Infection, Genetics and Evolution 86 (2020) 104592

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research Paper

Exploration of *Plasmodium vivax* merozoite surface proteins 1 and 7 genetic diversity in Brazilian Amazon and Rio de Janeiro Atlantic Forest

Natália Ketrin Almeida-de-Oliveira[a,b], Rebecca Abreu-Fernandes[a,b], Aline Rosa Lavigne[a,b], Anielle Pina-Costa[b,c,d], Daiana de Souza Perce-da-Silva[a,b,e], Marcos Catanho[f], Átila Duque Rossi[g], Patrícia Brasil[b,c], Cláudio Tadeu Daniel-Ribeiro[a,b], Maria de Fátima Ferreira-da-Cruz[a,b,1,*]

[a] Laboratório de Pesquisa em Malária, Instituto Oswaldo Cruz (IOC), Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro, RJ, Brazil
[b] Centro de Pesquisa, Diagnóstico e Treinamento em Malária (CPD-Mal), Reference Laboratory for Malaria in the Extra-Amazonian Region for the Brazilian Ministry of Health, SVS & Fiocruz, Rio de Janeiro, RJ, Brazil
[c] Laboratório de Pesquisa Clínica em Doenças Febris Agudas, Instituto Nacional de Infectologia Evandro Chagas, Fiocruz, Rio de Janeiro, RJ, Brazil
[d] Centro Universitário Serra dos Órgãos (UNIFESO), Teresópolis, RJ, Brazil
[e] Laboratório de Imunologia Básica e Aplicada, Faculdade de Medicina de Petrópolis - FMP/FASE, Petrópolis, RJ, Brazil
[f] Laboratório de Genética Molecular de Microrganismos, IOC, Fiocruz, Rio de Janeiro, RJ, Brazil
[g] Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, RJ, Brazil

ARTICLE INFO

Keywords:
P. vivax
Pvmsp-1
Pvmsp-7E
Pvmsp-7F
Genetic diversity

ABSTRACT

Plasmodium vivax merozoite surface proteins (*Pv*MSP) 1 and 7 are considered vaccine targets. Genetic diversity knowledge is crucial to assess their potential as immunogens and to provide insights about population structure in different epidemiological contexts. Here, we investigate the variability of *pvmsp-1*$_{42}$, *pvmsp-7E*, and *pvmsp-7F* genes in 227 samples from the Brazilian Amazon (BA) and Rio de Janeiro Atlantic Forest (AF). *pvmsp-1*$_{42}$ has 63 polymorphisms - 57 nonsynonymous - generating a nucleotide diversity of $\pi = 0.009$ in AF, and $\pi = 0.018$ in BA. In *pvmsp-7E,* 134 polymorphisms - 103 nonsynonymous - generate the nucleotide diversity of $\pi = 0.027$ in AF, and $\pi = 0.042$ in BA. The *pvmsp-7F* has only two SNPs - A610G and A1054T –, with nucleotide diversity of $\pi = 0.0004$ in AF, and $\pi = 0.0007$ in BA. The haplotype diversity of *pvmsp-1*$_{42}$, *pvmsp-7E*, and *pvmsp-7F* genes is 0.997, 1.00, and 0.649, respectively. None of the *pvmsp-1*$_{42}$ or *pvmsp*-7E sequences are identical to the Salvador 1 strain's sequence. Conversely, most of *pvmsp-7F* sequences (94/48%) are identical to Sal-1. We evaluated eight B-cell epitopes in *pvmsp-7E*, four of them showed higher nucleotide diversity compared to *pvmsp-7E's* epitopes. Positive selection was detected in *pvmsp-1*$_{42}$, *pvmsp-7E* central region, and *pvmsp-7F* with Tajima's D. In *pvmsp-7E*, the significant nucleotide and haplotype diversities with low genetic differentiation, could be indicative of balancing selection. The genetic differentiation of *pvmsp-1*$_{42}$ (0.315) and *pvmsp-7F* (0.354) genes between AF and BA regions is significant, which is not the case for *pvmsp-7E* (0.193). We conclude that *pvmsp-1*$_{42}$ and *pvmsp-7E* have great genetic diversity even in AF region, an enclosure area with deficient transmission levels of *P. vivax* zoonotic malaria. In both Brazilian regions, *pvmsp-1*$_{19}$, *pvmsp-7E*, and *pvmsp-7F* are conserved, most likely due to their roles in parasite survival, and could be considered potential targets for a “blood-stage vaccine”.

1. Introduction

Plasmodium vivax is the second most prevalent species causing malaria in humans and the most widespread *Plasmodium* parasite globally. In the Americas, the parasite is predominant, representing 80% of malaria cases (WHO, 2019). In Brazil, 90% of malaria cases registered in 2019 were caused by *P. vivax* (Ministério da Saúde, 2019).

Over the past 15 years, it seems that the prevalence of vivax malaria causing severe disease has been increasing (Lacerda et al., 2012; Price et al., 2014; Rahimi et al., 2014). This parasite displays singular

Abbreviations: WHO, World Health Organization; *P. vivax*, *Plasmodium vivax*; *pvmsp*, gene of *Plasmodium vivax* merozoite surface protein; AF, Atlantic Forest; BA, Brazilian Amazon; II, Nucleotide diversity; S, polymorphic sites; Hd, Haplotype diversity; Fst, Wright's fixation statistics; SNP, Single Nucleotide Polymorphism; An, Anopheles

* Corresponding author at: Laboratório de Pesquisa em Malária, Instituto Oswaldo Cruz (IOC), Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro, RJ, Brazil.
E-mail addresses: nataliak@ioc.fiocruz.br (N.K. Almeida-de-Oliveira), patricia.brasil@ini.fiocruz.br (P. Brasil), mffcruz@ioc.fiocruz.br (M.d.F. Ferreira-da-Cruz).
[1] Postal address: Laboratório de Pesquisa em Malária, IOC, Fiocruz. Avenida Brasil, 4365 - Manguinhos, Rio de Janeiro, RJ. CEP: 21040–360

https://doi.org/10.1016/j.meegid.2020.104592
Received 7 April 2020; Received in revised form 22 September 2020; Accepted 7 October 2020
Available online 13 October 2020

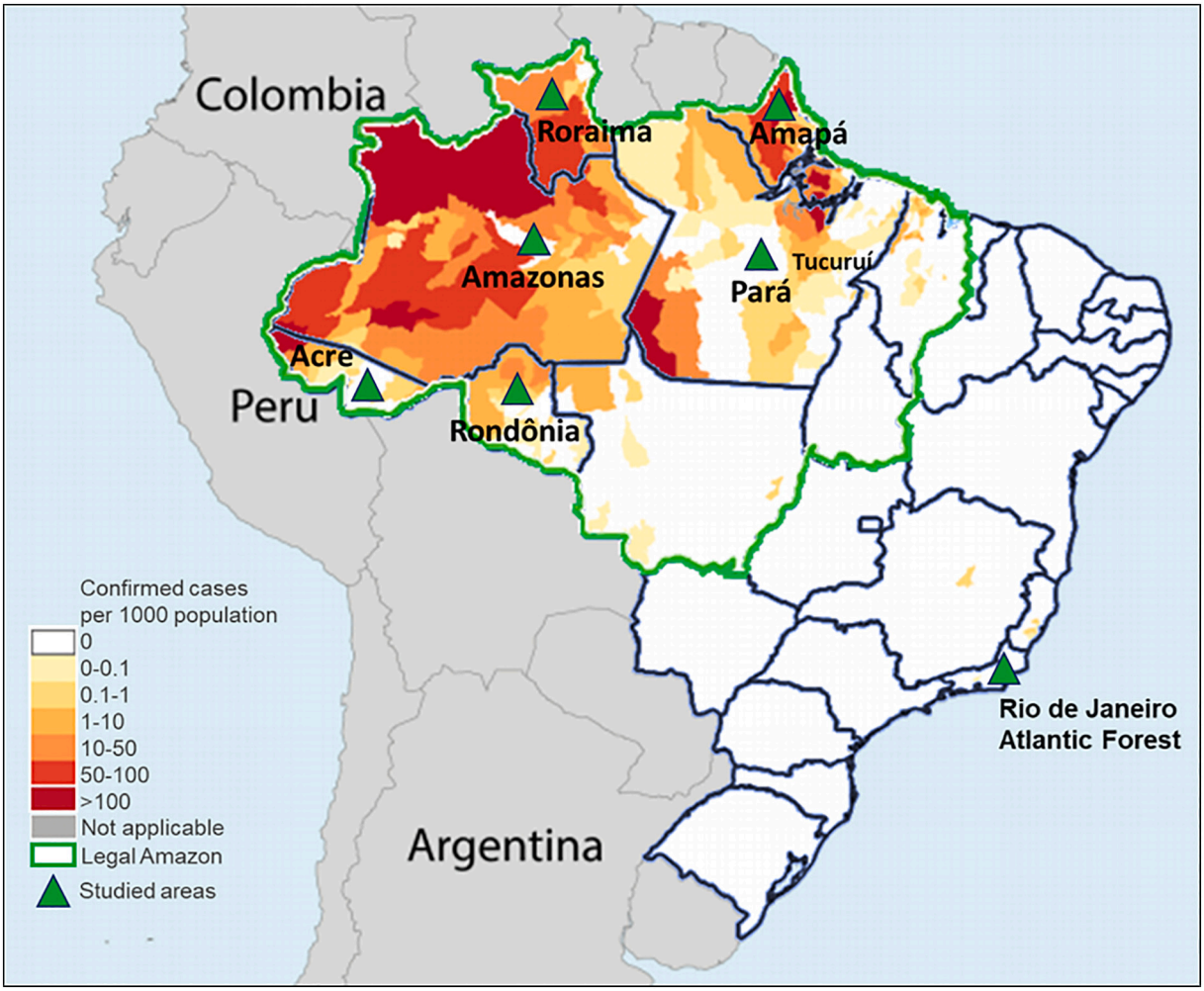

Fig. 1. Brazilian isolates including seven states represented by triangles in two distinct regions in Brazil: Brazilian Amazon and Atlantic Forest of Rio de Janeiro. Adapted from malaria report (WHO, 2018).
Map showing two regions of malaria sampling area: the Atlantic Forest of Rio de Janeiro and the Brazilian Amazon represented by seven states marked with triangles. Adapted from malaria report (WHO, 2018).

biological characteristics such as a hypnozoite stage - responsible for the relapsing illness - and rapid gametocyte formation. Together with the evidence of drug-resistant parasites, these facts have been hindering the control strategies (Kevin Baird, 2013a; Kevin Baird, 2013b).

