Genetic Diversity of Plasmodium vivax Reticulocyte Binding Protein 2b in Global Parasite Populations

xuexing zhang
China Medical University
Haichao Wei
China Medical University
Yangminghui Zhang
China Medical University
Yan Zhao
China Medical University
Lin Wang
China Medical University
Yubing Hu
China Medical University
Wang Nguitragool
Mahidol University
Jetsumon Sattabongkot
Mahidol University
John Adams
University of South Florida College of Public Health
Liwang Cui
University of South Florida
Yaming Cao
China Medical University
Qinghui Wang (guihui2010@163.com)
China Medical University

Research

Keywords: Plasmodium vivax, PvRBP2b, genetic diversity, vaccine

Posted Date: November 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1023914/v1

Abstract

Background

Plasmodium vivax reticulocyte binding protein 2b (PvRBP2b) plays a critical role in parasite invasion of reticulocytes by binding the transferring receptor 1. PvRBP2b is a vaccine candidate since the antibody titers against PvRBP2b recombinant proteins are negatively correlated with the parasitemia and risk of *vivax* malaria. This study aims to analyze the genetic diversity of the *PvRBP2b* gene in the global *P. vivax* populations.

Methods

The near full-length *PvRBP2b* nucleotide sequences (190-8349 bp) were obtained from 88 *P. vivax* isolates collected from the China–Myanmar border (n=44) and Thailand (n=44). Additional 224 sequences of *PvRBP2b* were retrieved from genome sequences from the global parasite populations. The genetic diversity, neutral selections, haplotypes distribution and genetic differentiation of *PvRBP2b* were examined.

Results

The genetic diversity of *PvRBP2b* was distributed unevenly with the peak in the reticulocyte binding region in the N-terminus and subjected to the balancing selection. Several amino acid variants were found in all or nearly all endemic fields. However, the critical residues responsible for reticulocyte binding were highly conserved. There was substantial population differentiation according to the geographical separation. The distribution of haplotypes in the reticulocyte binding region varied among regions; even the two major haplotypes Hap_6 and Hap_8 were found in only five populations.

Conclusions

Our data showed considerable genetic variations of *PvRBPb* in global parasite populations, and the geographic divergence may pose a challenge to PvRBP2b-based vaccine development.

Background

Malaria remains a threat to global health despite intensified control efforts in recent years. As the most widespread *Plasmodium* species, *Plasmodium v*ivax caused an estimated 7.9 million worldwide in 2018, with 53% of the *vivax* burden being in the Southeast Asia Region[1]. It is challenging to eliminate *P. vivax* due to its dormant liver stage, which gives rise to the relapses of malaria [2, 3, 4]. Integrated interventions, including novel tools such as vaccines, are urgently needed for malaria elimination.

Invasion of the red blood cells (RBCs) by the merozoites is an essential step for the asexual erythrocytic cycle of malaria parasites [5, 6]. Since merozoites are exposed to the host immune system, vaccine candidates are usually designed to target the merozoite surface proteins to block erythrocyte invasion [7,

8]. Compared to *P. falciparum*, much fewer vaccine candidates have been identified for *P. vivax*, partially owing to the absence of a long-term *in vitro* culture system for this parasite [9, 10]. *P. vivax* species requires the Duffy antigen receptor for chemokines (DARC) on the RBC surface for invasion. However, solid evidence of *P. vivax* infection in the DARC-negative individuals in Africa suggests that *P. vivax* may have evolved to explore alternative pathways for invasion [11, 12]. *P. vivax* shows a restricted tropism for reticulocytes with high levels of transferrin receptor 1 (TfR1 or CD71) [13]. Recently, TfR1 has been identified as the reticulocyte-specific receptor for *P. vivax* reticulocyte-binding protein 2b (PvRBP2b), a ternary complex expressed in the schizont stage [14, 15]. PvRBP2b belongs to the PvRBP family composed of at least 11 members with different binding preferences for normocytes or reticulocytes [5, 16, 17, 18, 19]. The crystal structure of the N-terminal domain of PvRBP2b has revealed a similar structural scaffold to that of *P. falciparum* reticulocyte-binding protein homolog 5 (PfRh5), a well-characterized vaccine candidate for *P. falciparum* [15, 20]. Monoclonal antibodies against PvRBP2b or TfR1 mutants that impede the binding of PvRBP2b to the reticulocytes successfully blocked the entry of *P. vivax* into the reticulocytes, suggesting PvRBP2b as a promising vaccine candidate targeting blood-stage infections [15].

The N-terminal domain of PvRBP2b is responsible for reticulocyte binding [14, 15], whereas the function of the C-terminal domain was not clear. With the recombinant PvRBP2b N-terminal domain, antibodies against PvRBP2b have been detected in *P. vivax* patient plasma samples, supporting that PvRBP2b contains immune recognition epitopes in the N-terminal domain [15, 18, 21]. Furthermore, the IgG levels against PvRBP2b were negatively correlated with parasitemia and the risk of vivax infections [18, 21, 22]. These studies highlight PvRBP2b as a promising target for *P. vivax* vaccine development [14, 15].

A major challenge for the efficacy of blood-stage malaria vaccines is the extensive genetic diversity of the target antigens. Therefore, understanding the genetic diversity of vaccine candidates is necessary for designing effective vaccines and predicting vaccine efficacy. In this study, we analyzed the genetic diversity, phylogenetic relationship, and population differentiation of *PvRBP2b* from 312 global *P. vivax* isolates, aiming to provide the necessary information for PvRBP2b-based vaccine development.

Methods

Sample collection and ethics statements

This study used 88 dried blood spots on filter papers collected from *P. vivax* patients attending the Laiza and Nabang hospitals in the China–Myanmar border area in 2014 (n=44) and the Tha Song Yang hospital in western Thailand in 2011-2012 (n=44). Malaria was diagnosed by microscopy, and finger-prick blood samples were collected.

PvRBP2b **gene amplification and sequencing**

Genomic DNA was extracted from dried blood spots on lter papers using the QIAamp DNA Mini kit (Qiagen, Hilden, Germany). Of the full-length protein-coding sequence of the *PvRBP2b* gene (8421bp), an

8160 bp fragment (*PvRBP2b*$_{190\text{-}8349}$), corresponding to the amino acids 64-2783 of the PvRBP2b protein was amplified from all the samples using KOD-Plus-Neo polymerase (Toyobo, Osaka, Japan). Given the large size of the gene, seven overlapping 1.5 kb fragments were amplified from each sample using seven pairs of primers (Table S1). PCR reactions were performed in 30 µl volume containing 1×KOD-Plus-Neo buffer, 200 µM dNTPs, 1 mM $MgSO_4$, 250 nM primer, 0.4 units KOD Plus polymerase, and 2 µl genomic DNA. The following cycling parameters were used: an initial denaturation at 94°C for 5 min, 35 cycles of denaturation at 94°C for 15 s, annealing at a determined temperature (Table S1) for 15 s, and extension at 68°C for 90 s, followed by a final extension at 68°C for 5 min. The PCR products were separated in 1% agarose gels and then subjected to DNA sequencing using the ABI BigDye™ Terminator Reaction Ready kit (Applied Biosystems, CA, USA).

Sequence assembly and retrieval

The 88 *PvRBP2b* sequences were assembled using the DNASTAR program (Lasergene). In addition, 224 *PvRBP2b* sequences from eleven global locations were obtained from previous whole-genome sequencing projects [23, 24, 25, 26, 27, 28]. Fastq files were downloaded from the Sequence Read Archive (SRA) of the National Center for Biotechnology Information. We used two parameters to exclude low-quality variants: quality ≤40, minor allele frequency less than 0.01. Only SNP variants in single infection were included according to separate criteria. After filtering out unqualified sequences, the global sample set includes Brazil (n=36), Colombia (n=28), Cambodia (n=35), China–Myanmar border (n=19), Ethiopia (n=18), Indonesia (n=3), Laos (n=2), Malaysia (n=4), Papua New Guinea (PNG) (n=20), Thailand (n=48), and Vietnam (n=11). Isolate codes and SRA accession numbers of samples used in this analysis are given in Table S2.

Analysis of genetic diversity and tests for detecting selections

A total of 312 *PvRBP2b* sequences were aligned with the reference sequence from the Salvador I (Sal I) strain (PVX_094255) using the Clustal W program in the MEGA7 software. For the evaluation of *PvRBP2b* genetic diversity, the nucleotide diversity (π), the number of haplotypes (H), and haplotype diversity (Hd) were computed using the DnaSP v5.10 software [29]. To test the departure from neutrality, Tajima's *D* test [30], Fu & Li's *F** test [31], and Fu & Li's *D** test [31] were computed using DnaSP v5.10 software. The McDonald-Kreitman (MK) test was performed to evaluate the departure from neutrality using the *P. cynomolgi* sequence (*PcyRBP2b*; PCYB_081060) as the outgroup [32]. Fisher's exact test was applied to assess statistical significance ($p < 0.05$).

Natural selection was determined by calculating the ratio of nonsynonymous (dN) to synonymous (dS) substitutions per nucleotide site (dN-dS), using the Nei-Gojobori method [33] with Jukes-Cantor correction for multiple substitutions. Statistical significance of the difference was estimated using the codon-based Z-test of selection implemented in MEGA 7 [33].

Finally, to determine the existence of specific codons targeted by selection in the global population [34], three maximum likelihood codon-based tests in the HyPhy package-based algorithms, SLAC [35], FEL [35], and FUBAR [36] implemented in the Datamonkey webserver (http://www.datamonkey.org) were performed.

