Published in final edited form as:

Vaccine. 2010 April 19; 28(18): 3112–3117. doi:10.1016/j.vaccine.2010.02.062.

Bottleneck Effects on Vaccine-Candidate Antigen Diversity of Malaria Parasites in Thailand

Somchai Jongwutiwes^{a,1}, Chaturong Putaporntip^a, and Austin L. Hughes^{b,1}

^aMolecular Biology of Malaria and Opportunistic Parasites Research Unit, Department of Parasitology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

^bDepartment of Biological Sciences, University of South Carolina, SC29208, USA

Abstract

A number of cell surface antigens of the infective stages of malaria parasites (genus *Plasmodium*) have been proposed as vaccine candidates, but high levels of polymorphism at the loci encoding these antigens are problematic for vaccine effectiveness. In order to test for the effects of anti-malarial control measures (including drugs and vector control) on polymorphism at antigen-encoding loci, we analyzed sequences of four antigen-encoding loci from *P. vivax* and two from *P. falciparum* collected in 2006–2007 from two areas of Thailand: (1) the NW, where malaria cases have remained high until recently; and (2) the South, where control measures have caused a dramatic decline in numbers of cases since 1990. Polymorphism in non-repeat regions of antigen-encoding loci was dramatically reduced in the South compared to the NW. These results suggest a two-pronged strategy for malaria eradication: (1) strenuous non-vaccine control measures that will cause a severe population bottleneck in the parasite; and (2) a subsequent local vaccine focused on one or a few locally occurring alleles at antigen-encoding loci.

Keywords

DNA sequence polymorphism; malaria eradication; *Plasmodium falciparum*; *Plasmodium vivax*; vaccine candidate antigens

1. Introduction

Malaria (caused by parasites of the genus *Plasmodium*) is responsible for deaths of 1–2 million humans a year, mostly children, making global eradication a public health priority and accelerating the search for an effective vaccine [1–2]. *Plasmodium* parasites express on surfaces of infective stages (the sporozoite and merozoite) a number of antigenic proteins that elicit an immune response on the part of the vertebrate host. These include the circumsporozoite protein (CSP) expressed on the surface of the sporozoite, the stage that infects the vertebrate host. In the 1960's, it was demonstrated that x-irradiated sporozoites confer protective immunity in mice [3]; and the cloning of the gene encoding CSP from the monkey malaria

© 2010 Elsevier Ltd. All rights reserved.

1Corresponding authors: A.L. Hughes (phone: 1-803-777-9186; fax: 1-803-777-4002; austin@biol.sc.edu) and S. Jongwutiwes (phone: 662-2564000 ext 3685; fax: 662-2525944; jongwutiwes@gmail.com).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

parasite *P. knowlesi* [4] led to hopes that the homologous protein might form the basis of a vaccine against human malaria parasites. The pace of clinical trials of vaccines based on CSP and other malaria surface proteins from the two most widespread human malaria parasites, *P. falciparum* and *P. vivax*, has increased dramatically in the past decade, but so far the results have been mixed [2].

One of the major challenges facing vaccine developers is the high level of naturally occurring polymorphism at several of the loci encoding surface proteins of *P. falciparum* and *P. vivax* [5]. In the case of the CSP of *P. falciparum*, polymorphic variants in epitopes for host CD4+T cell recognition have been shown not to be cross-reactive [6], implying that vaccines which rely on the use of these epitopes to stimulate an immune response will fail to provide protection against all naturally occurring parasite variants [5]. At the CSP locus of *P. falciparum*, there is evidence that the polymorphism in T cell epitopes is maintained by balancing selection driven by host T cell recognition [7–10]. Likewise, several other loci encoding malaria cell surface proteins show evidence of selectively maintained polymorphism [8,11–13].