In this scenario, efforts to develop an anti-*vivax* malaria vaccine have been made, and the studies are mainly orientated to antigens that correspond to *P. falciparum* vaccine targets, as the merozoite surface proteins (MSP) -1 and − 7. These antigens form a protein complex that participates in the parasite adhesion to the reticulocyte (Holder et al., 1992; Mello et al., 2002; Mongui et al., 2006).

The *pvmsp-1* gene is expressed as a 200 kDa protein that undergoes two proteolytic cleavage steps during the merozoite maturation. Before erythrocyte invasion, the 42 kDa fragment undergoes the second cleavage generating 33 and 19 kDa fragments (MSP-1_{33} and MSP-1_{19}); the latter remains on merozoite's surface during the invasion, and it is the critical portion in parasite red blood cell binding (Han et al., 2004). Both fragments were immunogenic in individuals naturally exposed to malaria infections (Park et al., 2001), including Brazilian areas (Riccio et al., 2013; Cassiano et al., 2016; Fonseca et al., 2016), and immunized MSP-1_{19} monkeys were protected from challenging doses (Collins et al., 1999).

MSP-7 is a multigene family comprising 12 members (Castillo et al., 2017). This gene plays diverse and critical roles in parasite-host interactions, including merozoite invasion (Kadekoppala et al., 2008). Since paralogous genes may diverge in their functions, not all MSP-7 family-members have a significant role in erythrocytes invasion (Castillo et al., 2017).

Like PfMSP-7, the PvMSP-7E has predicted binding domains to interact with MSP-1 (Weng Cheng et al., 2018). This gene displays much higher genetic diversity than *pvmsp-7F* (Garzón-ospina et al., 2014; Weng Cheng et al., 2018), suggesting that these *msp* paralogous genes may have evolved under different selective pressure.

Genetic diversity studies within and among *P. vivax* populations are crucial to assess the potential vaccine targets and provide insights into parasite population dynamics and transmission. Such studies contribute to *P. vivax* parasites' characterization and enable the comparison with global endemic regions (Neafsey et al., 2012). In Brazil, in addition to *P. vivax lato sensu* (l.s.) endemic parasites in the Brazilian Amazon, the so-called *P. vivax/P. simium* parasites occur exclusively in areas of the Atlantic Forest and involve a human (*P. vivax*), howler monkey (*P. simium*), and the mosquito vector (*Kerteszia* anophelines), which breeds in water stored in bromeliads.

Currently, bromeliad malaria appears in the Brazilian Atlantic forest biome almost exclusively as a zoonosis. Monkeys are the parasite's reservoir, and *Anopheles (Kerteszia) cruzii* is the mosquito vector biting monkeys in the forest canopy, as well as people at ground level (de Abreu et al., 2019; Brasil et al., 2017; Deane, 1992). *P. simium* is morphologically indistinguishable from *P. vivax* l.s. The near genetic

identity between human *P. vivax* l.s and non-human primate *P. simium* suggests that recent host transfers occurred (Tazi and Ayala, 2011; Rougeron et al., 2020).

Here, we analyze the haplotype networks, positive selection, and genetic diversity among *P. vivax l.s* and *P. vivax/P. simium* parasites based on three genes, *pvmsp-1*$_{42}$, *pvmsp-7E*, and *pvmsp-7F,* to investigate the possible role of different hosts and anopheles mosquito's vectors in the genetic diversity of these genes, in areas of Atlantic Forest (*P. vivax/ P. simium*) and Brazilian Amazon (*P. vivax l.s.*) regions.

2. Material and methods

2.1. Sample collection

The descriptive study of polymorphisms in *pvmsp-1*$_{42}$, *pvmsp-7E*, and *pvmsp-7F* genes were conducted with 227 Brazilian samples from Amazonian Basin (BA) and Atlantic Forest (AF) regions. 145 BA samples came from six states located in the North region of Brazil (Acre, Amapá, Amazonas, Pará, Rondônia, and Roraima), and 82 AF samples from Rio de Janeiro, in the Southeast region of Brazil (Fig. 1).

All blood samples were collected from *P. vivax* patients diagnosed with thick and thin blood smears and PCR (Torres et al., 2006). These samples were collected in two localities: 177 (95 from BA and 82 from AF) in the malaria reference center for diagnosis, teaching, and research (CPD-Mal) based at the Oswaldo Cruz Institution/ Rio de Janeiro (S 22° 54′ W 43° 12′) from 2011 to 2018 and; 50 samples in Tucuruí (S 3° 46 ‘W 49° 40’), Pará state, BA (Table 1). The study was approved by the Ethics Research Committee of Instituto Oswaldo Cruz, Fiocruz, Brazil (69,256,317.3.0000.5248). All volunteers' patients signed a written informed consent before collection of 4 mL of venous blood.

2.2. pvmsp-142, pvmsp-7E, and pvmsp-7F genes DNA genotyping

DNA was extracted from 1 mL blood sample using QIAamp™ DNA Blood Midi Kit (QIAGEN, Hilden, Germany), according to the manufacturer's instructions. The DNA was stored at −20 °C until use.

pvmsp-1$_{42}$-gene amplification was based on Parobek and colleagues' methodology (Parobek et al., 2014), while *pvmsp-7E* and *pvmsp-7F* genes amplification relied on Garzón-Ospina and colleagues' (Garzón-ospina et al., 2014).

All amplification reactions were carried out in a total volume of 20 μL, containing 2 μL (100–200 ng) of DNA, 1 μM primers (forward and reverse), 5× HOT FIREPol™ Blend Master Mix (Solis Biodyne, Hannover, Germany). Thermal conditions for all genes were set as follows: one hold of 95 °C/12 min; 30 cycles of 95 °C/20 s, 57 °C/45 s for *pvmsp-1*$_{42}$ and 60 °C/45 s for *pvmsp-7E/F*, 72 °C/1 min, followed by a final extension of 72 °C/7 min.

PCR products were analyzed in a 2% agarose gel containing 0.25 mg/mL of ethidium bromide, and the results were visualized in a UV transilluminator (DigiDoc-It, UVP, California, USA).

Amplicons were purified using Wizard™ SV Gel and PCR Clean-Up System (Promega, USA), following the manufacturer's instruction. Then, 1–2 μL of purified PCR products were submitted to thermal sequencing reaction using Big Dye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, California, USA), and 3.2 μM of forward and reverse PCR primers. Products were precipitated using isopropanol and resuspended in HI-DI formamide (Applied Biosystems California, USA); DNA samples were sequenced by capillary electrophoresis on the Genomic Platform PDTIS/RPT01A of Fiocruz using an ABI Prism DNA Analyzer™ 3730 (Applied Biosystems, California, USA).

Sequences were globally aligned using ClustalW after analyzing the electropherogram with NovoSNP software for quality control. The signal threshold value was set to QV = 10 for accurate base calling. Sequences with weak signal-to-noise ratio, as well as with overlapping peaks and base substitution present in only one strain (singleton), were discarded and resequenced. The corresponding genes in Salvador I strain were considered as references (accession numbers: *pvmsp-1*, PVX_099980; *pvmsp-7E,* PVX_082665; and *pvmsp-7F*, PVX_082670; available in plasmoDB.org). All sequences were deposited in GenBank with accession nnumbers. MT074744 - MT075302.

2.3. Genetic diversity, natural selection, and phylogenetic analyses of pvmsp-1$_{42}$, pvmsp-7E, and pvmsp-7F genes

The number of segregation sites (s), polymorphisms, haplotype diversity (*Hd*), and nucleotide diversity (π) were computed using the DnaSP 6.11 software (Rozas et al., 2017). Nucleotide diversity was assessed using the sliding window method with a window length of 50 bases and a step size of 25 sites.

Natural selection in *pvmsp-1* and *pvmsp*-7E genes were assessed by the *Z*-test and the Tajima's D test while in *pvmsp*-7F gene only by Tajima's D test. The Z-test method calculating the differences of the rates of nonsynonymous (dN) to synonymous (dS) substitutions (dN/dS) (1000 bootstrapping replicates) between isolates using the Nei and Gojobori's distance (Nei and Gojobori, 1986) and the Jukes and Cantor correction was performed in MEGA7 v7.0 (Nei and Kumar, 2000). Tajima's D neutrality test (Tajima, 1989) was calculated with the total number of segregating sites and the sliding window method in DnaSP 6.11 software, in which $p < 0.05$ was considered significant.

Phylogenetic analyses were performed in MEGA7 v7.0 applying the p-distance method, and tree reconstruction by the Neighbor-joining method (Saitou and Nei, 1987), with 1000 bootstrapping replicates. The branch length is proportional to time. The haplotype network was constructed on R with the “haploNet” function of the package “pegas” under an infinite site model (Paradis, 2010). Each haplotype is represented by a circle with size proportional to the frequency; the color provides information about the geographic distribution. The number of dots on links represents each haplotype's genetic distance, and pairwise differences evaluated it.

The degree of genetic differentiation among Brazilian regions was

Table 1
P. vivax sampling, according to localities and year of collection.

Region	State	Year of collection								
		2011	2012	2013	2014	2015	2016	2017	2018	Total
BA	Acre	2	3	1	2	1	0	2	1	12
n = 145	Amapá	3	1	1	1	2	0	1	0	9
	Amazonas	2	5	6	8	4	11	10	2	48
	Para	51	1	1	1	0	0	1	0	55
	Rondônia	9	2	1	0	0	2	3	0	17
	Roraima	1	2	0	1	0	0	0	0	4
AF	Rio de Janeiro	3	2	5	0	29	14	22	7	82
n = 82										
Total		71	16	15	13	36	27	39	10	227

BA: Brazilian Amazon; AF: Atlantic Forest.

estimated by the Wright's fixating index (Fst) (Hudson et al., 1992), using DnaSP 6.11. This value ranges from 0 to 1 and measures the amount of diversity between geographically distinct populations.

