

## Analysis of the genetic variability of PvMSP-3 $\alpha$ among *Plasmodium vivax* in Brazilian field isolates

Ricardo Souza Ribeiro<sup>1</sup>, Luisa Ladeira<sup>1</sup>, Antonio Mauro Rezende<sup>2</sup>, Cor Jesus Fernandes Fontes<sup>3</sup>,  
Luzia Helena Carvalho<sup>1</sup>, Cristiana Ferreira Alves de Brito<sup>1/+</sup>

<sup>1</sup>Laboratório de Malária <sup>2</sup>Laboratório de Parasitologia Celular e Molecular, Instituto de Pesquisa René Rachou-Fiocruz, Av. Augusto de Lima 1715, 30190-002 Belo Horizonte, MG, Brasil <sup>3</sup>Departamento de Clínica Médica, Universidade Federal do Mato Grosso, Cuiabá, MT, Brasil

*Reliable molecular markers are essential for a better understanding of the molecular epidemiology of Plasmodium vivax, which is a neglected human malaria parasite. The aim of this study was to analyze the genetic diversity of P. vivax isolates from the Brazilian Amazon using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis of the highly polymorphic merozoite surface protein-3 $\alpha$  (PvMSP-3 $\alpha$ ) gene. To accomplish this, 60 isolates of P. vivax from different endemic areas in the Brazilian Amazon were collected. The PvMSP-3 $\alpha$  gene was amplified by nested-PCR. Three major types of the PvMSP-3 $\alpha$  locus were detected at different frequencies: type A (68%), B (15%) and C (17%). A single sample showed two PCR fragments, which corresponded to infection with types A and C. PCR-RFLP analysis using the HhaI restriction enzyme for 52 isolates clearly identified 11 haplotypes, eight of which were from type A, two from type B and only one from type C. Seven other isolates did not show a clear pattern using PCR-RFLP. This result might be due to multiple clone infections. This study showed a high diversity of the PvMSP-3 $\alpha$  gene among P. vivax isolates from the Brazilian Amazon, but also indicated that the detection performance of PCR-RFLP of the PvMSP-3 $\alpha$  gene may not be sufficient to detect multiple clone infections.*

Key words: malaria - *Plasmodium vivax* - genetic variation - antigenic variability - MSP3 protein - restriction fragment length polymorphism

Malaria is one of the primary global human public health problems and it is also becoming an emerging disease in regions in which it has been controlled in the past. Malaria kills over one million people each year and approximately 3.2 billion people living in more than 100 countries are at risk for infection (Arez & do Rosario 2008).

Among the five *Plasmodium* species that cause malaria in humans, *Plasmodium falciparum* is the most virulent. However, *Plasmodium vivax* has the widest worldwide distribution and therefore contributes to the global burden of morbidity (Greenwood et al. 2008). *P. vivax* causes malaria in approximately 80 million cases annually. Outside Africa, *P. vivax* accounts for more than 50% of all malaria cases, the majority (80-90%) of which occur in the Middle East, Asia and the Western Pacific, as well as 10-20% in Central and South America. Malaria caused by *P. vivax* is usually considered to be a benign disease. However, there are reports of increasing clinical severity in Southeast Asia and South America (Mendis et al. 2001, Price et al. 2007, Kochar et al. 2009, Alexandre et al. 2010). The highly significant

economic impact of *P. vivax* malaria demands that more resources should be directed specifically to research of this parasite, including studies on population structure, pathogenesis and drug resistance.

The structure of the malaria parasite population has a significant impact on gene flow and thus on the rate of emergence and spread of new mutations, leading to drug resistance or escape to vaccine-induced immunity. These studies are important to better understand the role of parasite diversity in the transmission of malaria, as well as for designing and evaluating malaria vaccines. Compared to *P. falciparum*, little is known about the genetic variability and evolutionary history of *P. vivax*. The scarcity of studies for *P. vivax* is primarily due to its lower virulence and the absence of an efficient system for continuous cultivation of this parasite (Golenda et al. 1997). Different molecular markers, such as microsatellites and single-base polymorphisms (SNPs) within the surface antigens, which generally use orthologous genes previously identified in *P. falciparum*, have been used to study the genetic variability and structure of populations of *P. vivax* (Leclerc et al. 2004, Imwong et al. 2005, 2006).

One of the most-used molecular markers to study the variability of *P. vivax* isolates is the *P. vivax* merozoite surface protein-3 $\alpha$  (PvMSP-3 $\alpha$ ). The PvMSP-3 $\alpha$  gene encodes a MSP with a molecular weight ranging from 148-150 kDa and it also contains an alanine-rich central domain that is predicted to form a coiled-coil tertiary structure (Galinski et al. 1999). PvMSP-3 $\alpha$  is highly polymorphic and its variability has been assessed using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis in population

Financial support: FAPEMIG/PROBIC, Rede Malária/Pronex-CNPq, MS/DECIT, FAPEMIG, FAPEMAT, FAPERJ

CFAB and LHC were supported by CNPq research scholarships.

