The Geographical Distribution of Allele Polymorphisms of *Plasmodium vivax* in Different Regions of Thailand

Varakorn Kosaisavee DrPH*,**, Ian Hastings PhD***,
Alister Craig PhD***, Usa Lek-Uthai DrPH*

** Department of Parasitology and Entomology, Faculty of Public Health, Mahidol University, Bangkok, Thailand*
*** Doctor of Public Health Programme (Parasitology), Faculty of Graduate Studies, Mahidol University, Nakornpathom, Thailand*
**** Department of Molecular and Biochemical Parasitology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK*

Objective: *To investigate the genetic structure of Plasmodium vivax circumsporozoite surface protein (PvCSP) and P. vivax merozoite surface protein 1 (PvMSP1) genes from field isolates of malaria parasites from four different regions of Thailand.*
Material and Method: *The data was collected by cross-sectional survey, consisting of 273 P. vivax infected blood samples from malaria clinics in 4 different border regions of Thailand from February 2008 to February 2009. The dried blood spots were extracted for DNA and Plasmodium species confirmed by species-specific primer sets. PvCSP and PvMSP1 genes were amplified and their population genetics were analyzed by using the Heterozygosity (H_E) formula, F-STAT and LIAN programs.*
Result: *There was considerable variation in the PvMSP1 gene within 2 fragments for which H_E was 0.8303, whereas PvCSP showed low H_E at 0.1418. Significant differences in allele frequencies between sites were quantified by Fst, Linkage disequilibrium (LD). The results showed PvMSP1 F2; Fst = 0.063, p = 0.07; PvMSP1 F2 RFLP pattern; Fst = 0.154, p = 0.005; PvMSP1 F3; Fst = 0.23, p = 0.005 and the overall loci showed Fst = 0.151, p = 0.005 (Fisher's exact test). All values of Index association (I^S_A) were non-significant. There was no evidence of LD within the P. vivax populations.*
Conclusion: *H_E at each locus of the PvMSP1 gene showed significant differences in allele frequencies between sites.*

Keywords: *Plasmodium vivax, PvMSP1, PvCSP, Heterozygosity, Linkage disequilibrium*

J Med Assoc Thai 2012; 95 (Suppl. 6): S135-S140
Full text. e-Journal: *http://jmat.mat.or.th*

Malaria is a dangerous and very serious health concern in the border areas surrounding Thailand. *Plasmodium vivax* is one of the human malaria parasite species causing benign malaria. After the *P. vivax* genome project was completed in 2008[1], the present study of genetic diversity has been scrutinized in many countries. Data on genetic structure is important because the gene flow between malaria transmission sites and recombination of different allele types potentially have a severe effect upon the development of resistance to drugs and drug sensitivity. The investigation and comparison of parasite genetics between populations can quantify the geographical distribution of genetic polymorphisms, which can be used to monitor existing interventions or provide vaccine formulation information based on malaria endemicity.

The previous study on genetic diversity in *P. vivax* found that *P. vivax* in Asia were had greater diversity than the South-American populations[2], whereas African populations of *P. falciparum* have been recognized for some time as being highly genetically diverse[3]. Genetic diversity of *P. falciparum* has produced population structures and indicated that they show high levels of diversity, but no linkage disequilibrium in two different geographic regions which mean genotype in one region is independent in the another region[4]. In contrast, a study of *P. falciparum* in Brazil described low levels of diversity and significant linkage disequilibrium for three of the five populations[5]. The authors need to have an understanding about the structure of a *P. vivax* population. Once understood, this information assists us in defining the epidemiologic distribution of the parasite strains which will greatly aid in the development of an anti-*P. vivax* vaccine. The variety of