3. Results

3.1. *pvmsp-1*$_{42}$

pvmsp-1$_{42}$ sequences were obtained from 185 (81%) of the collected isolates: 63 from AF and 122 from BA. The BA samples cover six states: Acre (10), Amapá (6), Amazonas (43), Pará (48), Rondônia (12), and Roraima (3).

The 1267 bp *pvmsp-1*$_{42}$-gene fragment (nucleotides 3973–5239; codons 1325-1747) had 57 segregation sites. All nucleotide changes were detected in *pvmsp-1*$_{33}$ fragment (nucleotides 3973-4917), while the fragment *pvmsp-1*$_{19}$ (nucleotides 4918-5239) remained conserved. Most of them (53/91%) were biallelic, while four sites (7%) were triallelic and one (2%) quatre-allelic, comprising a total of 63 polymorphisms (Supplementary Material 1). Of these 63 polymorphisms (SNPs), 57 were nonsynonymous, and six synonymous, spanning 43 codons. The AF samples have 46 SNPs, and 42 are nonsynonymous, whereas BA has 61 SNPs, 55 nonsynonymous (Supplementary Material 2).

Among AF and BA samples, 44 polymorphisms were detected, and 40 (91%) nonsynonymous. Eleven (25%) are present in more than 50% of the samples: M1466L (96%), K1522E (82%), I1454L (75%), L1513I (66%), D1490T (62%), N1515S (61%), K1493N (61%), G1491A (60%), D1520T (59%), E1525Q (55%), and E1509K (54%) (Supplementary Material 2). Remarkably, these 11 polymorphisms are more frequent in AF than in BA (Supplementary Fig. 3). Additionally, two polymorphisms were detected only in AF samples (23; 36%): D1520A in seven (11%) samples, and K1523Q in 18 (29%) samples.

In BA, 111 samples (91%) have 17 polymorphisms not detected in AF, none of them displaying all these 17 polymorphisms. The number of samples containing such polymorphisms varied from one (with 10 SNPs) to 28 (with 1 SNP); Interestingly, the V1516G was detected only in Tucuruí, a locality of Pará state.

The nucleotide diversity in AF and BA regions is $\pi = 0.009$ and $\pi = 0.018$, respectively (Table 2). The highest peak of polymorphisms in the Brazilian isolates was detected in the segment positions 4524–4573 (midpoint 4548): AF ($\pi = 0.097$) and BA ($\pi = 0.199$) (Fig. 2); this segment accumulated 48% of the polymorphic amino acids.

The wide range of polymorphic sites in *pvmsp-1*$_{42}$ sequences are arranged in 148 haplotypes with a high haplotype diversity of 0.997: 48 haplotypes from AF (*Hd* = 0.989) and 100 from BA (*Hd* = 0.996) (Table 2).

We constructed a haplotype network tree with 148 haplotypes. Singleton haplotypes were detected in 87% of the isolates. Only 19 haplotypes (12%; H1-H19) are restricted to a particular geographic region and present in more than one sample: 6 in AF (21 samples) and 13 in BA (35 samples). Most AF haplotypes grouped together (Fig. 3).

The phylogenetic tree showed BA isolates spread over several clades, possibly correlating with the larger geographic area and the broad range of polymorphic haplotypes, varying from one to 29 SNPs per sequence, when compared to AF (Supplementary Fig. 4). Considering 95% confidence intervals (95% CI), the number of SNPs is 15–18 (mean 16) for BA and 22–25 (mean 23) for AF, though no statistically significant difference was verified ($p = 0.06$) by the Mann-Whitney test (Fig. 4). As expected, we observed extensive BA and AF interpopulation differentiation (Fst = 0.315).

We performed the Tajima's D and *Z*-test to determine whether natural selection was affecting the *pvmsp-1*$_{42}$-gene fragment. Both tests significantly rejected the null hypothesis showing positive selection in all AF and BA samples with 2.183 ($p < 0.01$) and 3.838 ($p < 0.001$) values for Z-test and Tajima's D, respectively. Segregated by region, the Z-test index for AF samples was higher (3.302; $p < 0.001$) than those for BA samples (1.621; $p < 0.05$) (Table 2) (Supplementary Material 5).

3.2. *pvmsp-7E*

pvmsp-7E-gene was successfully sequenced in 178 samples (79%): 54 from AF and 124 from BA. Analyzing the 1114 bp *pvmsp-7E*-gene product of each sample (nucleotides 18–1131; codons 6–372), 119 segregation sites, resulting in 134 polymorphisms (103 nonsynonymous and 33 synonymous), spanning 103 codons were identified (Supplementary Material 6). We observed 112 SNPs (82 nonsynonymous) in AF parasites, while in BA we found 133 SNPs (100 nonsynonymous). Five out of 119 segregation sites are triallelic (A415C/G, C432G/A, C478A/T, T514C/G, and T705A/C), the remaining 114 are biallelic.

Table 2
Genetic diversity and natural selection of *pvmsp-1*$_{42}$, *pvmsp-7E* and *pvmsp-7F* genes in Brazilian isolates.

Gene	Genetic diversity							Natural Selection	
	N of Isolates		S	P	π (SD)	Nh	*Hd* (SD)	*Z*-test	Tajima's D
pvmsp-1$_{42}$ (1266 bp)	BA	122	56	61	0.018 (± 0.0002)	100	0.996 (± 0.001)	1.621*	3.629***
	AF	63	48	46	0.009 (± 0.0010)	48	0.989 (± 0.006)	3.302***	0.436
	Total	185	57	63	0.018 (± 0.0002)	148	0.997 (± 0.001)	2.183**	3.838***
pvmsp-7E (1114 bp)	BA	124	118	133	0.042 (± 0.0010)	119	0.999 (± 0.001)	1.169	4.279 ***
	AF	54	108	112	0.027 (± 0.0020)	54	1.00 (± 0.004)	1.587	0.786
	Total	178	119	134	0.041 (± 0.0010)	173	1.00 (± 0.000)	1.118	4.326***
pvmsp-7F (1211 bp)	BA	126	2	2	0.0007 (± 0.0002)	3	0.675 (± 0.000)	NA	2.166*
	AF	70	2	2	0.0004 (± 0.0009)	3	0.004 (± 0.070)	NA	0.217
	Total	196	2	2	0.0007 (± 0.0002)	3	0.649 (± 0.016)	NA	2.357*

BA: Brazilian Amazon; AF: Atlantic Forest; S: Segregation sites; P: number of polymorphisms; π: Nucleotide diversity; SD: standard deviation; Nh: number of haplotypes; *Hd*: haplotypes diversity; ***($p < 0.001$); **($p < 0.01$); *($p < 0.05$); Natural selection in *pvmsp-1*$_{42}$ consider the fragment of 33 KDa and *pvmsp-7E* the central region; NA: not applicable.

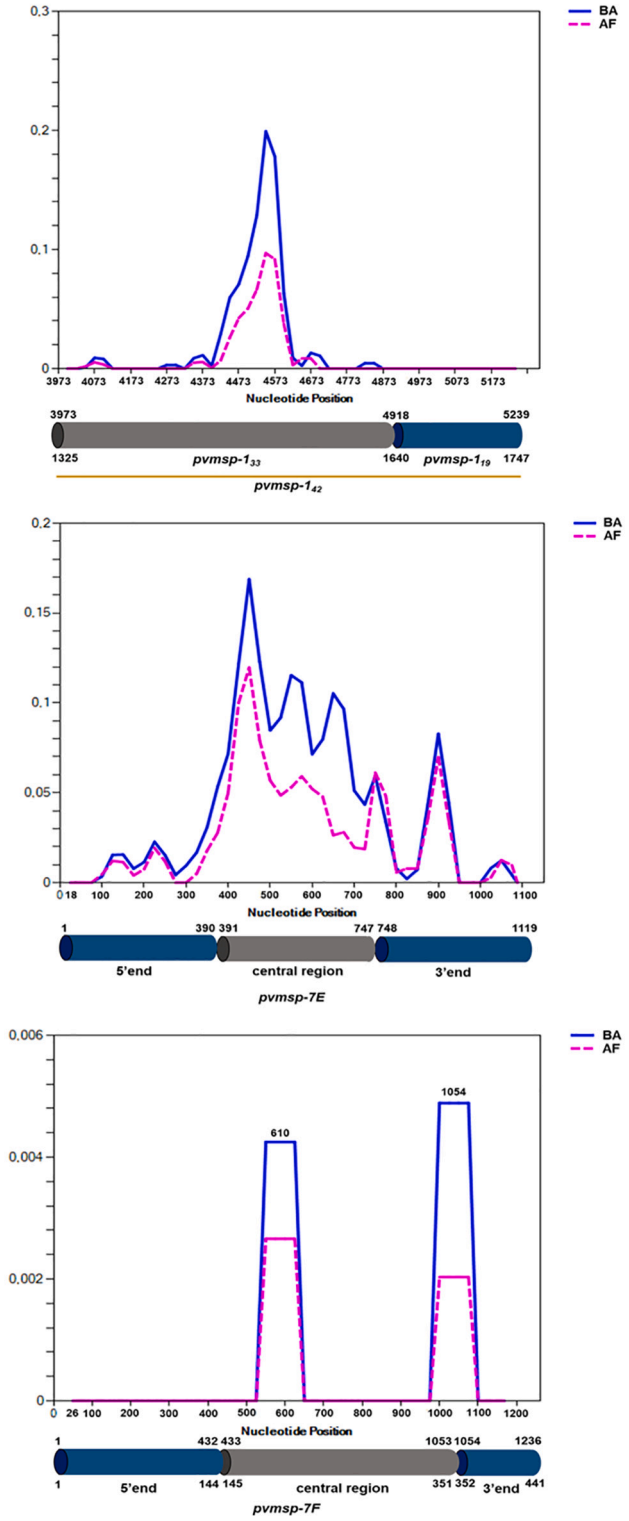

Fig. 2. Nucleotide diversity across the *pvmsp-1*$_{42}$, *pvmsp-7E*, and *pvmsp-7F* genes in *P. vivax* isolates from Brazilian Amazon (BA) and Rio de Janeiro Atlantic Forest (AF).
Sliding window of isolates from BA and AF was calculated with a length of 50 bp and step size of 25 bp; all coordinates are based on Sal-1 reference gene; *pvmsp-7E* and *pvmsp-7F* gene analyses were divided into three fragments, according to Garzón-Ospina et al. (2014).