+ Corresponding author: cristiana@cpqrr.fiocruz.br

Received 16 January 2011

Accepted 1 June 2011

studies using isolates from diverse geographical origins (Bruce et al. 1999, Cui et al. 2003, Mascorro et al. 2005, Ord et al. 2005, Zakeri et al. 2006, Veron et al. 2009).

The product of the PCR amplification of the PvMSP-3 $\alpha$  gene from field isolates demonstrates length polymorphism and digestion of the PCR products with restriction endonucleases indicates the presence of a great number of alleles (Bruce et al. 1999). Analysis of PvMSP-3 $\alpha$  has been performed using isolates from Papua New Guinea (PNG) (Bruce et al. 1999, Mueller et al. 2002), Thailand (Cui et al. 2003, Mascorro et al. 2005), Myanmar (Moon et al. 2009), Korea (Han et al. 2004), Pakistan and Iran (Zakeri et al. 2006, 2010, Khattoon et al. 2010), Venezuela (Ord et al. 2005), Peru (Sutton et al. 2009), India (Prajapati et al. 2010) and French Guiana (Veron et al. 2009). Here, we assess the genetic variability of geographically distinct *P. vivax* populations from Brazil using PCR-RFLP analysis of the PvMSP-3 $\alpha$  gene.

### SUBJECTS, MATERIALS AND METHODS

*Study area and sample collection* - Blood was obtained from 60 patients with an acute *P. vivax* infection in four different Brazilian Amazon regions: 10 patients from Cuiabá, state of Mato Grosso (MT), collected from July 2003-July 2004, 18 patients from Macapá, state of Amapá (AP), collected in November 2004, 16 patients from Augusto Corrêa, state of Pará (PA) collected in October 2005, and 16 patients from Manaus, state of Amazonas (AM), collected in May 2003. The incidence of malaria in Brazil is almost exclusively (99.8% of the cases) restricted to the Amazon Region. The transmission within the Amazon Basin Region is not regularly distributed because 7% of municipalities in this region account for 80% of the total registered malaria cases. The rate of transmission in Brazil is measured by the Annual Parasite Index (API), which reflects the number of positive blood smears/1,000 inhabitants. According to the guidelines of the Health Surveillance Secretariat of the Ministry of Health (SVS 2007), this index was used to stratify the areas based on the risk of malaria infection. High risk was defined as an API > 50 (Augusto Corrêa), medium risk as 10 > API < 50 (Manaus) and low risk as an API < 10 (Macapá), assuming that the API data are documented at the time of blood collection. Patients who came to the health centre in Cuiabá reported many locations of infection with different API. Because there is no malaria transmission in that locality, the area was classified as an area of variable risk. An acute *P. vivax* infection was confirmed by microscopic examination of Giemsa-stained blood smears. All patients were between 16-58 years old (mean 32 years). DNA was extracted from whole blood samples using the Puregene DNA isolation kit (Gentra Systems, Minneapolis, MN), according to the manufacturer's protocol.

*Amplification of PvMSP-3 $\alpha$  gene* - Nested-PCR amplification of the PvMSP-3 $\alpha$  gene was conducted in a reaction volume of 20  $\mu$ L using 2  $\mu$ L of DNA in the primary reaction and 2  $\mu$ L of the amplicon in the nested round. Primary and nested oligonucleotide PCR primers were previously described (Galinski et al. 1999). One unit of Taq polymerase (Promega) was used per reaction and oli-

gonucleotide primers were used at a final concentration of 0.1 mM with 0.15 mM of each dNTP in a reaction buffer (Promega) containing 2.5 mM magnesium chloride. Cycling conditions were as follows: primary reaction, one step of 94°C for 3 min and 40 cycles of 94°C for 30 s, 56°C for 30 s and 68°C for 90 s, and nested reaction, one step of 94°C for 3 min and 30 cycles of 94°C for 30 s, 57°C for 30 s and 72°C for 2.5 min. Products were visualized by ultraviolet (UV) illumination after electrophoresis on 1% agarose gels containing 0.25  $\mu$ g/mL ethidium bromide. Sizing of products was performed using a standard curve drawn from DNA markers run adjacently (1Kb plus DNA ladder, Invitrogen life technologies).

*Restriction of PCR fragments* - Nine microliters of PCR products were digested with five units of *Hha*I (Promega) in a 20  $\mu$ L reaction in a buffer supplied with the enzyme at 37°C for 3 h. DNA fragments were visualized by UV illumination after electrophoresis on 2% agarose gels containing 0.25  $\mu$ g/mL ethidium bromide.

*Population genetics analysis* - We measured intra-population diversity using expected heterozygosity ( $H_E$ ) analysis, which is the average probability that two alleles randomly obtained for each locus are different. We performed a Kruskal-Wallis test to verify whether the differences between the  $H_E$  of each locus for individual populations were significant. Between-population genetic diversity was quantified using pair wise fixation index ( $F_{ST}$ ), as implemented in the Arlequin 3.1 software (Excoffier et al. 2005). We also estimated the correlation between the matrices of pair wise genetic and geographic distances using the Mantel test (Mantel 1967) implemented by the Arlequin software.