Even under balancing selection, because of the role of genetic drift, the level of polymorphism that can be maintained is expected to be a function of the effective population size [14–15]. Consistent with theoretical expectations, there is evidence that population bottlenecks can effect the level of polymorphism at antigen-encoding loci of malaria parasites. For example, the locus encoding apical membrane antigen-1 (AMA-1) of *P. vivax* shows considerably reduced polymorphism in Brazil in comparison to the Old World, reflecting a bottleneck in colonization of the New World [10,16]. Likewise, studies of *P. falciparum* populations on Pacific islands have revealed relatively low levels of polymorphism at several antigen loci, as expected in the case of founder effects in the colonization of islands by the parasite [17–18]. On the other hand, local populations in Old World mainland areas where malaria has long been present, such as Southeast Asia, have revealed substantial levels of polymorphism at antigen-encoding loci [9,10,12,19]. Given these high levels of polymorphism, the design of a locality-specific vaccine that provides immunity against all locally occurring variants seems problematic.

In Thailand, where malaria remains a significant public health problem in rural areas, the Anti-Malaria Programme has employed a number of non-vaccine strategies have been employed to reduce malaria infection, including anti-malarial drugs; insecticides against mosquito vectors; and reduction of mosquito breeding habitat [20–21]. These strategies have produced striking reductions in the reported number of human malaria cases in Thailand over the past 30 years, although there have been regional differences with respect to the extent of the reduction. Epidemiological evidence of declining numbers of cases suggest that control measures may be able to produce substantial reductions in local parasite effective population sizes of malaria parasite species, which in turn might cause reduction in the level of parasite polymorphism. Thus, after extensive mobilization of non-vaccine control measures, a local population may have sufficiently reduced polymorphism that a location-specific vaccine might be feasible and effective.

We tested the hypothesis that control measures can induce a loss of polymorphism at antigen-encoding loci by examining data on numbers of *P. falciparum* and *P. vivax* infections and nucleotide sequence polymorphism at selected antigen-encoding loci in two areas of Thailand. We compared data from two different regions: (1) Tak Province, in northwestern Thailand, along the border of Myanmar (henceforth NW); and (2) from Yala and Narathiwat Provinces in southern Thailand (henceforth South; Figure 1). Reported cases of malaria have declined sharply in the South over the past three decades, but less sharply in the NW [19,21]. By comparing sequence polymorphism at antigen-encoding loci, we tested the hypothesis that the

more severe decline in malaria cases in the South has been accompanied by a reduction in polymorphism at these vaccine candidate loci.

2. Methods

2.1. Parasite populations

We randomly recruited blood samples from symptomatic malaria patients from northwestern (Tak Province) and southern Thailand (Yala and Narathiwat Provinces) collected during 1996–1997 for *P. falciparum* samples and 2006–2007 for both *P. falciparum* and *P. vivax* samples. The ethical aspects of this study have been approved by the Institutional Review Board of Faculty of Medicine, Chulalongkorn University.

2.2. DNA preparation, amplification and sequencing

DNA was extracted from either venous blood samples using QIAamp kit (Qiagen, Hilden, Germany) or finger-pricked blood spotted onto filter paper. We excluded multiple clone infections of *P. falciparum* isolates by genotyping of polymorphic block 2 of the merozoite surface protein-1 (*Pfmsp-1*) and the central repeat region of the merozoite surface protein-2 (*Pfmsp-2*) genes as described by others [22]. Likewise, genotyping of *P. vivax* isolates was performed using the highly polymorphic block 6 of the merozoite surface protein-1 (*Pvmsp1*) [23]. Further, samples showing superimposed electropherogram signals during DNA sequencing were also excluded from analysis. The complete nucleotide sequences of *P. falciparum* *csp* and *msp-2* and of *P. vivax* *msp-1*, *ama-1* and *msp-4* were obtained by using respective forward and reverse primers for each gene as described previously [10,12,19,23, 24]. Sequences of *P. falciparum* *csp* and *msp-2* from the NW and of *P. vivax* *ama-1*, *msp-4* and *msp-5* from both NW and South were from our previous analyses [10,12,19,24]. The complete 128 nucleotide sequences of *Pvmsp-1* were obtained following the methods as previously described [23]. The complete 126 *P. vivax* *msp-5* sequences spanning ~1.5 kb was amplified using a forward primer (PvMsp-5-F: TCTTCAATTTTCCGCTCAACC) and a reverse primer (PvMsp-5-R: CACAAGGTGAAGAGATCGAC) which were derived from 5' and 3' untranslated regions, respectively. DNA amplification was carried out 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 included the preamplification denaturation at 94°C for 1 min followed by 35 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 2 min, and a final extension at 72°C for 5 min. DNA amplification was performed by using a GeneAmp 9700 PCR thermal cycler (Applied Biosystems, Foster City, CA). The PCR product was purified by using QIAquick PCR purification kit (QIAGEN, Germany). DNA sequences were determined directly and bi-directionally from PCR-purified templates. Sequencing analysis was performed on an ABI3100 Genetic Analyzer using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA). Overlapping sequences were obtained by using sequencing primers. Whenever singleton substitution occurred, sequence was re-determined using PCR products from two independent amplifications from the same DNA template primers. Accession numbers for all sequences used in analyses are shown in Supplementary Table S1. Numbers of sequences for each locus from each endemic area are listed in Table 1.

