Review Committee of the Pasteur Institute of Iran.

**Molecular analysis of the *Pvmsp 1* gene.** Genomic DNA was purified from the infected blood samples by standard phenol/phenol-chloroform extraction and ethanol precipitation as previously described. The DNA was dissolved in TE buffer (10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA) and kept at -20°C until use.

Polymerase chain reaction (PCR) amplification of the fragment containing the ICB5-ICB6 region of *Pvmsp 1* (the block 6 region using nomenclature of Putaporntip and others), which corresponds to basepairs 2008–2426 (Belem strain, M60807) and excludes the primer sequences, was conducted using the oligonucleotide primers 5'-ACTACCTGATGAGTCCTC-3' (VS1F) and 5'-TTGTGACATGCCTAAAGCG-3' (VS1R). The reaction was carried out for 35 cycles at 94°C for 1 minute, 58°C for 1 minute, and 72°C for 2 minutes, and a final primer extension at 72°C for 10 minutes. The *Pvmsp 1* fragment corresponding to the C-terminus region (basepairs 4859–5162 of the Belem isolate M60807) was amplified using the oligonucleotide primers 5'-TTATTAACATGAGTCGAGGC-3' (VMCF) and 5'-TTAAAGCTCCATGCACAGGG-3' (VMCR). The reaction was carried out 95°C for 5 minutes, 30 cycles at 94°C for 1 minute, annealing at 60°C for 1 minute, and extension at 72°C for 1 minute, and a final primer extension at 72°C for 10 minutes.

For the variable block 5, the PCR products from 52 northern isolates and 94 southern isolates (69 from Chabahar, 10 from Hormozgan and 15 from Kerman) were directly sequenced in both directions. For the C-terminal region, PCR fragments from 20 isolates from the north and 60 from the southeast (53 from Chabahar isolates, 4 from Hormozgan and 3 from Kerman) were sequenced in both directions. An ABI 3100 DNA sequencer (Primm Company, Milan, Italy) was used.

Nucleotide and amino acid sequences were aligned with the corresponding Belem and/or Salvador sequences by using the CLUSTAL W. Major alleles were classified based on protein sequences alignment and the tree was constructed with the MEGA2.1 program (built with the neighbor-joining method, Kimura 2-parameter, pairwise deletion) based on nucleotide sequences of *Pvmsp 1* fragments amplified from Iranian isolates.

Nucleotide sequences of the *Pvmsp 1* variable block 5 reported in this article are available in the European Molecular Biology Laboratory, GenBank and DNA Data Bank of Japan databases under accession numbers AY162146–AY162159, AY192580, AY192598, AY192581–AY192587, AY337019–AY337022, AY632331–AY632354 (Pars Abad) for the northern isolates and AY642598–AY642666 (Chabahar), AY192588–AY192597 (Hormozgan), and AY162160–AY162174 (Kahnouj) for the southeastern isolates. The Gen-table 1

<table>
<thead>
<tr>
<th>Malaria cases</th>
<th>Male</th>
<th>Female</th>
<th>Mean age, years</th>
<th>Nationality</th>
<th>Travel history</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>174</td>
<td>51%</td>
<td>25</td>
<td>99.50%</td>
<td>0.50%</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>26%</td>
<td>24</td>
<td>62%</td>
<td>32% 6%</td>
</tr>
<tr>
<td>North</td>
<td>49%</td>
<td>51%</td>
<td>25</td>
<td>99.50%</td>
<td>0.50%</td>
</tr>
<tr>
<td>South</td>
<td>74%</td>
<td>26%</td>
<td>24</td>
<td>62%</td>
<td>32% 6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North</td>
<td>49%</td>
<td>51%</td>
<td>25</td>
<td>99.50%</td>
<td>0.50%</td>
</tr>
<tr>
<td>South</td>
<td>74%</td>
<td>26%</td>
<td>24</td>
<td>62%</td>
<td>32% 6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nationality</th>
<th>Travel history</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>15%</td>
</tr>
<tr>
<td>Pakse</td>
<td>14.50%</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bank accession numbers are AY925135–AY925116 for *Pvmsp 1* C-terminus fragments amplified from the 20 northern isolates and AY925095–AY925108 (Chabahar), AY925109–AY925112 (Hormozgan), and AY925113–AY925115 (Kahnouj) *Pvmsp 1* C-terminus fragments amplified from the 60 southern isolates.

