The genetic diversity of Plasmodium vivax - A Review

Wanessa Christina de Souza-Neiras*, Luciane Moreno Storti de Melo†, Ricardo Luiz Dantas Machado*

Departamento de Biologia, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, SP, Brasil *Centro de Investigação de Microrganismos, Departamento de Doenças Dermatológicas, Infecciosas e Parasitárias, Faculdade de Medicina de São José do Rio Preto, Av. Brigadeiro Faria Lima 5416, Bloco VI, 15090-000
São José do Rio Preto, SP, Brasil

The genetic diversity of Plasmodium vivax has been investigated in several malaria-endemic areas, including the Brazilian Amazon region, where this is currently the most prevalent species causing malaria in humans. This review summarizes current views on the use of molecular markers to examine P. vivax populations, with a focus on studies performed in Brazilian research laboratories. We emphasize the importance of phylogenetic studies on this parasite and discuss the perspectives created by our increasing understanding of genetic diversity and population structure of this parasite for the development of new control strategies, including vaccines, and more effective drugs for the treatment of P. vivax malaria.

Key words: malaria - Plasmodium vivax - genetic diversity - molecular markers - phylogeny - evolution

The importance of vivax Malaria

Currently, Plasmodium vivax is the most widely distributed human malaria species in the world causing an estimated 80-90 million cases each year. In the Americas and Asia, P. vivax is the most prevalent malaria species; in Brazil it represents more than 80% of clinical cases reported annually from the Amazon region (Brazilian Health Ministry 2002). The ability of this parasite to complete its sporogonic cycle at a temperature as low as 16ºC, compared with 21ºC for P. falciparum, has substantially contributed to its success in establishing stable foci of transmission in temperate zones. Vivax malaria is usually a non-lethal disease but its prolonged and recurrent infection can have major deleterious effects on personal well-being, growth and on the economic performance at individual, family, community, and national levels (Mendis et al. 2001). Alarmingly, an increasing number of clinical studies have shown the failure of treatment of the first-line P. vivax antimalarial agents, such as chloroquine (Baird 2004).

In response to the great importance of this type of infection, several studies have been proposed to investigate genetic diversity of P. vivax. The malaria parasite population structure has a significant influence on the gene flow and thus the rate at which new mutations leading to drug resistance or escape from vaccine-induced responses spread. Moreover, this information might give clues on the evolution and selection of pathogens (Rich et al. 1998). Thus, studies on malaria parasite population diversity are not only of academic interest to biologists and geneticists, but are of practical importance in the development and deployment of control strategies (Cui et al. 2003).

Polymorphic molecular markers and genetic diversity

The majority of studies on the genetic diversity of Plasmodium spp. have been based on P. falciparum, responsible for the most severe disease form, and on genes coding for antigenic determinants such as circumsporozoite surface protein (CSP) and merozoite surface protein (MSP). These antigenic genes are non-synonymous nucleotide polymorphisms and the multiple allelic forms differ in their ability to abrogate recognition by the host’s immune response (Rich & Ayala 2000). Similar approaches have been adopted to investigate P. vivax but this species has been less well studied at the molecular level when compared to P. falciparum (Cui et al. 2003).

Studies on antigenic diversity will help to predict and monitor the effectiveness of intervention strategies, such as the success of therapeutic regimens, the spread of drug resistance and the emergence of multidrug-resistant parasites. The diversity of P. vivax has been reported in terms of relapse patterns, morphology, and biochemistry (Figtree et al. 2000). Ortholog genes of P. falciparum associated to drug resistance, such as PvDHFR related to pyrimethamine resistance (de Pecoulas et al. 1998), PvCgl to chloroquine resistance (Nomura et al. 2001), and PvMDR1 to multiple drug resistance (Sá et al. 2005) have been characterized for P. vivax. However, analysis does not confirm if these polymorphic genes are really related to resistant phenotypes.

