E)- Student survey on malaria knowledge, malaria prevention and treatment seeking

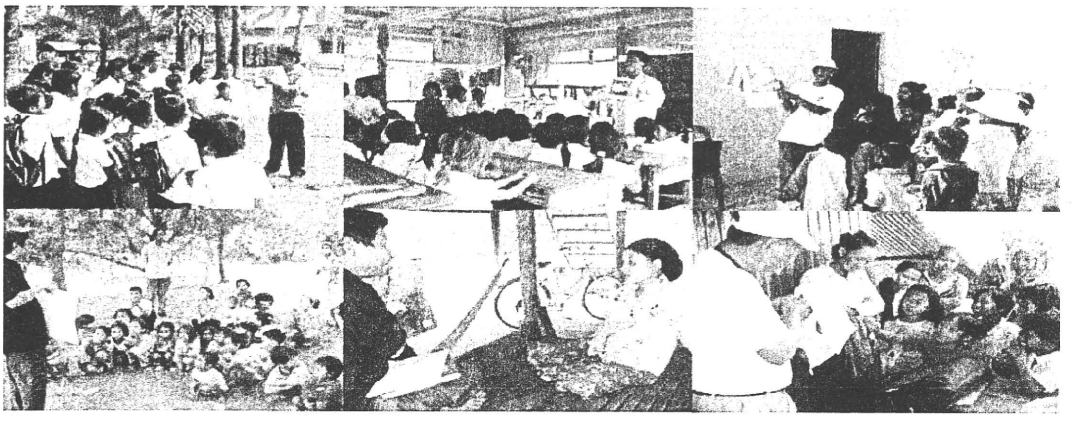
The survey team also randomly selected 40% (224 students) among the 560 students screened and administered the questionnaire to them. The questions were mainly asked related to the malaria prevention, sign and symptom of malaria and treatment seeking when gets malaria. 100% of the students recognized that malaria causes by mosquito bites. More than 90% of them knew the common malaria signs such as fever, chill and sweat. When asking where to go if they get malaria all most all said that they will go to health center and only few of them mentioned that they go to private provider. Regarding the question how to prevention from malaria, all of them stated that bed net use could protect them from malaria, especially the treated insecticide net. According to the interview with the school teachers and school director, the malaria topic was introduced in the study program and regularly educates them by their teachers and through the integrated health activities for instance vaccination campaign, malaria campaign, malaria education during the bed net distribution, especially the message and malaria information were disseminated by the village health volunteer etc....

F- Strengthen community's knowledge & practice through active health education

The health education has played a key role in malaria prevention for the remote and poor communities, particularly to strengthen and enhance the communities' knowledge and responsiveness for participating as the malaria control activity at the village level. Raising awareness through health education for malaria prevention is one of the important components in the implementation project's villages to prevent the villagers from mosquito bite and transmission to other people. Many IEC/BCC materials such as posters, leaflets, flipcharts, T-shirts, calendar etc... were distributed to the villagers during the implementation of the project. Weekly health educations have been conducted to the villagers throughout the personal contact at home, group education at school, in the filed or during the special events organized at the community.

During the project period, there is the total of 350 health education sessions were provided to the study villages and more than 13,000 villagers educated through the sessions. Leaflets, flipcharts and posters were also distributed via village malaria collectors as part of the health education strategy. Due to the current survey showed that more than 90% of the villagers know the malaria cause, their signs and how to present them selves from malaria.

Picture 23-28: Activities of the health education session provided to the community



VII. Conclusions

- Thanks to the project implementation and monitoring, the malaria epidemiological trend in the study villages has significantly reduced in the third year of the project, especially the Plasmodium Falciparum species compare to the first year and the same period of the project execution even more population movement has been observed during the last few years to the study villages.
- Malaria incidence in most of the study villages has also notably decreased; particularly in the new study villages compare to the old study villages with the average malaria infection around 2-4 malaria cases per village per month.
- No any malaria case has been reported since the onset of the project introduction until the third years of the projection implementation if compare to the situation before the project instigation that many malaria deaths were reported.
- The project implementation has extensively increased the community participation and considerably raised the awareness of the villagers about the malaria knowledge, prevention and treatment seeking behavior by strengthening the community's knowledge & practice through active health education and IEC distribution regularly. So all of the villagers come to village volunteers for diagnosis and treatment as soon as they can when they think or suspect that they get malaria.
- The project implementation has promoted the community responsiveness of the personal prevention and protection versus malaria transmission through active participation in the insecticide bed net distribution, especially make them active involve in the bed net re-impregnation sessions conducted by the volunteers that was proven by the increasing demand of the quantity of the insecticide bed nets from year to year.

VIII. Acknowledgements

- The National Center for Parasitology, Entomology and Malaria Control would like to express the profound and sincere thanks to the Ministry of Health, Welfare and Labor, Japan, for their kind and continuous assistance to our Center.
- A heartfelt gratefulness also to National Institute of Infectious Disease, Japan for their kind technical support and excellent facilitation for the grant application.
- A profuse appreciation as well to all the collaborators and staff at all levels who have contributed their time and endeavor to make the project functioning fruitfully.

ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Jan. 2011, p. 94–100
0066-4804/11/$12.00 doi:10.1128/AAC.01156-10

Vol. 55, No. 1

Spontaneous Mutations in the *Plasmodium falciparum* Sarcoplasmic/Endoplasmic Reticulum Ca^{2+}-ATPase (PfATP6) Gene among Geographically Widespread Parasite Populations Unexposed to Artemisinin-Based Combination Therapies▽†

Kazuyuki Tanabe,[1]* Sedigheh Zakeri,[2] Nirianne Marie Q. Palacpac,[3] Manada Afsharpad,[2] Milijaona Randrianarivelojosia,[4] Akira Kaneko,[5] Aung Swi Prue Marma,[6] Toshihiro Horii,[3] and Toshihiro Mita[7]

Laboratory of Malariology, Research Institute for Microbial Diseases, Osaka University, Osaka 565-0871, Japan[1]; Malaria and Vector Research Group, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran[2]; Department of Molecular Protozoology, Research Institute for Microbial Diseases, Osaka University, Suita, Osaka, Japan[3]; Malaria Research Unit, BP 1274-Antananarivo (101), Institut Pasteur de Madagascar, Antananarivo, Madagascar[4]; Island Malaria Group, Karolinska Institutet, Department of Microbiology, Tumor and Cell Biology, Nobels vägen 16, SE-171 77 Stockholm, Sweden[5]; Emerging and Re-emerging Diseases, Communicable Disease Control, Directorate General of Health Services, Mohakhali, Dhaka-1212, Bangladesh[6]; and Department of International Affairs and Tropical Medicine, Tokyo Women's Medical University, Tokyo, Japan[7]

Received 19 August 2010/Returned for modification 21 September 2010/Accepted 11 October 2010

Recent reports on the decline of the efficacy of artemisinin-based combination therapies (ACTs) indicate a serious threat to malaria control. The endoplasmic/sarcoplasmic reticulum Ca^{2+}-ATPase ortholog of *Plasmodium falciparum* (PfSERCA) has been suggested to be the target of artemisinin and its derivatives. It is assumed that continuous artemisinin pressure will affect polymorphism of the PfSERCA gene (*serca*) if the protein is the target. Here, we investigated the polymorphism of *serca* in parasite populations unexposed to ACTs to obtain baseline information for the study of potential artemisinin-driven selection of resistant parasites. Analysis of 656 full-length sequences from 13 parasite populations in Africa, Asia, Oceania, and South America revealed 64 single nucleotide polymorphisms (SNPs), of which 43 were newly identified and 38 resulted in amino acid substitutions. No isolates showed L263E and S769N substitutions, which were reportedly associated with artemisinin resistance. Among the four continents, the number of SNPs was highest in Africa. In Africa, Asia, and Oceania, common SNPs, or those with a minor allele frequency of ≥0.05, were less prevalent, with most SNPs noted to be continent specific, whereas in South America, common SNPs were highly prevalent and often shared with those in Africa. Of 50 amino acid haplotypes observed, only one haplotype (3D7 sequence) was seen in all four continents (64%). Forty-eight haplotypes had frequencies of less than 5%, and 40 haplotypes were continent specific. The geographical difference in the diversity and distribution of *serca* SNPs and haplotypes lays the groundwork for assessing whether some artemisinin resistance-associated mutations and haplotypes are selected by ACTs.

Artemisinin-based combination therapies (ACTs) are currently the first-line treatment for uncomplicated falciparum malaria in most areas of endemicity (23, 34). The deployment of ACT has greatly reduced malaria morbidity and mortality (8). However, recently there has been accumulating evidence which suggests a decline of the efficacy of ACTs and artemisinin monotherapy in western Cambodia (7, 22). Although the molecular mechanism of the antimalarial action of artemisinin and its derivatives (artemisinins) remains to be clarified, the endoplasmic and sarcoplasmic reticulum Ca^{2+}-ATPase ortholog of *Plasmodium falciparum* (PfSERCA or PfATP6) has been suggested to be the target of artemisinins (9, 15). A replacement of L at codon 263 of the PfSERCA gene (*serca*) with E (L263E) results in abrogation of inhibition of PfSERCA by artemisinin (32). A recent allelic exchange study also showed reduced (though not significantly) susceptibility to artemisinins in parasites expressing the L263E allele (33). Mutations(s) of *serca* has also been associated with *in vitro* artemether resistance, with field isolates from French Guiana having an artemether 50% inhibitory concentration (IC_{50}) of 1.7 nM whereas that for parasites having an S769N substitution was 79.4 nM (13, 15, 16). The association of these mutations with artemisinin resistance, however, has not been confirmed for other geographic areas (2, 3, 5–7, 11, 12, 14, 18, 20, 26, 35).

Limited sequence analyses have previously shown that *P. falciparum serca* contains a number of single nucleotide polymorphisms (SNPs) (6, 14, 28). Thus, a large number of field isolates would be required to properly assess whether drug pressure imposed by continuous deployment of ACTs causes a potential selection of a mutation(s) in *serca* that is associated

* Corresponding author. Mailing address: Laboratory of Malariology, Research Institute for Microbial Diseases, Osaka University, 3-1 Yamada-Oka, Suita, Osaka 565-0871, Japan. Phone: 81 6 6879 4260. Fax: 81 6 6879 4262. E-mail: kztanabe@biken.osaka-u.ac.jp.

† Supplemental material for this article may be found at http://aac.asm.org/.