To determine whether our samples could be grouped into genetic clusters and to infer the number of clusters that best fit the data according to allele frequencies, we used the Structure 2.3.3 software, which employs a Bayesian clustering method (Pritchard 2000). We assumed the admixture model, which considers that the genome of each individual parasite may have ancestries in more than one of the K parental populations, and a model of correlated allele frequencies, and we did not use prior information about population origin for each individual (PopFlag = 0). The log probability of the data was also considered. A certain value of K [LnP(D|K)] was calculated and compared across a range of K values to determine which one provided the best fit to the data. For each K value considered, we ran the program 10 times, each one with 100,000 burn-in and 1,000,000 iterations. K values ranged from 1-10.

The ethical and methodological aspects of this study were approved by the Ethical Committee of Research on Human Beings from the René Rachou Research Center, Oswaldo Cruz Foundation (Reports 05/2006 and 05/2010), according to the resolution of the Brazilian Council on Health (CNS 196/96).

### RESULTS

*PCR amplification of the PvMSP-3 $\alpha$  gene* - The PvMSP-3 $\alpha$  gene was successfully amplified in all *P. vivax* isolates obtained from 60 patients of four different geographic regions of the Brazilian Amazon. Based

on the length of PCR products, three different fragment sizes were found that corresponded to different locus types: type A (1.8-1.9 Kb), type B (1.5 Kb) and type C (1.1-1.2 Kb) (Fig. 1A). Type A was predominant in the four studied regions, with frequencies ranging from 44-80% (68% mean). The type B allele was detected in 15% of the total isolates and type C was found in 17% of isolates (Fig. 1B). One patient from AM showed a profile characteristic of different parasite genotypes with the presence of two different product sizes (alleles types A and C), corresponding to multiple infections and was not included in the results or further analyses.

**PCR-RFLP analysis of the PvMSP-3α gene** - The digestion of PCR products from single-infected patients with the *HhaI* restriction enzyme showed clear patterns in 52 out of 59 digested samples. Seven samples (14%) showed a complex and inconsistent pattern of fragments, possibly due to multiple-clone infections, and were excluded from the analysis. The PCR-RFLP pattern was formed by 11 fragments between 190-1,000 bp and revealed 11 different haplotypes: eight type A haplotypes (A1-A8), two type B (B1 and B2) and a unique type C (Table I). The three most prevalent haplotypes were A4, A2 and C, which were present in approximately 60% of isolates. Only three haplotypes seem to be region-specific: A5 and A8, both from PA, and A7 from MT (Fig. 2).

**Genetic diversity of the PvMSP-3α gene and population differentiation** - The genetic diversity of PvMSP-3α was measured based on the  $H_e$ , which ranged from 0.21-0.34 in different geographical populations (Table II). However, these differences were not statistically significant (Kruskal-Wallis test,  $p = 0.7526$ ). Analysis of  $F_{ST}$  values from geographic populations showed low to moderate differentiation among the populations, and the PA popu-

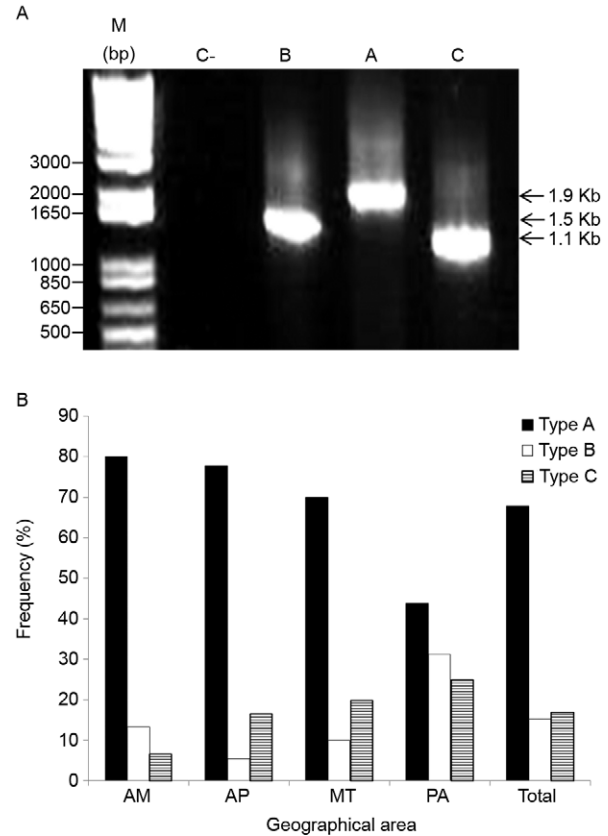


Fig. 1A: major *Plasmodium vivax* merozoite surface protein-3α (PvMSP-3α) alleles identified by a polymerase chain reaction in the *P. vivax* population from Brazilian Amazon. PvMSP-3α alleles are distinguished by size in three types: A (1.8-1.9 kb), B (1.5 kb) and C (1.2 kb); B: graphic representation of frequencies of PvMSP-3α types in isolates from different geographic areas from Brazilian Amazon. AM: Amazonas; AP: Amapá; M: DNA size maker; MT: Mato Grosso; PA: Pará.