CSP

The complete or partial nucleotide sequences of CSP, the most abundant polypeptide on the sporozoite surface, have been determined. It presents a central repeat domain flanked by non-repeated amino and carboxyl sequences containing highly conserved stretches, regions I and II (Fig. 1 illustrates the CSP gene). These flanking
regions also display some degree of polymorphism (Mann et al. 1994). The central repetitive domain varies in sequence and length among *Plasmodium* spp. Until some years ago, CSP was studied as the main target for antimalarial vaccine development, because it is present in the liver stages of the parasite and is the target of protective cellular and humoral immune responses. However, the existence of variations in the repetitive sequence of its central portion has made these studies impracticable.

Analysis of *P. vivax* CSP sequences revealed that parasites have repeats belonging to one of two types of nonapeptide repeat units, GDRA(A/D)QPQA or ANGA(G/D)(N/D)QPQA, named VK210 or VK247 respectively (Arnot et al. 1985, Rosenberg et al. 1989). In 1993, a new human malaria parasite from a *P. vivax*-infected person was identified by Qari et al. who named it *P. vivax*-like. The CSP sequence of this parasite has an 11-mer repeat sequence, APGANQ(E/G)GGAA, and is different to the two previously described genotypes. This CSP gene sequence is similar to that of *P. simiovale*, a monkey malaria parasite originally found in *Macaca sinica* in Sri Lanka (Qari et al. 1993a). All CSP variant genotypes have a worldwide distribution (Kain et al. 1992, Gonzáles et al. 2001, Imwong et al. 2005, Zakeri et al. 2006).

In Brazil, Machado and Póvoa (2000) using molecular diagnosis confirmed the presence of the variant types in the states of Rondônia, Amapá, and Pará, describing the occurrence of VK210 in pure infections, whereas the VK247 and *P. vivax*-like variants were only evident in mixed infections. Seroreactivity tests have identified the presence of three variants in samples from the state of São Paulo (Curado et al. 1995, 2006) and in indigenous communities of the Amazon region (Arruda et al. 1996, 1998). Oliveira-Ferreira et al. (2004a) confirmed the seroreactivity against synthetic peptides containing the CSP immunodominant epitope of *P. vivax* variants. The authors believe that antibody responses to the CSP repeats of these variants are modulated by HLA class II molecules in individuals naturally exposed to malaria.

Interestingly, studies have also reported differences in the infectivity of anophelines to the variants, indicating that *Anopheles darlingi* were more susceptible to infection by VK210 (da Silva et al. 2006). In malaria endemic areas in Mexico, Gonzales-Ceron et al. (1999) observed the susceptibility of *An. pseudopunctipennis* and *An. albimanus* mosquitoes to infections by the VK210 and VK247 genotypes, respectively. The VK247 genotype also was detected in anophelines originating from the states of São Paulo (Branquinho et al. 1997) and Acre (Marrelli et al. 1998). These findings could be a consequence of differences in the emergence of this genotype in specific geographical regions or suggest that the VK210 genotype is the best adapted variant in the world (Machado & Póvoa 2000).

**MSP**

Several *P. vivax* merozoite surface proteins (PvMSP) have been described, including PvMSP-1, the PvMSP-3bg family, PvMSP-4, PvMSP-5, and PvMSP-9 (del Portillo et al. 1991). One of the most promising vaccine candidates against the erythrocytic forms of malaria is MSP-1. The *MSP-1* gene is very large and has 10 relatively conserved blocks alternating with regions of much higher diversity (Crewther et al. 1996). The primary structure of PvMSP-1, originally characterized from two monkey-adapted *P. vivax* strains (Belém and Salvador-1), exhibits conserved, semi-conserved and polymorphic regions (del Portillo et al. 1991, Gibson et al. 1992).