Samples from AF and BA regions shared 111 polymorphisms, 32 (29%) were found in at least 50% of the samples. Similarly, we observed a higher frequency of the most common polymorphisms in *pvmsp-1*$_{42}$-gene of AF (22 SNPs) than in BA (10 SNPs) samples (Supplementary Material 2).

Conversely, 23 polymorphisms were detected only in one region: one SNP in 41% of AF samples, and 22 SNPs in 77% of BA samples. No exclusive SNP was found in BA samples (Supplementary Material 2).

The nucleotide diversity of *pvmsp-7E*-gene for all isolates was $\pi = 0.049$. However, when considered this diversity by region specific numbers for AF ($\pi = 0.027$) and BA ($\pi = 0.042$) samples were recorded (Table 2). The sliding window method recorded the highest peak in the gene's central region between residues 408 and 457 (midpoint 432) in both Brazilian regions, reaching values of $\pi = 0.120$ and $\pi = 0.169$, in AF and BA, respectively (Fig. 2). Most polymorphisms (96/72%) occur in the gene's central region (nucleotide 391 to 747). Multiple sequence alignment revealed at least four segregation sites in all isolates compared to the Sal-1 reference sequence. The haplotype diversity reached the maximum value in both Brazilian regions (H*d* = 1.00), indicating a substantial polymorphism arrangement in all the isolates, independently of the region (Table 2). Overall, the number of haplotypes is 173: 54 in AF and 119 in BA. Four haplotypes were detected only in BA samples: H7*E*-1 in three samples, and H7E-2, H7E-3, and H7E-4 in six samples. Therefore, the *pvmsp-7E*-gene network was more complicated than *pvmsp-1*$_{42}$, with a higher rate of substitution between haplotypes (Fig. 5).

Based on nucleotide polymorphisms, the neighbor-joining phylogenetic tree clustered the samples by region (Supplementary Material 4). BA isolates exhibited a broad range of polymorphic haplotypes, varying from two to 83 SNPs per sequence when compared to AF ($p = 0.002$) (Fig. 4). Considering 95% confidence intervals (95% CI), the number of SNPs was similar between BA (41 to 47; mean 44) and AF (38 to 42; mean 40).

A considerably low genetic differentiation was observed in Brazilian populations from AF and BA (Fst = 0.193).

Natural selection analyses comprised three fragments of the *pvmsp-7E*-gene: 5′-end (nucleotide 1 to 390), central region (nucleotide 391 to 747), and 3′-end (nucleotide 748 to 1119). 5′ and 3′ ends tend to be conserved in all studied areas. The central region of the gene had a positive value in *Z*-test (1.118; $p < 0.30$), as well as in Tajima's D test (4.326; $p < 0.001$). Tajima's D's sliding window plot showed diverse peaks across the *pvmsp7-E*-gene sequence, pehaps reflecting the maintenance of polymorphism by a balancing selection in AF and BA regions (Supplementary material 5).

Subsequently, we investigated the presence of polymorphisms in eight predicted B-cell epitopes in Sal-1 strain (Weng Cheng et al., 2018). Seven epitopes were polymorphic; six of these epitopes were found in more than 50% of samples, with the number of SNPs ranging from 8 to 20. The nucleotide diversity in epitopes B-3, B-4, B-5, and B-8 was higher than those observed in the entire *pvmsp-7E*-gene's amplified fragment (Table 3).

3.3. *pvmsp-7F*

196 out of 227 (87%) *P. vivax* isolates had the *pvmsp-7F*-gene fragment (1211 bp; nucleotides 26–1243; codon 9–411) successfully amplified. Among these isolates, 70 were from AF, and 126 from six BA states: Acre (10), Amapá (9), Amazonas (42), Pará (49), Rondônia (13), and Roraima (3).

Sequence analysis revealed two nonsynonymous SNPs occurring in all Brazilian regions, A610G and A1054T, in codons T204A and I352F, respectively (Supplementary Material 2). Multiple sequence alignment showed a high similarity between Brazilian *pvmsp-7F*-gene and the Sal-1

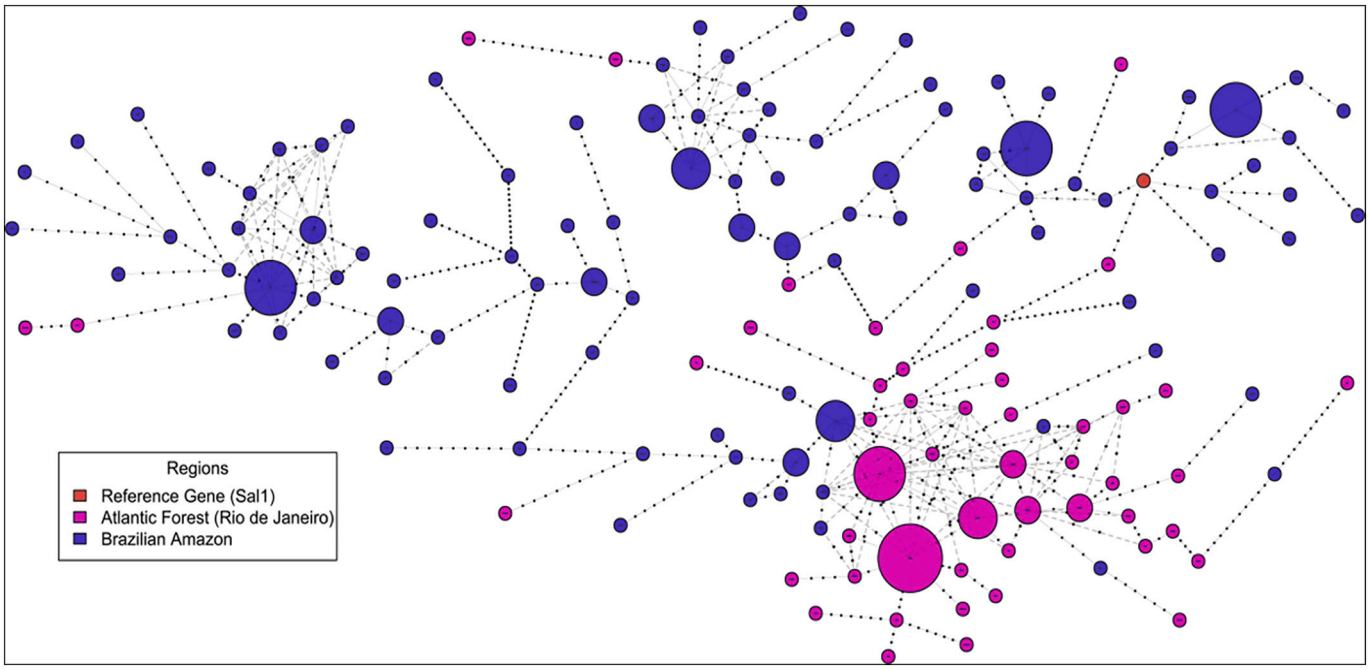

Fig. 3. Haplotype network of *P. vivax msp*-1_{42}-gene from Brazilian Amazon and Rio de Janeiro Atlantic Forest regions.
Node sizes are proportional to haplotype frequency, and branch lengths are indicative of the number of single nucleotide differences (SNPs) between sequences. Node colors indicate the geographic origin of the isolates.

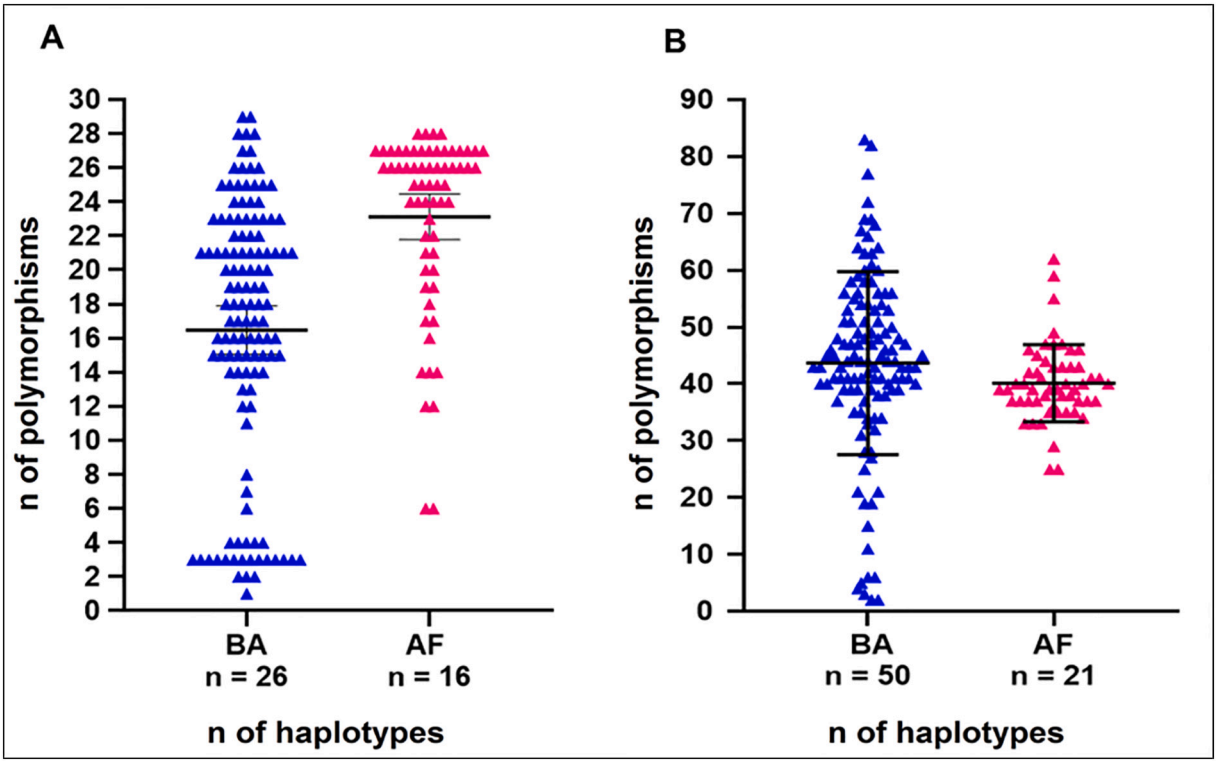

Fig. 4. The number of accumulated polymorphisms in haplotypes according to Brazilian Amazon and Rio de Janeiro Atlantic Forest regions.
A: *pvmsp-*1_{42} and B: *pvmsp-7E*.

reference sequence (Supplementary material 7). The nucleotide diversity for all *pvmsp-7F* sequences was $\pi = 0.0007$; similar results were achieved even when samples from distinct areas were analyzed separately: $\pi = 0.0004$ in AF and $\pi = 0.0007$ in BA (Fig. 2). The haplotype diversity was 0.675 for BA and 0.217 for AF isolates (Table 2).

