

RESEARCH ARTICLE

Population Genetics of *Plasmodium vivax* in Four Rural Communities in Central Vietnam

Nguyen Van Hong^{1,2,3}, Christopher Delgado-Ratto³, Pham Vinh Thanh¹, Peter Van den Eede², Pieter Guetens², Nguyen Thi Huong Binh¹, Bui Quang Phuc¹, Tran Thanh Duong¹, Jean Pierre Van Geertruyden³, Umberto D'Alessandro^{2,4}, Annette Erhart², Anna Rosanas-Urgell^{2*}

1 National Institute of Malariaology, Parasitology and Entomology (NIMPE), Hanoi, Vietnam, **2** Institute of Tropical Medicine (ITM), Antwerp, Belgium, **3** International Health Unit, Faculty of Medicine, University of Antwerp, Belgium, **4** Medical Research Council Unit (MRC Unit), Fajara, The Gambia

* arosanas@itg.be



 OPEN ACCESS

Citation: Hong NV, Delgado-Ratto C, Thanh PV, Van den Eede P, Guetens P, Binh NTH, et al. (2016) Population Genetics of *Plasmodium vivax* in Four Rural Communities in Central Vietnam. PLoS Negl Trop Dis 10(2): e0004434. doi:10.1371/journal.pntd.0004434

Editor: Joseph M. Vinetz, University of California San Diego School of Medicine, UNITED STATES

Received: October 23, 2015

Accepted: January 14, 2016

Published: February 12, 2016

Copyright: © 2016 Hong 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: Public availability of data would compromise patient privacy, therefore relevant data is available upon request to the contact person Jef Verellen: jverellen@itg.be.

Funding: This work was supported by FWO "Krediet aan Navorsers" (www.fwo.be, grant 1.5.197.13N), UBS Optimus Foundation (www.ubs.com/optimus, grant number Opt1648) and the Belgian Directorate-General for Development Cooperation (DGDC) within DGD-ITM Framework Agreement 3.2 and 3.3. The funders had no role in study design, data collection

Abstract

Background

The burden of malaria in Vietnam has drastically reduced, prompting the National Malaria Control Program to officially engage in elimination efforts. *Plasmodium vivax* is becoming increasingly prevalent, remaining a major problem in the country's central and southern provinces. A better understanding of *P. vivax* genetic diversity and structure of local parasite populations will provide baseline data for the evaluation and improvement of current efforts for control and elimination. The aim of this study was to examine the population genetics and structure of *P. vivax* isolates from four communities in Tra Leng commune, Nam Tra My district in Quang Nam, Central Vietnam.

Methodology/Principal Findings

P. vivax mono infections collected from 234 individuals between April 2009 and December 2010 were successfully analyzed using a panel of 14 microsatellite markers. Isolates displayed moderate genetic diversity ($He = 0.68$), with no significant differences between study communities. Polyclonal infections were frequent (71.4%) with a mean multiplicity of infection of 1.91 isolates/person. Low but significant genetic differentiation (F_{ST} value from -0.05 to 0.18) was observed between the community across the river and the other communities. Strong linkage disequilibrium ($I_A^S = 0.113$, $p < 0.001$) was detected across all communities, suggesting gene flow within and among them. Using multiple approaches, 101 haplotypes were grouped into two genetic clusters, while 60.4% of haplotypes were admixed.

Conclusions/Significance

In this area of Central Vietnam, where malaria transmission has decreased significantly over the past decade, there was moderate genetic diversity and high occurrence of polyclonal infections. Local human populations have frequent social and economic interactions that facilitate gene flow and inbreeding among parasite populations, while decreasing

and analysis, decision to publish, or preparation of the manuscript.

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

population structure. Findings provide important information on parasites populations circulating in the study area and are relevant to current malaria elimination efforts.

Author Summary

In Vietnam, *Plasmodium vivax* (*P. vivax*) is the second most frequent human malaria parasite and a major obstacle to countrywide malaria elimination. Knowing the local parasite structure is useful for elimination efforts. Therefore, we analyzed, with a panel of 14 microsatellite markers, 234 *P. vivax* mono infections in blood samples collected from 4 communities in central Vietnam. Genetic diversity in the population was moderate; a high occurrence of polyclonal infections and significant linkage disequilibrium were detected, suggesting inbreeding or recombination between highly related haplotypes. In addition, both genetic differentiation and population structure was low and only detected between communities at each side of the river. Those results suggest gene flow between study communities with the river defining a moderate geographical barrier. Future studies should determine how this genetic variation is maintained in an area of extremely low transmission.

Introduction

Vietnam has been extremely successful in decreasing the country's malaria burden, thanks to the large scale implementation of control interventions such as insecticide-treated bed nets, indoor residual spraying, and prompt, free-of-charge diagnosis and treatment; the number of cases fell from 130,000 in 2004 to 27,868 in 2014 [1]. Malaria has been virtually eliminated from Northern and Southern Vietnam [1, 2]. In 2014, 80% of malaria cases occurred in nine "hot provinces" where annual incidence peaked at 3.1 cases per 1000, indicating a highly heterogeneous transmission, with hot spots of transmission (mostly in mountainous and forested areas) surrounded by areas of low transmission [1–5]. Vietnam aims at eliminating malaria by 2030 [6]. Such ambitious goal is threatened by *P. vivax*, whose characteristics (dormant liver forms that relapse weeks or months after clearance of the primary infection and gametocytes production before the occurrence of symptoms) together with the relative high occurrence of sub-patent and asymptomatic infections that remain undetected and thus untreated [3, 7–10], make its transmission much more difficult to interrupt than that of *P. falciparum*. In addition, as already reported, elimination efforts are threatened by the emergence of drug resistance, for *P. falciparum* to artemisinin derivatives and partner drugs and for *P. vivax* to chloroquine (CQ) [2, 11, 12].

Current efforts to eliminate malaria are targeted to districts and communes reporting an increased number of malaria cases over time [2, 6]. In this context, understanding parasite genetic diversity and its population structure is relevant for (i) monitoring temporal changes in transmission following control efforts, (ii) elucidating the spatial distribution of parasite populations and predicting outbreaks, population resilience, and the spread of drug-resistant parasites, and (iii) identifying ecological and behavioral risk factors that can inform malaria control and elimination efforts [13–15].

A previous study conducted in Binh Thuan province in central Vietnam reported high levels of genetic diversity (average expected heterozygosity (He) = 0.86) and all infections being multi-clonal despite low transmission [13] (similar to what has been previously reported in South-East Asia) [16].

The aim of this study was to provide baseline data on the *P. vivax* parasite populations in four rural communities in the Vietnamese Quang Nam province.

Materials and Methods

Study site

Samples were collected from April 2009 to December 2010 in four communities (Fig 1) in the South Tra My district of Quang Nam, Central Vietnam during a prospective cohort study aiming to assess the short- and long-term efficacy of CQ and high-dose piperazine (PQ) for the treatment of *P. vivax* mono-infections [12]. Detailed sociodemographic characteristics of the local population have been already reported elsewhere [3]. In 2009, the prevalence of malaria by light microscopy was 7.8%, while by polymerase chain reaction (PCR) prevalence was estimated at 22.6% (ranging from 16.4 to 42.5%), with a high proportion of *P. vivax* mono infections (43%). Sub-patent infections accounted for 58.7% of all infections, evidencing the existence of a substantial hidden human reservoir of malaria [3].

Malaria transmission is seasonal, and peaks during the rainy season (May to November). Based on data from the Provincial Malaria Station, between 2009 and 2013 the mean prevalence of *P. falciparum*, *P. vivax*, and mixed malaria cases for all age groups in the study area was 64.1%, 31.5%, and 4.4%, respectively [17–18]. The main malaria vectors in the area are *Anopheles dirus sensu stricto* and *An. minimus*, though *An. vagus*, *An. aconitus*, and *An. philippinensis* are also present [17–18].

