Infection, Genetics and Evolution 11 (2011) 1327–1339

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Genetic diversity of *Plasmodium vivax* Duffy Binding Protein II (PvDBPII) under unstable transmission and low intensity malaria in Sri Lanka

Prasad H. Premaratne [a], B. Ramesh Aravinda [a], Ananias A. Escalante [b], Preethi V. Udagama [a,*]

[a] Department of Zoology, Faculty of Science, University of Colombo, Colombo 03, Sri Lanka
[b] School of Life Sciences, Arizona State University, Tempe, AZ, USA

ARTICLE INFO

Article history:
Received 5 January 2011
Received in revised form 21 March 2011
Accepted 21 April 2011
Available online 28 April 2011

Keywords:
Sri Lanka
Plasmodium vivax
Duffy Binding Protein
Genetic diversity

ABSTRACT

Elucidating the genetic diversity of the Duffy Binding Protein II (PvDBPII), a leading vaccine candidate for vivax malaria, in different geographical settings is vital. In Sri Lanka malaria transmission is unstable with low intensity. A relatively high level of allelic diversity, with 27 polymorphic nucleotides and 33 (aa) haplotypes was detected among the *PvdbpII* gene in 100 local *Plasmodium vivax* isolates collected from two hypoendemic areas, and from a non endemic area of the country. Mutations, recombination and balancing selection seem to maintain the observed local allelic diversity of *PvdbpII*. Lack of gene flow was evidenced by high *Fst* values between the two endemic study sites. Some of the aa polymorphisms may alter the binding and expression capacity of predicted T cell epitopes in PvDBPII. Of the 8 binding inhibitory linear B cell epitopes, 2 (H2 and M1) in the vicinity of the exact binding region of PvDBPII appeared to be highly conserved in Sri Lankan, Iran and Colombian isolates, while H3, M2, M3 and L3 neutralizing epitopes seem to be polymorphic globally, with H1 and L2 conserved in Colombian, South Korean and Iran isolates. In comparison to the reference Sal-1 strain, among 402 world-wide isolates (302 global and 100 local), 121 aa polymorphisms and 138 haplotypes were recorded of which 3 aa polymorphisms and 21 haplotypes seem to be unique to Sri Lanka. *PvdbpII* phylogeny suggests that local *P. vivax* parasites represent a sample of the global population. The ubiquitous presence of some PvDBPII aa haplotypes among both local and global *P. vivax* isolates may aid future vaccination strategies based on PvDBPII.

1. Introduction

With its global widespread predominance in the tropics, subtropics and temperate regions, and being responsible for 25–40% of the annual cases of malaria worldwide (reviewed by Price et al., 2007), *Plasmodium vivax* is the major cause of malaria outside of Africa, mainly afflicting Asia and the Americas with approximately 2.5 billion people worldwide at risk from this infection (reviewed by Baird, 2009; Guerra et al., 2010). The re-emergence of *P. vivax* in areas where it was considered eradicated, the emergence of drug resistance, and its association with severe and fatal malaria are evidence of its significant public health importance than traditionally considered (reviewed by Price et al., 2007; Tjitra et al., 2008). Given the substantial differences between *P. vivax* and *Plasmodium falciparum* in terms of their biology, pathogenesis, and epidemiology, it cannot be assumed that interventions developed for the control of *P. falciparum* will be similarly successful against *P. vivax*, and highlights the need of developing effective, long-term control strategies to reduce the impact of this disease (reviewed by Mueller et al., 2009). Due to the development of increasing resistance to both insecticides and anti-malarials, these strategies are increasingly becoming insufficient to reduce the global burden of malaria. An important part of any control strategy will be the implementation of a vaccine capable of inducing strain transcending immunity (Ntumngia et al., 2009); such a strategy seems particularly important in *P. vivax*.

Studies in humans and animal models have substantiated that immune responses targeting blood-stage merozoite antigens may hamper the parasite ability of invading the RBC and offer protection against clinical disease (reviewed by Richards and Beeson, 2009). However, extensive antigenic diversity associated with most of the vaccine candidate antigens of the *Plasmodium* merozoite, along with several other factors, hampers the progress of the development of blood stage vaccine(s).

The complex multistep RBC invasion process of *P. vivax* is dependent on the recognition of the Duffy blood group antigen

* Corresponding author at: Department of Zoology, Faculty of Science, University of Colombo, No. 94, Cumaratunga Munidasa Mawatha, Colombo 03, Sri Lanka. Tel.: +94 11 2503399; fax: +94 11 2503148.
E-mail address: dappvr@yahoo.com (P.V. Udagama).

1567-1348/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.meegid.2011.04.023

(DA) receptor on the host RBC surface by the merozoite ligand, the Duffy Binding Protein (DBP), exported to the merozoite surface during invasion (Adams et al., 1992; Chitnis and Miller, 1994). Two key observations, *i.e.* the complete resistance of Duffy negative individuals to *P. vivax* merozoite invasion (Miller et al., 1976), and reduced susceptibility to *P. vivax* of heterozygous carriers of a Duffy-negative allele compared with wild-type homozygotes (Kasehagen et al., 2007), provide strong evidence that complete or partial disruption of the expression of DA reduces invasion ability of the merozoite restricting the blood stage development of *P. vivax*. Although recent observations of transmission of *vivax* malaria in Duffy-negative populations (Ménard et al., 2010) suggest alternative invasion pathway(s) of *P. vivax*, how extended this phenomenon is remains obscure. Thus, *P. vivax* DBP (PvDBP) seems one of the most promising vaccine candidates against this disease.

Naturally acquired antibodies to PvDBP are prevalent in residents of areas where malaria is endemic (Fraser et al., 1997; Xainli et al., 2003; Tran et al., 2005; Cerávolo et al., 2008; Souza-Silva et al., 2010). Few studies demonstrate the presence of naturally acquired binding inhibitory antibodies directed against the PvDBP (Grimberg et al., 2007; Cerávolo et al., 2008; King et al., 2008; Souza-Silva et al., 2010) which contained both strain specific and strain transcending components (Cole-Tobian et al., 2009). Some such observations were confirmed in a recent study from Sri Lanka (P.V. Udagama-Randeniya, unpublished data). The fact that PvDBP specific binding inhibitory antibodies confer protection against blood stage *in vivo*, provide ample evidence that protective immune response to *P. vivax* is at least partially directed against PvDBP, and reiterate the importance of developing a vaccine based on this antigen (King et al., 2008).

The critical binding motif (CBM) was mapped to a central 170 aa stretch within the 330-aa residue binding motif (with 12 cysteines) located in region II of the DPB (PvDBPII). Six residues that lie between cysteines 4 and 6 come together in three dimensional space to form the exact DA recognition site of PvDBPII during RBC invasion (Mayor et al., 2005; Singh et al., 2006). Though the cysteines and the exact DA binding residues are conserved, extensive polymorphisms are associated with many other residues of the CBM (reviewed by Chitnis and Sharma, 2008). Many of these polymorphic residues, (i) were non-synonymous and organized into clusters of two contiguous stretches that lie opposite to the DA recognition site of PvDBPII (Singh et al., 2006), and (ii) can collectively alter the antigenic character of the molecule which can significantly change the sensitivity to inhibitory antibodies directed against PvDBPII (VanBuskirk et al., 2004). The pattern of polymorphism associated with the CBM imply the existence of selection pressure, suggesting that allelic variation functions as a mechanism for immune evasion (Cole-Tobian and King, 2003; Chootong et al., 2010). Such newly selected mutants will spread affecting the population structure, and may provide insight to the evolution and selection of parasite populations over time (Rich et al., 1997; Gosi et al., 2008). Many similar PvDBP alleles are widely distributed among different geographical areas worldwide (reviewed by Chitnis and Sharma, 2008). Although DBP represents an ideal vaccine target, the allelic variation and the associated strain specific immunity represent challenges for development of a broadly effective vaccine.

Given the complex geographic structure of *P. vivax* that may affect the observed genetic diversity of putative vaccine antigens (Cornejo and Escalante, 2006), the characterization of *PvdbpII* in different geographic regions will be particularly important in vaccine development and deployment. Such vital information was neither available from the Indian sub continent, nor from Sri Lanka where *P. vivax* is responsible for 65–80% of the total reported annual malaria incidence. Although, malaria conditions are described to be unstable transmission with low intensity (Mendis et al., 1990), the geographic isolation of Sri Lanka may impose unique selection constraints on the local parasite population. We thus analysed the nature of the genetic polymorphism, including that of the predicted linear B and T cell epitopes, of the *PvdbpII* gene in Sri Lanka by examining 100 local *P. vivax* clinical isolates, and the forces driving the maintenance of this genetic diversity. Also, the associations among local and global PvDBPII sequences were investigated by comparison of their phylogeny.

2. Materials and methods

2.1. Study sites and sample collection

This study protocol received approval from the Ethical Review Committee, Faculty of Medicine, University of Colombo, Sri Lanka (EC/04/103). Blood samples were collected between December 1998 and March 2000 (Wickramarachchci et al., 2010), with informed consent, from adults (age >15 years) with microscopically confirmed *P. vivax* infections, prior to anti malarial treatment. The patients were selected from two *P. vivax* endemic areas, Anuradhapura (8°22′N, 80°20′E; $N = 42$) and Kataragama (6°25′N, 81°20′E; $N = 73$), situated 250 km apart, and from residents of the non-endemic malaria area, Colombo (7°55′N, 79°50′E; $N = 52$), where the patients were acquired the disease only after visiting to the areas with natural malaria transmission in the island (Fig. 1). During sample collection, the annual parasite incidence due to *P. vivax* was 20–40 and 80–160 per 1000 individuals from Anuradhapura and Kataragama, respectively (Briët et al., 2003). Though the samples from Kataragama were collected from patients living in 7 contiguous villages comprising an area of only

Fig. 1. Map of Sri Lanka indicating the locations of the three study sites.

10 km^2 (Mendis et al., 1990), those from Anuradhapura represented the entire district with a significantly larger area, suggesting these two sets of isolates constitute discrete study populations. On the other hand, though the possibility of a relapse could not be excluded from the Colombo samples, the origin of these infections could be traced back to a recent visit to a *P. vivax* transmission area that included 11 different administrative districts within the country, including those two test endemic areas under study. Furthermore, two isolates were traced to be of South Indian origin.