The *MSP-1* gene, expressed as a protein of 200-190 kDa on the parasite surface (del Portillo 1988), has been cloned and sequenced (del Portillo et al. 1991). Studies performed mainly with *P. falciparum* indicate that MSP-1 is processed in two steps of proteolytic cleavage during merozoite maturation. First, it is cleaved into four major fragments of 83, 30, 38, and 42 kDa (MSP-1183, MSP-130, MSP-138, and MSP-142) then, before erythrocyte invasion the MSP-142 fragment undergoes a second cleavage resulting in 33 and 19 kDa (MSP-133 and MSP-119) fragments with the latter remaining on the merozoite surface during invasion (Pacheco et al. 2006). The MSP-142 and MSP-119 fragments have received special attention in *P. falciparum* and *P. vivax* as part of vaccine formulations as they are relatively well conserved and antibodies against these fragments inhibit parasite invasion into red blood cells (Stanisic et al. 2004). The critical role of the MSP-119 fragment in erythrocyte invasion is conserved even among distantly related species (O’Donnell et al. 2000). Studies on the naturally acquired responses against the MSP-1 protein of *P. vivax* were initiated after the primary structure of the gene encoding this antigen revealed the existence of conserved and polymorphic blocks among different *Plasmodium* species (del Portillo et al. 1991). Recombinant proteins representing conserved and polymorphic regions of the N-terminal of PvMSP-1 demonstrated that polymorphic regions, as opposed to conserved regions, are immunogenic in natural infections. The C-terminal of PvMSP-1 is the most immunogenic portion of the molecule and the presence of antibodies against it is associated with recent malaria attacks (Soares et al. 1997, 1999). Nogueira et al. (2006) demonstrated an association of clinical protection and reduced risk of infection with naturally acquired antibodies against the N-terminal but not the C-terminal portion of PvMSP-1 and that most asymptomatic individuals presented antibodies against the N-terminal. However, this association does not imply...

![Fig. 1: structure of the Plasmodium vivax CSP gene, with two highly-conserved terminal non-repeat regions (RI and RII); a central repetitive (CR) domain, with a variable number of tandem repeats, and a short IR (insertion region). In the CR are indicated below, the peptide variations from protein. According to Qari et al. (1993a).](image-url)
that antibody response to the PvMSP-1 N-terminal region is in itself the mechanism of protection or simply a marker of this mechanism (Nogueira et al. 2006).

The genetic polymorphism of the Plasmodium spp. MSP-1 appears to be maintained by positive natural selection, both in P. falciparum (Hughes 1992, Escalante et al. 1998, Conway et al. 2000) and P. vivax (Putaporntip et al. 2006). Similar observations have been made with other malaria antigens in which the host immune system is considered to be the selective driving force that allows the accumulation and frequent switch of suitable mutations in the parasite population (Escalante et al. 2004). There is extensive allelic diversity of MSP-1 among isolates, and so this polymorphism may hamper development of effective vaccines.

The PvMsp-1 and PfMsp-1 genes are basically similar. The organization of these genes are represented by a mosaic organization of several interallelic variable blocks flanked by conserved blocks, dimorphic substitutions in conserved blocks and rather polymorphic substitutions in variable blocks; with allelic recombination as a mechanism for the generation of new alleles (Putaporntip et al. 2002). Analysis of specific gene regions derived from parasite isolates from Sri Lanka (Premawansa et al. 1993), Colombia (Mancilla et al. 1994), and Thailand (Putaporntip et al. 1997) suggest interallelic recombination between the sequences typified as Bel and Sal-1, supporting the notion that the MSP-1 polymorphism in P. vivax is dimorphic in nature, as occurs in P. falciparum (Tanabe et al. 1987).

The variable blocks showed variation in repeats and non-repeat unique sequences. Numerous recombination sites were distributed throughout the PvMsp-1 gene in both conserved blocks and variable block unique sequences, without the distribution being uniform. A substantial percentage of replacements within conserved blocks result in changes of binding to HLA class II alleles. Polymorphisms in the T cell epitope regions could enable parasites to escape the host immune response. Therefore, host immune pressure plays a crucial role in the evolution and maintenance of polymorphisms in PvMsp-1 (Putaporntip et al. 2002). In this scenario, mutations are maintained longer in the parasite population than expected if genetic drift were the sole process acting on genetic polymorphisms.

A number the genes encoding for P. vivax MSPs have been identified. PvMsp-3a, PvMsp-3b and PvMsp-3g are members of a multi-gene family of related MSPs (del Portillo et al. 1991). The three encoded proteins share only 35-38% amino acid identity and 48-53% similarity in pair-wise comparisons. PvMsp-3a, similar to PvMsp-1, is very polymorphic and has been used as a genetic marker in population studies of isolates from diverse geographic localities and origins (Bruce et al. 1999). Even though PvMsp-3a is very polymorphic, this polymorphism is restricted to the N-terminal, while the C-terminal is well conserved (Rayner et al. 2002). Investigation of PvMsp-3b diversity was by cloning and sequencing full-length gene fragments from fifteen P. vivax isolates originating from Asia, South America and the Pacific region. PvMsp-3a and PvMsp-3b are extremely polymorphic paralogs with multiple gene sizes and sequences present in the different isolates (Rayner et al. 2004).

