

RESEARCH

Open Access

Genetic diversity and population structure of genes encoding vaccine candidate antigens of *Plasmodium vivax*

Stella M Chenet^{1,2,3}, Lorena L Tapia¹, Ananias A Escalante^{2,3}, Salomon Durand¹, Carmen Lucas¹ and David J Bacon^{1,4*}

Abstract

Background: A major concern in malaria vaccine development is genetic polymorphisms typically observed among *Plasmodium* isolates in different geographical areas across the world. Highly polymorphic regions have been observed in *Plasmodium falciparum* and *Plasmodium vivax* antigenic surface proteins such as Circumsporozoite protein (CSP), Duffy-binding protein (DBP), Merozoite surface protein-1 (MSP-1), Apical membrane antigen-1 (AMA-1) and Thrombospondin related anonymous protein (TRAP).

Methods: Genetic variability was assessed in important polymorphic regions of various vaccine candidate antigens in *P. vivax* among 106 isolates from the Amazon Region of Loreto, Peru. In addition, genetic diversity determined in Peruvian isolates was compared to population studies from various geographical locations worldwide.

Results: The structured diversity found in *P. vivax* populations did not show a geographic pattern and haplotypes from all gene candidates were distributed worldwide. In addition, evidence of balancing selection was found in polymorphic regions of the *trap*, *dbp* and *ama-1* genes.

Conclusions: It is important to have a good representation of the haplotypes circulating worldwide when implementing a vaccine, regardless of the geographic region of deployment since selective pressure plays an important role in structuring antigen diversity.

Keywords: Malaria, *Plasmodium vivax*, Vaccine candidates, Haplotypes

Background

Malaria is one of the major global public health problems that affect most tropical regions of the world. Even though *Plasmodium falciparum* is the most virulent, it is estimated that *Plasmodium vivax* produces around 80 to 300 million clinical cases per year [1]. Furthermore, there have been several reports of severe *P. vivax* malaria cases in the last few years [2-4]. In 2008, 560,221 malaria cases were reported in the Americas [5]; 74.2% of them caused by *P. vivax* and 25.7% by *P. falciparum* [1]. About 90% of these malaria cases originated in the Amazon basin shared by Bolivia, Brazil, Colombia, Ecuador, French Guiana, Guyana, Venezuela,

Suriname and Peru [5], whereas the other 10% was contributed by non-Amazon regions.

Developing a vaccine for *P. vivax* represents a major challenge especially considering the lack of *in vitro* cultures. Thus, current efforts focus on orthologs of *P. falciparum*. Over the past four decades, experiments performed in animals and human subjects have led to the development of several *Plasmodium* vaccine candidates. Antigenic surface proteins such as the Circumsporozoite protein (CSP), Thrombospondin related anonymous protein (TRAP), Duffy-binding protein (DBP), Merozoite surface protein-1 (MSP-1) and Apical membrane antigen-1 (AMA-1) are currently being evaluated as vaccine candidates in clinical trials [1,6-8]. The CSP is one of the most extensively studied antigens and it is involved in the motility and invasion of the sporozoite during its entrance in the hepatocyte [9]. *PvCSP*

* Correspondence: david.bacon@nrl.navy.mil

¹Parasitology Program, Naval Medical Research Unit No. 6, Lima, Peru
Full list of author information is available at the end of the article

has an immunogenic central repeat domain flanked by amino and carboxyl sequences containing highly conserved protein stretches (Regions I and II-plus) [10]. This protein displays two major types of nonapeptide repeats: type I (VK210), which contains repeats of GDRA(A/D)GQPA or related sequences and type II (VK247), which is composed of ANGAGNQP repeats [11,12]. A third type of variant, called *P. vivax*-like has repeat units of APGANQ(E/G)GAA, identical to that described for *Plasmodium simiovale* [13,14]. Currently, the *P. falciparum* RTS, S vaccine, which is based on CSP have showed to protect semi-immune adults [15] and children from *P. falciparum* natural infection in endemic areas [16-18]; this vaccine is being tested in Phase III clinical trials [8,19].

Another sporozoite surface protein that stands out as a good vaccine candidate is TRAP, which is a 90 kDa protein with six distinct regions. TRAP is essential for sporozoite gliding and hepatocyte invasion [20,21]. Genetic analysis of *Pv*TRAP reveals that most of the polymorphisms occur in Region II (von Willebrand factor A-domain) and III [22]. The potential role of TRAP as a malaria vaccine candidate has been tested in mice and monkeys, inducing high levels of specific antibodies [23]. Furthermore, a synthetic peptide-derived from a conserved region of the TRAP protein located in the N-terminal, proved to be immunogenic and provided partial protection in *Aotus* monkeys [1,24].

On the other hand, the asexual *P. vivax* parasite antigens DBP, MSP-1 and AMA-1 have also received attention for their role in the invasion process and/or their generation of antibody response. DBP interacts with the Duffy blood group antigen of reticulocytes, MSP-1 is expressed abundantly on the merozoite surface and AMA-1 plays a role in apical reorientation of merozoites following initial attachment during invasion [1]. Region II of DBP (DBP_{II}) functions as a ligand domain and consists of 330 aa with 12 cysteines and numerous aromatic residues. Studies have identified the central region of the DBP_{II} domain between cysteines 4 and 8, as the segment containing the conserved contact residues Tyr 94, Asn 95, Lys 96, Arg 103 and Ile 175 required for recognition of DARC (Duffy Antigen/Receptor for Chemokine) on human erythrocytes [25-28].

MSP-1 is part of a complex of proteins that is processed into smaller fragments by serine proteases during invasion of red blood cells [29]. It has been suggested that fragments *Pv*MSP-1₁₄ and *Pv*MSP-1₂₀ could be included in a *P. vivax* multi-antigen vaccine since they contain high affinity reticulocyte binding (HARB) cluster regions and confer protection in monkeys [30-32]. AMA-1 has also been evaluated for inclusion in a multi-sub-unit vaccine for both *P. falciparum* and *P. vivax* [33,34]. The complete AMA-1 protein contains three

domains, from which a higher rate of mutations and level of diversifying selection has been shown in domain I [35,36]. Indeed, AMA-1 is highly immunogenic, protecting against *Plasmodium chabaudi* infection in mice [37] and displaying significant antibody and T cell responses in endemic human populations [38-40].

Investigation of the sequence variation among malaria antigenic regions is necessary for the development of an effective malaria vaccine that includes haplotypes from different regions, considering that *Plasmodium* is highly diverse and these regions are under strong immunological pressure. In this study, important regions in *Pv*CSP, *Pv*TRAP, *Pv*DBP, *Pv*MSP-1 and *Pv*AMA-1 were genotyped using *P. vivax* Peruvian isolates collected during years 2006-2007. In addition, since natural selection plays a role in structuring diversity of these highly polymorphic antigens, the worldwide distribution of the main variants was determined by comparing the Peruvian samples with sequences previously reported from South America, Asia and Oceania.