Population differentiation, structure and phylogenetic relationship

To investigate population subdivision, Wright's fixation index (F_{ST}) representing inter-population variance in allele frequencies was calculated using DnaSP v5.10 [29, 37, 38]. The genetic structure of all the *vivax* parasite populations was then elucidated using the STRUCTURE v2.3.2 software based on the Bayesian analysis and admixture model [39, 40]. All samples were run at K = 2–7 (10 iterations each) with a burn-in period of 20000 iterations followed by 1200000 Markov Chain Monte Carlo (MCMC) iterations. Then the optimal number of grouping was determined by ΔK using the STRUCTURE HARVESTER v0.6.94 software [41, 42]. The partition of the clusters was presented using CLUMPP v1.1.2 [43] and the DISTRUCT 1.1 tools [44]. To determine the relationship among the parasites, phylogenetic analyses were performed using the Neighbor-Joining method implemented in MEGA7 [33]. Then the phylogenetic tree was optimized with the online tool ITOL [45]. A haplotype network based on the polymorphic sites in the reticulocyte-binding region of *PvRBP2b* was constructed using the PHYLOVIZ 2.0 software with the Neighbor-Joining method [46].

Prediction of linear B cell epitopes

The potential linear B-cell epitopes were predicted using the BCPreds prediction tool (http://ailab-projects1.ist.psu.edu:8080/bcpred/predict.html) [47]. Antigenicity was predicted using the VaxiJen v2.0 online tool (http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html) [48]. BCPreds predicts a peptide length of 12 consecutive amino acids with a threshold of 0.8, whereas VaxiJen sets a threshold at 0.5. The overlapped regions of predicted linear B cell epitopes by both methods were selected. The predicted 3D structures of the reticulocyte-binding region of PvRBP2b were constructed with the PHYRE2 algorithm [49] and further visualized and modeled using the molecular modeling tool PyMOL V2.3 [50]. The PvRBP2b amino acid sequence in the reference Sal I strain was used for prediction.

Results

Mutations revealed from global PvRBP2b sequences

Sequencing of the 8160 bp *PvRBP2b* fragment (190–8349 bp) was successful for the 88 *P. vivax* field isolates collected from the China–Myanmar and Thailand–Myanmar border areas. To gain a global perspective, we retrieved 224 *PvRBP2b* sequences from the whole genome sequences of *P. vivax* isolates collected in multiple *P. vivax*-endemic areas of the world. Alignment of all 312 *PvRBP2b* sequences with the Sal I reference identified 116 single nucleotide polymorphisms (SNPs), including 96 nonsynonymous and 20 synonymous mutations. All nonsynonymous mutations with allele frequencies of more than 1% in

different areas are shown in Fig. 1 and Table S3. The E136K, N349K, K363E, D366V/H, V395A/T, K412N, Q564R, D917E, N1529K, K1606E, E2265K, and E2746G amino acid mutations were found in all or nearly all endemic sites, reflecting the prevalent *PvRBP2b* polymorphisms in the world. Among them, D917E approached fixation (95.8%).

The reticulocyte binding region on the N-terminus of PvRBP2b is the critical portion for the receptor binding and RBC invasion. We refer it to the amino acid residues 168-633 since its flanking regions failed to be visualized in the PvRBP2b structure [14]. As shown in Table S3, from a total of 75 nonsynonymous nucleotide substitutions with allele frequencies > 1%, almost 50% (35/75) were accumulated in the reticulocyte binding region, generating 28 nonsynonymous amino acid mutations. Besides the 11 nonsynonymous mutations (R217H, R242T/S, K288P, K309Q, K363E, D366V/H, G382R/E, E497K, D558E, Q564R, and N591K) reported for the reticulocyte binding region [14], 17 additional nonsynonymous mutations (D220Y, T224R/K, S228P, E232K, L293V, N300K/D, D315Y, N349K, V395A/T, K412N, Q413E, K437E, H455Q, K575E, D578H, S586R, and E631K) were reported in this study (Fig. 2a). Among them, K363E and S586R are the residues interacting with the receptor TfR1 [14]. K363E was prevalent in all endemic areas with an allele frequency of 40.4%, whereas S586R was found only in Brazil (22.2%), Ethiopia (11.1%), and Thailand (1.1%) (Table S3).

Genetic diversity of PvRBP2b

We analyzed the population genetic indices to assess the nucleotide diversity of *PvRBP2b* (Table 1). The overall nucleotide diversity (π) from 312 sequences was 0.00196, with the highest found in the population from Malaysia (0.00203), followed by Thailand (0.00198), and Vietnam (0.00193). The overall haplotype diversity was high (0.997), with the highest found in Malaysia, Vietnam, PNG, Indonesia, and Laos (Table 1). PNG showed much lower nucleotide diversity (0.00114) but high haplotype diversity (1.000) (Table 1). The sliding window plots of nucleotide diversity revealed an uneven distribution with the peaks located at nucleotides 1015–1134 within the reticulocyte binding region (Fig. 3a). Similarly, 46.7% amino acid substitutions were clustered in the reticulocyte binding region (Fig. 3e). This result reflected relatively high polymorphism in the reticulocyte binding region of *PvRBP2b* in the worldwide *P. vivax* populations.

Table 1
Genetic diversity of *PvRBP2b* near-full length in global populations

Populations	No. isolates	Polymorphic sites	π ± SD	H	Hd±SD
Brazil	36	30	0.00121±0.00089	18	0.930±0.023
Colombia	28	29	0.00112±0.00091	18	0.950±0.025
Cambodia	35	55	0.00172±0.00167	30	0.987±0.012
China-Myanmar	63	81	0.00178±0.00213	45	0.973±0.012
Ethiopia	18	32	0.00120±0.00114	16	0.980±0.028
Laos	2	9	0.00110±0.00110	2	1.000±0.500
Malaysia	4	31	0.00203±0.00208	4	1.000±0.177
Papua New Guinea	20	39	0.00114±0.00135	20	1.000±0.016
Thailand	92	97	0.00198±0.00244	85	0.998±0.002
Vietnam	11	43	0.00193±0.00184	11	1.000±0.039
Indonesia	3	16	0.00131±0.00131	3	1.000±0.272
Total	312	147	0.00196±0.00299	248	0.997±0.0008
H: number of haplotypes, Hd: haplotype diversity, SD Standard deviation					

Evidence of potential selections

Neutrality tests were conducted to evaluate whether the *PvRBP2b* gene followed the neutral equilibrium model of molecular evolution. Although the overall *PvRBP2b* sequence did not significantly deviate from neutrality, a sliding window analysis identified significant positive values for the 1015–1034 bp region by Tajima's *D**, and Fu and Li's *F** tests, which paralleled the peak π value, reflecting balancing selection within the reticulocyte binding region in the global populations (Fig. 3b-d). In addition, some significant negative values for several C-terminal domains suggested the population expansion or excess of singletons (Fig. 3b-d).

Departure from neutrality was further evaluated using the MK test with the *P. cynomolgi RBP2b* sequence as the outgroup [32]. The results showed significantly more intraspecific nonsynonymous substitutions over synonymous substitutions than interspecific fixed differences in both the entire sequenced region (NI=1.874, p=0.005753) and reticulocyte-binding region (NI=6.134, p=0.004504) (Table 2). The dN-dS statistic was consistently positive for the reticulocyte-binding region among the global populations (Table 2), suggesting that polymorphisms found for this region of *PvRBP2b* were maintained by diversifying selection.

Table 2
Tests for selection in *PvRBP2b* gene from global samples

Gene fragment encoding *PvRBP2b*	N	MK test		Codon-based Z-test	
		NI	*p-value*	dN-dS	*p-value*
Near full-length	312	1.874	0.005753**	0.995	0.322
Reticulocyte binding region	312	6.134	0.004504**	2.801**	0.006
N: Number of isolates; NI: neutrality index; *, $p<0.05$ **, $p<0.01$					

Three likelihood-based algorithms were used to determine specific codons targeted by selection. Thirty-one positively selected and 12 negatively selected codons were identified by all three algorithms (Table 3). Twelve positively selected mutations are located in the reticulocyte-binding region, presumptively associated with parasite invasion and/or immune recognition (Table 3). K363 and S586 have been shown as the reticulocyte-binding sites [14]. In contrast to the positive selection at K363 confirmed by all three methods, S586 was supposed to be positively selected only by the FUBAR method (Table 3). Meanwhile, sites under purifying selection were scattered along the PvRBP2b (Table 3).