Two proteins, recently described in P. falciparum, MSP-4 and MSP-5, are found in tandem on chromosome 2 with the synthetic region of the genome being identified in rodent malaria. In P. vivax, which is quite phylogenetically distant from P. falciparum, both MSP-4 and MSP-5 homologues can be found with their relative arrangements in respect to surrounding genes being on the whole preserved. The PvMsp-4 and PvMsp-5 genes have two-exon structures with the sizes of the second exons being better conserved than the first exons in both genes, when compared to structure genes of P. falciparum (Black et al. 2001). PvMsp-9 is characterized along with ortholog genes from related simian malarias, are highly conserved. P. vivax samples from different geographical regions show that the N-terminal region of PvMSP-9 is the most conserved, while the tandem repeats have diverged only in the length and number of units (Vargas-Serrato et al. 2002). Both N-terminal and the tandem repeat regions of MSP-9 are immunogenic in mice (Oliveira-Ferreira et al. 2004b). Comparative interspecies investigations of the potential role of Plasmodium MSP-9 in merozoite invasion of erythrocytes and as a candidate for malaria vaccine may be useful.

Duffy binding protein

Duffy binding protein (DBP) erythrocyte invasion by P. vivax merozoites is dependent on the parasite ligand, the P. vivax DBP, binding to the Duffy antigen receptor for chemokines (DARC). Individuals that lack the Duffy surface antigen on their erythrocytes are naturally resistant to P. vivax malaria. Thus, P. vivax DBP provides an attractive target for vaccine-mediated immunity (Xainli et al. 2000, Souza et al. 2006). DBP is a 140 KDa protein located within the microthemes of Plasmodium spp. merozoites and characterized by two functionally conserved cysteine-rich regions. Region II and Region IV. Region II (DBPII) contains the binding motifs necessary for the adherence of DBP to DARc on the erythrocyte surface. Critical binding motifs in DBPII have been mapped to a region of a 170 amino-acid (aa) stretch that includes cysteines and where some hydrophobic amino acid residues are conserved and other amino acids are highly polymorphic in the ligand domain; this diversity varies geographically. The pattern of excessive polymorphisms occurring within the ligand domain and the high rate of non-synonymous polymorphisms suggest that this allelic variation functions as a mechanism of immune evasion. Previous studies have confirmed that this strong positive selection pressure in the DBPII ligand domain acts promoting greater diversity (Xainli et al. 2000, Souza et al. 2006).

Many factors may contribute to genetic diversity in malaria populations – mutations, intragenic recombination determined by multiplicity of infections and transmission intensity, natural selection, gene flow between different regions and population size. Recombination is likely to be important to maintain diversity for P. vivax and may be a critical source of variations in the dbpII gene. This origin of the variations may be more impor-
tant for malaria than previously appreciated as the recent use of PCR (polymerase chain reaction) has demonstrated that many people in endemic areas have chronic, asymptomatic infections often consisting of multiple parasites (Bruce et al. 2000). Meiotic recombination does not appear to be an important factor contributing to the diversity of the dbp gene. It is also possible that the dbp gene may contain a mutational “hot spot” that might mask the presence of recombination in \textit{P. vivax} dbpII (Cole-Tobian & King 2003).

