

## RESEARCH ARTICLE

Balancing selection and high genetic diversity of *Plasmodium vivax* circumsporozoite central region in parasites from Brazilian Amazon and Rio de Janeiro Atlantic Forest

Natália Ketrin Almeida-de-Oliveira<sup>1,2</sup>, Rebecca de Abreu-Fernandes<sup>1,2</sup>, Lidiane Lima-Cury<sup>1,2</sup>, Aline Rosa de Lavigne<sup>1,2</sup>, Anielle de Pina-Costa<sup>2,3,4</sup>, Daiana de Souza Perce-da-Silva<sup>1,2</sup>, Marcos Catanho<sup>5</sup>, Atila Duque Rossi<sup>6</sup>, Patrícia Brasil<sup>2,3</sup>, Cláudio Tadeu Daniel-Ribeiro<sup>1,2</sup>, Maria de Fátima Ferreira-da-Cruz<sup>1,2\*</sup>



**1** Laboratório de Pesquisa em Malária, Instituto Oswaldo Cruz (IOC), Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro, Brazil, **2** 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, Brazil, **3** Laboratório de Pesquisa Clínica em Doenças Febris Agudas, Instituto Nacional de Infectologia Evandro Chagas, Fiocruz, Rio de Janeiro, Brazil, **4** Centro Universitário Serra dos Órgãos (UNIFESO), Teresópolis, Rio de Janeiro, Brazil, **5** Laboratório de Genética Molecular de Microrganismos, IOC, Fiocruz, Rio de Janeiro, Brazil, **6** Laboratório de Virologia Molecular, Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, RJ, Brazil

\* [mffcruz28@gmail.com](mailto:mffcruz28@gmail.com)

## OPEN ACCESS

**Citation:** Almeida-de-Oliveira NK, Abreu-Fernandes Rd, Lima-Cury L, Lavigne ARd, Pina-Costa Ad, Perce-da-Silva DdS, et al. (2020) Balancing selection and high genetic diversity of *Plasmodium vivax* circumsporozoite central region in parasites from Brazilian Amazon and Rio de Janeiro Atlantic Forest. PLoS ONE 15(11): e0241426. <https://doi.org/10.1371/journal.pone.0241426>

**Editor:** Arnar Palsson, University of Iceland, ICELAND

**Received:** October 22, 2019

**Accepted:** October 14, 2020

**Published:** November 9, 2020

**Copyright:** © 2020 Almeida-de-Oliveira et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Sequences were deposited in Genbank™ with accession nos. MN417517-MN417735.

**Funding:** NKAO is recipient of a Doctoral Fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; <https://www.capes.gov.br/>); RAF is recipient of a Scientific Initiation with support of Scholarship from the Fiocruz/ Centro de Integração

## Abstract

Circumsporozoite protein (CSP) is the primary pre-erythrocytic vaccine target in *Plasmodium* species. Knowledge about their genetic diversity can help predict vaccine efficacy and the spread of novel parasite variants. Thus, we investigated *pvcsp* gene polymorphisms in 219 isolates (136 from Brazilian Amazon [BA], 71 from Rio de Janeiro Atlantic Forest [AF], and 12 from non-Brazilian countries [NB]). Forty-eight polymorphic sites were detected, 46 in the central repeat region (CR), and two in the C-terminal region. Also, the CR presents InDels and a variable number of repeats. All samples correspond to the VK210 variant, and 24 VK210 subtypes based on CR. Nucleotide diversity ( $\pi = 0.0135$ ) generated a significant number of haplotypes (168) with low genetic differentiation between the Brazilian regions ( $F_{st} = 0.208$ ). The haplotype network revealed similar distances among the BA and AF regions. The linkage disequilibrium indicates that recombination does not seem to be acting in diversity, reinforcing natural selection's role in accelerating adaptive evolution. The high diversity (low  $F_{st}$ ) and polymorphism frequencies could be indicators of balancing selection. Although malaria in BA and AF have distinct vector species and different host immune pressures, consistent genetic signature was found in two regions. The immunodominant B-cell epitope mapped in the CR varies from seven to 19 repeats. The CR T-cell epitope is conserved only in 39 samples. Concerning to C-terminal region, the Th2R epitope presented nonsynonymous SNP only in 6% of Brazilian samples, and the Th3R epitope remained conserved in all studied regions. We conclude that, although the uneven distribution of alleles may jeopardize the deployment of vaccines directed to a specific variable *locus*, a unique vaccine formulation could protect populations in all Brazilian regions.

Empresa/Escola (CIEE; <https://portal.ciee.org.br/rj>); CTDR (310445/2017-5) and MFFC (306025/2018-3) are recipients of Research Productivity Fellowships from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; <http://www.cnpq.br/>); CTDR (E-26/202.921/2018) and MFFC (E-26/203.295/2015) are “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/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** WHO, World Health Organization; *P. vivax*, *Plasmodium vivax*; CSP, Circumsporozoite protein; *pvcsp*, gene of *Plasmodium vivax* circumsporozoite protein; AF, Atlantic Forest; BA, Brazilian Amazon; NB, Non-Brazilian samples; blast, Basic Local Alignment Search Tool; CR, Central region;  $\pi$ , Nucleotide diversity; S, polymorphic sites; Hd, Haplotype diversity; Fst, Wright’s fixation statistics; SNP, Single Nucleotide Polymorphism; sSNP, Synonymous Single Nucleotide Polymorphism; nsSNP, Nonsynonymous Single Nucleotide Polymorphism; *An*, *Anopheles*.

## Introduction

Malaria is the most prevalent infectious disease globally, with an estimated 228 million cases and 405 thousand deaths in 2018 [1]. Although *Plasmodium falciparum* is responsible for most cases and deaths from malaria, *P. vivax* is by far the most widespread, accounting for almost half of non-African malaria cases [2]. More than a third of the world’s population is at risk of infection, posing a severe threat in Asia, Oceania, and South America, where *P. vivax* is responsible for about 80% of malaria cases. In Brazil, malaria remains a public health concern, and around 153 thousand cases were recorded in 2019, of which 90% caused by *P. vivax* [3].

Brazil presents two autochthonous profiles of malaria transmission: the most important corresponds to the Amazon rainforest, where more than 99% of the malaria cases occur; the second one takes place in non-endemic regions of Atlantic Forest (AF) and represents 0.08% of all malaria cases in Brazil [4, 5]. Although malaria cases have substantially decreased in Southeast Brazil after the malaria eradication campaign in 1960 [4], residual *P. vivax* cases persist in AF areas until today, probably as a zoonotic transmission [6]. The singular epidemiological aspect based on the bromeliad-malaria model sustains the autochthonous human cases in the past few decades in Rio de Janeiro state (RJ) [6, 7]. In RJ, the vectors are *Anopheles* mosquitoes, widely found in AF, belonging to the subgenus *Kerteszia*, mainly *An. K. cruzii*, that utilize bromeliads’ retained water as habitats for their immature stages [8]. The *An. cruzii*, known for its acrodendrophilic habit, bites monkeys in the canopy and, eventually, humans at ground level. This mosquito behavior makes possible zoonotic malaria transmission in Southeastern Brazil, where monkeys are the parasite reservoir [7–9]. Genetic and morphological similarities between *P. vivax*, which infects humans, and *P. simium*, which infects monkeys in the Atlantic Forest, is consistent with host switches between them in recent evolutionary times [6, 10–12].

*P. vivax* responds slower than *P. falciparum* to control strategies due to relapse episodes induced by the dormant liver stage—the hypnozoite—and early emergence of gametocytes soon after the pre-erythrocytic phase, increasing the chances of infecting mosquitoes and humans [4, 13, 14]. Besides that, the increasing number of *P. vivax* severe cases [15] and the spread of drug resistance [16] strengthen the priority to vaccine development against vivax malaria [17, 18].

Efforts to develop a pre-erythrocytic vaccine have focused on the circumsporozoite protein (CSP), which is abundantly expressed on the sporozoite surface, and responsible for the motility and invasion of the sporozoite into hepatocytes [19]. CSP-synthetic peptides induced a high and specific humoral response similar to those naturally acquired in endemic areas [20–22]. The structure of CSP is similar among *Plasmodium* species infecting rodents, primates, and humans. The CSP sequence comprises a central repeat region (CR) that is specific to each species and two flanking conserved domains identical among plasmodial species: Region I, located in the N terminal, participates in the binding on the salivary glands of the mosquito and Region II, at C-terminal, is a cell-adhesive motif to the hepatocyte [23]. The repetitive CR biological function of CSP is unknown. The *P. vivax* CSP has three types of nonapeptides in tandem that identify the three variants: VK210 (GDRA (A/D) GQPA), VK247 (ANGA (G/D) (N/D) QPG) and vivax-like (APGANQ (E/G) GGAA) [24]. VK210 is the most prevalent and widespread worldwide [25], whereas the VK247 is detected in regions where *P. vivax* and *P. falciparum* coexist, including Brazil [26–28].

VK247 seems to be more common in South-eastern Asia [29–32]. The vivax-like variant is identical to those described for *P. simiovale*, and it had been detected in Papua New Guinea, Brazil, Indonesia, and Madagascar [27, 33, 34]. In Brazil, the vivax-like variant had been disclosed in several Amazonian states [27] and an extra-Amazonian area in the state of Maranhão [34].

Although this molecule may represent a promising vaccine candidate, the *pvcsp* gene polymorphisms may affect the host immune response [34]. Consequently, polymorphism dynamics of circulating genotypes from different geographical areas, even in a low transmission setting, could be an obstacle to the development of effective vaccines against *P. vivax* [35–37].

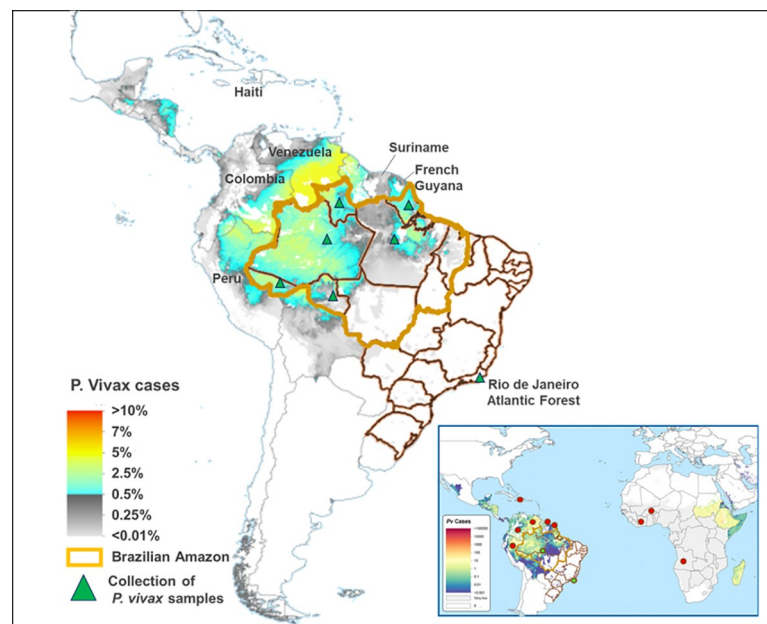
So far, there is no data on *pvcsp* gene diversity in the Atlantic Forest region. Thus, this study aimed to investigate the *pvcsp* polymorphisms in CR (286–825) and C-terminal (826–1,068) nucleotides, in parasites from the Brazilian Amazon, where the vast majority of cases occur, and also in RJ Atlantic forest that presents a particular zoonotic transmission malaria profile.