**RESULTS**

The *Pvmsp 1* variable block 5 was amplified from genomic DNA purified from 374 isolates collected from patients infected with *P. vivax*. The amplified fragments varied in size from 440 to 520 basepairs for northern isolates and from 430 to 550 basepairs for southern isolates. However, the 450-basepair and 520-basepair fragments were predominant among the northern and southern parasites, respectively. The PCR fragments were divided into groups that differ by 50 basepairs, and a subset of fragments was randomly selected from each group for further sequence analysis. The minor size variations observed within each group were taken into account. Fragments that were representative of all sizes and reflected their frequency were selected for sequencing. Thus, *Pvmsp-1* variable block 5 fragments derived from 52 *P. vivax* isolates from north Iran and 94 from southern Iran were analyzed. Thirty distinct variants were found among the 145 sequenced fragments (Figure 2). Amino acid sequence comparisons showed that the fragments could be classified into three basic sequence types as previously described. The number of repeated Gln residues (Q) of type 1 sequences (homologous to the Belem type) ranged between 16 and 21 (18–21 for the northern isolates and 16–21 for the southern isolates). Single nucleotide polymorphisms (SNPs) occurred at 11 positions. Synonymous mutations were observed in five positions, mostly within the Q repeat region, and mutations at the remaining six positions were non-synonymous. The SNPs that occurred in the region 5′ to the poly Q repeat coding segment were only in observed in southern samples. The type 2 sequences (homologous to the Sal 1 type) also showed amino acid variations (Figure 2). Synonymous SNPs were detected at four sites and non-synonymous SNPs were detected at nine sites. All sequences from the northern isolates contained a three-nucleotide insertion encoding Gln residue (CAA), which was not detected in the Sal 1 sequence type (M75674). This insertion occurred in 90.5% of southern

**Figure 2**. Alignment of the *Plasmodium vivax* merozoite surface protein 1 variable block 5 amino acid sequences for the 27 distinct allelic forms observed. **A**, Belem type 1 sequences; **B**, recombinant type 3 sequences; **C**, Sal 1 type 2 sequences. Sequences were compared with the published Belem (M60807) and Sal 1 (M756745) sequences. Dots and dashes represent identical residues and deletions, respectively. Amino acid changes resulting from nucleotide substitutions are shown in **bold**. **B** = Belem type 1; **S** = Salvador 1 type 2; **R** = recombinant type 3.
samples. Type 3 sequences, which represent a combination of a Sal 1-like sequence at the 5′-end and a Belem-like sequence at the 3′-end, including the poly Q segment, were found in 2 (4%) of the 52 sequences obtained from northern isolates and in 25 (26.5%) of the 94 sequences obtained from southern isolates, respectively. The number of poly Q segments was 19 for the northern type 3 isolates and ranged from 18 to 27 for southern isolates (Figure 2). A single residue was associated with a non-synonymous mutation, whereas five positions had synonymous mutations.

Representatives of the three types were detected in the parasites from both study areas, although their prevalences differed. For the northern isolates, 75% (39 of 52) were type 1, 21% (11 of 52) were type 2, and 4% (2 of 52) were type 3 sequences. For the southern isolates, 17.2% (16 of 94) were type 1, 56.3% (53 of 94) were type 2, and 26.5% (25 of 94) were type 3 sequences. Of the 30 distinct variants, 7 were derived from northern isolates and 27 from southern isolates. The frequency distribution for these variants differed for the two malaria-endemic areas (Figure 3). Three variants, type 1 B1 (71%) and type 2 S1 (11.5%), and S2 (7.6%), accounted for nearly all the variants detected in the northern parasites. In contrast, only 2 of the 27 distinct allelic forms detected in southern parasites exceeded a frequency of 10% (18% type 2 S1 and 18% type 3 R5). The remaining 25 variants were randomly distributed among the southern isolates. Only three variants, B1, S1, and S2, were detected in both northern and southern parasites (Figure 3).

The PCR amplification of the C-terminal region of Pvmsp 1 gene showed no size variations, with a 345-basepair fragment amplified from 52 isolates collected in north-ern Iran and 94 isolates collected in southern Iran. B = Belem type 1; S = Salvador I type 2; R = recombinant type 3.

