Infection, Genetics and Evolution 85 (2020) 104573

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research Paper

Genetic diversity of Merozoite surface protein 1–42 (MSP1-42) fragment of *Plasmodium vivax* from Indonesian isolates: Rationale implementation of candidate MSP1 vaccine

E. Elsa Herdiana Murhandarwati[a,b], E. Henny Herningtyas[c], Puspawati Puspawati[d], Fridolina Mau[e], Shen-Bo Chen[a,f], Hai-Mo Shen[a,f], Jun-Hu Chen[a,f,*]

[a] National Institute of Parasitic Diseases, Chinese Centre for Disease Control and Prevention, WHO Collaborating Centre for Tropical Diseases, National Centre for International Research on Tropical Diseases, Key Laboratory of Parasite and Vector Biology of the Chinese Ministry of Health, Shanghai 200025, People's Republic of China
[b] Department of Parasitology, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia
[c] Department of Clinical Pathology and Laboratory Medicine, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia
[d] District Hospital Ratu Zalecha Martapura, Kabupaten Banjar, Indonesia
[e] Institute of Research and Development Waikabubak, Sumba Barat, Indonesia
[f] The School of Global Health, Chinese Centre for Tropical Diseases Research, Shanghai JiaoTong University School of Medicine, Shanghai 200011, People's Republic of China

ARTICLE INFO

Keywords:
Plasmodium vivax
Merozoite surface protein-1
PvMSP1-42
Genetic diversity
Natural selection
Indonesia

ABSTRACT

Morbidity and mortality related to malaria in Indonesia are attributed to both *Plasmodium falciparum* and *P. vivax* parasites. In addition to vaccines for *P. falciparum*, vaccines against *P. vivax* are urgently needed for the prevention of the disease. An extensively studied antigen is the carboxyl-terminus of the 42 kDa region of *P. vivax* merozoite surface protein-1 (PvMSP1-42). The design of a vaccine based on this antigen requires an understanding of the extent of polymorphism. However, there is no information on the genetic diversity of the antigen in Indonesia. This study aimed to profile the diversity of PvMSP1-42 and its two subdomains (PvMSP1-33 and PvMSP1-19) among Indonesian *P. vivax* isolates. A total of 52 *P. vivax*-infected blood samples were collected from patients in two different endemic areas in Indonesia: Banjarmasin (Kalimantan) and Sumba Timur (Nusa Tenggara Timur). The polymorphic characteristics and natural selection of PvMSP1-42 were analyzed using the DnaSP, MEGA, and Structure software. Thirty distinct haplotypes of PvMSP1-42 were identified. They displayed amino acid changes compared to the reference PVP01 sequence. Most of the mutations were concentrated in the 33 kDa fragment. PvMSP1-42 of the Indonesian isolates appeared to be under positive selection. Recombination may also play a role in the resulting genetic diversity of PvMSP1. In conclusion, PvMSP1-42 of Indonesian isolates displayed allelic polymorphisms caused by mutation, recombination, and positive selection. These results will aid the understanding of the *P. vivax* population in Indonesia and to develop a PvMSP1 based vaccine against *P. vivax*.

1. Background

Malaria is a serious public health problem in most tropical and subtropical countries, including Indonesia. Despite the decreased incidence of malaria in Indonesia, 30% of the Indonesian population is still at risk, particularly those who reside in the eastern part of the country (Sitohang et al., 2018). Malaria is caused by five *Plasmodium* parasites and those of most concern are *Plasmodium falciparum* and *P. vivax*. The latter can cause severe manifestations (Anstey et al., 2009; Poespoprodjo et al., 2009). *Plasmodium vivax* infection can result in relapses (Imwong et al., 2007) and hemolysis caused by primaquine, as part of artemisinin combination therapy in patients deficient in glucose-6-phosphate dehydrogenase (Satyagraha et al., 2015). In Indonesia, *P. vivax* was responsible for close to 40% of all cases of malaria between 2010 and 2016 (WHO, 2017).

To reduce the impact of *P. vivax* malaria, prevention of the disease by vaccination must be considered. One of several promising and well-studied candidates is merozoite surface protein 1 (MSP1). MSP1 of *P. vivax* as well as *P. falciparum* is predicted to be involved in invasion of the parasite into erythrocytes. MSP1 fragments have been reported to

* Corresponding author at: National Institute of Parasitic Diseases, Chinese Centre for Disease Control and Prevention, Shanghai 200025, P.R. China.
E-mail address: chenjh@nipd.chinacdc.cn (J.-H. Chen).

https://doi.org/10.1016/j.meegid.2020.104573
Received 14 June 2020; Received in revised form 19 September 2020; Accepted 23 September 2020
Available online 26 September 2020

be immunogenic in populations in endemic areas (Chen et al., 2010; Pitabut et al., 2007). During infection by *P. vivax,* a naturally acquired antibody response to MSP1 is prevalent (Bastos et al., 2007).

The MSP1 protein undergoes two proteolytic steps during schizogony and release of merozoites. After the second proteolytic processing, MSP1 generates a C-terminal 42 kDa fragment (MSP1-42). The fragment is subsequently cleaved to 33 kDa (MSP1-33) and 19 kDa (MSP1-19) fragments (Blackman et al., 1990). The MSP1-42 and MSP1-19 fragments are considered promising candidates for malaria vaccines. These fragments are reportedly immunogenic (Riccio et al., 2013) and have been correlated with immunity to malaria in infants, children, and adults (Valderrama-Aguirre et al., 2005).

As PvMSP1-42 or PvMSP1-19 reportedly stimulate immune responses, clarification of their polymorphism is important. Genetic diversity is an important contribution impeding the development of malaria vaccines (Barry et al., 2012). Limited polymorphism is preferable when antigens are selected for vaccine development to ensure the inability of a strain to escape from the host immune response. The geography of Indonesia is also a challenge to vaccine development. Indonesia is an archipelago of thousands of islands traversed by Weber and Wallace lines. There are main three regions: western, central, and eastern. Each region is influenced by the natural environment including natural obstacles and movement of animals in the wild. In turn, each region has unique and characteristic diversities of wild animals and plants. The malaria vector (*Anopheles* spp.) may also influence the diversity of *Plasmodium* parasites.

This study profiled the diversity of PvMSP1-42 among Indonesian isolates collected from two different areas with different endemicity in Indonesia. One area consisted the Banjar and Banjarbaru Districts located in South Kalimantan Province (Kalimantan Island). The other area was Sumba Tengah District located in East Nusa Tenggara Province (East Nusa Tenggara/Flores Island). In particular, we aimed to analyze the diversity among the regions coding for PvMSP1-42, which includes PvMSP1-33 and PvMSP1-19.