## Material and methods

### Samples

A total of 239 isolates with *P. vivax* infection diagnosed by thick blood smear and PCR [38] were tested in this study. Samples were separated according to the site of infection: Brazilian Amazon rainforest (BA), Rio de Janeiro Atlantic Forest (AF), and non-Brazilian (NB) samples (Fig 1). Fifty samples were collected in Tucuruí Unit Health (S 3° 46' W 49° 40'), a municipality in the State of Pará (BA region), while the remaining 189 samples (95 from BA, 82 from AF, and 12 from NB) were collected from patients who have attended a reference center for malaria diagnosis in the Extra-Amazonian region (Centro de Referência para Tratamento e Diagnóstico da Malária—CPD-Mal/Fiocruz) (S 22° 54' W 43° 12'), between January 2011 and March 2018.

The study was approved by the Ethics Research Committee of Instituto Oswaldo Cruz, Fiocruz, Brazil (69256317.3.0000.5248). All volunteers signed a written informed consent before 5 mL of venous blood collection. The DNA was extracted from 1 mL blood samples using



**Fig 1. Sampling *P. vivax* collection in Brazil and foreign countries.** Rio de Janeiro Atlantic Forest and Brazilian Amazon (green). Non-Brazilian samples comprised of *P. vivax* isolates from Africa, Central America, and South America (red). This map was generated by the Malaria Atlas Project, University of Oxford (<https://malariaatlas.org/>). All maps are available under the Creative Commons Attribution 3.0 license (CC BY 3.0; <https://creativecommons.org/licenses/by/3.0>).

<https://doi.org/10.1371/journal.pone.0241426.g001>

QIAamp™ DNA Blood Midi Kit (QIAGEN, Hilden, Germany), according to the manufacturer's instructions.

### Primers design and *pvcsp* amplification

The primer set *pvcspf* 5' TCCTGTTGGTGGACTTGTTC 3' (forward), and *pvcspr* 5' GCCAGCACACTTATCCATTGT 3' (reverse), amplifying a fragment of 1,034 base pairs, was designed using the CPS of Salvador 1 strain (Sal-1; PVX\_119355; PlasmoDB: <http://www.plasmoDB.org>) as a reference, employing the software Primer3web (<http://bioinfo.ut.ee/primer3/>), and Oligo Analyzer™ 3.1 (Integrated DNA Technologies, USA). Amplification specificity was assessed comparing the *pvcsp* primer sequences with CSP nucleotide sequences using the Basic Local Alignment Search Tool (BLAST) algorithm at the National Center for Biotechnology Information website.

The amplification of the *pvcsp* gene fragment was performed in a final volume of 20 µL containing: 2 µL (100–200 ng) of DNA, 1 µM primers (forward and reverse), 5x HOT FIREPol™ Blend Master Mix (Solis Biodyne, Hannover, Germany); 1 mM triphosphate deoxynucleotides (dNTP); 12.5 mM of MgCl<sub>2</sub> plus enzyme buffer and dyes to increase sample density during the agarose gel electrophoresis. The PCR condition comprised an initial hold (95°C/12 min), followed by 35 cycles (95°C/30 sec, 55°C/45 sec, and 72°C/2 min), and a final hold (72°C/7 min). Negative and positive controls were included in all reactions. PCR products were visualized on ethidium bromide-stained 2% agarose gel electrophoresis under UV transilluminator (Digi-Doc-It, UVP, California, USA).

DNA templates were purified using Wizard™ SV Gel and PCR Clean-Up System (Promega, Wisconsin, USA), following the manufacturer's procedure. Subsequently, purified DNA was sequenced using Big Dye™ Terminator Cycle Sequencing Ready Reaction version 3.1 (Applied Biosystems, California, USA), with 3.2 µM of forward and reverse PCR primers. DNA sequences were determined using ABI Prism DNA Analyzer™ 3730 (Applied Biosystems, California, USA), at the Fiocruz Genomic Platform PDTIS/Fiocruz RPT01A.

The analysis of the genetic polymorphisms included single nucleotide polymorphism (SNP) and the number of nonapeptide repeats in CR (Insertion and Deletion / InDel).

Nucleotide sequences were aligned using ClustalW multiple sequence aligner in BioEdit software [19], and the electropherogram set to a 10-cutoff score was analyzed using NovoSNP software [39]. The Sal-1 strain was the reference sequence. DNA sequences with singleton mutation or containing overlapped peaks in the same locus were re-sequenced. DNA sequences were deposited in GenBank™ (NIH genetic sequence database; [www.ncbi.nlm.nih.gov/GenBank](http://www.ncbi.nlm.nih.gov/GenBank)) with accession nos. MN417517-MN417735.

Each nonapeptide repeat in CR corresponds to a “motif”. Each nucleotide sequence of the nonapeptide repeat produced the “allotypes”. The order of repeats at the CR identified the VK210 “subtypes”.

### Genetic diversity, natural selection, and statistical analysis

Genetic diversity of *pvcsp* gene was estimated using the DnaSP 6.11 software [40]. Within-population diversity was measured based on the number of segregating sites (*S*), nucleotide diversity ( $\pi$ ), and haplotype diversity (*H<sub>d</sub>*). The distribution of nucleotide diversity was investigated by the sliding window method, in which  $\pi$  is calculated on each window containing a DNA segment of 100 bp with 25 bp overlap.

The variance of allele frequencies between populations was estimated by the Wright's fixation statistics (*F<sub>st</sub>*) considering 1,000 pseudo-replicates bootstrapping [41]. To identify sequences that do not fit the neutral theory model at equilibrium between mutation and

genetic drift, Tajima's D test [42] based on the number of segregating sites ( $\theta$ ) was carried out with DnaSP 6.11 software [39].

The relationship between individual genotypes at the population level was assessed through a haplotype network constructed on R with the "haploNet" function of the package "pegas" under an infinite site model [43]. Arguments "size" of "plot" function were attributed to haplotype frequencies, and colors were defined as different regions of occurrence. Each haplotype's genetic distance was represented by the number of dots on links and was estimated by pairwise differences [44]. The statistical analyses and graphs were performed utilizing GraphPad Prism software version 8.1.2; differences were considered significant with p-value < 0.05.

Linkage disequilibrium (LD) to assess the nonrandom association of alleles located at different loci of the *pvcsp* gene was evaluated for all pairwise combinations by D, D' and  $R^2$  using the DnaSP v.5 software and chi-square test. The coefficient of determination  $R^2$  is the square of the correlation coefficient between two indicator variables. Monomorphic and non-diallelic variants were removed from analyzes.  $R^2$  for each pair of genetic polymorphisms was plotted on heatmap graphics with R using the "LDheatmap" package [45]. Genetic recombination was calculated using the ZZ statistic ( $Z_a - Z_{ns}$ ), where larger positive values are expected with increasing recombination [46]. The average of  $R^2$  overall pairwise comparisons represents  $Z_{ns}$  [46] and those among adjacent polymorphic sites corresponds to the  $Z_a$  [47].

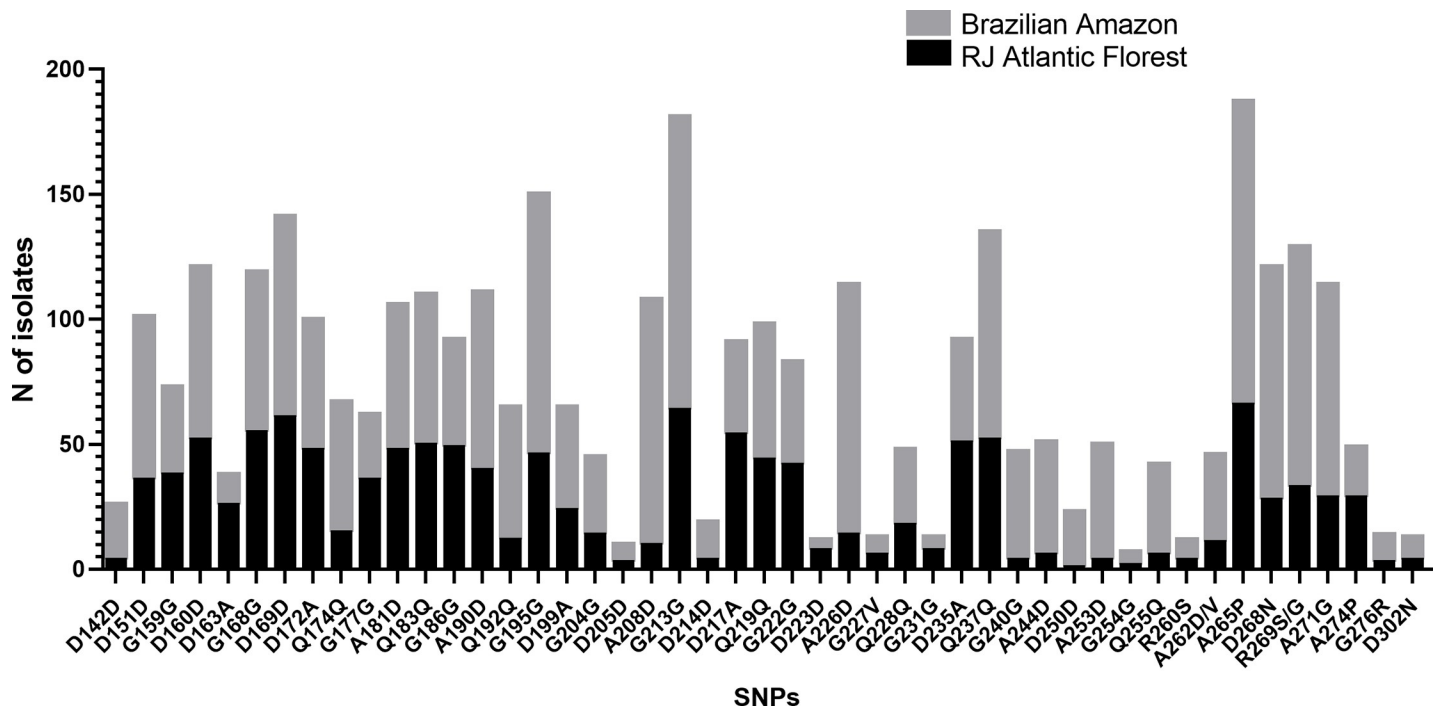
## Results

### Genetic diversity

PCR products revealed a DNA fragment band ranging from 980 to 1088 bp; unspecific amplifications were not verified. The *pvcsp* gene was successfully amplified in 219 isolates (92%): 71 from AF and 136 from BA, comprising the states of Pará (52), Amazonas (46), Rondônia (17), Acre (11), Amapá (8), and Roraima (2) (S1 Table). Additionally, 12 samples from nine countries spanning three continents were also tested (S1 Table): five from Africa [Burkina Faso (1), Ivory Coast (1), and Angola (3)]; one from Central America (Haiti), and six from South America [Colombia (1), French Guyana (1), Peru (1), Suriname (1), and Venezuela (2)].

Multiple sequence alignment of the amplified *pvcsp*-gene fragment of these 219 isolates disclosed 48 polymorphic sites, 26 synonymous and 22 nonsynonymous distributed in 47 codons (Fig 2 and S2 Table). Most of allelic codons differ by one nucleotide. The R269S/G codon displayed two nucleotide substitutions (A805G and A807T) (S1 File) while five codons presented triallelic positions (A504T/C, T531C/A, T558C/A, A612T/C, and C785A/T). CR had a higher polymorphic rate (46/96%) than C-terminal. Among the SNPs, seven were detected only in Brazilian isolates: six synonymous (A504C, T531A, T558A, T615C, A693T, and T750C) in CR, and one nonsynonymous (G904A) in C-terminal region. All other substitutions were observed at least in one NB sample. (S2 Table). InDels increasing or decreasing the number of nonapeptides from 22 to 18 repeats, respectively, were observed in CR. Insertions were more frequent in AF (38%), and deletions were only detected in BA (4%) (Table 1). The presence of InDels in CR does not seem to be related to those of SNPs in the non-repetitive region (C-terminal), since only one sample from AF presented insertion and nucleotide polymorphism simultaneously.

