Genotyping of Iranian *Plasmodium vivax* populations by using circumsporozoite protein gene (CSP) marker

Akram Abouie Mehrizi 1, Parin Naimi 1, Navid Diparast Djadid 1, Ahmad Raeisi 2, Shahnam Arshi 3, Georges Snounou 4, Sedigheh Zakeri 1*

1Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, Tehran, Iran. Pasteur Avenue, P.O.BOX 13164, Tehran, Iran, Tel: +98 21 6480780, Fax: +98 21 6465132

*Corresponding author: zakeris@yahoo.com

2Manager of Malaria Control Programme, Iran, Center for Diseases Management and Control (CDMC)

3Public Health Department, Ardebil University of Medical Science

4Unité de Parasitologie Biomédicale, Pasteur Institute of Paris, France

Abstract:

Although circumsporozoite protein of *Plasmodium vivax* (*Pvcsp*) has been considered to be a potential vaccine candidate, there was no information about the sequence variation in CSP of Iranian *P. vivax*. In current study, the prevalence of two types of *Pvcsp* was determined by PCR-RFLP in 374 *P. vivax* infected patients. In north, with one exception, all parasites were VK210, however, in south both types were present. Mixed infections were common in south (12%), but it has not been observed in northern isolates. We also examined the genetic diversity of *Pvcsp* gene by sequencing analysis of 137 out of 374 field isolates obtained from resurgent parasite in north and endemic areas, in south of Iran. We have also found very limited polymorphism in the non-repetitive region of the gene in all isolates. However, repeated region of all sequenced samples showed variation in number of repeat, and also base substitutions was occurred which suggested that point mutation have caused the emergence of the variant sequence in Iranian isolates. Also such variation was much more in VK210 type than in those of VK247 type and also in southern isolates comparing to north. We also reported one novel sequence with new mutation (A/P) from southern isolate. Also few isolates have 2 copies of 19 AA insertions in post-repeated region, which has not been reported previously. The sequence data reported here provide more genetic information on *P. vivax* populations particularly from central Asia for the first time, which has also important implications for both practical measures aimed at developing control programs, and on future field-based vaccine trials with *Pvc* protein antigen.

Keywords:

Malaria; *Plasmodium vivax*; Circumsporozoite; VK210; VK247; polymorphism; sequence; Iran

Introduction:

*Plasmodium vivax* is one of the most serious parasitic diseases to affect peoples in developing countries with tropical and subtropical climates and its prevalence has been increasing now. To control this parasite several efforts have been made in developing malaria vaccine and antimalarial drugs. One of the candidates for vaccine development is the circumsporozoite protein (Nussenzweig *et al.*, 1985). Although CSP based vaccine were found to induce complete protection in rodent models, their success in human vaccine trials have been limited and this is because of the presence of sequence variation in the CS protein (Udhayakumar *et al.*, 1998).

So far based on the consensus sequence of the repeats, two major variants, VK210 with a tandemly repeated nonapeptide GDRA (D/A) GQPA and for VK247 ANGA (G/D)(N/D) QPG, of this gene have been detected. A third sequence form of the *Pvcsp* was obtained from a parasite resembling morphologically to *P. vivax* with repeat sequence APGANQEGGAA. This was found to be similar to the simian malaria parasite *P. simi-ovale* (Qari *et al.*, 1993; Qari *et al.*, 1991; Rosenberg *et al.*, 1989; Arnot *et al.*, 1985).

Different study in malaria endemic areas showed that *P. vivax* parasite with the variant repeat and post repeat sequences of the *Pvc* protein have a wide geographic distribution (Qari *et al.*, 1992; Qari *et al.*, 1991; Kain *et al.*, 1991). Such studies of the diversity of specific genes provide information about how alleles are generated and maintained in the population. However, to date there is no published information about variants and also sequences data of the
P. vivax parasites from Iran and to some extent from Eastern Mediterranean Region (EMR). Therefore, we present DNA sequences data on the Pvcsp gene in isolates collected form a re-emerged malaria area in north and also malaria endemic areas in southern part of Iran, with considerable population movement between Iran, Afghanistan and Pakistan. In addition, we investigated synonymous and nonsynonymous changes in the repeat, pre and post regions of Pvcsp gene that may result in the emergence of variant sequences, also to determine the predominate types of P. vivax isolate in both resurgent area in north and endemic areas in southern Iran. The out coming results of this study have important implications for both practical measures aimed at developing control programs, and on future field-based vaccine trials with Pvcsp protein antigen.