Sample collection

A finger prick blood sample was collected at day 0 (before treatment) for diagnosis by light microscopy and two blood spots were collected on grade 3 filter paper (Whatman Ltd.,

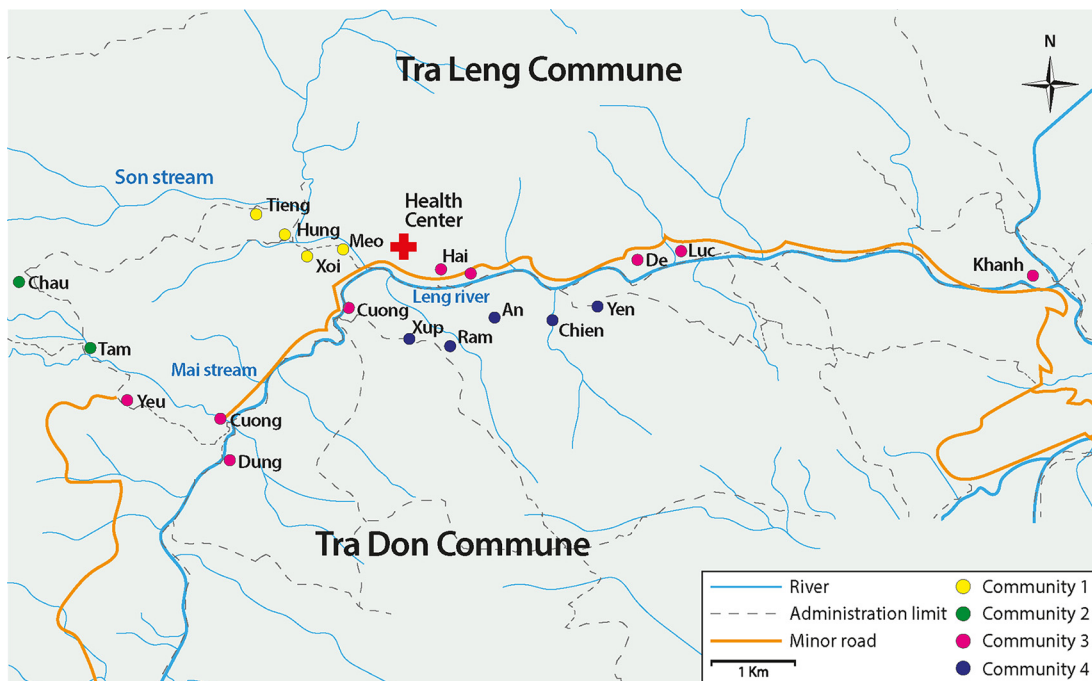


Fig 1. Study area in Nam Tra My district, Quang Nam, Vietnam.

doi:10.1371/journal.pntd.0004434.g001

Springfield Mill, Maidstone, United Kingdom) for molecular diagnosis and microsatellite (MS) genotyping.

Ethics statement

The study was approved by the National Institute of Malariology, Parasitology and Entomology in Hanoi, the Ministry of Health of Vietnam, and the review boards of the Institute of Tropical Medicine and Antwerp University Hospital (UZA) in Antwerp, Belgium. Adult participants (in case of minors one of the parents/guardians) provided written informed consent.

Light microscopy diagnosis

Thick and thin film blood slides were stained with a 3% Giemsa solution for 45 minutes, and the number of asexual parasites was calculated following World Health Organization (WHO) guidelines [19]. Parasite density was estimated by dividing the number of asexual parasites for 200 white blood cells (WBCs) counted and expressed as the number of asexual parasites per microliter of blood, assuming 8000 WBC/ μ L. All blood slides were double-read by two technicians and in case of disagreement, slides were read by a third senior technician. The final results were expressed as the mean of the two closest results. A slide was declared negative if no parasites were found after counting 1,000 WBCs.

Genotyping

DNA was extracted from filter paper blood spots cut into 5-mm-diameter disks with the QIAamp DNA Micro Kit following the manufacturer's recommendations (Qiagen, Hilden Germany). *P. vivax* mono infections were confirmed by species-specific multiplexed semi-nested PCR, as described by Rubio et al. [20]. All samples were genotyped with 14 MS (MS1-MS10, MS12, MS15, MS20, and PvSal1814) following previously described PCR protocols [13–14]. The PCR products of four MS were pooled and analyzed by capillary electrophoresis in a 3730 XL ABI sequencer (Applied Biosystems) and 1200 Liz was used as the internal size standard. Negative samples were repeated once.

Allele calling was performed using GeneMarker version 2.4.0. After pooling capillary electrophoresis fsa files from all samples, a standard cut-off value of 500 relative fluorescence units was defined and peaks below this limit were considered background noise. In addition, all samples were double-checked manually to confirm true alleles. At each locus the predominant allele (the one giving the highest peak) and minor alleles within at least two-thirds of the height of the predominant allele were scored [13, 15, 21]. Only predominant alleles were used to define haplotypes to ensure an unbiased estimate of minor allele frequency in polyclonal infections [15, 22–24].

Data analysis

Samples were defined as polyclonal if at least one locus presented more than one allele [13–14]. Polyclonal/locus (%) describes the percentage of samples identified as polyclonal by a given MS out of the total number of samples [13]. Multiplicity of infection (MOI), defined as the minimum number of different clones observed in a sample, was estimated by taking the maximum number of alleles at the two most polymorphic markers [13, 25]. Average MOI was defined as the sum of MOIs detected across all samples divided by the total number of samples. Average MOI and the proportion of monoclonal and polyclonal infections were compared with the Kruskal-Wallis and Pearson χ^2 test, respectively. A value of $p < 0.05$ was considered significant. The predominant alleles in each sample were used to calculate the number of

haplotypes by GenAEx 6.5 [26]. Haplotypes that appear only once in the population were defined as unique haplotypes.

Genetic diversity, defined as the probability of observing different genotypes at a given locus in two unrelated parasites, was assessed by calculating the expected heterozygosity (He) for each community using the formula: $He = [n/(n-1) * (1 - \sum p_i^2)]$, where n is the total number of alleles and p is the allele frequency. He ranges between 0 and 1, with values close to 1 indicating high genetic diversity [27]. Allelic richness, defined as the number of alleles per locus independently of sample size, was calculated using FSTAT v2.9.3 [28]. He and allelic richness were compared between communities using the Kruskal-Wallis test. The presence of bias due to false assignment of predominant haplotypes was investigated by comparing He in the database containing all the alleles and the database containing the predominant alleles [29].

A standardized index of association (I_A^s), calculated with LIAN v3.5, was used to assess the presence of multilocus linkage disequilibrium (LD) in the parasite population [30]. The significance of the I_A^s estimate was assessed with Monte Carlo simulation using 10,000 random permutations of the data. To differentiate between clonal propagation and epidemic expansion, we compared LD in the predominant allele dataset and the unique haplotype dataset [31]. Pairwise LD was used to evaluate the physical linkage between loci located within the same contig using the G statistic in FSTAT v2.3.9 [28, 32].

Genetic differentiation between pairs of communities was estimated using a pairwise unbiased estimator of F-statistics FSTAT v2.3.9 with no assumption of Hardy-Weinberg equilibrium within samples [28, 32]. A matrix of p values corresponding to each pairwise F_{ST} was calculated after Bonferroni correction, with a value of $p < 0.05$ considered significant. Crude F_{ST} values were adjusted for sample size using Recode Data v. 0.1 [33] and standardized F_{ST} estimates were obtained by dividing crude F_{ST} values by adjusted F_{ST} values. F_{ST} estimates ranged from 0 (no genetic differentiation between communities) to 1 (full differentiation).

As a complementary approach, population structure was investigated using the software programs STRUCTURE v2.3.2 [34], CLUMPP [35], DISTRUCT [36], and GENODIVE [37]. STRUCTURE was used to identify clusters of genetically related samples. The number of clusters (K) was set from 1 to 10 with 10 replications per K , and 150,000 Markov Chain Monte Carlo steps after a burn-in period of 50,000 iterations using the admixture model. The loc-prior model was used for accurate inference of population and individual ancestry. Next, we used STRUCTURE HARVESTER v0.6.94 [38] to calculate the most likely number of K clusters. Additional data parsing and formatting of the STRUCTURE output was performed using CLUMPP and DISTRUCT [38–39]. CLUMPP permutes the clusters' output by performing multiple replicate runs for the selected K . Samples with an average pairwise similarity (H value) of over 85% in one of the K clusters were considered to belong to that particular population; all other samples were considered admixed samples. DISTRUCT performs geographical displays of the aligned cluster assignment. We confirmed the optimal K , i.e. the K with the highest pseudo-F statistic, using AMOVA-based K -means clustering analysis in GENODIVE V2.0b23 (OS X 10.6 operating system). This method divides a number of individuals into an *a priori* assigned number of clusters (K) in such a way that minimizes within-group diversity and maximizes between-group diversity. The pseudo-F statistic was calculated by setting up simulated annealing runs with 150,000 steps and 50 algorithm repetitions to determine optimal clustering (highest *pseudo*- F-statistic).

eBURST v3 was used to identify clusters of closely related haplotypes, or haplogroups (HGs), which were defined as haplotypes sharing at least 9 loci from the 14 MS analyzed [40]. Haplotypes unrelated to any haplogroup (HG) were classified as singletons. The relationship between haplotypes following the defined K clusters was further analyzed by PHYLOViz [41].