2. Methods

2.1. Sample collection

Blood samples were collected from fever patients visiting a Primary Health Care who met the inclusion criteria of the study (i.e. *P. vivax* infected cases diagnosed by microscopic examination with asexual parasite density 1–100,000/ul, male and female aged > 1 year and without disease complication). The subjects were enrolled at Ratu Zalecha Hospital, which serves Banjar and Banjarbaru District, Banjarmasin (BJ), South Kalimantan Province, from 2012 to 2014 (44 samples), and the Primary Health Care at Sumba Tengah District (ST), East Nusa Tenggara, during a Mass Blood Survey performed in 2015 (20 samples) (Mau, 2016) (Fig. 1). Transmission was constant throughout the year. The annual incidence of parasites was approximately 1.5% for BJ and 7% for ST. *Plasmodium falciparum* and *P. vivax* were the dominant species. BJ had a medium to low Annual Parasite Incidence (API) of 1.43% in 2013 and 0.68% in 2015. The API was relatively high in ST at the same times (16.37% in 2013 and 7.04% in 2015) (Kementrian Kesehatan Republik Indonesia, 2015; Profil Kesehatan Indonesia Tahun, 2013). This study was approved by the Faculty of Medicine, Universitas Gadjah Mada (UGM) Ethical Committee (No. KE/FK/0052/EC/2017) and the Ethics Committee at the National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention. All adult subjects provided signed informed consent for blood collection. Informed consent for blood collection from younger subjects was provided by their legal guardians.

2.2. DNA isolation, amplification, and sequencing analysis

Blood samples stored with anticoagulant (2 mL) at 4 °C were sent to the Parasitology Laboratory of the Faculty of Medicine, Public Health and Nursing, UGM, for DNA extraction, amplification, and analysis. DNA was isolated by the Genomic DNA Mini Kit for blood/culture cells (Geneaid, New Taipei City, Taiwan) or the QIAamp blood kit (QIAGEN, Inc., Chatsworth, CA) according to the manufacturer's instructions. The DNA was stored at 4 °C until PCR was performed. To identify *Plasmodium* species, nested PCR was performed as previously described (Johnston et al., 2006).

The PvMSP1-42 fragment was amplified using the Pv1SF 5′-AGAA GAAAACGTAGCAGCAA-3′ and Pv1SR 5′-AAGCCCAGTTCAGAAC TCA-3′ primers (Zhou et al., 2017) or nested PCR primers P1: 5′-GAT GACGACGGGGAGGAAGACC-3′, P2: 5′-AAGCTCCATGCACAGGAGG-3′, followed by N1: 5′-GACCAAGTAACAACGGGAG-3′ and N2: 5′-GGACA AGCTTAGGAAGCTGG-3′ primers (Dias et al., 2011). Cycling parameters and primary and secondary reactions were performed as described previously (Dias et al., 2011). Ultraviolet visualization of the amplified product following 1% agarose gel electrophoresis revealed an approximately 1400 bp band using the Pv1SF and Pv1SR primers, and approximately 1200 bp using the N1 and N2 primers.

Purified PCR products were sequenced using the forward primers on an ABI PRISM 3700 DNA capillary sequencer (Genetika Science, Tangerang City, Malaysia). The sequences of PvMSP1-42 were confirmed with sequencing using forward and reverse primers. Sequencing was performed with internal forward and reverse primers to confirm inconsistent results (Fig. S1). The primers were: F1: 5′-TCCCGTTCCT GAATAGCC-3′ and/or F2: 5′ -GGAACGGAAGAATGGAGATGCTTGTT AAC-3′ and/or F3: 5′-ATACAAGCTGCTCGACTTGGAGAAGAAG-3′ and/or R1: 5′-TTCTTAGAAGTTCTG-3′, and/or R2: 5′-CTTTTGGAGGCTAT TCAGGAACGGGA-3′).

New sequences identified in this study were deposited in GenBank with accession numbers MK733842-MK733893.

2.3. Data analyses

We first performed a multiple sequence alignment of the MSP1 gene sequences using MEGA6 (Tamura et al., 2013). The *P. vivax* MSP1 sequence PVP01_0728900.1 (5205 bp) obtained from PlasmoDB (http://PlasmoDB.org) was used as the full MSP1 reference. All sequences were assembled and aligned using the Clustal W method. After all sequences were aligned, data were trimmed to 1131 bp, which consisted of MSP1-33 (867 nucleotides) and MSP1-19 fragments (264 nucleotides). To investigate the genetic polymorphism and natural selection of *P. vivax* MSP1-42 in BJ, ST, and between these two populations, we used DnaSP (Librado and Rozas, 2009) to calculate populations, number of haplotypes (H), mean value of nucleotide differences (K), nucleotide diversity (pi), and the ratio of nonsynonymous to synonymous substitution (dN/dS). Additionally, Tajima's D test was performed in DnaSP to evaluate the neutrality theory of evolution (with Fu and Li's test as a double-check).

To visualize the genetic relationships among the isolates, a maximum likelihood (ML) phylogenetic tree for each gene was constructed in MEGA6 with 1000 bootstrap replications (Tamura et al., 2013). The probability of recombination between adjacent nucleotides per generation and the minimum number of recombination events (Rm) were calculated using DnaSP. The linkage disequilibrium (LD) between different polymorphic sites was computed based on the D and R^2 indices.

Samples were collected from different regions of Indonesia with different endemicities. To assess allele ancestry, Structure software was used to assess clustering of isolates (Evanno et al., 2005). Ten iterations for the numbers of clusters (K) from 1 to 10 were run, each with a burn-in period of 5000 steps and 10,000 Markov chain Monte Carlo iterations.

To identify amino acid polymorphisms in the *P. vivax* MSP1-42 between two geographically diverse endemic areas, nucleotide sequences were translated into protein sequences.