Materials and methods

2.1. Blood sample collection and DNA preparation

The study was conducted in resurgent malaria area of Pars Abad in Ardebil province in north (174 samples) and also endemic areas in south including, chabahar district in Sistan and Baluchistan (155 samples), Minab in Hormozgan (17 samples) and Kahnouj in Kerman (28 samples) provinces, during 2000-2003. The diagnosis was made by light microscopy, then, 1 ml blood sample was collected after informed consent from all adults and also from parents or legal guardians of infants. The Ethical Review Committee of Research Council in Pasteur Institute of Iran approved this study.

Template of parasite deoxyribonucleic acid (DNA) was obtained by phenol/phenol-chloroform extraction and ethanol precipitation as described previously (Snounou et al., 1993).

2.2. Genotyping of CSP by PCR-RFLP:

The repeat region of the Pvcsp gene was genotyped by PCR-RFLP. The oligonucleotides used to amplify a fragment of the Pvcsp gene were designed using a published sequence of the Pvcsp gene. Nested PCR amplification strategies were adopted and in the primary reaction, the whole of the gene, 1100 bp was amplified. The product of this reaction was then used to initiate a second round of amplification in which the Region I and post repeat region of Pvcsp gene was amplified.

The PCR conditions were as follows: 95°C for 5 min, 25 cycles of 62°C for 1 min, 72°C for 2 min, 94°C for 1 min followed by 62°C for 1 min and 72°C for 5 min. To distinguish between two types of P. vivax variants, PCR products were digested with Alu I (5'AG/GT3') (Invitrogen) and Mva I (5'CC/WCG3') (Fermentas) using manufacturer’s recommendation conditions. The DNA was visualized on an ultraviolet transluminator following ethidium bromide staining.

2.3. DNA sequences

Sequence analysis was used to identify polymorphism in repeat and pre and post repeated regions of Pvcsp gene of P. vivax parasites from northern (58) and southern isolates (78) of Iran. The PCR products of the second round were loaded onto a 1% agarose gel and electrophoresed. The amplified DNA was excised and purified using QIAGEN DNA purification kit (Qiagen, Germany) following the manufacturer’s instructions. Sequencing of DNA was performed in both directions for each PCR product using ABI-3100 sequencer (Primme Company, Italy). Sequences were aligned and compared with previously published Pvcsp gene sequences from published data in GenBank. The DNA and amino acid sequences were aligned using Clustal in the Megalign software of DNASTAR package (1993-1995).

Results and Discussion:

In the past malaria was endemic in most parts of Iran; however, after 1958 by starting the national malaria control program, malaria was attacked in most parts except in southeast of
Iran in three provinces including Sistan and Baluchistan, Hormozgan and tropical part of Kerman. In these areas malaria is hypoendemic and characterized by the presence of two parasite species *P. vivax* and *P. falciparum*. Malaria transmission was also interrupted in all northern parts of Iran by 1977, but it reappeared in 1994 following large displacement of the population in the border areas of the Republic of Azerbaijan and to some extent from Armenia. According to microscopy detection method the only parasite in this region is *P. vivax*, however, by using molecular diagnosis method there were some mixed infection due to both *P. vivax* and *P. falciparum* species (Zakeri et al., 2003). In this study, we investigated the genetic diversity of *P. vivax* populations from resurgent malaria in north and hypoendemic areas in south.

For amplification of the *Pvcsp* gene, genomic DNA extracted from blood samples of the patients from north isolates produced a DNA fragment of 611-677 bp in VK210 variant and 740 bp for the only VK247 type. In south isolates, the size of PCR products was 638 to 719 bp and also 638-740 bp for VK210 and VK247 types, respectively.

