Vivax series:

The genetic diversity of *Plasmodium vivax* populations

Liwang Cui¹, Ananias A. Escalante², Mallika Imwong³ and Georges Snounou⁴

¹Department of Entomology, Penn State University, 501 ASI, University Park, PA 16802, USA
²Centers for Disease Control and Prevention, National Center for Infectious Disease, Division of Parasitic Diseases, Mail Stop F12, 4770 Buford Highway, Chamblee, GA 30341, USA
³Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok 10400, Thailand
⁴Unité de Parasitologie Biomédicale, Institut Pasteur, 25 Rue du Dr Roux 75724, Paris, France

Little is known of the genetic diversity and population structure of *Plasmodium vivax*, a debilitating and highly prevalent malaria parasite of humans. This article reviews the known polymorphic genetic markers, summarizes current data on the population structure of this parasite and discusses future prospects for using knowledge of the genetic diversity to improve control measures.

Until the middle of the 1900s, *Plasmodium vivax* was the most globally widespread, and arguably the most prevalent, of the four malaria parasite species that infect humans (*Plasmodium falciparum*, *P. vivax*, *Plasmodium ovale* and *Plasmodium malariae*). The ability of this parasite to complete its sporogonic cycle at a minimum lower temperature of 16°C, compared with 21°C for *P. falciparum*, has substantially contributed to its success in establishing stable foci of transmission in the temperate zones. The global eradication campaign, carried out by the WHO in the mid-1950s, was most successful in the temperate endemic regions of Europe, Asia and the Americas and, as a result, the range of *P. vivax* was temporarily curtailed. However, the prevalence of *P. vivax* has been increasing lately, and the ~70–80 million cases currently recorded annually are of global public health importance [1]. Despite this, research aimed at understanding this parasite species has lagged by comparison with that carried out for *P. falciparum*. Unlike *P. falciparum*, *P. vivax* is rarely directly lethal and causes a form of malaria termed benign tertian malaria. However, the often-protracted course of a *P. vivax* infection is by no means clinically benign, and infections are a very important cause of morbidity and social economic loss [1].

Even today, there are sporadic reports of severe and fatal *P. vivax* infections [1,2]. As resources for malaria control are currently often directed against the more deadly *P. falciparum* malaria, the hidden burden of *P. vivax* malaria is likely to become increasingly apparent. Moreover, the recent emergence of chloroquine-resistant *P. vivax* strains is of great concern [3,4]. Clearly, a more comprehensive research effort directed at *P. vivax* malaria is urgently required.

The parasites that make up a *Plasmodium* spp. are heterogeneous for several important parameters, including: innate or acquired resistance to a particular drug; course and duration of the infection; symptoms; immunological crossreactivity; and transmissibility by anopheline vectors. A substantial proportion of our knowledge of the biological diversity of *P. vivax* results from observations of experimentally infected humans (Box 1). Relapses caused by a latent liver-stage parasite, the hypnozoite, are characteristic of *P. vivax* infections, and their pattern and frequency serve to distinguish tropical from temperate strains. The description of Russian and Chinese strains, by Nikolaev in 1949 and by Jiang and colleagues in 1982, where the primary infection occurs eight months or more following inoculation by an infected mosquito, led to establishing them as two subspecies, *Plasmodium vivax hibernans* and *Plasmodium vivax multinucleatum*, respectively. Whether these represent the same subspecies is unknown since the Russian strains had disappeared by the 1980s. A recent suggestion based on molecular polymorphisms and susceptibility of mosquitoes to infection [5] that New World *P. vivax* parasites might represent a separate subspecies, *Plasmodium vivax collins*, requires further biological studies before confirmation.

Parasite diversity is commonly observed within an endemic area, where different parasite strains circulate in the human host. The ability to distinguish between the different populations is a pre-requisite to understanding the local and global epidemiology of a parasite species. Studies on malaria parasite population diversity are not only of academic interest to biologists or population geneticists, but are of practical significance for the strategic development and deployment of control measures. On one hand, the parasite population structure, usually displaying considerable geographical variations, could affect the distribution of existing parasite strains, and this in turn determines the usefulness of a particular vaccine or drug formulation in a malaria-endemic region. On the other hand, parasite population structure will have significant influence on the gene flow and thus the rate at which new mutations leading to drug resistance or escape
from vaccine-induced responses will spread. Extensive efforts in this field over the past two decades have contributed to our understanding of the genetic structure of *P. falciparum* populations (reviewed in Refs [6,7]). The genetic diversity of *P. vivax* parasites has been less investigated.