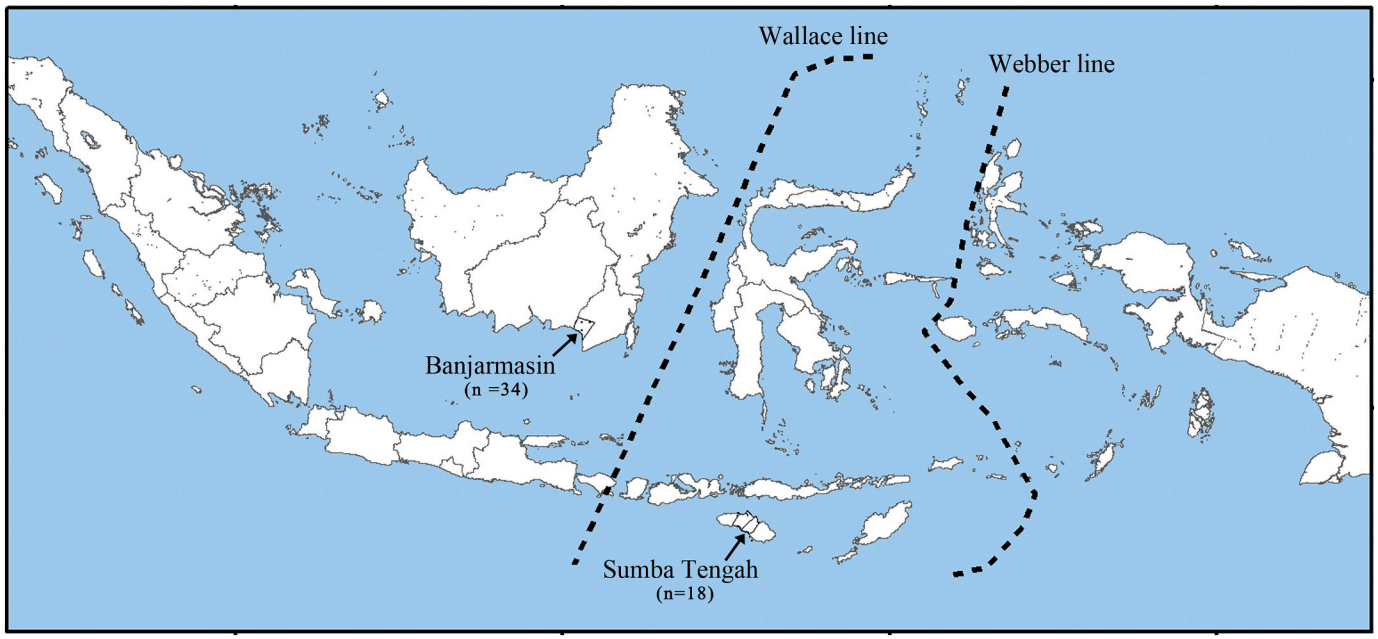

Fig. 1. Map of malaria sampling collection sites. Thirty-four samples of 44 *Plasmodium vivax* malaria samples and 18 of 20 *P. vivax* malaria samples from Banjarmasin, Kalimantan and Sumba Tengah, East Nusa Tenggara (black arrows), Indonesia, were successfully sequenced. The line separating the two areas is the Wallace line.

3. Results

Thirty-four of 44 BJ malaria samples were infected by *P. vivax*. The 34 samples were successfully sequenced for the PvMSP1-42 (1131 bp) corresponding to amino acid positions 1350 to 1726 of PvMSP1-42 of the PVP01 reference. For ST, 20 samples were infected with *P. vivax* and 18 were successfully sequenced.

Nucleotide sequence analysis of the 52 PvMSP1-42 sequences revealed no size variations between the sequences, but they showed polymorphic characteristics (Fig. 2). Of the 30 haplotypes found, 2 were shared in both the BJ and ST, 17 in BJ, and 11 in ST.

3.1. PvMSP1-42 fragment polymorphism

Further analysis of the sequences revealed that 64 polymorphic nucleotide positions and one distinct point mutation were evident in BJ and ST. Of the polymorphisms, 58 were dimorphic, 6 were trimorphic, and 1 was distinct at nucleotide position 1072 (Fig. S2).

The number of PvMSP1-42 haplotypes in both BJ and ST, BJ alone, and ST alone was 30, 19, and 14, respectively, with a haplotype diversity of 0.903, 0.890, and 0.935, respectively (Table 1). The PvMSP1-42 haplotypes in the studied populations were highly diverse in the total population (30 of 52, 58%). BJ displayed moderate diversity (19 of 34, 56%) and ST displayed high diversity (14 of 18, 78%). The nucleotide diversity of the PvMSP1-42 fragment for isolates in both BJ and ST, BJ alone, and ST alone was 0.01448, 0.01593, and 0.01168, respectively. Most of these variations occurred in the middle of the fragment, with the remainder of the fragment conserved (Fig. S2).

To examine the balancing selection of the PvMSP1-42 fragment, Tajima's D test was performed. Tajima's D value was 0.56477 for BJ alone, −0.83481 for ST alone, and 0.26001 for the total population ($p > 0.1$; Table 1). However, similar to the DNA polymorphism, Tajima's D test also showed opposite selection directions in different locations (Fig. 2). In the middle of the fragment (from nucleotide position 300 to 600), Tajima's D value was positive for BJ and negative for ST, consistent with results in Fu and Li's tests. For the rest of the PvMSP1-42 fragment, Tajima's D value was negative, which normally indicated directional selection.

The dN/dS in PvMSP1-42 for all 54 isolates was 1.66037. The dN/dS values were 1.653371 for BJ alone and 1.72118 for ST alone, suggesting a positive selection for PvMSP1-42 of *P. vivax* populations in BJ, ST, and both areas (Table 1).

The estimated minimum number of recombination events between adjacent polymorphic sites (Rm) in the isolates was 14 for both BJ and ST, 12 for BJ alone, and 7 for ST alone for PvMSP1-42. These high recombination parameter values suggested that meiotic recombination may occur between sites, resulting in genetic diversity. The LD index (R^2) also declined across PvMSP1-42, suggesting that intragenic recombination may also contribute to the PvMSP1-42 diversity (Fig. 4). Furthermore, we performed the LD test on each region, and found BJ sub-population showed similar pattern for LD SNP pairs with the whole population (Fig. S3A). But the ST sub-population showed less LD SNP pairs which may suggest more recombination ratio and complexity of origin (Fig. S3B). D < Tajima's D or Fu and Li's D* > was far greater than R^2, which suggested that most of these loci had low frequency alleles (Fig. 4).

Amino acid sequencing revealed 27 different proteins. Amino acid changes were dimorphic at 32 positions, trimorphic at 9, tetramorphic at 2, and distinct at amino acid position 358 (Fig. 3).

3.2. PvMSP1-33 fragment polymorphism

Thirty-four PvMSP1-33 sequences of the 52 PvMSP1-42 sequences were unique and were analyzed. Most of the nucleotide polymorphisms occurred in this region at 58 sites. They comprised 55 dimorphic and 3 trimorphic polymorphisms. Polymorphisms at distinct nucleotides were absent (Fig. S2).