Methods

Study area and sample collection

Blood samples from patients were collected during years 2006-2007 from endemic areas in the Peruvian Amazon region of Loreto Department using human use approved protocols (NMRCD.2005.0005). In Peru, *P. vivax* corresponds to more than 80% of the malaria cases and the majority of them are found in the Amazonian Department of Loreto [41]. Fifty-seven samples were obtained from Padrecocha, in Punchana district, a village situated five km from Iquitos (Loreto's capital) and forty-nine samples were collected from a village located in the San Juan district just south of Iquitos. All samples were confirmed to be positive for *P. vivax* by microscopic analysis on Giemsa-stained blood smears.

PCR and DNA sequencing

DNA was purified from 200 µL of whole blood using the QIAamp DNA Blood mini kit (QIAGEN, Valencia, CA). PCR amplification was performed in a 50 µL reaction mixture containing 0.2 µM each of forward and reverse primers, 200 µM each dNTP, 0.6 units of DNA polymerase recombinant (Invitrogen, Carlsbad, CA), 3 µL of 10× PCR buffer, 1.5 mM of MgCl₂ and 5 µL of DNA extracted from blood. To assess antigenic variability of *Pv*CSP, the Central Repetitive Region, Region I and Region II were amplified. The PCR conditions were: 5 min at 94°C; 1 min at 94°C, 1 min at 58°C, 2 min at 72°C for 35 cycles; and a final extension of 72°C for 2 min, using the primers VSC-OF (5'ATGTA-GATCTGTCCAAGGCCATAAA3') and VSC-OR (5'TAATTGAATAATGCTAGGACTAACAATATG 3') [42]. The amplification of the highly polymorphic

regions II and III of *PvTRAP* [22] was obtained using the primers *Pvssp*: 5' GTCGATTTATACCTCC-TAGTTGACGG 3' and *Pvsspr*: 5' CAGCTTGAACGTC-GACCTCCGC 3' with the following conditions: 5 min at 94°C; 1 min at 94°C, 1 min at 50°C, 2 min at 72°C for 35 cycles; and a final extension of 72°C for 2 min. The amplification of the central region of the *PvDBP_{II}* domain was done using primers *Liz-1* (5' TGGGGAG-GAAAAAGATG 3') and *Liz-2* (5' AACGGAA-CAAACGCATAAC 3') [43]. The PCR conditions used were as follows: 4 min at 94°C; 1 min at 94°C, 2 min at 58°C, 2 min at 72°C for 35 cycles; and a final extension of 72°C for 3 min. *PvMSP-1* primers described by Espinosa et al. (*HARB1_f*: 5' GGCAACAACGATGACGAC 3', *HARB1_r*: 5' CTTCTCAAGTCGAGCAG 3', *HARB2_f*: 5' TCACTGTAAAGAAATTGCAG 3' and *HARB2_r*: 5' ATACATGTGTGCTCGGAG 3') were used for amplifying the two HARB containing regions: *PvMSP-1₂₀* and *PvMSP-1₁₄* kDa [31] with the following conditions: 5 min at 94°C; 1 min at 94°C, 1 min at 50°C, 1 min at 72°C for 35 cycles; and a final extension of 72°C for 5 min for the first region and 5 min at 94°C; 1 min at 94°C, 1 min at 45°C, 1 min at 72°C for 35 cycles; and a final extension of 72°C for 5 min for the second region. Finally, a nested PCR was used to amplify domain I of *PvAMA-1* as described by Cheng and Saul [44]. PCR products were purified using the QIAquick PCR purifications kit (QIAGEN, Valencia, CA, USA) and sequenced using BigDye[®] Terminator v3.1 sequencing kit (Applied Biosystems, Foster City, CA, USA), with the primers previously mentioned and an ABI 3100 automated sequencer.

Sequence determination and data analysis

Sequences were analysed using the Sequencher[™] software version 4.7 (Gene Codes Corporation, Ann Harbor, MI, USA) and sequence alignments were performed using the MEGA software version 4.1 [45]. Comparison with previously deposited *P. vivax* sequences in the GenBank was performed using BLAST search. The number of haplotypes (h), haplotype diversity (hd) and nucleotide diversity (Pi) values were obtained using the DnaSP version 5.0 [46]. The effect of natural selection was also evaluated by estimating the difference of non-synonymous and synonymous substitutions (Dn-Ds) using the Nei and Gojobori's method with the Jukes and Cantor correction. The null hypothesis of neutrality was rejected if Dn:Ds \neq 1. Positive values for Dn-Ds indicated positive selection while negative values indicated negative selection [47]. Standard error was calculated using bootstrap with 1,000 replications and rates of Ds and Dn were compared by the Z-test of selection using Mega version 4.1 [45]. To test whether the haplotypes clustered according to their geographic origin, the

model-based clustering algorithm implemented in the Structure 2.1 software was used [48]. This software uses a Bayesian clustering approach to assign isolates to *K* populations characterized by a set of allele frequencies at each locus. The number of such populations may be either previously known or not. The program was run five times at different *K* values (from 2 to 10) with a burn-in period of 10,000 iterations followed by 50,000 iterations. The admixture model was used in all analyses, which allows for the presence of individuals with ancestry in two or more of the *K* populations. However, clustering algorithms incorporate stochastic simulation as part of the inference and, as a consequence, independent analyses of the same data may result in several distinct outcomes, even though the same initial conditions were used. To facilitate the interpretation of population-genetic clustering results, the computer program CLUMPP (Cluster Matching and Permutation Program) was used. CLUMPP strips away the 'label switching' heterogeneity so that the 'genuine multimodality' can be detected and quantified [49]. In addition, *distruct* 1.1 was used to graphically display the clustering results [50]. The genetic difference between clusters was measured by the fixation index (Fst) calculated from the haplotype frequencies (Hst) and pair-wise DNA sequence diversity (Kst*) using DnaSP version 5.0 [46]. The nearest-neighbour statistic (Snn) was also calculated between geographic populations. Snn is a measure of how often the "nearest neighbours" of sequences are from the same locality in geographic space. This statistic is applicable when genetic data are collected on individuals sampled from two or more localities. Snn is expected to be near one when the populations at the two localities are highly differentiated and near one-half when the populations at the two localities are part of the same pan-mictic population [51].

Results

PvCSP

A 1,100 bp PCR product corresponding to *Pvcsp* was amplified from 106 *P. vivax* samples from Peru. In addition, 326 complete CSP sequences from Brazil, Colombia, India, Iran and Korea were also included in the amino acid alignment. Sequences from Thailand, China and PNG were used to determine the percentage of strains circulating of each *PvCSP* type (Table 1); however, they were not considered in the haplotype analysis since the sequences were not complete. Most of the strains circulating worldwide belonged to the VK210 type and in most population sets analysed (except Colombia and Thailand) a greater percentage of VK210 haplotypes was found. The amino acid alignment of the sequences used in this study showed 97 different VK210 haplotypes and 28 VK247 haplotypes circulating

contained the repeated region ANGA(G/D)(D/N)QPG, 17 or 18 times. Region I of *PvCSP* showed a polymorphism in position 3 (L/V), whereas the 18 amino acid long linear peptide EWTPCSVTCGVGVRVRRR, located on Region II [53,54], was highly conserved. *PvCSP* analysis of Peruvian isolates also revealed high variability of VK210 and VK247 haplotypes compared to other populations from South America [55,56]. Region II-plus, which contains dominant B-cell epitopes [57,1] and block hepatocyte invasion by sporozoites [58], was highly conserved in Peruvian isolates, with the same amino acid sequence as the Sal I strain.