▽ Published ahead of print on 18 October 2010.

with artemisinin resistance. We consider that baseline information on *serca* polymorphism occurring before the implementation of ACTs would provide necessary information to infer whether some *serca* mutations are likely selected by ACTs. We have recently used *serca* as a genetic marker to study the geographical distribution of genetic diversity of *P. falciparum* and obtained 514 full-length *serca* sequences from nine *P. falciparum* populations in Africa, Asia, Oceania, and South America (30). In this study we newly obtained 139 full-length *serca* sequences, mostly from Malawi, Madagascar, Iran, and Bangladesh. Together with published sequences, a total of 656 sequences were analyzed for *serca* polymorphism. Importantly, all parasite isolates examined hereto were unexposed to ACTs. The results show that *serca* has many spontaneous mutations and haplotypes, which are geographically distinctive. It is predicted that further studies will reveal more SNPs/haplotypes. This information has several implications for inferring whether some *serca* mutations and haplotypes are selected by the current continuous deployment of ACTs.

MATERIALS AND METHODS

Parasite isolates. We collected *P. falciparum* isolates from Malawi, Madagascar, Iran, and Bangladesh. In Malawi, samples were collected from infected individuals in all age groups during cross-sectional surveys in June and July 2000 at two primary schools in the Salima District (4). The study was approved by the local ethics committee of the Malaria Control Programme and the Malawi Ministry of Health. In Madagascar, blood samples containing parasites were collected from consenting symptomatic outpatients in 2005 as part of the national network activities for the surveillance of drug-resistant *Plasmodium* spp. (24). Administrative authorizations and ethical clearances were provided by the Ministry of Health and the national ethics committee. Samples from rural areas of Ampasimpotsy and Saharevo were sent to the malaria research unit of the Institut Pasteur de Madagascar and kept frozen at −20°C until use. In Iran, blood samples were collected from *P. falciparum*-infected individuals, 1 to 70 years old, with symptomatic uncomplicated malaria attending the Malaria Health Center in Chabahar and the Public Health Department in Sistan and Baluchistan province, southeastern Iran, during 2001 and 2002. The study was approved by the Ethical Review Committee of Research of the Pasteur Institute of Iran. In Bangladesh, samples were collected from symptomatic malaria patients in all age groups at Bandarban district hospital from October to December 2007 (17). Approval of the study was obtained from the Bangladesh Medical Research Council and the local health regulatory body in Bandarban, Bangladesh. All isolates examined in this study were from parasite populations not exposed to ACTs (including rural areas in Bangladesh and Papua New Guinea [PNG] which have had no previous access to ACT, despite the countries' support for the WHO policy). We also used three cultured strains originally isolated from Sudan (29).

DNA sequences. Parasite genomic DNA was extracted using the QIAamp DNA blood minikit (Qiagen, Hilden, Germany). Full-length *serca* was amplified by PCR using Takara LA *Taq* polymerase (Takara Bio, Japan) in a 20-μl reaction mixture as previously described (28) with slight modifications: Primers U1 and 4099R were used (see Table S1 in the supplemental material). Forty cycles of amplification (20 s at 93°C and 5 min at 62°C) were preceded by denaturation at 93°C for 1 min and followed by a final elongation at 72°C for 10 min. The PCR product was diluted 10-fold, and a 2-μl aliquot was used as the template for a second PCR amplification of 20 cycles in a 50-μl reaction mixture using primers U1 and 4094R. The PCR products were purified using the QIAquick PCR purification kit (Qiagen). DNA sequencing was performed directly from two independent PCR products, using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems) and an ABI 3130 Genetic Analyzer (Applied Biosystems). Sequencing primers were designed to cover target regions in both directions (see Table S1 in the supplemental material). Mixed-genotype infections, judged from superimposed electropherogram peaks, were excluded from further analysis.

Sequence analyses. We obtained a total of 139 full-length sequences of *serca* coding regions from Malawi (n = 38), Madagascar (n = 19), Iran (n = 35), Bangladesh (n = 44), and Sudan (n = 3). Multiple infections were detected for 36 isolates (18.0%). Also, 514 full-length sequences that we recently published (30) were included for analysis; these were from Ghana (n = 38), Tanzania (n = 69), Thailand (n = 82), Philippines (n = 53), Papua New Guinea (PNG) (n = 89), Solomon Islands (n = 51), Vanuatu (n = 80), Brazil (n = 42), and Venezuela (n = 10). Table S2 in the supplemental material summarizes all *serca* sequences used in this study. Three additional sequences from cultured strains (3D7, Dd2, and HB3) (28) were also included. Nucleotide diversity was estimated by $\theta\pi$, the average pairwise nucleotide distance, and θ_S, the standardized number of polymorphic sites per site (Watterson's estimator), using DnaSP version 4.10 (25). Sequences were aligned using Clustal W (31) implemented in MEGA version 4.0 (27). Polymorphic sites and synonymous and nonsynonymous substitutions were determined using DnaSP.