2.2. Extraction of parasite genomic DNA and identification of single clone infections

Genomic parasite DNA was extracted from 5 ml of venous blood using phenol:chloroform method as described previously (Gunasekera et al., 2007). The final DNA extracts were re-suspended in Tris buffer and stored at −20 °C until further used.

The clonality of each *P. vivax* infection was investigated using a combined nested polymerase chain reaction (PCR) and restriction fragment length polymorphism (RFLP) analysis of the polymorphic msp-3α locus as described recently (Wickramarachchci et al., 2010). Only those confirmed as single clone infections were used for the present study.

2.3. PvDBPII PCR amplification and sequencing

A 224 aa stretch of the PvDBPII spanning 292–515 aa (numbered according to Fang et al., 1991) was amplified by nested PCR as previously described (Cole-Tobian et al., 2002), with minor modifications. Briefly, the final $MgCl_2$ concentration was adjusted to 15 mM in the reaction mixture, while 0.25 units of *Taq* polymerase (Promega, USA) with 100 μM of each of the four deoxynucleotide triphosphates was used per reaction. Each PCR product was visualized by electrophoresis on 1% agarose gels as described (Cole-Tobian et al., 2002).

PCR products were purified by Wizard@ SV Gel and PCR cleanup system (Promega, USA) and each amplicon was subjected to direct sequencing in an ABI prism automated sequencer using BigDye termination chemistry (Macrogen Inc., S. Korea). The forward and reverse primers used for nest 2 PCR amplification was also applied for sequencing. Three separate PCR products each from the initial 15 isolates were sequenced in both forward and reverse directions that resulted in sixfold sequence coverage, while the remaining isolates were subjected to twofold sequence coverage. Each mutation that resulted in only a single isolate was confirmed by sequencing of a new PCR product.

2.4. Editing and assembling of sequencing products

Sequences were assembled and edited using Seqman II and EditSeq of DNAstar programme (DNAstar, Madison, WI) and was aligned using Clustal W algorithm of MEGA ver 4.0 (Tamura et al., 2007). The complete 100 sequences of the present study was submitted to the GenBank; Colombo ($N = 36$: GU143914–GU143949), Anuradhapura ($N = 24$: GU143950–GU143973) and Kataragama ($N = 40$: GU143974–GU144013). Using *P. vivax* Salvador-1 strain (Sal-1), M37514, as the reference, sequences of the present study were compared with previously published *PvdbpII* sequences globally from regions with different *P. vivax* transmission intensities; Colombia ($N = 17$: U50575–U50591), Papua New Guinea (PNG) ($N = 111$: AF469515–AF469602, AF289480–AF289483, AF291096, AF289635–AF289639, AF289641–AF289653), South Korea ($N = 13$: AF215737–AF215738, AF220657–AF220667), Thailand ($N = 30$: EF219451, EF368159–EF368177, EF368179–EF368180, EF379127–EF379132, EF379134–EF379135), Iran ($N = 9$: EU860430–EU860438) and Brazil ($N = 122$: EU812839–EU812960).

2.5. Statistical analyses

Statistical analyses were performed using the options available in MEGA ver 4.0 (Tamura et al., 2007) and DnaSP ver 4.10.9 (Rozas et al., 2003).

The following genetic diversity indices were considered; π – the average number of nucleotide substitutions per site between any 2 sequences with Jukes and Cantor correction, S – number of segregating (polymorphic) sites, H and Hd – number of haplotypes and haplotype diversity, SP – the number of singleton positions and the PI – parsimony informative sites, and p – number of amino acid difference per site.

2.6. Test of neutrality and selection pressure

In order to explore the effect of natural selection, number of synonymous (dS) and non-synonymous (dNS) substitutions per site was investigated using Nei and Gojobori method (1986) with Jukes and Cantor correction. The dNS − dS difference test statistics were applied to test the null hypothesis of strict neutrality of this gene. Standard error was determined by 1000 bootstrap replications for dS and dNS, as well as two tail Z-test on the difference between dNS and dS, using MEGA 4.

The McDonald–Kreitman (MK) (McDonald and Kreitman, 1991) test was used as a second test of natural selection, using *Plasmodium knowlesi dbp α* gene [L14805-07] as an out group. Fisher's exact test was applied to the data to test for significant non randomness ($P < 0.05$), and the skew from randomness was calculated as the Neutrality index. Tajima's D-test (Tajima, 1989), D^* and F^* statistics of Fu's and Li's tests (Fu and Li, 1993), were performed for testing the hypothesis that the allele frequency range is compatible with the neutral model.

2.7. Recombination

Analysis of recombination in *PvdbpII* was performed on alignment of sequences by DnaSP to calculate the minimum number of recombination events (Rm) that have occurred through the sequence (Hudson and Kaplan, 1985). The recombination parameter (C) was estimated (Hudson, 1987), which include the effective population size and probability of recombination between adjacent nucleotides per generation.

2.8. Genetic differences between parasite populations

The degree of genetic differentiation of *PvdbpII* gene was analysed and compared within the sub-populations of the two endemic areas in Sri Lanka, and among populations of other countries by estimating *Fst* values using DnaSP version 4.10.9.

2.9. Phylogenetic analysis

Phylogenetic analysis was used to investigate the associations of PvDBPII alleles recorded in the present study with those previously published from different malaria endemic regions worldwide. Phylogenetic tree for a 672 bp region of *PvdbpII* was constructed using the Neighbor-Joining method (NJ) with 1000 bootstrap replicate, the Tamura's three-parameter distance model as implicated in the MEGA ver 4.0. Altogether 274 sequences were used to construct the phylogenetic tree, including 100 sequences from Sri Lanka, 111 from PNG, 17 Colombia, 13 South Korea, 30 Thailand and one each from India, Vietnam and Indonesia.

2.10. Polymorphism associated with B and T cell epitopes of PvDBPII

Using previously published major linear B and T-cell epitopes of the CBM of the PvDBPII (Xainli et al., 2002, 2003), we analysed the polymorphisms associated with B and T cell epitopes in Sri Lankan isolates. Further, the polymorphisms associated with eight linear B cell epitopes, which are targets of naturally acquired binding inhibitory antibody responses (Chootong et al., 2010), were also analysed in the Sri Lankan isolates.

The SYFPEITHI prediction tool (www.syfpeithi.de) (Rammensee et al., 1999) was used to predict the T cell epitopes associated with the CBM of the PvDBPII molecule (based on Sal-1 sequence). The differential binding ability of the predicted T cell epitopes of PvDBPII, with four most prevalent HLA-DRB1 alleles present in malaria endemic Asian and Afro-American populations (Saravia et al., 2008), was analysed. Subsequently, polymorphisms associated with the predicted T cell epitopes of the PvDBPII in natural Sri Lankan *P. vivax* isolates and their effect of binding ability with the HLA-DRB1 alleles were also predicted.

3. Results

Of 152 confirmed single clone *P. vivax* isolates, only 100 (Kataragama, $N = 40$; Anuradhapura, $N = 24$ and Colombo, $N = 36$) were successfully amplified for the 672 bp PvDBPII fragment.

3.1. Sequence polymorphism and diversity indices of the PvdbpII gene

Sequence analysis of the local *P. vivax* isolates revealed that, although size polymorphisms was not observed, single nucleotide polymorphisms were present within the sequenced PvdbpII gene fragment. Six residues surmised to make direct contact with the DA receptor (according to the Pv/Pk α DBL 3D structure), were invariant in local isolates. Compared to the Sal-1 strain, 27 dimorphic variable nucleotides which resulted in 27 mutations ($S = 27$) were recorded in the local isolates (Table 1). Only 3 of these polymorphisms (at positions 583, 651 and 669) were singleton (Si) variables while the rest (24) were parsimony informative (Pi) sites (Table 2). Pairwise nucleotide diversity (π) for overall local isolates was 0.00982 while those from the Anuradhapura recorded the highest values ($\pi = 0.0123$), followed by Colombo and Kataragama (Table 2). Further analysis with the sliding window option (window length 100 bp, step size 12) which allows the visualization of high polymorphic regions, resulted that within PvDBPII, π ranged from 0 to 0.02094, and the maximum diversity was found between cysteines 5 and 8 (data not shown). While 10 nucleotide mutations each occurred at the first and third base positions, 6 occurred at the second base position of the codons, that gave rise to 39 different nucleotide haplotypes (H = 39), with a haplotype diversity (Hd) of 0.922 (±0.014) (Table 2). Among the three test areas number of different haplotypes was highest in non-endemic Colombo (H = 19), while the haplotype diversity was highest among isolates from Anuradhapura (Hd = 0.942).

Of the 224 amino acids, 22 non synonymous and 4 synonymous mutations were found locally, which resulted in 33 different PvDBPII aa haplotypes (Table 3). Only 13 aa polymorphisms were confined within the CBM of the PvDBPII. Further, only 10 (48%) polymorphisms (371, 384, 385, 386, 390, 503,504, 505, 508 and 513) (Table 1) were placed in the two polymorphic stretches of the 3D structure where the other polymorphisms were distributed throughout the molecule with some of these lying close to the DA receptor binding site. Three dominant aa haplotypes, (SL-1, 2 and 3) which comprised 21%, 16% and 15% isolates, respectively, were found while another 21 haplotypes were represented by a single isolate each (Table 3). Only a single isolate of PvDBPII local haplotype, SL-18, was identical to that of the Sal-1 strain. The analysis of the distribution of the PvDBPII aa haplotypes within the country revealed that while SL-1 was found in 7 administrative districts, SL-2 and SL-3 were recorded from 4 different districts each. Of the rest, 6 haplotypes were distributed in 2 and the remaining 24 were restricted only to a single administrative district suggestive that PvDBPII haplotypes were widely distributed throughout the country (Table 3). However, certain PvDBPII haplotypes showed area restricted distribution in the two endemic areas (restriction of SL-4 and 6 to Kataragama, and SL-5, 7 and 8 to Anuradhapura). Further, an area preference distribution was observed for the three predominant haplotypes, with SL-1 and 3 predominantly found in Kataragama, and SL-2 more abundant in Anuradhapura (Fig. 2). In addition, the two isolates probably acquired from South India revealed that these sequences were identical to SL 1 and SL 2, the most frequent allele types in Sri Lanka.

Polymorphisms N417K, W437R and I503K, either in single or in combination, can alter the efficacy of an acquired binding

Table 1
Nucleotide and amino acid changes compared to the reference Sal-1 sequence within the 672 bp region of the *PvdbpII* gene among the Sri Lankan isolates.