Population genetic studies of malaria parasites

Two basic approaches have been followed in population genetic studies of malaria parasites: (1) those that aim to understand the parasite population genetic structure and intrahost dynamics, including the stability of strains and complexity of infections; and (2) studies assessing the diversity of specific genes encoding vaccine antigens and those associated with resistance to antimalarial drugs.

The studies on population structure quantify how alleles are spatially dispersed and how the overall genetic diversity is organized. These studies provide important indications as to the origin, dispersal and stability of multilocus genotypes. This information is essential for predicting and monitoring the effectiveness of intervention strategies such as: (1) the usefulness of combining specific drugs or therapeutic regimens; (2) the dispersion of drug resistance and the emergence of multidrug-resistant parasites; (3) the long-term impact of physical interventions such as bednets; and (4) the origin of genotypes that
might elude the immune response elicited by multivalent vaccines. Knowledge of the genetic diversity can also be used to assess geographical differentiation and explain the spatial distribution of alleles at a given loci. For example, it can provide baseline information for surveillance systems directed to identifying the origin of isolates in areas with occasional imported malaria cases.

Analysis of defined genes provides information about how different alleles are generated and maintained in the population. Specifically, it addresses the relevance of factors such as intragenic recombination and natural selection on the observed polymorphisms. This approach answers fundamental questions related to the association of specific alleles with drug resistance, and how the genetic variation at a given locus affects vaccine development and deployment.

Polymorphic molecular markers and \textit{P. vivax} population diversity

High levels of polymorphism have been described for many parasite molecules, particularly antigens that might be under selection by host immunity. Initially, techniques employing monoclonal antibodies and isozymes detected high levels of polymorphism in \textit{P. vivax} isolates collected in the field [8,9]. A major obstacle for this type of approach, and indeed for research on \textit{P. vivax} in general, is the difficulty in maintaining this pathogen in vitro. The recent demonstration of continuous culture of \textit{P. vivax} [10] is a major step forward, although the need for a constant supply of human reticulocytes is highly restrictive. With respect to genetic analyses, this limitation can be largely overcome by the detection of diversity at the nucleotide level using the powerful technique of polymerase chain reaction (PCR), alone or in combination with sequencing and restriction analysis. The lack of suitable genetic markers, such as microsatellites, single nucleotide polymorphisms (SNPs) or repetitive regions, has limited the number of field studies addressing basic molecular epidemiological questions. However, it is expected that the complete sequence of the parasite will increase the number of genetic markers when it is released. Nevertheless, polymorphic markers encoding parasite surface antigens have been used as genetic markers to study parasite diversity. Commonly used methods for large-scale population analysis, often developed for \textit{P. falciparum}, include the design of allele family-specific primers for PCR detection of different alleles and their size polymorphism [e.g. genes of the merozoite surface proteins (MSP), \textit{PvMSP1} and \textit{PvMSP2}], repeat region size polymorphism [genes for the glutamate-rich protein (GLURP) and the circumsporozoite protein (CSP) \textit{PvCSP}], or PCR/restriction fragment length polymorphism (e.g. \textit{PfMSP2} and \textit{PvMSP3a}). Occasionally, allele families are determined by hybridizing the PCR product to allele-specific probes (e.g. \textit{PfMSP2} and \textit{PvCSP}). Currently, DNA markers for genotyping \textit{P. vivax} include the genes encoding MSP1 and MSP3a, apical membrane antigen 1 (AMA1), CSP, GAM1 and dihydrofolate reductase (DHFR).