The cumulative polymorphisms, which Xainli et al. (2000) identified in isolates from Papua New Guinea, show that 124 of 133 (93%) individual mutations occur in the critical binding region between cysteines 4 and 7 (aa 460-291). All but one of these mutations is non-synonymous. Compared to the Sal-I isolate, a Belém strain from the Brazilian Amazon region, most polymorphisms are located in five residues: codons 308, 384, 390, 424, and 447 (Xainli et al. 2000). In Brazil, Souza et al. (2006) analyzed the DBP variability in isolates from the Amazon region and identified 14 polymorphic residues in the DBP ligand domain compared to the Sal-I sequence, with some of them being identical to those previously described in other regions of the world. The other seven polymorphic residues seem to be unique among Brazilian Amazon isolates. By grouping these residues, the authors constructed eight partial variant families representing the haplotypes present in Brazilian isolates. This large number of isolates, consisting of a diverse population of genetically distinct clones, may reflect random recombination which occurs during frequent mixed infections observed in distinct malaria endemic areas (Souza et al. 2006).

Diverse studies provide evidence for discrete allelic families in different \textit{P. vivax} endemic regions of the world. DBPPII diversity varies both within and between populations. Polymorphisms of this protein are under great selection pressure, and the presence of distinct allelic families in different geographic areas will complicate the development of a vaccine, emphasizing the need for better understanding on how genetic diversity is related to natural immunity. It will be critical to identify which regions of DBP are functionally constrained yet remain sufficiently immunogenic to stimulate a protective immunity so that a vaccine might not require multiple allelic forms for different geographic regions (Cole-Tobian & King 2003).

\textbf{vir genes}

Malaria parasites have clustered multigene families in subtelomeric regions of chromosomes, where high recombination rates facilitate their evolution and diversity. Thus, \textit{P. vivax} contains a major subtelomeric multigene superfamily termed \textit{vir} (\textit{P. vivax} variant genes), which corresponds to about 10% of the coding sequences (Fig. 2). The \textit{vir} genes are composed of different subfamilies (denominated A-F) organized by sequence similarities and expressed during intra-erythrocytic parasite development (del Portillo et al. 2001). Studies indicate that the \textit{vir} genes are probably related to mechanisms of antigenic variation; the host’s efforts to eliminate the parasite are counteracted by the parasite’s capability of constantly changing the specificity of its proteins (Brown & Brown 1965, Reeder & Brown 1996, Merino et al. 2006).

The \textit{vir} genes and pseudogenes have similar structures. They have three-exon structures, which contrast with the 10-exon structure of the \textit{sicavar} genes of \textit{P. knowlesi} and with the two-exon structures of the \textit{var}, \textit{rif} and \textit{stevor} genes of \textit{P. falciparum}. The second exon is highly variable containing a transmembrane domain and conserved cysteine residues. The region covering the point between exons 2 and 3 is extremely well conserved. Similar to the \textit{P. knowlesi} sicavar protein and the \textit{P. falciparum} PfEMP-1 protein, typical signal sequences are lacking in \textit{vir} antigens. The major role of \textit{vir} genes and their encoding of variant proteins in natural infections is currently unknown, although recently it has been proposed that they play a role in spleen-specific cytoadherence (macrophage-clearance escape) and in chronic infections. Thus, the function of \textit{vir} genes is probably not directly related to the antigenic variations in the strict sense (del Portillo et al. 2004). The fact that not all their proteins are exported to the surface of infected red blood cells (Fernandez-Becerra et al. 2005) reinforces the idea that \textit{vir} genes may have different functions related to immune evasion (Merino et al. 2006). The acquired immune responses to \textit{P. vivax} variant antigens were also evaluated in individuals living in Brazilian malaria-endemic areas. This study showed that there is a low frequency of individuals responding to each variant antigen in these regions, which may explain host susceptibility to new episodes of the disease (Oliveira et al. 2006).

Most of the studies on the repertoire of these multigene families in natural infections, in particular those from human malaria parasites, are related to \textit{P. falciparum} (Kirchgatter et al. 2000, Fowler et al. 2002). This knowledge is essential to understand the genetic diversity and evolution of these genes, which will contribute to elucidate chronicity in \textit{Plasmodium}. On the other hand, some population and genetic diversity parameters can be better investigated using neutral molecular markers or those that are not undergoing strong selection processes, such as microsatellite loci (Leclerc et al. 2004).
**Microsatellites**

Microsatellites are simple sequence tandem repeats that are generally hypervariable, codominant and locus specific. They are considered important neutral molecular markers as they are not directly subjected to host immunity. The neutral source of microsatellite polymorphisms is replication slippage, which is a commonly observed replication error in repetitive sequences that occurs when the new strand mispairs with the template strand (Russell et al. 2006).

