# Plasmodium vivax rhomboid-like protease 1 gene diversity in Thailand

Touchchapol Mataradchakul [a, f], Chairat Uthaipibull [b], Francois Nosten [c, d],
Joel Vega-Rodriguez [e], Marcelo Jacobs-Lorena [e], Usa Lek-Uthai [f, *]

[a] Doctor of Public Health Programme (Parasitology), Faculty of Graduate Studies, Mahidol University, Nakornpathom, Thailand
[b] Protein-Ligand Engineering and Molecular Biology Laboratory, National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Thailand Science Park, Pathum Thani, Thailand
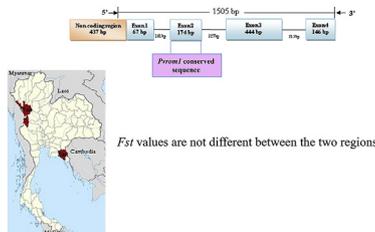[c] Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand
[d] Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine Research Building, University of Oxford Old Road Campus, Oxford, UK
[e] Malaria Research Institute and Department of Molecular Microbiology and Immunology, Johns Hopkins School of Public Health, Baltimore, MD, United States
[f] Department of Parasitology and Entomology, Faculty of Public Health, Mahidol University, Bangkok, Thailand

## GRAPHICAL ABSTRACT



Fst values are not different between the two regions

## ARTICLE INFO

## ABSTRACT

Background: Plasmodium vivax infection remains a major public health problem, especially along the Thailand border regions. We examined the genetic diversity of this parasite by analyzing single-nucleotide polymorphisms (SNPs) of the P. vivax rhomboid-like protease 1 gene (Pvrom1) in parasites collected from western (Tak province, Thai-Myanmar border) and eastern (Chanthaburi province, Thai-Cambodia border) regions.

Methods: Data were collected by a cross-sectional survey, consisting of 47 and 45 P. vivax-infected filter paper-spotted blood samples from the western and eastern regions of Thailand, respectively during September 2013 to May 2014. Extracted DNA was examined for presence of P. vivax using Plasmodium species-specific nested PCR. Pvrom1 gene was PCR amplified, sequenced and the SNP diversity was analyzed using F-STAT, DnaSP, MEGA and LIAN programs.

Results: Comparison of sequences of the 92 Pvrom1 831-base open reading frames with that of a reference sequence (GenBank acc. no. XM001615211) revealed 17 samples with a total of 8 polymorphic sites, consisting of singleton (exon 3, nt 645) and parsimony informative (exon 1, nt 22 and 39; exon 3, nt 336, 537 and 656; and exon 4, nt 719 and 748) sites, which resulted in six different deduced Pvrom1 variants. Non-synonymous to synonymous substitutions ratio estimated by the DnaSP program was 1.65 indicating positive selection, but the Z-tests of selection showed no significant deviations from neutrality for Pvrom1 samples from western region of Thailand. In addition McDonald Kreitman test (MK) showed not significant, and Fst values are not different between the two regions and the regions combined. Interestingly, only Pvrom1 exon 2 was the most conserved sequences among the four exons.

---

* Corresponding author. Department of Parasitology and Entomology, Faculty of Public Health, Mahidol University, 420/1 Ratchawithi Road, Bangkok, Thailand.
  E-mail address: usa.lek@mahidol.ac.th (U. Lek-Uthai).

*Conclusions:* The relatively high degree of *Pvrom1* polymorphism suggests that the protein is important for parasite survival in face of changes in both insect vector and human populations. These polymorphisms could serve as a sensitive marker for studying plasmodial genetic diversity. The significance of *Pvrom1* conserved exon 2 sequence remains to be investigated.