Position[a]	300	306	308	371	378[b]	384	385	386	390	398	404	417	419	424
Sal-1	AAT	TTT	AGG	AAA	CGC	GAT	GAA	AAG	CGT	TCT	ACA	AAT	ATA	TTA
Sri Lanka	**T**AT	TT**G**	AG**T**	**G**AA	CG**T**	G**G**T	**A**AA	AA**T**	C**A**T	**A**CT	A**G**A	AA**A**	AT**G**	**A**TA
Sal-1	N	F	R	K	R	D	E	K	R	S	T	N	I	L
Sri Lanka	Y	L	S	E	R	G	K	N	H	T	R	K	M	I
Percentage[d]	3%	7%	13%	34%		94%	20%	20%	66%	12%	12%	36%	18%	49%

Position[a]	437	465[b]	476[b]	486	489	503	504	505	508[c]	513	515[b]
Sal-1	TGG	ATC	CCA	CAA	ACC	ATA	ATG	GTA	GCA	ACG	GCA
Sri Lanka	**C**GG	AT**A**	CC**C**	**G**AA	**G**CC	A**A**A	A**G**G	**T**TA	**CG**A/**CGC**/**C**CA	A**A**G	GC**C**
Sal-1	W	I	P	Q	T	I	S	V	A	T	A
Sri Lanka	R	I	P	E	A	K	R	L	R/R/P	K	A
Percentage[d]	37%			1%	3%	55%	3%	3%	11%/1%/1%	24%	

[a] Number of the amino acid residue – according to Fang et al. (1991) (bold and underline nucleotides are the variants compared to reference Sal-1 sequence, within that particular codon).
[b] Positions of synonymous nucleotide changes.
[c] Dimorphic mutation (changed in to two amino acid types).
[d] Percentage of detection frequency of amino acid in the Sri Lankan isolates. Positions 306–437 confined within the CBM of the PvDBPII.

Table 2
Genetic diversity indices of the 672 bp region of the *PvdbpII* gene from world wide isolates.

Locality	S[a]	π (SD)[b]	H[c] Hd (SD)[d]	Sp[e]	Pi[f]	P(SE)[g]
Sri Lanka ($N=100$) (Total)	27	0.0098 (±0.0005)	39 0.922 (±0.014)	3	24	0.025 (0.006)
Anuradhapura[1] ($N=24$)	24	0.0123 (±0.0008)	15 0.942 (±0.031)	4	20	0.031 (0.007)
Kataragama[2] ($N=40$)	24	0.0075 (±0.0009)	16 0.868 (±0.035)	8	16	0.019 (0.005)
Papua New Guinea[3] ($N=111$)	69	0.0101 (±0.0005)	56 0.936 (±0.015)	42	27	0.024 (0.005)
Colombia[4] ($N=17$)	14	0.0090 (±0.0007)	15 0.985 (±0.025)	0	14	0.027 (0.007)
South Korea[5] ($N=13$)	33	0.0105 (±0.0033)	10 0.962 (±0.041)	26	7	0.026 (0.006)
Thailand[6] ($N=30$)	29	0.0112 (±0.0005)	24 0.983 (±0.015)	11	18	0.029 (0.007)
Iran[7] ($N=9$)	10	0.0063 (±0.0010)	7 0.944 (0.070)	2	8	0.018 (0.006)
Brazil[8] ($N=122$)	20	0.0082 (±0.0003)	34 0.934 (0.012)	0	20	0.021 (0.006)

Accession numbers: 1 = GU143950–GU143973; 2 = GU143974–GU144013; 3 = AF469515–AF469602, AF289480–AF289483, AF291096, AF289635–AF289639, AF289641–AF289653; 4 = U50575–U50591; 5 = AF215737–AF215738, AF220657–AF220668; 6 = EF219451, EF368159–EF368177, EF368180, EF379127–EF379132, EF379134–EF379135; 7 = EU860430–EU860438; 8 = EU812839–EU812960.

[a] Number of polymorphic (segregating) sites.
[b] Pairwise nucleotide diversity (standard deviation).
[c] Number of nucleotide haplotypes.
[d] Haplotype diversity (standard deviation).
[e] Number of singleton variables.
[f] Number of parsimony informative sites.
[g] Pairwise amino acid diversity (standard error).

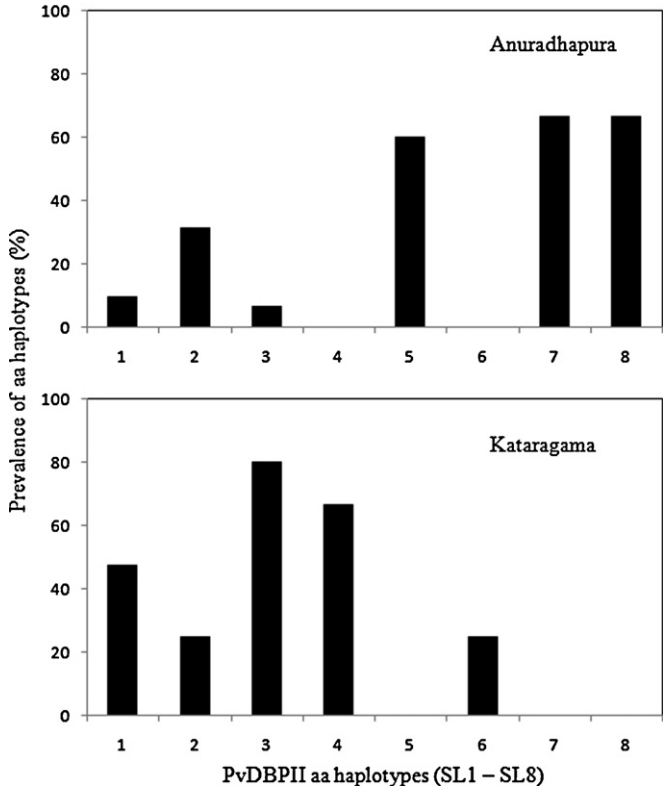

Fig. 2. Distribution of most commonly occurring PvDBPII amino acid haplotypes within the two malaria endemic areas in Sri Lanka. Each of the PvDBPII amino acid haplotypes with >3% prevalence of the total isolates were included. PvDBPII amino acid haplotypes 1–8 represent SL 1–SL 8 as in Table 3.

inhibitory antibody response (VanBuskirk et al., 2004). Locally, variants N417K, W437R and I503K were found at frequencies of 36%, 37% and 55%, respectively. Analyses of the combination of variants revealed that N417K-W437R occurred at a high frequency (99%), while N417K-I503K and W437R-I503K occurred at frequencies of 27% and 30%, respectively. This suggests that though a strong association occur between N417K-W437R, no such strong association was found between I503K with either N417K or W437R, among local PvDBPII alleles.

3.2. Comparison of PvDBPII Sri Lankan sequences with worldwide sequences

Sequence analysis of the 402 isolates from different malaria endemic areas locally and globally revealed that though direct DA binding residues were conserved, 121 variant aa were recorded compared to the reference Sal-1 strain. Of these, only 31 were placed within the two polymorphic clusters analogous to the Pv/Pk α DBL crystal structure. Eighteen of the polymorphisms (306, 308, 371, 384, 385, 386, 390, 398, 404, 417, 419, 424, 437, 486, 503, 504, 505 and 513) found in Sri Lanka confirmed previous records from other malaria endemic areas globally; where 16 were recorded from Thailand, 14 from Brazil, 10 each from PNG and South Korea, 9 from Iran and 8 from Colombia. The most commonly variant amino acids recorded in high frequencies from global isolates (D384G, K386N, N417K, L424I, W437R and I503K) also occurred at high frequencies in local isolates. Interestingly, 3 polymorphisms, N300Y, T489A and A508R or A508P, were exclusively detected in Sri Lanka.

All together, 138 different PvDBPII aa haplotypes were found in global sequence analysis. Of these, 33 were recorded locally of which 8 were also recorded from Brazil, 6 from Thailand, 4 from

Table 3
PvDBPII amino acid haplotypes found in the 100 Sri Lankan isolates.

Haplotype	Amino acid number																					%[a]	Area[b]		
	300	306	308	371	384	385	386	390	398	404	417	419	424	437	486	489	503	504	505	508	513		C	A	K
	N	F	R	K	D	E	K	R	S	T	N	I	L	W	Q	T	I	S	V	A	T				
SL-1	.	.	.	.	G	.	.	H	.	.	.	.	.	.	.	.	K	.	.	.	.	21	9 (6)	2	10
SL-2	.	.	.	E	G	.	.	.	.	.	K	M	I	R	.	.	.	.	.	.	.	16	7 (3)	5	4
SL-3	.	.	.	.	G	.	.	H	.	.	.	.	.	.	.	.	K	.	.	.	K	15	2 (2)	1	12
SL-4	.	.	.	E	G	K	N	H	.	.	K	.	I	R	.	.	.	.	.	.	.	6	2 (2)		4
SL-5	.	L	S	.	G	K	N	H	T	R	.	.	I	.	.	.	.	.	.	.	.	5	2 (1)	3	
SL-6	.	.	.	.	G	.	.	H	.	.	.	.	.	.	.	.	K	.	.	R	.	4	3 (2)		1
SL-7	.	.	.	E	G	.	.	.	.	.	K	.	I	R	.	.	K	.	.	.	.	3	1 (1)	2	
SL-8	.	.	S	.	G	K	N	H	T	R	.	.	I	.	.	.	.	.	.	.	.	3	1 (1)	2	
SL-9	.	.	.	.	G	.	.	H	.	.	.	.	.	.	.	.	K	.	.	R	K	2		2	
SL-10	.	.	.	.	G	.	.	H	.	.	.	.	.	.	.	.	.	.	.	.	.	2	1 (1)		
SL-11	.	.	S	.	G	K	N	H	T	R	.	.	I	.	E	.	.	.	.	.	.	1			1
SL-12	.	L	S	.	G	K	N	H	T	R	.	.	I	.	.	.	.	R	L	.	K	1		1	
SL-13	Y	L	S	.	G	K	N	H	T	R	.	.	I	.	.	.	.	.	.	.	.	1		1	
SL-14	.	.	.	.	.	.	.	.	.	.	K	.	I	R	.	.	K	.	.	.	.	1	1 (1)		
SL-15	.	.	.	.	G	.	.	H	.	.	.	.	.	.	.	A	.	.	.	.	.	1			1
SL-16	.	.	.	E	G	.	.	.	.	.	K	M	I	R	.	.	.	.	.	R	.	1	1 (1)		
SL-17	Y	.	.	.	G	.	.	H	.	.	.	.	.	.	.	.	K	.	.	.	.	1	1 (1)		
SL-18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1			1
SL-19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	L	.	.	1		1	
SL-20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	1			1
SL-21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	K	1	1 (1)		
SL-22	.	.	.	E	G	.	.	.	.	.	.	.	I	R	.	.	K	.	.	.	.	1	1 (1)		
SL-23	.	.	.	.	.	.	.	.	.	.	K	.	I	R	.	.	K	R	L	.	.	1		1	
SL-24	.	.	.	.	G	.	.	.	.	.	K	.	I	R	.	.	K	.	.	.	.	1		1	
SL-25	Y	.	.	E	G	.	.	.	.	.	K	.	I	R	.	A	K	.	.	R	.	1			1
SL-27	.	.	.	E	G	.	.	.	.	.	K	.	I	R	.	A	K	.	.	R	K	1		1	
SL-28	.	.	.	E	G	.	.	.	.	.	K	.	I	R	.	.	K	.	.	R	.	1			1
SL-29	.	.	.	.	G	.	.	H	.	.	.	.	.	.	.	.	K	.	.	P	K	1			1
SL-30	.	.	S	.	G	K	N	H	T	R	.	.	I	.	.	.	.	.	.	.	K	1	1 (1)		
SL-31	.	.	.	E	G	K	N	H	.	.	K	.	I	R	.	.	.	.	.	.	K	1	1 (1)		
SL-32	.	.	S	E	G	K	N	.	.	.	K	.	I	R	.	.	K	.	.	.	.	1	1 (1)		
SL-33	.	.	.	E	G	.	.	.	.	.	K	.	I	R	.	.	.	.	.	.	.	1			1