The numbers of PvMSP1-33 haplotypes were 26 in both BJ and ST, 15 in BJ alone, and 13 in ST alone with a haplotype diversity of 0.840, 0.809, and 0.902, respectively (Table 1). The PvMSP1-33 haplotypes in the studied *P. vivax* populations were highly diverse in both areas (26 of 52, 50%). BJ displayed moderate diversity (15 of 34, 44%) and ST displayed higher diversity (13 of 18, 72%). The nucleotide diversity of the MSP1-33 fragment for isolates in BJ and ST, BJ, and ST was 0.01838, 0.02024, and 0.01476, respectively. Most of these variations occurred in the middle of the fragment, with the remainder of the

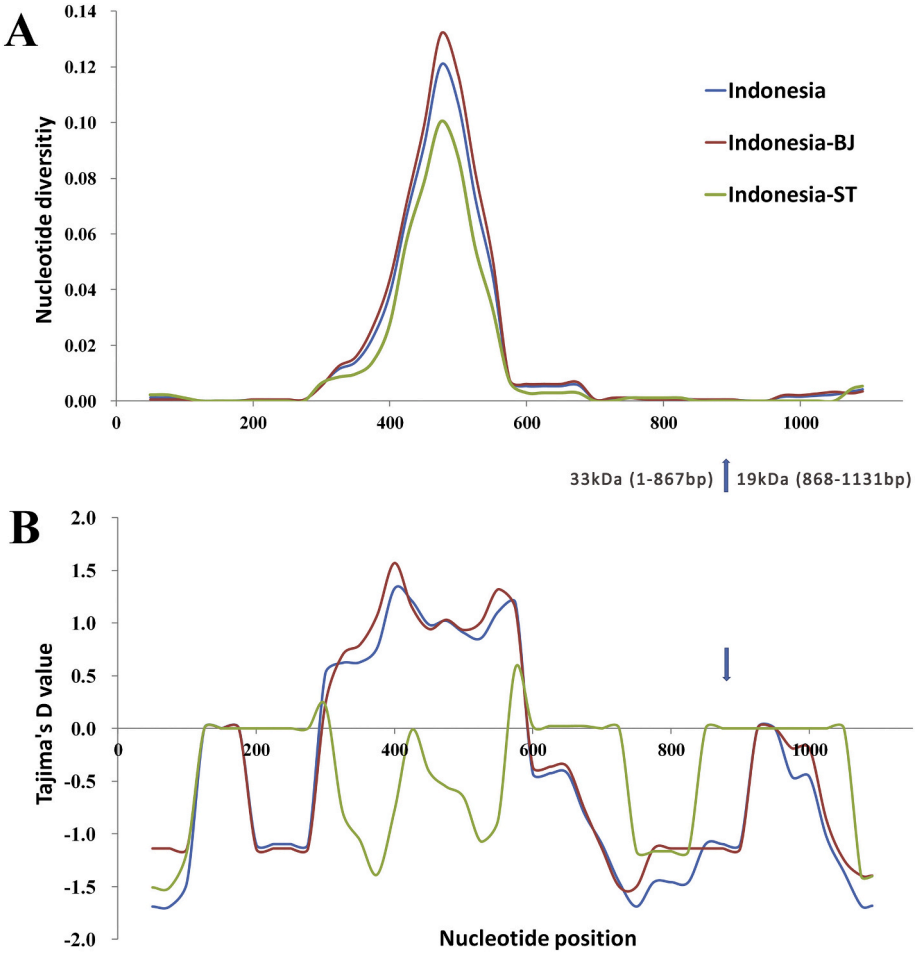

Fig. 2. Nucleotide diversity and Tajima's D value for PvMSP1-42 of Indonesian isolates. A. Position of PvMSP1-42 nucleotide diversities. B. Tajima's D value for PvMSP1-42 of Indonesian isolates. Red, green, and blue lines represent BJ, ST, and Indonesian isolates, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

fragment being conserved (Fig. 3).

The dN/dS in PvMSP1-33 for all 34 isolates was 0.55293. The dN/dS values were 1.808549 and 2.085427 for the isolates from BJ and ST, respectively, suggesting a positive selection for PvMSP1-33 of *P. vivax* populations from either or both areas (Table 1).

The Rm was estimated as 13, 11, and 7 for PvMSP1-33 isolated in both BJ and ST, BJ alone, and ST alone, respectively. These high recombination parameter values suggested that meiotic recombination may occur between sites, resulting in genetic diversity. Similar to the situation in PvMSP1-42, R^2 also declined across PvMSP1-33, suggesting that intragenic recombination may also contribute to the PvMSP1-33 diversity. In addition, D was far greater than R^2, which suggested that most of these loci with association were with low frequency alleles (Fig. 4).

Amino acid sequencing revealed 22 different proteins. Amino acid changes were dimorphic at 29 positions, trimorphic at 8, and tetramorphic at 1 (Fig. 3).

3.3. PvMSP1-19 fragment polymorphism

Eighteen PvMSP1-19 sequences of the 52 PvMSP1-42 sequences were unique and have been analyzed in this study. Nucleotide polymorphism in this region was less than that in the 33 kDa fragment, with only 7 mutations and 1 distinct nucleotide (G - > A at nucleotide position 1072) (Fig. S2).

Six polymorphic nucleotide positions and 3 distinct point mutations were evident from each of the two endemic areas. The number of MSP1-19 haplotypes in both areas, BJ alone and ST alone, was 7, 5, and 4, respectively, with a haplotype diversity of 0.840, 0.809, and 0.902, respectively (Table 1). The diversity of PvMSP1-19 was low in the total population (7 of 52, 13%). BJ displayed the lowest diversity (5 out of 34, 15%) and ST had slightly higher diversity than that of the total population (4 of 18, 22%). The nucleotide diversity of the PvMSP1-19 fragment for isolates in both areas, BJ, and ST was 0.00179, 0.00187, and 0.00162, respectively. Most of these variations occurred in the middle of the fragment, with the remainder of the fragment being conserved (Fig. 3A). Tajima's D value was negative, and the Fu and Li's tests confirmed its significance ($p < 0.05$).

The amino acid sequences revealed 5 different proteins. Amino acid changes were dimorphic at 5 positions with 1 distinct point at amino acid position 358 (Fig. 3).

3.4. Phylogenetic, Fst, and structure analyses of PvMSP1-42

ML phylogenetic tree of the 52 sequences with sequences published in Asian countries was performed. The results classified the samples into several clusters, but they did not completely follow the country-wise distribution (Fig. 5A). Despite the geographical distances, samples from maritime countries were more varied than those from continental countries. In our phylogenetic tree for each fragment, the 33 kDa fragment was similar to 42 kDa, but the 19 kDa fragment did not show enough geographic characteristics (Fig. S4A and S4B).

F_{st} values for PvMSP1-42 varied between countries, and there was a weak relationship with geographic origin. The F_{st} value for PvMSP1-42

Table 1
Estimation of nucleotide diversity and summary statistics of PvMSP1-42 in 52 *Plasmodium vivax* isolates grouping into their coding regions (MSP1-42, MSP1-33 and MSP1-19) and into two different regions Banjarmasin (BJ) and Sumba Tengah (ST).