The nucleotide diversity in all Brazilian isolates was  $\pi = 0.0135 (\pm 0.0002)$ . Similar diversity was observed when AF and BA regions were analyzed separately, with the highest peak of diversity occurring in CR (nucleotide positions 451 to 550) ( $\pi = 0.052$ ) (Fig 3). The great haplotype diversity in Brazilian and non-Brazilian regions was equally distributed: -0.99 for BA and AF-, and 1.00 for NB (Table 1).



**Fig 2. *pvcsp* codons in Brazilian isolates from two geographic regions: AF (71) and BA (136).** AF: Rio de Janeiro Atlantic Forest; BA: Brazilian Amazon. The first character represents the amino acid of the Sal-1 reference sequence, followed by the position of this residue (number), and the replacing amino acid observed at the same position (last character). Graph was constructed using GraphPad Prism software version 8.1.2.

<https://doi.org/10.1371/journal.pone.0241426.g002>

The positive value in Tajima’s D test was significant only in BA samples (2.694;  $p < 0.05$ ) (Table 1). These data together with the great nucleotide and haplotype diversities are suggestive of balancing selection.

### Variation in the central repeat region of *pvcsp* gene

All Brazilian and NB *pvcsp* sequences corresponded to the VK210 type. In CR, the polymorphic sites occurred from the sixth to the twentieth nonapeptide repeats: replacement variations began in the eighth and extended to the last (20<sup>th</sup>) nonapeptide (Table 2 and Fig 4).

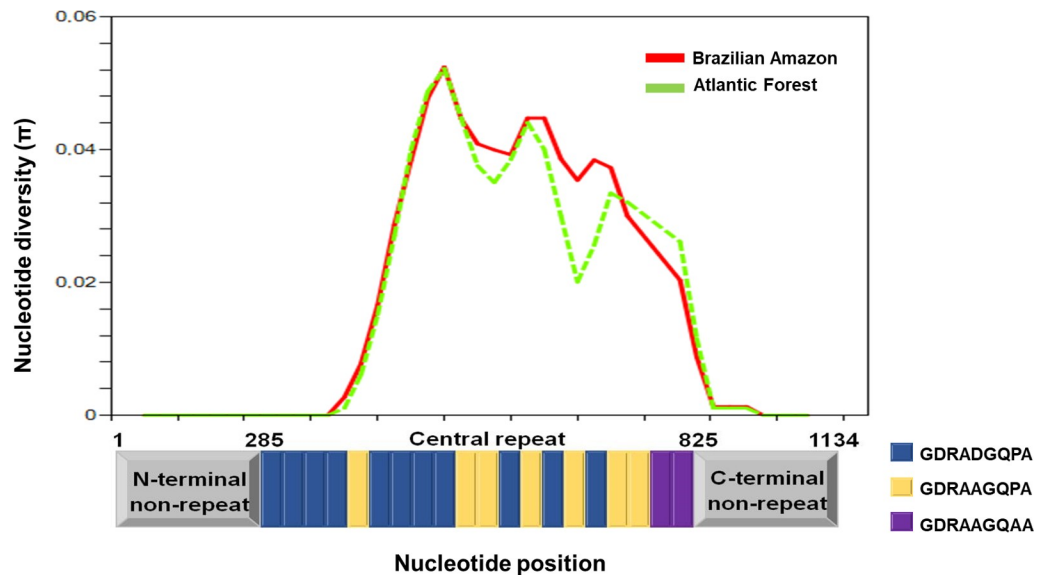
These SNPs spawned 13 nonapeptides allotypes in the CR. Among them, two nonapeptides—allotypes 1 and 2—were detected in all isolates (Fig 5). The allotype 1 was the most abundant—GDRADGQPA—varying from seven to 19 repeats, and the type 2—GDRAAGQPA—from one to 14 repeats. Among the CR motif, three repeats—15, 19, and 20—presented the largest number of nsSNPs generating 4, 7, and 5 nonapeptide types, respectively (Fig 4 and S3 Table).

**Table 1. *pvcsp* genetic diversity and natural selection in Brazilian and non-Brazilian parasite’s samples.**

Region	N	SNP	Diversity				Natural Selection
			In/Del	$\pi$ (SE)	nh	Hd (SE)	Tajima’s D
BA	136	48	22/18	0.0135 ( $\pm 0.0002$ )	100	0.99 ( $\pm 0.002$ )	2.694 ( $p < 0.05$ )
AF	71	48	22/0	0.0135 ( $\pm 0.0005$ )	61	0.99 ( $\pm 0.002$ )	1.784
NB	12	41	0/18	NA	12	1.00 ( $\pm 0.030$ )	0.885

BA: Brazilian Amazon; AF: Atlantic Forest; NB: Non-Brazilian samples; N: number of samples;  $\pi$ : nucleotide diversity; Nh: number of haplotypes; Hd: haplotype diversity; SE: standard error; SNP: Single Nucleotide Polymorphism; In/Del: Insertion/Deletion represented by the number of nonapeptide repeats; NA: not applicable.

<https://doi.org/10.1371/journal.pone.0241426.t001>



**Fig 3. Nucleotide diversity across the *pvmsp* gene in Brazilian isolates.** Sliding window representation (window length of 100 bp with a step size of 25 bp) of 136 isolates from Brazilian Amazon and 71 from Atlantic Forest. Sal-1 strain: the *pvmsp* reference sequence; nonapeptide types: 1 (blue), 2 (yellow) and 3 (purple). The  $\pi$  was calculated using the number of SNPs.

<https://doi.org/10.1371/journal.pone.0241426.g003>

Length polymorphism (InDels) at CR presented nonapeptide tandem repeats ranging from 18 to 22; the Sal-1 reference strain had 20 repeats (Table 3). In Brazil, 43 samples (21%) presented InDels: 37 (86%) with insertions and 6 (14%) with deletions. Insertions were observed in 27 samples from AF (38%) and 10 from BA (7%), most of them of Amazonas state (8/ 80%) (S1 Table). Of these samples, 8 (22%) had 21 repeats, and 29 (78%) had 22 repeats, corresponding to an increase of one (type 3: GDRAAGQAA or type 10: GNGAGGQAA) or two (type 2: GDRAAGQPA and; type 10: GNGAGGQAA) nonapeptide repeats. Insertions were not detected in NB.

Deletions were observed in Brazil only in 6 samples (3%) from BA: one sample had 19 repeats (nucleotides 799–825), and 5 samples had 18 repeats (nucleotides 772–825). In NB, deletions were detected in two samples (17%), one from Suriname and other Venezuela, both with 18 repeats (Table 3, S1 and S3 Tables).

Concerning VK210, 24 subtypes were identified (Fig 6). Subtype VK210-1 was the most common, accounting for 53% (116) of the isolates, and the frequencies of the remaining subtypes varied from 5% (11) to 0.5% (one) (Fig 6). Interestingly, 18 VK210 subtypes were found exclusively in Brazilian regions: five (VK210-4-5-6-7-8) in AF and BA, seven only in BA (VK210-9-11-13-16-17-21-24), and six only in AF (VK210-10-15-18-19-20-23). However, the lack of detection of VK210 subtypes in all studied locations may be related to different sampling number among them.

### Distribution of *pvmsp* haplotypes and phylogenetic analysis

Considering the arrangement of nonapeptide allotypes of VK210 subtypes and CR nonapeptide repeat length polymorphisms altogether, 168 haplotypes were identified (HCS01 – HCS168). Among them, 100 were detected in BA, 61 in AF, and 12 in NB. Only one isolate (BA/Amapá state) was identical to the reference Sal-1 strain (Fig 7 and S3 Table).

Four Brazilian haplotypes were also found in NB samples: HCS03 (Amazonas, Pará, Rondônia and Angola); HCS11 (AF and Colombia), HCS15 (Amapá and Peru), and HCS19 (Pará

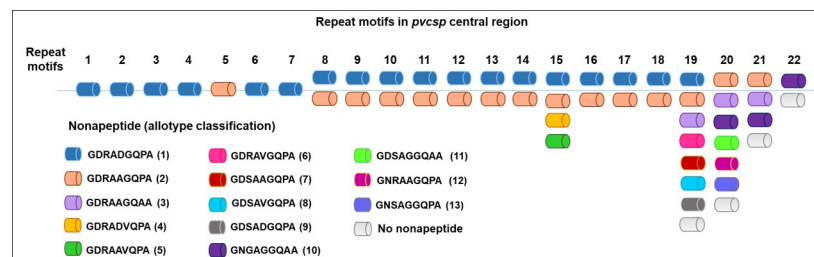
**Table 2. Synonymous and nonsynonymous substitutions in the nonapeptide tandem repeat of central region generating different allotypes.**

Repeat	Nonapeptide sequences	N of SNPs	Nucleotide sequences
6	<u>G</u> DRADGQPA	1	GGA GAT/C AGA GCA GAT GGA CAG CCA GCA
7	<u>G</u> DRAAGQPA	1	GGA GAC/T AGA GCA GAT GGA CAG CCA GCA
8	<u>G</u> DRA ( <b>D/A</b> ) GQPA	3	GGA/T GAC/T AGA GCA G (A/C) T GGA CAG CCA GCA
9	<u>G</u> DRA ( <b>D/A</b> ) <u>G</u> QPA	4	GGA/T/C GAC/T AGA GCA G (A/C) T GGA CAG/A CCA GCA
10	<u>G</u> DRA ( <b>A/D</b> ) <u>G</u> QPA	3	GGT/C/A GAT AGA GCA G (C/A) T GGA CAA/G CCA GCA
11	<u>G</u> DRA ( <b>A/D</b> ) <u>G</u> QPA	3	GGT/C/A GAT AGA GCA G (C/A) T GGA CAG/A CCA GCA
12	<u>G</u> DRA ( <b>D/A</b> ) GQPA	2	GGC/A GAT AGA GCA G (A/C) T GGA CAG CCA GCA
13	<u>G</u> DRA ( <b>A/D</b> ) GQPA	3	GGA/T/C GAT/C AGA GCA G (C/A) T GGA CAG CCA GCA
14	<u>G</u> DRA ( <b>D/A</b> ) <u>G</u> QPA	4	GGC/A GAT AGA GCA G (A/C) T GGA CAG/A CCA GCA
15	<u>G</u> DRA ( <b>A/D</b> ) ( <b>G/V</b> ) <u>Q</u> PA	5	GGA/T GAT/C AGA GCA G (C/A) T G (G/T) A CAA/G CAA GCA
16	<u>G</u> DRA ( <b>D/A</b> ) <u>G</u> QPA	3	GGA/T GAT AGA GCA G (A/C) T GGA CAA/G CCA GCA
17	<u>G</u> DRA ( <b>A/D</b> ) GQPA	2	GGA/T GAT AGA GCA G (C/A) T GGA CAG CCA GCA
18	<u>G</u> DRA ( <b>A/D</b> ) <u>G</u> QPA	4	GGA GAT/C AGA GCA G (C/A) T GGA/T CAG/A CCA GCA
19	GD ( <b>R/A</b> ) A ( <b>A/D/V</b> ) GQ ( <b>A/P</b> ) A	3	GGA GAT AG (A/T) GCA G (C/A/T) T GGA CAG (G/C) CA GCA
20	G ( <b>D/N</b> ) ( <b>R/G/S</b> ) A ( <b>A/G</b> ) GQ ( <b>A/P</b> ) A	5	GGA (G/A) AT (A/G) G (A/T) GCA G (C/G) T GGA CAG (G/C) CA GCA

Amino acids and nucleotides underlined represent synonymous substitution; amino acids and nucleotides in bold are nonsynonymous substitution.