## 1. Introduction

Malaria remains a serious public health problem along the Thai-Cambodia border region to the east and Thai-Myanmar border region to the west of Thailand, with *Plasmodium vivax* causing significant morbidity. The increasing incidence of *vivax* malaria and emergence of drug-resistant strains of *P. vivax* are causes of major concerns for malaria control in the country as well as in other malaria endemic areas worldwide (Mendis et al., 2001). Employment of an effective malaria vaccine is hampered by an incomplete understanding of malaria parasite genetic variations, gene flow and allelic recombination in a given parasite population. Much more progresses have been made in this regard with *P. vivax* on account of it's more severe pathology and fatal disease. (Gupta et al., 2016; Barber et al., 2015; Rahimi et al., 2014). However, since the completion of *P. vivax* genome project in 2008 (Carlton et al., 2008), It is now recognized that there is genetic diversity in *P. vivax* found in many regions of Asia, South America and Oceania, whereas African populations of *P. falciparum* are highly genetically diverse (Winter et al., 2015). Interestingly, the previous reports of genomic diversity in *P. vivax* were more diverse than *P. falciparum* species (Neafsey et al., 2012; Jennison et al., 2015). Similar trend can be found in Hupalo et al. (2016) that the diversity was higher in *P. vivax* more than *P. falciparum* species All of sampled *P. vivax* populations were significantly more diverse than *P. falciparum* ($P < 0.001$, Wilcoxon signed-rank test) especially in Papua New Guinea, Myanmar and Thailand. Several genetic diversity of *P. vivax* have since been studied, such as apical membrane antigen 1 (AMA1), merozoite surface protein (MSP), circumsporozoite protein (CSP), and Duffy binding protein (DBP) (Machado et al., 2004; Padley et al., 2003), which the most diverse genes were MSP3 (Rice et al., 2014), CSP ($H_E = 0.846$) and MSP1 ($H_E = 0.709$) (Maneerattanasak et al., 2016), genes important for immune evasion.

Rhomboids (ROMs) constitute a relatively recently discovered class of serine proteases, which have the unusual ability to cleave substrates within their transmembrane domains (Dowse et al., 2005). In all likelihood, plasmodial ROMs play vital roles during the entire parasite life cycle, in host liver and red blood cells and in mosquito vector (Lin et al., 2013). Parasite plasma membrane ROMs have been implicated in red blood cell invasion by *P. berghei*, *P. yoelii* and *P. falciparum* merozoites (Srinivasan et al., 2009; Vera et al., 2011; Singh et al., 2007).

The *P. falciparum* genome encodes nine distinct *rom* genes, including a *rom1* (Baker et al., 2006), however, no reports are available for the *Pvrom1* gene. Therefore, a study of *Pvrom1* gene is important to assist the malaria control in Thailand (Santos et al., 2012). This study investigated single-nucleotide polymorphisms (SNPs) in *Pvrom1* from field isolates obtained from two different malaria endemic border regions of Thailand. Comparison of *Pvrom1* SNPs between these two parasite populations may provide additional information on *P. vivax* genetic diversity in two widely separated *vivax* malaria endemic areas, which may have a bearing in future *vivax* malaria vaccine development in Thailand and possibly in neighboring countries.

## 2. Material and methods

### 2.1. P. vivax samples

A total of 92 finger prick *P. vivax*-infected blood samples (47 from Tak province, Thai-Myanmar border and 45 from Chanthaburi province, Thai-Cambodia border) (Fig. 1) collected from September 2013 to May 2014 were identified by light microscopy and individually spotted on filter papers (Whatman, USA). Dried blood spots were kept in a refrigerator until processed. Presence of only *P. vivax* in the collected blood samples was subsequently verified by means of *Plasmodium* species-specific multiplex PCR (Padley et al., 2003). Permission for the present study was obtained from the Ethical Review Committee for Research in Human Subjects, Ministry of Public Health, Thailand (permission no. 15/2556). Written prior consents were obtained.



**Fig. 1. *P. vivax* collecting regions in Thailand.** West, Tak province; East, Chanthaburi province.
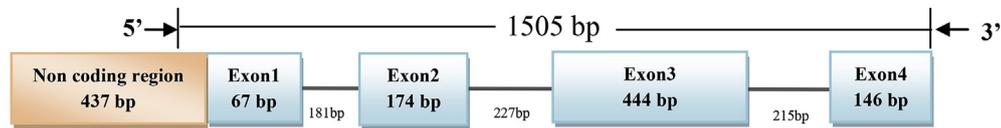
**Fig. 2. Schematic diagram of *Pvrom1* gene.** Arrows indicate locations of PCR primers.