Correspondence to:
Lek-Uthai U, Department of Parasitology and Entomology, Faculty of Public Health, 420/1 Rajvithi Road, Rajthewee, Bangkok 10400, Thailand.
Phone & Fax: 0-2644-5130
E-mail: phulu@mahidol.ac.th

parasitic forms of the *P. vivax* is necessary to formulate an effective vaccine. Stated simply, if these two genes are conserved, or evolve more slowly, this would help to make a case for further study in the development of a vaccine. If, however, they are prone to rapid mutation, vaccine development would not be cost effective. Knowledge of *P. vivax* population structure is needed to define the epidemiologic distribution of *P. vivax* strains and to aid the development of an anti-*P. vivax* vaccine because the presence of variant forms of *P. vivax* parasites has implications for the efficacy of vaccine formulation: if these two genes are seen to be more conserved, this would strengthen the case for further study for vaccine development and treatment in Thailand. The aim of the present study was to determine the genetic structure of *PvCSP* and *PvMSP1* genes from field isolates of malaria parasites from four different regions of Thailand.

Material and Method

Study sites

The data was collected by cross sectional survey from February 2008 to February 2009. The authors collected 273 infected *P. vivax* blood samples from malaria clinics in four different border regions of Thailand which included the Western region [Tak Province] (85 isolates), Eastern region [Chanthaburi Province] (130 isolates), Southern region [Prachuap Kriri Khun Province] (39 isolates) and Northern region [Mae Hong Sorn Province] (19 isolates) (location at each sites are shown in Fig. 1). All malaria patients were symptomatic and the malaria species had been identified by microscopy. Blood samples were obtained by finger prick and blotted onto filter paper (Protein saver card, Whatman Inc). The genomic DNA were confirmed by Multiplex Speciation PCR[6] before PvCSP and PvMSP1 amplification. Mixed species infections were not included in the present study. All dried blood spots were kept in a refridgerator until processed.

Permission for the present study was obtained from the ethical review committee for research in human subjects, Ministry of Public Health, Thailand (Reference no. 101/2550).

Gene amplifications

The amplification methods used in the present study followed the previous study of Imwong M et al and Alves RT et al[7,8] for PvMSP1 and PvCSP genes, respectively. Hotstar Taq polymerase (QIAGEN, Germany) was used for reaction activating. Ten micro-litres of PCR product was digested with 1 U of Alu I

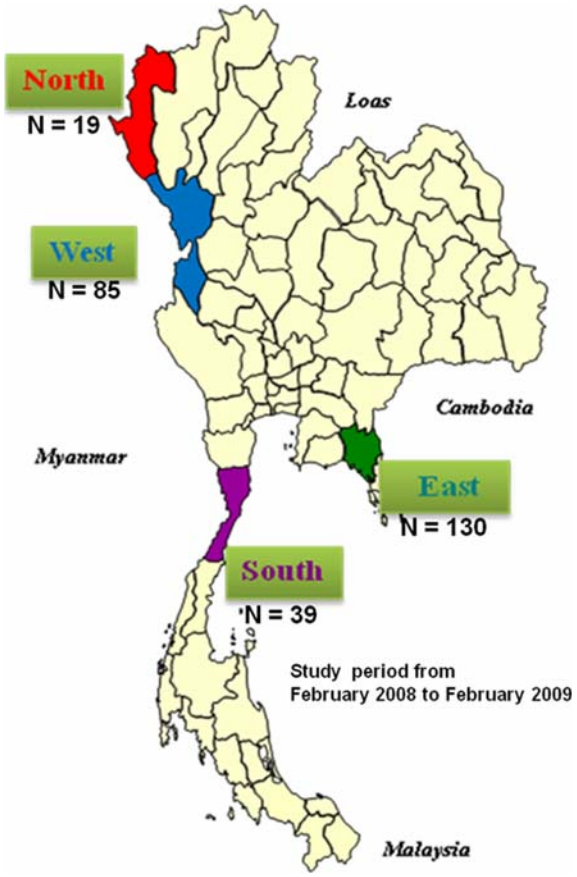

Fig. 1 Four study regions (north = Mae Hong Sorn; n = 19, west = Mae Sod; n = 85, east = Chanthaburi; n = 130, south = Prachup Khiri Khun; n = 39)

(Fermentas, Life Sciences) at 37°C for 1 hour. The DNA fragments obtained following PCR were analyzed by electrophoresis in 1.5% agarose gels. Electrophoresis was performed in a TBE buffer and the DNA was stained with ethidium bromide and then was viewed with an ultraviolet trans-illuminator. The size of the amplified fragments was estimated by comparison to a 100 bp ladder. All mixed genotypes were excluded before PCR-RFLP.

