Spatial Variation in Genetic Diversity and Natural Selection on the Thrombospondin-Related Adhesive Protein Locus of Plasmodium vivax (PvTRAP)

Rattiporn Kosuwin¹, Chaturong Putaporntip¹, Hiroshi Tachibana², Somchai Jongwutiwes¹*

1 Molecular Biology of Malaria and Opportunistic Parasites Research Unit, Department of Parasitology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, 2 Department of Infectious Diseases, Tokai University School of Medicine, Kanagawa, Japan

Abstract
Thrombospondin-related adhesive protein (TRAP) of malaria parasites is essential for sporozoite motility and invasions into mosquito’s salivary gland and vertebrate’s hepatocyte; thereby, it is a promising target for pre-erythrocytic vaccine. TRAP of Plasmodium vivax (PvTRAP) exhibits sequence heterogeneity among isolates, an issue relevant to vaccine development. To gain insights into variation in the complete PvTRAP sequences of parasites in Thailand, 114 vivax malaria patients were recruited in 2006-2007 from 4 major endemic provinces bordering Myanmar (Tak in the northwest, n = 30 and Prachuap Khiri Khan in the southwest, n = 25), Cambodia (Chanthaburi in the east, n = 29) and Malaysia (Yala and Narathiwat in the south, n = 30). In total, 26 amino acid substitutions were detected and 9 of which were novel, resulting in 44 distinct haplotypes. Haplotype and nucleotide diversities were lowest in southern P. vivax population while higher levels of diversities were observed in other populations. Evidences of positive selection on PvTRAP were demonstrated in domains II and IV and purifying selection in domains I, II and VI. Genetic differentiation was significant between each population except that between populations bordering Myanmar where transmigration was common. Regression analysis of pairwise linearized Fst and geographic distance suggests that P. vivax populations in Thailand have been isolated by distance. Sequence diversity of PvTRAP seems to be temporally stable over one decade in Tak province based on comparison of isolates collected in 1996 (n = 36) and 2006-2007. Besides natural selection, evidences of intragenic recombination have been supported in this study that could maintain and further generate diversity in this locus. It remains to be investigated whether amino acid substitutions in PvTRAP could influence host immune responses although several predicted variant T cell epitopes drastically altered the epitope scores. Knowledge on geographic diversity in PvTRAP constitutes an important basis for vaccine design provided that vaccination largely confers variant-specific immunity.

Introduction
In low- and middle-income countries in tropical areas, malaria remains one of the leading ten causes of morbidity and mortality, resulting in an estimated economic loss of nearly 40 million disability-adjusted life-years (DALYs) [1]. Although Plasmodium falciparum is the most virulent and prevalent species, the significance of P. vivax should not be underappreciated because it can cause chronic relapsing illness due to reactivation of hypnozoites and it can potentially lead to severe complications similar to those caused by P. falciparum [2]. Furthermore, vivax malaria occupies a broader geographical range than falciparum malaria whilst the emergences of chloroquine- and primaquine-resistant P. vivax are of particular concern if they would be widespread as chloroquine-resistant P. falciparum [3]. The progress of malaria control and elimination has been hampered by not only the emergence of drug resistant parasites but also insecticide resistant mosquito vectors; thereby, development of a malaria vaccine offers an important preventive measure [4]. Because P. vivax and P. falciparum co-circulate in several endemic areas outside Africa where co-infections of both species are not uncommon [5], effective malaria control requires vaccines against both species.

To date, malarial circumsporozoite (CS) protein is a prime candidate for pre-erythrocytic vaccine development [4]. Although CSP-derived immunogens could elicit immunity against sporozoites, the subunit vaccines derived from this molecule such as RTS, S recombinant vaccine against P. falciparum has resulted in limited clinical efficacy in field studies [6]. Because vaccines derived from irradiation-attenuated or live sporozoites consistently outperform vaccines incorporating single sporozoite proteins, a more effective pre-erythrocytic stage vaccine may require combination of multiple protective immunogens [7,8]. In a murine model, co-immunization of CSP with thrombospondin-related...
adhesive protein (TRAP), a protein mobilized from microneme to the surface of sporozoite [9], has conferred complete protection against parasite challenge whereas vaccination using each of these immunogens could elicit only partial protection [10]. It is important to note that TRAP-specific CD8+ T lymphocytes are prime mediators for protection against sporozoite challenge in mouse vaccination trials, resulting in significant reduction in liver stage parasites [11]. Furthermore, seroepidemiological study has shown that anti-TRAP antibodies were negatively correlated with parasite density among infected individuals in malaria endemic areas [12].

TRAP has been shown to mediate gliding motility and invasion processes of malarial sporozoites into vertebrate’s hepatocyte and mosquito’s salivary gland [13,14]. TRAP contains a hydrophobic N-terminal peptide (domain I), an integrin-like magnesium binding (or von Willebrand factor) A domain (domain II), thrombospondin type I repeats (domain III), an acidic proline/asparagine-rich region (domain IV), hydrophobic transmembrane domain (domain V) and a cytoplasmic tail (domain VI) [15,16]. The locomotion of sporozoites is mediated by the subpellicular actomyosin system that linked to the cytoplasmic tail of TRAP [17]. Despite functional importance of TRAP in parasite survival, analysis of the TRAP loci of P. falciparum (PfTRAP) and of P. vivax (PvTRAP) from clinical isolates revealed microheterogeneity of sequence that has been maintained by positive selective pressure [18–20]. Although it remains to be explored whether polymorphism in T cell epitopes of malarial TRAP could alter host cell immune recognition as that observed in CSP of P. falciparum [21], a rationale design of a malaria subunit vaccine targeting TRAP requires knowledge on the extent and pattern of sequence diversity in this molecule from diverse endemic areas.

Both P. vivax and P. falciparum have been circulating in Thailand with almost comparable prevalence since the past 2 decades. However, regional prevalence of P. vivax relative to P. falciparum seems to differ across major endemic areas of Thailand [22,23]. Our recent studies have shown that P. vivax populations in this country exhibited spatial variation in the extent of sequence diversity of major malaria vaccine candidate loci [24,25]. An effective malaria vaccine design undoubtedly requires knowledge on this issue to circumvent possible vaccine escape variants. Herein, we analyzed sequence diversity in the PvTRAP locus of clinical isolates from 4 major malaria endemic areas of Thailand.