<https://doi.org/10.1371/journal.pone.0241426.t002>

and Suriname); only one Brazilian haplotype—HCS01 –was detected in AF and BA. All other 163 haplotypes were segregated according to BA, AF, and NB localities (Fig 7 and S3 Table). Ninety-six out of 100 BA haplotypes were exclusively detected in this area, and among the 61 AF haplotypes, 59 were exclusive of this region. In NB samples eight out of 12 haplotypes were not detected in Brazilian samples: Angola (HCS51 and HCS80), Ivory Coast (HCS22), Burkina



**Fig 4. Schematic representation of the nonapeptides in central region of the *pvmsp* gene.** Each color represents one of the 13 allotypes, according to the nonapeptide motifs. Motifs 1 to 5 are identical to Sal-1 reference strain; motifs 6 and 7 synonymous SNPs; motifs 21 and 22 insertions and motifs 19 and 20 deletions.

<https://doi.org/10.1371/journal.pone.0241426.g004>



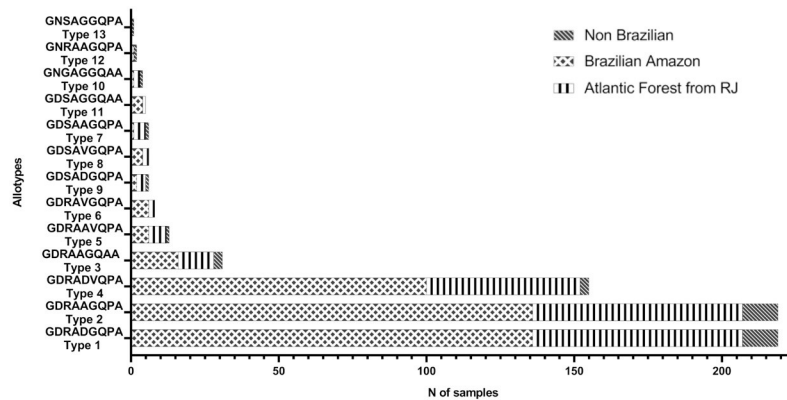


Fig 5. Allotypes distribution of *P. vivax* in Brazilian's and non-Brazilian's samples.

<https://doi.org/10.1371/journal.pone.0241426.g005>

Faso (HCS86), Haiti (HCS62), French Guyana (HCS101), and Venezuela (HCS67 and HCS110).

The haplotype network tree showed that BA samples were closely related to Sal-1 reference sequence (Fig 7). Low genetic differentiation, consistent with a weak population subdivision, was observed between AF and BA populations ( $F_{st} = 0.208/ p < 0.001$ ).

### Linkage disequilibrium

Linkage disequilibrium was assessed for all possible pairwise combinations. The results considering all samples, showed an overall low occurrence of LD. Higher signals ( $R^2 \geq 0.75$ ) were observed between G802A, A805G, and A807T, suggesting a small LD block containing these three polymorphisms. High LD levels were also found between A807T/C812G and between insertions 826::27/854::27 (Fig 8A). When split by regions, results from BA showed high levels of LD only between G802A, A805G, and A807T, while AF samples displayed a large LD block including G802A, A805G, A807T, and C812G. Perfect LD ( $R^2 = 1$ ) was found for G802A and A805G, and almost perfect for 826::27/854::27 insertions in AF samples (Fig 8B and 8C). The ZnS, Za, and ZZ had low values across the *pvcsp* sequence, suggesting that meiotic recombination has not played a crucial role in shuffling nucleotide variation among DNA sequences (Table 4).

### Polymorphism in B- and T- cell epitopes

When analyzed four B-cell epitopes [19, 20], no SNPs were found in N- and C- terminal regions coding peptides P8 (GDAKKKKDGKKAEPKNPREN), P24 (CSVTGCVGVRRVRRVNAANK) and P25 (VRRRVNAANKKPEDLTNLNDL), but P11 (GDRADGQPA; allotype 1), in

Table 3. Number of InDel polymorphisms in *pvcsp* gene central region, compared to the Sal-1 reference strain's *pvcsp* gene.

Polymorphism	Number of Repeats	BA (136)	AF (71)	NB (12)	Total
Insertion	22	3	26	0	29
	21	7	1	0	8
Deletion	19	1	0	0	1
	18	5	0	2	7
Sal-1 Type	20	120	44	10	174

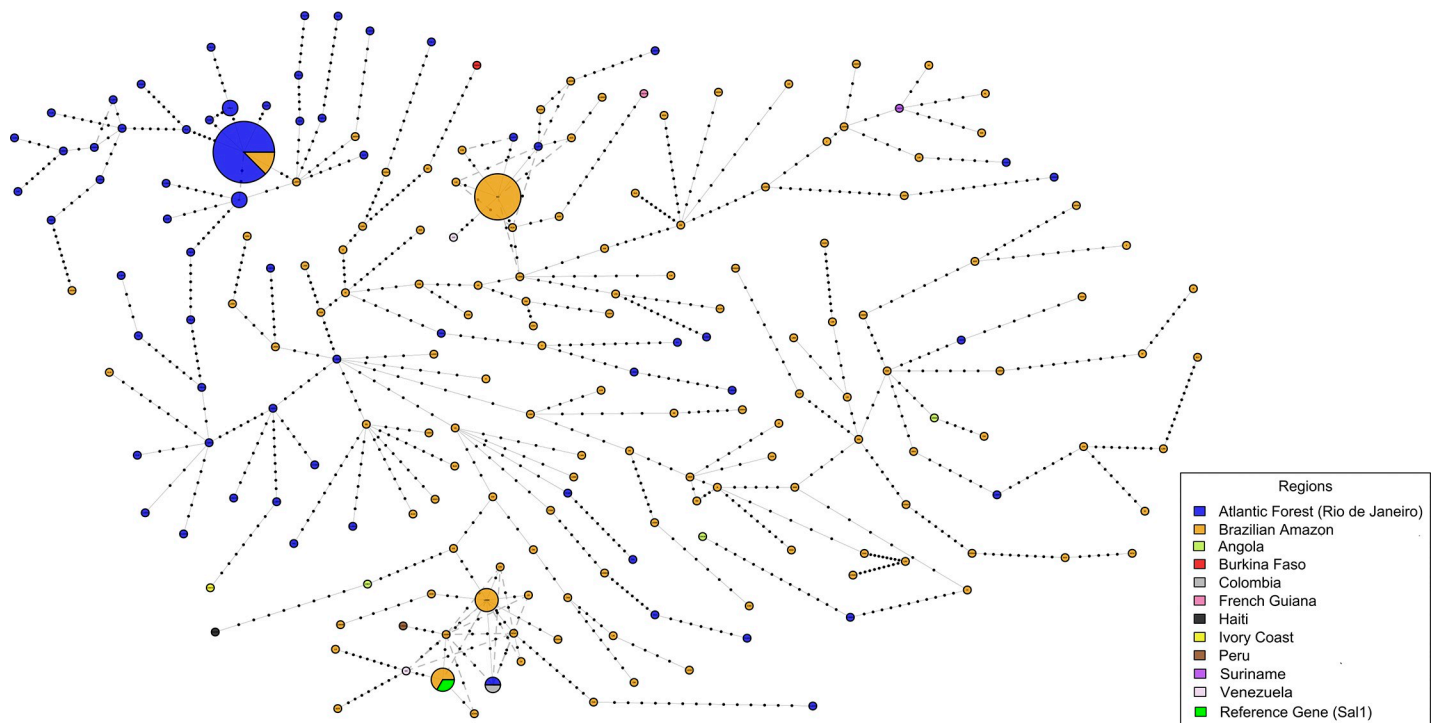
BA: Brazilian Amazon; AF: Atlantic Forest; NB: Non-Brazilian; Sal-1: reference strain

<https://doi.org/10.1371/journal.pone.0241426.t003>

VK210 Subtypes	Nonapeptide types													N of Isolates			Total	%
	1	2	3	4	5	6	7	8	9	10	11	12	13	BA	AF	NB		
1	■	■								■				84	30	2	116	53,0
2		■												11	10	4	25	11,4
3*			■											19	5	2	26	11,9
4				■						■				2	9	0	11	5,0
5					■									4	6	0	10	4,6
6						■								2	1	0	3	1,4
7							■							1	2	0	3	1,4
8								■						2	1	0	3	1,4
9									■					2	0	0	2	0,9
10										■				0	2	0	2	0,9
11											■			2	0	0	2	0,9
12												■		1	0	1	2	0,9
13													■	2	0	0	2	0,9
14														0	0	1	1	0,5
15														0	1	0	1	0,5
16														1	0	0	1	0,5
17														1	0	0	1	0,5
18														0	1	0	1	0,5
19														0	1	0	1	0,5
20														0	1	0	1	0,5
21														1	0	0	1	0,5
22														0	0	1	1	0,5
23														0	1	0	1	0,5
24														1	0	0	1	0,5
25														0	0	1	1	0,5
<b>Total n of subtypes per region</b>														<b>16</b>	<b>14</b>	<b>7</b>	<b>219</b>	

**Fig 6. Distribution of VK210 subtypes according to regions.** BA: Brazilian Amazon; AF: Atlantic Forest, NB: Non Brazilian; (\*) Salvador 1 strain VK210 type; nonapeptide allotypes: 1 (GDRAAGQPA); 2 (GDRAAGQPA); 3 (GDRAAGQAA); 4 (GDRADVQPA); 5 (GDRAAVQPA); 6 (GDRAVGQPA); 7 (GDSAAGQPA); 8 (GDSAVGQPA); 9 (GDSADGQPA); 10 (GNGAGQAA); 11 (GDSAGQAA); 12 (GNRAAGQPA); 13 (GNSAGQPA).

<https://doi.org/10.1371/journal.pone.0241426.g006>



**Fig 7. Haplotype network of *pvcsp* gene.** Node sizes are proportional to haplotype frequency, and branch lengths are indicative of the number of single nucleotide differences between sequences. Node colors indicate sampling geographic origin.

<https://doi.org/10.1371/journal.pone.0241426.g007>

CR, presented a nsSNP in five of the nonapeptide repeats (motifs 8, 9, 12, 14 and 16) generating, therefore, allotype 2 (Table 5).

In relation to the three TCD8<sup>+</sup> cytotoxic cell epitopes (P1, P3 and P5) [48] located in the C-terminal region, only P1 presented a nsSNP [YL(D/N)KVRATV] in 6% (13) of the Brazilian samples: 7% (9) from BA and 6% (4) from AF (Table 5).