Four haplotypes are shared between BA and AF. Most samples (94/48%) display gene sequences identical to the Sal-1 reference gene: 58 (83%) in AF and 36 (29%) in BA. The H7F-1 haplotype (T204**A**/I352**F**) was detected in 57 (29%) isolates: seven (10%) in AF and 50 (40%) in BA. The H7F-2 (T204**A**/**I**352F) was found in 42 (21%) samples: four (6%) in AF and 38 (30%) in BA. The H7F-3 (**T**204A/I352**F**) was detected only in three (2%) samples, one from AF and two from BA (Fig. 6

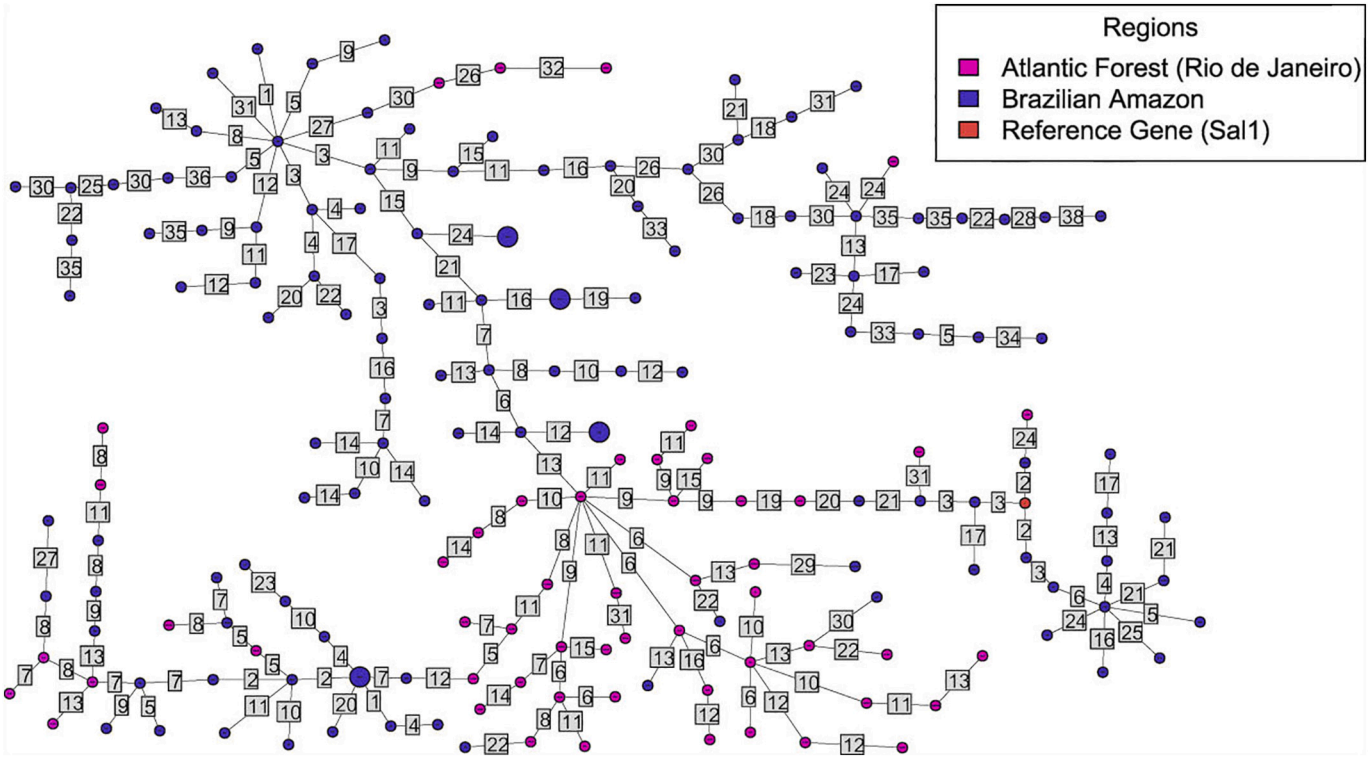

Fig. 5. Haplotype network of *P. vivax msp-7E*-gene of *P. vivax* from Brazilian Amazon and Rio de Janeiro Atlantic Forest regions.
Node sizes are proportional to haplotype frequency, and branch lengths are indicative of the number of single nucleotide differences (SNPs) between sequences. Node colors indicate the geographic origin of the isolates.

Table 3
Polymorphisms in *pvmsp-7E*-gene coding B-cell epitopes.

Epitope	Peptide	Position	SNP n (S/NS)	π	Samples N (%)
B-1	SEKLGVQKKKKNLEQDATHA	15-34	0	NA	0
B-2	IGQSKG(K/N)IKGQA(D/V)(T/A)DN(Q/E)AQR	119-138	9 (5S/ 4 NS)	0.037 (±0.003)	76 (41%)
B-3	AA(Q/P)(P/H/Q)G(G/R)(V/E)(S/L)(P/S)(S/A)(T/A)(S/R/G)(A/G)(R/Q)(P/S)(Q/R)(E/D)(P/S/T)(G/A)(K/R)	143-162	21 (4S/ 17NS)	**0.148** (±0.004)	178 (97%)
B-4	(Q/E)NVG(P/A/H/D)NGQ(R/P)(A/T/V/I)(A/V)D(P/T/R/S)Q(P/S)(G/R)(R/P)AA(N/T)	184-203	12 (2S/ 10NS)	**0.123** (±0.002)	175 (96%)
B-5	N(D/G)(P/H)QQG(G/E)(S/R)(E/A)(S/P)TE(G/R)(P/T)AVTP(R/P/S)P	211-230	10 (1S/ 9NS)	**0.078** (±0.002)	170 (93%)
B-6	DEVLTT(S/T)D(N/S)T(N/K)GIHVPDYHS	252-271	8 (5S/ 3NS)	0.046 (±0.002)	157 (86%)
B-7	NTIRQKYEYSMNPVEYEIVK	274-293	1S	0.003 (±0.002)	22 (12%)
B-8	LNVGFK(N/K)(D/E)G(A/G)(A/T)(S/P)S(D/A)A(T/A)(P/S)LVD	296-315	11 (3S/ 8NS)	**0.066** (±0.002)	108 (56%)

Predicted B-cell epitopes (Weng Cheng et al., 2018); nonsynonymous substitutions between parenthesis in red; synonymous substitutions in bold blue; n: maximum number of polymorphisms in the peptide sequence; S: synonymous; NS: nonsynonymous; π: nucleotide diversity. Nucleotide diversity values in bold were higher than those of entire fragment in all the samples ($\pi = 0.041$); N: number of isolates that presented at least one polymorphism.