**Materials and Methods**

**Human Ethics Statement**

The study protocol was approved by the Institutional Review Board on Human Research of Faculty of Medicine, Chulalongkorn University (IRB303/56). Written informed consent was obtained from all participants.

**Parasite populations**

Blood samples were obtained from 114 P. vivax-infected patients in northwestern (Tak province, n = 30), southwestern (Prachup Khiri Khan province, n = 25), eastern (Chanthaburi province, n = 29) and southern (Yala and Narathiwat provinces, n = 30) Thailand during 2006 and 2007. Venous blood samples were taken from each patient and were preserved in EDTA anticoagulant. These isolates have been previously determined to contain single clone infection based on sequence analysis of the merozoite surface protein-1 gene (Pomsp-1) as described previously [24,26].

DNA extraction, amplification and sequencing

DNA was prepared from blood samples using QIAamp kit (Qiagen, Hilden, Germany). The complete PvTRAP sequence was amplified by nested PCR using a forward outer primer (PvTRAP-F0: 5’-ATGTGTTGTACATTGGCTG-3’, nucleotides −442 to −422 before the start codon of the Salvador I sequence, GenBank accession number XM_001614097) and a reverse outer primer (PvTRAP-R0: 5’-TCATGAGGTGGGAAACACAC-3’, nucleotides 1961 to 1968 from the start codon). The inner pairs of primers were PvTRAP-F1 (5’-ATGTCTGTTGAGCNGCGGT-3’, nucleotides −398 to −378) and PvTRAP-R1 (5’-TCATCAAGACAGTTCCAGC-3’, nucleotides 1791 to 1798). PCR amplification was performed in a total volume of 30 μl of the reaction mixture containing template DNA, 2.5 mM MgCl₂, 300 mM each deoxynucleoside triphosphate, 3 μl of 10x ExTaq PCR buffer, 0.3 μM of each primer and 1.25 units of ExTaq DNA polymerase (Takara, Seta, Japan). Thermal cycling profile for primary PCR included the preamplification denaturation at 94°C for 2 min followed by 35 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 2 min, and a final extension at 72°C for 5 min. Nested PCR was done using the same thermal profile except that amplification was done for 25 cycles. DNA amplification was carried out using a GeneAmp 9700 PCR thermal cycler (Applied Biosystems, Foster City, CA). ExTaq DNA polymerase reportedly possesses efficient 5’→3’ exonuclease activity to increase fidelity and no strand displacement (Takara, Japan). The PCR product was examined by electrophoresis in a 1% agarose gel and was purified by using QIAquick PCR purification kit (QIAGEN, Germany). The PvTRAP sequences were determined directly and bi-directionally from PCR-purified templates and were performed on an ABI3100 Genetic Analyzer using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA). Whenever singleton substitution occurred, sequence was re-determined using PCR products from two independent amplifications from the same DNA template. Sequences in this study have been deposited in the GenBank Database under the accession numbers KJ807186 to KJ807299.

**Statistical analyses**

Sequence alignment with the Clustal X program [27] was done by using the PvTRAP sequence of the Salvador I strain as reference. The PvTRAP sequences of our previous Thai and Brazilian isolates previously published [19] and those reported by others [28] were included for analysis. For interspecific comparison, all PvTRAP sequences were aligned with the TRAP sequences of P. knowlesi (PKTRAP) and P. cynomolgi (PcTRAP) (GenBank accession numbers XM_002259951 and XM_004223600, respectively). Haplotype diversity (h) and its sampling variance were computed according to equations 8.4 and 8.12 but replacing 2n by n [29]. Nucleotide diversity (π) was computed from the average number of pairwise sequence differences in the sample [29] using the MEGA version 6 program [30]. Evidence of genetic recombination was determined by 2 different approaches: (i) estimation of the minimum number of recombination events (Rm) that can be parsimoniously inferred from a sample of sequences with Monte Carlo simulations [31] using the DnaSP version 5 software [32] and (ii) searching for recombination breakpoints by phylogenetic approach by the Genetic Algorithm Recombination Detection (GARD) method. Goodness of fit for the GARD method was calculated by Akaike Information Criterion derived from a maximum likelihood model fit to each segment (AICc) using the HyPhy package [33,34].

The rates of synonymous substitutions per synonymous site (dS) and nonsynonymous substitutions per nonsynonymous site (dN)
were computed by using Nei and Gojobori’s method [35] with Juke and Cantor correction [36] and their standard errors of these parameters were estimated by the bootstrap method with 1,000 pseudoreplicates as implemented in the MEGA 6.0 program [30]. A statistical significance level of differences in $d_{st}$ and $d_{s}$ was set at 5%.

Tests for departure from neutral evolution based on intraspecific comparisons were done by Tajima’s $D$, Fu and Li’s $D^*$ and $F^*$ statistics. Tajima’s $D$ test determines the differences between the average number of nucleotide differences and an estimate of 0 from the number of segregating sites where $\theta = 4N_e\mu$, in which $N_e$ and $\mu$ are the effective population size and the mutation rate, respectively [37]. Fu and Li’s statistics measure the differences between the number of singletons and the total number of mutations ($D^*$ test) and the differences between the number of singletons and the average number of nucleotide differences between pairs of sequences ($F^*$ test) [38]. To examine whether deviation from neutrality occurs in particular fragments of $PvTRAP$, a sliding window analysis of Tajima’s $D$ across the coding sequences was done using a window size of 250 bp and a step size of 10 bp [32].

Evidence of departure from neutrality based on interspecific comparison included Fu and Li’s $D$ and $F$ tests, and McDonald-Kreitman tests. Fu and Li’s $D$ statistics measure the differences between the number of mutations in external branches of the genealogy and the number of mutations while their $F$ statistic compares the differences between the number of mutations in external branches of the genealogy and the average number of nucleotide differences between pairs of sequences [38]. Coalescence simulation with 10,000 pseudoreplicates was done to estimate significance levels of these parameters as implemented in the DnaSP program [32]. The McDonald-Kreitman test contrasts intraspecific polymorphism with interspecific divergence of closely related species. Statistical departure from neutral expectation was computed by Fischer’s exact test [39].