Table 3
Codon-based tests for selection in *PvRBP2b* gene

	SLAC	FEL	FUBAR	By all three algorithms
positive/diversifying selection sites	136, 217, 224, 242, 288, 349, 363, 382, 395, 412, 564, 575, 591, 1181, 1381, 1510, 1529, 1606, 1870, 2073, 2190, 2200, 2221, 2250, 2261, 2265, 2278, 2318, 2741, 2746, 2750	136, 217, 220, 224, 242, 288, 293, 315, 349, 363, 382, 395, 412, 413, 558, 564, 575, 591, 917, 1181, 1239, 1381, 1510, 1529, 1577, 1606, 1870, 1984, 2073, 2190, 2200, 2221, 2236, 2250, 2261, 2265, 2278, 2318, 2642, 2741, 2746, 2750	136, 217, 220, 224, 228, 232, 242, 288, 293, 300, 309, 315, 349, 363, 382, 395, 412, 413, 437, 455, 497, 558, 564, 575, 578, 586, 591, 631, 666, 917, 1168, 1173, 1181, 1239, 1289, 1381, 1510, 1520, 1529, 1577, 1606, 1870, 1907, 1984, 2073, 2190, 2200, 2201, 2221, 2236, 2250, 2261, 2265, 2272, 2278, 2318, 2382, 2393, 2517, 2605, 2613, 2642, 2738, 2741, 2746, 2747, 2750	136, 217, 224, 242, 288, 349, 363, 382, 395, 412, 564, 575, 591, 1181, 1381, 1510, 1529, 1606, 1870, 2073, 2190, 2200, 2221, 2250, 2261, 2265, 2278, 2318, 2741, 2746, 2750
negative/purifying selection sites	885, 1113, 1644, 1834, 1886, 2149, 2225, 2238, 2247, 2255, 2292, 2329, 2391, 2465	243, 670, 885, 1113, 1644, 1834, 1886, 1917, 1987, 2149, 2225, 2238, 2247, 2255, 2292, 2329, 2391, 2465	885, 1113, 1644, 1834, 1886, 1987, 2149, 2247, 2255, 2292, 2329, 2391, 2465	885, 1113, 1644, 1834, 1886, 2149, 2247, 2255, 2292, 2329, 2391, 2465

The mutations within the 3D structure and predicted B-epitopes in the reticulocyte-binding region

The 3D model structure with positively selected amino acid mutations mapped in the reticulocyte-binding region of PvRBP2b showed that most mutated residues were on the surface of α helices except V395 hidden inside the protein and K412 located in the flexible loop structure (Fig. 2b). Most mutations were close to the hydrophobic binding region. Three peptides (270–281, 519–530, and 566–577) were predicted to be linear B-epitopes with the prediction scores of 0.839, 0.987, and 0.959, respectively, by the BCPreds and VaxiJen software (Table 4). However, only one polymorphic residue, K575, was presented in the predicted B-epitope of 566–577.

Table 4
Predicted *PvRBP2b* B cell epitopes in the reticulocyte binding region

Start Position	Stop Position	Sequence	Score
270	281	KLRQYEEKKEAF	0.839
519	530	NEFKKDYDNNVE	0.987
566	577	NIPANSNAQKKV	0.959

Haplotype network analysis

The 17 SNPs with minor allele frequencies of >5% in the *PvRBP2b* reticulocyte-binding region were chosen for haplotype analysis, which yielded a total of 114 haplotypes amongst the global dataset of 312 sequences (Table S4 and Fig. 4-5). Of the 114 haplotypes, 58 (50.9%) were region-specific, with a frequency between 0.32% and 1.60%. There were many rare haplotypes; 92.1% (105/114) of the haplotypes were shared by no more than five parasite isolates, among which 70/114 were represented by single parasite isolates. Even the two major haplotypes Hap_6 and Hap_8, with a frequency of 14.74% and 12.82%, respectively, were only found in five regional populations. Hap_8, the same as the Sal I haplotype, was most abundant in the China–Myanmar border (39.7%) and Colombia (42.9%). Hap_6, the predominant haplotype identified in Thailand, was also shared among the parasite populations from the China–Myanmar border, Cambodia, Brazil, and Colombia. The haplotypes in Asia were distributed in all global parasite populations, and Thailand harbored the highest haplotype diversity (52/114) (Table S4 and Fig. 4, 5). In contrast, the distribution pattern shown in the haplotype network was different from parasite populations from South America (Brazil and Colombia), Africa (Ethiopia), and Oceania (PNG) (Fig. 5). Sixteen out of 20 parasite isolates in PNG were from a unique haplotype restricted to this specific region (Table S4 and Fig. 4).

Population structure and differentiation

Analysis of the *PvRBP2b* sequences showed that the global parasites were optimally grouped into three clusters (K = 3) (Fig. 6a). These clusters were unevenly distributed among different geographic regions. The parasites from PNG and Indonesia were mainly represented by the purple cluster, whereas the parasites from Ethiopia, Brazil, and Colombia occupied the green cluster. The red cluster represented parasites predominantly from the China–Myanmar border and Thailand. Interestingly, the remaining parasites from the China–Myanmar border and Thailand and parasites from other countries of the Greater Mekong Subregion (GMS) (Cambodia, Laos, and Vietnam) showed genetic mixing of the green and purple clusters. When structure analysis was conducted using the *PvRBP2b* reticulocyte-binding region, six clusters were identified with overlapping, worldwide distribution (Fig. 6b). It is noteworthy that only the parasites from the PNG and Indonesia formed genetically distinct clusters from other geographic regions.

The genetic differentiation between two parasite populations was also evaluated via Wright's fixation index (F_{ST}) using the entire *PvRBP2b* sequence and reticulocyte-binding region, respectively. The heatmap of the F_{ST} values from both analyses revealed population differentiation patterns consistent with the Structure analysis (Fig. 7). Consistent with the principle of isolation by distance, parasite populations were genetically similar within each continent (e.g., Brazil vs. Colombia, PNG vs. Indonesia, countries within the GMS). In contrast, considerable differentiation was detected between populations from different continents. Notably, the PNG and Indonesia *P. vivax* populations had high levels of genetic differentiation from the rest of the global parasite populations. In comparison, parasite populations from the GMS were moderately differentiated from the South American parasite populations. Interestingly, parasites from Africa (represented by Ethiopia) showed little genetic differentiation from the GMS and South American parasites.

Discussion

This study evaluated the genetic diversity of *PvRBP2b* as a *P. vivax* malaria blood-stage vaccine candidate. The overall nucleotide diversity of *PvRBP2b* from the global samples was modest (0.00196), much lower than that of the highly polymorphic surface antigens, such as *PvAMA1* and *PvMSP1* [51, 52]. The genetic polymorphisms of *PvRBP2b* were geographically heterogeneous, with higher diversity found in the GMS countries than in South America (Brazil and Colombia), Africa (Ethiopia), and Oceania (PNG). Such a genetic diversity distribution pattern was also observed in *PvAMA1*, and it seemed to correlate with the larger effective population size in Southeast Asia [53]. It is speculated that *P. vivax* in Southeast Asia was the possible source population [54]. The high transmission intensity and the frequent migrations of infected people among Southeast Asian countries might result in the large effective population size in this continent. Recently, continuous malaria control and elimination strategies had successfully reduced the parasite incidence but seemed to have little short-term effect on the *P. vivax* population size [55, 56]. The relatively high levels of asymptomatic vivax infections in endemic populations possibly contributed to the maintenance of the effective population size [57, 58].

PvRBP2b-mediated the reticulocyte invasion depends on the *PvRBP2b*–TfR1–Tf ternary complex [14]. Consistent with the previous reports [15], the reticulocyte binding region was under the balancing selection. It accumulated most of the nonsynonymous amino acid substitutions and had the highest diversity. Interestingly, most of the residues on the reticulocyte binding region were conserved. At least three residues (Y542, K600, and Y604) on PvRBP2b were critical to the reticulocyte binding and complex formation because mutations at these sites reduced binding affinity by around 80% [14]. These amino acids were absolutely conserved among all 312 samples, reflecting their essential function in reticulocyte binding. Most of the identified residues making contact with TfR1/Tf from structure analysis [14] were conserved except S586R and K363E which had low (3.5%) and moderate (40.26%) allele frequencies, respectively. S586R had a limited distribution in Brazil, Ethiopia, and Thailand, whereas K363E occurred in all endemic areas. Additional nonsynonymous SNPs of low-modest frequencies (1.6–51%) were detected surrounding the interaction sites. Several residues within the reticulocyte binding region were under

positive selection, raising the possibility that they were the result of the host immune selection [59]. Moreover, most of the positively selected residues are located on the α helices, some being close to the hydrophobic binding region. The K412N mutation had a frequency of 51% and resided in the flexible loop structure. Only one mutation, K575E, was mapped to the predicted linear B-cell epitope. However, it is still unclear whether these specific polymorphic residues would change protein structure by altering protein polarity and hydrophilicity, therefore adapting to the different TfR1 mutants (e.g., L212V, N348A, and S412G) [14, 60, 61]. More functional investigations are required to answer these questions.

One of the obstacles hindering the development of a successful malaria vaccine is the extensive genetic diversity of blood-stage antigens in *P. vivax* [62]. Naturally acquired antibodies against the reticulocyte binding region of PvRBP2b showed a strong association with the reduced parasitemia [18, 21, 22], highlighting the vaccine potential of PvRBP2b. Although the overall genetic diversity of *PvRBP2b* was not as high as that of leading blood-stage vaccine candidates, *PvAMA1* and *PvMSP1*, which have advanced to clinical trials [63], the reticulocyte binding region of PvRBP2b deserves further attention. This region carried almost 50% of the mutations in the entire PvRBP2b, and analysis of the global parasite samples identified 114 haplotypes. Furthermore, most haplotypes were region-specific and represented by a single parasite isolate, and very few haplotypes were shared worldwide. Even the predominant haplotypes Hap_6 (14.74%) and Hap_8 (12.82%) were distributed in only five regions. Of note, PNG showed distinct haplotype patterns from other worldwide populations. This enormous haplotype diversity may present a challenge to developing a PvRBP2b-based vaccine. Since an effective vaccine should include most of the common alleles relevant to the induction of immune responses to ensure sufficient coverage of the genetic diversity [64, 65], it is essential to determine whether the major PvRBP2b alleles confer strain-transcending immunity.

This study provides several lines of evidence confirming the geographical separation of global *P. vivax* populations. Pairwise F_{ST} comparison identified the considerably high levels of differentiation between the Oceania (PNG and Indonesia) parasite populations from other endemic regions, whereas parasite populations in the GMS were much less differentiated. Interestingly, populations from South America appeared to be closely related to those from Africa (Ethiopia). Population structure analysis further reinforced such a finding on population relatedness. The population differentiation pattern identified here was in agreement with other population studies using individual genes [51], microsatellites [54, 66], or whole genome sequences [23, 24, 67]. The distinct parasite genetic structure in PNG may reflect the limited gene flow between PNG and the rest of the world, while the unique RBC polymorphisms in the human populations may have also contributed to this difference [53, 66, 68, 69]. In the GMS, however, the high transmission intensity in some border areas and frequent host migrations among the different countries are likely responsible for the panmixia of parasite populations. Population genetics can help assess the effects of malaria control strategies, track the source of imported infections, and inform the vaccine design. Therefore, continuous monitoring of the population genetic structures will allow longitudinal evaluation of the efficacy of the malaria control strategy both locally and globally.