TABLE I

Haplotype fragment patterns and their frequencies based on polymerase chain reaction-restriction fragment length polymorphism of *Plasmodium vivax* merozoite surface protein-3α analysis from 52 isolates of *P. vivax* in the Brazilian Amazon

Haplotype	Fragment size (bp) <sup>a</sup>											Freq (%)
	1	2	3	4	5	6	7	8	9	10	11	
A1	190	210	-	-	-	-	-	-	-	500	1,000	8
A2	-	210	230	-	-	-	410	-	-	-	1,000	17
A3	-	210	230	-	-	360	-	-	-	-	1,000	8
A4	-	210	230	-	320	-	-	-	-	-	1,000	23
A5	-	210	230	-	-	-	-	460	-	-	1,000	2
A6	-	210	230	-	320	-	410	-	-	-	1,000	6
A7	-	210	230	-	-	360	-	460	480	-	-	2
A8	-	-	-	-	320	-	-	-	480	-	1,000	2
B1	-	-	-	260	320	-	-	-	-	-	1,000	12
B2	-	-	-	-	-	-	-	-	-	500	1,000	4
C	190	-	-	-	-	-	-	-	-	-	1,000	17

<sup>a</sup>: fragment size estimated in 2% agarose gel ethidium bromide stained after digestion using *HhaI*; Freq: allelic frequencies in the whole isolates.

lation was the most distinct genetic population (Table II). No correlation was observed between genetic differentiation and geographical distance (Mantel test,  $p = 0.4587$ ).

**Population structure** - We analysed the clustering of PvMSP-3 $\alpha$  haplotypes based on allele frequencies using the Structure software, which determines whether the isolates are clustered according to their geographic origins. The K value which best fit the data was 4, indicating that our isolates were composed of four ancestral populations (Fig. 3A). Based on this K value, the clustering of *P. vivax* isolates from different areas was observed (Fig. 3B). However, this clustering occurred independently of the geographic origin of the isolates. Based only on PvMSP-3 $\alpha$  genotyping by PCR-RFLP analysis, 42% of the isolates from different areas had originated from one ancestral population (> 70%) and 58% of the isolates had originated from one main ancestral populations.

## DISCUSSION

The diversity of Brazilian isolates has not been previously studied by PCR-RFLP analysis of the PvMSP-3 $\alpha$  gene. Among 60 Brazilian isolates, we identified three previously described by fragment sizes. Type A was the most prevalent, which is consistent with data in the literature from different parts of world (Supplementary data). The only exception was the report from Northern Iran where type C has been the predominant PvMSP-3 $\alpha$  allele (Zakeri et al. 2006). Around the world, the prevalence of type A ranged from 52% (India)-96.4% (Colombia), with an average of 76%, excluding the data from Northern Iran (outlier) (Supplementary data). In the present study, the prevalence of types B and C ranged from 6-33% (mean 12%) and from 3-23% (mean 8%), respectively. The lower prevalence of type B and C (less than 33%) is in agreement with their reported prevalence in many different regions, with the exception of Northern Iran as discussed above (Zakeri et al. 2006). Rayner et al. (2002) identified 32.8% of type B *P. vivax* isolates while studying Asian and American samples, most of them monkey-adapted strains. However, when taking into account only American isolates, this frequency decreased to 22%. When comparing the PvMSP-3 $\alpha$  gene variability of Brazilian isolates with isolates solely from

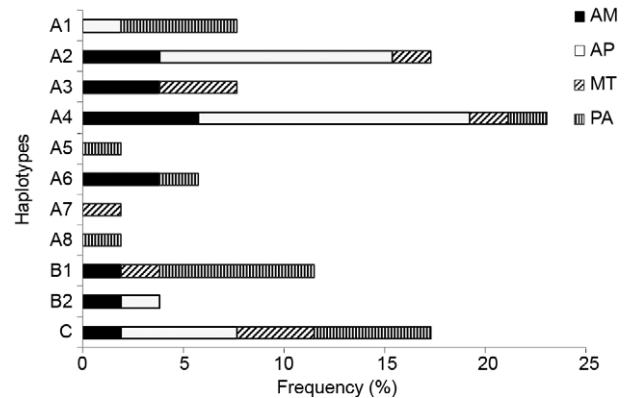


Fig. 2: graphic representation of frequencies of *Plasmodium vivax* merozoite surface protein-3 $\alpha$  haplotypes after *HhaI* restriction in each geographical population from Brazilian Amazon. Eight haplotypes were from type A (A1-A7), two from type B (B1 and B2) and only one from type C. AM: Amazonas; AP: Amapá; MT: Mato Grosso; PA: Pará.

other South America locations (Peru, Venezuela, Colombia and French Guiana), the prevalence of PvMSP-3 $\alpha$  gene types were more similar to those reported in Venezuela (Ord et al. 2005) and were more distinct from data collected in Colombia (Cristiano et al. 2008). In two localities (India and Pakistan), a rare type D allele (0.3-0.5 kb) was identified. This type occurred only in one isolate from each locality (Khatoun et al. 2010, Prajapati et al. 2010). The low prevalence of parasites with PvMSP-3 $\alpha$  genotypes other than type A identified in our study and in the large majority of previous studies supports the hypothesis that this sequence deletion, despite not being essential, may result in a reduction in parasite fitness (Cui et al. 2003). However, the explanation for this hypothesis still needs further confirmation.