Tests of departure from neutrality at specific codons were performed based on estimation of the global ratio of the rate of non-synonymous to synonymous substitutions ($dN/dS$) across the $PvTRAP$ gene. The single-likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), random effects likelihood (REL) and mixed effects model of evolution (MEME) methods implemented in the HyPhy package [33] were used for analysis. SLAC model is highly conservative based on the maximum likelihood reconstruction of the ancestral sequences and the counts of synonymous and non-synonymous changes at each codon position in a phylogeny under the assumption of neutral evolution. FEL model compares the ratio of non-synonymous to synonymous substitution on a site-by-site basis, without assuming an a priori distribution of rates across sites whereas REL model first fits a distribution of rates across sites and then infers the substitution rate for individual sites. MEM algorithm detects codons under episodic positive selection unmasked by the abundance of purifying selection along the lineages [40]. Significance level settings for SLAC, FEL and MEME were $p$ values $< 0.1$ and Bayes Factor $> 1000$ for REL followed the default values available on the Datamonkey Web Server [41].

The population genetic structure was analyzed by molecular variance approach (AMOVA) using the Arlequin 3.5 software which is similar to the Weir and Cockerham’s method but it takes into account the number of mutations between haplotypes [42]. The implemented algorithm calculates the fixation index $F_{ST}$ identical to the weighted average $F$-statistic over loci, $\theta w$ [43] and the significance levels of the fixation indices were estimated by a non-parametric permutation [42]. Pairwise Slakin’s linearized $Fst$ was calculated based on $Fst/(1-Fst)$ to address the relationship between the magnitude of population differentiation and distance between endemic areas [44].

The genetic structure was also determined by STRUCTURE 2.3.4 program that deploys the Bayesian model-based clustering approach [45]. The most probable number of populations (K) was estimated using an admixture model. All sample data were run for values K = 1–10, each with a total of 15 iterations. We used 500,000 Markov Chain Monte Carlo generations for each run following a burn-in of 50,000 steps. The most likely number K in the data was estimated by calculating $\Delta K$ values [46] and identifying the K value that maximizes the log probability of data, ln(PD) [45] as implemented in STRUCTURE HARVESTER program [47].

The phylogenetic tree of the $PvTRAP$ haplotypes was constructed by using the Maximum Likelihood method based on the model with the lowest Bayesian Information Criterion (BIC) scores [30]. All positions containing gaps and missing data were eliminated. Initial tree(s) for the heuristic search were obtained by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The final tree topology was selected based on superior log likelihood value as implemented in the MEGA version 6 program [30].

Amino acid properties were characterized based on polarity and charge. Polarity of each residue is categorized as polar (S, Y, C, W, H, Q, T and N) or nonpolar (the remainder). Charge property includes positive (H, R and K), negative (D and E) or neutral (the remainder). Prediction of HLA-I binding peptides was based on the method taken into account proteasomal C terminal cleavage and transporter associated with antigen processing (TAP) transport efficiency. The predicted scores were determined by the NetCTL program [48].

### Results

#### Genetic diversity in $PvTRAP$

Of 114 complete $PvTRAP$ sequences, 109 isolates contained 1668 bp and 5 isolates from Tak province whose sequences were identical had 1677 bp. Sequence alignment with the Salvador I strain revealed 36 nucleotide substitutions within 33 codons, resulting in 26 non-synonymous and 7 synonymous changes. Comparing with previous data from Brazilian, Thai [19], South Korean isolates [28], 9 new amino acid substitutions were identified (Table 1). Like our previous findings, the majority of amino acid changes occurred in domains II and IV [19]. Three codons contained more than two substituted nucleotides: codon 206 (after the Salvador 1 sequence) having 4 different substitutions: c.616C>T (P206S), c.617G>C (P206S), c.617G>T (P206L) and c.616_617delinsTT (P206L), codon 268 having double mutations: c.802_803delinsGG (K268G), and codon 424 containing 2 changes: c.1271A>C (N424T) and c.1272C>A (N424K). Among isolates examined herein, no discernable bias in frequency of nucleotide substitutions in each position of the codons occurred, i.e. 11, 11 and 13 substitutions in the first, second and third positions of the codons, respectively. In total, 44 distinct haplotypes were identified in this analysis, one of which had a duplication of 9 nucleotides encoding PDS in domain IV as previously noted [19] and responsible for size variation (Tables 1 and S1).

The distribution of $PvTRAP$ haplotypes in southern Thailand (Yala and Narathiwat provinces) was rather skewed towards few haplotypes as shown by a relatively lower value of haplotype diversity ($h = 0.653 \pm 0.063$) whereas parasite populations from
### Table 1. Distribution of PvTRAP haplotypes from 4 major malaria endemic areas of Thailand.

| Amino acid residue | Prachuap Khiri Khan | Chanthaburi | Yala & Narathiwat | Total |
|--------------------|---------------------|-------------|-------------------|-------|
| Domain             | I                   | II          | II                | IV    | III          | IV | IV | IV | IV |
| Salvador-1         | L                   | L           | T                 | D     | E            | Q  | R  | V  | R  |
|                     | 1                   | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 2                   | 2           | 3                 | 3     | 4            | 4  | 4  | 4  | 4  |
|                     | 3                   | 5           | 0                 | 2     | 4            | 5  | 6  | 7  | 8  |
|                     | 4                   | 8           | 8                 | 8     | 9            | 9  | 2  | 3  | 4  |
|                     | 5                   | 6           | 6                 | 5     | 8            | 5  | 7  | 6  | 0  |
|                     | 6                   | 4           | 4                 | 3     | 1            | 4  | 1  | 1  | 2  |
|                     | 7                   | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 8                   | 2           | 2                 | 2     | 2            | 2  | 2  | 2  | 2  |
|                     | 9                   | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 10                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 11                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 12                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 13                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 14                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 15                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 16                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 17                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 18                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 19                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 20                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 21                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 22                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 23                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 24                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 25                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 26                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 27                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 28                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
|                     | 29                  | 1           | 1                 | 1     | 1            | 1  | 1  | 1  | 1  |
Table 1. Cont.