Conclusion

In conclusion, this study revealed a remarkably high level of genetic polymorphisms in the *PvRBP2b* among the global *P. vivax* populations, with mutations clustered in the reticulocyte binding region. The genetic differentiation of parasite populations among different continents were remarkable, suggesting a potential need to cover major protein variants in case of strain-transcending immunity. Future studies addressing the functions of antibodies against different PvRBP2b variants are warranted.

Abbreviations

RBCs: Red blood cells; DARC: Duffy antigen receptor for chemokines; TfR1: Transferrin receptor 1; PfRh5: *P. falciparum* reticulocyte-binding protein homolog 5; PNG: Papua New Guinea; PCR: Polymerase chain reaction; SNPs: Single nucleotide polymorphisms; SRA: Sequence Read Archive; π: The nucleotide diversity; H: the number of haplotypes; Hd: haplotype diversity; MK: McDonald-Kreitman; F_{ST}: Wright's fixation index; PvAMA1: *Plasmodium vivax* apical membrane antigen; PvMSP: *Plasmodium vivax* Merozoite surface protein 1; GMS: Greater Mekong Subregion.

Declarations

Acknowledgments

We would like to thank the patients participating in this study and providing the blood samples.

Funding

This work was supported by grants (U19AI089672 and U01AI155361) from the National Institute of Allergy and Infectious Diseases, NIH, USA.

Availability of data and materials

The data supporting the conclusion of this article are included within the article.

Authors' contributions

XZ performed the experiment study and drafted the manuscript. HW and YMZ, participated in data analisis.YZ, LW, YH and JA participated in sample collection and sorting. WN, JS, LC,YC and QW conceived the study and revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Written informed consent was obtained from the adult patients and the guardians of minor participants. The study protocol was approved by the ethics review boards of relevant institutions, and the use of the anonymized samples was approved by the institutional review board of China Medical University.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing intrests.

Author details

[1] Department of Immunology, College of Basic Medical Science, China Medical University, Shenyang 110122, Liaoning, China. [2] Department of Blood Transfusion Medicine, General Hospital of Northern Theater Command, Shenyang 110015, Liaoning, China. [3] Department of Blood Transfusion, Yantaishan Hospital, Yantai 264000, Shandong , China. [4] Central Laboratory, the first hospital of China Medical University, Shenyang 110001, Liaoning, China. [5] Mahidol Vivax Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. [6] Global Health Infectious Disease Research (GHIDR) Program, College of Public Health, Tampa, Florida, USA. [7] Department of Internal Medicine, Morsani College of Medicine, University of South Florida, 3720 Spectrum Boulevard, Suite 304, Tampa, FL 33612, USA.

References

1. WHO. World malaria report 2019. Geneva: World Health Organization; 2019. https://www.who.int/publications-detail-redirect/9789241565721.

2. Baird JK. Resistance to therapies for infection by Plasmodium vivax. Clin Microbiol Rev. 2009;22 3:508-34; doi: 10.1128/CMR.00008-09. https://www.ncbi.nlm.nih.gov/pubmed/19597012.

3. Imwong M, Snounou G, Pukrittayakamee S, Tanomsing N, Kim JR, Nandy A, et al. Relapses of Plasmodium vivax infection usually result from activation of heterologous hypnozoites. J Infect Dis. 2007;195 7:927-33; doi: 10.1086/512241. https://www.ncbi.nlm.nih.gov/pubmed/17330781.

4. Price RN, von Seidlein L, Valecha N, Nosten F, Baird JK, White NJ. Global extent of chloroquine-resistant Plasmodium vivax: a systematic review and meta-analysis. Lancet Infect Dis. 2014;14 10:982-91; doi: 10.1016/S1473-3099(14)70855-2. https://www.ncbi.nlm.nih.gov/pubmed/25213732.

5. Singh S, Chitnis CE. Molecular Signaling Involved in Entry and Exit of Malaria Parasites from Host Erythrocytes. Cold Spring Harb Perspect Med. 2017;7 10. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=28507195&query_hl=1.

6. Singh S, Alam MM, Pal-Bhowmick I, Brzostowski JA, Chitnis CE. Distinct external signals trigger sequential release of apical organelles during erythrocyte invasion by malaria parasites. PLoS

Pathog. 2010;6 2:e1000746. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=20140184&query_hl=1.

7. Mueller I, Shakri AR, Chitnis CE. Development of vaccines for Plasmodium vivax malaria. Vaccine. 2015;33 52:7489-95; doi: 10.1016/j.vaccine.2015.09.060. https://www.ncbi.nlm.nih.gov/pubmed/26428453.

8. Tham WH, Beeson JG, Rayner JC. Plasmodium vivax vaccine research - we've only just begun. Int J Parasitol. 2017;47 2-3:111-8; doi: 10.1016/j.ijpara.2016.09.006. https://www.ncbi.nlm.nih.gov/pubmed/27899329.

9. Noulin F, Borlon C, Van Den Abbeele J, D'Alessandro U, Erhart A. 1912-2012: a century of research on Plasmodium vivax in vitro culture. Trends Parasitol. 2013;29 6:286-94; doi: 10.1016/j.pt.2013.03.012. https://www.ncbi.nlm.nih.gov/pubmed/23623759.

10. Udomsangpetch R, Kaneko O, Chotivanich K, Sattabongkot J. Cultivation of Plasmodium vivax. Trends Parasitol. 2008;24 2:85-8; doi: 10.1016/j.pt.2007.09.010. https://www.ncbi.nlm.nih.gov/pubmed/18180202.

11. Howes RE, Reiner RC, Jr., Battle KE, Longbottom J, Mappin B, Ordanovich D, et al. Plasmodium vivax Transmission in Africa. PLoS Negl Trop Dis. 2015;9 11:e0004222; doi: 10.1371/journal.pntd.0004222. https://www.ncbi.nlm.nih.gov/pubmed/26587988.

12. Mercereau-Puijalon O, Menard D. Plasmodium vivax and the Duffy antigen: a paradigm revisited. Transfus Clin Biol. 2010;17 3:176-83; doi: 10.1016/j.tracli.2010.06.005. https://www.ncbi.nlm.nih.gov/pubmed/20655790.

13. Malleret B, Li A, Zhang R, Tan KS, Suwanarusk R, Claser C, et al. Plasmodium vivax: restricted tropism and rapid remodeling of CD71-positive reticulocytes. Blood. 2015;125 8:1314-24; doi: 10.1182/blood-2014-08-596015. https://www.ncbi.nlm.nih.gov/pubmed/25414440.

14. Gruszczyk J, Huang RK, Chan LJ, Menant S, Hong C, Murphy JM, et al. Cryo-EM structure of an essential Plasmodium vivax invasion complex. Nature. 2018;559 7712:135-9; doi: 10.1038/s41586-018-0249-1. https://www.ncbi.nlm.nih.gov/pubmed/29950717.

15. Gruszczyk J, Kanjee U, Chan LJ, Menant S, Malleret B, Lim NTY, et al. Transferrin receptor 1 is a reticulocyte-specific receptor for Plasmodium vivax. Science. 2018;359 6371:48-55. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=29302006&query_hl=1.

16. Galinski MR, Medina CC, Ingravallo P, Barnwell JW. A reticulocyte-binding protein complex of Plasmodium vivax merozoites. Cell. 1992;69 7:1213-26. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=1617731&query_hl=1.

17. Gruszczyk J, Lim NT, Arnott A, He WQ, Nguitragool W, Roobsoong W, et al. Structurally conserved erythrocyte-binding domain in Plasmodium provides a versatile scaffold for alternate receptor engagement. Proc Natl Acad Sci U S A. 2016;113 2:E191-200. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=26715754&query_hl=1.

18. He WQ, Karl S, White MT, Nguitragool W, Monteiro W, Kuehn A, et al. Antibodies to Plasmodium vivax reticulocyte binding protein 2b are associated with protection against P. vivax malaria in populations living in low malaria transmission regions of Brazil and Thailand. PLoS Negl Trop Dis. 2019;13 8:e0007596. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=31425514&query_hl=1.

19. Miller LH, Mason SJ, Clyde DF, McGinniss MH. The resistance factor to Plasmodium vivax in blacks. The Duffy-blood-group genotype, FyFy. N Engl J Med. 1976;295 6:302-4. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=778616&query_hl=1.

20. Crosnier C, Bustamante LY, Bartholdson SJ, Bei AK, Theron M, Uchikawa M, et al. Basigin is a receptor essential for erythrocyte invasion by Plasmodium falciparum. Nature. 2011;480 7378:534-7; doi: 10.1038/nature10606. https://www.ncbi.nlm.nih.gov/pubmed/22080952.

21. Franca CT, He WQ, Gruszczyk J, Lim NT, Lin E, Kiniboro B, et al. Plasmodium vivax Reticulocyte Binding Proteins Are Key Targets of Naturally Acquired Immunity in Young Papua New Guinean Children. PLoS Negl Trop Dis. 2016;10 9:e0005014. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=27677183&query_hl=1.