Amino acid positions were numbered according to Fang et al. (1991).

[a] % = percentage of isolates observed within the country.

[b] Area = number of isolates collected from the three study areas (C = Colombo, A = Anuradhapura, K = Kataragama) number in parentheses denotes the different number of administrative districts traced to the origin of the infections (parasite isolates) of residents of Colombo.

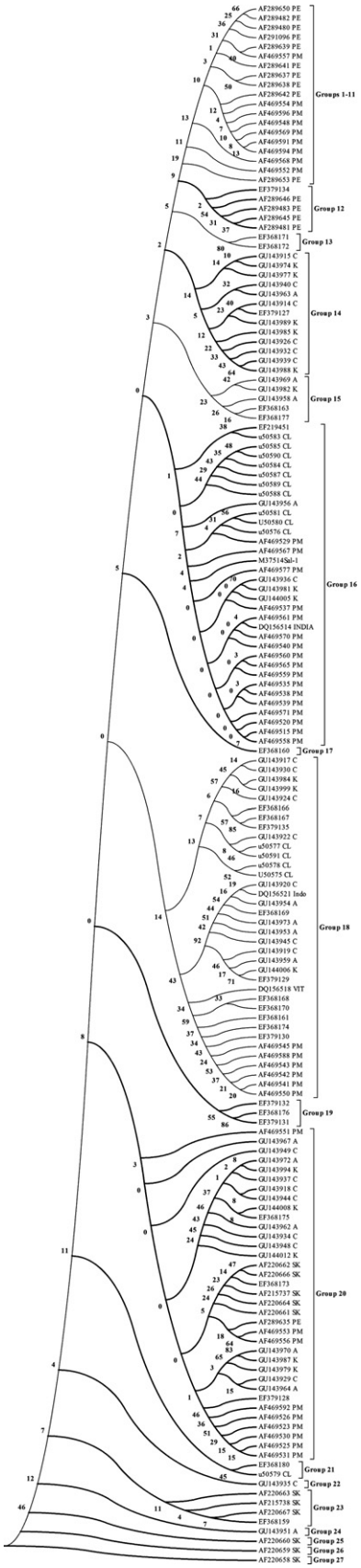

Fig. 3. Neighbor-Joining tree of *PvdbpII* nucleotide sequences (using Kimura 2 parameter distance) showing the phylogenetic relationship between *P. vivax* isolates from different geographical areas globally. Only representative single isolates from different countries globally were used to construct the gene tree. Each isolate number is the relevant GenBank accession number. GU – Sri Lanka (GUxxxxxxC = Colombo; GUxxxxxxA = Anuradhapura; GUxxxxxxK = Kataragama), PNG – Papua New Guinea, THA – Thailand, SK – South Korea, CL – Colombia, Ind – India, INDO – Indonesia and VIT – Vietnam, SAL-1 – *P. vivax* Salvador 1. Numbers at nodes indicate percentage support of 1000 bootstrap replicates.

Iran, each from PNG and South Korea, but none of the local haplotypes matched those recorded from Colombia. Importantly, 21 of the 33 local PvDBPII aa haplotypes seem to be unique to Sri Lanka, since these were not detected in global isolates. Interestingly SL-1, the most common PvDBPII aa haplotype in Sri Lanka found as the most common isolate in Brazil also.

Comparison of the global and local sequences of *PvdbpII* revealed that Thailand, South Korea and PNG recorded high nucleotide diversity than Sri Lanka while Colombia, Iran and Brazil recorded low diversity (Table 2). However haplotype diversity (Hd) was lowest in Sri Lankan isolates. While isolates from Sri Lanka, Thailand, Colombia, Iran and Brazil contained more Pi sites, those from PNG and South Korea contained more Si sites (Table 2).

3.3. Test of neutrality and natural selection

Though not significant, dNS − dS was >1 for all 100 isolates, suggesting that this part of the PvDBP protein may be subjected to positive selection to maintain the polymorphism in the form of balancing selection (Table 4). However, the *Z*-test did not result in significant values for the probability of rejecting the null hypothesis of strict neutrality (dNS = dS).

The 100 local sequences as a whole showed high excess of NS polymorphisms ($P = 0.0465$ based on Fisher's exact test) within *P. vivax* sequences relative to NS polymorphisms from the out group species, *P. knowlesi*, which was significant only in the endemic area Anuradhapura (Table 4). This may evidence departure from neutrality and indicative of balancing selection occurring at the *PvdbpII* gene. Both Tajima's D and Fu and Li's D and F statistics resulted in non significant positive values for sequences as a whole and for Anuradhapura, while for Kataragama a non significant negative value was obtained.

3.4. Recombination

The four gamete test resulted at least in 9 recombination events that occurred among the 100 sequences, while C was estimated at 0.0206 between adjacent nucleotides and 13.8 across the whole gene sequence. It is hard to assess recombination at a locus suspected to be under selection as there is no means to separate convergence from a real recombination event. Both R^2 and D', indices of linkage disequilibrium (LD), declined with increasing molecular map distances between pair of sites. Both LD parameters were significantly larger among sites <200 bp apart on the chromosome with 43% of these pairwise comparisons in LD compared to sites >200 bp apart with a lower R^2 and D' and significantly fewer ($P < 0.0001$; $\chi^2 = 17.55$) sites in LD (21%). These data provide evidence that meiotic recombination may play a role in generating haplotype diversity across *PvdbpII* gene among the Sri Lankan *P. vivax* parasites. At least 5 and 6 recombination events were found with recombination parameter, C, was estimated at 0.0246 and 0.0073 for the endemic test areas Anuradhapura and Kataragama, respectively.

3.5. Genetic differentiation between parasite populations

The degree of genetic differentiation of the *PvdbpII* gene within the two endemic populations in Sri Lanka and among populations of other countries, estimated by *Fst* values, indicated that highest difference with Sri Lankan isolates was observed in Colombia (*Fst* = 0.1828) followed by South Korea (*Fst* = 0.1682), PNG

Table 4
Selection pressure of the *PvdbpII* gene of the clinical *P. vivax* isolates collected from Sri Lanka.

	Rates of synonymous and non-synonymous mutations				McDonald–Kreitman test			Tajima's test	Fu and Li's test	
	dNS[a]	dS[b]	dNS-dS[c] (±SE)	*Z* test	Within Spp. (S/NS)[d]	Between Spp. (S/NS)[e]	Neutrality index (P)[f]	Tajima's D	D*	F*
Total (*N* = 100)	0.0108 (0.003)	0.0058 (0.003)	0.005 (0.004)	dNS > dS *P* = 0.112	5/21	47/70	2.821 (0.046*)	0.7643 *P* > 0.10	0.8144 *P* > 0.10	0.9517 *P* > 0.10
Anuradhapura (*N* = 24) (malaria endemic)	0.0133 (003)	0.0087 (0.005)	0.0046 (0.006)	dNS > dS *P* = 0.221	4/22	47/70	3.641 (0.023*)	1.0222 *P* > 0.10	0.6632 *P* > 0.10	0.9050 *P* > 0.10
Kataragama (*N* = 40) (malaria endemic)	0.0085 (0.002)	0.0039 (0.002)	0.0046 (0.003)	dNS > dS *P* = 0.070	6/18	47/70	2.014 (0.175)	−0.3748 *P* > 0.10	−0.6583 *P* > 0.10	−0.6658 *P* > 0.10

[a] Nucleotide diversity of non-synonymous mutations per non-synonymous site (Nei–Gojobori with Jukes–Cantor correction).
[b] Nucleotide diversity of synonymous mutations per synonymous site (Nei–Gojobori with Jukes–Cantor correction).
[c] Differences of dNS and dS with standard deviation estimated by bootstrap with 1000 pseudo replicates.
[d] Fixed difference within species (synonymous/non-synonymous).
[e] Fixed difference between species (synonymous/non-synonymous).
[f] *P* value for Fisher's exact test.

(*Fst* = 0.148), Iran (Fst = 0.0989), Thailand (*Fst* = 0.0860) and Brazil (0.0303). Interestingly, the two endemic areas within Sri Lanka, Anuradhapura (*N* = 24 sequences) and Kataragama (*N* = 40), showed a higher *Fst* value (0.1038) than that between Sri Lanka with Thailand, Iran and Brazil.