| Amino acid residue | Prachuap Khirikhan | Chanthaburi | Yala & Narathiwat | Total |
|--------------------|--------------------|-------------|------------------|-------|
| Tak                |                    |             |                  |       |
| 1                  | 1                  | 1           | 1                | 1     |
| 2                  | 1                  | 1           | 1                | 1     |
| 3                  | 1                  | 1           | 1                | 1     |
| 4                  | 1                  | 1           | 1                | 1     |
| 5                  | 1                  | 1           | 1                | 1     |
| Haplotype          | 1                  | 4           | 2                | 1     |
| 5                  | 5                  | 0           | 2                | 4     |
| 7                  | 3                  | 0           | 6                | 0     |
| 6                  | 2                  | 6           | 0                | 6     |
| 5                  | 8                  | 7           | 4                | 6     |
| Domain             | I                  | II          | II               | II    |
| Salvador-1         | L                  | L           | T                | D     |
| #30                | . . . . S . . . . . | K . I . G . R | . - - . . G . - | -   |
| #31                | . . . S . S . E . . | I T K . I . G . R | . - - K . D . | -   |
| #32                | . . . S . S . D . . | I T . I . G . R | . - - K . . . | -   |
| #33                | . . . S . . . . . . | K . I . G . R | T - - . . G . | -   |
| #34                | . . . S . S . . . . | I T . I . G . R | . - - K . D . | -   |
| #35                | . . . S . E . . . Q | I T K . I . G . R | - - K . D . | -   |
| #36                | . . . S . S . . . . | K . T I . G . R | . - - . . . . | -   |
| #37                | . . . S . S . . . . | I T K . I . G . R | - - . . G . | -   |
| #38                | . . . S . S . . . . | K . I . G . S | . - - . . . . | -   |
| #39                | . . . S . S . . . . | K . I . G . R | T - - . . . . | -   |
| #40                | . . . S . S . . . . | K . I . G . R | T - - . . . . | -   |
| #41                | . . . S . S . . . . | K . I . G . R | T - - . . . . | -   |
| #42                | . . . S . S . . . . | K . I . G . R | . - - K . . . | -   |
| #43                | . . . S . . . . . . | K . I . G . R | - - K . . . | -   |
| #44                | . . . S . . . . . . | K . I . G . R | T - - . . . . | -   |
| Protein Structure* | β1 α3 β4 α4 α5 α5 | 31 24 29 30 | 114             |

Newly identified amino acid changes are in bold. *Protein secondary structure and segment are after Song, et al. 2012 [51].

doi:10.1371/journal.pone.0110463.t001
Tak, Prachuap Khirikhan and Chanthaburi provinces had significantly higher haplotype diversity values, suggesting a more even distribution of haplotype frequencies in these latter endemic regions (Table 2). It is noteworthy that the predominant haplotypes differed between endemic areas. For example, haplotypes #39 (n = 15) and #42 (n = 10) were exclusively found in Yala and Narathiwat provinces whereas haplotypes #28 (n = 10) and #30 (n = 6) were unique for Chanthaburi province. In contrast, several haplotypes were shared between isolates from Tak and Prachuap Khirikhan provinces (Table 1). The nucleotide diversity (π) of P. vivax populations from Brazil, South Korea and Thailand were significantly different from one another, indicating strong differentiation and reproductive isolation among populations. These results were in line with analysis using STRUCTURE version 2.3.4 software in which the subpopulation likelihood reached a plateau at K = 3 to K = 6. Consistently, a greatest delta K value was also obtained at K = 3 (Figure 1). The genetic structure at K = 3 could differentiate Brazilian isolates from Thai and South Korean parasite populations while a separation of South Korean and southern Thai P. vivax populations were not perceived. Similar findings were observed at K = 4. A more concordant results between those obtained by STRUCTURE program and the Fst parameters occurred when K = 5 in which isolates from Tak collected during different periods and from Prachuap Khirikhan were indistinguishable whereas South Korean and southern Thai populations were distinct (Figure 1). Regression analysis has shown positive correlation between pairwise linearized Fst and the natural logarithm of linear distance between endemic areas (Mantel test, r = 0.4619, p = 0.024). Likewise, positive correlation was found when geographic linear distance between endemic areas was considered (Mantel test, r = 0.4503, p = 0.020), implying that the majority of P. vivax populations in Thailand were genetically isolated by distance (Figure 2).