## 2.2. PCR amplification and sequencing of Pvrom1

DNA was extracted from dried blood samples using a DNA mini blood kit (QIAGEN, Germany) according to the manufacturer's instructions. Primers employed were *Pvrom1* F (5′-CCCCAAAT-TAACGTCCAAGTTCCAAG-3′) and *Pvrom1* R (5′-CGCAAGTGATTCA AATCGACGGAAC-3′) located in exon 1 and exon 4 respectively (Fig. 2) (GenBank acc. no. NC_009914). PCR was conducted in a 50 μL mixture containing 10 mM Tris-HCl pH 7.4, 1.5 mM MgCl₂, 0.2 μM each primer, 200 μM each of the four dNTPs, 2 μL of DNA and 0.02 U/μL *Taq* DNA polymerase (New England BioLabs, USA). Thermal cycling conditions were as follows: 95 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 52 °C for 1 min and 68 °C for 90 s, and a final heating step of 5 min at 68 °C. Amplicons (1505 bp) were analyzed by 1.5% agarose gel electrophoresis and stained with Gel red (Biotium, USA). Amplicons were purified using QIAquick PCR purification kit (QIAGEN, Germany) and sequenced in both directions using primers 5′-AATTAAGAATGGGATTTACATTGGAG-3′ and 5′-GTTTCCGTA TATGCCTGTCAAG-3′. DNA sequencing experiment was performed using ABI's 3500XL Genetic Analyzer (Ramathibodi Hospital, Thailand). *Pvrom1* sequence was deposited as GenBank acc. no. XM_001615211.

## 2.3. Genbank accession numbers

The *Pvrom 1* gene sequences of the 92 *P. vivax* isolates determined in this study were deposited in the GeneBank database under accession numbers KX422390 to KX422481.

## 2.4. Statistical analysis

Percent similarity and rate of nonsynonymous and synonymous substitutions were calculated using BLASTN (NCBI, Bethesda, MD) and DnaSP ver. 5.10.01, respectively (Librado and Rozas., 2009).

*Pvrom1* sequences were analyzed using Bioedit ver. 7.0.5.3. The haplotype diversity (Hd), number of haplotypes (H), number of segregating sites (S), nucleotide diversity (π), average number of pairwise nucleotide differences within the population (K), linkage disequilibrium (LD), recombination parameters (R and Rm) were calculated using the DnaSP package ver. 5.10.01. In addition, the positively selected genes ($d_N/d_S$) were computed using the method of Kimura M (Kimura, 1977), Goldman N and Yang Z (Goldman and Yang., 1994). The rates of synonymous (dS) and non-synonymous (dN) mutations were computed using MEGA version 6.0 (Tamura et al., 2013) with the method of Nei and Gojobori (1986) with the Jukes and Cantor (JC) correction (Jukes and Cantor, 1969). The standard errors of dS and dN estimates were calculated with 1000 bootstrap replicates and the significance of the difference dS−dN was tested using a two-tailed Z test (Nei and Kumar 2000). The null hypothesis is that Ds = Dn; thus we assumed as null hypothesis that the observed polymorphism was neutral. The neutrality test Tajima's *D* (Tajima, 1989), Fu and Li's *F* and Fu and Li's *D* (Fu and Li, 1993) were calculated using the DnaSP package ver. 5.10.01. The McDonald Kreitman test (MK) (McDonald and kreitman, 1991) was applied by using a single *P. cynomolgi rom-like 1* sequence (PlasmoDB_092040) (Tachibana et al., 2012); as the outgroup for comparison with the *P. vivax* populations. Genetic differentiation

**Table 2**
Neutrality test of *Pvrom1* obtained from two regions of Thailand.

| Study site | No | S | Tajima's *D* | Fu and Li's *F* | Fu and Li's *D* | MK |
|---|---|---|---|---|---|---|
| West | 47 | 6 | −1.56349 | −0.30990 | 0.32972 | 0.674 |
| East | 45 | 6 | −1.55986 | −1.65426 | −1.34580 | 0.375 |
| West + East | 92 | 8 | −1.61027 | −0.28173 | 0.46820 | 0.703 |

No, number of sequences; S, number of segregating sites. MK, McDonald and Kreitman test. *P-value < 0.05, ** P-value < 0.001 were considered significant; The *P. cynomolgi* sequence (PlasmoDB_092040) was used as an out group for McDonald and Kreitman test.