In the present study, it was possible to identify 11 haplotypes of PvMSP-3 $\alpha$  using PCR-RFLP analysis among the 52 *P. vivax* field isolates that were analysed, most of which were of type A. The total number of haplotypes was similar in Brazil ( $n = 11$ , herein), French Guiana ( $n = 11$ ) (Veron et al. 2009), Colombia ( $n = 9$ ) (Cristiano et al. 2008) and Venezuela ( $n = 9$ ) (Ord et al. 2005), but it was

TABLE II

Analysis of genetic diversity and population differentiation of *Plasmodium vivax* isolates from different Brazilian Amazon regions using polymerase chain reaction-restriction fragment length polymorphism of *P. vivax* merozoite surface protein-3 $\alpha$

Population	$H_E \pm SD$	$F_{ST}$			
		Amazonas	Amapá	Mato Grosso	Pará
Amazonas	$0.24 \pm 0.19$	-	-	-	-
Amapá	$0.21 \pm 0.21$	-0.04336	-	-	-
Mato Grosso	$0.34 \pm 0.16$	-0.01840	0.02292	-	-
Pará	$0.28 \pm 0.20$	0.13181 <sup>a</sup>	0.13820 <sup>a</sup>	0.05788	-
Mean	$0.27 \pm 0.19$	-	-	-	-

<sup>a</sup>:  $p < 0.05$ ;  $F_{ST}$ : fixation index between populations;  $H_E$ : expected heterozygosity; SD: standard deviation.



slightly higher in Peru ( $n = 17$ ) (Sutton et al. 2009). The haplotype comparison is more difficult because of minor differences between fragment sizes in each study. However, we could identify approximately seven haplotypes, which were previously described, including the C haplotype, which has the same digestion profile in all previous studies. When comparing the fragment profiles, the haplotypes similar to those described here were as follows: one from Colombia (Cristiano et al. 2008), three from Peru (Sutton et al. 2009), four from PNG (Bruce et al. 2000) and six from Thailand (Cui et al. 2003). These findings suggest a global distribution of parasites containing similar PvMSP-3 $\alpha$  genotypes. The PCR-RFLP patterns of all three genotypes included a band (approximately 1.0 kb) that was slightly polymorphic. Consistent with Veron et al. (2009), this band was not used to distinguish patterns because the size differences were not easily resolved by agarose gel electrophoresis. Here, for the first time, we identified a haplotype without this band (haplotype A7), which was identified in a single isolate. This fragment pattern was consistent with the presence of a mutation that generated an additional recognition site for the restriction enzyme.

As previously demonstrated, high genetic diversity can be found even in areas of low *P. vivax* transmission (Ferreira et al. 2007, Rezende et al. 2010). Consequently, it was not surprising that the Brazilian isolates described here were equally diverse as those described in the hyperendemic areas of PNG (Bruce et al. 2000). For *P. falciparum*, a much clearer correlation between malaria endemicity and genetic variability has been demonstrated (Anderson et al. 2000). The most plausible explanation seems to be the biological differences between these two parasites, with *P. vivax* presenting relapses and early gametocytogenesis. These factors can favour meiotic recombination of different parasite haplotypes within the mosquito vector (Cui et al. 2003, Mascorro et al. 2005).

In this study, populations of *P. vivax* from different geographical areas in Brazil showed some structuring, with the population of PA being the most divergent. This result could neither be explained by geographic distance alone nor by geographic isolation. The only possible explanation is the rate of transmission. Notably, the API is higher in PA compared to other localities studied herein. Moreover, using different molecular markers, such as tandem repeats or microsatellites, some structuring of the populations was observed (Supplementary data). However, the most divergent population was not the same in different studies. This finding suggests that transmission is local and restricted to each geographic area. Some studies have reported high rates of microsatellite haplotype replacement in a longitudinal analysis (Ferreira et al. 2007, Van den Eede et al. 2011). Thus, the variability seen here could be the result of both spatial and temporal effects because the samples were collected over an interval of more than two years. However, because the geographical areas were very far from each other and the expected gene flow was low (population structuring), we believe that the temporal variability seems to be less important in our study and impossible to be evaluated with our methodological design.

Surprisingly, we found only a single *P. vivax* isolate with two PvMSP-3 $\alpha$  gene types and seven other isolates from which we could not discern the components of multiple-clone infections after PCR-RFLP analysis. Sutton et al. (2009) described the same problem and suggested the presence of multiallele detections, as reported here. Previous studies from our group using different molecular markers, such as tandem repeats and microsatellites, showed much higher rates of multiple infections upon analysis of the same isolates (Supplementary data) (Rezende et al. 2009, 2010). Because PCR-RFLP analysis detects only polymorphisms in the restriction sites of the enzyme chosen, its efficiency to identify multiple-clone infections is low, which leads to an underestimation of this phenomenon. Using this same protocol, Rayner et al. (2002) found high levels of multiple infections (40%). However, in Rayner's study, multiple infections could be identified only after sequencing the cloned PCR products. This result suggested that in multiple infections, the secondary haplotype might be in very low prevalence. The slightly lower sensitivity of detection observed for PvMSP-3 $\alpha$  PCR-RFLP analysis in field samples (approximately 400 parasites/ $\mu$ L of blood) hampered the detection of rare parasite variants (Bruce et al. 1999). We believe that the major allelic types could be assessed using PCR-RFLP analysis because the sequencing of this fragment showed few main groups with