\	N	S	π	θ	h	Hd	k	dN	dS	dN/dS	Tajima's D	*p*-value	Fu & Li D*	p-value	Fu & Li F*	p-value
MSP1-42	52	64	0.01448	0.01348	30	0.903	16.40724	0.01584	0.00954	1.660377	0.26001	$p > 0.10$	−0.05715	$p > 0.10$	0.07253	$p > 0.10$
BJ and ST																
BJ region	34	59	0.01593	0.01382	19	0.89	18.04635	0.01741	0.01053	1.653371	0.56477	$p > 0.10$	0.48695	$p > 0.10$	0.60851	$p > 0.10$
ST region	18	54	0.01168	0.01463	14	0.935	13.22876	0.01284	0.00746	1.72118	−0.83481	$p > 0.10$	−1.21457	$p > 0.10$	−1.28145	$p > 0.10$
MSP1-33	52	58	0.01838	0.01608	26	0.840	15.93137	0.01123	0.02031	0.55293	0.49597	$p > 0.10$	0.31857	$p > 0.10$	0.45890	$p > 0.10$
BJ and ST																
BJ region	34	55	0.02024	0.01693	15	0.809	17.54902	0.02031	0.01123	1.808549	0.72152	$p > 0.10$	0.64098	$p > 0.10$	0.79196	$p > 0.10$
ST region	18	51	0.01476	0.01811	13	0.902	12.79739	0.01660	0.00796	2.085427	−0.76355	$p > 0.10$	−1.16799	$p > 0.10$	−1.21895	$p > 0.10$
MSP1-19	52	6	0.00179	0.00499	7	0.346	0.47587	0.00125	0.00391	0.319693	−1.62064	$0.10 > p > 0.05$	−2.2901	$0.10 > p > 0.05$	−2.43722	$p < 0.05$
BJ and ST																
BJ region	34	4	0.00187	0.00368	5	0.367	0.49733	0.00159	0.00300	0.53000	−1.22167	$p > 0.10$	−1.02774	$p > 0.10$	−1.26153	$p > 0.10$
ST region	18	3	0.00162	0.00328	4	0.314	0.431	0.00053	0.00581	0.091222	−1.40138	$p > 0.10$	−1.19315	$p > 0.10$	−1.43299	$p > 0.10$

N, number of samples; S, number of polymorphic sites; π, nucleotide diversity; θ, Watterson estimator; H, the number of haplotypes; Hd, haplotype diversity; K, the average number of nucleotide differences; dN, the rates of nonsynonymous substitutions; dS, the rates of synonymous substitutions.

between BJ and ST was −0.00015. The F_{st} value for PvMSP1-33 between BJ and ST was −0.00028, which was less than the value for PvMSP1-19 (0.00438). Indonesian isolates were phylogenetically close to Cambodia, Thailand, and India, but were distant from Myanmar, South Korea, and Singapore (Table 2).

The structure results suggested five major groups among the Asian samples with optimized cluster values of K = 8. One group comprised individually associated or anomalous loci (Fig. 5B). Indonesian and Singaporean samples generally lacked some k components that were commonly found in continental samples. Furthermore, some of isolates from BJ exhibited regional characteritics (k7) may due to geographic isolation. A potential ancestor characterized by a set of allele frequencies was shared by BJ and Singaporean samples, but was rarer in ST.

4. Discussion

The PvMSP1 gene codes for an antigen that has been implicated as a major malaria vaccine candidate. However, the polymorphic nature of the antigen hinders the design for a protective vaccine (Galinski and Barnwell, 2008; Suphakhonchuwong et al., 2018; Valencia et al., 2011). This is the first study to report a polymorphism of PvMSP1 in Indonesia. We provide two sets of genetic information of PvMSP1 from the indigenous population of Kalimantan East Nusa Tenggara. We found 30 new haplotypes of PvMSP1-42 with 19 and 14 haplotypes, respectively, for BJ and ST from samples collected during 2014 and 2015. The mutation loci were mainly found in the nucleotide position 350 to 550 bp fragment and the peak nucleotide diversity value was 0.12. The haplotype diversity was lower in BJ and higher in ST. The findings demonstrated that the disease transmitting parasite populations are highly heterogenic and dynamic in these endemic areas.

Early studies suggested that the diversity of MSPs in *P. vivax* was presumably associated with parasite immune evasion and was important for the rationale of malaria vaccine designs (Han et al., 2018; McCaffery et al., 2019; Shen et al., 2017). Limited LD was detected in 52 isolates at different polymorphic sites. For those LD SNP pairs, the longest distance was < 200 bp. This finding suggests that meiotic recombination might occur between sites for a long time, and most regions of PvMSP1-42 are randomized. As an important part of the invasion process, the damaging mutations were suppressed by selective pressure, resulting in lower overall LD levels.

The dN/dS of PvMSP1-42 was > 1, indicating that PvMSP1-42 is affected by positive selection. The overall Tajima's D value of PvMSP1-42 was 0.26 for all 52 isolates, and the Tajima's D values were 0.56 and − 0.8 for the BJ and ST isolates, respectively, indicating different kinds of selection for PvMSP1-42. The dN/dS value indicated positive selection, which suggests polymorphisms of this gene in favor of avoiding host immune pressure.

Positive selection on PvMSP1 means that the advantageous allele increases as a consequence of superiority in survival. Genetic diversity results showed that the majority of polymorphic sites were in the middle portions, while other regions remained highly conserved as a result of positive selection. These polymorphic sites gave Indonesian isolates geographical characteristics, but they still could not be distinguished clearly in the ML tree. The structure results also showed substantial variability in the origin of the parasites. Presumably, all Asian *P. vivax* isolates originated from the same potential ancestor. However, analysis of these isolates showed distinct ancestors from the high prevalence of isolates among different countries.

High diversity observed in Indonesian isolates is similar to that in other studies from Asia (Dias et al., 2011; Thakur et al., 2008; Zhou et al., 2017) with a high dN/dS. The high diversity of PvMSP1-42 in Indonesia might be caused by the large and highly mobile effective population size. Gene flow could be introduced to different locations through movement from one island to another or that from overseas through sea trade on maritime routes. However, the different structures