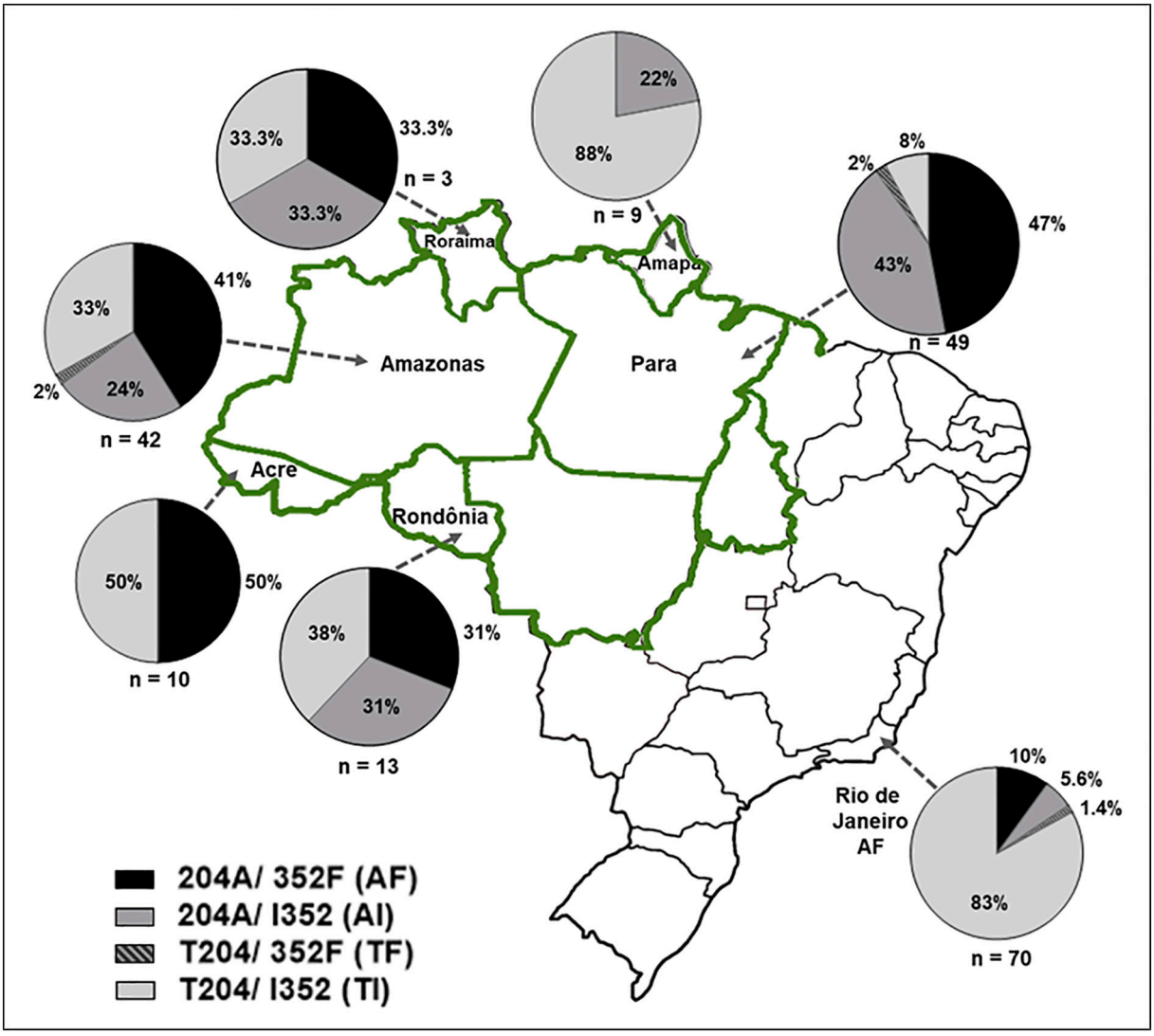

Fig. 6. Geographic distribution of the *pvmsp-7F* haplotypes in Brazilian Amazon and Rio de Janeiro Atlantic Forest. Haplotypes are represented by distinct colors (grayscale and black) or pattern (grid). The haplotype frequency is indicated in the circles. The BA municipalities include Acre, Amapá, Amazonas, Para, Rondônia, and Roraima.

and Supplementary Fig. 8).

The selective pressure estimated by Tajima's D test indicated positive selection in all 196 isolates (2.357; $p < 0.05$); stratified by region, only BA showed a significant value (2.166; $p < 0.05$) (Table 2 and Supplementary Material 6).

The genetic differentiation between BA and AF populations produced a high Fst value (0.354), but the low number of *pvmsp-7F* polymorphisms prevented a phylogenetic tree reconstruction.

4. Discussion

The investigation of nucleotide diversity and natural selection of *pvmsp-1*$_{42}$, *pvmsp-7E*, and *pvmsp-7F* genes is crucial to recognize the population structure and in-country parasite lineages, perhaps helping select potential vaccine targets in the future. It is the first report on the genetic diversity of *pvmsp-1*$_{42}$, *pvmsp-7E*, and *pvmsp-7F* genes in Brazilian regions.

This work involves *Plasmodium* parasites from distinct Brazilian areas (BA and AF), separated by a considerable geographic distance. The “Legal Amazon” is a social-economic division that includes BA, representing approximately 60% of the Brazilian territory, a malaria-endemic area comprising more than 99% of the cases (De Pina-costa et al., 2014). Approximately 90% of malaria cases occurring in BA corresponds to the ubiquitous *P. vivax s.l.* Rio de Janeiro's Atlantic Forest has a markedly smaller area than BA, corresponding only to 0.1% of the national territory, and the autochthonous malaria cases in AF regions correspond to a zoonotic bromeliad malaria transmission (Brasil et al., 2017).

The genetic diversity of *pvmsp-1*$_{42}$ and *pvmsp-7 E/F* genes from BA parasite's populations was higher than of AF's. Possibly, the transmission intensity, together with the most extensive geographic distribution of BA samples, promoting human migration in-between, could increase the levels of sexual recombination in this region, generating greater diversity in BA. The analysis of *pvmsp-1*$_{42}$ ($\pi = 0.018$) and *pvmsp-7E* ($\pi = 0.041$) DNA sequences revealed a larger number of nonsynonymous polymorphisms and absence of Sal-1 type sequence, contrasting with *pvmsp-7F* ($\pi = 0.0007$) that displayed a small number of polymorphisms and, therefore, a high similarity with Sal-1 type strain.

The *pvmsp-1*$_{42}$ and *pvmsp-7E* polymorphisms were somehow expected, once genes encoding surface proteins - usually immune response targets - accumulate polymorphisms by the genetic evolution process, which could alter the protein sequence, thereby allowing parasite evasion of the host's immune response. Nucleotide diversity of *pvmsp-7E*-gene was slightly lower than Colombia ($\pi = 0.055$), Thailand

(π = 0.061) (Garzón-ospina et al., 2014; Weng Cheng et al., 2018), North Korean, India VII, Mauritania I, and Brazil I (π = 0.051) (Neafsey et al., 2012).

pvmsp-7F-gene high conservation corroborates previous findings from Colombia (Garzón-ospina et al., 2014), in which samples presented the same two polymorphisms and nucleotide diversity, as well as from *P. vivax* sequences available in PlasmoDB (π = 0.001) (Castillo et al., 2017). Conversely, 19KDa portion of *pvmsp-1*$_{42}$, 5′ and 3′ ends in *pvmsp-7E*, as well as *pvmsp-7F* are highly conserved in *P. vivax* isolates, indicating that these fragments may have been exposed to different immune system's selective pressure and could play relevant roles in parasite's survival and transmission (Garzón-Ospina et al., 2012; Weng Cheng et al., 2019).

The nucleotide diversity of *pvmsp-1*$_{42}$C-terminal fragment in Brazil is similar to that found in Nicaragua (π = 0.020) (Gutiérrez et al., 2016), Cambodia (π = 0.020), India (π = 0.021), Thailand (π = 0.020) (Parobek et al., 2014), Korea (π = 0.016) (Kang et al., 2012), and China-Myanmar (π = 0.019) (Zhou et al., 2017). All polymorphic sites are in 33 KDa fragment, unevenly distributed, mostly concentrated at 1475 and 1541 amino acid positions, as previously observed in Africa and Asia (Dias et al., 2011b; Pacheco et al., 2007; Thakur et al., 2008).

We then focus on T-cell 33 kDa's region epitopes to investigate if these polymorphisms could affect the antibody response. The promiscuous immunodominant T-cell epitope (DYDVVYLKPLAGMYK) (Rosa et al., 2006) is highly conserved in all isolates independently of the studied parasite. These data seem to indicate that the helper T-cell epitope could promote an effective response to the MSP-1-based vaccine to face both the dominant *P. vivax* in the rainfall Amazon Basin and the unusual *P. vivax* in the Atlantic Forest regions.

The lack of polymorphism in the 19 KDa fragment was also reported in parasites from French Guiana (Bonilla et al., 2006), Nicaragua (Gutiérrez et al., 2016), Mexico (González-Cerón et al., 2015), Cambodia (Parobek et al., 2014), Korea (Kang et al., 2012), and Thailand (Putaporntip et al., 2002). Since the 19 KDa fragment plays a role in reticulocyte invasion and can induce anti-19 kDa IgG response, including specific IgG3 antibodies (Cassiano et al., 2016), and a high cellular activation (Soares et al., 1997; Riccio et al., 2013) in naturally exposed individuals, this fragment also has potential to be a vaccine target for Brazilian populations, as well.

pvmsp-1$_{33}$ and *pvmsp-7E* display significant differences in their rates of nonsynonymous substitutions compared to synonymous substitutions rates, indicating positive selection. Tajima's D and *Z*-test indicate positive selection acting on the *pvmsp-1*$_{42}$ fragment. However, when segregated by geographic region, *P. vivax* samples from AF were considered under positive selection only by the Z-test. These contradictory results on *pvmsp-1*$_{42}$-gene could reflect the fact that polymorphisms in the AF region are not fixed in this parasite population and, therefore, such mutations on these antigens could be in balancing selection as a result of transient directional selection.

Since all polymorphic sites of *pvmsp-1*$_{42}$-gene were at the *pvmsp-1*$_{33}$, this fragmented was, in the same way, under positive selection and *pvmsp-1*$_{19}$ keep on neutrality, as previously reported in China, Myanmar (Zhou et al., 2017), Cambodia (Parobek et al., 2014), Korea (Kang et al., 2012), and Sri Lanka (Dias et al., 2011a). This may reflect the protein's tri-dimensional structure, since the 33 kDa portion covers the 19 kDa one, limiting its exposure to the human immune system (Parobek et al., 2014).

The central region of *pvmsp-7E*-gene presented positive values in the Z-test and Tajima's D; these results, together with high nucleotide and haplotype diversities and low Fst, indicate balancing selection in parasite's populations of AF and BA. Conversely, 5′ and 3′ ends are under negative selection according to Z-test. Positive and negative selections acting along the *pvmsp-7E*-gene were also highlighted in parasite's populations from Colombia (Garzón-Ospina et al., 2014), Thailand (Weng Cheng et al., 2018), in clinical isolates from several geographic regions, as well as in PlasmoDB reference sequences (Salvador I, North Korean, India VII, Mauritania I, and Brazil I) (Neafsey et al., 2012). The restrictive polymorphism pattern in *pvmsp-7E* 5′ and 3′ ends, leading to negative selection, could be due to the need for essential amino acid conservation to maintain protein function.

For *pvmsp-7F*-gene, positive selection was verified only in BA parasites with Tajima's D, which could be an indication that the two SNPs found could be fixed in these populations. Conversely, in AF parasites, the *pvmsp-7F* polymorphisms indicate neutral selection according to Tajima's D test, which is somewhat expected due to the lowest frequency of SNPs in AF parasite's populations. Since the measurement of the dN/dS ratio is not indicated for analysis of sequences with low diversity, the Z-test analysis was not performed in this gene.

The haplotype network and phylogenetic tree confirmed the extensive genetic diversity observed in *pvmsp-1*$_{42}$ and *pvmsp-7E* sequences. These findings are consistent with the quite different levels of endemicity in the AF and AB regions. AF isolates exhibited a robust clustering, while BA's samples showed dispersion throughout the clades, which could be related to the higher level of transmission in this region, promoting parasite recombination, consequently increasing the genetically diverse population, differing from AF's parasites, an inbreeding population.