Data analysis

Allele frequencies were estimated from PCR products in each gene. The proportion of single and mixed infections was calculated and confidence intervals were obtained using the modified Wald method available at http://www.graphpad.com/quickcalcs/ConfInterval1.cfm.

The expected Heterozygosity (H_E) was quantified at each locus using the formula $H_E = [n/(n-$

1)][1-$\sum p_i^2$], where p_i are the allele frequencies, n is the number of isolates[9]. H_E is therefore the probability that pair of alleles randomly sampled from a population are different and its value placed between 0 and 1, with values close to 1 indicating high genetic diversity levels in the population.

The genetic differentiation between populations was measured by Fixation index (Fst) obtained from F-STAT version 2.9.3.2 (Feb 2002)[10]. The standardized Fst (Fst = Fst/Fst_{Max}) was calculated, where Fst_{Max} was recoded from Fst input to obtain the proper maximum divergence among populations[11]. Standardized Fst corrects for different levels of genetic diversity in the marker loci thereby allowing valid comparison of results across loci, populations and published studies.

Linkage disequilibrium (LD) was quantified by statistical independence of alleles at all loci (Index association; I^S_A) computed using LIAN ver 3.0 program[12] (also available on http://adenine.biz.fh-weihenstephan.de/cgi-bin/lian/lian.cgi.pl). The variance of allele distance (V_D) was calculated from the results of 10,000 simulated data sets in which alleles at each locus were randomly reshuffled among genotypes. The null hypothesis H_0: $V_D = V_E$ (V_E, the expected variance) is tested by Monte Carlo simulation and by a new parametric test[13].

Results

The genetic diversity pooled for four geographic regions indicated the heterozygosity (H_E) at 0.5012, 0.8303, 0.6459 and 0.1418 for PvMSP1 F2 (size polymorphism), F2 (RFLP pattern), F3 and PvCSP, respectively. The results showed that there was high diversity in the PvMSP1 gene. Comparison of the 2 fragments analyzed by RFLP (Fig. 2) indicated that fragment 2 is slightly more diverse than PvMSP1 F3.

Significant differences in allele frequencies between sites were identified by exact test and quantified by Fst. The results are 0.063 (p = 0.07), 0.154 (p = 0.005), and 0.23 (p = 0.005) for PvMSP1 F2 (size polymorphism), F2 (RFLP polymorphism), and F3, respectively. The overall loci showed significant Fst = 0.151, p = 0.005 so authors carried out pairwise tests between the populations to investigate whether the entire population differed, or whether the result was due to a single, genetically-distinct population. The results (Table 1) indicate that there were significant differences between the Eastern site and the other sites, with the latter being genetically indistinct from one another.

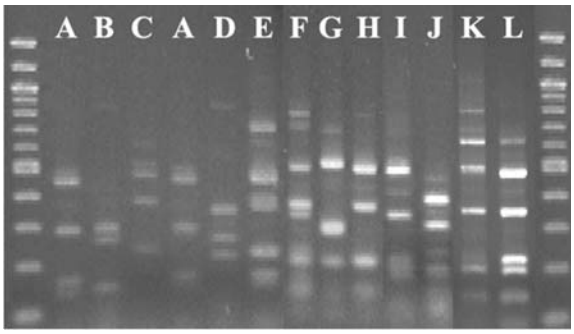

Fig. 2 Gel electrophoresis of PCR-RFLP for PvMSP1 fragment 2

Table 1. Pairwise comparison of Fst between geographic sites using all loci pooled. Fst values are given above the diagonal and standardized Fst below the diagonal