\textbf{MSP}

To date, genes encoding seven MSPs have been cloned from \textit{P. vivax}, all named in accordance with the \textit{P. falciparum} orthologues. \textit{PvMSP1} and \textit{PvMSP3a} are highly polymorphic and are currently used for genotyping \textit{P. vivax} field isolates [11,12]. The primary structure of \textit{PvMSP1}, originally characterized from two monkey-adapted \textit{P. vivax} strains (Belém and Salvador-1), exhibits conserved, semi-conserved and polymorphic regions [11,13]. This gene has a similar genetic organization to that of \textit{PfMSP1} and \textit{P. yoelii} MSP1, with several regions highly conserved among species (termed ‘interspecies conserved blocks’ (ICBs)). Sequence analysis of the region between ICB5 and ICB6 from parasite isolates from Brazil [14], Sri Lanka [15], Colombia [16], Papua New Guinea (PNG) [17] and Thailand [18] demonstrates the dimorphic nature of this region and suggests intragenic recombination as a mechanism of generating antigenic diversity. Other regions (between ICB2 and ICB4, ICB4 and ICB5, ICB8 and ICB10) also exhibit varying degrees of polymorphism [19,20]. Although most studies with \textit{PvMSP1} show evidence of genetic recombination of different alleles, and some even detect geographical clusters of specific variations [19], sample sizes are too small to allow the accurate determination of allele frequencies.

In addition to the length polymorphism of the PCR product, extensive polymorphism at the nucleotide level has been observed by PCR-RFLP in the \textit{PvMSP3a} gene in \textit{P. vivax} isolates from diverse geographical regions [21,22]. The polymorphism and deletions are largely restricted to the central domain containing coiled-coil heptad repeats [23,24]. In malaria-hyperendemic Papua New Guinea, 24 alleles with varying frequencies are observed from 74 samples from asymptomatic patients [22].

\textbf{AMA1}

AMA1, a protein essential for erythrocyte invasion, is highly conserved among \textit{Plasmodium} spp. Similar to \textit{PfAMA1}, \textit{PvAMA1} shows limited sequence polymorphism [25,26]. Moreover, \textit{PvAMA1} might have only a few predominant haplotypes and display very limited genetic diversity within a geographic region, although a more comprehensive study of the whole gene is necessary to confirm this observation.

\textbf{CSP}

CSP genes have been cloned from numerous \textit{Plasmodium} spp, including \textit{P. vivax} [27]. Being one of the prime targets for anti-infection vaccines, CSP has been studied extensively in terms of antigenicity and polymorphism. In all \textit{Plasmodium} spp., it is a single-copy gene encoding structurally similar CSP, where a central domain of tandemly repeated sequences is flanked by predominantly nonrepeated conserved sequences. The central repetitive domain varies in sequence and length among \textit{Plasmodium} spp. Analysis of PVcSP sequences revealed that parasites have repeats belonging to one of two types of nonapeptide repeat units (GDRAA/DGPQA or ANGAGNQPG), named respectively VK210 or VK247, for the sample number from which they were originally described [28]. Both variants have a worldwide distribution [29]. Analysis of the \textit{PvCSP} sequences from southeast Asian isolates indicates very limited polymorphisms outside of the repeat region [30]. Currently, \textit{PvCSP} has been employed successfully in
epidemiological studies of \textit{P. vivax} malaria using two well-established methods: molecular hybridization analysis of parasite DNA isolated from blood samples [31] and ELISA to detect sporozoite proteins in mosquitoes artificially infected or collected from the field [32,33].

Other genes
The \textit{PvGAM1} gene, named because of its expression in gametocytes, displays polymorphic deletions near the 3' end in a limited number of Sri Lankan isolates [34]. Its potential as a molecular marker for genotyping is compromised due to artefacts associated with amplification of this region [35]. The reason for the artefacts is not clear; the presence of GC-rich regions, especially with consecutive runs of G, close to the site of the deletion might be implicated.

In addition to point mutations that are associated with resistance to pyrimethamine, the \textit{P. vivax} DHFR gene also has a size polymorphism resulting from the deletion of five or six amino acids. Four allelic variants have been observed to date [36,37]. 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.

Population structure of \textit{P. vivax}
Given that meiotic recombination takes place in the vector as part of the \textit{Plasmodium} life cycle, the rate of genetic recombination and the effective population size are both expected to be linked to the transmission intensity. In agreement, studies carried out on \textit{P. falciparum} have shown high genetic diversity and failed to detect multilocus associations in Africa with a high transmission rate, and have found low genetic diversity and strong linkage disequilibria in areas with low transmission [38,39].