PvTRAP

A total of 104 sequences of regions II and III of *Pvtrap* were obtained from the Peruvian isolates. In addition, 63 Thailand sequences from GenBank were included in the amino acid alignment. The N-terminal region (209-256aa), which has been proved to be immunogenic and to provide partial protection in *Aotus* monkeys [24] was highly conserved with only three amino acid substitutions in positions 216 (R/K), 219 (H/Q) and 229 (G/K). In general, more diversity was found in Thailand ($h = 34$, $\pi = 0.007 \pm 0.00029$ and $Hd = 0.961 \pm 0.01$) (Table 2) compared to Peru ($h = 7$, $\pi = 0.00235 \pm 0.00012$ and $Hd = 0.641 \pm 0.026$) and according to the Dn-Ds test, regions II and III of *Pvtrap* were under positive selection (Table 3).

PvDBP

A 1,150 bp DNA fragment of *Pvdbp* was amplified from all 106 Peruvian isolates and 601 *P. vivax* sequences from Brazil, Colombia, India, Sri Lanka, Thailand, Korea, Iran and Papua New Guinea were also included in the analysis. The alignment of 510 bp from region II of *dbp* showed 150 different haplotypes, from which 39 were found in South America, 42 in Asia, 56 in Oceania and 13 were found in South America and Asia. In the Peruvian isolates 26 different haplotypes were found with genetic diversity values of $\pi = 0.008 \pm 0.00048$ and $Hd = 0.903 \pm 0.014$. In general, the average genetic and haplotype diversity in South America and Asia was very similar (Table 2). On the other hand, the analysis of DNA sequence polymorphisms showed a significant difference between synonymous and non-synonymous substitutions, indicating balancing selection in *dbp_{II}* (Table 3). In addition, several substitutions in the positions underlined were found in DBPII regions corresponding to peptides H1: FHRDITFRKLYLKRKL, H3: DEKA QQRRKQW/WNESK and 1653: RDYVSELPTVEVQKLKE KC, previously reported as immunogenic [27,59].

PvMSP-1

According to the amino acid alignment of *PvMSP-1₂₀* and *PvMSP-1₁₄*, a total of 28 haplotypes were found

worldwide, 11 in South America, 14 in Asia and three in both continents. When comparing genetic diversity between populations, Peru ($\pi = 0.00286 \pm 0.00029$, $Hd = 0.665 \pm 0.041$) and Brazil ($\pi = 0.00293 \pm 0.00061$, $Hd = 0.75 \pm 0.139$) showed the highest levels of nucleotide and haplotype diversity, while in Korea, Thailand and Turkey these HARBs were highly conserved. In general, values of genetic diversity for *Pvmsp-1* were low when compared with the other genes and no clear signature of selection was found according to the Dn-Ds test (Table 3).

PvAMA-1

The polymorphic region of *Pvama-1* (Domain I of protein) was amplified and sequenced from 98 of the 106 Peruvian isolates. A total of 774 *Pvama-1* partial sequences were used in the analysis, including samples from Brazil, Venezuela, India, Sri Lanka, China, Korea, Myanmar, Philippines, Thailand, Papua New Guinea and Solomon Islands. A total of 168 haplotypes were identified; 135 were exclusively found in Asia, nine in Oceania and 12 in South America. In addition, seven haplotypes were shared in America and Asia and five in Asia and Oceania. According to the analysis of polymorphisms, the dN-dS test indicated balancing selection in this domain (Table 3) with several haplotypes in low frequency. In addition, values of nucleotide diversity were very similar between geographic regions and populations (Table 2).

Population genetic analysis

When analysing *Plasmodium* isolates, samples are usually grouped by geographic location; nevertheless, this approach assumes that geographic barriers or distance are the underlying factors that limit gene flow. It also assumes that genetic exchange within each population geographically defined is random. To better estimate the distribution of haplotypes of different *P. vivax* vaccine candidate antigens, a Bayesian clustering algorithm [48] was used to sub-group samples according to related SNPs haplotypes. The Structure model assumes that loci are independent within populations; however, this assumption is likely to be violated for sequence data, or data from non-recombining regions. It is a matter of debate whether it is valid to group sequences containing potentially linked polymorphisms in a single gene under balancing selection [60]. The main cost of the dependence within regions will be that Structure underestimates the uncertainty in the assignment of particular individuals. However, if there is enough independence across regions than LD within regions, then Structure may actually perform reasonably well [48]. Moreover, our main goal is not to describe the demographic

Table 2 Estimates of *Plasmodium vivax* genetic diversity using GenBank sequences