22. Hietanen J, Chim-Ong A, Chiramanewong T, Gruszczyk J, Roobsoong W, Tham WH, et al. Gene Models, Expression Repertoire, and Immune Response of Plasmodium vivax Reticulocyte Binding Proteins. Infect Immun. 2015;84 3:677-85. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=26712206&query_hl=1.

23. Diez Benavente E, Campos M, Phelan J, Nolder D, Dombrowski JG, Marinho CRF, et al. A molecular barcode to inform the geographical origin and transmission dynamics of Plasmodium vivax malaria. PLoS Genet. 2020;16 2:e1008576; doi: 10.1371/journal.pgen.1008576. https://www.ncbi.nlm.nih.gov/pubmed/32053607.

24. Brashear AM, Fan Q, Hu Y, Li Y, Zhao Y, Wang Z, et al. Population genomics identifies a distinct Plasmodium vivax population on the China-Myanmar border of Southeast Asia. PLoS Negl Trop Dis. 2020;14 8:e0008506; doi: 10.1371/journal.pntd.0008506. https://www.ncbi.nlm.nih.gov/pubmed/32745103.

25. Ford A, Kepple D, Abagero BR, Connors J, Pearson R, Auburn S, et al. Whole genome sequencing of Plasmodium vivax isolates reveals frequent sequence and structural polymorphisms in erythrocyte

binding genes. PLoS Negl Trop Dis. 2020;14 10:e0008234; doi: 10.1371/journal.pntd.0008234. https://www.ncbi.nlm.nih.gov/pubmed/33044985.

26. Auburn S, Getachew S, Pearson RD, Amato R, Miotto O, Trimarsanto H, et al. Genomic Analysis of Plasmodium vivax in Southern Ethiopia Reveals Selective Pressures in Multiple Parasite Mechanisms. J Infect Dis. 2019;220 11:1738-49; doi: 10.1093/infdis/jiz016. https://www.ncbi.nlm.nih.gov/pubmed/30668735.

27. de Oliveira TC, Rodrigues PT, Menezes MJ, Goncalves-Lopes RM, Bastos MS, Lima NF, et al. Genome-wide diversity and differentiation in New World populations of the human malaria parasite Plasmodium vivax. PLoS Negl Trop Dis. 2017;11 7:e0005824; doi: 10.1371/journal.pntd.0005824. https://www.ncbi.nlm.nih.gov/pubmed/28759591.

28. Hupalo DN, Luo Z, Melnikov A, Sutton PL, Rogov P, Escalante A, et al. Population genomics studies identify signatures of global dispersal and drug resistance in Plasmodium vivax. Nat Genet. 2016;48 8:953-8; doi: 10.1038/ng.3588. https://www.ncbi.nlm.nih.gov/pubmed/27348298.

29. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25 11:1451-2. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=19346325&query_hl=1.

30. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123 3:585-95. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=2513255&query_hl=1.

31. Fu YX, Li WH. Statistical tests of neutrality of mutations. Genetics. 1993;133 3:693-709. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=8454210&query_hl=1.

32. McDonald JH, Kreitman M. Adaptive protein evolution at the Adh locus in Drosophila. Nature. 1991;351 6328:652-4; doi: 10.1038/351652a0. https://www.ncbi.nlm.nih.gov/pubmed/1904993.

33. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33 7:1870-4. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=27004904&query_hl=1.

34. Poon AF, Frost SD, Pond SL. Detecting signatures of selection from DNA sequences using Datamonkey. Methods Mol Biol. 2009;537:163-83; doi: 10.1007/978-1-59745-251-9_8. https://www.ncbi.nlm.nih.gov/pubmed/19378144.

35. Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol. 2005;22 5:1208-22; doi: 10.1093/molbev/msi105. https://www.ncbi.nlm.nih.gov/pubmed/15703242.

36. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol Biol Evol. 2013;30 5:1196-205; doi: 10.1093/molbev/mst030. https://www.ncbi.nlm.nih.gov/pubmed/23420840.

37. Cockerham CC. Analyses of gene frequencies. Genetics. 1973;74 4:679-700. https://www.ncbi.nlm.nih.gov/pubmed/17248636.

38. Cockerham CC, Weir BS. Covariances of relatives stemming from a population undergoing mixed self and random mating. Biometrics. 1984;40 1:157-64. https://www.ncbi.nlm.nih.gov/pubmed/6733226.

39. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155 2:945-59. https://www.ncbi.nlm.nih.gov/pubmed/10835412.

40. Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour. 2009;9 5:1322-32; doi: 10.1111/j.1755-0998.2009.02591.x. https://www.ncbi.nlm.nih.gov/pubmed/21564903.

41. Earl DA, Vonholdt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012;4 2:359-61; doi: 10.1007/s12686-011-9548-7. <Go to ISI>://WOS:000303536400036.

42. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14 8:2611-20; doi: 10.1111/j.1365-294X.2005.02553.x. <Go to ISI>://WOS:000229961500029.

43. 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 14:1801-6; doi: 10.1093/bioinformatics/btm233. <Go to ISI>://WOS:000249248300012.

44. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes. 2004;4 1:137-8; doi: 10.1046/j.1471-8286.2003.00566.x. <Go to ISI>://WOS:000189159500042.

45. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49 W1:W293-W6; doi: 10.1093/nar/gkab301. https://www.ncbi.nlm.nih.gov/pubmed/33885785.

46. Bandelt HJ, Forster P, Rohl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999;16 1:37-48. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=10331250&query_hl=1.

47. El-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. J Mol Recognit. 2008;21 4:243-55; doi: 10.1002/jmr.893. https://www.ncbi.nlm.nih.gov/pubmed/18496882.

48. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8:4; doi: 10.1186/1471-2105-8-4. https://www.ncbi.nlm.nih.gov/pubmed/17207271.

49. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10 6:845-58; doi: 10.1038/nprot.2015.053. https://www.ncbi.nlm.nih.gov/pubmed/25950237.

50. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A. 2001;98 18:10037-41; doi: 10.1073/pnas.181342398. https://www.ncbi.nlm.nih.gov/pubmed/11517324.

51. Bittencourt NC, Silva A, Virgili NS, Schappo AP, Gervasio J, Pimenta TS, et al. Plasmodium vivax AMA1: Implications of distinct haplotypes for immune response. PLoS Negl Trop Dis. 2020;14 7:e0008471; doi: 10.1371/journal.pntd.0008471. https://www.ncbi.nlm.nih.gov/pubmed/32639964.

52. Ruan W, Zhang LL, Feng Y, Zhang X, Chen HL, Lu QY, et al. Genetic diversity of Plasmodium Vivax revealed by the merozoite surface protein-1 icb5-6 fragment. Infect Dis Poverty. 2017;6 1:92; doi: 10.1186/s40249-017-0302-6. https://www.ncbi.nlm.nih.gov/pubmed/28578709.

53. Arnott A, Mueller I, Ramsland PA, Siba PM, Reeder JC, Barry AE. Global Population Structure of the Genes Encoding the Malaria Vaccine Candidate, Plasmodium vivax Apical Membrane Antigen 1 (PvAMA1). PLoS Negl Trop Dis. 2013;7 10:e2506; doi: 10.1371/journal.pntd.0002506. https://www.ncbi.nlm.nih.gov/pubmed/24205419.

54. Rougeron V, Elguero E, Arnathau C, Acuna Hidalgo B, Durand P, Houze S, et al. Human Plasmodium vivax diversity, population structure and evolutionary origin. PLoS Negl Trop Dis. 2020;14 3:e0008072; doi: 10.1371/journal.pntd.0008072. https://www.ncbi.nlm.nih.gov/pubmed/32150544.

55. Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGready R, Phyo AP, et al. Population genetic correlates of declining transmission in a human pathogen. Mol Ecol. 2013;22 2:273-85; doi: 10.1111/mec.12099. https://www.ncbi.nlm.nih.gov/pubmed/23121253.

56. Li Y, Hu Y, Zhao Y, Wang Q, Ngassa Mbenda HG, Kittichai V, et al. Dynamics of Plasmodium vivax populations in border areas of the Greater Mekong sub-region during malaria elimination. Malar J. 2020;19 1:145; doi: 10.1186/s12936-020-03221-9. http://www.ncbi.nlm.nih.gov/pubmed/32268906.

57. Liu Z, Soe TN, Zhao Y, Than A, Cho C, Aung PL, et al. Geographical heterogeneity in prevalence of subclinical malaria infections at sentinel endemic sites of Myanmar. Parasit Vectors. 2019;12 1:83; doi: 10.1186/s13071-019-3330-1. https://www.ncbi.nlm.nih.gov/pubmed/30777127.

58. Zhao Y, Zeng J, Zhao Y, Liu Q, He Y, Zhang J, et al. Risk factors for asymptomatic malaria infections from seasonal cross-sectional surveys along the China-Myanmar border. Malar J. 2018;17

1:247; doi: 10.1186/s12936-018-2398-y. https://www.ncbi.nlm.nih.gov/pubmed/29973194.

59. Weedall GD, Conway DJ. Detecting signatures of balancing selection to identify targets of anti-parasite immunity. Trends Parasitol. 2010;26 7:363-9; doi: 10.1016/j.pt.2010.04.002. https://www.ncbi.nlm.nih.gov/pubmed/20466591.

60. Demogines A, Abraham J, Choe H, Farzan M, Sawyer SL. Dual host-virus arms races shape an essential housekeeping protein. PLoS Biol. 2013;11 5:e1001571; doi: 10.1371/journal.pbio.1001571. https://www.ncbi.nlm.nih.gov/pubmed/23723737.

61. Abraham J, Corbett KD, Farzan M, Choe H, Harrison SC. Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. Nat Struct Mol Biol. 2010;17 4:438-44; doi: 10.1038/nsmb.1772. https://www.ncbi.nlm.nih.gov/pubmed/20208545.