The allele frequency of SNPs was calculated using Arlequine version 3.1 (10). We categorized SNPs as either common or uncommon, defined as those with a minor allele frequency of ≥0.05 or with a minor allele frequency of <0.05, respectively (19). (Excluding samples with multiple infection did not affect the frequency of common SNPs in this study, because analysis of 36 multiply infected samples showed a frequency of common SNPs [27/39 = 69%] that was comparable to that for 139 singly infected samples [103/139 = 74%] [P = 0.88, chi-square test].) Nucleotide and amino acid positions were numbered according to the 3D7 sequence (PlasmoDB gene identification no. PFA0310c). Amino acid haplotype diversity (h) was calculated using the formula $h = [n/(n-1)] \times (1 - \Sigma pi^2)$, where pi is the frequency of the ith *serca* amino acid haplotype (21). The variance (V) of h was calculated using a formula modified from Nei's formula for a haploid genome: $V = [2/n(n-1)]\{2[n-2][\Sigma pi^3 - (\Sigma pi^2)^2] + \Sigma pi^2 - [\Sigma pi^2]^2\}$.

Nucleotide sequence accession numbers. The sequences reported in this study have been deposited in the DNA Data Bank of Japan/European Molecular Biology Laboratory/GenBank database (accession numbers AB576210 to AB576348).

RESULTS

An alignment of 656 sequences (3,687 to 3,693 bp) revealed 62 polymorphic nucleotide sites with 64 SNPs, of which 38 resulted in amino acid substitutions (Fig. 1; see Table S3 in the supplemental material). Among the 64 SNPs, 43 SNPs (20 synonymous SNPs and 23 nonsynonymous SNPs) were newly identified in this study. Together with partial *serca* sequences obtained by other investigators, there were 110 SNPs (42 synonymous SNPs and 68 nonsynonymous SNPs) in total. In addition to SNPs, a variation in the number of Asn residues (9 to 11) in the Asn-tandem repeat region at codons 457 to 465 (of the 3D7 sequence) and a deletion of the Gly residue at codon 844 were also noted (Fig. 1; see Table S3 in the supplemental material). In our samples, we found no polymorphism at codons 263 and 769; these polymorphisms have been shown to affect the PfSERCA activity or to be associated with increased artemisinin IC_{50}s (13, 32). An E431K SNP, which has been reported to be associated with increased artesunate IC_{50}s in Senegal (13), was detected in Africa, Asia, and South America. I89T and N465K, which have been inferred not to be associated with artemisinin resistance (7), were observed only in Asia and Oceania. Amino acid replacements were largely clustered in cytoplasmic domain 3 (Fig. 1). In contrast, 10 transmembrane domains contained only two amino acid changes (at codons 67 and 1169), which had similar residue properties (i.e., basic residues K67R and hydrophobic residues V1169I).

The number of SNPs was relatively high in Africa (44 SNPs with 27 nonsynonymous SNPs) compared to other continents (10 to 21 SNPs with 5 to 10 nonsynonymous SNPs) (Table 1 and Fig. 2a). Two nucleotide diversity indices, i.e., θ_S, the number of polymorphic sites per site, and $\theta\pi$, the average number of pairwise nucleotide differences, showed substantially different levels in the four continents. Overall, θ_S was higher than $\theta\pi$ in Africa, Asia, and Oceania, whereas $\theta\pi$ was somewhat higher than θ_S in South America. This indicates that the majority of the alleles in Africa, Asia, and Oceania are