Antigen	Region	Country	N	GenBank Accession Number	H	π (s.d.)	Hd (s.d.)
<i>TRAP</i>	South East Asia	Thailand	63	AF312559-AF312661	34	0.00707 (0.00029)	0.961 (0.01)
<i>DBPII</i>	South America	Brazil	126	DQ156520, EU812839-EU812960, EU870443-EU870445	24	0.00919 (0.00042)	0.883 (0.018)
		Colombia	17	DQ156513, U50575-U50590	12	0.00812 (0.00081)	0.956 (0.033)
	South Asia	India	101	DQ156516, FJ491142-FJ491241	28	0.01011 (0.00063)	0.86 (0.028)
		Sri Lanka	100	GU143914-GU144013	14	0.00864 (0.00063)	0.738 (0.039)
	South East Asia	Thailand	30	EF219451, EF368159-EF368178, EF379127-EF379135	19	0.01628 (0.00354)	0.972 (0.018)
	East Asia	Korea	16	EU395592, DQ156515, DQ156522-DQ156523, AF215737-AF215738, AF220657, AF220659-AF220667	11	0.00817 (0.00149)	0.933 (0.048)
	Western Asia	Iran	11	EU860428-EU860438	8	0.01047 (0.00205)	0.945 (0.054)
	Oceania	Papua New Guinea	200	AF291096, AF289653, AF469515-AF469602, DQ156512-DQ156523, AY970837-AY970925	59	0.01166 (0.00041)	0.888 (0.015)
<i>MSP-1</i>	South America	Brazil	8	AF435622-AF435625, AF435627, AF435629-AF435631	4	0.00293 (0.00061)	0.75 (0.139)
	South East Asia	India	28	EU430452-EU430479	6	0.0172 (0.00049)	0.489 (0.112)
		Thailand	170	GQ890872-GQ891041	2	0.00013 (0.00005)	0.068 (0.026)
		Singapore	50	GU971656-GU971705	10	0.0019 (0.00032)	0.407 (0.09)
	Western Asia	Turkey	30	AB564559-AB564588	2	0.00073 (0.00034)	0.186 (0.088)
	East Asia	Korea	8	HQ171941-HQ171934	1	0	0
<i>AMA-1</i>	South America	Brazil	65	EF031154-EF031216, EF031183, EF031184	12	0.01358 (0.00056)	0.866 (0.02)
		Venezuela	73	EU346015-EU346087	12	0.01368 (0.00053)	0.847 (0.019)
	South Asia	India	189	EU282774-EU282822, FJ490963-FJ491062, EF025187-EF025197, GU086369-GU086382, AF171250-AF171254, AF171256, AF171258-AF171260, AF171262-AF171267	68	0.01427 (0.00036)	0.945 (0.008)
		Sri Lanka	23	EF218679-EF218701	8	0.00722 (0.00148)	0.83 (0.054)
	East Asia	China	7	AF171243-AF171249	6	0.01211 (0.00193)	0.893 (0.111)
		Korea	6	AF357212, AF357213, GU476488, EU395599, AF402593, AF402594	4	0.00812 (0.00328)	0.9 (0.161)
	South East Asia	Myanmar	39	FJ157247-FJ157285	36	0.01833 (0.00092)	0.996 (0.007)
		Philippines	110	AF171273-AF171382	20	0.01474 (0.0005)	0.802 (0.029)
		Thailand	235	FJ784891-FJ85121, AF17453-AF171456	44	0.01488 (0.00039)	0.899 (0.011)
	Oceania	PNG	22	AF171421-AF171426, AF171428-AF171443	13	0.01379 (0.00163)	0.913 (0.045)
Solomon Islands		5	AF171444, AF171445, AF171447, AF171449, AF171450	5	0.01855 (0.00318)	1 (0.126)	

history of the population but to get a graphically representation of the vaccine candidate's haplotype distribution. Genetic differentiation between clusters was calculated using F_{st} from haplotype-frequencies and pairwise DNA sequence diversity. The haplotype frequency-based statistics are more sensitive for small sample sizes, while the sequence-based statistic is a more sensitive method for detecting population structure in highly polymorphic loci [61,62].

The results obtained with *PvTRAP*, *PvDBP* and *PvAMA-1* showed three different clusters. Significant differentiation was observed in haplotype and pair-wise comparisons between all clusters according to the F_{st} values (Table 4), except for *Pvmsp-1*, from which a significant number of clusters it was not possible to identify. Although admixed haplotypes were observed in *Pvdbp* and *Pvama-1*, the estimate of $K = 3$ best explained the distribution of haplotypes from all

Table 3 Summary of genetic data of the *Plasmodium vivax* vaccine antigens analysed including Sal I sequence.

Antigen	Expression	Domain	Sites	N	H	π (s.d.)	Hd (s.d.)	dN-dS (s.d.)	Ztest (dN \neq dS)
<i>Dbp</i>	Merozoite	Region II	510	707	150	0.01103 (0.00025)	0.9319 (0.0053)	0.009 (0.003)	$P < 0.05$
<i>Msp-1</i>	Merozoite	MSP1 ₁₄ and MSP1 ₂₀	512	400*	28	0.00129 (0.00014)	0.358 (0.031)	0.001 (0.001)	n.s.
<i>Ama-1</i>	Merozoite	Domain I	345	872	168	0.01653 (0.00016)	0.9691 (0.0017)	0.012 (0.004)	$P < 0.05$
<i>Trap</i>	Sporozoite	Regions	837	168	38	0.0059 (0.00029)	0.854 (0.018)	0.007 (0.002)	$P < 0.05$

*Belem sequence included

geographic populations; besides, the defined sub-groups were not geographically or even regionally restricted (Figure 2). Snn values between the different geographic populations were also calculated for each of the antigens. We observed that for most of the comparisons in DBP, AMA-1 (Table 5) and TRAP (Snn = 0.86803*), Snn values were near one, which indicates that the localities are highly differentiated. However, close geographic populations such as Sri Lanka and India in South Asia and Peru and Brazil in South America showed almost not genetic differentiation (Snn = 0.58, $p < 0.01$ and Snn = 0.65, $p < 0.01$). Thus, relative low Snn values were found between Brazil, India and Sri Lanka, since there are shared haplotypes circulating in South America and South Asia. Moreover, previous work also indicated low genetic differentiation using Fst analysis between *dbpII* sequences from Brazil and Iran, India or Sri Lanka [63]. On the other hand, at least seven of the pairwise comparisons made using MSP-1 resulted in non-significant values.

Discussion

When designing an efficient anti-malarial vaccine, worldwide information of the circulating antigenic variants is necessary for the formulation of a polyvalent vaccine, which would be effective in different regions. Because of the complex parasite life cycle and the co-existence of *P. falciparum* and *P. vivax* infections in most endemic areas, it is accepted that a multi-antigen