**Figure 3.** Frequency distribution of representative allelic variants of the Plasmodium vivax merozoite surface protein 1 variable block 5. Data were derived from direct sequencing of polymerase chain reaction products amplified from 52 isolates collected in northern Iran and 94 isolates collected in southern Iran. B = Belem type 1; S = Salvador I type 2; R = recombinant type 3.

**DISCUSSION**

Malaria was eradicated from many subtropical countries and significantly reduced in vast zones of tropical countries through the application of vector control measures and the deployment of mass treatment of febrile individuals in disease-endemic areas. However, climate changes, socioeconomic and political instability, and population displacements have contributed to the re-emergence of the infection in some malaria-free areas. The re-appearance of malaria in the northern provinces of Iran 20 years after its eradication from these regions was due to the large population movements from Azerbaijan and Armenia. The re-introduction of malaria to the Republic of Azerbaijan in 1994 coincides with its emergence in the province of Ardebil in Iran in the same year. In the Iranian southeastern provinces, war and political instability in the neighboring countries have driven refugees across the border. The potential to introduce malaria in the tropical southern region of Iran is exacerbated by the lengthier transmission season. It is extremely important that malaria in the largely non-immune Iranian populations is brought under control, and that the spread of any drug-resistant parasites is contained or prevented. Moreover, it is particularly important to monitor other regions of Iran that are at risk for introduced malaria because competent anopheline vectors and permissive climatic conditions occur over many areas of the country. The choice of adequate drug treatments and eventually effective vaccines would benefit from a knowledge of the genetic diversity of the parasites transmitted in the different malaria-endemic areas. This study was conducted to characterized Iranian P. vivax populations in the two malaria areas of contrasting endemicity with respect to PvMSP 1, a leading vaccine candidate.

The data indicated that the extent of genetic diversity in the P. vivax populations from southern Iran is higher than that of those from northern Iran. This is consistent with the concept that genetic diversity decreases as levels of transmission decrease. The fact that the parasites in the northern regions are relatively isolated compared with those in the southern area probably helps to maintain the lower degree of diversity. All but one of the patients from northern Iran were Iranian nationals who had not traveled outside the province within the two months prior to the survey. In contrast, patients recruited in southern Iran were Iranian, Afghan, and Pakistani nationals and many had traveled to and from these three countries,
increasing the likelihood that any *P. vivax* infections they might have harbored could have been disseminated.

Thirty distinct variants of *Pvmsp 1* variable block 5 were found in 146 sequences derived from fragments amplified from samples collected from both study areas. The fact that only seven of these variants were observed in the northern *P. vivax* isolates, compared with 27 in those from the south, and that their frequency distribution was highly biased in contrast to the random distribution in southern isolates, emphasize the lower genetic diversity of northern parasites. This further suggests a founder effect linked to an introduction of malaria from Azerbaijan and Armenia to northern part of Iran. As in Azerbaijan, the type 1 allelic variants were predominant. Type 2 sequences had the highest prevalence in southern isolates; an observation consistent with the sole presence of type 2 sequences in samples collected in India, although the sample size in that study was low. Evidence of increased transmission and thus recombination for the southern *P. vivax* populations was derived from the high frequency (27%) of recombinant type 3 sequences, which presumably arise from intragenic recombination in the mosquito vector in southern isolates in contrast to the low type 3 frequency (4%) in the *P. vivax* populations introduced into northern Iran.

Phylogenetic analysis of the *Pvmsp 1* variable block 5 sequences showed that some were identical with those observed in other locations, while others were specific to Iran. The presence of some of the *P. vivax* strains in the two geographically distinct regions in Iran (CHBel1, NBel1; CHSal1, NSal1; CHSal2, NSal2; 100% similarity), raises the possibility that these particular populations might have the same origin. This might have resulted from regular movement of immigrants from Afghanistan not only to southeastern Iran but also to Azerbaijan and Armenia and subsequently to northern Iran, although no record of Afghani refugees in Pars Abad were noted in this study. The dissemination of these parasites could have also occurred through population movement within Iran, but again no evidence of internal migration between the north and south was noted, although focal epidemics could be