**Table 1**
Genetic diversity of *Pvrom1* obtained from two regions of Thailand.

| Study site | No | S | Eta | H | Hd | K | π | Rm (R) | $d_N$ | $d_S$ | $d_N/d_S$ | p (Z-stat) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| West | 47 | 6 | 6 | 8 | 0.314 | 0.53469 | 0.00064 | 2 | 0.00071 | 0.00043 | 1.65 | 0.599 (0.528) |
| Exon 1 | 47 | 2 | 2 | 3 | 0.084 | 0.126 | 0.00188 | (<0.0001) | 0.0025 | 0 | | 0.194 (1.306) |
| Exon 2 | 47 | 0 | 0 | 1 | 0 | 0 | 0 | | 0 | 0 | | 1.000 (0.000) |
| Exon 3 | 47 | 2 | 2 | 3 | 0.163 | 0.167 | 0.00038 | | 0.00048 | 0 | | 0.157 (1.426) |
| Exon 4 | 47 | 2 | 2 | 4 | 0.201 | 0.242 | 0.00166 | | 0.00152 | 0.00219 | | 0.830 (−0.215) |
| | | | | | | | | | | | | |
| East | 45 | 6 | 6 | 8 | 0.395 | 0.54949 | 0.00066 | 1 | 0.00053 | 0.00112 | 0.47 | 0.464 (−0.734) |
| Exon 1 | 45 | 1 | 1 | 2 | 0.087 | 0.087 | 0.00130 | (0.0089) | 0.00171 | 0 | | 0.323 (0.993) |
| Exon 2 | 45 | 0 | 0 | 1 | 0 | 0 | 0 | | 0 | 0 | | 1.000 (0.000) |
| Exon 3 | 45 | 3 | 3 | 4 | 0.171 | 0.216 | 0.00049 | | 0.00062 | 0 | | 0.123 (1.555) |
| Exon 4 | 45 | 2 | 2 | 3 | 0.206 | 0.246 | 0.00169 | | 0.00193 | 0.00116 | | 0.719 (0.360) |
| | | | | | | | | | | | | |
| West + East | 92 | 8 | 8 | 12 | 0.352 | 0.5413 | 0.00065 | 3 | 0.00062 | 0.00078 | 0.79 | 0.787 (−0.271) |
| Exon 1 | 92 | 2 | 2 | 4 | 0.085 | 0.107 | 0.00159 | (<0.0001) | 0.00211 | 0 | | 0.152 (1.442) |
| Exon 2 | 92 | 0 | 0 | 1 | 0 | 0 | 0 | | 0 | 0 | | 1.000 (0.000) |
| Exon 3 | 92 | 4 | 4 | 5 | 0.166 | 0.192 | 0.00043 | | 0.00055 | 0 | | 0.056 (1.933) |
| Exon 4 | 92 | 2 | 2 | 4 | 0.201 | 0.242 | 0.00166 | | 0.00171 | 0.00168 | | 0.958 (0.053) |

No, number of sequences; S, number of segregating sites; Eta, total number of mutations, H, number of haplotypes; Hd, haplotype diversity; K, average number of pair-wise nucleotide differences, π, nucleotide diversity; Rm, minimum number of recombination events; R, estimate of recombination, between adjacent sites; $d_N$, number of non-synonymous substitutions per non-synonymous sites, $d_S$, number of synonymous substitutions per synonymous sites. Z-stat, dn-ds is significant at *p < 0.05.

**Table 3**
Pairwise *Fst* estimates for *Pvrom1* obtained from two regions of Thailand.

| Site | (West) | (East) |
|------|--------|--------|
| West | | 0.54133 |
| East | 0 | |

*Fst* values are shown in the lower left quadrant and average numbers of pair-wise nucleotide differences between populations are shown in the upper right quadrant.

between populations was measured using a Fixation index (*Fst*) obtained from F-STAT ver. 2.9.3.2 (Feb 2002) (Goudet, 1995) *Fst* value at each locus of 0, >0−0.05, >0.05−0.15, and >0.15−0.25 is considered as no differentiation, low genetic differentiation, moderate differentiation, and great differentiation, respectively (Meirmans, 2006).

## 3. Results

The *Pvrom1* gene was successfully PCR amplified from the extracted DNA of all 92 samples *P. vivax*-infected blood samples and a 1505-bp band was visualized on agarose gel (supplement 1).