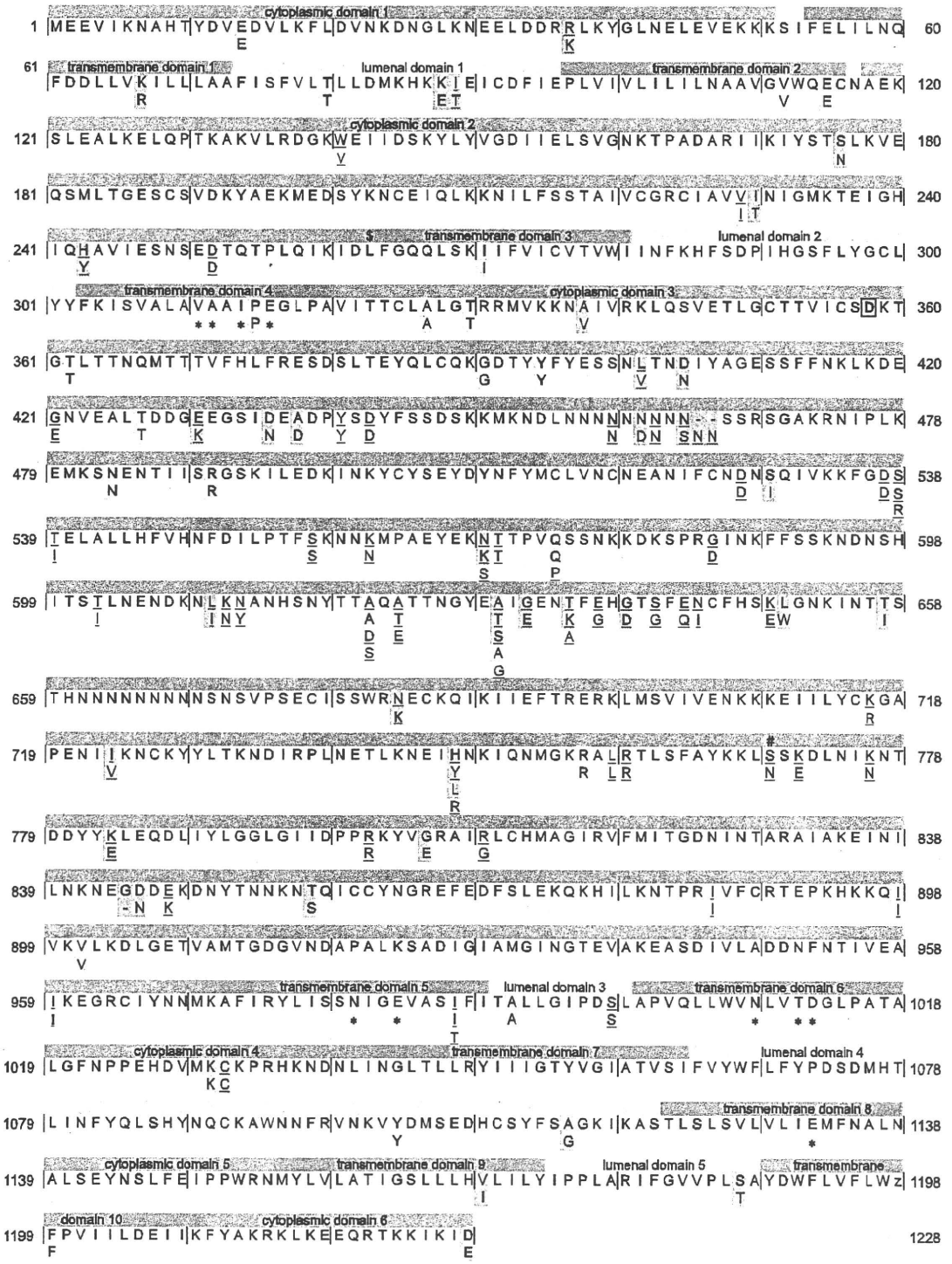


FIG. 1. Polymorphism in the *P. falciparum* SERCA (PfATP6) genes from worldwide parasite populations unexposed to artemisinin-based combination therapies. Nonsynonymous substitutions and synonymous substitutions observed in 656 samples examined in this study are highlighted in pink and blue, respectively, alongside the 3D7 sequence (PlasmoDB, PFA0310c). Substitutions reported by other investigators are underlined. Sequence regions for 10 transmembrane domains, 6 cytoplasmic domains, 5 luminal domains, putative calcium-binding sites (*), and a phosphorylation site (boxed D at 358) were inferred from the rabbit *serca* gene (Swiss-Prot Protein Data Bank [PDB] code P04191). The L at position 263, where experimental substitution to E results in abrogation of inhibition of PfSERCA by artemisinin (32), is indicated by $. The S769N substitution, which has been shown to be associated with an increased artemether IC_{50} in French Guiana (13), is indicated by #. Dashes between positions 465 and 466 and at position 844 denote deletions.

TABLE 1. SNPs and amino acid haplotypes in 656 full-length *serca* sequences from 13 *P. falciparum* populations (656 isolates)

Geographic area	*n*	No. of SNPs				Nucleotide diversity (mean ± SD)		No. of amino acid haplotypes			Haplotype diversity, *h* (mean ± SD)
		Total	Nonsynonymous	Common[b]	Continent specific	$\theta\pi$	θ_S	Total	Common[b]	Continent specific	
Worldwide[a]	656	64	38	61	50	0.00045 ± 0.00002	0.00238 ± 0.00030	50	48	40	0.579 ± 0.023
Africa	164	44	27	38	35	0.00061 ± 0.00003	0.00206 ± 0.00031	34	30	27	0.760 ± 0.033
Ghana	38	13	7	9		0.00041 ± 0.00006	0.00084 ± 0.00023	9	5		0.563 ± 0.094
Tanzania	69	31	18	26		0.00067 ± 0.00005	0.00170 ± 0.00031	19	15		0.771 ± 0.051
Malawi	38	14	10	7		0.00056 ± 0.00006	0.00090 ± 0.00024	13	8		0.801 ± 0.058
Madagascar	19	10	6	NA		0.00074 ± 0.00010	0.00078 ± 0.00025	10	NA		0.918 ± 0.036
Asia + Oceania	434	24	11	21	16	0.00028 ± 0.00002	0.00098 ± 0.00020	19	17	12	0.420 ± 0.030
Asia	214	21	10	15	9	0.00037 ± 0.00003	0.00096 ± 0.00021	17	15	8	0.581 ± 0.036
Iran	35	11	5	4		0.00070 ± 0.00005	0.00073 ± 0.00022	7	2		0.805 ± 0.032
Bangladesh	44	8	5	4		0.00025 ± 0.00004	0.00050 ± 0.00018	7	4		0.432 ± 0.091
Thailand	82	11	7	8		0.00027 ± 0.00004	0.00060 ± 0.00018	9	7		0.404 ± 0.068
Philippines	53	4	2	1		0.00024 ± 0.00003	0.00024 ± 0.00012	4	1		0.565 ± 0.040
Oceania	220	11	5	8	3	0.00018 ± 0.00003	0.00050 ± 0.00015	6	4	2	0.217 ± 0.036
PNG	89	9	5	7		0.00016 ± 0.00004	0.00048 ± 0.00016	6	5		0.211 ± 0.058
Solomon Islands	51	4	2	2		0.00010 ± 0.00004	0.00024 ± 0.00012	2	1		0.077 ± 0.050
Vanuatu	80	3	2	0		0.00024 ± 0.00005	0.00016 ± 0.00009	2	0		0.292 ± 0.055
South America	52	10	7	3	3	0.00083 ± 0.00005	0.00060 ± 0.00019	7	2	3	0.798 ± 0.024
Brazil	42	9	6	2		0.00072 ± 0.00006	0.00057 ± 0.00019	6	1		0.783 ± 0.035
Venezuela	10	4	2	NA		0.00026 ± 0.00013	0.00038 ± 0.00019	2	NA		0.200 ± 0.154