Table 1. Baseline characteristics of study patients.

Community	Sample	EM	Age groups		Sex		Hamlets	HH	Season of collection		AS	MP	MG
			≤15y	>15 y	M	F			Dry	Rainy			
			Community 1 (%)	92 (39.3)	Cadong	55 (23.5)			37 (15.8)	51 (21.8)			
Community 2 (%)	57 (24.4)	Cadong	34 (14.5)	23 (9.8)	35 (15.0)	22 (9.4)	2 (14.2)	30 (24.2)	11 (4.7)	46 (19.7)	36 (15.4)	5829	1162
Community 3 (%)	32 (13.7)	Cadong	26 (11.1)	6 (2.6)	21 (8.9)	11 (4.7)	5 (35.8)	23 (18.5)	5 (2.1)	27 (11.5)	19 (8.1)	7144	490
Community 4 (%)	53 (22.6)	M'nong	29 (12.4)	24 (10.3)	33 (14.1)	20 (8.5)	3 (21.4)	26 (21.0)	6 (2.6)	47 (20.1)	24 (10.3)	5297	516
Total (%)	234 (100)		144 (61.5)	90 (38.5)	140 (59.8)	94 (40.2)	14 (100)	124 (100)	38 (16.2)	196 (83.8)	138 (59.0)	6691	723

EM, ethnic minority; HH, household; AS, asymptomatic infection (defined as no fever at enrolment), MP, mean number of asexual parasites per μL ; MG, mean number of gametocytes per μL .

doi:10.1371/journal.pntd.0004434.t001

Finally, to investigate the relationship between geographic and genetic distances in the study population, we performed principal coordinate analysis with the Mantel test for matrix correspondence in GenAlEx 6.5 [26]. Allele frequency was calculated in GenAlEx 6.5 using the predominant allele data set with 14 loci. The existence of a recent population bottleneck was investigated by evaluating the allele frequency distribution in the population (alleles at low frequencies are less abundant in populations with a recent bottleneck) [42–43].

Results

Baseline characteristics

In total, 234 individuals with *P. vivax* mono infection from the 260 recruited in the original study [12] were successfully genotyped and included in the analysis. Baseline characteristics of study participants are described in Table 1.

Genetic diversity

Successful genotyping, with at least 12 of the 14 MS, was achieved in 194 patients (82.9%). Allele data were successfully recovered in more than 83% of the samples for all MS except MS20 and Pvsal1814, for which successful amplification was achieved in 66% and 75% of samples respectively.

The MS characteristics are described in Table 2. Overall, genetic diversity was moderate, with an average $He = 0.68$ (95%CI 0.58–0.77) for all MS. MS3 and MS9 were the least polymorphic markers ($He = 0.45$ and 0.31 respectively), while MS10 and Pvsal1814 were the most polymorphic markers ($He = 0.99$ and 0.92 respectively), which were therefore used to calculate MOI. The number of alleles per MS ranged from 3 to 14. All MS had non-significant differences in He values in the database containing all alleles per locus and the predominant allele datasets ruling out bias in the construction of haplotypes from polyclonal infections ($p = 0.68$). The average number of alleles per locus was 5.5 (95%CI 3.82–7.17), the average number of alleles detected in a sample by any locus was 1.14 (95%CI 1.0–1.27) and the average allelic richness was 5.05 (95%CI 4.11–5.98).

The proportion of polyclonal infections, similar in all communities ($p > 0.05$), was 71.4% (167/234) when all 14 MS were used, but 64.1% (141/220) ($N = 220$ as 14 samples had missing

Table 2. Characteristics of the 14 microsatellite loci used in *P. vivax* populations from four communities in Quang Nam, Vietnam (n = 234).

Locus	Repeat sequence	ASR	No of alleles	H _E	A*	PI/locus (%)	hMOI	Average alleles/ locus
MS1	(GAA)11	211–239	5	0.664	4.94	3.41	3	1.04
MS2	(TAAA)2TATA (TAAA)6 TATA (TAAA)19	179–215	4	0.686	4.00	5.12	2	1.06
MS3	(GAA)11	185–199	3	0.446	3.00	4.27	2	1.11
MS4	(AGT)18	189–237	6	0.728	5.94	20.94	3	1.24
MS5	CCTCTT(CCT)11	164–206	5	0.697	4.94	1.70	2	1.01
MS6	(TCC)2(TCT)3(CCT)2(TTC)2 GCTTCT(TCC)10	242–254	5	0.655	5.00	5.55	2	1.06
MS7	(GAA)9	142–157	3	0.601	3.00	5.55	2	1.06
MS8	(CAG)2(CAA)11	198–268	6	0.703	5.94	5.55	2	1.06
MS9	(GGA)18	155–170	3	0.311	3.00	2.56	2	1.03
MS10	GAA(GGA)2AGA(GGA)9AGA(GGA)4AGAGGAAGA(GGA)3AGAGGAAGA(GGAAA)4(GGA)2(AGA)11(GGA)3(AGA)2GGAAGA(GGA)2	184–222	9	0.993	8.94	9.82	3	1.12
MS12	(TTC)10(TGC)4	207–224	5	0.618	5.00	5.55	2	1.06
MS15	(TCT)10	241–281	5	0.525	5.00	5.12	2	1.05
MS20	(GAA)11GAG(GAA)13(CAA)4GAA(CAA)5	148–206	4	0.680	4.00	3.41	2	1.05
Pvsal 1814	(AGA)44	546–704	14	0.919	7.99	53.84	3	2.01
All samples		142–704	77	0.676	5.05	71.37	2.2	1.14

ARS, Allele size range; A, Allelic richness; hMOI, highest multiplicity of infection; PI, Polyclonal infection.

* Predominant alleles only

doi:10.1371/journal.pntd.0004434.t002

data for MS10 and Pvsal1814) when only Pvsal1814 and MS10, the most polymorphic markers were used. The same two MS were used to calculate MOI (based on 220 samples with completed data), whose mean in the four communities was 1.91 (95% CI 1.81–2.02), with no significant differences between communities ($p = 0.52$), age groups ($MOI_{\leq 15\text{years}} = 2.0$ vs $MOI_{> 15\text{years}} = 1.9$, $p = 0.50$), gametocyte carriage ($MOI_{\text{gametocytes present}} = 1.93$ vs $MOI_{\text{gametocytes absent}} = 1.80$, $p = 0.42$), sex ($MOI_{\text{male}} = 1.89$ vs $MOI_{\text{female}} = 1.95$, $p = 0.55$), symptomatic vs asymptomatic (defined as fever at enrolment vs no fever at enrolment, $MOI = 1.98$ vs $MOI = 1.80$, respectively) ($p = 0.10$), season ($MOI_{\text{rainy season}} = 1.90$ vs $MOI_{\text{dry season}} = 1.97$, $p = 0.64$), and ethnic minority ($MOI_{\text{Cadong}} = 1.89$ vs $MOI_{\text{M' nong}} = 2.0$, $p = 0.46$). No significant differences were found between the four communities for either level of He ($p = 0.08$) or allelic richness ($p = 0.31$).

We identified 101 haplotypes from 144 samples of which 84 haplotypes were defined as unique haplotype with complete genotyping data for 13 MS. MS20 was excluded because it had the lowest successful genotyping rate (66%). Of these haplotypes, 25.7% (26/101) were found in monoclonal infections and 16.8% (17/101) were found in both monoclonal and polyclonal infections; 6.93% of haplotypes (7/101) had a frequency of over 2 in 40 samples and the two most frequent haplotypes were detected in 7.6% (11/144) and 6.2% (9/144) of samples. One haplotype was shared between the four communities, 4 haplotype found in community 1 were also present in community 2 and one haplotype shared between community 3 and community 4.

To evaluate the existence of a recent population bottleneck we analyzed the allele frequency distribution in the population. Fig 2 shows an L-shaped distribution of allele frequencies, as would be expected from neutral evolution.