2.3. Statistical methods

Non-repeat portions of coding sequences were aligned using the CLUSTAL X program [25]. Alignment in repeat regions of malaria antigens is uncertain because of rapid expansion and contraction of repeat arrays, apparently by a slipped-strand mispairing-like mechanism [9,10, 12]. Therefore, we excluded from sequence comparisons repeat regions of *P. vivax* *msp1*, *P. vivax* *msp4*, *P. vivax* *msp5*, *P. falciparum* *csp*, and *P. falciparum* *msp2*. The excluded repeat

regions of *P. vivax msp1* corresponded to blocks 2, 6, 8 and 9 as defined by Putaporntip et al. [23]. The excluded repeat regions of *P. vivax msp4* were repeat array 1 (in exon 1) and repeat array 2 (in exon 2) identified by Putaporntip et al. [24]. The excluded region of *P. vivax msp5* was the single central charged amino acid residue-rich repeat region [26]. In the case of *P. falciparum csp*, the excluded region corresponded to the central array of NANP repeats; thus, the CD4 T-cell epitopes in the C-terminal non repeat portion of the protein were included [7,10]. In *P. falciparum msp2*, the repeat regions differ between the two major allelic families, 3D7 and FC27 [12]; therefore, members of these two allelic families were analyzed separately, excluding the family-specific repeat regions in each case.

In the non-repeat regions, we used Nei and Gojobori's [27] method to estimate the number of synonymous substitutions per synonymous site (d_S) and the number of nonsynonymous substitutions per nonsynonymous site (d_N). In preliminary analyses, more complicated methods [28,29] yielded essentially identical results, as expected because the number of substitutions per site was low in this case [30]. We computed the mean of all pairwise d_S values, designated the synonymous nucleotide diversity (π_S); and the mean of all pairwise d_N values, designated the nonsynonymous nucleotide diversity (π_N). Standard errors of π_S and π_N were estimated by the bootstrap method [30]; 1000 bootstrap samples were used. In computing π_S and π_N , we excluded from all pairwise comparisons any codon at which the alignment postulated a gap in any sequence.

We estimated the haplotype diversity in non-repeat regions of the antigen-encoding loci by the formula:

$$1 - \sum_{i=1}^n x_i^2$$

where n is the number of distinct haplotypes and x_i is the sample frequency of the i^{th} haplotype (Ref. ³¹, p. 177). We used a randomization method to test whether the numbers of haplotypes and haplotype diversity differed between the NW and South. For a given locus, let N be the number of sequences available from the NW and M be the number of sequences available from the South. We created 1000 pseudo-data sets by sampling (with replacement) M sequences from the N sequences collected from the NW. We then computed the numbers of haplotypes and the haplotype diversity for each pseudo-data set, and compared the real values with those computed for the pseudo-data sets.