---

**Figure 4.** Phylogenetic tree constructed with the MEGA2.1 program based on the nucleotide sequences of the polymorphic region variable block 5 of *Plasmodium vivax* merozoite surface protein 1 (*Pvmsp 1*) from areas with re-emerged and endemic malaria in northern and southern Iran and from previously published sequences. The northern samples with the different allelic forms are indicated as NBel (Belem type), NSal (Salvador I type), and NRec (recombinant type). Southern isolates are indicated as CHBel, CHSal, and CHRec (Chabahar isolates Belem, Sal I, and recombinant types), HBel, HSal, and HRec (Hormozgan isolates Belem, Sal I, and recombinant types). The geographic origin of the 31 *Pvmsp 1* published sequences is as follows: Azerbaijan (Azer1:AY789680, Azer2:AY789683, Azer3:AY789691, Azer4:AY789658, Azer5:AY789679, and Azer6:AY789669), Turkey (Turkey1:AJ494827, Turkey2:AJ494828, Turkey3:AJ494830, Turkey4:AJ494831, Turkey5:AJ494832, Turkey6:AJ494829, Turkey7:AJ494826), India (India1:AJ15852, India2:AY229867, India3:AJ494986, India4:AJ494997), Bangladesh (Bang1:AF435619, Bang2:AF435616), Thailand (Thail1:AF435611, Thail2:AF435604, Thail3:AF435615), South and North Korea (Skorea1:AF435635, Skorea2:AF435049, and NKorea:AF216677), China (China1:AY465380, China2:AY465407, and China3:AY465379), Brazil (AF435630), Belem (M60807), and Sal-1 (M75674). The length of the bottom line is proportional to the genetic differences (%). The bootstrap values are shown on the branches and indicate the number of times in 1,000 replications. Only bootstrap values about 50% are shown.
initiated from a small number of Plasmodium-carrying migrants. Nonetheless, speculation about the origin of P. vivax populations will have to await large-scale molecular surveys of this parasite in neighboring countries and other Asian malaria-endemic regions. Moreover, meaningful conclusions require that the analysis should be extended to the other variable blocks of Pvmsp 1, as well as to other polymorphic genes such as those coding for the circumsporozoite protein, and the merozoite surface proteins 3 alpha and beta, and microsatellite markers. The lack of variation in the sequence of the C-terminus of Pvm sp 1 is highly encouraging with respect to the potential of experimental vaccines based on this polypeptide to be effective against P. vivax infection in the eastern Mediterranean region, whether these occur in the temperate or tropical zones.

Received May 12, 2005. Accepted for publication October 31, 2005.

Acknowledgments: We thank the Malaria Division, Center for Diseases Management and Control, Iran, especially Dr. A. Raeisi, for their cooperation; Dr. M. T. Tabatabai (Zahedan University of Medical Sciences), the staff in the Public Health Department of Sistan and Baluchistan, Dr. Medhi Zand and Mr. Khodadad Gorgji (Chabahar District), Dr. Naseri Nejad (Hormozgan Province), Mr. Fandandari Mahmoudi (Kerman Province), and Dr. Shahnam Arshi and Mr. Kourosh Emdadi ( Pars-Abad and Ardebil) for assistance in collecting blood samples; and S. Gholizadeh for editing the figures.

Financial support: This study was supported in part by grants from the United National Development Program/World Bank/World Health Organization Special Program for Research and Training in Tropical Diseases (No. A30133) and the Pasteur Institute of Iran (No. 210).

Authors’ addresses: Sedigheh Zakeri, Akram Abouie Mehrizi, Shadi Mamaghani, Somayeh Noorizadeh, and David Dinparast Djaidj, Malaria Research Group, Biotechnology Department, Pasteur Institute of Iran, Pasteur Avenue, P.O. Box 13164, Tehran, Iran, Telephone: 98-21-648-0780, Fax: 98-21-646-5132, E-mails: zakeris@yahoo.com and azad@institute.pasteur.ac.ir. Georges Snounou, Parasitologie Comparée et Modèles Expi rimentaux, Unite Scientifique Museum 307, Centre de la Recherche Scientifique, Institut Federatif de Recherche 101, Museum National d’Histoire Naturelle, CP52, 61 Rue Buffon, 75231 Paris Cedex 5, France, Telephone: 33-1-40-79-34-98, Fax: 33-1-40-79-34-99.

Reprint requests: Sedigheh Zakeri, Malaria Research Group, Biotechnology Department, Pasteur Institute of Iran, Tehran, Iran.

REFERENCES