Considering now the three TCD4<sup>+</sup> helper cell epitopes (P6, P15 and P25) [19, 49], only P15 (GDRAAGQAA/allotype 3), located in CR, was polymorphic. Of the 219 samples, only 39 (18%) presented the P15 epitope sequence conserved but the number of repeats varied from 1 to 3: six samples, three from BA and three from AF, had two repeats, as Sal-1 reference strain; two samples from BA had three repeats and; 31 samples, 16 from BA, 12 from AF and 3 from NB had one repeat. In relation to the polymorphic sequences at CR, the P15 epitope presented in motif 19 three nsSNPs [GD (R/A) A (A/D/V) GQ (A/P) A] in 201 (92%) samples. The motif 20 disclosure four nsSNPs [G (D/N) (R/G/S) A (A/G) GQ (A/P) A] in 177 (81%) samples (Table 5).

In the Th2R epitope [PNEKSVKEYL (D/N) KVRATVG] [33], located in the C-terminal region, a nsSNP was identified in 13 (6%) of Brazilian samples: 4 (6%) from AF and 9 (7%) from BA. No polymorphisms were detected in Th3R epitope.

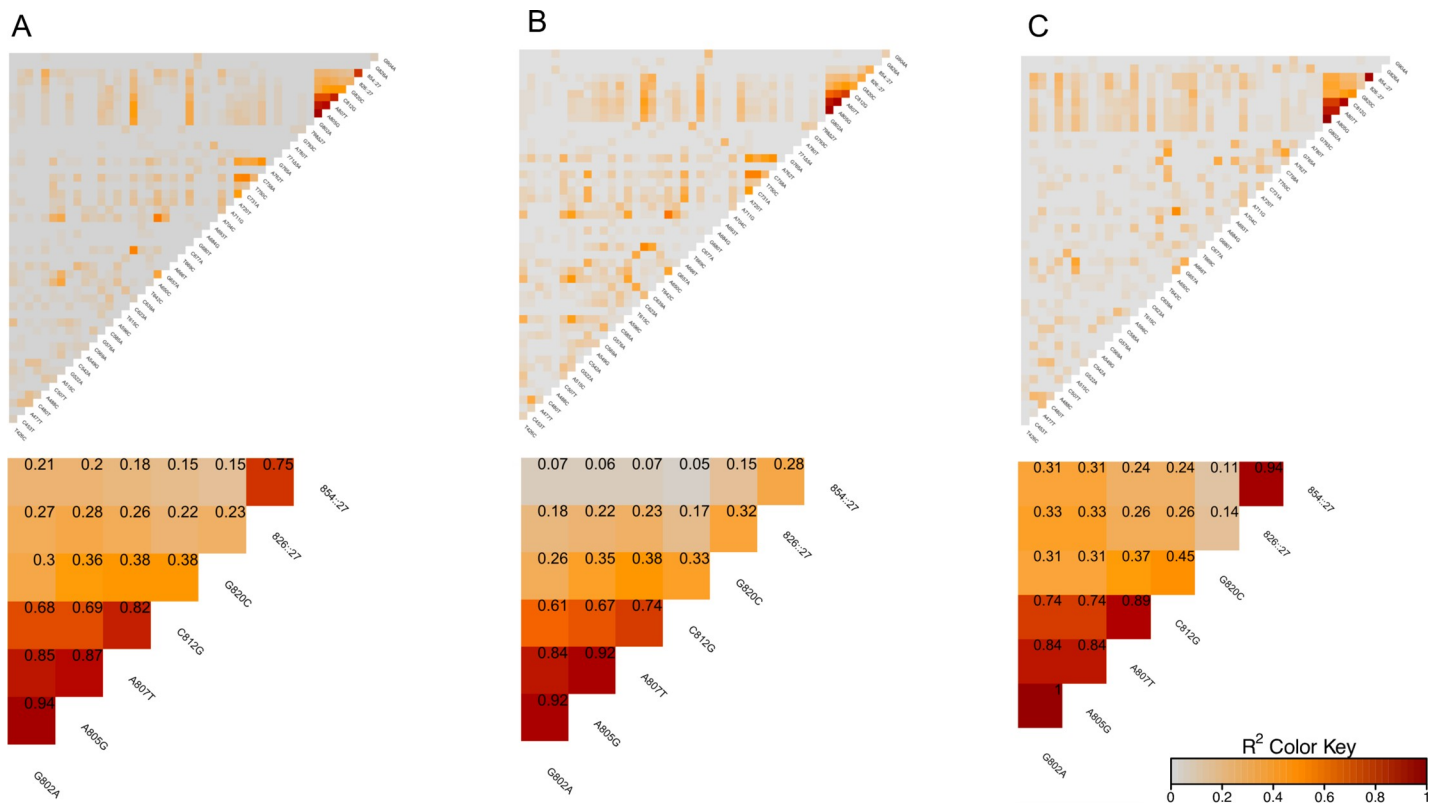
## Discussion

*pvcsp* is a single copy gene located in chromosome 8 of *P. vivax*. It is highly expressed on sporozoites' surface, playing an essential role in adhesion and invasion of hepatocytes [50]. Due to its surface localization, along with the antibody protective response, this protein has been considered as a malaria pre-erythrocytic phase vaccine candidate [51]. The knowledge regarding the nature and extent of *pvcsp* genetic diversity in Brazil is limited since the majority of *pvcsp* studies had been focused on the immunological response triggered by CSP epitopes [22, 52] or the characterization of *pvcsp* variants associated with the response to treatment, or even in the detection of such variants in *Anopheles* mosquitoes [26, 34]. Now, we report for the first time the genetic diversity, linkage disequilibrium, and natural selection of *pvcsp* gene in two different epidemiological sites, the Brazilian Amazon rainforest (BA) and the Rio de Janeiro Atlantic Forest (AF).

Only *pvcsp*'s gene variant VK210 was detected in both BA and AF; consequently, no association with the presence of VK210 and the geographic origin of the parasites could be made. Although other two variants—VK247 and vivax-like—have already been detected in *P. vivax* parasites from the Brazilian Amazon [22, 26, 28, 53] and pre-Amazon [34], the cosmopolitan variant VK210 also predominate in French Guyana [54], Nicaragua [55], Pakistan [30], Sri Lanka [56], Myanmar [57], Cambodia [29], Iran [58], Vanuatu Island [59], and China [25].

It has been claimed that the presence and extension of *pvcsp* variants differ according to transmission intensity [60]. However, this study and those performed in countries with different transmission intensities, where only VK210 type was detected, as Sudan, Azerbaijan, Sri Lanka [61], Honduras [62], and China [63], do not support this proposition. Additionally, the fixation of this variant in Brazil, and perhaps around the world, seems not to be associated with mosquito vectors infection, since malaria transmission in BA occurs mainly by *An. Darlingi* and *An. albitarsis*, while in AF malaria transmission occurs almost exclusively by *An. (Kerteszia) cruzii* vector [7]. Altogether, these findings indicate that VK210 may be the best-adapted variant globally, possibly, due to the more successful evasion of such *P. vivax* parasite variant from the immune system.

Sequence analyses revealed high CR diversity, contrasting to none or low diversity in the N- and C- terminals (5' and 3' ends), respectively. Indeed, the pre-central (Region I) and the post-



**Fig 8. Linkage disequilibrium across the *pvmsp* gene of parasite populations from Brazilian Amazon and Rio de Janeiro Atlantic Forest.** A: Parasites from Brazilian Amazon and Rio de Janeiro Atlantic Forest; B: Brazilian Amazon; C: Rio de Janeiro Atlantic Forest; 771Δ54 deletion of 2 repeats; 798Δ27 deletion of 1 repeat; 826::27 insertion of 1 repeat; 854::27 insertion of 2 repeats.

<https://doi.org/10.1371/journal.pone.0241426.g008>

central (Region II) repeat sequences that contain both B- and T-cell epitopes [64] remained conserved in this study. The motif known as Region II was also found conserved in previous studies on the Brazilian Amazon [65], Peru [66], Iran, Sri Lanka, India [67], Cambodia, and Colombia [29]. This constrained variation in terminal regions reinforces such protein structures importance for its interaction with *Anopheles* mosquito salivary gland and human hepatocytes during the invasion processes [52]. Conversely, CR diversity could reflect the host's immune system's evasion mechanism, as already observed for other surface antigens as *pvd* gene [12, 68].

A large number of VK210's substitutions and length polymorphisms have been found worldwide [26, 69]. However, its nonapeptide repeat structure remained conserved, probably by intragenic shuffling, as previously suggested for *P. falciparum* parasites [70]. In this study, all *P. vivax* sequences contain one or more copies of two nonapeptide motifs displaying three synonymous substitutions (allotypes 1 and 2). The persistence of these two common allotypes suggests its maintenance by selective pressure. These silent polymorphisms could protect the

**Table 4. Linkage disequilibrium test across the *pvmsp* sequence by ZnS, Za, and ZZ statistics.**

	AF	BA	Total
Za	0.125	0.072	0.060
Zns	0.047	0.054	0.038
ZZ	0.078	0.018	0.022

<https://doi.org/10.1371/journal.pone.0241426.t004>

Table 5. Polymorphisms in B- and T-cell epitope regions of *pvcsp*.

Peptide	Residues	Position	Epitope	n of SNPs	n of isolates
Th2R	PNEKSVKEYL (D/N) KVRATVG	292–309	T-helper	1 NS	13
P1	YL (D/N) KVRATV	300–308	T-cytotoxic	1 NS	13
P11	GDRA (D/A) GQPA	96–104	B	4 (3S / 1NS)	219
P15 (motif 19)	GD (R/A) A (A/D/V) GQ (A/P) A	258–266	T-helper	3 NS	201
P15 (motif 20)	G (D/N) (R/G/S) A (A/G) GQ (A/P)	267–275	T-helper	4 NS	177

Epitopes previously identified [19, 20, 33, 48, 49]; NS: nonsynonymous; S: synonymous.

<https://doi.org/10.1371/journal.pone.0241426.t005>

parasites against stochastic reductions in variability, such as those resulting from bottlenecks. On the other hand, the allotype 3, considered a T-helper cell epitope, revealed an “excess” of nonsynonymous substitutions in the great majority of the samples. The low Sal-1 allotype 3 frequency has already been reported in China [63], Mexico [69], Sudan [61], and also in Brazil [27, 28, 71]. Thus, it is reasonable to speculate that such nonsynonymous polymorphisms are not essential for *P. vivax* parasites’ survival, but they might help to evade the host immune response.

The allotype 10 (GNGAGGQAA), also present in Belem 1 strain (GenBank: EU401923), a Brazilian reference strain, was detected in high frequency (71%) mostly in BA isolates. Other investigations in Korea [72], Cambodia [29], Vanuatu [59], and Iran [58] also revealed the presence of this allotype, demonstrating its global distribution. In contrast, among the other nine low-frequency allotypes, only one—allotype 12 (GNRAAGQPA)—was already reported in Sri Lanka [56] and Cambodia [29]. The low prevalence of these novel VK210 subtypes might suggest a relatively recent introduction and/or expansion of these genotypes into Brazilian territory, as previously observed in Iran’s malaria transmission areas [58]. Inversely, allotype 13 (GNSAGGQPA), only detected in French Guyana, suggested that this allotype is not circulating in Brazil. Unfortunately, the number of samples from Amapá state, which borders French Guyana, was too small to confirm the lack of allotype 13 in Brazil.

Concerning InDel polymorphisms, the CR also varied in the number of tandem nonapeptide repeats (18 to 22), which is a significant factor leading to global isolates’ genetic *pvcsp* diversity. Insertions, resulting in 21 nonapeptide repeats, were mostly detected in BA, while 22 repeats predominated in AF isolates. This number of insertions was similar to those observed in China [63], Cambodia [29], and even in Brazilian Amazon [28]. Interestingly, the deletion of one or two motifs, reducing the number of nonapeptide in tandem to 18 and 19, respectively, was unique detected in BA and bordering countries (Suriname and Venezuela).