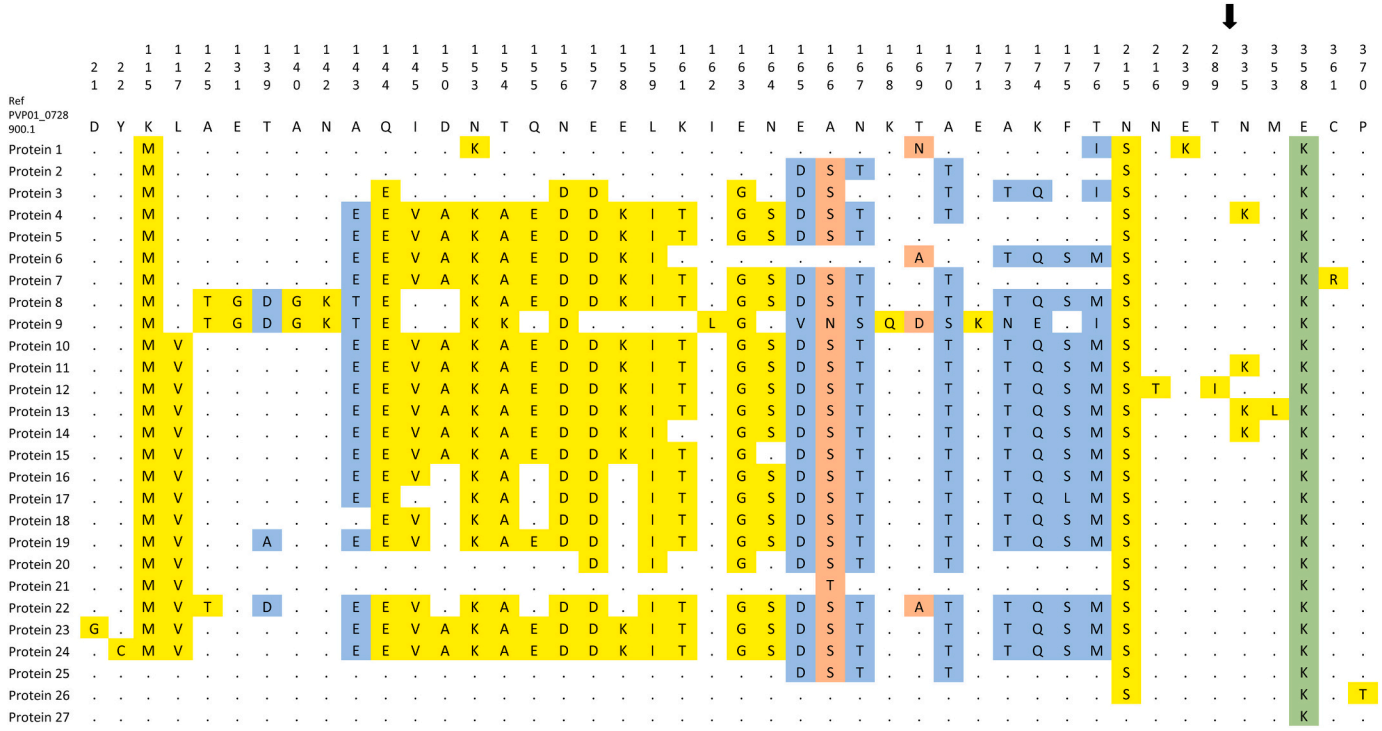

Fig. 3. Amino acid sequence polymorphism of PvMSP1-42 in Indonesian *Plasmodium vivax* isolates. Polymorphic amino acids were compared to the PvP01 (0728900) reference sequence and are listed for each haplotype. Dot (.) indicates that the amino acid is the same as PvP01_0728900.1. Amino acid sequences revealed 27 different proteins. Amino acid changes were dimorphic at 34 positions (highlighted in yellow), trimorphic at nine (highlighted in blue), tetramorphic at two (highlighted in red), and one distinct at amino acid 358 (highlighted in green). The arrow divides the protein into PvMSP1-33 (amino acids 1–289) and PvMSP1-19 protein (amino acids 290–377). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the BJ and ST sub-populations indicate that geographic isolation plays a role, as suggested by studies in Columbia and Turkey (Valderrama-Aguirre et al., 2011; Zeyrek et al., 2010). Furthermore, the high diversity was not expected as the samples were taken from isolated rural areas, such as ST, and from indigenous populations, such as BJ. We suspect that the diversity among Indonesian samples might have been inherited historically when Indonesia had a strategic geographical position along international maritime routes with India and China during the Silk Road period (Sulistiyono, 2013). The F_{st} results indicating that Indonesian isolates are closer to India rather than to other countries also supports this notion.

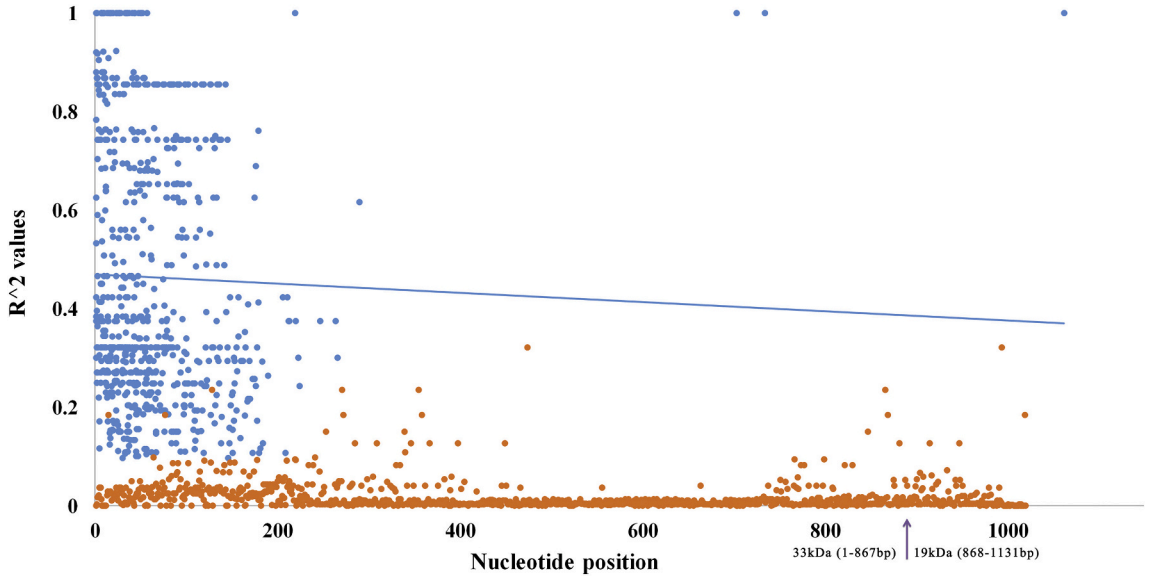

Fig. 4. Linkage disequilibrium (LD) of PvMSP1-42 of Indonesian isolates. LD across the PvMSP1-42 gene of Banjarmasin (Kalimantan) and Sumba Tengah (East Nusa Tenggara) isolates, Indonesia, was calculated using R^2. Significant LD values among samples are shown as calculated by Fisher's exact test. Trace line represents the regression line. Orange and blue dots represent non-significant and significant R^2 values, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