Site	West	East	South	North
West		0.16*	0.006	0.095
East	0.290		0.22*	0.27*
South	0.013	0.383		0.20
North	0.196	0.462	0.379	

* Significant differences at the 0.05 level after Bonferonni correction *i.e.* p-value less than 0.05/6 = 0.00833 (significance is the same for both Fst and standardised Fst values)
West; n = 85, East; n = 130, South; n = 39, North; n = 19
No. of alleles of PvMSP1 F2 and F3 were 12 and 8, respectively

Significant LD was detected if the observed V_D exceeded 95% of the values generated in the reshuffled data sets. This method relies on knowing the 2-locus haplotypes in each malaria clone so only blood samples with a single clone infection were incorporated into this analysis. Pooled samples from West, South and North gave I^S_A = 0.0132 (p = 0.22) and data from the Eastern site alone gave I^S_A = 0.0142 (p = 0.20). All values of I^S_A were non-significant (*i.e.* $p > 0.05$). We therefore concluded that there was no evidence of LD within the *P. vivax* populations. The LD in pooled data was due to the differences between geographic sites (Table 2).

Discussion

This report is one of a series of studies designed to bring our understanding of *P. vivax* genetic structure up to the level of that for *P. falciparum*. The results for PvMSP1 indicated that there was genetic

Table 2. The linkage disequilibrium (LD) was measured within single clone samples in each different site by Index of association (I^S_A)

Study Site	Single Clone	
	n	I^S_A*
West	47	0.0190
East	51	0.0142
South	18	0.0198
North	4	-0.2500

*p-value = 0.05

diversity in Thailand. The previous study has confirmed that parasite population structure can result from the adaptation of malaria parasites to their vectors[14]. The H_E was high in the present study, as found in many countries in Asia[2]. In the present study, the geographic differentiation was low, but it showed significant differences in the eastern site against other sites. This observation was similar to other reports[7,15,16]. The ideal *P. vivax* population's geographic structure study was the microsatellite diversity (tandem repeats) study for large-scale population. However, there have been local-level analyses, like the present study, which reported H_E were 0.7517 in Ethiopia, 0.8450 in Myanmar and 0.8610 in Sri Lanka[17]. Furthermore, pairwise Fst showed only Eastern isolates that were found to be different in parasite population structure. This may be affected by the geographical area, as in the Eastern region, which is close to Cambodia, while the other three sites are close to Myanmar (west).

There was no evidence of LD within the *P. vivax* populations in the present study. The LD in pooled data was due to the differences between geographic sites. Nevertheless, the markers for the evaluation between microsatellite variation and repeat array length could be observed in the substantial variability of *P. vivax* isolates from different regions with high levels of multiple-clones and high LD[18]. The present study of population and transmission dynamics of *P. vivax* in Amazonia stated that small-area genetic diversity under low-level transmission is not severely constrained by the low rates of effect meiotic recombinant, thus providing clear public health implications[19]. The present study among *P. vivax* populations with high genetic diversity have been mapped and revealed that low densities of markers, which were SNPs and LD patterns in South America and Asia, were found to be the feasibility support of genome-wide association finding[20].

These studied genes are recognized by the host's immune system. Analyses of polymorphism and population diversity of *P. vivax* have focused on parasite molecules that are under selection by host immunity, particularly antigens homologous to those of *P. falciparum*, such as MSP, CSP or erythrocyte binding antigens[21]. However, some population parameters are better determined by using DNA markers that are neutral, such as microsatellite loci[22].

Conclusion

There was great diversity in PvMSP1 genes but low diversity was found in PvCSP's. The Eastern isolates have a different genetic structure when compared with the other sites. There was no linkage disequilibrium between the presented genetic markers. The neutral genetic markers, such as a microsatellite marker, should be used for future genetic diversity studies. Little is known about the importance of relapse on *P. vivax* epidemiology and population structure. Further research is needed to better understand the dynamics for before and after treatment.