Remarkably, samples with deletion showed the most conserved DNA sequence, and, inversely, those with insertion were more polymorphic. Therefore, most allotypes had been arising in the final CR portion (motifs 19 and 20), suggesting more significant selective pressure on these motifs, most likely mediated by the host immune system.

As the sexual recombination during meiosis in the mosquito, CR diversity could be raised by intrahelical strand-slippage events during mitotic DNA replication [28, 70], including inserting or deleting of a repeated motif without changing non-repetitive flanking sequences. However, it is not trivial to discriminate against the influence of these two possibilities unless in low endemic areas, where meiotic recombination rates are far more limited. Thus, the finding of *P. vivax* populations with great genomic diversity in several low endemic regions across the world [29, 36, 73], including the AF data here reported for the first time in Brazil, may reinforce the role of mitotic recombination in accelerating CR evolution. However, although human malaria has a very low endemicity in the AF region, parasites are continuously

circulating in monkeys/mosquitoes and facing the immune response of this animal host over and over and, consequently, meiotic recombination could also play a role in genetic diversity.

The genetic diversity in South America has been attributed to different parasite lineages originating from geographically diverse regions [12, 36, 73], showing that the demographic history of *P. vivax* could also affect the *pvcsp* population structure. Despite the considerable distance between the Brazilian Amazon and Atlantic Forest, the low mobility of infected people between these areas, and the smaller parasite population's size in AF that generally could reduce gene flow between BA and AF, identical polymorphisms were found in parasite populations of these regions.

The low genetic differentiation between AF and BA populations was also verified with a haplotype network, in which sequences of both regions are interlinked and have similar distances between their haplotypes. AF isolates were formed more clusters than the BA ones. Thus, the isolates of the same geographic region may share similar evolutionary histories. The extensive haplotype distance in BA parasite populations could be a result of i) BA geographic area is much larger than the AF; ii) BA area concentrates approximately 99% of Brazilian *P. vivax* cases, increasing the chances of recombination; and iii) BA parasites showed more variation in SNP numbers than AF.

Although AF and BA have distinct epidemiological profiles, including different vector species, transmission rates, and hosts with different immune pressure, migration rates, and regional-temporal fluctuations, we did not detect any fixed nucleotide differences in isolates from these geographic areas. Remarkably, AF autochthonous cases, indicative of zoonotic transmission to humans [6, 10], where parasite populations suffered a bottleneck effect, followed by selective expansion to adapt to non-human host, and different species of anopheline, no systematic genomic changes leading to any consistent signature was recorded.

Contrary to the expectations, all sampled *P. vivax* populations harbored high genetic diversity and low level of genetic differentiation between AF and BA parasite populations. These findings were already observed in parasite populations from Colombian geographically distinct regions [74].

The low values of linkage disequilibrium across the *pvcsp* sequence suggest that recombination has not played a crucial role in shuffling the nucleotide variations. Alternatively, the central region of CSP may contain a mutational "hot spot" that could mask the presence of recombination in the *pvcsp* [75].

On the other hand, insertion positions of a small block at the end of the central region with high LD in AF samples could suggest a recent bottleneck and/or clonal expansions, as expected in places with a reduction in transmission rates.

To better understand if selection acts upon the *pvcsp* gene, Tajima's D neutrality test was performed, and the results suggest that diversity was mediated by balancing selection, probably by the presence of multiple B- and T- cell immunodominant epitopes [19, 20, 48, 49].

It is well known that a single amino acid change or clustered replacements in B- or T- cell epitopes may potentially reduce the antibody response and the peptide-binding pocket of HLA, respectively. Here, we investigated eleven epitopes previously described to trigger the host immune response [20, 33, 48, 49]. Nine of them were in the N- and C- terminal regions, and two in the CR one. The Th2R and Th3R T-cell *P. vivax* peptides are located in the C-terminal region and they are orthologous to those of *P. falciparum* CSP [76]. The Th2R peptide sequence comprised the P1 T-cytotoxic epitope; polymorphisms in these peptides were present in a small number of Brazilian isolates (13/6%) from AF (4/6%) and BA (9/7%). The Th3R remained conserved in the parasite populations here studied, similar to those of two sub-species of *P. ovale* [77], but different to *P. malariae* [78], *P. knowlesi* [79] and *P. falciparum* [76] parasites, in which both Th2R and Th3R epitopes were polymorphic.

To summarize, this study provides information on the genetic polymorphisms of *pvcsp*-gene in isolates from different Brazilian endemic areas, showing high CR diversity and low geographic *pvcsp* population's structure, probably modulated by natural selection and host immunity. Although the uneven distribution of allelic genes may jeopardize vaccines' deployment directed to a specific variable *locus*, a unique vaccine formulation could protect human populations in all Brazilian regions.

## Supporting information

**S1 Table. Number of collected and amplified samples per region/locality.**

(PDF)

**S2 Table. Polymorphisms in *pvcsp*-gene fragment observed in Brazilian and foreign country isolates.**

(PDF)

**S3 Table. Distribution of allotypes in the CR, and haplotypes.**

(XLSX)

**S1 File. Multiple sequence alignment of proteins encoded by *pvcsp*-gene.**

(PDF)

## Acknowledgments

We want to thank all the patients for their participation in this study. We also acknowledge the health staff's technical assistance of for helping with data collection and microscopic diagnosis, and the staff from the Genomic Platform for DNA sequencing facility RPT01A/PDTIS/Fiocruz.

## Author Contributions

**Conceptualization:** Natália Ketrin Almeida-de-Oliveira, Maria de Fátima Ferreira-da-Cruz.

**Data curation:** Natália Ketrin Almeida-de-Oliveira, Anielle de Pina-Costa, Patrícia Brasil, Cláudio Tadeu Daniel-Ribeiro.

**Formal analysis:** Natália Ketrin Almeida-de-Oliveira, Daiana de Souza Perce-da-Silva, Marcos Catanho, Atila Duque Rossi.

**Funding acquisition:** Cláudio Tadeu Daniel-Ribeiro, Maria de Fátima Ferreira-da-Cruz.

**Investigation:** Rebecca de Abreu-Fernandes, Lidiane Lima-Cury, Anielle de Pina-Costa, Daiana de Souza Perce-da-Silva, Marcos Catanho, Patrícia Brasil.

**Methodology:** Natália Ketrin Almeida-de-Oliveira, Rebecca de Abreu-Fernandes, Lidiane Lima-Cury, Aline Rosa de Lavigne, Daiana de Souza Perce-da-Silva, Maria de Fátima Ferreira-da-Cruz.

**Project administration:** Maria de Fátima Ferreira-da-Cruz.

**Resources:** Patrícia Brasil, Cláudio Tadeu Daniel-Ribeiro.

**Software:** Daiana de Souza Perce-da-Silva, Marcos Catanho.

**Supervision:** Cláudio Tadeu Daniel-Ribeiro, Maria de Fátima Ferreira-da-Cruz.

**Writing – original draft:** Natália Ketrin Almeida-de-Oliveira, Maria de Fátima Ferreira-da-Cruz.

**Writing – review & editing:** Anielle de Pina-Costa, Daiana de Souza Perce-da-Silva, Marcos Catanho, Patrícia Brasil, Cláudio Tadeu Daniel-Ribeiro, Maria de Fátima Ferreira-da-Cruz.

## References

1. World Health Organization. World Malaria Report 2019. 2019. Available: <https://www.who.int/publications-detail/world-malaria-report-2019>
2. George MT, Schloegel JL, Ntumngia FB, Barnes SJ, King CL, Casey JL, et al. Identification of an Immunogenic Broadly Inhibitory Surface Epitope of the *Plasmodium vivax* Duffy Binding Protein Ligand Domain. *mSphere*. 2019; 4: e00194–19. <https://doi.org/10.1128/mSphere.00194-19> PMID: 31092602
3. Ministério da Saúde. Epidemiological report of malaria cases in Brazil 2019. Ministério da Saúde. 2019. Available: [https://public.tableau.com/profile/mal.ria.brasil#!/casos\\_notificados\\_2019\\_regio\\_Amaznica](https://public.tableau.com/profile/mal.ria.brasil#!/casos_notificados_2019_regio_Amaznica)
4. Pina-costa A De, Brasil P, Maria S, Santi D, Araujo MP De, Suárez-mutis MC, et al. Malaria in Brazil: what happens outside the Amazonian endemic region. 2014; 109: 618–633. <https://doi.org/10.1590/0074-0276140228> PMID: 25185003
5. Almeida-de-Oliveira N, Abreu-Fernandes R, Lavigne A, Pina-Costa A, Perce-da-Silva D, Catanho M, et al. Exploration of *Plasmodium vivax* Merozoite Surface Proteins 1 and 7 genetic diversity in Brazilian Amazon and Rio de Janeiro Atlantic Forest. *Infect Genet Evol*. 2020;in press. <https://doi.org/10.1016/j.meegid.2020.104592> PMID: 33059085
6. Brasil P, Zalis MG, de Pina-Costa A, Siqueira AM, Bianco Júnior C, Silva S, et al. Outbreak of human malaria caused by *Plasmodium simium* in the Atlantic Forest in Rio de Janeiro: a molecular epidemiological investigation. *Lancet Glob Heal*. 2017; 5: e1038–46. [https://doi.org/10.1016/S2214-109X\(17\)30333-9](https://doi.org/10.1016/S2214-109X(17)30333-9)
7. Abreu FVS de, Santos E Dos, Mello ARL, Gomes LR, Alvarenga DAM de, Gomes MQ, et al. Howler monkeys are the reservoir of malarial parasites causing zoonotic infections in the Atlantic forest of Rio de Janeiro. *PLoS Negl Trop Dis*. 2019; 13: e0007906. <https://doi.org/10.1371/journal.pntd.0007906> PMID: 31815937
8. Deane LM. Simian malaria in Brazil. *Mem Inst Oswaldo Cruz*. 1992; 87 Suppl 3: 1–20. <https://doi.org/10.1590/s0074-02761992000700001> PMID: 1343676
9. Deane LM. Malaria vectors in Brazil. *Mem Inst Oswaldo Cruz*. 1986; 81: 5–14. <https://doi.org/10.1590/s0074-02761986000600002>
10. Cerqueira Buery J, Thihara Rodrigues P, Natal L, Camoese Salla L, Carolina Loss A, Rachel Vicente C, et al. Mitochondrial genome of *Plasmodium vivax/simium* detected in an endemic region for malaria in the Atlantic Forest of Espírito Santo state, Brazil: do mosquitoes, simians and humans harbour the same parasite? *Malar J*. 2017; 16: 437. <https://doi.org/10.1186/s12936-017-2080-9> PMID: 29084553
11. Miguel RB, Peiter PC, Albuquerque H De, Coura JR, Moza PG, Costa ADP, et al. Malaria in the state of Rio de Janeiro, Brazil, an Atlantic Forest area: an assessment using the health surveillance service. 2014; 109: 1–7. <https://doi.org/10.1590/0074-0276130558> PMID: 25185004
12. Rodrigues PT, Valdivia HO, De Oliveira TC, Alves JMP, Duarte AMRC, Cerutti-Junior C, et al. Human migration and the spread of malaria parasites to the New World. *Sci Rep*. 2018; 8: 1993. <https://doi.org/10.1038/s41598-018-19554-0> PMID: 29386521
13. Howes RE, Battle KE, Kamini MN, Smith DL, Cibulskis RE, Baird JK, et al. Global Epidemiology of *Plasmodium vivax*. *Am J Trop Med Hyg*. 2016; 95: 15–34. <https://doi.org/10.4269/ajtmh.16-0141> PMID: 27402513
14. Taylor WRJ, Thriemer K, von Seidlein L, Yuentrakul P, Assawariyathipat T, Assefa A, et al. Short-course primaquine for the radical cure of *Plasmodium vivax* malaria: a multicentre, randomised, placebo-controlled non-inferiority trial. *Lancet*. 2019; 394: 929–938. [https://doi.org/10.1016/S0140-6736\(19\)31285-1](https://doi.org/10.1016/S0140-6736(19)31285-1) PMID: 31327563
15. Lacerda MVG, Mourão MPG, Alexandre MA a, Siqueira AM, Magalhães BML, Martinez-Espinosa FE, et al. Understanding the clinical spectrum of complicated *Plasmodium vivax* malaria: a systematic review on the contributions of the Brazilian literature. *Malar J*. 2012; 11: 12. <https://doi.org/10.1186/1475-2875-11-12> PMID: 22230294
16. Cowell AN, Winzeler EA. The genomic architecture of antimalarial drug resistance. *Brief Funct Genomics*. 2019; 00: 1–15. <https://doi.org/10.1093/bfpg/elz008> PMID: 31119263
17. Tham WH, Beeson JG, Rayner JC. *Plasmodium vivax* vaccine research—we've only just begun. *Int J Parasitol*. 2017; 47: 111–118. <https://doi.org/10.1016/j.ijpara.2016.09.006> PMID: 27899329
18. Garrido-Cardenas JA, Cebrián-Carmona J, González-Cerón L, Manzano-Agugliaro F, Mesa-Valle C. Analysis of global research on malaria and *Plasmodium vivax*. *Int J Environ Res Public Health*. 2019;16. <https://doi.org/10.3390/ijerph16111928> PMID: 31159165