3.6. Phylogenetic analysis

A gene tree using the 274 *PvdbpII* sequences revealed that, though some of the isolates from different malaria endemic areas tend to cluster according to geographic locations, most isolates from different geographical areas were interspersed and distributed throughout the phylogenetic tree (Fig. 3). Of the 27 groups evident on the gene tree, the first 11 contained isolates from PNG only, providing evidence for geographic isolation. However, the rest of the PNG isolates were mixed and distributed among four other groups consisting of isolates from different geographical areas. Groups 13, 17 and 19 exclusively contained isolates from Thailand. Groups 16 and 18 both (where Sal-1 reference strain grouped in the former), contained isolates from Sri Lanka, PNG, Colombia, Thailand but interestingly lacked South Korean isolates. Further the group 16 contained a single isolate from India while group 18 contained single isolate each from Vietnam and from Indonesia. Group 20 contained isolates from Sri Lanka, PNG, Thailand and South Korea.

Seven different groups of the gene tree included Sri Lanka isolates. A majority of these concentrated among groups 14, 18 and 20 with no clustering detected according to the sample collection sites (data not shown). Interestingly, of the 27 groups, groups 22 and 24 exclusively contained a single isolate from Sri Lanka. Groups 14 and 15 included Sri Lankan isolates in combination with Thailand isolates. The Sri Lankan sequences thus appeared to represent a sample from *PvdbpII* worldwide genetic diversity, rather than from any particular lineage.

3.7. Polymorphism associated with B and T cell epitopes of PvDBPII

According to the each of the 5 described major linear B (Xainli et al., 2003) and T-cell (Xainli et al., 2002) epitopes associated with the critical binding motif of the *PvdbpII* gene, we found that only 3 contained polymorphic residues among Sri Lankan isolates (Table 5). Of these, peptides 1, 5 and 6 contained 3, 4 and a single, polymorphic residue(s), respectively. All these polymorphic peptides had higher nucleotide diversity compared to that of the entire 672 bp region of the *PvdbpII*, and all were non-synonymous in nature.

Among the two universal PvDBPII epitopes described recently (Saravia et al., 2008), one (RDYVSELPTEVQKLKEKCDG) was conserved among Sri Lanka while (ISVKNAEKVQTAGIVTPYDI), only 13 residues of the other peptide were sequenced in the present study, of which 39% (5/13) were polymorphic among local isolates.

Ten putative continuous B cell linear epitopes, located within the PvDBP molecule, which function as potential inhibitory B cell epitopes were recently described (Chootong et al., 2010). Those B cell linear epitopes were categorized as high/H [*N* = 3], medium/M [*N* = 3] and low/L [*N* = 4]) epitopes based on their ability to inhibit the binding of PvDBPII to its RBC receptor, the DA. Eight of these B cell epitopes located within PvDBPII and their associated polymorphisms among the Sri Lankan and other global isolates were

Table 5
Polymorphisms observed in the predicted linear B and T-cell epitopes of PvDBPII in Sri Lankan isolates.

Peptide	Residues[a]	Epitope[b]	Sequence[c]	π[d]
1	299–313	T/B	V **N** (Y) N T D T N **F**(L) H **R**(S) D I T F R	0.0271
2	315–329	T	L Y L K R K L Y D A A V E G	0
3	321–335	T/B	L I Y D A A V E G D L L L K L	0
4	329–343	T/B	G D L L L K L N N Y R Y N K D	0
5	380–391	B	I F G T **D**(G) **E**(K) **K**(N) A Q Q **R**(H) K Q	0.0299
6	421–435	T	I C K **L**(I) N V A V N I E P Q I Y	0.0113
7	445–459	B	Y V S E L P T E V Q K L K E K	0

[a] Residues were numbered according to Fang et al. (1991).
[b] T = T-cell epitope (Xainli et al., 2002); B = B-cell epitope (Xainli et al., 2003).
[c] Bold underlined residues correspond to polymorphisms and those in parentheses are the alternative amino acid found in the Sri Lankan field isolates as compared to the sequence of the reference Sal-1 strain.
[d] Pairwise nucleotide diversity.

Table 6
Polymorphisms observed in the potential binding inhibitory B cell linear epitopes of the PDBPII ligand domain in local and global *P. vivax* isolates.

Epitope[a]	Residues[b]	Sequence (based on Sal-1 strain)	Country[c]						
			SL	PNG	COL	SK	THAI	IRAN	BRA
H 1	306–321	F H R D I T F R K L Y L K R K L	P[d] (13%)	P	C[e]	C	P	C	P
H 2	328–341	E G D L L L K L N N Y R Y N	C	P	C	P	P	C	P
H 3	384–399	D E K A Q Q R R K Q W W N E S K	P (32%)	P	P	P	P	P	P
M 1	344–355	F C K D I R W S L G D F	C	P	C	P	P	C	C
M 2	414–429	L K G N F I W I C K L N V A V N	P (19%)	P	P	P	P	P	P
M 3	432–447	P Q I Y R W I R E W G R D Y V S	P (6%)	P	P	P	P	P	P
L 2	400–411	A Q I W T A M M Y S V K	P (8%)	P	C	C	P	P	P
L 3	364–377	M E G I G Y S K V V E N N L	P (7%)	P	P	P	P	p	P

[a] PvDBPII binding inhibitory B cell linear epitopes; H1 to H3 high-inhibition, M1 to M3 moderate-inhibition, and L2 and L4 low-inhibition B-cell epitopes (Chootong et al., 2010).
[b] Amino acid residues were numbered according to Fang et al. (1991).
[c] Country of origin of the isolates; SL = Sri Lanka, PNG = Papua New Guinea, COL = Colombia, SK = South Korea, THAI = Thailand. Numbers within the parentheses next to the SL, indicates the percentage of isolates containing polymorphic residues of that particular sequence.
[d] P = polymorphic epitope.
[e] C = conserved epitope.

analysed (Table 6). Among the local isolates, highest percentage of polymorphism (32%) was recorded in the H3 epitope followed by M2 (19%), H1 (13%), L2 (8%), L3 (7%) and M3 (6%). However, H2 and M1 epitopes located in an area that overlaps with the receptor recognition site of the PvDBPII molecule, is highly (100%) conserved among the 100 Sri Lankan isolates which was also observed in all 17 isolates from Colombia and 9 isolates from Iran. Conversely, these two epitopes seem to be subjected to a low level of polymorphism in PNG (single polymorphic site each in H2 and M1 giving rise to 8% and 1% polymorphism in 111 isolates), South Korea (2 and 1 polymorphic sites in H2 and M1 with 8% polymorphisms detected in 13 isolates) and Brazil (M1 is conserved while a single polymorphic site in H2 with 6% polymorphism in 122 isolates). However, the 30 Thai isolates showed 47% and 3% polymorphism due to single variant site each in H2 and M1. Epitopes H1 and L2 were conserved in Colombian, South Korean and Iran isolates. Nevertheless epitopes H3, M2, M3 and L3 were polymorphic to varying degrees in all different geographical areas.

Using T-cell epitope prediction algorithm (SYFPEITHI software – access via www.syfpeithi.de) (Rammensee et al., 1999), 15 mer peptides of the PvDBPII Sal-1 sequence were predicted and evaluated to their ability to bind by four most prevalent HLA-DRB1 alleles present in the malaria endemic population of Asia and Afroamerica. The four HAL-DRBI alleles considered was (HLA-DRB1*0101, HLA-DRB1*0401, HLA-DRB1 *0701 and HLA-DRB1*1101) (Southwood et al., 1998). The binding capacity of peptides was evaluated based on a scoring system predicted by the software in which only a score of $\geq$25 was considered to give successful binding. Of the 224 amino acid region analysed, HLA-DRB1*0101 allele was predicted to be able to successfully binding to 7 peptides, HLA-DRB1*0401 to 6 peptides and both DRB1*0701 and DRB1*1101 alleles to a single peptide each (Table 7). Importantly, local amino acid polymorphisms associated with some of the predicted peptides were able to change the binding capacity of those peptides to class II MHC allotypes. Binding score of DRB1*0101 peptide 1 (residues 492–506) reduced from 27 to 19 by a replacement of I to K, replacement of K to E of the peptide 3 (residues 366–380) reduced the binding score from 26 to 24 and two replacements of peptide 7 (residues 500–514) I to K and A to R reduced the binding score 25 to 17. Replacement of F to L associated with DRB1*0401 binding peptide 8 (residues 303–317) reduced the binding capacity from 28 to 26. Further two HLA-DRB1*0101 binding peptides, 3 (residues 366–380) and 7 (residues 500–524) and a single DRB1*0401 binding peptide 8 (residues 303–317) contained polymorphisms in the anchor residues of those peptides (Table 7). Of those anchor residue polymorphisms, only the polymorphism of DRB1*0401 peptide 8 reduces the

Table 7
SYFPEITHI (www.syfpeithi.de) predicted peptides of the CBM of PvDBPII (Sal-1 sequence) that bind with four selected HLA-DRB1 alleles.

HLA-DRB1 allele type[a]	Peptide number	Residues[b]	Sequence[c]	Binding score[d]
DRB1*0101	1.	492–506	**K** N Q **W** D **V** L S **N** K F *I* *S* *V* K	27
	2.	353–367	**G** D F **G** D **I** I M **G** T D M E G I	26
	3.	366–380	**G** I G **Y** S *K* V V **E** N N L R S I	26
	4.	442–456	**G** R D **Y** V **S** E L **P** T E V Q K L	26
	5.	374–388	**E** N N **L** R **S** I F **G** T *D* *E* *K* A Q	25
	6.	408–422	**Y** S V **K** K **R** L K **G** *N* F *I* W I C	25
	7.	500–514	**N** K F *I* *S* *V* K N *A* E K V Q *T* A	25
DRB1*0401	8.	303–317	**D** T N *F* H *R* **D** I **T** F R K L Y L	28
	9.	341–355	**N** K D **F** C **K** **D** I **R** W S L G D F	28
	10.	492–506	**K** N Q **W** D **V** L S **N** K F *I* *S* *V* K	28
	11.	499–513	**S** N K **F** *I* *S* *V* K **N** *A* E K V Q *T*	28
	12.	443–457	**R** D Y **V** S **E** **L** P **T** E V Q K L K	26
	13.	446–460	**V** S E **L** P **T** **E** V **Q** K L K E K C	26
DRB1*0701	14.	320–334	**K** L I **Y** D **A** A V **E** G D L L L K	30
DRB1*1101	15.	404–418	*T* A M **Y** S **V** K K **R** L K G *N* F *I*	26

[a] Four of the 10 most prevalent HLA-DRB1 alleles present in malaria endemic populations of Asia and Afroamerica (Southwood et al., 1998).
[b] The amino acids corresponding to positions 1, 4, 6 and 9 (and 7 in some cases) were marked by SYFPEITHI in bold since they corresponded to residues fitting into the main HLA-DRB1 allele's pockets on the peptide-binding region – anchor residues (Saravia et al., 2008). Italicized underlined residues are polymorphic sites found in the Sri Lankan isolates.
[c] SYFPEITHI predicted binding score (only scores equal to or more than 25 were considered).

binding capacity while the anchor residue polymorphisms of the other two peptides do not change the binding capacities.