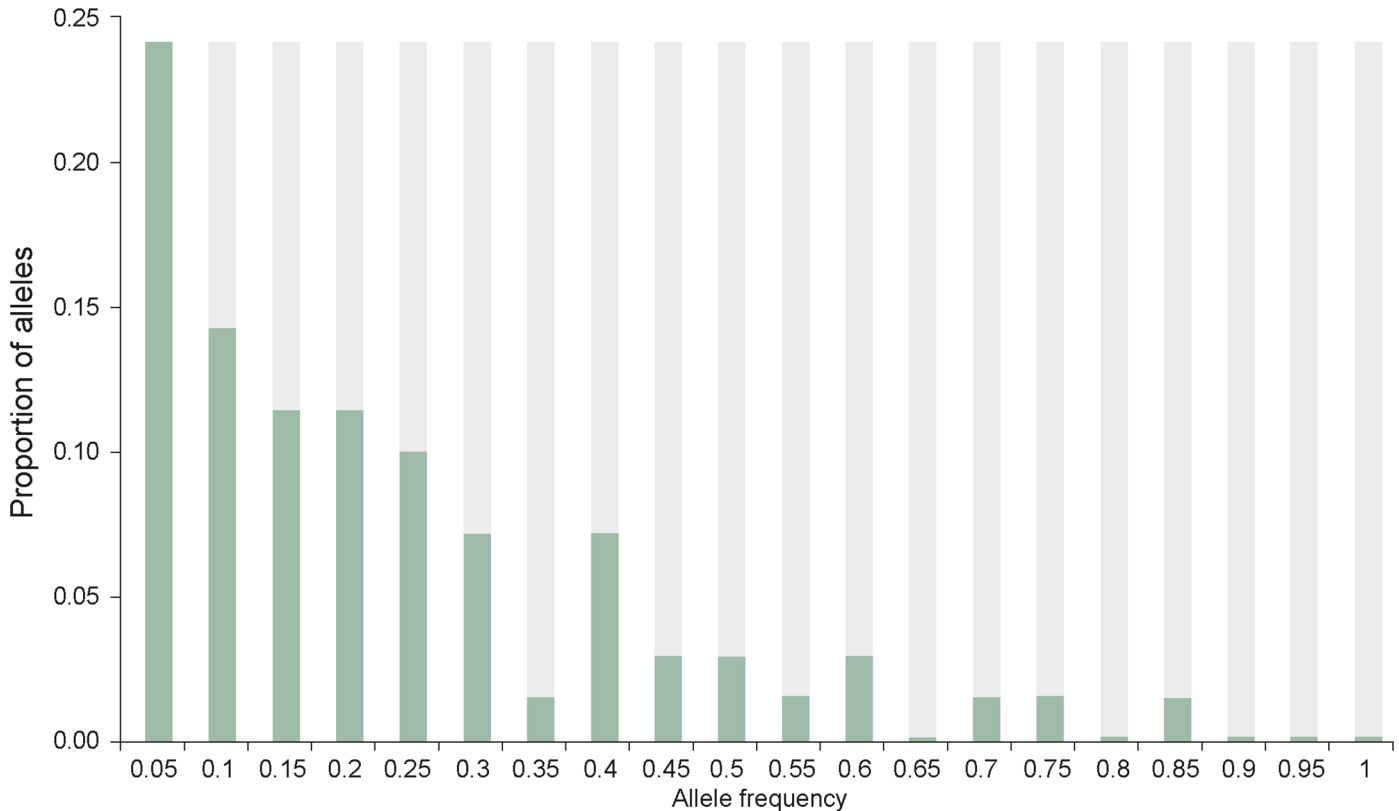


Fig 2. Allele frequency distribution in four communities in Quang Nam (n = 234). Predominant allele database and all loci combined are included.

doi:10.1371/journal.pntd.0004434.g002

Linkage disequilibrium

MS20 was also excluded from the LD analysis to maximize sample size and avoid bias due to an imbalanced number of samples between communities. Hence at least 25% of samples per community were included in the analysis. Significant LD was observed in each community (I_A^s ranged from 0.10 to 0.17) and in the overall study population ($I_A^s = 0.113, p < 0.001$). LD remained significant ($I_A^s = 0.059, p < 0.001$) when only the unique haplotypes were used. We then examined patterns of LD between pairs of MS (Fig 3). Pairwise LD was observed between loci located within the same contigs (MS4-MS5, MS7-MS8, and MS12-MS15) and also within different contigs. Even though lower pairwise LD was observed in communities 3 and 4, the fact that the overall LD was significant ($p = 0.008$) suggests the existence of a clonal parasite population.

Genetic differentiation

We first compared the datasets containing only monoclonal infections (n = 67) with the predominant allele (n = 234) and found low genetic differentiation ($F_{ST} = 0.05$), indicating absence of bias. Then, we calculated F_{ST} values for pairwise genetic differentiation between the four communities (Table 3). We observed moderate genetic differentiation between community 4 and the other communities ($F_{ST} = 0.15-0.18$) and low differentiation for the other combinations ($F_{ST} < 0.1$) (n = 144), indicating that the parasite population in community 4 is moderately, although significantly, differentiated from parasite populations in communities 1, 2, and

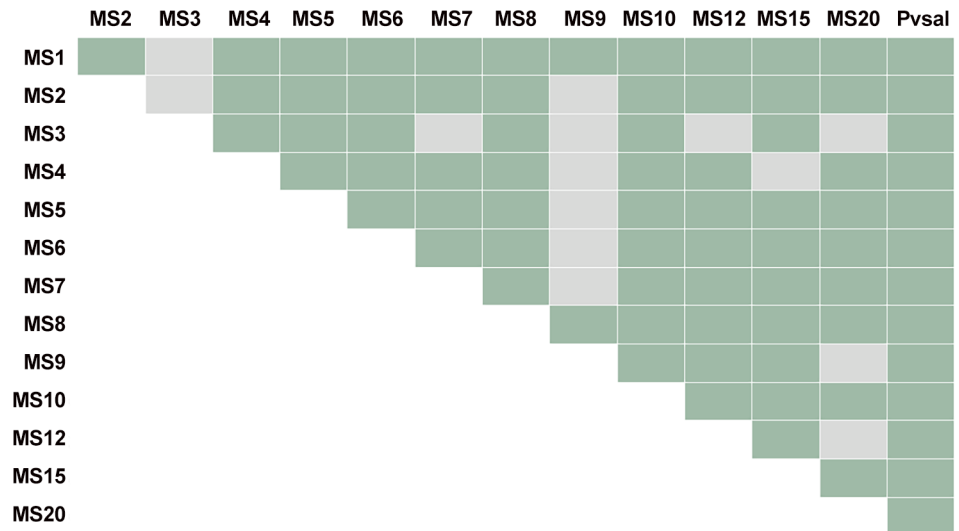


Fig 3. Pair-wise loci linkage disequilibrium analysis (N = 234). Significant association between alleles at pairs of loci in each sample was tested using FSTAT. Green color denotes LD at 5% based on 36400 permutations. Markers within the same contig are: MS4-MS5, MS7-MS8, MS12-MS15.

doi:10.1371/journal.pntd.0004434.g003

3 ($p = 0.008$). Similar F_{ST} values were obtained when only unique haplotypes ($n = 84$) were used.

Population structure

Structure analysis identified the most likely clusters in the population to be (i) $K = 7$ ($\Delta K = 9.4$), (ii) $K = 2$ ($\Delta K = 5.2$), and (iii) $K = 3$ ($\Delta K = 3.2$) ($n = 144$). The AMOVA-based K -means clustering analysis identified $K = 2$ as the optimal number of clusters (pseudo- $F = 30.5$). We further analyzed the parasite population divided by $K = 2$ (cluster 1 and cluster 2) with CLUMPP and DISTRUCT (Fig 4), and found that 33.3% (48/144) of the samples (with complete haplotypes) observed in the study population belonged to cluster 1 and 20.1% (29/144) belonged to cluster 2 and 46.6% (67/144) were admixed samples. Community 4 had the highest proportion of admixed samples (62.1%), followed by community 3 (53.9%), while community 1 and 2 had similar rates (40.9% and 41.3% respectively). The proportion of admixed samples remained high when the number of clusters was set to $K = 3$ and $K = 7$ (48.6% and 38.9%, respectively). Of note, cluster 1 samples were absent from community 4, which supports a moderate degree of population structure between communities 1–3 and community 4.

Table 3. Pairwise genetic differentiation.

Standardize F_{ST}	Community 1	Community 2	Community 3	Community 4
Community 1		0.046*	-0.028	0.153*
Community 2	0.043		-0.005	0.176*
Community 3	0.026	-0.035		0.168*
Community 4	0.174*	0.080*	0.145*	

The upper-right section shows the standardized fixation index (F_{ST}) values obtained from 13 microsatellite markers ($n = 144$, MS20 excluded). The lower-left section shows F_{ST} values obtained from unique haplotypes from 13 microsatellite markers ($n = 84$).