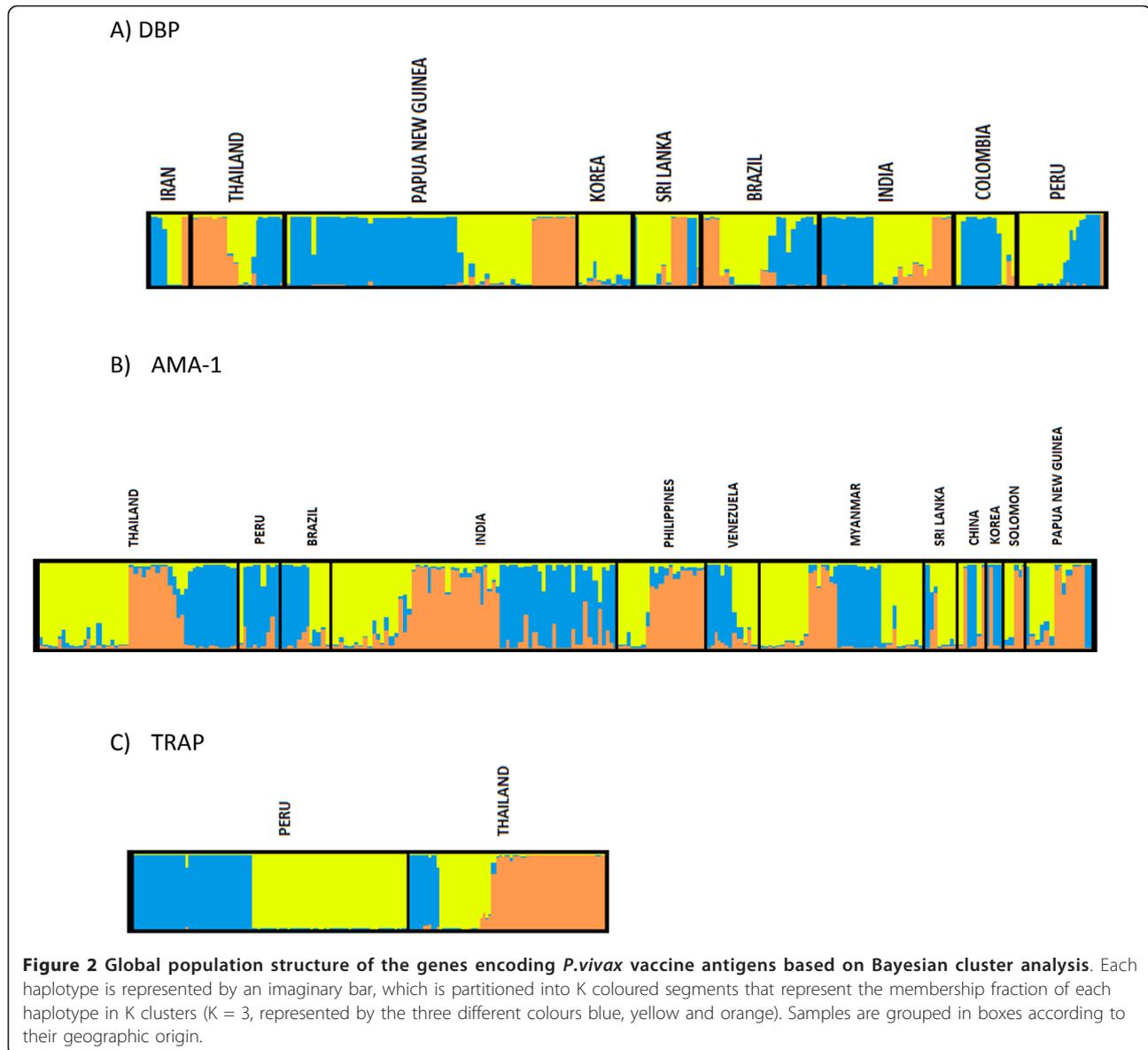
and multi-species vaccine should be developed for most endemic regions [1]. In *P. falciparum*, the majority of haplotypes upon which current vaccines are based have been found to be present at extremely low frequencies in the global parasite population, except for *lsa-1* and *msp-1* haplotypes [62]. Since immune selection might be playing a role in structuring the diversity of highly polymorphic antigens, population genetic studies in endemic regions are essential for the next generation of vaccines. In this study, the information obtained from the Peruvian isolates contribute to get a better understanding of the haplotypes circulating in the Amazon basin and to compare these results with other previously reported from different parts of the world.

For the analysis of Peruvian isolates, samples from Padrecocha and San Juan were considered as a single geographic population since these areas are relatively closed and they present similar haplotype distribution in all antigens with little or none genetic differentiation between them. In general, data from a two-year period (2006-2007) in Peru revealed high nucleotide and haplotype diversity compared to other geographic populations included in the final analysis. These results suggest that the extent of allelic diversity in this region is determined by complex demographic processes that include heterogeneity of transmission intensities in time and space, as well as the number of recurrent infections that are common in the region, allowing for a high turnover of parasite genotypes [41]. The presence of asymptomatic cases [64] participating in the infection process might also be contributing to the persistence of genetic variation since they are a “reservoir” of untreated patients. In addition, we need to take in consideration possible *P. vivax* multi clonal infections that might be difficult to observe by direct sequencing. This limitation could potentially affect the overall estimates of genetic diversity and population structure; however, for the goal of this study, the data collected is useful to obtain a broad picture of how these vaccine candidate haplotypes are worldwide distributed. If a fine scale study is considered in which the goal is to understand the population structure of *Plasmodium* in a specific geographic area then, unlinked microsatellite loci should be used to get a better idea of the parasite’s population history.

Table 4 Estimation of genetic differentiation between clusters of populations obtained by Structure.

A)			
Antigen	1 vs 2	1 vs 3	2 vs 3
DBP	0.13742*	0.18583*	0.11845*
AMA-1	0.09891*	0.06873*	0.06239*
TRAP	0.2478*	0.3393*	0.12982*
B)			
Antigen	1 vs 2	1 vs 3	2 vs 3
DBP	0.34601*	0.4555*	0.41168*
AMA-1	0.50583*	0.31637*	0.177725*
TRAP	0.71399*	0.60509*	0.64183*

F_{ST} calculated from A) haplotype frequency and B) sequence diversity with * $P < 0.001$



When considering worldwide data, genes encoding non-merozoite antigens in *P. falciparum* showed variation in diversity [62]. In *P. vivax*, the same pattern was observed using CSP and TRAP. Although, limited population samples were available for TRAP sequences. On the other hand, the genes encoding merozoite antigens in *P. falciparum* and *P. vivax* showed different patterns. *P. falciparum* genes encoding AMA-1, EBA-175 and MSP-1, MSP-3 and MSP-4 showed similar levels of diversity among countries and regions [62]. In *P. vivax*, genes *dbpII* and *ama-1* also showed this trend; however, *msh-1* exhibited very low but different polymorphism and haplotype diversity values among regions and countries. These results support the hypothesis that maintaining the structural conformation of the HARBs

regions (MSP-1₂₀ and MSP-1₁₄) is important for recognition and binding of reticulocytes [31]. The limited variation of amino acid substitutions in these fragments makes them strong candidates to be tested in immunological studies. Also, TRAP sequences showed low levels of nucleotide and haplotype diversity with a highly conserved region between positions 209 to 256, previously reported to be immunogenic.

Contrary to *P. falciparum*, a largely structured diversity in all *P. vivax* populations was not found on the basis of geography. The clusters determined by Structure were found in most countries except in those with small number of samples, suggesting that haplotypes corresponding to those clusters were underrepresented. Moreover, a high level of admixture in all countries was

Table 5 Estimation of genetic differentiation between geographic populations using the Snn statistics with *P < 0.01

A)										
	Brazil	Colombia	Peru	India	Sri Lanka	Thailand	Korea	Iran		
Colombia	0.93864*									
Peru	0.65322*	0.9486*								
India	0.63332*	0.94986*	0.74331*							
Sri Lanka	0.61225*	0.94952*	0.67073*	0.58268*						
Thailand	0.8613*	0.81977*	0.91711*	0.80463*	0.8734*					
Korea	0.92369*	0.91515*	0.90485*	0.9886*	0.80846*	0.80833*				
Iran	0.85693 ns	0.82143*	0.84720*	0.84885*	0.84003*	0.62883*	0.76684*			
PNG	0.91672*	0.9749*	0.93263*	0.90407*	0.96110*	0.93994*	0.9817*	0.95887*		

B)					
	Brazil	India	Korea	Thailand	Turkey
India	0.68639 ns				
Korea	0.6184*	0.61948 ns			
Thailand	0.95065*	0.81079*	0.895 ns		
Turkey	0.0210*	0.52367 ns	0.625 ns	0.76791*	
Singapore	0.84141*	0.60742*	0.7 ns	0.7*	0.113 ns