3. Results

3.1. Malaria incidence

Numbers of cases of both *P. falciparum* and *P. vivax* showed an overall downward trend in both the NW and the South between 1979 and 2008, interrupted by several sharp peaks (Figure 2). For example, there were peaks of *P. falciparum* cases in both the NW and the South in 1984; and *P. falciparum* cases peaked again in the NW in 1990 and in the South in 1989 (Figure 2A). Likewise, in the case of *P. vivax*, there were peaks in the NW in 1989–1991 and 1997–2001, while in the South there was a sharp peak in 1989 (Figure 2B). In spite of fluctuations, in the South both *P. falciparum* and *P. vivax* had declined to less than 5000 cases per year by 1990, and this level was maintained every year through 2008 (Figure 2). On the other hand, in the NW, infections with both parasites fell below 5000 only in 2004 (Figure 2). Thus, the sharp reduction in cases of both *P. falciparum* and *P. vivax* malaria occurred over a decade earlier in the South than in the NW and was thus sustained for a much longer time.

In the South, the patterns of fluctuation in the two parasites were very similar (Figure 2). In fact, in the South the correlation between the number of *P. falciparum* cases and the number of *P. vivax* cases was remarkably close ($r = 0.927$; $P < 0.001$; Figure 3B). The corresponding correlation in the case of the NW was also positive and significant ($r = 0.615$; $P < 0.001$; Figure 3A). However, the relationship showed considerably more scatter in the NW (Figure 3A) than in the South (Figure 3B); as a consequence, there was a significant difference between the two correlation coefficients (2-tailed $P = 0.001$).

3.2. Sequence polymorphism

We examined polymorphism at four antigen-encoding loci of *P. vivax* and two of *P. falciparum* in collections from the NW and South from 2006–2007. In the non-repeat regions of *P. vivax* *ama-1*, *msp1*, *msp4*, and *msp5*, the numbers of haplotypes and haplotype diversities were strikingly reduced in the South compared to the NW (Table 1). In every case except *msp4*, the number of haplotypes was significantly lower in the South than in the NW; and at *msp4*, there were only two haplotypes in the South, as compared to six in the NW (Table 1). The relatively low diversity at *msp4* is consistent with relatively low diversity at this locus elsewhere in the world [22]. Moreover, at all four *P. vivax* loci, the haplotype diversity was significantly lower in the South than in the North (Table 1).

At the *csp* and *msp2* loci of *P. falciparum*, both the number of haplotypes and the haplotype diversity were significantly reduced in the South in comparison to the NW (Table 1). In fact there was only a single haplotype in the non-repeat region of *msp2* in the South. Alleles at *msp2* of *P. falciparum* fall into two very distinctive allelic families, designated 3D7 and FC27, which have very divergent sequences both in the repeats and in portions of the non-repeat region [10]. Only the 3D7 family was found in 82 sequences sampled from the South.

In *P. vivax*, nonsynonymous nucleotide diversity (π_N) in non-repeat regions was significantly lower in the South than in the NW at all loci except *msp4* (Table 2). In *P. falciparum*, π_N at *csp* was significantly lower in the South than in the NW (Table 2). Only the 3D7 family of *P. falciparum* *msp2* alleles was found in both NW and South, and there were no synonymous or nonsynonymous polymorphisms at this locus in the South (Table 2). Among the 3D7 family alleles at *msp2* in the NW, π_N was significantly greater than zero and thus significantly than the value for the South (Table 2). At *msp1* of *P. vivax*, synonymous nucleotide diversity (π_S) was significantly lower in the South than in the NW (Table 2). The other loci examined did not show significant differences between NW and South with respect to π_S , but π_S in the NW was in every case greater than or equal to that in the South (Table 2).

Reduction of diversity in sequences from the South was seen in repeat regions as well as non-repeat regions. For example, of the 88 *P. falciparum* *msp2* sequences from the South, all of which had the same haplotype in the non-repeat regions, only two showed distinct sequences in the repeat regions. These two divergent sequences differed from the others only by the number of repeats in the region previously designated as the Gly-Ala-rich repeats, differences potentially caused by recent events of slipped-strand mispairing [12]. The more abundant of the two haplotypes in the non-repeat regions of *P. falciparum* *csp* was associated with identical NANP repeats at the amino acid level in all 85 sequences from the South that showed this haplotype. The only difference among the repeat regions seen in these 85 sequences was a single synonymous point mutation seen in just one sequence.