Nucleotide sequences in the four *Pvrom1* exons (831 nt) from the 92 *P. vivax* samples [47 from Thai-Myanmar border (West) and 45 from Thai-Cambodia border (East)] were compared with that of the GenBank reference (accession no. XM001615211). SNPs were present in 8 West and 9 East samples adding to a total of 8 polymorphic sites, consisting of singleton (at exon 3, nt 645) and parsimony informative (at exon 1, nt 22 and 39; exon 3, nt 336, 537 and 656; and exon 4, nt 719 and 748) sites. These have resulted in changes in the deduced amino acid sequences of 7 West *Pvrom1* samples (M240R (3 samples), H219R, A8T/D13E, M240R/M250L, and A8T/H219R/M250L) and 7 East samples (D13E (2 samples), M240R (4 samples) and M240R/M250L) as shown in supplement 2.

Nucleotide diversity ($\pi$) was 0.00064, 0.00066 and 0.00065 for West, East and West + East samples, respectively. Nonsynonymous to synonymous substitution (dN/dS) ratio was 1.65 ($p = 0.599$), 0.47 ($p = 0.464$) and 0.79 ($p = 0.787$) for West, East and West + East samples, respectively, the p-value for the Z-tests were no significant deviations from neutrality in all samples (Table 1). Furthermore, Tajima's *D* as well as Fu and Li's *F* and *D* statistics and McDonald Kreitman test (MK) were not significant (Table 2). *Fst*
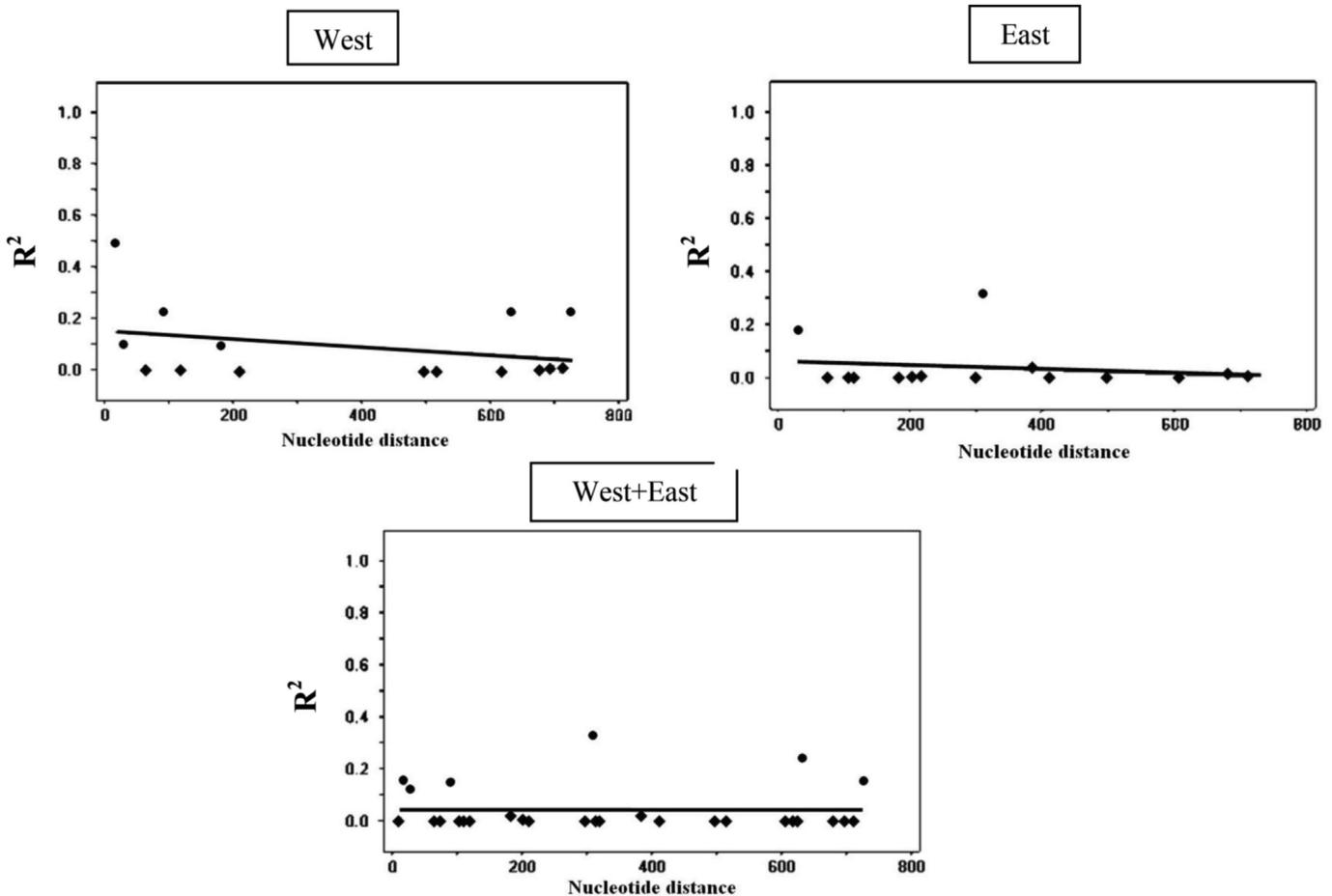