* P value <0.05

doi:10.1371/journal.pntd.0004434.t003

However, principal coordinate analysis failed to detect geographical clustering (per community) in the population.

Then, we investigated genetic relatedness, defined as haplotypes sharing at least 10/13 loci ($n = 101$) by eBURST. Eight different HGs and 13 singletons were identified. HG1 contained 51.5% (52/101) of all haplotypes in the four communities, while the other 7 HGs contained between 2.0% (2/101) to 11.9% (12/101) haplotypes. However, when relatedness was defined as haplotypes sharing at least 9/13 loci, only 2 HGs and 2 singletons were identified. HG1 included 95.0% (96/101) of all haplotypes detected. PHYLOVIZ analysis supported the existence of related haplotypes among all study communities (with slightly clustering of community 4 samples) (Fig 5A) and confirmed the absence of cluster 1 haplotypes in community 4 (Fig 5B).

Discussion

We analyzed the genetic diversity and population structure of 234 *P. vivax* pre-treatment clinical isolates collected in a forested area of Central Vietnam between April 2009 and December 2010 [12]. We observed moderate levels of heterozygosity in all four study communities, with a high proportion of polyclonal infections and significant LD, suggestive of inbreeding across parasite populations circulating in the study communities. Genetic differentiation and population structure between study communities was low but present between villages at each side of the river defining a moderate geographical barrier to gene flow.

In this study we used eight MS (MS1, MS4, MS6, MS9, MS10, MS12, MS15, and MS20) with balanced diversity, three (MS2, MS5 and Pvsal1814) with unbalanced diversity, one (MS8) with significant excess diversity, and two (MS3 and MS7) with significant reduced diversity [25]. Mean He in the study population ($He = 0.68$) using those 14 MS was non-significantly different to He when only MS with balanced diversity (recommended for measuring population diversity) were used. Therefore, all MS were kept in the analysis to assess both diversity parameters and polyclonal infections, but only the two most polymorphic markers (MS10 and Pvsal1814) were kept to investigate MOI.

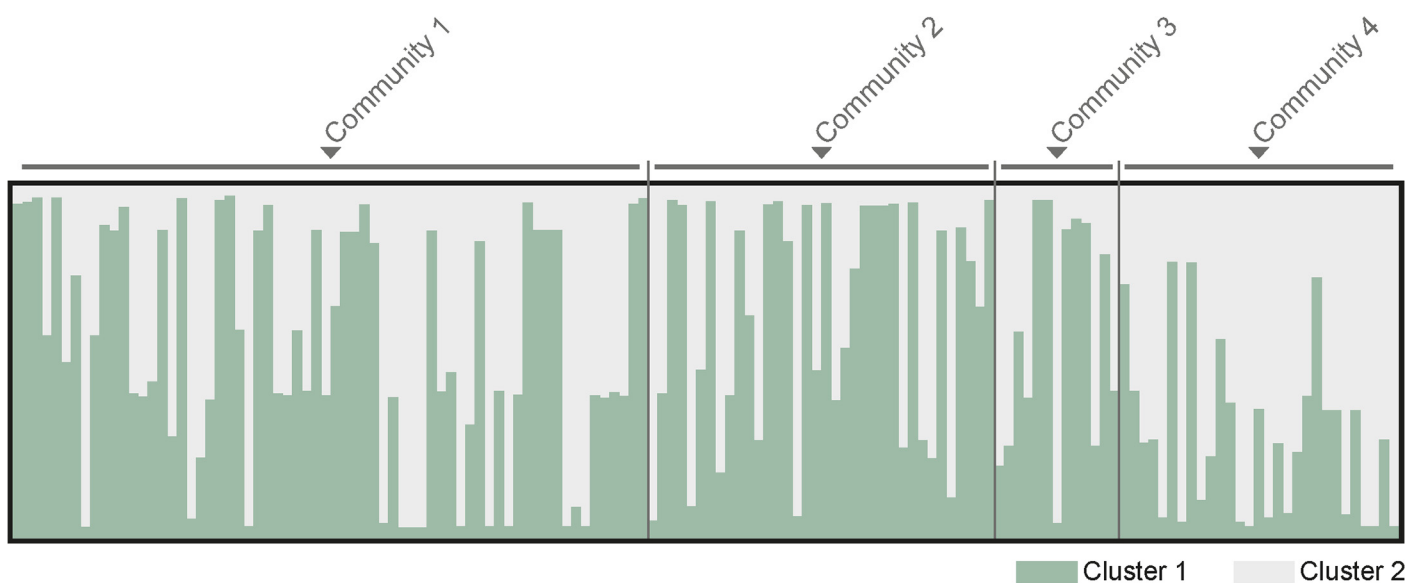


Fig 4. Genetic clustering analysis by STRUCTURE. The graph represent the clustering model when the parasite population was grouped into 2 clusters ($K = 2$). In the bar plot, each isolate is represented by a single vertical line divided into K colors.

doi:10.1371/journal.pntd.0004434.g004

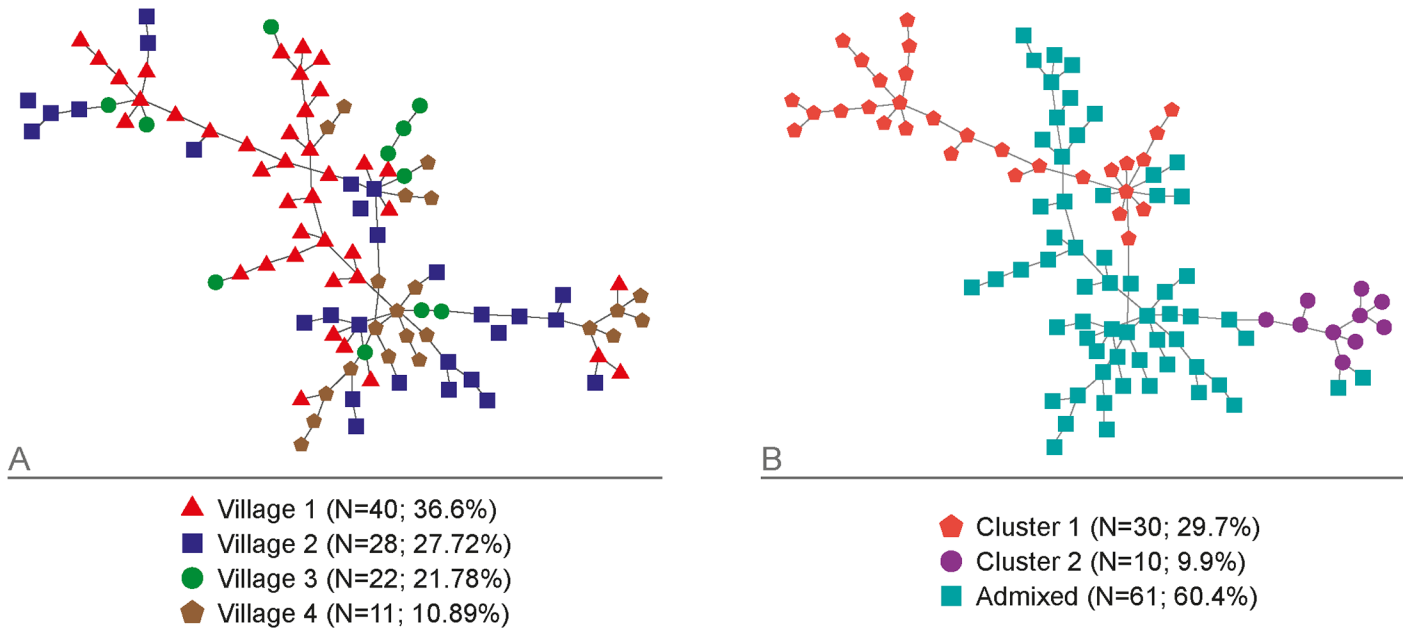


Fig 5. Phylogenetic relationship of 101 unique *P. vivax* haplotypes using PHYLOViZ. Haplotypes relationships by (A) community, and (B) clustering $K = 2$. Each colored figure represents a unique haplotype and grey lines indicate shared alleles among individual haplotypes.

doi:10.1371/journal.pntd.0004434.g005

The mean He in our study population ($He = 0.68$) was similar to figures seen in areas of north-west Brazil with similar transmission intensities ($He = 0.74$ and $He = 0.68$) [43–44]; higher than those observed in South Korea ($He = 0.43$) [45] and the Loreto district, Peru ($He = 0.37$) [16], and lower than those seen in Sri Lanka ($He = 0.89$) [46], Pursat, Cambodia ($He = 0.84$) [16], and Binh Thuan, Central Vietnam ($He = 0.88$) [13]. In our study MS9, MS3, and MS7 displayed the lowest number of alleles per locus ($n = 3$) and MS9 and MS3 had the lowest He values ($He_{MS9} = 0.31$ and $He_{MS3} = 0.45$). MS3, MS7, and MS9 would therefore appear to be poorly informative markers in the study area and their use in future studies is not recommended.