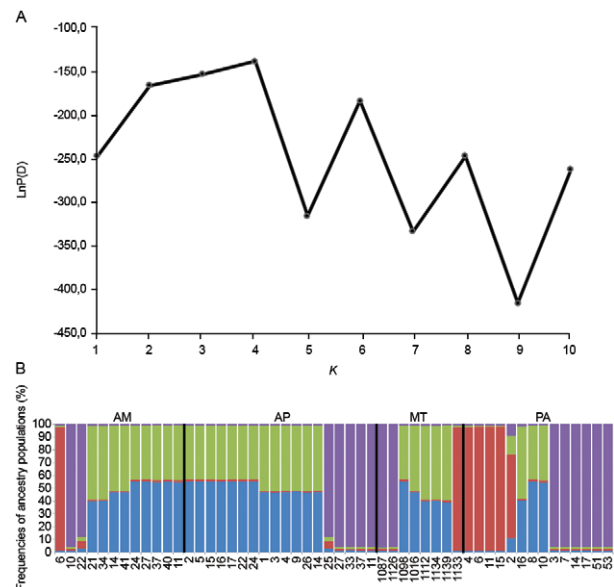


Fig. 3A: log probability plot of the data [LnP(D)] given values for  $K$  of 1-10. Circles represent the mean value of variation of 10 replicate runs at each  $K$  value; B: population structure of *Plasmodium vivax* plotted according to geographical origin of each isolate inferred from polymerase chain reaction-restriction fragment length polymorphism analysis of *P. vivax* merozoite surface protein 3- $\alpha$  at  $K = 4$ . Graphs of structure results were produced by using the Excel program. Individuals are represented by a thin vertical line from each geographic population shown above the graphic: 12 isolates from Amazonas (AM), 18 isolates from Amapá (AP), eight isolates from Mato Grosso (MT) and 14 isolates from Pará (PA). Each colour represents a genetic ancestry population.

some SNPs (Han et al. 2004, Ord et al. 2005). However, we must keep in mind that the turnover of predominant and rare genotypes can be quite fast, as has been mentioned previously. An identified genotype may depend on the day of blood retrieval. This effect has been shown by Bruce et al. (2000) at the species level and by Jafari et al. (2004) at the clonal level (*P. falciparum*).

In conclusion, our results indicate that the PvMSP-3 $\alpha$  gene is a suitable marker for analysis of the molecular epidemiology of *P. vivax* field isolates, having in its favour its simplicity and low cost compared to methods that analyse tandem repeats, microsatellites or SNPs. In addition, there is no interference when a *P. falciparum* coinfection is present. Analyses using this marker may even be enhanced when the amplified fragments are cloned and sequenced, specifically when the characterization of minor parasite populations in mixed infections is desired.