There are no studies on \textit{P. vivax} using this type of approach, but the available evidence suggests that the situation could be similar. In a hyperendemic area such as Papua New Guinea, intragenic recombination and high genetic diversity have been reported in genes encoding malarial vaccine antigens such as MSP1 and MSP3α [17,22]. Surprisingly, even in hypoendemic areas such as Thailand and Brazil, \textit{PcMSP1}, \textit{PcMSP3α} and \textit{P. vivax} thrombospondin-related adhesive protein (\textit{PvTRAP}) also display high levels of diversity and evidence of genetic recombination [24,40]. By contrast, relatively low genetic diversity has been detected in the re-emerging \textit{P. vivax} malaria focus in Korea [41], probably owing to the epidemic nature where few strains might be involved. For \textit{PvCSP}, parasite isolates appear to cluster geographically; for example, a 36-base insertion 3' to the central repeats is found only in parasite samples from China and the Korean Peninsula [42,43]. Unfortunately, since the information available is derived from antigens, geographical differentiation caused by positive natural immune selection rather than reduced gene flow cannot be ruled out. Recently, studies on the susceptibility of mosquitoes to infection have provided grounds to speculate about the genetically varying frequencies of the two \textit{PvCSP} repeat variants. Laboratory-bred \textit{Anopheles albimanus} and \textit{Anopheles pseudopunctipennis} from southern Mexico were found to be differentially susceptible to infections by VK210- and VK247-bearing \textit{P. vivax} [44], and the prevalence of the two parasite types correlated with the distribution of the two mosquito vectors [45]. It remains to be determined whether ookinete destruction and oocyst development arrest in the susceptible vector [46] or a polymorphic gene genetically closely linked to \textit{PvCSP} accounts for these observations.

Multiple clonal infections (i.e. the simultaneous infection of a host by more than one strain of the same parasite species) are common in malaria. For \textit{P. falciparum}, the levels of multiplicity of infection are partially correlated with the levels of transmission intensities. For \textit{P. vivax} malaria, the proportion of mixed-strain infections estimated in Papua New Guinea, India and Thailand ranges from ~30% to 65% [17,22,24,47]. In Papua New Guinea, as many as six parasite clones have been found in a single host [22]. Although not directly comparable, since different sampling strategies and genotyping methods were used, these values are strikingly high in regard to the low endemicity of \textit{P. vivax} malaria in certain regions such as Thailand. This could be because of biological features of the \textit{P. vivax} parasite, such as earlier gametocytogenesis and relapse. The production of gametocytes during the presymptomatic period might allow for a more efficient transmission to the mosquito vector before drug treatment is initiated, while regular relapses would also enhance transmission and increase the probability of detecting mixed infections as further inoculations occur. Consequently, the multiplicity of infections is likely to facilitate genetic recombination of parasites and the generation of novel strains. Similar to \textit{P. falciparum} [48], the natural populations of \textit{P. vivax} also appear to have random mating as demonstrated using the two \textit{PvCSP} repeat types [33]. This might also contribute to the high levels of genetic diversity of \textit{P. vivax} populations in low-endemic areas [24].

Molecular evolution of antigens
For malarial vaccine antigens, the observed high levels of genetic diversity have been attributed to positive natural selection. This interpretation is supported, among other lines of evidence, by the observation that nonsynonymous nucleotide substitutions are more common than synonymous substitutions [49,50]. Synonymous substitutions are likely to be neutral, or nearly so, whereas nonsynonymous substitutions could be functionally constrained and, thus, subject to negative natural selection and are eliminated from the population. It is expected that the genetic polymorphism resulting from random accumulation of mutations will exhibit more synonymous than nonsynonymous substitutions. An excess of nonsynonymous substitutions can be taken as evidence that these mutations might confer an adaptive advantage so they are maintained by natural selection. Under this scenario, the action of the host immune system on the parasite population allows the accumulation and frequent switch of suitable mutations.