62. Laurens MB. The Promise of a Malaria Vaccine-Are We Closer? Annu Rev Microbiol. 2018;72:273-92; doi: 10.1146/annurev-micro-090817-062427. https://www.ncbi.nlm.nih.gov/pubmed/30200856.

63. Ntege EH, Takashima E, Morita M, Nagaoka H, Ishino T, Tsuboi T. Blood-stage malaria vaccines: post-genome strategies for the identification of novel vaccine candidates. Expert Rev Vaccines. 2017;16 8:769-79; doi: 10.1080/14760584.2017.1341317. https://www.ncbi.nlm.nih.gov/pubmed/28604122.

64. Remarque EJ, Faber BW, Kocken CH, Thomas AW. A diversity-covering approach to immunization with Plasmodium falciparum apical membrane antigen 1 induces broader allelic recognition and growth inhibition responses in rabbits. Infect Immun. 2008;76 6:2660-70; doi: 10.1128/IAI.00170-08. https://www.ncbi.nlm.nih.gov/pubmed/18378635.

65. Osier FH, Weedall GD, Verra F, Murungi L, Tetteh KK, Bull P, et al. Allelic diversity and naturally acquired allele-specific antibody responses to Plasmodium falciparum apical membrane antigen 1 in Kenya. Infect Immun. 2010;78 11:4625-33; doi: 10.1128/IAI.00576-10. https://www.ncbi.nlm.nih.gov/pubmed/20732997.

66. Koepfli C, Rodrigues PT, Antao T, Orjuela-Sanchez P, Van den Eede P, Gamboa D, et al. Plasmodium vivax Diversity and Population Structure across Four Continents. PLoS Negl Trop Dis. 2015;9 6:e0003872; doi: 10.1371/journal.pntd.0003872. https://www.ncbi.nlm.nih.gov/pubmed/26125189.

67. Pearson RD, Amato R, Auburn S, Miotto O, Almagro-Garcia J, Amaratunga C, et al. Genomic analysis of local variation and recent evolution in Plasmodium vivax. Nat Genet. 2016;48 8:959-64; doi: 10.1038/ng.3599. https://www.ncbi.nlm.nih.gov/pubmed/27348299.

68. Rosanas-Urgell A, Lin E, Manning L, Rarau P, Laman M, Senn N, et al. Reduced risk of Plasmodium vivax malaria in Papua New Guinean children with Southeast Asian ovalocytosis in two

cohorts and a case-control study. PLoS Med. 2012;9 9:e1001305; doi: 10.1371/journal.pmed.1001305. https://www.ncbi.nlm.nih.gov/pubmed/22973182.

69. Muller I, Bockarie M, Alpers M, Smith T. The epidemiology of malaria in Papua New Guinea. Trends Parasitol. 2003;19 6:253-9; doi: 10.1016/s1471-4922(03)00091-6. https://www.ncbi.nlm.nih.gov/pubmed/12798082.

Figures

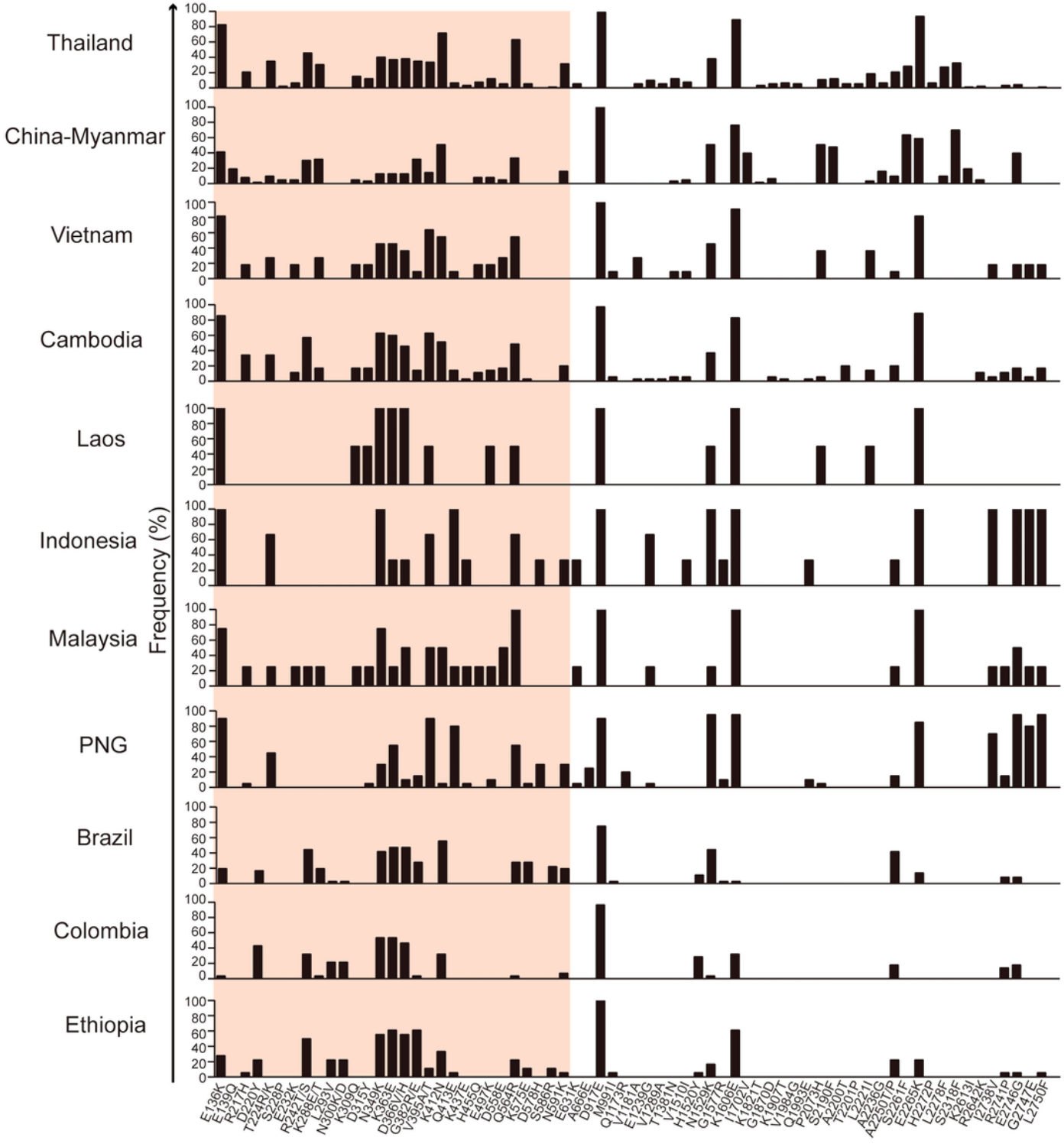

Figure 1

Prevalence of amino acid substitutions in PvRBP2b among worldwide P. vivax populations. Positions and frequencies of amino acid changes in PvRBP2b among different populations are shown. The PvRBP2b reticulocyte binding region (168–633 aa) is shadowed.

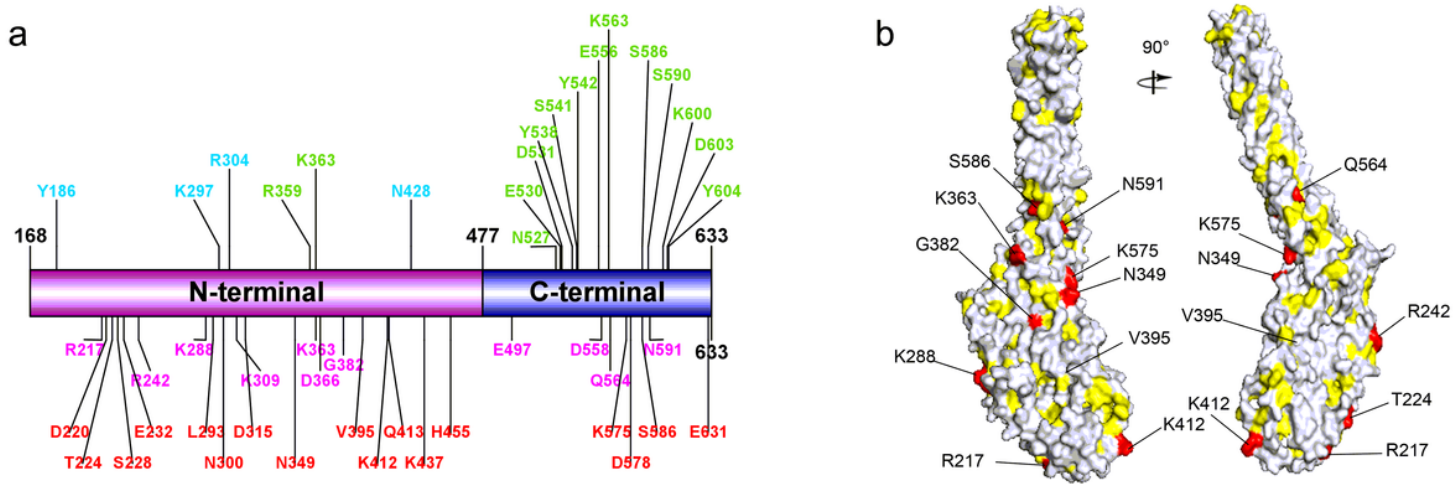

Figure 2

3D structure of PvRBP2b reticulocyte binding region. (a) A modified schematic diagram from reference [14] shows the important residues in the reticulocyte binding region composed of the N-terminal and C-terminal domains. The polymorphic residues previously reported were labelled in magenta, whereas the newly identified polymorphic residues in this study were labelled in red. Residues interacting with TfR1 and Tf are shown in green and cyan, respectively. (b) Two orthogonal views of the 3D model structure of PvRBP2b reticulocyte binding region shows the 12 residues positively selected by all three codon-based tests in the Datamonkey webserver and S586 positively selected by FUBAR. Hydrophobic binding region is shown in yellow.