## REFERENCES

- Alexandre MA, Ferreira CO, Siqueira AM, Magalhães BL, Mourão MPG, Lacerda MV, Alecrim MGC 2010. Severe *Plasmodium vivax* malaria, Brazilian Amazon. *Emerg Infect Dis* 16: 1611-1614.
- Anderson TJC, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurwa S, French N, Whitworth J, Velez ID, Brockman AH, Nosten F, Ferreira MU, Day KP 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Mol Biol Evol* 17: 1467-1482.
- Arez AP, do Rosario VE 2008. The relevance of molecular markers in the analysis of malaria parasite populations. *Transbound Emerg Dis* 55: 226-232.
- Bruce MC, Galinski MR, Barnwell JW, Donnelly CA, Walmsley M, Alpers MP, Walliker D, Day KP 2000. Genetic diversity and dynamics of *Plasmodium falciparum* and *P. vivax* populations in multiply infected children with asymptomatic malaria infections in Papua New Guinea. *Parasitology* 121: 257-272.
- Bruce MC, Galinski MR, Barnwell JW, Snounou G, Day KP 1999. Polymorphism at the merozoite surface protein-3 alpha locus of *Plasmodium vivax*: global and local diversity. *Am J Trop Med Hyg* 61: 518-525.
- Cristiano FA, Pérez MA, Nicholls RS, Guerra AP 2008. Polymorphism in the *Plasmodium vivax msp 3 $\alpha$*  gene in field samples from Tierralta, Colombia. *Mem Inst Oswaldo Cruz* 103: 493-496.
- Cui LW, Mascorro CN, Fan Q, Rzomp KA, Khuntirat B, Zhou G, Chen H, Yan GY, Sattabongkot J 2003. Genetic diversity and multiple infections of *Plasmodium vivax* malaria in western Thailand. *Am J Trop Med Hyg* 68: 613-619.
- Evanno G, Regnaut S, Goudet J 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611-2620.
- Excoffier L, Laval G, Schneider S 2005. Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform* 1: 47-50.
- Ferreira MU, Karunaweera ND, Nunes MS, Silva NS, Wirth DF, Hartl DL 2007. Population structure and transmission dynamics of *Plasmodium vivax* in rural Amazonia. *J Infect Dis* 195: 1218-1226.
- Galinski MR, Corredor-Medina C, Pova M, Crosby J, Ingravallo P, Barnwell JW 1999. *Plasmodium vivax* merozoite surface protein-3 contains coiled-coil motifs in an alanine-rich central domain. *Mol Biochem Parasitol* 101: 131-147.
- Golenda CF, Li J, Rosenberg R 1997. Continuous *in vitro* propagation of the malaria parasite *Plasmodium vivax*. *Proc Natl Acad Sci USA* 94: 6786-6791.
- Greenwood BM, Fidock DA, Kyle DE, Kappe SHI, Alonso PL, Collins FH, Duffy PE 2008. Malaria: progress, perils, and prospects for eradication. *J Clin Invest* 118: 1266-1276.
- Han ET, Song TE, Park JH, Shin EH, Guk SM, Kim TY, Chai JY 2004. Allelic dimorphism in the merozoite surface protein-3 $\alpha$  in Korean isolates of *Plasmodium vivax*. *Am J Trop Med Hyg* 71: 745-749.
- Imwong M, Pukrittayakamee S, Gruner AC, Renia L, Letourneur F, Looareesuwan S, White NJ, Snounou G 2005. Practical PCR genotyping protocols for *Plasmodium vivax* using Pvcs and PvMSP1. *Malar J* 4: 20.
- Imwong M, Sudimack D, Pukrittayakamee S, Osorio L, Carlton JM, Day NPJ, White NJ, Anderson TJC 2006. Microsatellite variation, repeat array length, and population history of *Plasmodium vivax*. *Mol Biol Evol* 23: 1116-1018.
- Jafari S, Le Bras J, Bouchaud O, Durand R 2004. *Plasmodium falciparum* clonal population dynamics during malaria treatment. *J Infect Dis* 189: 195-203.
- Khatoun L, Baliraine FN, Bonizzoni M, Malik SA, Yan GY 2010. Genetic structure of *Plasmodium vivax* and *Plasmodium falciparum* in the Bannu district of Pakistan. *Malar J* 9: 112.
- Kochar DK, Das A, Kochar SK, Saxena V, Sirohi P, Garg S, Kochar A, Khatri MP, Gupta V 2009. Severe *Plasmodium vivax* malaria: a report on serial cases from Bikaner in Northwestern India. *Am J Trop Med Hyg* 80: 194-198.
- Leclerc MC, Durand P, Gauthier C, Patot S, Billotte N, Menegon M, Severini C, Ayala FJ, Renaud F 2004. Meager genetic variability of the human malaria agent *Plasmodium vivax*. *Proc Natl Acad Sci USA* 101: 14455-14460.
- Mantel N 1967. Detection of disease clustering and a generalized regression approach. *Cancer Res* 27: 209-220.
- Mascorro CN, Zhao K, Khuntirat B, Sattabongkot J, Yan G, Escalante AA, Cui L 2005. Molecular evolution and intragenic recombination of the merozoite surface protein MSP-3  $\alpha$  from the malaria parasite *Plasmodium vivax* in Thailand. *Parasitology* 131: 25-35.
- Mendis K, Sina BJ, Marchesini P, Carter R 2001. The neglected burden of *Plasmodium vivax* malaria. *Am J Trop Med Hyg* 64: 97-106.
- Moon SU, Lee HW, Kim JY, Na BK, Cho SH, Lin K, Sohn WM, Kim TS 2009. High frequency of genetic diversity of *Plasmodium vivax* field isolates in Myanmar. *Acta Trop* 109: 30-36.
- Mueller I, Kaiok J, Reeder JC, Cortés A 2002. The population structure of *Plasmodium falciparum* and *Plasmodium vivax* during an epidemic of malaria in the eastern highlands of Papua New Guinea. *Am J Trop Med Hyg* 67: 459-464.
- Ord R, Polley S, Tami A, Sutherland CJ 2005. High sequence diversity and evidence of balancing selection in the PvMSP3  $\alpha$  gene of *Plasmodium vivax* in the Venezuelan Amazon. *Mol Biochem Parasitol* 144: 86-93.
- Prajapati SK, Joshi H, Valecha N 2010. *Plasmodium vivax* merozoite surface protein-3  $\alpha$ : a high-resolution marker for genetic diversity studies. *J Vector Borne Dis* 47: 85-90.
- Price RN, Tjitra E, Guerra CA, Yeung S, White NJ, Anstey NM 2007. *Vivax* malaria: neglected and not benign. *Am J Trop Med Hyg* 77: 79-87.
- Pritchard JK 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Rayner JC, Corredor V, Feldman D, Ingravallo P, Iderabdullah F, Galinski MR, Barnwell JW 2002. Extensive polymorphism in the