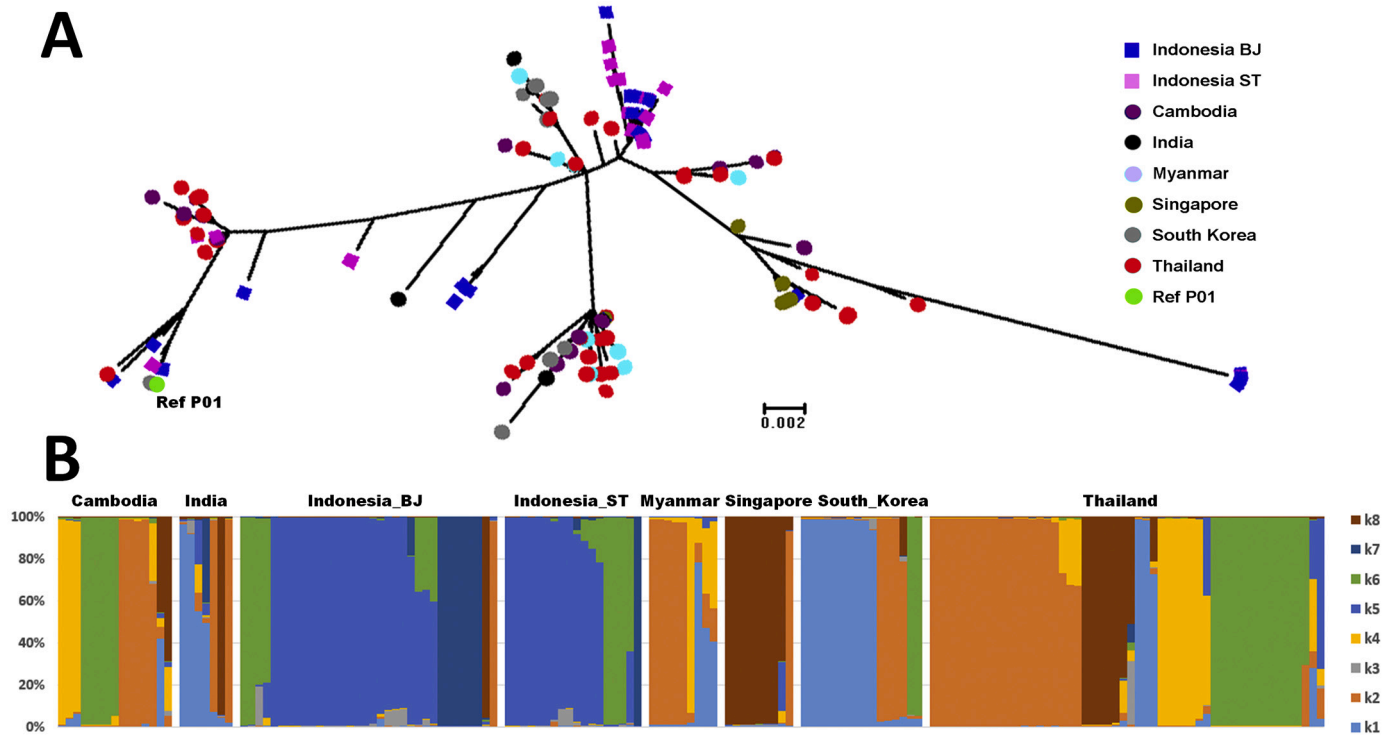

Fig. 5. Phylogenetic tree and structure analysis of PvMSP1-42 in Indonesian *Plasmodium vivax* isolates. A. Maximum likelihood phylograms of PvMSP1-42 genes constructed from the sequences. Samples are colored according to geographic origin. B. Structure analysis of the full set of variation loci from all isolates. Cluster for each isolate was assessed according to an optimized cluster value of K = 8.

Table 2
F_{st} analysis of PvMSP1-42 from Indonesia with that from other Asian countries.

	Cambodia	India	Indonesia	Myanmar	Singapore	South Korea	Thailand
Cambodia	–						
India	0.10707	–					
Indonesia	0.0874	0.0781	–				
Myanmar	0.17494	0.05615	0.19872	–			
Singapore	0.50415	0.47304	0.32791	0.81869	–		
South Korea	0.19799	0.00638	0.21987	0.18938	0.76343	–	
Thailand	0	0.04498	0.09725	0.07944	0.36268	0.13872	–

5. Conclusion

This study is the first to reveal an in-depth understanding of genetic diversity of PvMSP1-42 in Indonesia. Mutations, natural selection, and recombination seem to fuel and sustain evasion of the host immunity.

Funding

This work was supported by the Higher Education Research Excellence Grant from the Indonesian Ministry of Higher Education, Research and Technology (Grant No. LPPM-UGM/1345/LIT/2013), the National Sharing Service Platform for Parasite Resources (Grant No. TDRC-2019-194-30), the Project of Shanghai Science and Technology Commission (Grant No. 18490741100), the National Natural Science Foundation of China (Grant No. 81101266), and the Open Project of Key Laboratory of Parasite and Vector Biology, Ministry of Health (Grant No. WSBKFKT-201605).

Availability of data and materials

All supporting data are contained within the manuscript. New sequences identified in this study are deposited in GenBank with the accession numbers MK733842-MK733893.

Authors' contributions

EEHM and JHC were involved in the conceptualization, design, execution, and first draft of the manuscript. FM and PP collected samples and performed the investigations. EHH contributed to consulting, data analysis, interpretation of the results, and drafting of the manuscript. SBS and HMS contributed to consulting, data analysis, and interpretation of the results. All authors were involved in the preparation of the manuscript. All authors read and approved the final manuscript.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgments

EEHM and EHH received fellowship from the 2016 Open Project Application Form of Key Laboratory of Parasite and Vector Biology of the Chinese Ministry of Health. We thank Charles Ma for his invaluable input to the quality control of sequence readings and proofreading.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2020.104573.

References

Anstey, N.M., Russell, B., Yeo, T.W., Price, R.N., 2009. The pathophysiology of vivax malaria. Trends Parasitol. 25 (5), 220–227.

Barry, A.E.B.J., Reeder, J.C., Fowkes, F.J.I., Arnott, A., 2012. Using population genetics to guide malaria vaccine design. Malaria Parasites 231–264. https://www.intechopen.com/books/malaria-parasites/using-population-genetics-to-guide-malaria-vaccine-design Accessed 28 Oct 2018.

Bastos, M.S., da Silva-Nunes, M., Malafronte, R.S., Hoffmann, E.H., Wunderlich, G., Moraes, S.L., Ferreira, M.U., 2007. Antigenic polymorphism and naturally acquired antibodies to *Plasmodium vivax* merozoite surface protein 1 in rural Amazonians. Clin. Vaccine Immunol. 14 (10), 1249–1259.

Blackman, M.J., Heidrich, H.G., Donachie, S., McBride, J.S., Holder, A.A., 1990. A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies. J. Exp. Med. 172 (1), 379–382.

Chen, J.H., Jung, J.W., Wang, Y., Ha, K.S., Lu, F., Lim, C.S., Takeo, S., Tsuboi, T., Han, E.T., 2010. Immunoproteomics profiling of blood stage *Plasmodium vivax* infection by high-throughput screening assays. J. Proteome Res. 9 (12), 6479–6489.

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

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14 (8), 2611–2620.

Galinski, M.R., Barnwell, J.W., 2008. Plasmodium vivax: who cares? Malar J. 7 (Suppl. 1), S9.