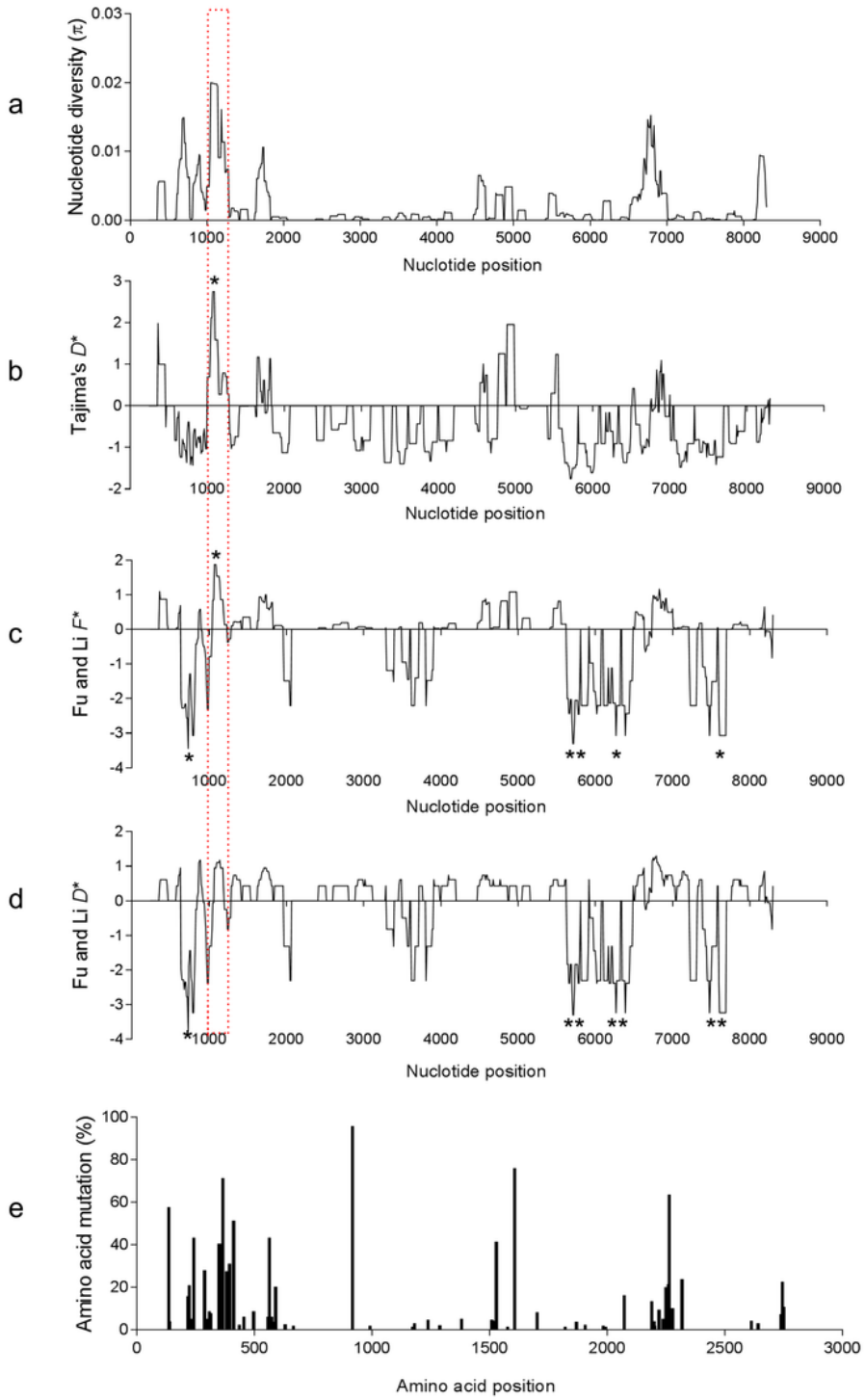

Figure 3

Nucleotide diversity, neutrality tests and amino acid polymorphisms. Sliding window plots of nucleotide diversity (a), Tajima's D* (b), Fu and Li's F* (c) and Fu and Li's D* (d) for PvRBP2b sequences are shown with a window size of 90 bp and a step size of 3 bp. Amino acid polymorphisms (e) are also visualized to the corresponding locations. The asterisk (*) depicts sites of statistical significance under balancing or directional selections ($p<0.05$).

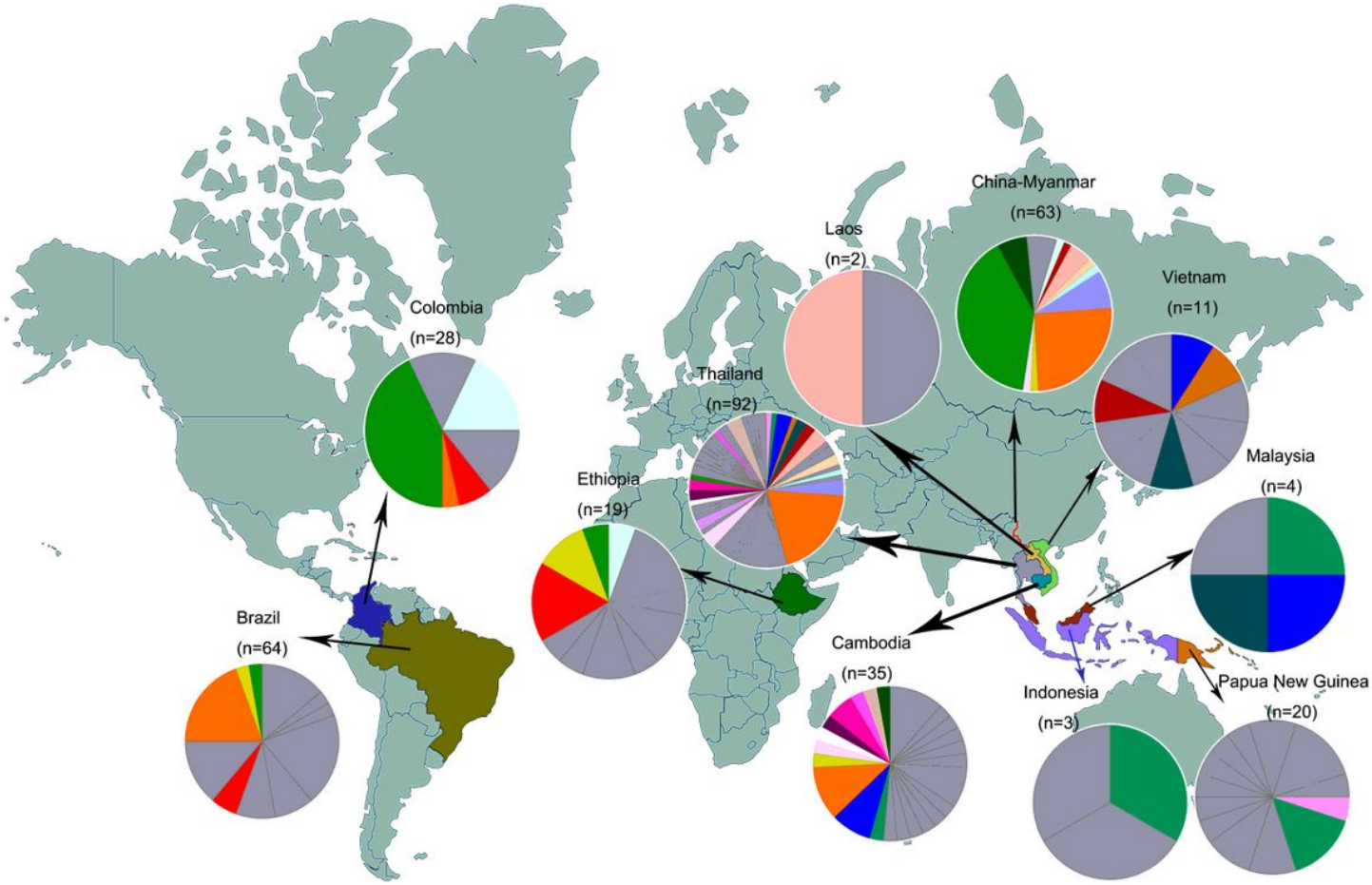

Figure 4

Map showing the distribution of PvRBP2b haplotypes. The frequencies of 114 haplotypes based on the 17 nonsynonymous amino acid mutations (>5%) in the PvRBP2b reticulocyte binding region are depicted as pie charts and mapped to their geographic origins. The red line represents the region along the Myanmar-China border. Shared haplotypes are shown in color and unique haplotypes are shown in grey.