- Plasmodium vivax* merozoite surface coat protein MSP-3 *alpha* is limited to specific domains. *Parasitology* 125: 393-405.
- Rezende AM, Tarazona-Santos E, Couto ADA, Fontes CJF, Souza JM, Carvalho LH, Brito CFA 2009. Analysis of genetic variability of *Plasmodium vivax* isolates from different Brazilian Amazon areas using tandem repeats. *Am J Trop Med Hyg* 80: 729-733.
- Rezende AM, Tarazona-Santos E, Fontes CJF, Souza JM, Couto AD, Carvalho LH, Brito CFA 2010. Microsatellite loci: determining the genetic variability of *Plasmodium vivax*. *Trop Med Int Health* 156: 718-726.
- Sutton PL, Neyra V, Hernandez JN, Branch OH 2009. *Plasmodium falciparum* and *Plasmodium vivax* infections in the Peruvian Amazon: propagation of complex, multiple allele-type infections without super-infection. *Am J Trop Med Hyg* 816: 950-960.
- SVS - Secretaria de Vigilância em Saúde, Ministério da Saúde, Brazil 2007. *Malaria epidemiologic situation in Brazil*, SVS, Brasília, 1 pp.
- Van den Eede PV, Soto-Calle VE, Delgado C, Gamboa D, Grande T, Rodriguez H, Llanos-Cuentas A, Anne J, D'Alessandro U, Erhart A 2011. *Plasmodium vivax* sub-patent infections after radical treatment are common in Peruvian patients: results of a 1-year prospective cohort study. *Plos ONE* 6: 16257.
- Veron V, Legrand E, Yrinesi J, Volney B, Simon S, Carme B 2009. Genetic diversity of msp3 and msp1\_b5 markers of *Plasmodium vivax* in French Guiana. *Malar J* 8: 40.
- Zakeri S, Barjesteh H, Djadid ND 2006. Merozoite surface protein-3 *alpha* is a reliable marker for population genetic analysis of *Plasmodium vivax*. *Malar J* 5: 53.
- Zakeri S, Raeisi A, Afsharpad M, Kakar Q, Ghasemi F, Atta H, Zamani G, Memon MS, Salehi M, Djadid ND 2010. Molecular characterization of *Plasmodium vivax* clinical isolates in Pakistan and Iran using PvMSP-1, PvMSP-3 *alpha* and pvcsp genes as molecular markers. *Parasitol Int* 591: 15-21.

## SUPPLEMENTARY DATA

Frequencies of *Plasmodium vivax* merozoite surface protein 3-*alpha* gene types of *P. vivax* isolates of different genetic population studies around the world

	Types <sup>a</sup>				Geographical region	Reference
	A	B	C	D		
Frequencies (%)	58.2	32.8	9	0	Asia and Americas	Rayner et al. (2002)
	<b>67.8</b>	<b>15.2</b>	<b>17</b>	<b>0</b>	<b>Brazil</b>	<b>This study</b>
	<b>96.4</b>	<b>1.8</b>	<b>3.6</b>	<b>0</b>	<b>Colombia</b>	<b>Cristiano et al. (2008)</b>
	<b>85.1</b>	<b>9.6</b>	<b>5.3</b>	<b>0</b>	<b>French Guiana</b>	<b>Veron et al. (2009)</b>
	51.7	17	27.6	3.4	India	Prajapati et al. (2010)
	70/24 <sup>b</sup>	12/6 <sup>b</sup>	18/70 <sup>b</sup>	0	Iran	Zakeri et al. (2006)
	77.3	16.7	4	0	Iran	Zakeri et al. (2010)
	82	6	8	2	Pakistan	Khatoon et al. (2010)
	72.7	16	6.4	0	Pakistan	Zakeri et al. (2010)
	<b>86.9</b>	<b>10.1</b>	<b>3</b>	<b>0</b>	<b>Peru</b>	<b>Sutton et al. (2009)</b>
	70.5	6.7	22.8	0	Papua New Guine	Bruce et al. (2000)
	74.8	6.5	18.7	0	Thailand	Cui et al. (2003)
	<b>59.3</b>	<b>21.9</b>	<b>18.8</b>	<b>0</b>	<b>Venezuela</b>	<b>Ord et al. (2005)</b>
	Mean	76	12	8	0.4	-

*a*: mean of frequencies types; *b*: data from Southern and Northern Iran, respectively. In bold the data from South America *P. vivax* isolates.

## SUPPLEMENTARY DATA

Comparison of genetic population parameters of Brazilian *Plasmodium vivax* isolates using different molecular markers

	$H_E \pm SD$	MI (%)	$F_{ST}$	Reference
PvMSP-3 $\alpha$ (n = 52) <sup>a</sup>	0.24 $\pm$ 0.19	2	0.003-0.23	This study
TR (n = 44)	0.56 $\pm$ 0.18	27	0.002-0.24	Rezende et al. (2009)
MS (n = 50)	0.71 $\pm$ 0.18	57	0.07-0.24	Rezende et al. (2010)

*a*: number of *P. vivax* infected patients included in each study, being 33 isolates used in all studies;  $F_{ST}$ : fixation index between populations;  $H_E$ : expected heterozygosity; MI: multiple-clone infection; MS: microsatellites; PvMSP-3 $\alpha$ : *P. vivax* merozoite surface protein 3-*alpha*; SD: standard deviation; TR: tandem repeats.