C)										
	Brazil	Peru	Venezuela	China	India	Korea	Myanmar	Thailand	Solomon	PNG
Peru	0.6455*									
Venezuela	0.6200*	0.7909*								
China	0.9833*	0.9925*	0.9735*							
India	0.8711*	0.9092*	0.8919*	0.9446*						
Korea	0.9851*	0.9903*	0.9744*	0.4744 ns	0.9787*					
Myanmar	0.8850*	0.9520*	0.9096*	0.7959 ns	0.8443*	0.8704 ns				
Thailand	0.9444*	0.9694*	0.9366*	0.9569*	0.8352*	0.9744*	0.7896*			
Solomon	0.9713*	0.9810*	0.9760*	0.9321*	0.9625*	0.8667*	0.9025*	0.9429*		
PNG	0.9792*	0.9846*	0.9784*	0.8444*	0.9505*	0.8914*	0.8628*	0.9368*	0.6270 ns	
Philippines	0.9637*	0.9958*	0.9648*	0.9746*	0.9228*	0.9913*	0.8857*	0.8689*	0.9669*	0.978*

A) DBP B) MSP-1 and C) AMA-1

observed, which could be explained by natural and high fluctuations over time, as a result of frequency-dependent selection or the discrepancy due to the sampling used [62], since the worldwide sequences obtained from GenBank came from different time points and locations. On the other hand, the lack of geographic clustering could be explained by balancing selection acting on the genes studied and as a result high worldwide antigen diversity was observed not only in Asia but also in South America.

Conclusion

In conclusion, it would be ideal to have a good representation of the haplotypes circulating worldwide when implementing a vaccine regardless of the geographic region of deployment. Considering that selective pressure plays an important role in structuring the diversity of antigens, it cannot be assumed that a single strain could accurately represent the haplotypes circulating worldwide and as a consequence be effective in all endemic regions. In addition, *Plasmodium* population

definition should be revised since parasite sub-groups are not necessarily geographically restricted and treatment or vaccination in a region cannot be considered as if it was isolated but a dynamic map needs to be acknowledged. Thus, important epidemiological units should be established according to the parasite's gene flow, which would lead to a more effective malaria control strategy.

Acknowledgements

Several authors of this manuscript are employees of the US Government. This work was prepared as part of their duties. Title 17 U.S.C. § 105 provides that 'Copyright protection under this title is not available for any work of the United States Government.' Title 17 U.S.C. § 101 defines US Government work as a work prepared by a military service member or employee of the US Government as part of that person's official duties. The views expressed in this article are those of the authors only and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the US Government. This research was supported in part by the grant R01GM80586 from the National Institute of Health to AE.

Author details

¹Parasitology Program, Naval Medical Research Unit No. 6, Lima, Peru.

²Arizona State University, School of Life Sciences, Tempe, AZ, USA. ³Center

for Evolutionary Medicine and Informatics, The Biodesign Institute, Arizona State University, Tempe, AZ, USA. ⁴Naval Research Laboratory, Washington, DC, USA.

Authors' contributions

SC conducted the molecular genetic studies, performed the genetic diversity experiments, analysed the data and wrote the first draft of the manuscript. LT performed part of the genetic diversity experiments. AE, CL, SD contributed to the write up and reviewed the final draft. DB designed the project, supervised and directed the research team and contributed to the writing and editing of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 19 December 2011 Accepted: 14 March 2012