Several molecular evolutionary studies have been performed on the \textit{PfAMA1} gene [50–52], where the rate of nonsynonymous substitution is found to be much higher than that of synonymous substitution, which is suggestive of positive diversifying selection probably by...
host immunity. This is particularly evident in the regions that serve as T helper (Th)-cell epitopes [51]. There is also indication of high recombination rates of the molecule, demonstrated by the high recombination parameter value and the rapid decline in linkage disequilibrium with increasing nucleotide distance [52]. In the specific case of \textit{P. vivax}, the available information suggests that there is no such overwhelming pattern of more nonsynonymous than synonymous substitutions as in \textit{P. falciparum}. A plausible explanation is that its genome does not have the confounding factor of a strong codon bias. Specifically, the average effective number of codons, a statistic that measures codon bias, is 54.44 for \textit{P. vivax}, based on 15 genes. This value contrasts with the reported 36.82 for \textit{P. falciparum}. The available evidence shows that there are more synonymous than nonsynonymous substitutions in partial sequences of genes such as \textit{PvAMA1} and \textit{PvMSP1} [26], and evidence for positive natural selection has been found in \textit{PvTRAP} [40]. This might also be true for \textit{PvCSP}, where amino acid substitutions are mostly localized in the regions that are probably under strong immune selection (e.g. regions that harbor T-cell determinants) [30].

It has been suggested that synonymous substitutions might not be neutral in the case of \textit{PvAMA1}, in that its rate is lower than the one observed on \textit{PvMSP1} [26]. The interpretation of this result demands a further clarification of what neutrality means. The neutral theory is compatible with the widely observed phenomenon that genes evolve at different rates. The rate of substitutions is affected by the overall forces of selection (negative or positive) acting at that locus and closely linked loci, the population history and its effective population size, among other factors. These arguments support the comparison among alleles at a given locus of the synonymous and nonsynonymous substitution rates. The observation that the rate of synonymous substitutions in one part of a gene is lower when compared with partial sequences from other genes does not imply that the silent substitutions are more or less neutral; indeed, the available theory explains such patterns. However, we agree that the effect of factors such as local codon bias needs to be explored. As we stated earlier, although the \textit{P. vivax} genome is not particularly biased, the effective number of codons for \textit{PvAMA1} is 44.3, which is lower than observed for other merozoite genes such as \textit{PvMSP1} (average 55.04 with 61 for MSP1). Studies using complete sequences are needed to address these issues. It is worth noticing that the comparison of synonymous and nonsynonymous substitutions is not the only approach for studying the effect of selection. There are several methods available that have been used mostly in \textit{P. falciparum} [50,53,54]. For \textit{P. vivax}, there is an advantage over \textit{P. falciparum} given that comparative studies are possible due to the availability of isolates from closely related species (e.g. \textit{P. simium}) [55,56].

**Genetic diversity and malaria control**

The efficacy of control measures is generally assessed on their impact on the number of infected individuals/-mosquitoes, clinical cases and treatment failures, and so on. However, it is possible to envisage situations where intervention only marginally affects such measurable outcomes over the period of observation. Parasite genetic diversity could then be exploited to detect subtle changes in the population of the pathogen.

Changes in the overall genetic composition of the population or the multiplicity of infection observed in individuals can provide an indication of an effect resulting from an intervention such as the deployment of bednets or vaccination. An increasing number of \textit{P. vivax} antigens are now under development as vaccine candidates: CSP, MSP1, AMA1, Pvs25, Pvs28, TRAP and the Duffy binding protein (DBP). Many of these antigens are polymorphic and the induced immune responses might be more effective against parasites carrying particular allelic variants. Determination of allelic frequencies should be included in future field-based trials involving vaccination with polymorphic antigens.

The emergence and spread of drug resistance are major problems in the control of malaria. Early detection and monitoring of resistant \textit{P. vivax} parasites is needed to avoid the disastrous humanitarian problems that have resulted from the spread of drug-resistant \textit{P. falciparum} parasites. Direct monitoring requires knowledge of the genetic basis of resistance. The enzyme targets of anti-folate drugs pyrimethamine and sulfadoxine are DHFR and dihydropteroate synthase (DHPS). Rapid molecular monitoring of anti-folate resistance in \textit{P. falciparum} has been made possible with the determination of the point mutations associated with such resistance. Recent analysis of \textit{pvdhfr} has brought to light five residues putatively associated with resistance [37]. Protocols for the detection of these mutations in field samples have been elaborated. However, a similar molecular approach for monitoring resistance to chloroquine or primaquine in \textit{P. vivax} awaits the discovery of the target proteins of these drugs. Since \textit{in vitro} culture of \textit{P. vivax} is not easily available, monitoring of drug resistance relies exclusively on \textit{in vivo} drug efficacy trials. When conducted in an endemic area, as is frequently the case, these studies suffer from a major limitation: the inability to distinguish true recrudescences (which denotes a treatment failure) from a re-infection during the follow-up period. This has been largely solved by the development of PCR-based practical protocols for parasite genotyping [57]. Standard protocols, similar to those developed for \textit{P. falciparum} and based on the \textit{PvCSP} and \textit{PvMSP1} genes, have been developed (M. Imwong et al. unpublished) and should be included in future drug efficacy trials.