4. Discussion

In the South of Thailand (Yala and Narathiwat Provinces), where there has been an approximately two decade-long reduction in the number of reported cases of both *P. falciparum* and *P. vivax* as a result of a concerted anti-malaria campaign, our results showed

that there is also reduced nucleotide sequence diversity at antigen-encoding loci. Haplotype diversities in non-repeat regions were dramatically lower in the South than in the NW of Thailand (Tak Province), significantly lower than expected if the former represented a random sample of the latter. In the South, all antigen-encoding loci showed only a small number of haplotypes in non-repeat regions. Most strikingly, at *msp2* of *P. falciparum*, only a single haplotype was found in 83 sequences sampled from the South, whereas there were 40 haplotypes in 195 sequences sampled from the same locus in the NW.

Several lines of evidence suggest that reduced sequence diversity in the South compared to the NW is due to population bottlenecks in the parasites caused by control measures. First, epidemiological data showed a decline in numbers of cases of both *P. falciparum* and *P. vivax* that began a decade earlier and thus has persisted longer in the South than in the NW. Second, the numbers of cases per year for *P. falciparum* and *P. vivax* were highly correlated in the South, suggesting that populations of both parasites were responding to the same environmental factors. Moreover, epidemiological studies have previously noted the relatively slow progress of anti-malaria measures in the NW, which have been attributed largely to population movement across border with Myanmar, exacerbated by unstable political situations [21,32]. Since insecticides have played a major role in the malaria control measures in Thailand [21], a population bottleneck in their vectors has likely been the major factor in causing population bottlenecks in *P. falciparum* and *P. vivax*.

Our evidence that genetic diversity in the NW has not been reduced is consistent with the epidemiological data and thus supports the conclusion that parasite genetic diversity can be impacted by control measures. Data on the numbers of malaria cases showed evidence that the anti-malaria campaign had begun to have a major impact in the NW after about 2004, representing about a decade and a half time lag relative to the South. Thus, the South had experienced a bottleneck for over a decade longer than the NW. Since a prolonged bottleneck is required in order to cause a substantial reduction in genetic polymorphism [33], the absence of an appreciable reduction in polymorphism in the NW is consistent with the hypothesis that the more prolonged bottleneck in the South has been the major factor in reducing polymorphism of *P. falciparum* and *P. vivax* in the latter region.

An alternative explanation for reduction in polymorphism at these loci might involve a selective sweep due to newly arisen directional selection favoring one or a few alleles over others [34]. Because the circumsporozoite protein is expressed in the sporozoite stage involved in transmission from the insect to vertebrate host, it has been proposed that interactions with the mosquito host may exert selective pressure on the *csp* locus [35]. Because of the use of insecticides in Thailand [21], there may have been changes in allele frequency within vector populations at loci that affect parasite-vector interactions. On the other hand, the fact that positive selection on the *P. falciparum* *csp* gene is focused on T-cell epitopes supports the hypothesis that polymorphism in that region is maintained largely by interactions with the human host, not the vector [7,10]. Note that the T-cell epitopes contain most of the polymorphism in non-repeat regions of the *csp* gene, where we found substantially reduced polymorphism in the South of Thailand.

In addition, the pattern of reduced polymorphism in the South was seen also at loci encoding merozoite proteins, which are not expressed in the mosquito host. Thus, interactions with the immune system of the host seem the most plausible source of balancing selection maintaining polymorphism at the loci examined here [8,11–12]. It is difficult to imagine any factor that could have caused directional selection to replace balancing selection at these loci just since 1990 and only in the South. Moreover, it seems very unlikely that selective sweeps would have occurred independently at the same time at numerous different loci in two different *Plasmodium* species. Thus, the overall pattern is much more easily explained on the bottleneck

hypothesis than on the hypothesis of selective sweeps, although we cannot rule out the possibility that the latter may have occurred at certain individual loci.