19. Hernández-Martínez MÁ, Escalante AA, Arevalo-Herrera M, Herrera S. Antigenic diversity of the *Plasmodium vivax* circumsporozoite protein in parasite isolates of Western Colombia. *Am J Trop Med Hyg*. 2011; 84: 51–57. <https://doi.org/10.4269/ajtmh.2011.09-0785> PMID: 21292878
20. Arévalo-Herrera M, Roggero MA, Gonzalez JM, Vergara J, Corradin G, Lopez JA, et al. Mapping and comparison of the B-cell epitopes recognized on the *Plasmodium vivax* circumsporozoite protein by immune Colombians and immunized Aotus monkeys. *Ann Trop Med Parasitol*. 1998; 92: 539–551. <https://doi.org/10.1080/00034989859230> PMID: 9797827
21. Herrera S, Bonelo A, Perlaza BL, Valencia AZ, Cifuentes C, Hurtado S, et al. Use of long synthetic peptides to study the antigenicity and immunogenicity of the *Plasmodium vivax* circumsporozoite protein. *Int J Parasitol*. 2004; 34: 1535–1546. <https://doi.org/10.1016/j.ijpara.2004.10.009> PMID: 15582530
22. Pereira VA, Sánchez-Arcila JC, Vasconcelos MPA, Ferreira AR, de Souza Videira L, Teva A, et al. Evaluating seroprevalence to circumsporozoite protein to estimate exposure to three species of *Plasmodium* in the Brazilian Amazon. *Infect Dis Poverty*. 2018;7. <https://doi.org/10.1186/s40249-018-0388-5> PMID: 29391070
23. Coppi A, Natarajan R, Pradel G, Bennett BL, James ER, Roggero MA, et al. The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host. *J Exp Med*. 2011; 208: 341–356. <https://doi.org/10.1084/jem.20101488> PMID: 21262960
24. Arnot DE, Barnwell JW, Tam JP, Nussenzweig V, Nussenzweig RS E V. Circumsporozoite protein of *Plasmodium vivax*: gene cloning and characterization of the immunodominant epitope. *Science*. 1985; 230: 815–8. <https://doi.org/10.1126/science.2414847> PMID: 2414847
25. Ling Zhang L, Nong Yao L, Liang Chen H, Yi Lu Q, Ruan W. Genetic diversity analysis of PvCSP and its application in tracking of *Plasmodium vivax*. *Exp Parasitol*. 2018; 188: 26–35. <https://doi.org/10.1016/j.exppara.2018.03.001> PMID: 29526573
26. Machado RL, Póvoa MM. Distribution of *Plasmodium vivax* variants (VK210, VK247 and P. vivax-like) in three endemic areas of the Amazon region of Brazil and their correlation with chloroquine treatment. *Trans R Soc Trop Med Hyg*. 2000; 94: 377–381. [https://doi.org/10.1016/s0035-9203\(00\)90110-x](https://doi.org/10.1016/s0035-9203(00)90110-x) PMID: 11127238
27. Machado RLD, Filho AF de, Calvosa VSP, Figueredo MC, Nascimento JM, Póvoa MM. Correlation between *Plasmodium vivax* variants in Belém, Pará State, Brazil and symptoms and clearance of parasitaemia. *Brazilian J Infect Dis*. 2003; 7: 175–177. <https://doi.org/10.1590/s1413-86702003000300002> PMID: 14499040
28. Patil A, Orjuela-Sánchez P, da Silva-Nunes M, Ferreira MU. Evolutionary dynamics of the immunodominant repeats of the *Plasmodium vivax* malaria-vaccine candidate circumsporozoite protein (CSP). *Infect Genet Evol*. 2010; 10: 298–303. <https://doi.org/10.1016/j.meegid.2010.01.006> PMID: 20097310
29. Parobek CM, Bailey J a., Hathaway NJ, Socheat D, Rogers WO, Juliano JJ. Differing Patterns of Selection and Geospatial Genetic Diversity within Two Leading *Plasmodium vivax* Candidate Vaccine Antigens. *PLoS Negl Trop Dis*. 2014;8. <https://doi.org/10.1371/journal.pntd.0002796> PMID: 24743266
30. Zakeri S, Raeisi A, Afsharipad M, Kakar Q, Ghasemi F, Atta H, et al. Molecular characterization of *Plasmodium vivax* clinical isolates in Pakistan and Iran using pvmsp-1, pvmsp-3α and pvcsp genes as molecular markers. *Parasitol Int*. 2010; 59: 15–21. <https://doi.org/10.1016/j.parint.2009.06.006> PMID: 19545647
31. Raza A, Ghanchi NK, Thaver AM, Jafri S, Beg M a. Genetic diversity of *Plasmodium vivax* clinical isolates from southern Pakistan using pvcsp and pvmsp1 genetic markers. *Malar J*. 2013; 12: 16. <https://doi.org/10.1186/1475-2875-12-16> PMID: 23311628
32. Ruan W, Ling Zhang L, Feng Y, Zhang X, Liang Chen H, Yi Lu Q, et al. Genetic diversity of *Plasmodium vivax* revealed by the merozoite surface protein-1 icb5-6 fragment. *Infect Dis Poverty*. 2017; 6. <https://doi.org/10.1186/s40249-017-0302-6> PMID: 28578709
33. Qari SH, Goldman IF, Póvoa MM, di Santi S, Alpers MP, Lal AA. Human recognition of T cell epitopes on the *Plasmodium vivax* circumsporozoite protein. *Mol Biochem Parasitol*. 1992; 55: 105–113. [https://doi.org/10.1016/0166-6851\(92\)90131-3](https://doi.org/10.1016/0166-6851(92)90131-3) PMID: 1279418
34. Ribeiro B de P, Cassiano GC, de Souza RM, Cysne DN, Grisotto MAG, de Azevedo Dos Santos, et al. Polymorphisms in *Plasmodium vivax* Circumsporozoite Protein (CSP) Influence Parasite Burden and Cytokine Balance in a Pre-Amazon Endemic Area from Brazil. *PLoS Negl Trop Dis*. 2016; 10: e0004479. <https://doi.org/10.1371/journal.pntd.0004479> PMID: 26943639
35. Arnott A, Barry AE, Reeder JC. Understanding the population genetics of *Plasmodium vivax* is essential for malaria control and elimination. *Malar J*. 2012; 11: 14. <https://doi.org/10.1186/1475-2875-11-14> PMID: 22233585

36. Neafsey DE, Galinsky K, Jiang RHY, Young L, Sykes SM, Saif S, et al. The malaria parasite *Plasmodium vivax* exhibits greater genetic diversity than *Plasmodium falciparum*. *Nat Genet.* 2012; 44: 1046–1050. <https://doi.org/10.1038/ng.2373> PMID: 22863733
37. Arévalo-Herrera M, Vásquez-Jiménez JM, Lopez-Perez M, Vallejo AF, Amado-Garavito AB, Céspedes N, et al. Protective Efficacy of *Plasmodium vivax* Radiation-Attenuated Sporozoites in Colombian Volunteers: A Randomized Controlled Trial. *PLoS Negl Trop Dis.* 2016;10. <https://doi.org/10.1371/journal.pntd.0005070> PMID: 27760143
38. Torres KL, Figueiredo DV, Zalis MG, Daniel-Ribeiro CT, Alecrim W, De Fátima Ferreira-da-Cruz M. 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.* 2006; 98: 519–524. <https://doi.org/10.1007/s00436-005-0085-8> PMID: 16416292
39. Weckx Stefan, Jurgen Del-Favero Rosa Rademakers LC, Cruts Marc, Peter De Jonghe Christine Van Broeckhoven PDR. NovoSNP, a novel computational tool for sequence variation discovery. *Genome Res.* 2005; 15: 436–442. <https://doi.org/10.1101/gr.2754005> PMID: 15741513
40. Rozas J, Ferrer-Mata A, Carlos Sanchez-DelBarrio JS, Guirao-Rico S, Librado P, and Ramos-Onsins SE, et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol Biol Evol.* 2017; 34: 3299–3302. <https://doi.org/10.1093/molbev/msx248> PMID: 29029172
41. Hudson RR, Slatkint M, Maddison WP. Estimation of Levels of Gene Flow From DNA Sequence Data. *Genetics.* 1992; 132: 583–589. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1205159/pdf/ge1322583.pdf> PMID: 1427045
42. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics.* 1989; 123: 585–595. PMC1203831 PMID: 2513255
43. Paradis E, Barrett J. pegas: an R package for population genetics with an integrated-modular approach. *Bioinform Appl NOTE.* 2010; 26: 419–420. <https://doi.org/10.1093/bioinformatics/btp696> PMID: 20080509
44. Saitou N, Nei M. The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol Biol Evol.* 1987; 4: 406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454> PMID: 3447015
45. Shin J-H, Blay S, Graham J, McNeney B. LDheatmap: An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. *J Stat Softw.* 2006; 16. <https://doi.org/10.18637/jss.v016.i11> PMID: 21451741
46. Rozas J, Gullaud M, Blandin G, Aguadé M. DNA variation at the rp49 gene region of *Drosophila simulans*: Evolutionary inferences from an unusual haplotype structure. *Genetics.* 2001; 158: 1147–1155. Available: <http://flybase.bio.indiana.edu> PMID: 11454763
47. Kelly JK. A test of neutrality based on interlocus associations. *Genetics.* 1997; 146: 1197–1206. PMID: 9215920
48. Arévalo-Herrera M, Valencia AZ, Vergara J, Bonelo A, Fleischhauer K, González JM, et al. Identification of HLA-A2 restricted CD8+ T-lymphocyte responses to *Plasmodium vivax* circumsporozoite protein in individuals naturally exposed to malaria. *Parasite Immunol.* 2002; 24: 161–169. <https://doi.org/10.1046/j.1365-3024.2002.00449.x> PMID: 12078650
49. Herrera S, Escobar P, De Plata C, Avila GI, Corradin G, Herrera MA. Human recognition of T cell epitopes on the *Plasmodium vivax* circumsporozoite protein. *J Immunol.* 1992; 148: 3986–3990. Available: <http://www.jimmunol.org/content/148/12/3986> PMID: 1376347
50. Kappe SHI, Buscaglia CA, Nussenzweig V. *Plasmodium* Sporozoite Molecular Cell Biology. *Annu Rev Cell Dev Biol.* 2004; 20: 29–59. <https://doi.org/10.1146/annurev.cellbio.20.011603.150935> PMID: 15473834
51. López C, Yepes-Pérez Y, Hincapié-Escobar N, Díaz-Arévalo D, Patarroyo MA. What is known about the immune response induced by *Plasmodium vivax* malaria vaccine candidates? *Front Immunol.* 2017;8. <https://doi.org/10.3389/fimmu.2017.00008> PMID: 28144241
52. Santos E de A, Primo DG, Sucupira IMC, Cassiano GC, Barbosa DRL, Viana GMR, et al. Frequency of *Plasmodium vivax* circumsporozoite protein genotypes in humans and anopheline mosquitoes in an endemic area of southeastern Pará State, Brazil. *Rev Pan-Amazônica Saúde.* 2016; 7: 57–64. <https://doi.org/10.5123/S2176-62232016000300007>
53. Santos E de A, Sucupira IMC, Martins BMDO, Guimarães RJDPS, Catete CP, Lessa De Souza RT, et al. VK210 and VK247 genotypes of *Plasmodium vivax* in anopheline mosquitoes from Brazilian Amazon. *Nat Sci Reports.* 2019; 9. <https://doi.org/10.1038/s41598-019-45809-5> PMID: 31253819
54. Bonilla JA, Validum L, Cummings R, Palmer CJ. Genetic diversity of *Plasmodium vivax* PVCSP and PVMSP1 in Guyana, South America. *Am J Trop Med Hyg.* 2006; 75: 830–835. PMID: 17123973