4. Discussion

Assessment of the level of genetic polymorphism associated with *PvdbpII* within and between populations from different malaria endemic geographical regions is a prerequisite for its evaluation as a vaccine candidate. In the present study for the first time, we investigated the genetic polymorphism of *PvdbpII* in Sri Lanka and compared those with global *PvdbpII* sequences.

Conservation of the key DA contact residues and cysteines among the local isolates further confirmed that structurally and functionally important residues of the PvDBPII are conserved among different parasite populations (reviewed by Chitnis and Sharma, 2008). Contrary to the predictions based on the proposed Pk/Pv DBL 3D structure that polymorphism was concentrated within two contiguous stretches that lie opposite to the DA recognition site of PvDBPII (Singh et al., 2006), a majority of polymorphic residues were scattered within the intact 3D structure in local as well as global isolates which reiterate the suggestion of a more extensive and widely dispersed *PvdbpII* polymorphisms being present in nature (McHenry and Adams, 2006).

Both the pairwise nucleotide diversity and amino acid sequence diversity observed in the local isolates were comparable with those from other geographical areas. Although the exact reason was unclear, both the π and *P* was highest in Anuradhapura, an area with comparatively low malaria transmission. The differences between sampling and sequencing techniques used in previous global attempts and the restriction of the size of the sample collection area in Kataragama, compared to Anuradhapura, may explain this disparity. The π value observed for the *PvdbpII* in the present study was similar to that of *Pvama*-1 ($\pi = 0.0095$) (P.V. Udagama-Randeniya, unpublished data) and was considerably lower than *Pvmsp-1* ($\pi = 0.023$) (Dias et al., 2011) for the battery isolates used for the present study. Low nucleotide diversity values of the *PvdbpII* gene than that of the *Pvmsp*-1 may reflect a major functional restriction in DBP protein, due to its crucial role in the invasion process as described previously (Martinez et al., 2004). Further, PvDBPII is more diverse than *eba-175*, its homologue in *P. falciparum* (Escalante et al., 1998; Baum et al., 2003).

Non-synonymous polymorphism in all three bases of different codons created a high level of microheterogeneity in the local PvDBP protein sequences that resulted in 33 different amino acid haplotypes. This number of haplotypes is unexpectedly high under existing unstable and low transmission malaria condition in Sri Lanka. However, these results were in agreement with the general inference that proteins involved in parasite recognition by the host's immune system are under balancing selection given the pressure for accumulating polymorphism as a strategy for evading the host's immune attack (Escalante et al., 2004). Although the 3 residues (N417K, W437R, I503K) described as those directly involved in resisting binding inhibitory antibodies, those were found in different frequencies among the local and global isolates (Sousa et al., 2006; Gosi et al., 2008; Babaeekho et al., 2009). The strong association between N417K and W437R (99%) found among isolates from Sri Lanka, PNG and Brazil, together with their functional effect may further support that these two residues may constitute a discontinuous epitope in PvDBPII involved in antibody mediated immune evasion as described previously (Sousa et al., 2006).

The presence of three dominant PvDBPII amino acid haplotypes reported in Sri Lanka endorsed previous records from other malaria endemic geographic areas (Cole-Tobian et al., 2007). Two plausible reasons for this dominance as described by Cole-Tobian et al. (2007) may be that; (i) certain dominant alleles may bind better to DA and may increase the fitness of the parasite, and (ii) several dominant PvDBPII haplotypes may be less immunogenic, thereby escaping the host immune response. More extensive analysis with higher numbers of isolates is required to elucidate the exact reason/s for the prevalence of high frequencies of certain PvDBPII haplotypes, locally as well as globally. The area restricted occurrence of certain PvDBPII haplotypes (*i.e.* SL – 4, 5, 6, 7 and 8) and high prevalence of some (*i.e.* SL – 1, 2 and 3) were found within both endemic areas of the country. This may be interpreted as the effect of demographic processes that need to be investigated using neutral loci. Geographical isolation of certain *P. vivax* haplotypes of the two endemic areas within the country may be a key reason to have relatively high genetic differentiation (*Fst* = 0.1038) between Anuradhapura and Kataragama as observed in the present study. Interestingly, the area-wise isolation observed for certain PvDBPII haplotypes in the present study, was recently demonstrated for certain genotypes of the *Pvmsp-3α*, from the same parasite populations used for the present study (Wickramarachchci et al., 2010). This further indicates that demography is playing an important role in the observed pattern of genetic diversity.

As high level of polymorphism at nucleotide (74%) and amino acid (85%) levels seen locally is reflected globally, and the worldwide distribution of common variant amino acid haplotypes of *PvdbpII* insinuates that a common vaccine construct based on PvDBP may be useful in different malaria endemic areas, globally. Conversely, the presence of unique amino acid polymorphisms ($N = 3$) and high numbers of unique PvDBPII amino acid haplotypes ($N = 21$) locally, as well as globally clearly indicates the importance of in-depth analysis of molecular epidemiological data in areas with different malaria endemicity. Existing of such unique geographical sequences, may be due to biogeographic limitations such as in Sri Lanka being an island, or may be due to the possible recent introduction of the parasite to that geographic area (Sousa et al., 2010).

Many previous PvdbpII diversity studies, indicated high rates of non-synonymous relative to synonymous mutations (dNS/dS ratio ranged from 1.9 to 3.4) that reflected a positive selection promoting greater diversity of PvDBPII, presumably to avoid host immune responses, irrespective of the geographical distribution (Cole-Tobian and King, 2003). Though the local situation reflected a dNS/dS ratio of 1.86 this difference was not significant. This is further supported by the diverse pattern and increased rates of non-synonymous to synonymous mutations observed in B- and T-cell epitopes both locally and globally. Further, though not significant, the positive values obtained for Tajima's and Fu and Li's D and F test statistics indicated that local PvDBPII alleles occur at more intermediate frequencies than expected with few alleles being rare or near to fixation, which is consistent with the action of balancing selection maintaining the allelic variation in the population, as indicated previously for EBA-175 (Baum et al., 2003). Conversely, the MK test indicated a high ratio of non-synonymous to synonymous mutations significantly smaller between species than within *P. vivax*, which according to Cole-Tobian and King (2003) suggests that purifying selection is operating to reduce the number of non-synonymous substitutions between species. However, the contrasting detection of different types of selective pressure applied on *PvdbpII* observed in local isolates, was previously reported for both functionally important regions II and VI of this molecule (Cole-Tobian and King, 2003; Martinez et al., 2004). This may imply a balance between high functional restriction for maintaining structural constraint due to the key role of the Duffy Binding Protein in the erythrocyte invasion process and the search for diversity as a response against host immune pressure (Martinez et al., 2004). However, the

presence of recombination influences the ability to detect selection since it breaks up the associations between sites under selection and linked variation. Further, recombination is predicted to make Tajima's and Fu and Li's tests conservative in identifying significant positive deviations from neutrality (Polley and Conway, 2001).

The existence of 9 recombination events and decline of the LD with increasing distance between nucleotide sites suggest that in addition to natural selection, meiotic recombination may also contribute to maintain the observed diversity of *PvdbpII* gene among isolates of hypoendemic settings in Sri Lanka, as recorded from other geographical areas as Brazil and Colombia (Sousa et al., 2010). However, absence of the recombination events in areas with higher *P. vivax* transmission intensity such as PNG and South Korea (Cole-Tobian and King, 2003), may suggest that even though recombination forces appear to be acting on *PvdbpII*, their influence may be less important, as suggested for Colombian isolates, which also demonstrated the presence of recombination sites within the *PvdbpII* gene (Martinez et al., 2004). Recombination was found to be a factor of maintaining genetic diversity of two other merozoite proteins, *Pvama-1* and *Pvmsp-1* in the same battery of parasite isolates used for the present study (Gunasekera et al., 2007; Dias et al., 2011, P.V. Udagama-Randeniya – unpublished data) and few other isolates collected from the endemic area Kataragama (Manamperi, 2002). The results of the present and past studies may indicate that rare recombinant haplotypes generated within the Sri Lankan population are highly selectively advantageous or that new haplotypes generated in areas of higher endemicity may sweep into the Sri Lankan population (Gunasekera et al., 2007). The latter was evident as several aa haplotypes (SL-1, 2, 3, 5, 10, 14,18) found also in Sri Lanka were recorded in other *P. vivax* endemic global settings (SL-1 from India and Brazil, SL-2, 3, 5, 10, 11 and 14 from Thailand, SL-18 in PNG and SL-14 in South Korea, also). However, the existence of 20% multiple clone *P. vivax* infections locally (Wickramarachchci et al., 2010), also increases the chance of sexual out crossing, suggests that occurrence of rare recombinant events is also plausible within the local *P. vivax* population.

Local *PvdbpII* genes were distributed throughout the cladogramme interspersed with worldwide alleles. The overall observations indicate that despite the barrier to gene flow imposed mainly due to its island status, Sri Lankan PvDBPII alleles appeared to represent a sample from the worldwide population. The present study reiterate recent local studies on *Pvama-1* (Gunasekera et al., 2007), *Pvmsp-3α* (Wickramarachchci et al., 2010), *Pvmsp-1* (Dias et al., 2011) and microsatellite diversity of *P. vivax* (Karunaweera et al., 2008) where no significant geographic clustering was evident in *P. vivax* population in Sri Lanka, compared to worldwide *P. vivax* populations. Thus it is plausible that although there is strong geographic structure within the country as evidenced by *Fst*, it has not been maintained long enough to allow the emergence of most commonly distributed local alleles that are exclusively found in Sri Lanka. The *Fst* values and phylogenies could be affected by differences in the strength of selection worldwide is important so the pattern of geographic differentiation may not be the result solely of genetic drift. More studies are required in order to separate the effect from drift and selection. Nevertheless, describing the geographic patterns is still important whereas defining the exact individual contribution of each factor (*i.e.* drift, selection, recombination, etc.) may not be possible under many circumstances.