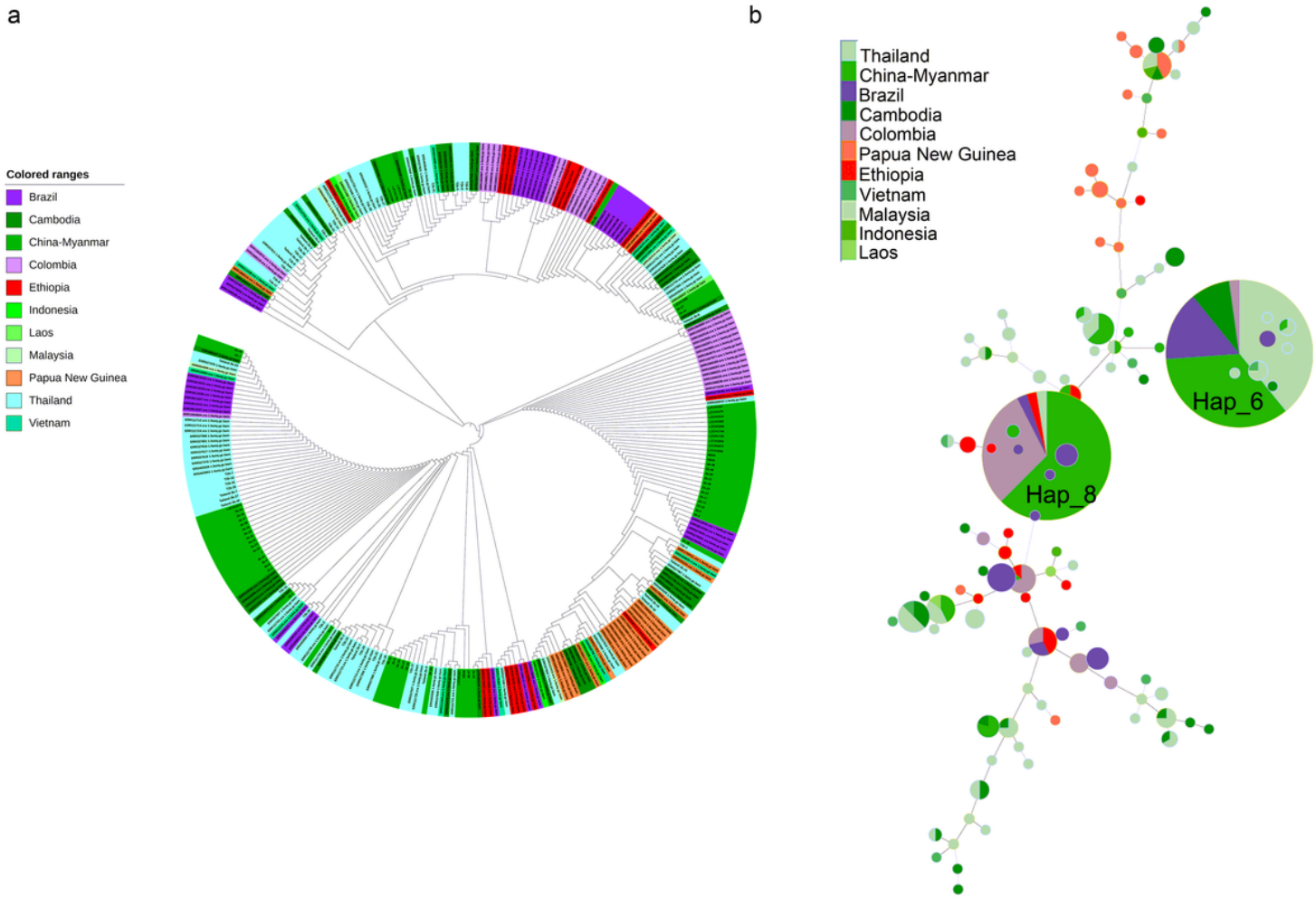

Figure 5

Phylogenetic relationship of the parasite isolates based on the PvRBP2b reticulocyte binding region. (a) An unrooted phylogenetic tree. The tree was constructed using the Neighbor-Joining method with bootstrap supports from 1000 replicates. The origins of the parasite isolates are represented by different colors. (b) A haplotype network. The size of the pies reflects the frequency of the specific haplotype. The lengths of the lines connecting the pies measured from their centers are in proportion to the number of base pair substitutions separating the haplotypes. Endemic regions are represented by same colors used in the phylogenetic tree.

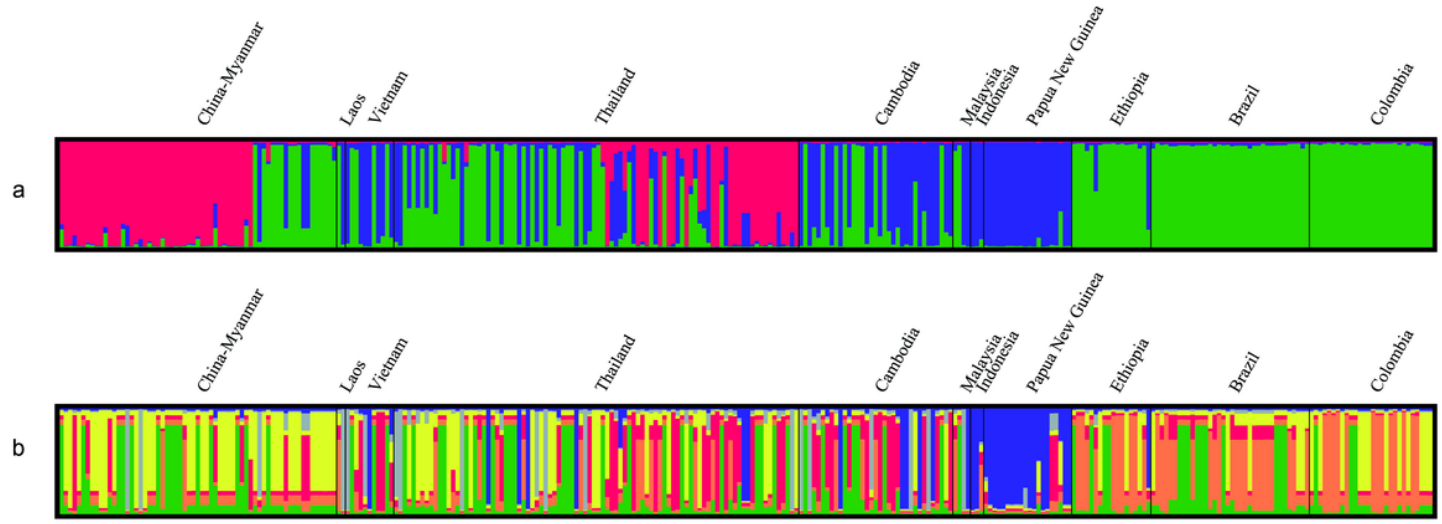

Figure 6

Structure analysis of PvRBP2b sequences from the global P. vivax populations. Plots represent the genetic structures of the near full-length sequences of PvRBP2b at K = 3 (a) and of the reticulocyte binding region at K = 6 (b) in parasite populations from different endemic regions.

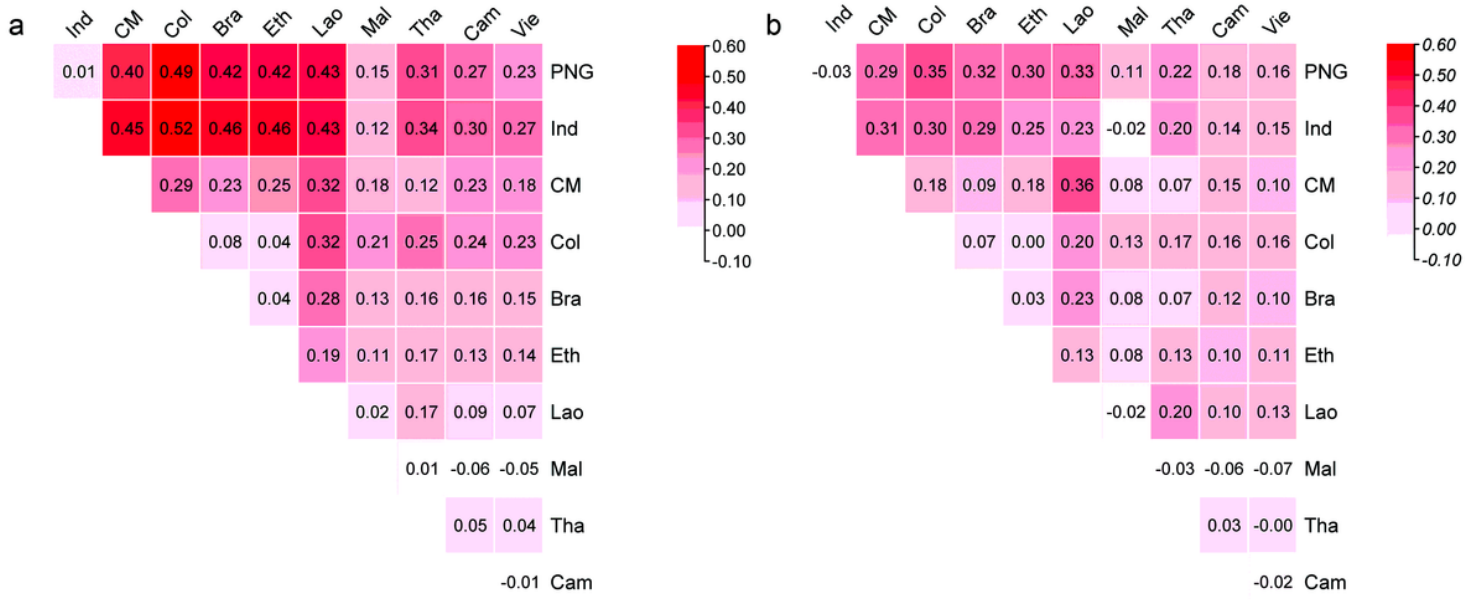

Figure 7

FST analysis of global P. vivax populations. Heatmaps show pairwise comparison for the near full-length sequence (a) and the reticulocyte binding region (b) in PvRBP2b among worldwide P. vivax populations. The numbers in the cells are the FST values between two endemic areas.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GraphicAbstractree.tif
- Supplementaltables1420211020.docx