Polyclonal infections were frequent (71.4%) when the results from 14 MS were combined and moderately lower (64.1%) when just Pvsal1814 and MS10 were used [47]. Mean MOI was 1.91, with similar MOI observed in symptomatic and asymptomatic study participants, possibly because of the high parasite density (mean 3,919/μL; 95%CI 2,852–4,986) detected in asymptomatic participants at day 0. In the literature, the proportion of polyclonal infections vary considerably depending on the MS markers used [16, 48], highlighting the need for a standardized methodology that allows comparison between studies and geographical regions. High proportions of polyclonal infections have also been reported in hypo-endemic areas in Sri Lanka (60%) [49], Colombia (60–80%) [48], the Amazon Basin in Brazil (50%) [43], and more recently, in a pre-elimination context in Sri Lanka (69%) [46]. It is noteworthy in a study carried out (1999–2000) in Binh Thuan province, central-south Vietnam, where the entomological inoculation rate was estimated at 1 infective bite/person/year, 100% of vivax infections were polyclonal with a mean MOI of 3.7 [13]. The high levels of genetic diversity and polyclonal infections in low transmission areas [13,23,48] can be, at least partially, explained by the unique biology of *P. vivax* which result in (i) a high prevalence of asymptomatic and low parasite density infections (which last longer because are difficult to detect, increasing the likelihood of repeated infections with divergent clones, resulting in increased polyclonality) and (ii) relapse from dormant liver stages (the reactivation of heterologous clones increases the likelihood of

peripheral superinfections). Since a high proportion of study participants were asymptomatic at recruitment (59.0%) and poor adherence to PQ radical cure is known in the study area [3], the high proportion of polyclonal infections found in this study may reflect peripheral superinfection fed by heterologous clones from both relapses and reinfections.

Despite those high rates of polyclonal infections, we observed a significant LD ($I_A^s = 0.113$, $p < 0.001$) in the overall study population. Asexual clones present in one infection produce gametocytes that, taken by the vector, recombine during meiosis and generate new haplotypes in a process known as outcrossing. Consequently, the breakdown of pre-existing associations between unlinked loci would reduce LD to low levels [50] as opposed to recombination between gametes from the same parasite [51]. As transmission decreases, fewer parasite types will be present in the population and recombination will often occur between related parasites, increasing the level of inbreeding in the population. This is supported by the fact that 53.8% of all polyclonal infections were identified by multiple alleles at just one locus. Indeed, LD remained significant in the analysis using only unique haplotypes, indicating that it is a result of inbreeding rather than expansion of few haplotypes due to outbreaks or epidemics [31]. Closely related parasites in hypoendemic areas have been previously reported [52, 53]. In addition, inbreeding was further supported by overall significant pairwise LD [21, 31]. LD combined with high levels of polyclonality has been reported in rural Amazonia [54] and more recently in Sri Lanka [46]. The authors of these studies offered two alternative interpretations for this phenomenon. First, the MS may not be strictly neutral (10/14 MS mapping to loci encoding either hypothetical or annotated proteins may be subject to natural selection) [22–23]. And second, replication-slippage events during mitotic (asexual) replication could result in the generation of new alleles due to the addition or deletion of repeats [49, 55]. If the replication-slippage rate is higher than that of effective recombination (the probability of producing a recombinant genome), the clones generated would increase polyclonality, without altering LD.

It has been previously reported that replication-slippage events (and therefore number of alleles per locus and He) correlate positively with increasing repeat length and non-perfect repeats motifs, i.e. interrupted or compound motifs [25, 56]. Pvsal1814 MS used in this study, which had an (AGA)₄₄ motif structure with an interrupted/compound motif, $He = 0.91$ and 14 different alleles with frequencies ranging from 1.3% to 16%, identified 53.8% of all polyclonal samples in the study population. Indeed, inherent mutability in this MS has been described to produce excess diversity, which in turn is recommended to identify MOI [25].

We identified 101 haplotypes, of which 84 appeared only once in the population. Ninety percent of them were grouped in a single haplogroup (HG1), defined by identical alleles in at least 9/13 loci, indicating a high degree of relatedness among parasites across the communities. These results support the view that despite a high level of polyclonality, inbreeding among highly related haplotypes maintains LD.

The adjusted genetic differentiation was low between communities 1, 2, and 3 ($F_{ST} < 0.05$) and moderate when community 4 was included ($F_{ST} = 0.15–0.18$), indicating limited geographical boundaries between neighboring communities 1–3 but higher differentiation with the community across the river. In concordance with the F_{ST} values, the STRUCTURE analysis detected two main parasite populations. Two clusters of haplotypes, with a high proportion of mixture haplotypes (60.4%) were observed in all four communities. The fact that a majority of haplotypes found in community 4 belonged to cluster 2, which was the minor cluster in the other 3 communities, supports a certain degree of differentiation between communities 1–3 and 4. Moderate population differentiation between these communities can be explained by geographical proximity and socioeconomic relationships between the communities' inhabitants as previously described [3]. Inhabitants of community 1–3 (located at one side of the river) belong to the Cadong ethnic group and therefore share some degree of kinship,

facilitating social exchange. Conversely, community 4, whose inhabitants belong to the M'ngong ethnicity, is located at the other side of the river with limited access during the rainy season.

Malaria incidence in the Quang Nam province has dropped by 78.0% over the last decade thanks to the implementation of efficient control strategies [1,17,57]. At the time of the study, malaria prevalence in the study area was 7.8% as assessed by light microscopy and 23.6% as estimated by PCR [3]. Therefore, the moderate-to-high levels of genetic diversity detected, together with the high polyclonality and low population structure are consistent with an epidemiological context of transition from moderate to low endemicity [58–59].

Future studies aiming at identifying changes in genetic diversity and population structure to support the development or improvement of control and elimination interventions should include isolates collected at several time points from all areas where malaria is prevalent (or has been recently eliminated). Ideally, a molecular surveillance system should be implemented within the existing network of sentinel sites for drug resistance across the country to support evaluation of interventions and improve response strategies at the provincial level.

Parasite populations with strong LD and the presence of gene flow could fuel the spread of resistant parasites in the event of the emergence of drug resistance, threatening current treatment efforts and achievements towards malaria elimination in Central Vietnam. Temporal analysis to investigate haplotype persistence and the risk of clonal expansion is urgently needed in order to inform decision makers.

Acknowledgments

We would like to thank all study participants as well as the local and provincial authorities for their support and trust.

Author Contributions

Conceived and designed the experiments: NVH CDR PVdE NTHB BQP JPVG UD AE ARU. Performed the experiments: NVH CDR PG. Analyzed the data: NVH CDR PG ARU. Wrote the paper: NVH CDR PVT PVdE PG NTHB BQP TTD JPVG UD AE ARU.

References

1. National Institute of Malariaology, Parasitology and Entomology (2015) Annual Report of the National Malaria Control Program in Vietnam 2014, Hanoi, Vietnam. National Institute of Malariaology, Parasitology and Entomology.
2. TT Duong, ND Thang (2015) Malaria epidemiological stratification in Vietnam 2014. National report in Malariaology, Parasitology and Entomology, National Institute of Malariaology, Parasitology and Entomology Hanoi, Vietnam 11:21, Medical Publisher INSB 978-604-66-0950-6.
3. Pham VT, Nguyen VH, Nguyen VV, Malderen CV, Obsomer V, Rosanas-Urgell A, Grietens KP, Banccone G, Chowwiwat N, Tran TD, Nguyen XX, D'Alessandro U, Speybroeck N, Erhart A (2015) Epidemiology of forest malaria in Central Vietnam: the hidden parasite reservoir. *Malar J* 14:86. doi: [10.1186/s12936-015-0601-y](https://doi.org/10.1186/s12936-015-0601-y) PMID: [25880664](https://pubmed.ncbi.nlm.nih.gov/25880664/)
4. Erhart A, Ngo DT, Phan VK, Ta TT, Van Overmeir C, Speybroeck N, Obsomer V, Le XH, Le KT, Coosemans M, D'Alessandro U (2005) Epidemiology of forest malaria in central Vietnam: a large scale cross-sectional survey. *Malar J* 4:58. PMID: [16336671](https://pubmed.ncbi.nlm.nih.gov/16336671/)
5. Trung HD, Van Bortel W, Sochantha T, Keokenchanh K, Quang NT, Cong LD, Cooseman M (2004) Malaria transmission and major malaria vector in different geographical areas of southeast Asia. *Trop Med Int Health* 9:230–237. PMID: [15040560](https://pubmed.ncbi.nlm.nih.gov/15040560/)
6. National Institute of Malariaology, Parasitology and Entomology (2011) National strategy for malaria control and elimination for the period of 2011–2020 and orienting to 2030. Government of socialist republic Vietnam, (http://www.chinhphu.vn/portal/page/portal/chinhphu/noidungchienluocphattrienkinhtexahoi?piref135_16002_135_15999_15999.strutsAction=ViewDetailAction.do&piref135_16002_135_15999_15999.docid=1160&piref135_16002_135_15999_15999.substract=).