Polymorphism in T-cell epitope regions of parasite antigens may well enable parasites to escape host immune responses, as polymorphism in T-cell epitopes can up or down regulate a T cell response to the index peptide, or completely arrest an immune response assisting the parasite escape the host immune system (Tanabe et al., 2007). Although the number of peptides that included B and T cell epitopes of the PvDBPII among Sri Lankan isolates ($N = 2$ of each epitope type) were lower than the global status described previously (Cole-Tobian and King, 2003), the high nucleotide diversity and the existence of all non-synonymous substitutions suggest that these polymorphisms arose to avoid the host immune response. Importantly, of 8 linear B-cell epitopes identified to induce binding inhibitory antibody responses (Chootong et al., 2010), 6 contained polymorphic residues among the Sri Lankan isolates. Of the epitopes H2 and M1–M3, located in the vicinity to the exact DA binding site though able to induce binding inhibitory antibody responses (Chootong et al., 2010). Importantly H2 and M1 epitopes were found to be highly conserved in Sri Lankan, Columbia and Iran isolates. Both these epitopes were polymorphic to a lower degree in PNG, South Korea and Brazil, as was M1 in Thailand. However, H2 was relatively highly polymorphic in Thailand. Epitopes M2 and M3 on the other hand showed varying degrees of polymorphism in all these different geographical areas. Thus, the relatively conserved binding inhibitory epitopes H2 and M1 seem to be vital in vaccine development. In this vein, it is interesting to note on the ability of *P. vivax* merozoites to enter Duffy negative reticulocytes (Ménard et al., 2010).

Of the two peptides described as universal epitopes for PvDBPII (Saravia et al., 2008), one was found to be conserved among the local isolates, while the higher percentage of polymorphism (39%) recorded for the other peptide in Sri Lankan as well as global sequences, limits its ability to be considered as a universal epitope of PvDBPII. According to the results of SYFPEITHI T-cell epitope prediction software, all four HLA-DRB-1 alleles were predicted to be able to bind and express PvDBPII Sal-1 peptides with a binding score ranging from 25 to 30, implied that HLA-DRB-1 alleles expressed different binding profiles. Some of the polymorphisms changed the binding and expression capacity of that particular peptide and changed its ability to be expressed as a T-cell epitope. Since most of these polymorphisms reduced the binding capacity of that particular epitopes to the HLA-DRB molecules it may be suggested that such polymorphism arose as an evasion mechanism to overcome the host's immunity. However, since the confidence level for predicting epitopes to MHC class II molecules is usually not very high (approximately 50%), also as these molecules accept a wider range of peptides in size and binding capacity, the predicted results must be interpreted with caution.

In conclusion, the present study for the first time revealed that the *PvdbpII* gene among Sri Lankan isolates is genetically diverse which seems to arise as a mechanism of immune evasion, where natural selection and recombination maintain the observed diversity in the form of balancing selection. The conservation of the DA binding site/residues of the PvDBPII molecule and the regular occurrence of the most common and high prevalent polymorphisms and the variant residues of the B and T cell epitopes (including those epitopes that target binding inhibitory antibodies) in diverse *P. vivax* endemic geographical areas worldwide, suggests that a vaccine based on the PvDBPII would be effective against diverse *P. vivax* isolates. However, the presence of unique PvDBPII alleles within Sri Lanka and in different malaria endemic areas globally, indicate the necessity of in-depth molecular epidemiological analyses to be conducted in areas with different malaria transmission intensities worldwide, and those highly prevalent, common alleles should be incorporated into developing a vaccine based on PvDBPII to be deployed universally.

Acknowledgements

This study was financially supported by the National Science Foundation (NSF/RG/2005/HS/06 and NSF/Sch/2004/07) and the National Research Council (NRC-05-34), Sri Lanka. AAE is

supported by the National Institutes of Health, US. All patients who donated blood for this study, the medical superintendents, physicians, housemen and the nursing staff of the General Hospital Anuradhapura and the National Hospital, Sri Lanka in Colombo, are acknowledged for their cooperation. Drs. S.M. Handunnetti and T. Wickramarachchi and Messers L. Perera and S. Bandara are acknowledged for collection of blood samples (through PVUR's grant No. F/3008-1 from IFS, Sweden) are acknowledged for collection of blood samples (for a previous study) through the Malaria Research Unit, Department of Parasitology, Faculty of Medicine, University of Colombo.

References

Adams, J.H., Sim, B.K.L., Dolan, S.A., Fang, X., Kaslow, D.C., Miller, L.H., 1992. A family of erythrocyte binding proteins of malaria parasites. Proc. Natl. Acad. Sci. 89, 7085–7089.

Babaeekho, L., Zakeri, S., Djadid, N.D., 2009. Genetic mapping of the Duffy Binding Protein (DBP) ligand domain of *Plasmodium vivax* from unstable malaria region in the Middle East. Am. J. Trop. Med. Hyg. 80 (1), 112–118.

Baird, J.K., 2009. Resistance to therapies for infection by *Plasmodium vivax*. Clin. Microbiol. Rev. 22, 508–534.

Baum, J., Thoms, A.W., Conway, D.J., 2003. Evidence for diversifying selection on erythrocyte binding antigens of *Plasmodium falciparum* and *P. vivax*. Genetics 163, 1327–1336.

Briët, O.J.T., Gunawardena, D.M., van der Hoek, W., Amerasinghe, F.P., 2003. Sri Lanka malaria maps. Malar. J. 2 (22), www.malariajournal.com/content/2/1/22.

Cerávolo, I.P., Souza-Silva, F.A., Fontes, C.J.F., Braga, E.M., Madureira, A.P., Krettli, A.U., Souza, J.M., Brito, C.F.A., Adams, J.H., Carvalho, L.H., 2008. Inhibitory properties of the antibody response to *Plasmodium vivax* Duffy binding protein in an area with unstable malaria transmission. Scand. J. Immunol. 67, 270–278.

Chitnis, C.E., Miller, L.H., 1994. Identification of the erythrocyte binding domain of *Plasmodium vivax* and *Plasmodium knowlesi* proteins involved in erythrocyte invasion. J. Exp. Med. 180, 497–506.

Chitnis, C.E., Sharma, A., 2008. Targeting the *Plasmodium vivax* Duffy-binding protein. Trends Parasitol. 24, 29–34.

Chootong, P., Ntumngia, F.B., VanBuskirk, K.M., Xainli, J., Cole-Tobian, J.L., Campbell, C.O., Fraser, T.S., King, C.L., Adams, J.H., 2010. Mapping epitopes of the *Plasmodium vivax* Duffy Binding Protein with naturally acquired inhibitory antibodies. Infect. Immun. 78, 1089–1095.

Cole-Tobian, J., Cortes, A., Biasor, M., Kastens, W., Xainli, J., Bockarie, M., Adams, J.H., King, C.L., 2002. Age-acquired immunity to a *Plasmodium vivax* invasion ligand, the Duffy binding protein. J. Infect. Dis. 186, 531–539.

Cole-Tobian, J.L., Michon, P., Bisor, M., Richads, J.S., Beeson, J.G., Mueller, I., King, C.L., 2009. Strain-specific Duffy Binding Protein antibodies correlate with protection against infection with homologous compared to heterologous *Plasmodium vivax* strains in Papua New Guinean children. Infect. Immun. 77, 4009–4017.

Cole-Tobian, J., King, C.L., 2003. Diversity and natural selection in *Plasmodium vivax* Duffy binding protein gene. Mol. Biochem. Parasitol. 127, 121–132.

Cole-Tobian, J.L., Zimmerman, P.A., King, C.L., 2007. High throughput identification of the predominant malaria clone in complex blood stage infections using a multi-SNP molecular haplotyping assay. Am. J. Trop. Med. Hyg. 76, 12–19.

Cornejo, O.E., Escalante, A.A., 2006. The origin and age of *Plasmodium vivax*. Trends Parasitol. 22, 558–563.

Dias, S., Longacre, S., Escalante, A.A., Udagama-Randeniya, P.V., 2011. Genetic diversity and recombination at the C-terminal fragment of the merozoite surface protein-1 of *Plasmodium vivax* (PvMSP-1) in Sri Lanka. Infect. Genet. Evol. 11, 145–156.

Escalante, A.A., Lal, A.A., Ayala, F.J., 1998. Genetic polymorphism and natural selection in the malaria parasite *Plasmodium falciparum*. Genetics 149, 189–202.

Escalante, A.A., Cornejo, O.E., Rojas, A., Udhayakumar, V., Lal, A.A., 2004. Assessing the effect of natural selection in malaria parasites. Trends Parasitol. 20, 388–395.

Fang, X., Kaslow, D.C., Adams, J.H., Miller, L.H., 1991. Cloning of the *Plasmodium vivax* Duffy receptor. Mol. Biochem. Parasitol. 44, 125–132.

Fraser, T., Michon, P., Barnwell, J.W., Noe, A.R., Al-Yaman, F., Kaslow, D.C., Adams, J.H., 1997. Expression and serologic activity of a soluble recombinant *Plasmodium vivax* Duffy binding protein. Infect. Immun. 65, 2772–2777.

Fu, Y.X., Li, W.H., 1993. Statistical tests of neutrality of mutations. Genetics 133, 693–709.

Gosi, P., Khusmith, S., Khalambaheti, T., Lanar, D.E., Schaecher, K., Fukuda, M.M., Miller, S.R., 2008. Polymorphic patterns of Duffy-binding protein among Thai *Plasmodium vivax* isolates. Malar. J. 7, 112, doi:10.1186/1475-2875-7-112.

Grimberg, B.T., Udomsangpetch, R., Xainli, J., McHenry, A., Panichakul, T., Sattabongkot, J., Cui, L., Bockarie, M., Chitnis, C., Adams, J., Zimmerman, P.A., King, C.L., 2007. *Plasmodium vivax* invasion of human erythrocytes inhibited by antibodies directed against the Duffy Binding Protein. PLoS Med. 4 (12), e337, doi:10.1371/journal.pmed.0040337.