The reduction of polymorphism at antigen-encoding loci supports the prediction that, even where balancing selection acts to maintain polymorphism at antigen-encoding loci of malaria parasites, bottleneck effects can severely limit diversity of local populations [16–18]. Moreover, our evidence that anti-malarial measures can cause dramatic population bottlenecks with subsequent loss of genetic diversity at vaccine-candidate loci suggests a two-pronged strategy for malaria eradication: (1) strenuous non-vaccine control measures that will cause a severe population bottleneck in the parasite; and (2) a subsequent local vaccine focused on one or a few locally occurring alleles at antigen-encoding loci. By reducing or eliminating parasite polymorphism prior to vaccine deployment, such an approach would be able to overcome the problems posed for antigen-based vaccines by polymorphism of parasite populations [5].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to all patients who donated their blood samples for this study and to Thongchai Hongsriruang, Sunee Seethamchai, Pannadhat Areekul, Ratiporn Kosuwin, Urassaya Pattanawong, Teerayot Kobasa and the staff of the Bureau of Vector Borne Disease, Department of Disease Control, Ministry of Public Health, Thailand, for assistance in field work. This research was supported by grants from the National Research Council of Thailand and the Thai Government Research Budget to S.J and C.P.; The Thailand Research Fund (RMU5080002) to C.P.; and from the National Institutes of Health (GM43940) to A.L.H.

References

1. Breman JG. Eradicating malaria. *Sci Prog* 2009;92:1–38. [PubMed: 19544698]
2. Sharma S, Pathak S. Malaria vaccine: a current perspective. *J Vector Borne Dis* 2008;45:1–20. [PubMed: 18399312]
3. Nussenzweig R, Vanderberg J, Most H. Protective immunity produced by the injection of x-irradiated sporozoites of *Plasmodium berghei*. IV. Dose response, specificity and humoral immunity. *Mil Med* 1969;134:1176–1182. [PubMed: 4987036]
4. Ellis J, Ozaki LS, Gwadz RW, Cochrane AH, Nussenzweig V, et al. Cloning and expression in *E. coli* of the malarial sporozoite surface antigen gene from *Plasmodium knowlesi*. *Nature* 1983;302:536–538. [PubMed: 6339949]
5. Good MF. The hope but challenge for developing a vaccine that might control malaria. *Eur J Immunol* 2009;39:939–943. [PubMed: 19350585]
6. Zevering Y, Khamboonruang C, Good MF. Natural amino acid polymorphisms of the circumsporozoite protein of *Plasmodium falciparum* abrogate specific human CD4+ T cell responsiveness. *Eur J Immunol* 1994;24:1418–1425. [PubMed: 7911423]
7. Hughes AL. Circumsporozoite protein genes of malaria parasites (*Plasmodium* spp.): evidence for positive selection on immunogenic regions. *Genetics* 1991;127:345–353. [PubMed: 1706291]
8. Hughes MK, Hughes AL. Natural selection on *Plasmodium* surface proteins. *Mol Biochem Parasitol* 1995;71:99–113. [PubMed: 7630387]
9. Jongwutiwes S, Tanabe K, Hughes MK, Kanbara H, Hughes AL. Allelic variation in the circumsporozoite protein of *Plasmodium falciparum* from Thai field isolates. *Am J Trop Med Hyg* 1994;51:659–668. [PubMed: 7985759]
10. Putaporntip C, Jongwutiwes S, Hughes AL. Natural selection maintains a stable polymorphism at the circumsporozoite protein locus of *Plasmodium falciparum* in a low endemic area. *Infect Genet Evol* 2009;9:567–573. [PubMed: 19460323]
11. Hughes AL. Positive selection and interallelic recombination at the merozoite surface antigen-1 (MSA-1) locus of *Plasmodium falciparum*. *Mol Biol Evol* 1992;9:381–393. [PubMed: 1584009]