To determine the prevalence of the two *Pvcsp* variants in 374 infected patients, PCR-RFLP was used, which showed high sensitivity in this study (Fig. 1, 2, 3). These data showed that there are only two major *Pvcsp* phenotypes, VK210 and VK247 in Iran and no *P. vivax* like parasite was detected in this region of world; however, it was found in Brazil, Madagascar and Indonesia isolates (Qari et al., 1993). In north, parasites were pure and with one exception the predominant species (99.5%) was VK210 variant. However, in south both variants alone (70.5% VK210 and 17.5% VK247) or in mixed forms (12%) were present (Fig.4). Also the prevalence of two parasite types may be correlated with the distribution of different vector species in each malaria endemic areas. Comparison of the *Pvcsp* gene sequences of field isolates from distant geographic regions in north and south of Iran revealed that each of these regions had distinct types of polymorphic CS protein particularly based on repeat region of the gene. However, there was high degree of similarity in region I and post repeated regions of all in both study areas. In this study we have found that both repeat and post repeat regions of the *Pvcsp* gene are polymorphic (Fig.5). We found 17-18 copies of the nonapeptide repeat sequence, in *Pvcsp* gene of all VK247, which was found also in the VK247 isolate from Thailand (Burkot et al., 1989). Variation in the repeated region of all is only restricted to the 5th and 6th positions. This polymorphism was detected in Papua New Guinea and Brazilian but not in Thai isolates (Qari et al., 1991). Therefore, polymorphism in the *Pvcsp* gene is present in both repeat and non-repeat regions that point mutations have led to the emergence of variant repeat sequences, and different clonal types of *P. vivax* in the malaria endemic regions. Such study is essential before malaria vaccine programs can be implemented in malaria endemic regions. Except two isolates from south all 135 sequenced samples contained a 24 base insert 3′ (Fig.5) to the repeats (ANKKAEDA), Which have previously seen only in a North Korean and Chinese isolates (Lim et al., 2001; Mann et al., 1994) and this may suggest that circulating parasites in Iran were originated from Far East.

Hence, in summary It seems that *P. vivax* parasite populations in south of Iran display high levels of diversity than re-emerging *P. vivax* malaria in north. The reason for such diversity may be related to the increased population movement in south between malaria endemic areas and neighboring countries. During this movement people can carry the new allelic form from abroad and adds to the gene pool of the parasites in malaria setting of Iran and this parasite can establish and transmitted by local Anophelines. More importantly, we found that the repeat and non-repeat polymorphic sequences of the *Pvcsp* protein are similar in parasites from geographically distant regions of the world, thus indicating that distribution of polymorphic parasites would be limited. This limited universal polymorphism implies that a CSP-based
vaccine alone or in combination with other antigens from other parasite stages may be universally effective.

**Nucleotide sequences data for Pvcsp gene reported in this paper are available in the EMBL, GenBank and DDJB databases under the accession numbers AY367278-AY367303, AY443726, AY632256-AY632287 (Pars Abad) from north and AY443700-AY443725, AY632288-AY632318, AY632243-AY632251, AY632319-AY632330 from south.**

**Acknowledgments:**
We thank the Malaria Division, CDC, Iran, for their cooperation. We are grateful for the hospitality and generous collaboration of Zahedan University of Medical Sciences (Dr. M.T. Tabatabai), staff in Public Health Department, Sistan and Baluchistan province, Chabahar district (Dr. Zand, and Mr. Gorgij) and also Pars-Abad, Ardebil (Mr. Emdadi) for their assistance in collecting blood samples from the field. This work was supported by a grant from Pasteur Institute of Iran (No. 210 to Dr. S. Zakeri).

**References**


Mann VH, Huang T, Cheng Q, Saul A. Sequence variation in the circumsporozoite protein gene of *Plasmodium vivax* appears to be regionally biased. Mol Biochem Parasitol 1994; 68: 45-52.


Fig.1: Representative of PCR-RFLP of csp alleles for VK210 (right) and VK247 (left) variants after Alu I and Mva I digestion in P. vivax isolates from north.

Fig.2: PCR-RFLP typing of the P. vivax isolates south using Alu I.

Fig.3: PCR-RFLP typing of the P. vivax isolates from south using Mva I (VK247 Variant) southern Iran.

Fig.4: Frequency distribution of Pvcsp sequence types in wild P. vivax isolates from northern and southern Iran.

Fig.5: Representative Amino Acid sequence alignment of the CSP protein of variant Iranian P. vivax isolates. All sequenced samples of VK210 as well as VK240 variants from north and south contained ANKKAEDAG insertion in the non-repeat region, found in North Korean and Chinese isolates.