**Perspective**

The molecular analysis of parasite populations is a relatively recent research tool. Although these studies have so far been limited to a few polymorphic genes, several important insights in the biology and immunology of the parasite have accrued. With the advent of relatively cheap and rapid sequencing of whole genomes, as well as the development of high-throughput post-genomic analyses, the forthcoming completion of the \textit{P. vivax} genome should substantially increase the potential for discovery. Genome-wide comparisons among the \textit{Plasmodium} spp. might then provide a clearer indication of the phylogenetic position of \textit{P. vivax}. Wholesale comparison of the genomes
of different P. vivax lines should provide a broader array of genetic markers, including microsatellite and noncoding DNA stretches. In this context, a challenge for the future is the development of strategies to analyze on a population level the parasite multigene families that are associated with pathogenicity and virulence, for example the var family of P. falciparum, the py235 family of P. yoelii and the recently described vir genes of P. vivax [58]. Ultimately, in the absence of acceptable experimental models of the infection, parasite genotyping strategies are crucial for deriving knowledge of parasite survival strategies from observations of natural infections.

References
23 Rayner, J.C. et al. (2002) Extensive polymorphism in the Plasmodium vivax merozoite surface coat protein MSP-3a is limited to specific domains. Parasitology 125, 393–405
42 Mann, V.H. et al. (1994) Sequence variation in the circumsporozoite protein gene of Plasmodium vivax appears to be regionally biased. Mol. Biochem. Parasitol. 68, 45–52

http://parasites.trends.com

46 Gonzalez-Ceron, L. et al. (2001) Plasmodium vivax: ookinet destruction and oocyst development arrest are responsible for Anopheles albimanus resistance to circumsporozoite phenotype VK247 parasites. Exp. Parasitol. 98, 152–161


**Endeavour**

the quarterly magazine for the history and philosophy of science

Online access to *Endeavour* is FREE to BioMedNet subscribers, providing you with a collection of beautifully illustrated articles on the history of science, book reviews and editorial comment.

featuring

The pathway to the cell and its organelles: one hundred years of the GoGi apparatus by M. Bentivoglio and P. Mazzarello

Joseph Fourier, the ‘greenhouse effect’ and the quest for a universal theory of terrestrial temperatures by J.R. Fleming

The hunt for red elixir: an early collaboration between fellows of the Royal Society by D.R. Dickson

Art as science: scientific illustration 1490–1670 in drawing, woodcut and copper plate by C.M. Pyle

The history of reductionism versus holistic approaches to scientific research by H. Andersen

Reading and writing the Book of Nature: Jan Swammerdam (1637–1680) by M. Cobb

The role of museums in history of science, technology and medicine by L. Taub

The ‘internal clocks’ of circadian and interval timing by S. Hinton and W.H. Meck

The troubled past and uncertain future of group selectionism by T. Shanahan

A botanist for a continent: Ferdinand Von Mueller (1825–1896) by R.W. Home

Rudolf Virchow and the scientific approach to medicine by L. Benaroyo

Darwinism and atheism: different sides of the same coin? by M. Ruse

Alfred Russel Wallace and the flat earth controversy by C. Garwood

John Dalton: the world’s first stereochemist by Dennis H. Rouvray

Forensic chemistry in 19th-century Britain by N.G. Coley

Owen and Huxley: unfinished business by C.U.M. Smith

Characteristics of scientific revolutions by H. Andersen

and much, much more . . .


Log on to http://reviews.bmn.com, hit the ‘Browse Journals’ tab and scroll down to *Endeavour*