12. Putaporntip C, Jongwutiwes S, Hughes AL. Differential selective pressures on the merozoite surface protein 2 locus of *Plasmodium falciparum* in a low endemic area. *Gene* 2008;427:51–57. [PubMed: 18840512]
13. Verra F, Hughes AL. Evidence for ancient balanced polymorphism at the Apical Membrane Antigen-1 (AMA-1) locus of *Plasmodium falciparum*. *Mol Biochem Parasitol* 2000;105:149–153. [PubMed: 10613707]
14. Maruyama T, Nei M. Genetic variability maintained by mutation and overdominant selection in finite populations. *Genetics* 1981;98:441–459. [PubMed: 17249094]
15. Takahata N, Nei M. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* 1990;124:967–978. [PubMed: 2323559]
16. Grynberg P, Fontes CJF, Hughes AL, Braga ÉM. Polymorphism at the Apical Membrane Antigen-1 locus reflects the world population history of *Plasmodium vivax*. *BMC Evol Biol* 2008;8:123. [PubMed: 18445274]
17. Tanabe K, Sakihama N, Kaneko A. Stable SNPs in malaria antigen genes in isolated populations. *Science* 2004;303:493. [PubMed: 14739451]
18. Sakihama N, Ohmae H, Bakote'e B, Kawabata M, Hirayama K, Tanabe K. Limited allelic diversity of *Plasmodium falciparum* merozoite surface protein gene 1 from populations in the Solomon Islands. *Am J Trop Med Hyg* 2006;74:31–40. [PubMed: 16407343]
19. Putaporntip C, Jongwutiwes S, Grynberg P, Cui L, Hughes AL. Nucleotide sequence polymorphism at the apical membrane antigen-1 locus reveals population history of *Plasmodium vivax* in Thailand. *Infect Genet Evol* 2009;9:1295–1300. [PubMed: 19643205]
20. Malikul S. The current situation of the anti-malaria programme in Thailand. *Southeast Asian J Trop Med Pub Health* 1988;19:355–359. [PubMed: 3064319]
21. Chareonviriyaphap T, Bangs MJ, Ratanatham S. Status of malaria in Thailand. *Southeast Asian J Trop Med Pub Health* 2000;31:225–237. [PubMed: 11127318]
22. Snounou G, Zhu X, Siripoon N, Jarra W, Thaithong S, Brown KN, Viriyakosol S. Biased distribution of msp1 and msp2 allelic variants in *Plasmodium falciparum* populations in Thailand. *Trans R Soc Trop Med Hyg* 1999;93:369–374. [PubMed: 10674079]
23. Putaporntip C, Jongwutiwes S, Sakihama N, Ferreira MU, Kho WG, Kaneko A, Kanbara H, Hattori T, Tanabe K. Mosaic organization and heterogeneity in frequency of allelic recombination of the *Plasmodium vivax* merozoite surface protein-1 locus. *Proc Natl Acad Sci USA* 2002;99:16348–16353. [PubMed: 12466500]
24. Putaporntip C, Jongwutiwes S, Ferreira MU, Kanbara H, Udomsangpeth R, Cui L. Limited global diversity of the *Plasmodium vivax* merozoite surface protein 4 gene. *Infect Genet Evol* 2009;9:821–826. [PubMed: 19409511]
25. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Diggins DG. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997;25:4876–4882. [PubMed: 9396791]
26. Gomez A, Suarez CF, Martinez P, Saravia C, Patarroyo MA. High polymorphism in *Plasmodium vivax* merozoite surface protein-5 (MSP5). *Parasitology* 2006;133:661–672. [PubMed: 16978450]
27. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 1986;3:418–426. [PubMed: 3444411]
28. Li WH. Unbiased estimates of the rates of synonymous and nonsynonymous substitution. *J Mol Evol* 1993;36:96–99. [PubMed: 8433381]
29. Yang Z, Nielsen R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* 2000;17:32–43. [PubMed: 10666704]
30. Nei, M.; Kumar, S. Molecular evolution and phylogenetics. New York: Oxford University Press; 2000.
31. Nei, M. Molecular evolutionary genetics. New York: Columbia University Press; 1987.
32. Komchom S, Singhasivanon P, Kaewkungwal J, Chupraphawan S, Thimasarn K, Kidson C, Rojanawatsirivet C, Yimsamran S, Looareesuwan S. Trend of malaria incidence in highly endemic provinces along the Thai borders, 1991–2001. *Southeast Asian J Trop Med Public Health* 2003;34:486–494. [PubMed: 15115117]