Published: 14 March 2012

References

1. Arevalo-Herrera M, Chitnis C, Herrera S: **Current status of *Plasmodium vivax* vaccine.** *Hum Vaccin* 2010, **6**:124-132.
2. Tjitra E, Anstey NM, Sugianto P, Warikan N, Kenangalem E, Karyana M, Lampah DA, Price RN: **Multidrug-resistant *Plasmodium vivax* associated with severe and fatal malaria: a prospective study in Papua, Indonesia.** *PLoS Med* 2008, **5**:e128.
3. Rogerson SJ, Carter R: **Severe vivax malaria: newly recognised or rediscovered.** *PLoS Med* 2008, **5**:e136.
4. Price RN, Douglas NM, Anstey NM: **New developments in *Plasmodium vivax* malaria: severe disease and the rise of chloroquine resistance.** *Curr Opin Infect Dis* 2009, **22**:430-435.
5. **Report on Situation of Malaria in the Americas, PAHO.** 2008.
6. Tediosi F, Maire N, Penny M, Studer A, Smith TA: **Simulation of the cost-effectiveness of malaria vaccines.** *Malar J* 2009, **8**:127.
7. Sharma S, Pathak S: **Malaria vaccine: a current perspective.** *J Vector Borne Dis* 2008, **45**:1-20.
8. Crompton PD, Pierce SK, Miller LH: **Advances and challenges in malaria vaccine development.** *J Clin Invest* 2010, **120**:4168-4178.
9. Mota MM, Rodriguez A: **Migration through host cells: the first steps of *Plasmodium* sporozoites in the mammalian host.** *Cell Microbiol* 2004, **6**:1113-1118.
10. Mann VH, Huang T, Cheng Q, Saul A: **Sequence variation in the circumsporozoite protein gene of *Plasmodium vivax* appears to be regionally biased.** *Mol Biochem Parasitol* 1994, **68**:45-52.
11. Rosenberg R, Wirtz RA, Lanar DE, Sattabongkot J, Hall T, Waters AP, Prasittisuk C: **Circumsporozoite protein heterogeneity in the human malaria parasite *Plasmodium vivax*.** *Science* 1989, **245**:973-976.
12. Kain KC, Brown AE, Webster HK, Wirtz RA, Keystone JS, Rodriguez MH, Kinahan J, Rowland M, Lanar DE: **Circumsporozoite genotyping of global isolates of *Plasmodium vivax* from dried blood specimens.** *J Clin Microbiol* 1992, **30**:1863-1866.
13. Qari SH, Shi YP, Goldman IF, Udhayakumar V, Alpers MP, Collins WE, Lal AA: **Identification of *Plasmodium vivax*-like human malaria parasite.** *Lancet* 1993, **341**:780-783.
14. Goldman IF, Qari SH, Millet PG, Collins WE, Lal AA: **Circumsporozoite protein gene of *Plasmodium simiu*, a *Plasmodium vivax*-like monkey malaria parasite.** *Mol Biochem Parasitol* 1993, **57**:177-180.
15. Bojang KA, Milligan PJ, Pinder M, Vigneron L, Allouche A, Kester KE, Ballou WR, Conway DJ, Reece WH, Gothard P, Yamuah L, Delchambre M, Voss G, Greenwood BM, Hill A, McAdam KP, Tornieporth N, Cohen JD, Doherty T, S Malaria Vaccine Trial Team: **Efficacy of RTS, S/AS02 malaria vaccine against *Plasmodium falciparum* infection in semi-immune adult men in The Gambia: a randomised trial.** *Lancet* 2001, **358**:1927-1934.
16. Alonso PL, Sacarlal J, Aponte JJ, Leach A, Macete E, Milman J, Mandomando I, Spiessens B, Guinovart C, Espasa M, Bassat Q, Aide P, Ofori-Anyinam O, Navia MM, Corachan S, Ceuppens M, Dubois MC, Demoitie MA, Dubovsky F, Menéndez C, Tornieporth N, Ballou WR, Thompson R, Cohen J: **Efficacy of the RTS, S/AS02A vaccine against *Plasmodium falciparum* infection and disease in young African children: randomised controlled trial.** *Lancet* 2004, **364**:1411-1420.
17. Moorthy VS, Reed Z, Smith PG: **Clinical trials to estimate the efficacy of preventive interventions against malaria in paediatric populations: a methodological review.** *Malar J* 2009, **8**:23.
18. Abdulla S, Oberholzer R, Juma O, Kubhoja S, Machera F, Membri C, Omari S, Urassa A, Mshinda H, Jumanne A, Salim N, Shomari M, Aebi T, Schellenberg DM, Carter T, Villafana T, Demoitie MA, Dubois MC, Leach A, Lievens M, Vekemans J, Cohen J, Ballou WR, Tanner M: **Safety and immunogenicity of RTS, S/AS02D malaria vaccine in infants.** *N Engl J Med* 2008, **359**:2533-2544.
19. Cohen J, Nussenzweig V, Nussenzweig R, Vekemans J, Leach A: **From the circumsporozoite protein to the RTS, S/AS candidate vaccine.** *Hum Vaccin* 2010, **6**:90-96.
20. Sultan AA, Thathy V, Frevert U, Robson KJ, Crisanti A, Nussenzweig V, Nussenzweig RS, Menard R: **TRAP is necessary for gliding motility and infectivity of plasmodium sporozoites.** *Cell* 1997, **90**:511-522.
21. Morahan BJ, Sallmann GB, Huestis R, Dubljevic V, Waller KL: ***Plasmodium falciparum*: genetic and immunogenic characterisation of the rhoptry neck protein PfRON4.** *Exp Parasitol* 2009, **122**:280-288.
22. Putapornpit C, Jongwutiwes S, Tia T, Ferreira MU, Kanbara H, Tanabe K: **Diversity in the thrombospondin-related adhesive protein gene (TRAP) of *Plasmodium vivax*.** *Gene* 2001, **268**:97-104.
23. Rogers WO, Gowda K, Hoffman SL: **Construction and immunogenicity of DNA vaccine plasmids encoding four *Plasmodium vivax* candidate vaccine antigens.** *Vaccine* 1999, **17**:3136-3144.
24. Castellanos A, Arevalo-Herrera M, Restrepo N, Gullosio L, Corradin G, Herrera S: ***Plasmodium vivax* thrombospondin related adhesion protein: immunogenicity and protective efficacy in rodents and Aotus monkeys.** *Mem Inst Oswaldo Cruz* 2007, **102**:411-416.
25. Xainli J, Adams JH, King CL: **The erythrocyte binding motif of *Plasmodium vivax* duffy binding protein is highly polymorphic and functionally conserved in isolates from Papua New Guinea.** *Mol Biochem Parasitol* 2000, **111**:253-260.
26. Ocampo M, Vera R, Rodriguez LE, Curtidor H, Urquiza M, Suarez J, Garcia J, Puentes A, Lopez R, Trujillo M, Torres E, Patarroyo ME: ***Plasmodium vivax* Duffy binding protein peptides specifically bind to reticulocytes.** *Peptides* 2002, **23**:13-22.
27. Chootong P, Ntumngia FB, VanBuskirk KM, Xainli J, Cole-Tobian JL, Campbell CO, Fraser TS, King CL, Adams JH: **Mapping epitopes of the *Plasmodium vivax* Duffy binding protein with naturally acquired inhibitory antibodies.** *Infect Immun* 2010, **78**:1089-1095.
28. Singh SK, Hora R, Belrhali H, Chitnis CE, Sharma A: **Structural basis for Duffy recognition by the malaria parasite Duffy-binding-like domain.** *Malar J* 2006, **493**:741-744.
29. Koussis K, Withers-Martinez C, Yeoh S, Child M, Hackett F, Knuepfer E, Juliano L, Woehlbier U, Bujard H, Blackman MJ: **A multifunctional serine protease primes the malaria parasite for red blood cell invasion.** *EMBO J* 2009, **28**:725-735.
30. Caro-Aguilar I, Rodriguez A, Calvo-Calle JM, Guzman F, De la Vega P, Patarroyo ME, Galinski MR, Moreno A: ***Plasmodium vivax* promiscuous T-helper epitopes defined and evaluated as linear peptide chimera immunogens.** *Infect Immun* 2002, **70**:3479-3492.
31. Espinosa AM, Sierra AY, Barrero CA, Cepeda LA, Cantor EM, Lombo TB, Guzman F, Avila SJ, Patarroyo MA: **Expression, polymorphism analysis, reticulocyte binding and serological reactivity of two *Plasmodium vivax* MSP-1 protein recombinant fragments.** *Vaccine* 2003, **21**:1033-1043.
32. Sierra AY, Barrero CA, Rodriguez R, Silva Y, Moncada C, Vanegas M, Patarroyo MA: **Splenectomised and spleen intact Aotus monkeys' immune response to *Plasmodium vivax* MSP-1 protein fragments and their high activity binding peptides.** *Vaccine* 2003, **21**:4133-4144.
33. Gentil F, Bargieri DY, Leite JA, Francoso KS, Patricio MB, Espindola NM, Vaz AJ, Palatnik-de-Sousa CB, Rodrigues MM, Costa FT, Soares IS: **A recombinant vaccine based on domain II of *Plasmodium vivax* Apical Membrane Antigen 1 induces high antibody titres in mice.** *Vaccine* 2010, **28**:6183-6190.
34. Cech PG, Aebi T, Abdallah MS, Mpina M, Machunda EB, Westerfeld N, Stoffel SA, Zurbriggen R, Pluschke G, Tanner M, Daubenberger C, Genton B, Abdulla S: **Virosome-formulated *Plasmodium falciparum* AMA-1 & CSP derived peptides as malaria vaccine: randomized phase 1b trial in semi-immune adults & children.** *PLoS One* 2011, **6**:e22273.
35. Cortes A, Mellombo M, Mueller I, Benet A, Reeder JC, Anders RF: **Geographical structure of diversity and differences between**