Recently, several authors have been discussing the true degree of allelic diversity presented by the microsatellites (Gomez et al. 2003, Leclerc et al. 2004, Imwong et al. 2006, Russell et al. 2006). In direct contrast to the results of Gomez et al. (2003), Leclerc et al. (2004) found a low number of polymorphisms in the microsatellites they examined. In this study, 13 microsatellite sequences were isolated with 9/13 of these being completely monomorphic in eight analyzed populations, while of the remaining four loci only one showed extensive polymorphism. Using the draft of the unpublished *P. vivax* genome, Imwong et al. (2006) designed primer sets for 11 di-nucleotide microsatellites. Their data showed that *P. vivax* had a high allelic diversity. In the data of Leclerc et al. (2004), the isolated microsatellite sequences have very short repeat arrays (median = 5.5) and so would not be expected to show high levels of variation. Imwong et al. (2006), on the other hand, analyzed microsatellites with 12-18 repeats (median = 16). For this reason, the explanation given by Imwong et al. (2006) for the stark difference in their results compared to Leclerc et al. (2004) was that microsatellite variation is dependent on the length of the repeating sequence. Longer arrays are more diverse than shorter ones because slippage mutations become exponentially more common with an increase in array length. If microsatellites are stratified on the basis of length, *P. vivax* has a similar level of polymorphism to *P. falciparum* (Russell et al. 2006). These studies highlight the importance of measures standardizing genetic variations by repeat array length (Petit et al. 2005). In this context, Ferreira et al. (2007) recently standardized the use of tri- and tetranucleotide microsatellites, which usually yield more accurate allele scoring than dinucleotide markers to analyze the population structures of *P. vivax* from rural Amazonia. The authors showed that *P. vivax* isolates obtained during cross-sectional and longitudinal surveys in rural Amazonia display extensive genetic diversity and frequent multiple clone infections coexisting with strong multilocus linkage disequilibrium.

**Other polymorphic molecular markers**

The gene of *PvAMA-1* encodes a very important protein for erythrocyte invasion which is highly conserved in *Plasmodium* spp. However, this gene has few predominant haplotypes, displaying very limited genetic diversity within any geographic region; a more comprehensive study of the whole gene is necessary to confirm this observation. On the other hand, the presence of single nucleotide polymorphism (SNP) has been identified in some gene regions (Cui et al. 2003). This provides many potential alleles thereby making this gene a useful marker for typing parasite populations (Figtree et al. 2000).

The antifolate agents, sulfadoxine and pyrimethamine (SP), are commonly used to treat *P. falciparum* malaria. Nevertheless, they can also affect the *P. vivax* parasite if it co-exists with *P. falciparum* as both species have common drug targets. The dihydrofolate reductase (DHFR) enzyme of *Plasmodium* spp. is the therapeutic target of the pyrimethamine component of SP. In addition to point mutations in the target enzymes of the respective parasite, the *PvDHFR* gene also has a size polymorphism resulting from the deletion of five or six amino acids. Four allelic variants have been observed to date. However, the fact that the gene is under drug selection must be taken into account if this polymorphism is to be used as a genetic marker (Cui et al. 2003). Even less is known about associations between these specific alleles in *P. vivax* populations in Brazil.

**New tools for studies of the *P. vivax* genome**

Unfortunately, due to the necessity of continuously maintaining this parasite in culture, the low parasitemias associated with natural infections and the difficulty of adapting field isolates to grow in monkeys, research on *P. vivax* remained largely neglected, up to a decade ago (Merino et al. 2003). This situation has changed dramatically in recent years. Despite the limited availability of *P. vivax* biological material, recent research has led to the construction of *P. vivax* genome in yeast artificial chromosomes (YAC). This new tool has revolutionized studies on the structure and function of the *P. vivax* genome, with the search of new polymorphic molecular markers (Camargo et al. 1997, del Portillo et al. 2001, Merino et al. 2003, 2006, Feng et al. 2003, Gomez et al. 2003, Fernandez-Becerra et al. 2005). Moreover, the availability of the *P. vivax* telomeric YAC clones has helped in the identification of many antigen-encoding genes mapped to the dynamic subtelomeric domains of plasmodial chromosomes; this fact has led to a suggestion that the recombination within these domains has been recruited as a novel mechanism to generate antigenic diversity (Camargo et al. 1997).