Independently of the locality, no variation in genetic diversity of genes *pvmsp-1*$_{42}$, *pvmsp-7E*, and *pvmsp-7F* was observed for years in parasites from BA or AF, which could suggest spatial but not temporal genetic variations, similarly to previous studies with these genes in Thailand (Weng Cheng et al., 2018) and Cambodia (Parobek et al., 2014).

Nonsynonymous polymorphisms detected in Brazilian regions are reported for the first time: *pvmsp-1*$_{42}$ SNPs in AF (K1523Q) and BA (V1516G; K1612R; G1482R; N1517G), as well as, *pvmsp-7E* in AF (S258T) and BA (T139A, S154R; S231T). Indeed, the interaction between the parasite and different hosts has a profound impact on parasite's genome, and could generate characteristic “signatures” of selection, most likely by selective pressures, as the host immunity response (Parobek et al., 2014), through numerous mutational steps, generates different haplotypes in both BA and AF. The knowledge of the complex evolutionary history of non-human primate malaria could help in this issue.

We conclude that *pvmsp-1*$_{42}$ 33KDa and *pvmsp-7E*-gene central region fragments have tremendous genetic diversity, even in an enclosed area with deficient transmission levels as Rio de Janeiro's AF, most likely due to its roles in parasite's survival and transmission. On the other hand, 19KDa *pvmsp-1*$_{42}$, *pvmsp-7E* 5′-3′ ends, and *pvmsp-7F* were conserved in isolates from both Brazilian regions and, therefore, could be considered potential targets to a *P. vivax* blood-stage vaccine.

Ethics approval and consent to participate

The study was approved by the Ethics Research Committee of Oswaldo Cruz Institute, Fiocruz, Brazil. Reference number CAEE: 69256317.3.0000.5248.

Acknowledgements

We would like to thank the patients for agreeing to participate in this study and, the health staff for helping with data collection and microscopic diagnosis. We also thank the staff from the Genomic Platform for DNA sequencing Network Technology Platforms RPT01A/PDTIS/Fiocruz.

We would like to express our thanks to the funding organizations: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; https://www.capes.gov.br/) by Doctoral Fellowship that NKAO is recipient; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; http://www.cnpq.br/) by Research Productivity Fellowships whom CTDR (310445/2017-5) and MFFC (306025/2018-

3) are recipients;"Cientistas do Nosso Estado" from the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado de Rio de Janeiro (FAPERJ; http://www.faperj.br/), number grant CTDR (E-26/202.921/2018) and MFFC (E-26/203.295/2015) are. This work was also supported by the IOC/FIOCRUZ of Health ministry of Brazil, CPD-Mal-Center of reference malaria diagnoses at Oswaldo Cruz Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2020.104592.

References

Bonilla, J.A., Validum, L., Cummings, R., Palmer, C.J., 2006. Genetic diversity of *Plasmodium vivax* PVCSP and PVMSP1 in Guyana, South America. Am. J. Trop. Med. Hyg. 75, 830–835.

Brasil, P., Zalis, M.G., de Pina-Costa, A., Siqueira, A.M., Bianco Júnior, C., Silva, S., Areas, A.L.L., Pelajo-Machado, M., Alvarenga, D.A., Santelli, A.C.F., Albuquerque, H.G., Cravo, P., Abreu, F.V.S., Leonel Peterka, C., Maria Zanini, G., Suárez Mutis, M.C., Pissinatti, A., Lourenço-de-Oliveira, R., Brito, C.F.A., Ferreira-da-Cruz, M. de F., Culleton, R., Daniel-Ribeiro, C.T., 2017. Outbreak of human malaria caused by *Plasmodium simium* in the Atlantic Forest in Rio de Janeiro: a molecular epidemiological investigation. Lancet Glob. Health 5, e1038–e1046. https://doi.org/10.1016/S2214-109X(17)30333-9.

Cassiano, G.C., Furini, A.A.C., Capobianco, M.P., Storti-Melo, L.M., Almeida, M.E., Barbosa, D.R.L., Póvoa, M.M., Nogueira, P.A., Machado, R.L.D., 2016. Immunogenetic markers associated with a naturally acquired humoral immune response against an N-terminal antigen of *Plasmodium vivax* merozoite surface protein 1 (PvMSP-1). Malar. J. 15, 306. https://doi.org/10.1186/s12936-016-1350-2.

Castillo, A.I., Andreína Pacheco, M., Escalante, A.A., 2017. Evolution of the merozoite surface protein 7 (msp7) family in *Plasmodium vivax* and *P. falciparum*: A comparative approach. Infect. Genet. Evol. 50, 7–19. https://doi.org/10.1016/j.meegid.2017.01.024.

Collins, W.E., Kaslow, D.C., Sullivan, J.S., Morris, C.L., Galland, G.G., Yang, C., Saekhou, A.M., Xiao, L., Lal, A.A., 1999. Testing the efficacy of a recombinant merozoite surface protein (MSP- 119) of *Plasmodium vivax* in Saimiri boliviensis monkeys. Am. J. Trop. Med. Hyg. 60, 350–356. https://doi.org/10.4269/ajtmh.1999.60.350.

de Abreu, F.V.S., Santos, E. Dos, Mello, A.R.L., Gomes, L.R., de Alvarenga, D.A.M., Gomes, M.Q., Vargas, W.P., Bianco-Júnior, C., de Pina-Costa, A., Teixeira, D.S., Romano, A.P.M., de Manso, P.P.A., Pelajo-Machado, M., Brasil, P., Daniel-Ribeiro, C.T., de Brito, C.F.A., de Ferreira-da-Cruz, M.F., Lourenço-de-Oliveira, R., 2019. Howler monkeys are the reservoir of malarial parasites causing zoonotic infections in the Atlantic forest of Rio de Janeiro. PLoS Negl. Trop. Dis. 13, e0007906. https://doi.org/10.1371/journal.pntd.0007906.

De Pina-costa, A., Brasil, P., Maria, S., Santi, D., De Araujo, M.P., Suárez-mutis, M.C., Carolina, A., Santelli, S., Oliveira-ferreira, J., Lourenço-de-oliveira, R., Daniel-ribeiro, C.T., 2014. Malaria in Brazil : What Happens outside the Amazonian Endemic Region 109. pp. 618–633. https://doi.org/10.1590/0074-0276140228.

Deane, L.M., 1992. Simian malaria in Brazil. Mem. Inst. Oswaldo Cruz 87 (Suppl. 3), 1–20. https://doi.org/10.1590/S0074-02761992000700001.

Dias, S., Longacre, S., Escalante, A.a., Udagama-Randeniya, P.V., 2011a. 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. https://doi.org/10.1016/j.meegid.2010.09.007.

Dias, S., Somarathna, M., Manamperi, A., Escalante, A.A., Gunasekera, A.M., Udagama, P.V., 2011b. Evaluation of the genetic diversity of domain II of *Plasmodium vivax* Apical Membrane Antigen 1 (PvAMA-1) and the ensuing strain-specific immune responses in patients from Sri Lanka. Vaccine 29, 7491–7504. https://doi.org/10.1016/j.vaccine.2011.07.029.

Fonseca, J.A., Cabrera-Mora, M., Singh, B., Oliveira-Ferreira, J., da Costa Lima-Junior, J., Calvo-Calle, J.M., Lozano, J.M., Moreno, A., 2016. A chimeric protein-based malaria vaccine candidate induces robust T cell responses against *Plasmodium vivax* MSP119. Sci. Rep. 6, 34527. https://doi.org/10.1038/srep34527.

Garzón-Ospina, D., López, C., Forero-Rodríguez, J., Patarroyo, M.A., 2012. Genetic diversity and selection in three *Plasmodium vivax* merozoite surface protein 7 (Pvmsp-7) genes in a Colombian population. PLoS One 7, 1–10. https://doi.org/10.1371/journal.pone.0045962.

Garzón-ospina, D., Forero-rodríguez, J., Patarroyo, M.A., 2014. Heterogeneous genetic diversity pattern in *Plasmodium vivax* genes encoding merozoite surface proteins (MSP) -7E, -7F and -7L. Malar. J. 13, 495. https://doi.org/10.1186/1475-2875-13-495.

González-Cerón, L., Cerritos, R., Corzo-Mancilla, J., Santillán, F., 2015. Diversity and evolutionary genetics of the three major *Plasmodium vivax* merozoite genes participating in reticulocyte invasion in southern Mexico. Parasit. Vectors 8, 651. https://doi.org/10.1186/s13071-015-1266-7.

Gutiérrez, S., González-Cerón, L., Montoya, A., Sandoval, M.A., Tórres, M.E., Cerritos, R., 2016. Genetic structure of Plasmodium vivax in Nicaragua, a country in the control phase, based on the carboxyl terminal region of the merozoite surface protein-1. Infect. Genet. Evol. 40, 324–330. https://doi.org/10.1016/j.meegid.2015.08.040.

Han, H.J., Park, S.G., Kim, S.H., Hwang, S.Y., Han, J., Traicoff, J., Kho, W.G., Chung, J.Y., 2004. Epidermal growth factor-like motifs 1 and 2 of *Plasmodium vivax* merozoite surface protein 1 are critical domains in erythrocyte invasion. Biochem. Biophys. Res. Commun. 320, 563–570. https://doi.org/10.1016/j.bbrc.2004.06.008.

Holder, A.A., Blackman, M.J., Burghaus, P.A., Chappel, J.A., Ling, I.T., McCallum-Deighton, N., Shai, S., 1992. A malaria merozoite surface protein (MSP1)-structure, processing and function. Mem. Inst. Oswaldo Cruz. https://doi.org/10.1590/S0074-02761992000700004.

Hudson, R.R., Slatkint, M., Maddison, W.P., 1992. Estimation of levels of gene flow from DNA sequence data. Genetics 132, 583–589.

Kadekoppala, M., O'Donnell, R.a., Grainger, M., Crabb, B.S., Holder, A.a., 2008. Deletion of the *Plasmodium falciparum* merozoite surface protein 7 gene impairs parasite invasion of erythrocytes. Eukaryot. Cell 7, 2123–2132. https://doi.org/10.1128/EC.00274-08.