55. González-Cerón L, Martínez-Barnetche J, Montero-Solís C, Santillán F, Soto AM, Rodríguez MH, et al. Molecular epidemiology of *Plasmodium vivax* in Latin America: Polymorphism and evolutionary relationships of the circumsporozoite gene. *Malar J*. 2013; 12: 1. <https://doi.org/10.1186/1475-2875-12-1> PMID: 23282136
56. Dias S, Wickramarachchi T, Sahabandu I, Escalante AA, Udagama P V. Population genetic structure of the *Plasmodium vivax* circumsporozoite protein (Pvcsp) in Sri Lanka. *Gene*. 2013; 518: 381–387. <https://doi.org/10.1016/j.gene.2013.01.003> PMID: 23333606
57. Moon SU, Lee HW, Kim JY, Na BK, Cho SH, Lin K, et al. High frequency of genetic diversity of *Plasmodium vivax* field isolates in Myanmar. *Acta Trop*. 2009; 109: 30–36. <https://doi.org/10.1016/j.actatropica.2008.09.006> PMID: 18851938
58. Shabani SH, Zakeri S, Mehrizi AA, Mortazavi Y, Djadid ND. Population genetics structure of *Plasmodium vivax* circumsporozoite protein during the elimination process in low and unstable malaria transmission areas, southeast of Iran. *Acta Trop*. 2016; 160: 23–34. <https://doi.org/10.1016/j.actatropica.2016.04.006> PMID: 27102931
59. Chan CW, Sakihama N, Tachibana S-II, Idris ZM, Lum JK, Tanabe K, et al. *Plasmodium vivax* and *Plasmodium falciparum* at the crossroads of exchange among Islands in Vanuatu: Implications for malaria elimination strategies. *PLoS One*. 2015; 10: e0119475. <https://doi.org/10.1371/journal.pone.0119475> PMID: 25793260
60. Cui L, Mascorro CN, Fan Q, Rzomp KA, Khuntirat B, Zhou G, et al. Genetic diversity and multiple infections of *Plasmodium vivax* malaria in western Thailand. *Am J Trop Med Hyg*. 2003; 68: 613–619. <https://doi.org/10.4269/ajtmh.2003.68.613> PMID: 12812356
61. Talha AA, Pirahmadi S, Mehrizi AA, Djadid ND, Nour BYMM, Zakeri S. Molecular genetic analysis of *Plasmodium vivax* isolates from Eastern and Central Sudan using pvcsp and pvmsp-3 $\alpha$  genes as molecular markers. *Infect Genet Evol*. 2015; 32: 12–22. <https://doi.org/10.1016/j.meegid.2015.02.004> PMID: 25721363
62. Lopez AC, Ortiz A, Coello J, Sosa-Ochoa W, Torres REM, Banegas EI, et al. Genetic diversity of *Plasmodium vivax* and *Plasmodium falciparum* in Honduras. *Malar J*. 2012; 11: 1. <https://doi.org/10.1186/1475-2875-11-1> PMID: 22212246
63. Liu Y, Zhou RM, Zhang YL, Wang DQ, Li SH, Yang CY, et al. Analysis of polymorphisms in the circumsporozoite protein gene of *Plasmodium vivax* isolates from Henan Province, China. *Malar J*. 2018; 17. <https://doi.org/10.1186/s12936-017-2160-x> PMID: 29316929
64. Takala SL, Plowe C V. Genetic diversity and malaria vaccine design, testing and efficacy: Preventing and overcoming “vaccine resistant malaria.” *Parasite Immunol*. 2009; 31: 560–573. <https://doi.org/10.1111/j.1365-3024.2009.01138.x> PMID: 19691559
65. Souza-Neiras WC, Storti-Melo LM, Cassiano GC, SCA Couto V, ARA Couto Á, Soares IS, et al. *Plasmodium vivax* circumsporozoite genotypes: a limited variation or new subspecies with major biological consequences? *Malar J*. 2010. Available: <http://www.malariajournal.com/content/9/1/178> <https://doi.org/10.1186/1475-2875-9-178> PMID: 20573199
66. Chenet SM, Tapia LL, Escalante AA, Durand S, Lucas C, Bacon DJ. Genetic diversity and population structure of genes encoding vaccine candidate antigens of *Plasmodium vivax*. *Malar J*. 2012; 11. <https://doi.org/10.1186/1475-2875-11-11> PMID: 22230255
67. Kaur H, Sehgal R, Kumar A, Sehgal A, Bharti PK, Bansal D, et al. Exploration of genetic diversity of *Plasmodium vivax* circumsporozoite protein (Pvcsp) and *Plasmodium vivax* sexual stage antigen (Pvs25) among North Indian isolates. *Malar J*. 2019; 18: 308. <https://doi.org/10.1186/s12936-019-2939-z> PMID: 31492135
68. Almeida-De-Oliveira NK, Lima-Cury L, De Abreu-Fernandes R, De Rosa Lavigne A, De Pina-Costa A, De Souza Perce-Da-Silva D, et al. Extensive genetic diversity of *Plasmodium vivax* dbp-II in Rio de Janeiro Atlantic Forest and Brazilian Amazon Basin: Evidence of positive selection. *Malar J*. 2020; 19: 81. <https://doi.org/10.1186/s12936-020-03159-y> PMID: 32075659
69. González-Cerón L, Rodríguez MH, Nettel-Cruz JA, Hernández-Ávila JE, Malo-García IR, Santillán-Valenzuela F, et al. *Plasmodium vivax* CSP-Pvs25 variants from southern Mexico produce distinct patterns of infectivity for *Anopheles albimanus* versus *An. pseudopunctipennis*, in each case independent of geographical origin. *Parasites and Vectors*. 2019; 12. <https://doi.org/10.1186/s13071-018-3194-9> PMID: 30616670
70. Rich SM, Hudson RR, Ayala FJ. *Plasmodium falciparum* antigenic diversity: Evidence of clonal population structure. *Proc Natl Acad Sci U S A*. 1997; 94: 13040–13045. <https://doi.org/10.1073/pnas.94.24.13040> PMID: 9371796
71. Pratt-Riccio LR, de OliveiraBaptista B, Rodrigues Torres V, Bianco-Junior C, de SouzaPerce-Da-Silva D, Pratt Riccio EK, et al. Chloroquine and mefloquine chemoresistance profiles are not related to the Circumsporozoite Protein (CSP) VK210 subtypes in field isolates of *Plasmodium vivax* from Manaus,

- Brazilian Amazon. Mem Inst Oswaldo Cruz. 2019;114. <https://doi.org/10.1590/0074-02760190054> PMID: 31411308
72. Lim CS, Kim YK, Lee KN, Kim SH, Hoffman KJ, Song KJ, et al. The analysis of circumsporozoite-protein gene sequences from South Korean isolates of *Plasmodium vivax*. Ann Trop Med Parasitol. 2001; 95: 229–235. <https://doi.org/10.1080/00034980120053997> PMID: 11339882
  73. de Oliveira TC, Rodrigues PT, Menezes MJ, Gonçalves-Lopes RM, Bastos MS, Lima NF, et al. Genome-wide diversity and differentiation in New World populations of the human malaria parasite *Plasmodium vivax*. Sinnis P, editor. PLoS Negl Trop Dis. 2017; 11: 1–24. <https://doi.org/10.1371/journal.pntd.0005824> PMID: 28759591
  74. Pacheco MA, Schneider KA, Céspedes N, Herrera S, Arévalo-Herrera M, Escalante AA. Limited differentiation among *Plasmodium vivax* populations from the northwest and to the south Pacific Coast of Colombia: A malaria corridor? PLoS Negl Trop Dis. 2018;13. <https://doi.org/10.1371/journal.pntd.0007310> PMID: 30921317
  75. Cole-Tobian J, King CL. Diversity and natural selection in *Plasmodium vivax* Duffy binding protein gene. Mol Biochem Parasitol. 2003; 127: 121–132. [https://doi.org/10.1016/s0166-6851\(02\)00327-4](https://doi.org/10.1016/s0166-6851(02)00327-4) PMID: 12672521
  76. Waitumbi JN, Anyona SB, Hunja CW, Kifude CM, Polhemus ME, Walsh DS, et al. Impact of RTS,S/AS02A and RTS,S/AS01B on genotypes of *P. falciparum* in adults participating in a malaria vaccine clinical trial. PLoS One. 2009; 4: 7849. <https://doi.org/10.1371/journal.pone.0007849> PMID: 19924281
  77. Lu F, Ahmed MA, Xu S, Xu S, Han JH, Liu Q, et al. *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* in Chinese travelers: Prevalence of novel genotypes of circumsporozoite protein in the African continent. Infect Genet Evol. 2019; 70: 9–14. <https://doi.org/10.1016/j.meegid.2019.02.008> PMID: 30763775
  78. Saralamba N, Mayxay M, Newton PN, Smithuis F, Nosten F, Archasuksan L, et al. Genetic polymorphisms in the circumsporozoite protein of *Plasmodium malariae* show a geographical bias. Malar J. 2018; 17: 269. <https://doi.org/10.1186/s12936-018-2413-3> PMID: 30012172
  79. Fong MY, Ahmed MA, Wong SS, Lau YL, Sitam F. Genetic diversity and natural selection of the *Plasmodium knowlesi* circumsporozoite protein nonrepeat regions. PLoS One. 2015;10. <https://doi.org/10.1371/journal.pone.0137734> PMID: 26379157