7. Whilte NJ (2011) Determinant of relapse periodicity in *Plasmodium vivax* malaria. *Malar J* 10:297. doi: [10.1186/1475-2875-10-297](https://doi.org/10.1186/1475-2875-10-297) PMID: [21989376](https://pubmed.ncbi.nlm.nih.gov/21989376/)
8. Mueller I, Galinsky MR, Baird JK, Carlton MJ, Kochar DK, Alonson PL, Adel Portillo H (2009) Key gap in the knowledge of *Plasmodium vivax*: a neglected human malaria parasite. *Lancet Infect Dis* 9:555–66. doi: [10.1016/S1473-3099\(09\)70177-X](https://doi.org/10.1016/S1473-3099(09)70177-X) PMID: [19695492](https://pubmed.ncbi.nlm.nih.gov/19695492/)
9. Nguyen HV, Van den Eede P, Van Overmeir C, Thang ND, Hung Le X, D'Alessandro U, Erhart A (2012) Marked age-dependent prevalence of symptomatic and patent infections and complexity of distribution of human *Plasmodium* species in central Vietnam. *Am J Trop Med Hyg* 87:989–995. doi: [10.4269/ajtmh.2012.12-0047](https://doi.org/10.4269/ajtmh.2012.12-0047) PMID: [23128294](https://pubmed.ncbi.nlm.nih.gov/23128294/)
10. Imwong M, Snounou G, Pukrittayakamee S, Tanomsing N, Kim JR, Nandy A, Guthmann JP, Nosten F, Carlton J, Looareesuwan S, Nair S, Sudimack D, Day NP, Anderson TJ, White NJ (2007) Relapses of *Plasmodium vivax* infection usually result from activation of heterologous hypnozoites. *J Infect Dis* 195:927–933. PMID: [17330781](https://pubmed.ncbi.nlm.nih.gov/17330781/)
11. World Health Organization (2012) *Plasmodium vivax* control and Elimination: development of Global Strategy and Investment Case. Malaria policy Committee Meeting 11–13 September 2012, WHO HQ session 8.
12. Thanh PV, Hong NV, Van NV, Melva L, Baird K, Xa NX, Hung LX, Duong TT, Rosanas-Urgell A, Speybroeck N, D'Alessandro U, Erhart A (2015). Confirmed *Plasmodium vivax* resistance to chloroquine in Central Vietnam. *Antimicrob. Agents Chemother.* 00791–15
13. Van den Eede P, Erhart A, Van der Auwera G, Van Overmeir C, Thang ND, Hung le X Anne J, D'Alessandro U (2010) High complexity of *Plasmodium vivax* infections in symptomatic patients from a rural community in central Vietnam detected by microsatellite genotyping. *Am J Trop Med Hyg* 82:223–227. doi: [10.4269/ajtmh.2010.09-0458](https://doi.org/10.4269/ajtmh.2010.09-0458) PMID: [20133996](https://pubmed.ncbi.nlm.nih.gov/20133996/)
14. Van den Eede P, Soto-Calle VE, Delgado C, Gamboa D, Grande T, Rodriguez H, Llanos-Cuentas A, Anne J, D'Alessandro U, Erhart A (2011) *Plasmodium vivax* sub-patent infections after radical treatment are common in Peruvian patients: results of a 1-year prospective cohort study. *PLoS One* 6:e16257. doi: [10.1371/journal.pone.0016257](https://doi.org/10.1371/journal.pone.0016257) PMID: [21297986](https://pubmed.ncbi.nlm.nih.gov/21297986/)
15. Christopher D, Veronica ES, Peter V DE, Dionica G, Angle R, Emmanuel N A, Hugo R, Alejandro L, Jean-Pierre V G, Annette E, Umberto D (2014): Population structure and spatio-temporal transmission dynamics of *Plasmodium vivax* after radical cure treatment in a rural village of the Peruvian Amazon. *Malaria J* 13:8.
16. Orjuela- Sanchez P, Sá JM, Brandi MC, Rodrigues PT, Bastos MS, Amaratunga C, Duong S, Fairhurst RM, Ferreira MU (2013) Higher microsatellite in *Plasmodium vivax* than in sympatric *Plasmodium falciparum* population in Pursat western Cambodia. *Exp Parasitol* 134(3):318–326. PMID: [23562882](https://pubmed.ncbi.nlm.nih.gov/23562882/)
17. Phuc BQ, Manh ND, Binh NTH Ngoc NTH, Tan HV, Hanh NV, Hanh TV, Dao LD, Tinh TT. (2011) PCR investigation of malaria parasite structure in Nam Tra My District, Quang Nam province. National report in Malariology, Parasitology and Entomology, Hanoi, Vietnam 162: 171. National Institute of Malariology, Parasitology and Entomology.
18. Quang Nam provincial Malaria Station (2009–2013) Annual report on malaria entomological collections in Quang Nam province, Internal Report. Quang Nam provincial Malaria Station Tam Ky, Quang Nam, Vietnam.
19. World Health Organization (2010). Guidelines for the treatment of malaria. Second Edition, Geneva, Switzerland. [<http://apps.who.int/medicinedocs/en/d/Js19105en/>].
20. Rubio JM, Post RJ, van Leeuwen WM, Henry MC, Lindergard G, Hommel M (2002) Alternative polymerase chain reaction method to identify *Plasmodium* species in human blood samples: the semi-nested multiplex malaria PCR (SnM-PCR). *Trans R Soc Trop Med Hyg* 96: S199–S204. PMID: [12055839](https://pubmed.ncbi.nlm.nih.gov/12055839/)
21. Karunaweera ND, Ferreira MU, Hartl DL, Wirth DF (2007) Fourteen polymorphic microsatellite DNA markers for the human malaria parasite *Plasmodium vivax*. *Molecular Ecology Notes* 7:172–175.
22. Anderson TJC, Su X-Z, Bockarie M, Lagog M, Day KP (2009) Twelve microsatellite markers for characterization of *Plasmodium falciparum* from finger prick blood samples. *Parasitology* 119:113–25.
23. Ferreira MU, Karunaweera ND, da Silva-Nunes M, da Silva NS, Wirth DF, Hartl DL (2007) Population structure and transmission dynamics of *Plasmodium vivax* in rural Amazonia. *J Infect Dis* 195: 1218–1226. PMID: [17357061](https://pubmed.ncbi.nlm.nih.gov/17357061/)
24. Koepfli C, Timinao L, Antao T, Barry AE, Siba P, Mueller Ivo, Felger I (2013) A Large Reservoir and Little Population Structure in the South Pacific. *PLoS One* 8: e66041. PMID: [23823758](https://pubmed.ncbi.nlm.nih.gov/23823758/)
25. Sutton P.L. (2013) A call to arms: on refining *Plasmodium vivax* microsatellite marker panel for comparing global diversity. *Malar J* 12:447. doi: [10.1186/1475-2875-12-447](https://doi.org/10.1186/1475-2875-12-447) PMID: [24330329](https://pubmed.ncbi.nlm.nih.gov/24330329/)