Acknowledgement

The authors wish to thank staffs of the Malaria Clinics of all study areas, for their valuable time collecting samples. This work was supported by "Strategic Scholarships Fellowships Frontier Research Network" from the Commission on Higher Education (Thailand) (Grant No. CHE-PhD-SW-INV-20060169) and also partially supported for publication by the China Medical Board, Faculty of Public Health, Mahidol University, Bangkok, Thailand.

Potential conflicts of interest

None.

References

1. Carlton JM, Adams JH, Silva JC, Bidwell SL, Lorenzi H, Caler E, et al. Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. Nature 2008; 455: 757-63.
2. Imwong M, Nair S, Pukrittayakamee S, Sudimack D, Williams JT, Mayxay M, et al. Contrasting genetic structure in *Plasmodium vivax* populations from Asia and South America. Int J Parasitol 2007; 37: 1013-22.
3. Jeffares DC, Pain A, Berry A, Cox AV, Stalker J, Ingle CE, et al. Genome variation and evolution of

the malaria parasite *Plasmodium falciparum*. Nat Genet 2007; 39: 120-5.
4. Anthony TG, Conway DJ, Cox-Singh J, Matusop A, Ratnam S, Shamsul S, et al. Fragmented population structure of *plasmodium falciparum* in a region of declining endemicity. J Infect Dis 2005; 191: 1558-64.
5. Machado RL, Povoa MM, Calvosa VS, Ferreira MU, Rossit AR, dos Santos EJ, et al. Genetic structure of *Plasmodium falciparum* populations in the Brazilian Amazon region. J Infect Dis 2004; 190: 1547-55.
6. Padley D, Moody AH, Chiodini PL, Saldanha J. Use of a rapid, single-round, multiplex PCR to detect malarial parasites and identify the species present. Ann Trop Med Parasitol 2003; 97: 131-7.
7. Imwong M, Pukrittayakamee S, Gruner AC, Renia L, Letourneur F, Looareesuwan S, et al. Practical PCR genotyping protocols for *Plasmodium vivax* using Pvcs and Pvmsp1. Malar J 2005; 4: 20.
8. Alves RT, Povoa MM, Goldman IF, Cavasini CE, Rossit AR, Machado RL. A new polymerase chain reaction/restriction fragment length polymorphism protocol for *Plasmodium vivax* circumsporozoite protein genotype (VK210, VK247, and *P. vivax*-like) determination. Diagn Microbiol Infect Dis 2007; 59: 415-9.
9. Anderson TJ, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, et al. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol Biol Evol 2000; 17: 1467-82.
10. Goudet J. Fstat version 1.2: a computer program to calculate Fstatistics. J Hered 1995; 86: 485-6.
11. Meirmans PG. Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution 2006; 60: 2399-402.
12. Haubold B, Hudson RR. LIAN 3.0: detecting linkage disequilibrium in multilocus data. Linkage Analysis. Bioinformatics 2000; 16: 847-8.
13. Haubold B, Travisano M, Rainey PB, Hudson RR. Detecting linkage disequilibrium in bacterial populations. Genetics 1998; 150: 1341-8.
14. Joy DA, Gonzalez-Ceron L, Carlton JM, Gueye A, Fay M, McCutchan TF, et al. Local adaptation and vector-mediated population structure in *Plasmodium vivax* malaria. Mol Biol Evol 2008; 25: 1245-52.
15. Kim JR, Imwong M, Nandy A, Chotivanich K, Nontprasert A, Tonomsing N, et al. Genetic diversity of *Plasmodium vivax* in Kolkata, India. Malar J 2006; 5: 71.
16. Rosenberg R, Wirtz RA, Lanar DE, Sattabongkot J, Hall T, Waters AP, et al. Circumsporozoite protein heterogeneity in the human malaria parasite *Plasmodium vivax*. Science 1989; 245: 973-6.
17. Gunawardena S, Karunaweera ND, Ferreira MU, Phone-Kyaw M, Pollack RJ, Alifrangis M, et al. Geographic structure of *Plasmodium vivax*: microsatellite analysis of parasite populations from Sri Lanka, Myanmar, and Ethiopia. Am J Trop Med Hyg 2010; 82: 235-42.
18. Rezende AM, Tarazona-Santos E, Fontes CJ, Souza JM, Couto AD, Carvalho LH, et al. Microsatellite loci: determining the genetic variability of *Plasmodium vivax*. Trop Med Int Health 2010; 15: 718-26.
19. Ferreira MU, Karunaweera ND, Silva-Nunes M, da Silva NS, Wirth DF, Hartl DL. Population structure and transmission dynamics of *Plasmodium vivax* in rural Amazonia. J Infect Dis 2007; 195: 1218-26.
20. Orjuela-Sanchez P, Karunaweera ND, Silva-Nunes M, da Silva NS, Scopel KK, Goncalves RM, et al. Single-nucleotide polymorphism, linkage disequilibrium and geographic structure in the malaria parasite *Plasmodium vivax*: prospects for genome-wide association studies. BMC Genet 2010; 11: 65.
21. Cui L, Mascorro CN, Fan Q, Rzomp KA, Khuntirat B, Zhou G, et al. Genetic diversity and multiple infections of *Plasmodium vivax* malaria in Western Thailand. Am J Trop Med Hyg 2003; 68: 613-9.
22. Russell B, Suwanarusk R, Lek-Uthai U. *Plasmodium vivax* genetic diversity: microsatellite length matters. Trends Parasitol 2006; 22: 399-401.