[a] Three sequences from Sudan and three sequences from cultured parasites are included. See Materials and Methods for details.
[b] Common SNPs are those with a minor allele frequency of ≥5%. Common haplotypes are those with a frequency of ≥5%. Countries for which there were fewer than 20 isolates are excluded (NA).

uncommon (allele frequency of <5%) but that this is not so in South America. The rank order of nucleotide diversity (θ_S and $\theta\pi$) was Africa > Asia > Oceania ~ South America, with the exception of $\theta\pi$ in South America (Table 1).

The geographical distribution of SNPs in *serca* was remarkably different, particularly between Africa/Asia/Oceania and South America; there were two notable features. First, common SNPs with allele frequencies of ≥0.05 were less prevalent in Africa, Asia, and Oceania, at 6/44 SNPs (14%), 6/21 SNPs (29%), and 3/11 SNPs (27%), respectively (Table 1 and

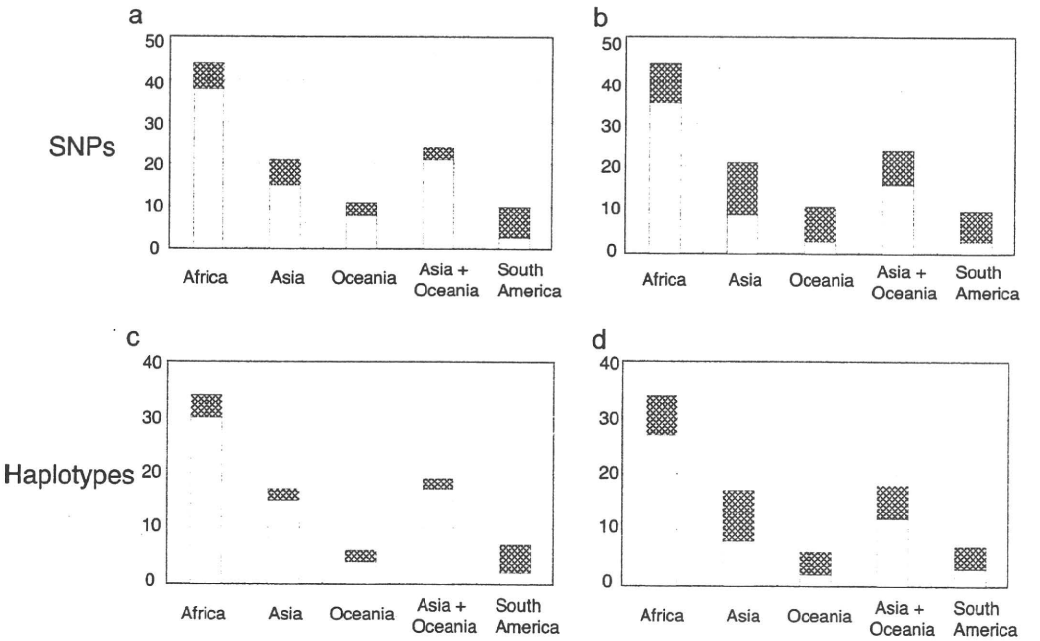


FIG. 2. Frequency distributions of SNPs and amino acid haplotypes in *serca* of *P. falciparum* isolates from Africa, Asia, Oceania, and South America. (a) Alleles in SNPs are divided into those with a minor allele frequency of <5% (open bars) and those with a minor allele frequency of ≥5% (shaded bars). (b) SNPs are divided into those that are continent specific (open bars) or not (shaded bars). (c) Amino acid haplotypes are divided into those with a minor haplotype frequency of <5% (open bars) and those with a minor haplotype frequency of ≥5% (shaded bars). (d) Amino acid haplotypes are divided into those that are continent specific (open bars) or not (shaded bars). The vertical axes represent the number of isolates.