26. Peakall R, Smouse PE (2012) GenAlEx 6.5 genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 10.1093.
27. Gunawardena S, Karunaweera ND, Ferreira MU, Phonr-Kyaw M, Pollack RJ, Alifrangis M, Rajakaruna RS, Konaradsen F, Amerasinghe PH, Schousboe ML, Galappaththy GN, Abeyasinghe RR, Hartl DL, Wirth DF (2010) Geographic structure of *Plasmodium vivax*: microsatellite analysis of parasite populations from Sri Lanka, Myanmar and Ethiopia. *Am J Trop Med Hyg* 82:235–242. doi: [10.4269/ajtmh.2010.09-0588](https://doi.org/10.4269/ajtmh.2010.09-0588) PMID: [20133999](https://pubmed.ncbi.nlm.nih.gov/20133999/)
28. Goudet J (1995) FSTAT version 1.2: a computer program to calculate F-Statistic. *J Hered* 86(6); 485–486.
29. Barry AE, Schultz L, Senn N, Nake J, Kiniboro B, Siba PM, Mueller I, Reeder J (2013) High levels of genetic diversity of *Plasmodium falciparum* populations in Papua New Guinea despite variable infection prevalence. *Am J Trop Med Hyg* 88: 718–725. doi: [10.4269/ajtmh.12-0056](https://doi.org/10.4269/ajtmh.12-0056) PMID: [23400571](https://pubmed.ncbi.nlm.nih.gov/23400571/)
30. Haubold B, Hudson RR (2000) LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics* 16:847–848. PMID: [11108709](https://pubmed.ncbi.nlm.nih.gov/11108709/)
31. Smith JM, Smith NH, O'Rourke M, Spratt BG (1993) How clonal are bacteria? *Proc Natl Acad Sci USA* 90:4384–4388. PMID: [8506277](https://pubmed.ncbi.nlm.nih.gov/8506277/)
32. Gout. J FSTAT. <http://www2.unil.ch/popgen/software/fstat.htm>. Last access 15th July 2015
33. Meirmans. P Recode Data v.0.1 (http://www.bentleydrummer.nl/software/recodeData_manual.pdf). Last access 15th July 2015
34. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. PMID: [15969739](https://pubmed.ncbi.nlm.nih.gov/15969739/)
35. Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 123 (14): 1801–6.
36. Noah A. Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Note* 4: 137–138.
37. Meirman PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4: 792–794.
38. Earl Dent A, vonHoldt Bridgett M (2014) Structure harvester: a website and program for visualizing STRUCTURE out put and empling Evano method. *Conversation Genet Resour* 4: 359–361.
39. Evano G, Regnaut S, Goudet J (2005). Detecting number of clusters of individual using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. PMID: [15969739](https://pubmed.ncbi.nlm.nih.gov/15969739/)
40. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG (2004) eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* 186:1518–1530. PMID: [14973027](https://pubmed.ncbi.nlm.nih.gov/14973027/)
41. Fansisco AP, Vaz C, Monterio PT, Cristino JM, Ramirez M, Carrrio JA (2012) Phyloviz: Phylogentic inference and data visualization for sequence base typing methods. *BMC Bioinformatics* 13–87.
42. Gluikart G, Allendorf FM, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J Heredity* 89: 238–247.
43. Batista CL, Barbosa S, Da Silva Bastos M, Viana SA, Ferreira MU (2015) Genetic diversity of *Plasmodium vivax* over time and space: a community-based study in rural Amazonia. *Parasitology* 142:374–384. doi: [10.1017/S0031182014001176](https://doi.org/10.1017/S0031182014001176) PMID: [25068581](https://pubmed.ncbi.nlm.nih.gov/25068581/)
44. Orjuela-Sanchez P, DaSilva NS, da Silva- Nunes M, Ferreira MU (2009) Recurrent parasitemias and population dynamic of *Plasmodium vivax* polymorphisms in rural Amazonia. *Am J Trop Med Hyg* 81:961–968. doi: [10.4269/ajtmh.2009.09-0337](https://doi.org/10.4269/ajtmh.2009.09-0337) PMID: [19996423](https://pubmed.ncbi.nlm.nih.gov/19996423/)
45. Iwagami M, Fukumoto M, Hwang S, Kim S, Kho W, Kano S (2012) Population structure and transmission dynamic of *Plasmodium vivax* in the republic of Korea based on microsatellite DNA analysis. *PLoS One* 6:e1592.
46. Gunawardena S, Ferreira MU, Kapilananda GM, Wirth DF, Karunaweera ND (2014) The Sri Lankan paradox: high genetic diversity in *Plasmodium vivax* populations despite decreasing level of malaria transmission. *Parasitology* 141:880–890. doi: [10.1017/S0031182013002278](https://doi.org/10.1017/S0031182013002278) PMID: [24533989](https://pubmed.ncbi.nlm.nih.gov/24533989/)
47. Arnott A, Barry AE, Reeder JC (2012) Understanding the population genetic of *Plasmodium vivax* is essential for malaria control and elimination. *Malar J* 11:14. doi: [10.1186/1475-2875-11-14](https://doi.org/10.1186/1475-2875-11-14) PMID: [22233585](https://pubmed.ncbi.nlm.nih.gov/22233585/)
48. Imwong M, Nair S, Pukrittayakamee S, Sudimack D, Williams JT, Mayxay M, Newton PN, Kim JR, Nandy A, Osorio L, Carlton JM, White NJ, Day NP, Anderson TJ (2007) Contracting genetic structure in *Plasmodium vivax* population from Asia and South America. *Int J Parasitol* 37: 1013–1022. PMID: [17442318](https://pubmed.ncbi.nlm.nih.gov/17442318/)

49. Karunaweera ND, Ferreira MU, Munasinghe A, Barnwell JW, Collins WE, King CL, Kawamoto F, Hartl DL, Wirth DF (2008) Extensive microsatellite diversity in the human malaria parasite *Plasmodium vivax*. *Gene* 410: 105–112. doi: [10.1016/j.gene.2007.11.022](https://doi.org/10.1016/j.gene.2007.11.022) PMID: [18226474](https://pubmed.ncbi.nlm.nih.gov/18226474/)
50. http://www.as.wvu.edu/~kgarbutt/QuantGen/Gen535_2_2004/Linkage_disequilibrium.htm, Last access on 17th July, 2015.
51. Mzilahowa T, McCall PJ, Hastings IM (2007) “Sexual” population structure and genetic of the malaria agent *P. falciparum*. *PloS One* 2(7):e613. PMID: [17637829](https://pubmed.ncbi.nlm.nih.gov/17637829/)
52. Chenet SM, Schneider KA, Villgeas L, Escalante AA (2012) Local population structure of *Plasmodium*: impact on malaria control and elimination. *Malar J* 11:412. doi: [10.1186/1475-2875-11-412](https://doi.org/10.1186/1475-2875-11-412) PMID: [23232077](https://pubmed.ncbi.nlm.nih.gov/23232077/)
53. Obaldia N, Barol NK, Calzada JE, Santamiaria AM, Daniel R, Wong W, Chang H, Hamilton EJ, Arevalo-Herrera M, Herrera S, Wirth DF, Hartl DL, Marti M, Wolkman SK (2015) Clonal outbreak of *Plasmodium falciparum* in eastern Panama. *J Infect Dis*. 211 (7):1087–1096. PMID: [25336725](https://pubmed.ncbi.nlm.nih.gov/25336725/)
54. Ferreira MU, Karunaweera ND, Silva- Nunes Md, Silva MSd, With FD, Hartl DL (2007) Population structure and transmission dynamics of *Plasmodium vivax* in Rural Amazonia. *J Infect Dis* 195:1218–26. PMID: [17357061](https://pubmed.ncbi.nlm.nih.gov/17357061/)
55. Levinson G (1987) Slipped–Strand mispairing: amajor mechanism for DNA sequence evolution. *Mol Biol Evol* 4: 203–21 PMID: [3328815](https://pubmed.ncbi.nlm.nih.gov/3328815/)
56. Pearson CE, Nichol Edamura K, Cleary JD (2005) Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet*. 6:729–742 PMID: [16205713](https://pubmed.ncbi.nlm.nih.gov/16205713/)
57. NV Van, NK Huy, NT Nhan, BQ Phuc (2015) Some epidemiological aspect of malaria confirmed cases at Nam Tra My district, Quang Nam province in 2013. National report in Malariology, Parasitology and Entomology, Hanoi, Vietnam 22–30. National Institute of Malariology, Parasitology and Entomology.
58. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, Kabaria CW, Manh BH, Elyazar IRF, Bookrt S, Smith DL, Moyeed RA (2009) A world malaria map: *Plasmodium falciparum* endemicity in 2007. *PloS One* 6:e1000048.
59. Erhart E, Thang ND, Hung NQ, Toi LV, Hung LX, Tuy TQ, Cong LD, Speybroeck N, Coosemans M, D’Alessandro U (2004) Forest malaria in Vietnam: A challenge for control. *Am J Trop Med Hyg* 70:110–118 PMID: [14993619](https://pubmed.ncbi.nlm.nih.gov/14993619/)