**Fig. 3. Linkage disequilibrium index ($R^2$) across *Pvrom1* open reading frame.** Analysis was performed on all polymorphic sites (filled diamonds and circles) in the 831-bp *Pvrom1* open reading frame using *P. vivax* Sal-1 rhomboid-like protease 1 partial mRNA sequence (GenBank accession no.XM_001615211) as reference. Nucleotide that shows statistically significant $R^2$ value (Chi-square test, *P*-value < 0 0.05) is depicted as filled circle.

values showed no significant difference between West and East *P. vivax* samples (Table 3).

The minimum number of recombination events (Rm) in the *Pvrom1* samples was 3 between adjacent polymorphic sites, while the values of recombination parameter (R) were <0.0001 between adjacent sites. In addition, Rm value was 2 and 1 for West and East samples respectively (Table 1). Based on LD analysis, 6/28, 6/15 and 2/15 pairwise comparisons have significant $R^2$ values over the entire *Pvrom1* open reading frame for West + East, West and East samples, respectively (Fig. 3).

## 4. Discussion

The genetic diversity of *Plasmodium vivax* has been investigated in several malaria-endemic areas, including Thailand, where this is currently the most prevalent species causing malaria in humans. Genetic diversity of *P. vivax* population in Thailand has previously been studied in circumsporozoite protein (CSP), Merozoite surface protein 1 (MSP1), Merozoite surface protein 3 alpha (MSP3$\alpha$), Duffy blinding Protein II (DBP$_{II}$) and Apical membrane antigen 1 (AMA1).

High degree of genetic diversity was observed in *PvCSP* (H$_E$ = 0.846), *PvMSP1* (H$_E$ = 0.709) along the Thai-Myanmar border of Thailand (Western regions) but not *MSP3$\alpha$* (Maneerattanasak et al., 2016) (Rice et al., 2014). Additionally, the nucleotide diversity of *PvMSP3$\alpha$* BlockII, a potential vaccine candidate region, was 0.019 ($\theta\pi$ = 0.019) worldwide samples, but nucleotide diversity was relatively low in Thailand ($\theta\pi$ = 0.017) as well as Suan Oi in Tak Province, Thai-Myanmar border ($\theta\pi$ = 0.013) (Gupta et al., 2015) that are comparable to previous report (Maneerattanasak et al., 2016). Compare to the *PvDBP$_{II}$*, a potential vaccine candidate region, worldwide nucleotide diversity was 0.0101 ($\theta\pi$ = 0.0101) but it was relatively high in Thailand ($\theta\pi$ = 0.0109) (Nóbrega de Sousa et al., 2011). Moreover, the evidence of high levels of nucleotide diversity was also seen in *PvAMA1* in Thai populations ($\pi$ = 0.009) (Zhu et al., 2016). The observations of high genetic diversity were found in Thailand, especially the Thai-Myanmar border (Western regions). Our results is the first report of use of *Pvrom1* gene in this type of investigation reveal that the nucleotide diversity of *Pvrom1* between parasites obtained from the western and eastern malaria endemic regions of Thailand were 0.00064 and 0.00066, respectively. Comparison of SNPs in four exons of *Pvrom1* between western and eastern regions of Thailand highlighted the *Pvrom1* exon 2 was the most conserved sequences among the four exons. On the contrary, The *Pvrom1* exon 1, 3 and 4 were non-conserved sequences of western and eastern region of Thailand. The dN/dS ratio of *Pvrom1* for western region of Thailand was 1.65 indicating positive selection, and the ratio of *Pvrom1* for eastern region of Thailand was 0.47 indicating purifying selection, but the Z-tests of selection showed no significant deviations from neutrality. The same conclusion was reached when both *P. vivax* populations were analyzed as a single population. The values of Rm and decline in LD index ($R^2$) with increasing distance between pairs of nucleotide sites across the entire *Pvrom1* 831-base open reading frame indicate that intragenic recombination may have occurred among *P. vivax* parasites, particularly those in western Thailand.