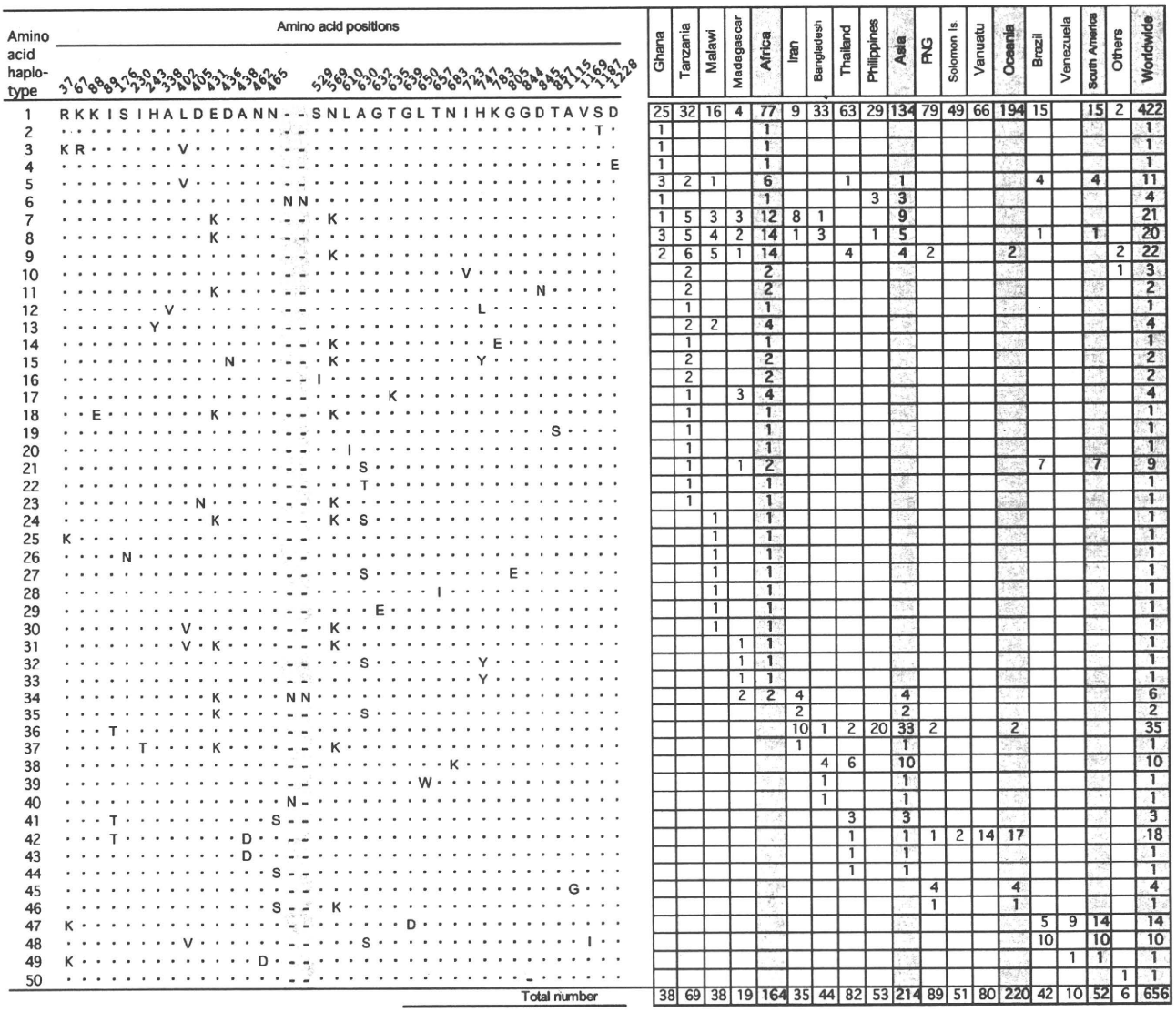


FIG. 3. Amino acid haplotypes of *P. falciparum serca* observed in 656 worldwide samples. Amino acid positions are numbered according to the 3D7 *serca* sequence (PFA0310c). Deletions of N next to position 465 are dashed. "Others," haplotypes of cultured parasites: 1, 3D7 and Dd2; 9, two Sudanese isolates; 10, one Sudanese isolate; 50, HB3.

Fig. 2a). However, in South America the prevalence of common SNPs was high (7/10 SNPs [70%]). Second, substantial numbers of SNPs were continent specific in Africa and Asia: 35/44 SNPs (80%) were found only in Africa and 9/21 SNPs (43%) in Asia (Table 1 and Fig. 2b). Most SNPs found in Oceania (8/11 [73%]) were also found in Asia, and thus when SNPs in Asia and Oceania are combined, 16/24 SNPs (67%) are specific to these geographic areas. Limited numbers of SNPs are shared between Africa and Asia/Oceania: of 24 SNPs (33%) found in Asia/Oceania, only 8 are shared with Africa. In contrast, in South America, only 3/10 SNPs (30%) found are specific to the continent, and the remaining 70% are shared with Africa.

In total, there are 50 *serca* amino acid haplotypes in our sample set (Fig. 3). Haplotype 1 (3D7 strain) is the most prevalent (422/656 [64%]) and is the only haplotype found in all four continents. Next was haplotype 36 (35/656 [5%]), which was found in Asia and Oceania. All other haplotypes are rare (<5%), with numerous singleton haplotypes (n = 28), and 40 haplotypes are continent specific (Fig. 2c and d). These results indicate that there are unexpectedly numerous *serca* haplotypes that are specific to a geographic area. The frequencies and geographical distribution of haplotypes differ greatly among geographic areas (Fig. 3). The number of haplotypes and haplotype diversity (h) are high in Africa (34 and 0.760, respectively) compared to Asia and Oceania (17 haplotypes and h = 0.581 in Asia; 6 haplotypes and h = 0.217 in Oceania) (Table 1). In South America, the value for h (0.798) was comparable to that in Africa.

DISCUSSION

Continuous implementation of ACTs as first-line treatment for uncomplicated malaria is likely imposing artemisinin pres-

sure to *P. falciparum* populations in wide geographic areas. At present, however, no alternative classes of antimalarial drugs are available to replace the artemisinin derivatives. We therefore believe it important to see whether and how the polymorphism profiles of *serca* change over time and spatially and in particular whether specific mutations/haplotypes are selected for under artemisinin pressure. The present sequence data provide baseline information on spontaneous mutations in *serca* from global *P. falciparum* populations that were unexposed to artemisinin or its derivatives. Natural variations in *serca* are characterized by abundance of geographic area-specific SNPs, the majority of which are observed at frequencies of less than 5%. The highest diversity of geographic area-specific haplotypes is observed in African parasite populations. The low prevalence of common SNPs in all parasite populations, except in South America, underscores several implications for detecting potential artemisinin-driven selection of resistant parasites and lends valuable insight on whether some *serca* haplotypes are selected by ACTs in natural parasite populations.