### Table 2. Variation in the PvTRAP sequences of Thai isolates from diverse endemic areas.

|                | Tak              | Prachuap Khirikhan | Chanthaburi | Yala and Narathiwat | All   |
|----------------|------------------|--------------------|-------------|---------------------|-------|
| N              | 30               | 25                 | 29          | 30                  | 114   |
| M              | 14               | 24                 | 18          | 7                   | 32    |
| S              | 14               | 23                 | 18          | 7                   | 31    |
| H              | 17               | 18                 | 12          | 6                   | 43    |
| h ± S.D.       | 0.956 ± 0.018    | 0.957 ± 0.027      | 0.818 ± 0.053| 0.653 ± 0.063       | 0.954 ± 0.008 |
| π ± S.D.       | 0.00220 ± 0.00017| 0.00272 ± 0.00032  | 0.00212 ± 0.00032| 0.00117 ± 0.00014### | 0.00272 ± 0.00009 |
| dS ± S.E. (all)| 0.00018 ± 0.00017| 0.00074 ± 0.00043  | 0.00069 ± 0.00052| 0.00201 ± 0.00142  | 0.00118 ± 0.00081 |
| Domain I       | 0.00000 ± 0.00000| 0.00397 ± 0.00467   | 0.00000 ± 0.00000| 0.00000 ± 0.00000  | 0.00008 ± 0.00097 |
| Domain II      | 0.00000 ± 0.00000| 0.00029 ± 0.00027   | 0.00048 ± 0.00046| 0.00176 ± 0.00130  | 0.00067 ± 0.00038 |
| Domain III     | 0.00000 ± 0.00000| 0.00000 ± 0.00000   | 0.00000 ± 0.00000| 0.00000 ± 0.00000  | 0.00000 ± 0.00000 |
| Domain IV      | 0.00000 ± 0.00000| 0.00002 ± 0.00002   | 0.00000 ± 0.00000| 0.00000 ± 0.00000  | 0.00000 ± 0.00000 |
| Domain V       | 0.00000 ± 0.00000| 0.00000 ± 0.00000   | 0.00000 ± 0.00000| 0.00000 ± 0.00000  | 0.00000 ± 0.00000 |
| Domain VI      | 0.00250 ± 0.00281| 0.00575 ± 0.00593   | 0.00721 ± 0.00781| 0.01906 ± 0.00000  | 0.01245 ± 0.01187 |
| dS ± S.E. (all)| 0.00278 ± 0.00090** | 0.00327 ± 0.00086** | 0.00252 ± 0.00077* | 0.00091 ± 0.00052 | 0.00314 ± 0.00090 |
| Domain I       | 0.00000 ± 0.00000| 0.00233 ± 0.00232   | 0.00000 ± 0.00000| 0.00000 ± 0.00000  | 0.00053 ± 0.00055 |
| Domain II      | 0.00264 ± 0.00134* | 0.00428 ± 0.00161*  | 0.00406 ± 0.00161* | 0.00014 ± 0.00014  | 0.00396 ± 0.00173 |
| Domain III     | 0.00423 ± 0.00426 | 0.00442 ± 0.00439   | 0.00067 ± 0.00071 | 0.00000 ± 0.00000  | 0.00261 ± 0.00266 |
| Domain IV      | 0.00381 ± 0.00162* | 0.00314 ± 0.00132*  | 0.00248 ± 0.00118* | 0.00227 ± 0.00141  | 0.00373 ± 0.00157* |
| Domain V       | 0.00000 ± 0.00000| 0.00000 ± 0.00000   | 0.00000 ± 0.00000| 0.00000 ± 0.00000  | 0.00000 ± 0.00000 |
| Domain VI      | 0.00000 ± 0.00000| 0.00000 ± 0.00000   | 0.00000 ± 0.00000| 0.00000 ± 0.00000  | 0.00000 ± 0.00000 |

N, number of sequences; M, number of mutations; S, number of segregating sites; H, number of haplotypes.

Z-tests of the hypothesis that π equals the corresponding value for each population: ### p < 0.0001.
Z-tests of the hypothesis that mean dS equals that of mean dN: * p < 0.05; ** p < 0.01.
doi:10.1371/journal.pone.0110463.t002
Tests for departure from neutral evolution

Analysis of patterns of nucleotide substitutions in the \textit{PvTRAP} sequences of isolates in this study reveals that the rate of nonsynonymous substitutions per nonsynonymous sites or $d_N$ significantly exceeds that of synonymous substitutions per synonymous sites or $d_S$ in all parasite populations ($p < 0.05$ or $p < 0.01$) except \textit{P. vivax} isolates from Yala and Narathiwat provinces (Table 2). In contrast, $d_S$ was greater than $d_N$ for isolates from Yala and Narathiwat provinces although the difference was not statistically meaningful ($p = 0.467$). Closer looks into the rate of nucleotide substitutions in each domain of \textit{PvTRAP} has revealed that $d_N$ significantly outnumbered $d_S$ in domains II and IV of...
isolates from Tak, Prachuap Khirikhan and Chanthaburi provinces. Although no significant difference in $d_S$ and $d_N$ for the entire coding region of the *PvTRAP* gene was observed when all parasite populations were considered, $d_N$ in domain IV was obviously greater than $d_S$ (Table 2). Meanwhile, statistics based on Tajima’s $D$ and Fu & Li’s $D^*$ and $F^*$ did not show evidences of departure from neutral expectation (Table 4). However, sliding window analysis reveals significant negative Tajima’s $D$ in domain II of *P. vivax* populations from Prachuap Khirikhan and Chanthaburi provinces, suggesting purifying selection at certain residues in this gene. In contrast, evidence of balancing selection was detected in domain IV of Yala and Narathiwat isolates as shown by significant

Figure 2. Regression analysis of pairwise linearized $Fst$ regressed on natural logarithm geographic distances between *P. vivax* populations in Thailand (A) and pairwise linearized $Fst$ regressed on geographic distances between populations (B).

doi:10.1371/journal.pone.0110463.g002
Table 4. Intraspecific and interspecific neutrality tests for the PvTRAP locus from each endemic area of Thailand.

| Province/Endemic area | Out group: P. cynomolgi | Out group: P. knowlesi | Fu & Li's D | Fu & Li's F | Tajima's D |
|-----------------------|-------------------------|------------------------|-------------|-------------|-----------|
| Tak (n = 30)           | 0.097                   | 0.097                  | 0.080       | -0.090      | -0.173    |
| Prachuap Khirikhan (n = 25) | 0.350                 | 0.350                  | 0.967       | 0.967       | 0.350     |
| Chanthaburi (n = 29)   | 0.818                   | 0.818                  | 2.080       | 2.080       | 1.429     |
| Yala and Narathiwat (n = 30) | 0.278                | 0.278                  | 1.097       | 1.097       | 0.698     |
| All (n = 114)          | 0.698                   | 0.698                  | 1.572       | 1.572       | 1.055     |

All values did not have significance departure from zero. doi:10.1371/journal.pone.0110463.t004

Phylogenetic relationship

Model test reveals that the Tamura 3-parameter with the rate variation model that allowed for some sites to be evolutionarily invariable gave the lowest BIC value. The tree with the highest log likelihood (−3285.15) reveals that the majority of Brazilian isolates were placed in a separate lineage whereas Chanthaburi isolates and southern Thai P. vivax populations had equal number of potential recombination sites (Rm = 2). Locations of potential recombination sites for each population are listed in Table S3, most of which span domain IV. These results suggest that intragenic recombination could shape diversity at the P. vivax TRAP locus.
Yala and Narathiwat provinces seem to be located in separate branches from the remaining Thai isolates. However, there was no clear lineage for each parasite population because several isolates from different endemic areas shared or was placed in closely related branches. More importantly, all branches in this phylogenetic inference received bootstrap values less than 80%, suggesting no distinct lineage of the \( P_vTRAP \) haplotypes from worldwide origins (Figure 4).