Few polymorphic molecular markers, in particular, orthologs of previously identified *P. falciparum* antigen genes, have been used for population studies of *P. vivax*. SNPs markers have been mapped to a chromosome segment by sequencing a large insert of the YAC library (Feng et al. 2003) and as SNPs are often present at a high frequency, they are ideal genetic markers. Identification and development of large numbers of genetic markers such as SNPs from *P. vivax* will provide a framework on which studies of molecular evolution and genetic mapping can be based.

Gomez et al. (2003) identified, for the first time, a locus in a *P. vivax* telomeric YAC clone that contains simple sequence repeats. These sequences known as microsatellites were located within the second exon of the *vir13* pseudogene. In 2003, Feng et al. analyzed a contiguous chromosome segment of approximately 100
kb from five isolates, revealing 191 SNPs and 44 size polymorphisms. These analyses showed that *P. vivax* has a highly diverse genome that may represent some challenges for drug and vaccine development. This study also showed that SNPs tend to cluster in intergenic regions or even in specific genes that may be under selection (Feng et al. 2003).

Starting from genomic library constructed in YACs, using parasites obtained directly from human patients, Merino et al. (2003) described the first expressed sequence tags (ESTs), parasite genes expressed during *P. vivax* pathogenesis development. The description of EST is also a valuable resource to validate gene predictions and to create a gene index for this malaria parasite.

Recently, Imwong et al. (2006) described patterns of variations of 11 dinucleotide microsatellites in *P. vivax* populations from Colombia, India and Thailand using the YACs libraries as an important tool to better understand the parasite’s genome. These data show that *P. vivax* has a highly diverse genome and provided useful information to further understand the genome diversity of the parasite.

**Phylogenetic characteristics of the *P. vivax***

Almost 60 years ago, Haldane proposed that human malaria might act as a selective force on human populations. Since then, several studies have been proposed to test this hypothesis, starting with attempts of phylogenetic origin reconstruction of malaria parasites (Escalante et al. 1995).

The first phylogenetic reports using molecular biology analyses have focused primarily on *P. falciparum*, an agent of malignant malaria. A previous study suggested that *P. falciparum* is older than *P. vivax* (Escalante et al. 1995), and the virulence of *P. falciparum* has been attributed to the fact that it recently became a human parasite due to a host switch (probably, birds) between 5,000 to 10,000 years ago (Boyd 1949, Livingstone 1958, Snowin et al. 1991, Escalante et al. 1995). Low microsatellite and tandem repeat variability may indicate that *P. vivax* has only infected humans recently (10,000 years ago) (Leclerc et al. 2004), however a different study based on polymorphisms of two nuclear and one plastid gene places the origin at between 45,000 and 81,000 years ago (Escalante et al. 2005). Through these preliminary studies, two major conclusions were reached: (a) malaria parasites arose independently as human pathogens and, (b) *P. reichenowi*, a chimpanzee parasite, is the species that shares the most recent common ancestor with *P. falciparum* (Escalante et al. 1995). Thus, questions about the age and geographic origin of *P. vivax*, the most prevalent human malaria parasite in the world, remain largely unresolved. Recent studies have suggested possible origins for *P. vivax* but most are not based on strong phylogenetic data. Two hypotheses have been systematically discussed. One of these placed the origin of *P. vivax* in Southeast Asia, together with other *Plasmodium* parasitic species in non-human primates (Carter & Mendis 2002, Carter 2003, Jongwutiwes et al. 2005). Hence, it is possible that *P. vivax* might have fixed in this area through some hominoid lineage that had its origin there, even modern humans, since host switches seem to be common phenomena among malaria parasites. This argument was also supported by the abundance of simian malaria parasite species in Asia and the presence of several macaque parasites that shared similar morphological and biological characteristics with *P. vivax* (Carter 2003). This hypothesis has been well accepted in recent years (Escalante et al. 2005, Mu et al. 2005, Jongwutiwes et al. 2005).