33. Nei M, Maruyama T, Chakraborty R. The bottleneck effect and genetic variability in populations. *Evolution* 1975;29:1–10.
34. Maynard Smith J, Haigh J. The hitch-hiking effect of a favorable gene. *Genet. Res* 1974;23:23–35. [PubMed: 4407212]
35. Kumkhaek C, Phra-ek K, Rénia L, Singhasivanon P, Looareesuwan S, Hirunpetcharat C, White NJ, Brockman A, Grüner AC, Lebrun N, Allouche A, Nosten F, Khusmith S, Snounou G. Are extensive T cell epitope polymorphisms in the *Plasmodium falciparum* circumsporozoite antigen, a leading sporozoite vaccine candidate, selected by immune pressure? *J Immunol* 2005;175:3935–3939. [PubMed: 16148140]

Figure 1.
Map of Thailand indicating the two regions (NW and South) from which sequences were collected.

Figure 2.

Reported numbers of indigenous human cases of (A) *P. falciparum* and (B) *P. vivax* in the NW and South from 1978–2008 (source: Division of Vector-Borne Diseases, Ministry of Public Health, Thailand).

A)

B)

Figure 3.

Plots of the number of *P. vivax* cases vs. the number of *P. falciparum* cases in (A) NW ($r = 0.615$; $P < 0.001$); and (B) South ($r = 0.927$; $P < 0.001$).

Table 1
Numbers and diversity of haplotypes in non-repeat regions of *Plasmodium* surface antigen loci in two regions (NW and South) of Thailand.

Species	Locus	NW				South		
		No. sequences	No. haplotypes	Haplotype diversity	No. sequences	No. haplotypes	Haplotype diversity	
<i>P. vivax</i>	<i>ama-1</i>	102	75	0.964	73	4***	0.436***	
	<i>msp1</i>	66	48	0.965	67	6***	0.335***	
	<i>msp4</i>	77	6	0.236	53	2	0.037**	
	<i>msp5</i>	73	11	0.573	53	3***	0.358***	
<i>P. falciparum</i>	<i>csp</i>	223	9	0.489	88	2***	0.065***	
	<i>msp2</i>	195	40	0.880	82	1***	0.000***	

Randomization tests of the hypothesis that the number of haplotypes or haplotype diversity in the South equals that in the NW:

** P < 0.01;

*** P < 0.001.

Table 2

Synonymous (π_S) and nonsynonymous (π_N) nucleotide diversity ($\pm$ S.E.) in non-repeat regions of *Plasmodium* surface antigen loci in two regions (NW and South) of Thailand.

Species	Locus	NW		South	
		π_S	π_N	π_S	π_N
<i>P. vivax</i>	<i>ama-1</i>	0.0084 $\pm$ 0.0027	0.0077 $\pm$ 0.0017	0.0036 $\pm$ 0.0018	0.0035 $\pm$ 0.0010 *
	<i>msp1</i>	0.0425 $\pm$ 0.0043	0.0200 $\pm$ 0.0015	0.0151 $\pm$ 0.0020 ***	0.0067 $\pm$ 0.0007 ***
	<i>msp4</i>	0.0000 $\pm$ 0.0000	0.0006 $\pm$ 0.0004	0.0000 $\pm$ 0.0000	0.0001 $\pm$ 0.0001
	<i>msp5</i>	0.0025 $\pm$ 0.0016	0.0026 $\pm$ 0.0010	0.0024 $\pm$ 0.0023	0.0000 $\pm$ 0.0000 *
	<i>csp</i>	0.0000 $\pm$ 0.0000	0.0035 $\pm$ 0.0013	0.0000 $\pm$ 0.0000	0.0005 $\pm$ 0.0002 *
<i>P. falciparum</i>	<i>msp2</i>				
	<i>3D7</i>	0.0006 $\pm$ 0.0004	0.0143 $\pm$ 0.0041	0.0000 $\pm$ 0.0000	0.0000 $\pm$ 0.0000 ***
	<i>FC27</i>	0.0010 $\pm$ 0.0007	0.0045 $\pm$ 0.0021	--	--

Z-tests of the hypothesis that π_S or π_N in the South equals the corresponding value in the NW:

* $P < 0.05$;
*** $P < 0.001$.