Kang, J.-M., Ju, H.-L., Kang, Y.-M., Lee, D.-H., Moon, S.-U., Sohn, W.-M., Park, J.-W., Kim, T.-S., Na, B.-K., 2012. Genetic polymorphism and natural selection in the C-terminal 42 kDa region of merozoite surface protein-1 among *Plasmodium vivax* Korean isolates Jung-Mi. Malar. J. 11. https://doi.org/10.1016/j.meegid.2016.07.040.

Kevin Baird, J., 2013a. Evidence and implications of mortality associated with acute *Plasmodium vivax* malaria. Clin. Microbiol. Rev. 26, 36–57. https://doi.org/10.1128/CMR.00074-12.

Kevin Baird, J., 2013b. Malaria caused by *Plasmodium vivax*: recurrent, difficult to treat, disabling, and threatening to life - averting the infectious bite preempts these hazards. Pathog. Glob. Health 107, 475–479. https://doi.org/10.1179/2047772413Z.000000000179.

Lacerda, M.V.G., Mourão, M.P.G., Alexandre, M.A.A., Siqueira, A.M., Magalhães, B.M.L., Martinez-Espinosa, F.E., Santana Filho, F.S., Brasil, P., MRS Ventura, A., Tada, M.S., SCD Couto, V., Silva, A.R., Silva, R.S.U., Alecrim, M.G.C., Filho, F.S.S., Brasil, P., Ventura, A.M.R.S., Tada, M.S., Couto, V.S.C.D., Silva, A.R., Silva, R.S.U., Alecrim, M.G.C., 2012. Understanding the clinical spectrum of complicated *Plasmodium vivax* malaria: a systematic review on the contributions of the Brazilian literature. Malar. J. 11, 12. https://doi.org/10.1186/1475-2875-11-12.

Mello, K., Daly, T.M., Morrisey, J., Vaidya, A.B., Long, C.A., Bergman, L.W., 2002. A multigene family that interacts with the amino terminus of *Plasmodium* MSP-1 identified using the yeast two-hybrid system. Eukaryot. Cell 1, 915–925. https://doi.org/10.1128/EC.1.6.915-925.2002.

Ministério da Saúde, 2019. Epidemiological Report of Malaria Cases in Brazil 2019. Ministério da Saúde.

Mongui, A., Perez-Leal, O., Soto, S.C., Cortes, J., Patarroyo, M.A., 2006. Cloning, expression, and characterisation of a Plasmodium vivax MSP7 family merozoite surface protein. Biochem. Biophys. Res. Commun. 351, 639–644. https://doi.org/10.1016/j.bbrc.2006.10.082.

Neafsey, D.E., Galinsky, K., Jiang, R.H.Y., Young, L., Sykes, S.M., Saif, S., Gujja, S., Goldberg, J.M., Young, S., Zeng, Q., Dash, A.P., Anvikar, A.R., Sutton, P.L., Birren, W., Escalante, A.A., Barnwell, J.W., Carlton, J.M., 2012. The malaria parasite *Plasmodium vivax* exhibits greater genetic diversity than *Plasmodium falciparum*. Nat. Genet. 44, 1046–1050. https://doi.org/10.1038/ng.2373.The.

Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426. https://doi.org/10.1093/oxfordjournals.molbev.a040410.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. NY, New York.

WHO, 2018. World Malaria Report 2018. World Health Organization, Geneva.

Pacheco, M.A., Poe, A.C., Collins, W.E., Lal, A.A., Tanabe, K., Kariuki, S.K., Udhayakumar, V., Escalante, A.A., 2007. A comparative study of the genetic diversity of the 42kDa fragment of the merozoite surface protein 1 in *Plasmodium falciparum* and *P. vivax*. Infect. Genet. Evol. 7, 180–187. https://doi.org/10.1016/j.meegid.2006.08.002.

Paradis, E., Barrett, J., 2010. pegas: an R package for population genetics with an integrated-modular approach. Bioinforma. Appl. NOTE 26, 419–420. https://doi.org/10.1093/bioinformatics/btp696.

Park, J.W., Moon, S.H., Yeom, J.S., Lim, K.J., Sohn, M.J., Jung, W.C., Cho, Y.J., Jeon, K.W., Ju, W., Ki, C.S., Oh, M.D., Choe, K., 2001. Naturally acquired antibody responses to the C-terminal region of merozoite surface protein 1 of *Plasmodium vivax* in Korea. Clin. Diagn. Lab. Immunol. 8, 14–20. https://doi.org/10.1128/CDLI.8.1.14-20.2001.

Parobek, C.M., Bailey, J.a., Hathaway, N.J., Socheat, D., Rogers, W.O., Juliano, J.J., 2014. Differing patterns of selection and geospatial genetic diversity within two leading *Plasmodium vivax* candidate vaccine antigens. PLoS Negl. Trop. Dis. 8. https://doi.org/10.1371/journal.pntd.0002796.

Price, R.N., von Seidlein, L., Valecha, N., Nosten, F., Baird, J.K., White, N.J., 2014. Global extent of chloroquine-resistant *Plasmodium vivax*: A systematic review and meta-analysis. Lancet Infect. Dis. 14, 982–991. https://doi.org/10.1016/S1473-3099(14)70855-2.

Putaporntip, C., Jongwutiwes, S., Sakihama, N., Ferreira, M.U., Kho, W.-G., Kaneko, A., Kanbara, H., Hattori, T., Tanabe, K., 2002. Mosaic organization and heterogeneity in frequency of allelic recombination of the *Plasmodium vivax* merozoite surface protein-1 locus. Proc. Natl. Acad. Sci. U. S. A. 99, 16348–16353. https://doi.org/10.1073/pnas.252348999.

Rahimi, B.A., Thakkinstian, A., White, N.J., Sirivichayakul, C., Dondorp, A.M., Chokejindachai, W., 2014. Severe vivax malaria: a systematic review and meta-analysis of clinical studies since 1900. Malar. J. 13, 481. https://doi.org/10.1186/1475-2875-13-481.

Riccio, E.K.P., Totino, P.R.R., Pratt-Riccio, L.R., Ennes-Vidal, V., Soares, I.S., Rodrigues, M.M., de Souza, J.M., Daniel-Ribeiro, C.T., Ferreira-da-Cruz, M.D.F., 2013. Cellular and humoral immune responses against the *Plasmodium vivax* MSP-1_{19} malaria vaccine candidate in individuals living in an endemic area in north-eastern Amazon region of Brazil. Malar. J. 12, 326. https://doi.org/10.1186/1475-2875-12-326.

Rosa, D.S., Iwai, L.K., Tzelepis, F., Bargieri, D.Y., Medeiros, M.A., Soares, I.S., Sidney, J.,

Sette, A., Kalil, J., Mello, L.E., Cunha-Neto, E., Rodrigues, M.M., 2006. Immunogenicity of a recombinant protein containing the *Plasmodium vivax* vaccine candidate MSP119 and two human CD4+ T-cell epitopes administered to non-human primates (*Callithrix jacchus jacchus*). Microbes Infect. 8, 2130–2137. https://doi.org/10.1016/j.micinf.2006.03.012.

Rougeron, V., Elguero, E., Arnathau, C., Acuña Hidalgo, B., Durand, P., Houze, S., et al., 2020. Human *Plasmodium vivax* diversity, population structure and evolutionary origin. PLoS Negl. Trop. Dis. 14 (3), e0008072. https://doi.org/10.1371/journal.pntd.0008072.

Rozas, J., Ferrer-Mata, A., Carlos Anchez-DelBarrio, J.S., Guirao-Rico, S., Librado, P., An Ramos-Onsins, S.E., Anchez-Gracia, A.S., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34, 3299–3302. https://doi.org/10.1093/molbev/msx248.

Saitou, N., Nei, M., 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454.

Soares, I.S., Levitus, G., Souza, J.M., Del Portillo, H.A., Rodrigues, M.M., 1997. Acquired immune responses to the N- and C-terminal regions of *Plasmodium vivax* merozoite surface protein 1 in individuals exposed to malaria. Infect. Immun. 65, 1606–1614. https://doi.org/10.1128/iai.65.5.1606-1614.1997.

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

Tazi, L., Ayala, F., 2011. Unresolved direction of host transfer of *Plasmodium vivax* v. *P. simium* and *P. malariae* v. *P. brasilianum.* Infection. Genet. Evol. 11, 209–221. https://doi.org/10.1016/j.meegid.2010.08.007.

Thakur, A., Alam, M.T., Sharma, Y.D., 2008. Genetic diversity in the C-terminal 42kDa region of merozoite surface protein-1 of *Plasmodium vivax* (PvMSP-1_{42}) among Indian isolates. Acta Trop. 108, 58–63. https://doi.org/10.1016/j.actatropica.2008.08.011.

Torres, K.L., Figueiredo, D.V., Zalis, M.G., Daniel-Ribeiro, C.T., Alecrim, W., De Fátima Ferreira-da-Cruz, M., 2006. Standardization of a very specific and sensitive single PCR for detection of *Plasmodium vivax* in low parasitized individuals and its usefulness for screening blood donors. Parasitol. Res. 98, 519–524. https://doi.org/10.1007/s00436-005-0085-8.

Weng Cheng, C., Putaporntip, C., Jongwutiwes, S., 2018. Polymorphism in merozoite surface protein-7E of *Plasmodium vivax* in Thailand: natural selection related to protein secondary structure. PLoS One 13, e0196765. https://doi.org/10.1371/journal.pone.0196765.

Weng Cheng, C., Jongwutiwes, S., Putaporntip, C., Jackson, A.P., 2019. Clinical expression and antigenic profiles of a *Plasmodium vivax* vaccine candidate: merozoite surface protein 7 (PvMSP-7). Malar. J. 18, 197. https://doi.org/10.1186/s12936-019-2826-7.

World Health Organization, 2019. World Malaria Report 2019.

Zhou, X., Tambo, E., Su, J., Fang, Q., Ruan, W., Chen, J.H., Yin, M.B., Zhou, X.N., 2017. Genetic diversity and natural selection in 42 kDa region of *Plasmodium vivax* merozoite surface protein-1 from China-Myanmar endemic border. Korean J. Parasitol. 55, 473–480. https://doi.org/10.3347/kjp.2017.55.5.473.