*Pvrom1* exon 2 sequence was the most conserved among the four exons. *Pfrom1* also contains four exons (Singh et al., 2007) and the deduced amino acid sequence shares 85.14% similarity with that of PvROM1 as shown in supplement 3. Until *Pvrom1* is cloned and expressed allowing PvROM1 structure, function and localization in the parasite to be elucidated, the properties of this enzyme and its six variants remain in the realm of speculation.

The major limitation of this study is the limited number of *P. vivax* samples that were obtained from the study sites. Due to the long distances from the laboratory, parasite DNA needed to be extracted from dried blood samples stored on filter paper. Given the low parasitemias of *P. vivax* infection, in order to ensure availability of sufficient parasite DNA, only blood samples with parasitemias >2% were selected.

In conclusion, this study adds plasmodial *rom*s to the list of polymorphic genes that can provide information on evolution of malaria parasite population structure, relative to changes in both insect vector and human host populations.

## Appendix A. Supplementary data

Supplementary data related to this chapter can be found at http://dx.doi.org/10.1016/j.exppara.2017.06.007.

## References

Baker, R., Wijetilaka, R., Urban, S., 2006. Two *Plasmodium* rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. PLoS Pathog. 2, 922–932.

Barber, B.E., William, T., Grigg, M.J., Parameswaran, U., Piera, K.A., Price, R.N., et al., 2015. Parasite biomass-related inflammation, endothelial activation, microvascular dysfunction and disease severity in *vivax* malaria. PLoS Pathog. 11, e1004558.

Carlton, J.M., Adams, J.H., Silva, J.C., Bidwell, S.L., Lorenzi, H., Caler, E., et al., 2008. Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. Nature 455, 757–763.

Dowse, T.J., Pascall John, C., Brown Kenneth, D., Soldati, D., 2005. Apicomplexan rhomboids have a potential role in microneme protein cleavage during host cell invasion. Int. J. Parasitol. 35, 747–756.

Fu, Y.X., Li, W.H., 1993. Statistical tests of neutrality of mutations. Genetics 133, 693–709.

Goldman, N., Yang, Z., 1994. A codon-based model of nucleotide substitution for protein- coding DNA sequences. Mol. Biol. Evol. 11, 725–736.

Goudet, J., 1995. F-stat version 1.2: a computer program to calculate F-statistics. J. Hered. 86, 485–486.

Gupta, B., Reddy, B.P., Fan, Q., Yan, G., Sirichaisinthop, J., Sattabongkot, J., et al., 2015. Molecular evolution of PvMSP3alpha block II in *Plasmodium vivax* from diverse geographic origins. PLoS One 10, e0135396.

Gupta, P., Sharma, R., Chandra, J., Kumar, V., Singh, R., Pande, V., et al., 2016. Clinical manifestations and molecular mechanisms in the changing paradigm of *vivax* malaria in India. Infect. Genet. Evol. 39, 317–324.

Hupalo, D.N., Luo, Z., Melnikov, A., Sutton, P.L., Rogov, P., Escalante, A., et al., 2016. Populationgenomics studies identify signatures of global dispersal and drug resistance in *Plasmodium vivax*. Nat. Genet. 48, 953–958.

Jennison, C., Arnott, A., Tessier, N., Tavul, L., Koepfli, C., Felger, I., et al., 2015. Plasmodium vivax populations are more genetically diverse and less structured than sympatric Plasmodium falciparum populations. PLoS Negl. Trop. Dis. 9, e0003634.

Jukes, T.H., Cantor, R.C., 1969. Evolution of protein molecules. In: Munro, H.N. (Ed.), Mammalian Protein Metabolism. Academic Press, New York, pp. 21–132.

Kimura, M., 1977. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature 267, 275–276.