It seems very probable that Duffy-negative homozygote individuals (lack of the Duffy blood group antigen), are completely protected against *P. vivax* infection (Escalante et al. 2005). Thus, the high prevalence of these genotypes among African populations, has been used to supplant *P. vivax* origin on this continent, showing that a selective pressure might have been generated in favor of the survival of the parasite organism, because the presence of the Duffy protein is essential for the merozoite invasion of red blood cells. Nevertheless, for that to have really happened, the presence of a strong selection factor would have be necessary, which seems unlikely as *P. vivax* does not exhibit high levels of virulence in terms of mortality rates. For this reason, it is probable that the Duffy-negative genotype present on the African continent, could have been fixed by another process (selection due to another pathogen or chance) and then became a barrier against the introduction of *P. vivax* (Carter 2003, Escalante et al. 2005). Also attempts to test this hypothesis starting from an investigation of two nuclear genes, b-tubulin and cell division cycle 2 and a gene from the plastid genome, the elongation factor *Tu* (*TufA*) were made. In this investigation, most of the data was not compatible with the explanation that *P. vivax* was a *Homo* parasite before the expansion of the hominoid populations out of Africa. Hence, the authors suggest that *P. vivax* was probably derived from ancestral macaque parasites when hominoids colonized Southeast Asia (Escalante et al. 2005).

One the other hand, the first evidence of the origin of *P. vivax* based on molecular data using mitochondrial *cytochrome b* gene analysis showed strong similarities of *P. vivax* with a species of a simian malaria parasite from Asia, provided by an apparently recent species radiation. This phylogenetic study also suggests that *P. vivax* could have originated from a malaria parasite of non-human primates as a result of a host switch, probably from a macaque (Escalante et al. 2005, 2006). These analyses received additional support when the combined data of complete mitochondrial genomes were reported in two independent studies (Mu et al. 2005, Jongwutiwes et al. 2005). This information can maybe explain the fact that several isolates of *P. vivax* are practically identical to *P. simium* and the previous suggestion that this South American species originated from a host switch from humans to monkeys (Escalante et al. 1998, 2005, Mu et al. 2005).

As in other phylogenetic studies (Fig. 3), the *P. vivax* variant genotypes belong to a same clade that includes several types of primate *Plasmodium* species (Escalante et al. 1995, Qari et al. 1996). Analyses of the non-repetitive portion of *P. vivax* CSP (*PvCSP*) showed homology with CSPs genes of two simian malaria para-
sites, *P. cynolmolgi* and *P. knowlesi* (Arnot et al. 1985). In contraposition, the *P. vivax*-like gene is different to the CSPs genes of other human malaria parasites. However, the CSP gene sequence of this variant genotype is practically identical to *P. simiovale CSP* and has similar morphologically to two other malaria parasites, *P. simium* and *P. ovale* (Qari et al. 1993b, Escalante et al. 1995). The similarity between *P. simiovale* and *P. vivax*-like has been driving intense discussions, regarding the evolutionary origin of this variant, as well as on clinical epidemiology and the vectors involved in its transmission (Qari et al. 1994). Escalante et al. (1995) also confirmed these findings through phylogenetic studies based on CSP genes, showing that the analogies between these two parasites determine important epidemiological consequences; primate hosts can serve as reservoirs for human malaria parasites. However, phylogenetic studies have been developed with the CSP gene as the only molecular marker so they cannot explain the evolutionary relationship among the three variant genotypes of CSP, or the genetic distance of these with other primate parasites that possess molecular similarities with *P. vivax* (Escalante et al. 1995).

In summary, understanding of the genetic recombination patterns and sequence variation may help to design vaccines, which represent the worldwide repertoire of polymorphic malarial surface antigens and elucidate the selection events associated to drug resistance.

**ACKNOWLEDGEMENTS**

To David Andrew Hewitt for the English revision of this manuscript.

**REFERENCES**


Genetic diversity and phylogeny of *P. vivax* • Wanessa Christina de Souza Neiras et al.


Genetic diversity and phylogeny of *P. vivax* • Wanessa Christina de Souza Neiras et al.