Guerra, C.A., Howes, R.E., Patil, A.P., Gething, P.W., Van Boeckel, T.P., Temperley, W.H., Kabaria, C.W., Tatem, A.J., Manh, B.H., Elyazar, I.R., Baird, J.K., Snow, R.W., Hay, S.I., 2010. The international limits and population at risk of *Plasmodium vivax* transmission in 2009. PLoS Negl. Trop. Dis. 4 (8), e774, doi:10.1371/journal.pntd.0000774.

Gunasekera, A.M., Wickramarachchi, T., Neafsey, D.E., Ganguli, I., Perera, L., Premaratne, P.H., Hartl, D., Handunnetti, S.M., Udagama-Randeniya, P.V., Wirth, D.F., 2007. Genetic diversity and selection at the *Plasmodium vivax* apical membrane antigen-1 (*PvAMA-1*) locus in a Sri Lankan population. Mol. Biol. Evol. 24, 939–947.

Hudson, R.R., Kaplan, N.L., 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111, 147–164.

Hudson, R.R., 1987. Estimating the recombination parameter of a finite population model without selection. Gen. Res. 50, 245–250.

Kasehagen, L.J., Mueller, I., Kiniboro, B., Bockarie, M.J., Reeder, J.C., Kazura, J.W., Kastens, W., McNamara, D.T., King, C.H., Whalen, C.C., Zimmerman, P.A., 2007. Reduced *Plasmodium vivax* erythrocyte infection in PNG Duffy negative heterozygous. PLoS ONE 2 (3), e336, doi:10.1371/journal.pone.0000336.

Karunaweera, N.D., Ferreira, M.U., Munasinghe, A., Barnwell, J.W., Collins, W.E., King, C.L., Kawamoto, F., Hartl, D.L., Wirth, D.F., 2008. Extensive microsatellite diversity in the human malaria parasite *Plasmodium vivax*. Gene 410, 105–112.

King, C.L., Michon, P., Shakri, A.R., Marcotty, A., Stanisic, D., Zimmerman, P.A., Cole-Tobian, J.L., Mueller, I., Chitnis, C.E., 2008. Naturally acquired Duffy-binding protein-specific binding inhibitory antibodies confer protection from blood stage *Plasmodium vivax* infection. PNAS 105, 8363–8368.

Manamperi, A., 2002. Preliminary preparation towards malaria vaccine trials in Sri Lanka; studies on genetic polymorphism and immune responses against the Cterminal merozoite surface protein-1. PhD Thesis. University of Colombo, Colombo, Sri Lanka.

Martinez, P., Suarezi, C.F., Cardenas, P.P., Patarroyo, M.A., 2004. *Plasmodium vivax* Duffy binding protein: a modular evolutionary proposal. Parasitology 128, 353–366.

Mayor, A., Bir, N., Sawhney, R., Singh, S., Pattnaik, P., Singh, S.K., Sharma, A., Chitnis, C.E., 2005. Receptor-binding residues lie in central regions of Duffy-binding like domains involved in red cell invasion and cytoadherence by malaria parasites. Blood 105, 2557–2563.

McDonald, J.H., Kreitman, M., 1991. Adaptive protein evolution at the Adh locus in Drosophila. Nature 351, 652–654.

McHenry, A.M., Adams, J.H., 2006. The crystal structure of *P. knowlesi* DBPα DBL domain and its implication for immune evasion. Trends Biochem. Sci. 31, 487–491.

Ménard, D., Barnadas, C., Bouchier, C., Henry-Halldin, C., Gray, L.R., Ratsimbasoa, A., Thonier, V., Carod, J., Domarle, O., Coling, Y., Bertrand, B., Picot, J., King, C.L., Grimberg, B.T., Mercereau-Puijalon, O., Zimmerman, P.A., 2010. *Plasmodium vivax* clinical malaria is commonly observed in Duffy-negative Malagasy people. PNAS 107, 5967–5971.

Mendis, C., Gamage-Mendis, A.C., De Zoysa, A.P.K., Abhayawardena, T.A., Carter, R., Herath, P.R.J., Mendis, K.N., 1990. Characteristics of malaria transmission in Kataragama, Sri Lanka: a focus for immuno-epidemiological studies. Am. J. Trop. Med. Hyg. 42, 298–308.

Miller, L.H., Manson, S.J., Clyde, D.F., McGinnis, M.H., 1976. The resistance factor to *Plasmodium vivax* in blacks: Duffy blood group genotype, Fy Fy. N. Engl. J. Med. 295, 302–304.

Mueller, I., Galinski, M.R., Baird, J.K., Carlton, J.M., Kochar, D.K., Alonso, P.L., del Portillo, H.A., 2009. Key gaps in the knowledge of *Plasmodium vivax*, a neglected human malaria parasite. Lancet Infect. Dis. 9, 555–566.

Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.

Ntumngia, F.B., McHenry, A.M., Barnwell, J.W., Cole-Tobian, J., King, C.L., Adams, J.H., 2009. Genetic variation among *Plasmodium vivax* isolates adapted to non-human primates and the implication for vaccine development. Am. J. Trop. Med. Hyg. 80, 218–227.

Polley, S.D., Conway, D.J., 2001. Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene. Genetics 158, 1505–1512.

Price, R.N., Tjitra, E., Guerra, C.A., Yeung, S., White, N.J., Anstey, N.M., 2007. Vivax malaria: neglected and not benign. Am. J. Trop. Med. Hyg. 77, 79–87.

Rammensee, H., Bachmann, J., Emmerich, N.P., Bachor, O.A., Stevanovic, S., 1999. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50, 213–219.

Rich, S.M., Hudson, R.R., Ayala, F.J., 1997. *Plasmodium falciparum* antigenic diversity: evidence of clonal population structure. PNAS 94, 13040–13045.

Richards, J.S., Beeson, J.G., 2009. The future for blood stage vaccine against malaria. Immunol. Cell Biol. 87, 377–390.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Saravia, C., Martinez, P., Granados, D.S., Lopez, C., Reyes, C., Patarroyo, M.A., 2008. Identification and evaluation of universal epitopes in *Plasmodium vivax* Duffy binding protein. Biochem. Biophys. Res. Commun. 377, 1279–1283.

Singh, S.K., Hora, R., Belrhali, H., Chitnis, C.E., Sharma, A., 2006. Structural basis for Duffy recognition by the malaria parasite Duffy-binding-like domain. Nature 439, 741–744.

Southwood, S., Sidney, J., Kondo, A., del Guercio, M.F., Appella, E., Hoffman, S., Kubo, R.T., Chesnut, R.W., Grey, H.M., Sette, A., 1998. Several common HLA-DR types share largely overlapping peptide binding repertoires. J. Immunol. 160, 3363–3373.

Sousa, T.N., Cerávolo, I.P., Fontes, C.J.F., Couto, A., Carvalho, L.H., Brito, C.F., 2006. The pattern of major polymorphisms in the Duffy binding protein ligand domain among *Plasmodium vivax* isolates from the Brazilian Amazon area. Mol. Biochem. Parasitol. 146, 251–254.

Souza-Silva, F.A., Silva-Nunes, M., Sanchez, B.A.M., Cerávolo, I.P., Malafronte, R.S., Brito, C.F.A., Ferreira, M.U., Carvalho, L.H., 2010. Naturally acquired antibodies to *Plasmodium vivax* Duffy Binding Protein (DBP) in Rural Brazilian Amazon. Am. J. Trop. Med. Hyg. 82 (2), 185–193.

Sousa, T.N., Tarazona-Santos, E.M., Wilson, D.J., Madureira, A.P., Falcão, P.R.K., Fontes, C.J.F., Gil, L.H.S., Ferreira, M.U., Carvalho, L.H., Brito, C.F.A., 2010. Genetic variability and natural selection at the ligand domain of the Duffy binding protein in Brazilian *Plasmodium vivax* populations. Malar. J. 9 (334), www.malariajournal.com/content/9/1/334.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. *MEGA4:* Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tanabe, K., Escalante, A., Sakihama, N., Honda, M., Arisue, N., Horii, T., Culleton, R., Hayakawa, T., Hashimoto, T., Longacre, S., Pathirana, S., Handunnetti, S., Kishino, H., 2007. Recent independent evolution of msp1 polymorphism in *Plasmodium vivax* and related simian malaria parasites. Mol. Biochem. Parasitol. 156, 74–79.

Tjitra, E., Nicholas, M., Anstey, N.M., Sugiarto, P., Warikar, N., Kenangalem, E., Karyana1, M., Lampah, D.A., Price, R.N., 2008. Multidrug-resistant *Plasmodium vivax* associated serve and fatal malaria: a prospective study in Papua, Indonesia. PLoS Med. 5 (6), e128, doi:10.1371/journal.pmed.0050128.

Tran, T.M, Oliveira-Ferreira, J., Moreno, A., Santos, F., Yazdani, S.S., Chitnis, C.E., Altman, J.D., Meyer, E.V., Barnwell, J.W., Galinski, M.R., 2005. Comparison of IgG reactivities to *Plasmodium vivax* merozoite invasion antigens in a Brazilian Amazon population. Am. J. Trop. Med. Hyg. 73, 244–255.

VanBuskirk, K.M., Cole-Tobian, J.L., Baisor, M., Sevova, E.S., Bockarie, M., King, C.L., Adams, J.H., 2004. Antigenic drift in the ligand domain of *Plasmodium vivax* Duffy binding protein confers resistance to inhibitory antibodies. J. Infect. Dis. 190, 1556–1562.

Wickramarachchci, T., Premaratne, P.H., Dias, S., Handunnatti, S.M., Udagama-Randeniya, P.V., 2010. Genetic complexity of *Plasmodium vivax* infections in Sri Lanka, as reflected at the merozoite-surface-protein-3α locus. Ann. Trop. Med. Parasitol. 104 (2), 95–108.

Xainli, J., Baisor, M., Kastens, W., Bockarie, M., Adams, J.H., King, C.L., 2002. Age dependent cellular immune responses to *Plasmodium vivax* Duffy binding protein in humans. J. Immunol. 169, 3200–3207.

Xainli, J., Cole-Tobian, J.L., Baisor, M., Kastens, W., Bockarie, M., Yazdani, S.S., Chitnis, C.E., Adams, J.H., King, C.L., 2003. Epitope-specific humoral immunity to *Plasmodium vivax* Duffy Binding Protein. Infect. Immun. 71, 2508–2515.