**Discussion**

Host adhesion and motility are fundamental properties of malarial sporozoites to establish infection that involves sporozoite-specific
proteins, one of which is TRAP. It has been shown that TRAP-deficient sporozoites were severely impaired in their host cell adherence property and gliding motility; thereby, invasion of mosquito’s salivary glands and hepatocytes was interrupted [14,49]. Two extracellular portions of malarial TRAP, the von Willebrand A-domain and the thrombospondin repeats located in domains II and IV, respectively, are crucial for initial host cell adherence and stabilization of adhesion/deadhesion during gliding motility of sporozoites [49,50]. Crystal structure analyses spanning the von Willebrand A-domain and the thrombospondin repeats of PvTRAP and PTRAP reveal two conformational states, open and closed structures. Such structural phase transition is possibly responsible for ‘stick-and-slip’ or gliding motility of sporozoites through the actomyosin motility apparatus [51,52]. The sequence motif of metal-ion-dependent adhesion sites (MIDAS) in von Willebrand A-domain and the extensible β ribbon that implicates in ligand binding during the open high-affinity state [51] were perfectly conserved in all PvTRAP haplotypes as shown in this study. Although amino acid substitutions were observed in the β1 and β3 domains, charge and polarity of these substituted residues were retained, suggesting structural constraint in these portions of the protein (Table 1). Experiment studies have shown that artificially engineered mutations of the MIDAS-coordinating amino acid residues that were conserved across malaria species could impair gliding motility and sporozoite infectivity to both mosquito and mammalian hosts [49,50]. On the other hand, alteration in both charge and polarity profiles of some substituted amino acids (residues 143, 166, 172 and 176, Tables 1 and S2) were identified in β4 and α5 domains, implying that changes in these residues may not drastically affect functional conformation of PvTRAP. Nevertheless, further experimental studies are required to address this issue.

The extent of sequence diversity of the PvTRAP locus among parasite populations in Thailand seems to be comparable among endemic areas except that from southern region (Table 2). The results from this study are in line with our recent analyses of other major vaccine antigens, i.e. apical membrane antigen-1 (Pvama-1), merozoite surface protein-1 (Pvmsp-1), merozoite surface protein-4 (Pvmsp-4) and merozoite surface protein-5 (Pvmsp-5) of P. vivax showing that southern parasite population possesses significantly lower levels of nucleotide diversity, number of haplotypes and haplotype diversity than those of northwestern parasite population [24,25,53]. Heterogeneity in the PvTRAP sequences could have been arisen from selective pressures and intragenic recombination. Recombination confers evolutionary advantages because creation of adaptive traits or removal of deleterious mutants is enabled. Importantly, most recombination sites were identified within or spanning domain IV of the PvTRAP locus where signature of positive selection was identified. The relatively higher levels of nucleotide diversity of PvTRAP from Thai-Myanmar border (Tak and Prachuap Khirikhan isolates) than that from Thai-Cambodia border (Chanthaburi isolates) seems to be in line with Rm in each population. It is likely that transmigration of people could maintain and spread of P. vivax harboring variant PvTRAP haplotypes while recombination could further generate novel alleles. Intriguingly, it seems that reduction in genetic diversity of southern parasitie population may not be simply explained by clonal population structure of P. vivax in this region because recombination event could be traced despite the paucity of potential recombination sites observed (Rm = 2). In contrast, several decades of extensive malaria control in this region and a lack of remarkable re-introduction of malaria cases from Malaysia could have resulted in population bottleneck as previously suggested [24]. Therefore, after bottleneck effect and relaxation of malaria control in southern Thailand because of difficulty in implementation due to local political unrest, it could be possible that recombination between different PvTRAP haplotypes may gradually increase the level of genetic diversity in this population.

Our analyses have shown that domains II and IV of the PvTRAP locus had significantly different in the rate of nonsynonymous substitutions than that of synonymous substitutions. Evidence of positive or balancing selection was also reaffirmed by the McDonald-Kreitman test. Although departure from neutrality in the entire coding PvTRAP sequence was not apparent by using Tajima’s D and related statistics, sliding window analysis has identified significant negative Tajima’s D values in domains II of P. vivax population from Prachuap Khirikhan and Chanthaburi provinces while significant positive values occurred in domain IV among isolates from Yala and Narathiwat provinces. Differential patterns of departure from neutrality among populations could reflect different evolutionary forces exerted on PvTRAP of each population. Alternatively, intrinsic statistical property of Tajima’s D statistics can be confounded by demographic processes [54]. Therefore, significant positive Tajima’s D observed in southern Thai isolates could imply balancing selection or a recent bottleneck because an excess of intermediate frequency alleles as rare alleles are lost from the population due to genetic drift can be found immediately after the bottleneck. Meanwhile, significant negative Tajima’s D value in domains II could imply purifying selection and selective sweep while migration and population growth may be the confounders [55]. Nevertheless, positively selected codons were found based on analyses using SLAC, FEL, REL and MEME methods while negatively selected residues were detected in domains I, II and VI (Table 6). Taken together, it seems that co-existence of both

| Population               | Polymorphic changes within P. vivax | Fixed differences between species | Neutrality index | p value |
|--------------------------|-------------------------------------|----------------------------------|------------------|---------|
|                          | Synonymous | Nonsynonymous | Synonymous | Nonsynonymous |                       |                   |
| Tak                      | 1          | 13            | 116       | 140          | 10.77     | 0.0045               |
| Prachuap Khirikhan       | 3          | 21            | 117       | 138          | 5.93      | 0.0011               |
| Chanthaburi              | 2          | 16            | 117       | 140          | 6.69      | 0.0053               |
| Yala and Narathiwat      | 3          | 4             | 115       | 145          | 1.06      | 1.0000               |
| All                      | 6          | 23            | 117       | 136          | 3.298     | 0.0096               |

Note: Repeats are excluded from all analyses.