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

Lin, J.W., Meireles, P., Prudencio, M., Engelmann, S., Annoura, T., Sajid, M., et al., 2013. Loss-of- function analyses defines vital and redundant functions of the *Plasmodium* rhomboid Protease family. Mol. Microbiol. 88, 318–338.

Maneerattanasak, S., Gosi, P., Krudsood, S., Tongshoob, J., Lanteri, C.A., Snounou, G., et al., 2016. Genetic diversity among *Plasmodium vivax* isolates along the Thai-Myanmar border of Thailand. Malar. J. 15, 75.

Machado, R.L., Povoa, M.M., Calvosa, V.S., Ferreira, M.U., Rossit, A.R., dos Santos, E.J., et al., 2004. Genetic structure of *Plasmodium falciparum* populations in the Brazilian Amazon region. J. Infect. Dis. 190, 1547–1555.

McDonald, J.H., Kreitman, M., 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. Nature 351, 652–654.

Meirmans, P.G., 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution 60, 2399–2402.

Mendis, K., Sina, B.J., Marchesini, P., Carter, R., 2001. The neglected burden of

Plasmodium vivax malaria. Am. J. Trop. Med. Hyg. 64, 97–106.

Neafsey, D.E., et al., 2012. The malaria parasite *Plasmodium vivax* exhibits greater genetic diversity than *Plasmodium falciparum*. Nat. Genet. 44, 1046–1050.

Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitution. Mol. Biol. Evol. 3, 418–426.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Nobrega de Sousa, T., Carvalho, L.H., Alves de Brito, C.F., 2011. Worldwide genetic variability of the Duffy binding protein: insights into *Plasmodium vivax* vaccine development. PLoS One 6, e22944.

Padley, D., Moody, A.H., Chiodini, P.L., Saldanha, J., 2003. Use of a rapid, single-round, multiplex PCR to detect malarial parasites and identify the species present. Ann. Trop. Med. Parasitol. 97, 131–137.

Rahimi, B.A., Thakkinstian, A., White, N.J., Sirivichayakul, C., Dondrop, A.M., Chokejindachai, W., 2014. Severe vivax malaria: a systematic review and meta-analysis of clinical studies since 1900. Malar. J. 13, 481.

Rice, B.L., Acosta, M.M., Pacheco, M.A., Carlton, J.M., Barnwell, J.W., Escalante, A.A., 2014. The origin and diversification of the merozoite surface protein 3 (msp3) multi-gene family in *Plasmodium vivax* and related parasites. Mol. Phylogenet Evol. 78, 172–184.

Santos, J.M., Graindorge, A., Soldati-Favre, D., 2012. New insights into parasite rhomboid proteases. Mol. Biochem. Parasitol. 182, 27–36.

Singh, S., Plassmeyer, M., Gaur, D., Miller, L., 2007. Mononeme: a new secretory organelle in P. falciparum merozoites identified by localization of rhomboid-1 protease. Proc. Natl. Acad. Sci. U. S. A. 104, 20043–20048.

Srinivasan, P., Coppens, I., Marcelo, J., 2009. Distinct roles of *Plasmodium* rhomboid 1 in parasite development and malaria pathogenesis. PLoS Pathog. 5, 1–10.

Tachibana, S., Sullivan, S.A., Kawai, S., Nakamura, S., Kim, H.R., Goto, N., et al., 2012. *Plasmodium cynomolgi* genome sequences provide insight into Plasmodium vivax and the monkey malaria clade. Nat. Genet. 44, 1051–1055.

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

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

Vera, I.M., Beatty, W.L., Sinnis, P., Kim, K., 2011. *Plasmodium* protease ROM1 important for proper formation of the parasitophorous vacuole. PLoS Pathog. 7, 1–18.

Winter, D.J., Pacheco, M.A., Vallejo, A.F., Schwartz, R.S., Arevalo-Herrera, M., Herrera, S., et al., 2015. Whole genome sequencing of field isolates reveals extensive genetic diversity in *Plasmodium vivax* from Colombia. PLoS Negl. Trop. 9, e0004252.

Zhu, X., Zhao, P., Wang, S., Liu, F., Liu, J., Wang, J., et al., 2016. Analysis of *Pvama1* genes from China-Myanmar border reveals little regional genetic differentiation of *Plasmodium vivax* populations. Parasite Vectors 9, 614.