การกระจายทางภูมิศาสตร์ของการเปลี่ยนแปลงอัลลีลของเชื้อพลาสโมเดียมไวแวกซ์ในพื้นที่ประเทศไทย

วรากร โกศัยเสวี, Ian Hastings, Alister Craig, อุษา เล็กอุทัย

***วัตถุประสงค์**: เพื่อตรวจสอบโครงสร้างทางพันธุกรรมของจีน PvCSP และ PvMSP1 ในการติดเชื้อมาลาเรียชนิดไวแวกซ์ในพื้นที่ 4 จังหวัดในประเทศไทย*

***วัสดุและวิธีการ**: เก็บตัวอย่างเลือดผู้ป่วยมาลาเรียชนิดไวแวกซ์ 273 ตัวอย่างลงบนกระดาษกรองตั้งแต่เดือนกุมภาพันธ์ พ.ศ. 2551 ถึงเดือนกุมภาพันธ์ พ.ศ.2552 ทำการสกัดดีเอ็นเอของตัวอย่างเลือดเพื่อนำไปตรวจยืนยันการติดเชื้อชนิดไวแวกซ์เท่านั้นโดยวิธี PCR หลังจากนั้นทำการศึกษาความหลากหลายทางพันธุกรรมของจีน PvCSP และ PvMSP1 โดยวิธี PCR – RFLP และวิเคราะห์โครงสร้างทางพันธุกรรมประชากรของเชื้อโดยใช้โปรแกรม heterozygosity (H_E), F-STAT (Fst) และ LIAN*

***ผลการศึกษา**: พบจีน Pvmsp1 F2 มีความหลากหลายมากเมื่อเทียบกับอีกสองส่วน (F2, F3) โดยค่า H_E = 0.8303 แต่พบค่า H_E ต่ำในจีน PvCSP (0.1418) ส่วนการศึกษาโครงสร้างทางพันธุกรรมประชากร ของเชื้อไวแวกซ์พบค่า Fst ต่ำ แต่พบความแตกต่างอย่างมีนัยสำคัญ ระหว่างตัวอย่างจากทางตะวันออกและพื้นที่อื่นๆ (Fst = 0.16-0.27) ไม่พบหลักฐานของ linkage disequilibrium ภายในประชากรของเชื้อไวแวกซ์*

***สรุป**: ค่า H_E ที่พบมีความแตกต่างระหว่างพื้นที่ที่ศึกษาสามารถสรุปได้ว่า จีน PVMSP1 มีความหลากหลายทางพันธุกรรมมากแต่จีน PvCSP มีความหลากหลายทางพันธุกรรมน้อย*