doi:10.1371/journal.pone.0110463.t005
| Domain | Codon | SLAC | FEL | REL | MEME | Consensus* |
|--------|-------|------|-----|-----|------|------------|
|        |       | d\_N - d\_S | p value | d\_N - d\_S | p value | d\_N - d\_S | Bayes Factor | \(\omega^+\) | p value | SLAC | FEL | REL | MEM |
| II     | 120   | 93.598 | 0.213 | 2821.14 | 0.091 | 73.567 | 9897.75 | >100 | 0.115 | + | + |
| II     | 134   | 74.642 | 0.496 | 2675.58 | 0.208 | 73.381 | 195736 | >100 | 0.230 | + |
| II     | 206   | 84.790 | 0.280 | 3760.95 | 0.105 | 73.323 | 70283 | >100 | 0.051 | + | + |
| II     | 255   | 149.777 | 0.112 | 3884.15 | 0.063 | 73.501 | 2.27475 \times 10^3 | >100 | 0.084 | + | + | + |
| IV     | 410   | 95.051 | 0.401 | 3087.63 | 0.178 | 73.377 | 8780900 | >100 | 0.203 | + |
| IV     | 421   | 55.136 | 0.594 | 1821.01 | 0.252 | 73.389 | 248453 | >100 | 0.267 | + |
| IV     | 439   | 83.134 | 0.304 | 2728.63 | 0.104 | 73.541 | 267258 | >100 | 0.129 | + |
| I      | 21    | -102.062 | 0.151 | -2637.21 | 0.063 | -38.418 | 35182.5 | <100 | >0.100 | - | - |
| II     | 30    | -204.123 | 0.023 | -5228.32 | 0.011 | -38.410 | 29828.7 | <100 | >0.100 | - | - |
| II     | 100   | -46.299 | 0.333 | -1399.13 | 0.131 | -38.407 | 29051.4 | <100 | >0.100 | - |
| II     | 133   | -31.970 | 0.483 | -1019.35 | 0.210 | -38.100 | 855841 | <100 | >0.100 | - |
| VI     | 519   | -92.599 | 0.111 | -2873.15 | 0.032 | -38.407 | 29054.6 | <100 | >0.100 | - | - |
| VI     | 539   | -83.719 | 0.199 | -3571.07 | 0.051 | -38.391 | 27525.8 | <100 | >0.100 | - |

*positively selected site, +; negatively selected site, -.

Default significance levels in Datamonkey program: SLAC, \(p<0.1\); FEL, \(p<0.1\); REL, Bayes Factor >1000; MEME, \(p<0.1\).

doi:10.1371/journal.pone.0110463.t006
negatively selected amino acids and residues under positive selection could stem from functional constraint of this molecule and probably driven by host immune pressure, respectively. Meanwhile, it seems that intragenic recombination was pronounced in populations from Tak and Prachuap Khirikhan provinces while relatively few recombinations occurred in parasites from Chanthaburi province and southern isolates. These findings are in accord with our previous analysis on \(PvMSP4\) and \(PvMSP5\) in which frequency of recombination seems to correlate with endemicity of each area [25,53].

Despite low level of sequence diversity, genetic distances inferred from the \(PvTRAP\) locus among \(P. \text{ vivax}\) populations could be deployed for analysis of population structure. Pairwise comparison of populations revealed that South Korean isolates displayed high and significant \(Fst\) indices when compared with malaria populations in other endemic areas. Likewise, slightly lower but significant \(Fst\) indices were observed when isolates from Brazil were compared with populations elsewhere (Table 3). Therefore, gene flow between \(P. \text{ vivax}\) populations in Thailand, Brazil and Korea may not occur because geographic distance has precluded intra- and intercontinental spread of parasites. Although malaria can be transferred between remote endemic areas through infected travelers, this situation could be inefficient due to rarity in numbers of such cases to contribute to genetic admixture of parasites in these three countries. However, closer look into genetic differentiation between 4 major malaria endemic areas of Thailand, significant genetic differentiation of \(P. \text{ vivax}\) population between endemic areas was observed except parasites from Tak province collected over a decade apart (1996 and 2006–2007) having low and non-significant \(Fst\) indices. Therefore, diversity in the \(PvTRAP\) locus of parasites in Thailand seems to exhibit spatial but not temporal variation. It is noteworthy that regression analysis of linear geographic as well as normal logarithm of geographic distances and pairwise linearized \(Fst\) values yielded significant correlation (Mantel test, \(p = 0.024\) and \(0.028\), respectively), indicating that \(P. \text{ vivax}\) populations in these 4 endemic areas have been isolated by distance (Figure 2). Therefore, it seems likely that the flight range of mosquito vectors was the major determinant for magnitude of genetic differentiation of \(P. \text{ vivax}\) populations between endemic areas. However, there remain some other factors that could influence and shape genetic structure of malaria in Thailand. Although geographic distance between Tak and Prachuap Khirikhan provinces is about 160 kilometers longer than that between Tak and Chanthaburi provinces, significant genetic differentiation occurred in the latter but not the former. It should be noted that transmigration of people along Thai-Myanmar border where both Tak and Prachuap Khirikhan provinces are located is intense during the past decades whereas transmigration of gem miners who, not uncommonly, carried malaria parasites in their circulation between Tak and Chanthaburi provinces had ceased for over 3 decades ago [56]. Furthermore, the abundance of malaria vectors along Thai-Myanmar
| HLA | Domain | Predicted epitopes | Predicted score | %Reduction | No. isolates |
|-----|--------|-------------------|----------------|------------|-------------|
| A1  | I      | YLVLVFELLY        | 2.1645         |            | 112         |
| A1  | II     | YLVLVFELL        | 1.6240         | 24.97      | 2           |
| A1  | I      | VCNESVSDLFY       | 0.9376         |            | 113         |
| A1  | II     | VCNESVSDL        | 0.7571         | 19.25      | 2           |
| A2  | I      | FLLYSIYFA        | 1.1395         |            | 112         |
| A2  | II     | FLLYSIYF        | 0.5416         | 52.47      | 2           |
| A2  | I      | NMATAALEEV        | 1.1851         |            | 47          |
| A2  | II     | NMATAALDEV      | 1.1104         | 63.00      | 65          |
| A2  | II     | NMATAALDDV      | 0.7819         | 34.02      | 2           |
| A3  | II     | YTALEVAKK        | 0.8093         |            | 63          |
| A3  | II     | YRALEVAKK        | 0.2512         | 68.96      | 51          |
| A24 | I      | VFLLVSIF         | 1.7214         |            | 112         |
| A26 | II     | ELKYTPSYP        | 2.0070         |            | 12          |
| B7  | II     | KLKQRNVSL         | 0.8786         |            | 112         |
| B7  | II     | KLKQRNVSL        | 0.8400         | 4.39       | 1           |
| B7  | II     | KLKQRNTL         | 0.7263         | 17.33      | 1           |
| B7  | II     | KQRNVTLAV        | 0.9688         |            | 1           |
| B7  | II     | KQRNVSALV        | 0.8745         | 9.73       | 1           |
| B7  | II     | KQRNVSALV        | 0.7908         | 18.37      | 112         |
| B7  | IV     | LPVPAPLPA        | 1.3094         |            | 2           |
| B8  | II     | RPRENCKF         | 1.9566         |            | 1           |
| B8  | II     | RPRELNCCKF       | 1.3892         | 29.00      | 2           |
| B8  | II     | RPRESNCKF        | 1.0950         | 44.04      | 2           |
| B8  | II     | RREPNCCKF        | 1.0440         | 46.64      | 109         |
| B27 | II     | KQRNVSALV        | 1.2095         |            | 1           |
| B27 | II     | KQRNVSALV        | 0.9591         | 20.70      | 112         |
| B27 | II     | KQRNVTLAV        | 0.9591         | 20.70      | 1           |
| B27 | II     | KQRNVSALV        | 0.7781         |            | 112         |
| B27 | II     | KQRNVTLAV        | 0.7781         | 0          | 1           |
| B27 | III    | GGGTHSSR         | 1.2775         |            | 96          |
| B39 | I      | SYLLVFL          | 0.4904         | 61.61      | 18          |
| B44 | IV     | NEKVFPNPL        | 1.4650         |            | 110         |
| B58 | I      | KSYLLVFLF        | 1.4989         |            | 2           |
| B62 | II     | KQRNVTLAV        | 1.1504         |            | 1           |
|     |        | KQRNVSALV        | 1.1443         | 0.53       | 112         |
border could facilitate gene flow of *P. vivax* in these areas. In contrast, ecological niches of *Anopheles* vectors between Tak and Chanthaburi provinces are largely interrupted by local deforestation and urbanization, precluding natural spread of malaria between these areas.