- symptomatic and asymptomatic infections for *Plasmodium falciparum* vaccine candidate AMA1. *Infect Immun* 2003, **71**:1416-1426.
36. Garg S, Alam MT, Das MK, Dev V, Kumar A, Dash AP, Sharma YD: Sequence diversity and natural selection at domain I of the apical membrane antigen 1 among Indian *Plasmodium falciparum* populations. *Malar J* 2007, **6**:154.
37. Crewther PE, Matthew ML, Flegg RH, Anders RF: Protective immune responses to apical membrane antigen 1 of *Plasmodium chabaudi* involve recognition of strain-specific epitopes. *Infect Immun* 1996, **64**:3310-3317.
38. Thomas AW, Narum D, Waters AP, Trape JF, Rogier C, Goncalves A, Rosario V, Druilhe P, Mitchell GH, Dennis D: Aspects of immunity for the AMA-1 family of molecules in humans and non-human primates malarias. *Mem Inst Oswaldo Cruz* 1994, **89**(Suppl 2):67-70.
39. Wickramarachchi T, Premaratne PH, Perera KL, Bandara S, Kocken CH, Thomas AW, Handunnetti SM, Udagama-Randeniya PV: Natural human antibody responses to *Plasmodium vivax* apical membrane antigen 1 under low transmission and unstable malaria conditions in Sri Lanka. *Infect Immun* 2006, **74**:798-801.
40. Bueno LL, Fujiwara RT, Soares IS, Braga EM: Direct effect of *Plasmodium vivax* recombinant vaccine candidates AMA-1 and MSP-1₁₉ on the innate immune response. *Vaccine* 2008, **26**:1204-1213.
41. Van den Eede P, Soto-Calle VE, Delgado C, Gamboa D, Grande T, Rodriguez H, Llanos-Cuentas A, Anne J, D'Alessandro U, Erhart A: *Plasmodium vivax* sub-patent infections after radical treatment are common in Peruvian patients: results of a 1-year prospective cohort study. *PLoS One* 2011, **6**:e16257.
42. Imwong M, Pukrittayakamee S, Gruner AC, Renia L, Letourneur F, Looareesuwan S, White NJ, Snounou G: Practical PCR genotyping protocols for *Plasmodium vivax* using *Pcsp* and *Pmsp1*. *Malar J* 2005, **4**:20.
43. Ampudia E, Patarroyo MA, Patarroyo ME, Murillo LA: Genetic polymorphism of the Duffy receptor binding domain of *Plasmodium vivax* in Colombian wild isolates. *Mol Biochem Parasitol* 1996, **78**:269-272.
44. Cheng Q, Saul A: Sequence analysis of the apical membrane antigen I (AMA-1) of *Plasmodium vivax*. *Mol Biochem Parasitol* 1994, **65**:183-187.
45. Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular evolutionary genetics analysis (mega) software version 4.0. *Mol Biol Evol* 2007, **24**:1596-1599.
46. Librado P, Rozas J: DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 2009, **25**:1451-1452.
47. Escalante AA, Cornejo OE, Rojas A, Udhayakumar V, Lal AA: Assessing the effect of natural selection in malaria parasites. *Trends Parasitol* 2004, **20**:388-395.
48. Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 2000, **155**:945-959.
49. Jakobsson M, Rosenberg NA: CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 2007, **23**:1801-1806.
50. Abacassamo F, Enosse S, Aponte JJ, Gomez-Olive FX, Quinto L, Mabunda S, Barreto A, Magnussen P, Ronn AM, Thompson R, Alonso PL: Efficacy of chloroquine, amodiaquine, sulphadoxine-pyrimethamine and combination therapy with artesunate in Mozambican children with non-complicated malaria. *Trop Med Int Health* 2004, **9**:200-208.
51. Hudson RR: A new statistic for detecting genetic differentiation. *Genetics* 2000, **155**:2011-2014.
52. Zakeri S, Abouie Mehrizi A, Djadid ND, Snounou G: Circumsporozoite protein gene diversity among temperate and tropical *Plasmodium vivax* isolates from Iran. *Trop Med Int Health* 2006, **11**:729-737.
53. Seth RK, Bhat AA, Rao DN, Biswas S: Acquired immune response to defined *Plasmodium viva* antigens in individuals residing in northern India. *Microbes Infect Institut Pasteur* 2010, **12**:199-206.
54. Sinnis PCP, Fenyo D, Chait BT, Cerami C, Nussenzweig V: Structural and functional properties of region II-plus of the malaria circumsporozoite protein. *J Exp Med* 1994, **180**:297-306.
55. Gonzalez JM, Hurtado S, Arevalo-Herrera M, Herrera S: Variants of the *Plasmodium vivax* circumsporozoite protein (VK210 and VK247) in Colombian isolates. *Mem Inst Oswaldo Cruz* 2001, **96**:709-712.
56. Santos-Ciminera PD, Alecrim MG, Roberts DR, Quinnan GV Jr: Molecular epidemiology of *Plasmodium vivax* in the State of Amazonas, Brazil. *Acta Trop* 2007, **102**:38-46.
57. Chatterjee S, Wery M, Sharma P, Chauhan VS: A conserved peptide sequence of the *Plasmodium falciparum* circumsporozoite protein and antipeptide antibodies inhibit *Plasmodium berghei* sporozoite invasion of Hep-G2 cells and protect immunized mice against *P. berghei* sporozoite challenge. *Infect Immun* 1995, **63**:4375-4381.
58. Tewari R, Spaccapelo R, Bistoni F, Holder AA, Crisanti A: Function of region I and II adhesive motifs of *Plasmodium falciparum* circumsporozoite protein in sporozoite motility and infectivity. *J Biol Chem* 2002, **277**:47613-47618.
59. Martinez P, Lopez C, Saravia C, Vanegas M, Patarroyo MA: Evaluation of the antigenicity of universal epitopes from PvDBPII in individuals exposed to *Plasmodium vivax* malaria. *Microbes Infect* 2010, **12**:1188-1197.
60. Takala SL, Coulibaly D, Thera MA, Batchelor AH, Cummings MP, Escalante AA, Ouattara A, Traoré K, Niangaly A, Djimé AA, Doumbo OK, Plowe CV: Extreme polymorphism in a vaccine antigen and risk of clinical malaria: implications for vaccine development. *Sci Transl Med* 2009, **1**:2ra5.
61. Hudson RR, Boos DD, Kaplan NL: A statistical test for detecting geographic subdivision. *Mol Biol Evol* 1992, **9**:138-151.
62. Barry AE, Schultz L, Buckee CO, Reeder JC: Contrasting population structures of the genes encoding ten leading vaccine-candidate antigens of the human malaria parasite, *Plasmodium falciparum*. *PLoS One* 2009, **4**:e8497.
63. Nobrega de Sousa T, Carvalho LH, Alves de Brito CF: Worldwide genetic variability of the Duffy binding protein: insights into *Plasmodium vivax* vaccine development. *PLoS One* 2011, **6**:e22944.
64. Branch O, Casapia WM, Gamboa DV, Hernandez JN, Alava FF, Roncal N, Alvarez E, Perez EJ, Gotuzzo E: Clustered local transmission and asymptomatic *Plasmodium falciparum* and *Plasmodium vivax* malaria infections in a recently emerged, hypoendemic Peruvian Amazon community. *Malar J* 2005, **4**:27.

doi:10.1186/1475-2875-11-68

Cite this article as: Chenet *et al.*: Genetic diversity and population structure of genes encoding vaccine candidate antigens of *Plasmodium vivax*. *Malaria Journal* 2012 **11**:68.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