Han, J.H., Cho, J.S., Cheng, Y., Muh, F., Yoo, W.G., Russell, B., Nosten, F., Na, S., Ha, K.S., Park, W.S., et al., 2018. *Plasmodium vivax* Merozoite surface protein 1 paralog as a mediator of parasite adherence to reticulocytes. Infect. Immun. 86 (9) e00239-00218.

Imwong, M., Snounou, G., Pukrittayakamee, S., Tanomsing, N., Kim, J.R., Nandy, A., Guthmann, J.P., Nosten, F., Carlton, J., Looareesuwan, S., et al., 2007. Relapses of *Plasmodium vivax* infection usually result from activation of heterologous hypnozoites. J. Infect. Dis. 195 (7), 927–933.

Johnston, S.P., Pieniazek, N.J., Xayavong, M.V., Slemenda, S.B., Wilkins, P.P., da Silva, A.J., 2006. PCR as a confirmatory technique for laboratory diagnosis of malaria. J. Clin. Microbiol. 44 (3), 1087–1089.

Kementrian Kesehatan Republik Indonesia. Profil Kesehatan Indonesia Tahun 2015. Available online:. *http://wwwdepkesgoid/resources/download/pusdatin/profil-kesehatan-indonesia/profil-kesehatan-Indonesia-2015pdfAccessed 28 Oct 2018*.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25 (11), 1451–1452.

Mau, F.M.E., 2016. Keragaman Genetik dari Msp 1, Msp 2, dan Glurp pada *Plasmodium Falciparum* di Kabupaten Sumba Tengah, Nusa Tenggara Timur. Bul Penelit Kesehat 44, 77–84.

McCaffery, J.N., Fonseca, J.A., Singh, B., Cabrera-Mora, M., Bohannon, C., Jacob, J., Arévalo-Herrera, M., Moreno, A., 2019. A multi-stage *Plasmodium vivax* malaria vaccine candidate able to induce long-lived antibody responses against blood stage parasites and robust transmission-blocking activity. Front. Cell. Infect. Microbiol. 9, 135.

Pitabut, N., Panichakorn, J., Mahakunkijcharoen, Y., Hirunpetcharat, C., Looareesuwan, S., Khusmith, S., 2007. IgG antibody profile to c-terminal region of *Plasmodium vivax* merozoite surface protein-1 in Thai individuals exposed to malaria. Southeast Asian J. Trop. Med. Public. Health. 38 (1), 1–7.

Poespoprodjo, J.R., Fobia, W., Kenangalem, E., Lampah, D.A., Hasanuddin, A., Warikar, N., Sugiarto, P., Tjitra, E., Anstey, N.M., Price, R.N., 2009. Vivax malaria: a major cause of morbidity in early infancy. Clin. Infect. Dis. 48 (12), 1704–1712.

Profil Kesehatan Indonesia Tahun 2013. Available online: . *http://wwwdepkesgoid/resources/download/pusdatin/profil-kesehatan-indonesia/profil-kesehatan-indonesia-2013pdfAccessed28 Oct 2018*.

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

Satyagraha, A.W., Sadhewa, A., Baramuli, V., Elvira, R., Ridenour, C., Elyazar, I., Noviyanti, R., Coutrier, F.N., Harahap, A.R., Baird, J.K., 2015. G6PD deficiency at Sumba in eastern Indonesia is prevalent, diverse and severe: implications for primaquine therapy against relapsing vivax malaria. PLoS Negl. Trop. Dis. 9 (3), 1–17.

Shen, H.M., Chen, S.B., Wang, Y., Xu, B., Abe, E.M., Chen, J.H., 2017. Genome-wide scans for the identification of *Plasmodium vivax* genes under positive selection. Malar. J. 16 (1), 238.

Sitohang, V., Sariwati, E., Fajariyani, S.B., Hwang, D., Kurnia, B., Hapsari, R.K., Laihad, F.J., Sumiwi, M.E., Pronyk, P., Hawley, W.A., 2018. Malaria elimination in Indonesia: halfway there. Lancet Glob. Health 6 (6), e604–e606.

Sulistiyono, S.T.R.Y., 2013. Contest for hegemony : the dynamics of inland and maritime cultures relations in the history of Java island, Indonesia. J. Mar. Isl. Cult 2, 115.

Suphakhonchuwong, N., Chaijaroenkul, W., Rungsihirunrat, K., Na-Bangchang, K., Kuesap, J., 2018. Evaluation of *Plasmodium vivax* isolates in Thailand using polymorphic markers *Plasmodium* merozoite surface protein (*PvMSP*) 1 and *PvMSP3*. Parasitol. Res. 117 (12), 3965–3978.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30 (12), 2725–2729.

Thakur, A., Alam, M.T., Sharma, Y.D., 2008. Genetic diversity in the C-terminal 42 kDa region of merozoite surface protein-1 of *Plasmodium vivax* (PvMSP-1(42)) among Indian isolates. Acta Trop. 108 (1), 58–63.

Valderrama-Aguirre, A., Quintero, G., Gómez, A., Castellanos, A., Pérez, Y., Méndez, F., Arévalo-Herrera, M., Herrera, S., 2005. Antigenicity, immunogenicity, and protective efficacy of *Plasmodium vivax* MSP1 PV200l: a potential malaria vaccine subunit. Am J Trop Med Hyg 73 (5 Suppl), 16–24.

Valderrama-Aguirre, A., Zúñiga-Soto, E., Mariño-Ramírez, L., Moreno, L., Escalante, A.A., Arévalo-Herrera, M., Herrera, S., 2011. Polymorphism of the Pv200L fragment of merozoite surface protein-1 of *Plasmodium vivax* in clinical isolates from the Pacific coast of Colombia. Am J Trop Med Hyg 84 (2 Suppl), 64–70.

Valencia, S.H., Rodríguez, D.C., Acero, D.L., Ocampo, V., Arévalo-Herrera, M., 2011. Platform for Plasmodium vivax vaccine discovery and development. Mem. Inst. Oswaldo Cruz 106 (Suppl. 1), 179–192.

WHO, 2017. World Malaria report. World Health Organization.

Zeyrek, F.Y., Tachibana, S., Yuksel, F., Doni, N., Palacpac, N., Arisue, N., Horii, T., Coban, C., Tanabe, K., 2010. Limited polymorphism of the *Plasmodium vivax* merozoite surface protein 1 gene in isolates from Turkey. Am J Trop Med Hyg 83 (6), 1230–1237.

Zhou, X., Tambo, E., Su, J., Fang, Q., Ruan, W., Chen, J.H., Yin, M.B., Zhou, X.N., 2017. Genetic diversity and natural selection in 42 kDa region of *Plasmodium vivax* Merozoite surface Protein-1 from China-Myanmar endemic border. Korean J. Parasitol. 55 (5), 473–480.