Analysis of hierarchical splitting of gene pools by STRUCTURE has shown that *K* = 3 is an optimum value. Intriguingly, *K* = 5 could not subdivide South Korean and southern Thai isolates that in fact differed biologically and genetically. A very long incubation period of *vivax* malaria from 230 to 300 days has been recognized among South Korean isolates [20] but, to our knowledge, has never been recorded in Thai strains. Furthermore, phylogenetic analysis has placed all South Korean isolates to a separate cluster from southern Thai isolates, albeit without significant bootstrap support (Figure 3). In contrast, the partition at *K* = 5 is more consistent with results from analysis using *Fst* indices in which distinct subpopulation structure was perceived between isolates from southern Thailand and those from South Korea whilst at *K* ≤ 5, no discernible difference was observed between these populations. Meanwhile, phylogenetic inference does not support distinct clusters of isolates relating to their geographic origins because no branch received remarkably high bootstrap values, consistent with our previous study [19]. Therefore, genetic diversity of the *Pv*TRAP locus displays tendency towards geographic variation (Figure 4).

Protective immunity against malarial TRAP involves both humoral and cell mediated immune responses. Antibodies against TRAP targeting domains involved in gliding motility and hepatocyte binding ligands could prevent sporozoite invasion into hepatocytes [13] while TRAP-specific cytotoxic T cells could destroy sporozoite-infected hepatocytes [57]; thereby, exoerythrocytic development does not ensue. The roles of anti-TRAP antibodies in partial protection from or significantly reduced risk of *falciparum* malaria have been demonstrated in African endemic areas [13,58,59]. With analogy to the thrombospondin-related motif (TRM) of PfTRAP that binds heparin sulfate proteoglycan of hepatocytes [13], while TRAP-specific cytotoxic T cells could destroy sporozoite-infected hepatocytes [57]; thereby, exoerythrocytic development does not ensue. The roles of anti-TRAP antibodies in partial protection from or significantly reduced risk of *falciparum* malaria have been demonstrated in African endemic areas [13,58,59]. With analogy to the thrombospondin-related motif (TRM) of PfTRAP that binds heparin sulfate proteoglycan of hepatocytes, the TRM in domain III of PfTRAP is relatively conserved with two amino acid substitutions, i.e. WTACSVTC-

**Table 7.** Cont.

| HLA Domain | Predicted epitopes | Predicted score | %Reduction | No. isolates |
|------------|-------------------|----------------|------------|-------------|
| KKRIN\$S$AV | 0.2828 | 75.42 | 1 |

*Based on the C-terminal cleavage and the transporter associated with antigen processing efficiency [48].

Supporting Information

**Table S1** Haplotypes of *Pv*TRAP from Thai isolates.

**Table S2** Characteristics of amino acid substitutions in *Pv*TRAP of worldwide and Thai isolates.

**Table S3** Minimum number of recombination events (Rm) and recombination sites in the *Pv*TRAP gene of each parasite population in Thailand.

Acknowledgments

We are grateful to all patients who donated their blood samples for this study. We thank the staﬀ of the Bureau of Vector Borne Disease, Department of Disease Control, Ministry of Public Health, Thailand, for assistance in field work.

Author Contributions

Conceived and designed the experiments: SJ CP RK. Performed the experiments: RK CP SJ. Analyzed the data: SJ CP RK. Contributed reagents/materials/analysis tools: SJ CP HT. Contributed to the writing of the manuscript: SJ GP. Reviewed the paper: SJ CP RK HT.