First, the occurrence of numerous SNPs in *serca* may make the rapid detection of artemisinin-driven selection of resistant parasites difficult. Resistant parasites, when they appear, would be geographically restricted, and their prevalence would initially be low. It is therefore rather difficult to distinguish a mutation associated with artemisinin resistance from abundantly present resistance-unrelated mutations, which are low in frequency and also geographic area specific. This is particularly true for African parasite populations. Identification of a resistance-associated mutation(s) is likely to be missed until such a time that the resistance-associated mutation becomes considerably prevalent. In this context, it should also be mentioned that some SNPs, though limited in number, are already highly prevalent. Those SNPs are probably unrelated to artemisinin resistance; for example, an I89T SNP was common in Asia, but the mutation has been suggested to be unrelated to artemisinin resistance (7). Such highly prevalent SNPs found in populations unexposed to ACTs therefore must be excluded from candidate mutations for consideration.

Second, the detection of potential artemisinin-driven selection would require a better understanding of the parasite population structure. This is particularly true for South America. The characteristics of *serca* polymorphism in South America are distinct from those in other continents: overall, the number of SNPs (and haplotypes) is limited, and SNPs with a minor allele frequency of ≥ 0.05 are prevalent. In such a scenario, distinguishing between an increase in allele frequency of a mutation due to artemisinin resistance and an increase of spontaneous mutations unrelated to artemisinin resistance would not be straightforward. South American parasite populations are strongly structured and genetic differentiation among local populations is remarkably high, probably due to epidemic expansion of some parasite genotypes (1). Consistently, *serca* SNPs/haplotypes were found to be limited in numbers but multiply represented (Table 1 and Fig. 3). Thus, the population structure must be taken into account to properly assess artemisinin-driven selection, preferably using neutral markers such as microsatellites and synonymous SNPs.

Third, the abundance of minor haplotypes may confound potential artemisinin-driven selection in some cases. If an artemisinin resistance-associated mutation(s) was generated in the major wild genotype (3D7 type), the identification of that mutation would not be difficult. A selective sweep of resistance parasites during continuous exposure to artemisinins would greatly reduce within-population diversity of *serca* haplotypes, with a fixation or predominance of a resistance-associated haplotype. A simple comparison of haplotypes from artemisinin-sensitive and -resistant parasites should reveal the resistance-associated mutation(s). However, if the resistance-associated mutation(s) was selected from one of the less prevalent haplotypes, the identification could not be readily made. Of the 48 minor haplotypes observed in this study (Table 1), 21 haplotypes have two to four amino acid substitutions compared with the major (3D7) haplotype. Since these haplotypes are in most cases continent specific, if resistant haplotypes were generated independently in several geographic areas, minor haplotypes would possess multiple SNPs, most of which might be unrelated to resistance (in some cases continent specific) but possibly selected by genetic hitchhiking linked to a nearby resistance-affording mutation. In these cases, comparison of resistance haplotypes originating from different continents with different evolutionary histories of *serca* would lead to identification of resistance-conferring mutations with high confidence.

In summary, the present analysis of a total of 656 full-length *serca* sequences identified numerous SNPs and haplotypes from geographically widespread *P. falciparum* populations that were unexposed to ACTs. The SNPs and haplotypes observed were in most cases present at a low frequency and were geographic area specific, with the exception of sequences from South America. The geographical difference in the diversity and distribution of *serca* SNPs and haplotypes observed in this study lays the groundwork for assessing whether some artemisinin resistance-associated mutations and haplotypes are selected by the current continuous and increasing deployment of ACTs.

ACKNOWLEDGMENTS

We thank all the people who participated in the epidemiological studies for their kind cooperation.

This work was supported by Grant-in-Aids for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (18073013), from the Japan Society for Promotion of Sciences (18GS03140013, 20390120, 22406012), and from the Ministry of Health, Labor and Welfare (H20-Shinkou-ippan-013).

REFERENCES

1. **Anderson, T. J., B. Haubold, J. T. Williams, J. G. Estrada-Franco, L. Richardson, R. Mollinedo, M. Bockarie, J. Mokili, S. Mharakurwa, N. French, J. Whitworth, I. D. Velez, A. H. Brockman, F. Nosten, M. U. Ferreira, and K. P. Day.** 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol. Biol. Evol. **17**:1467–1482.
2. **Bacon, D. J., A. M. McCollum, S. M. Griffing, C. Salas, V. Soberon, M. Santolalla, R. Haley, P. Tsukayama, C. Lucas, A. A. Escalante, and V. Udhayakumar.** 2009. Dynamics of malaria drug resistance patterns in the Amazon Basin region following changes in Peruvian national treatment policy for uncomplicated malaria. Antimicrob. Agents Chemother. **53**:2042–2051.
3. **Bertaux, L., L. H. Quang, V. Sinou, N. X. Thanh, and D. Parzy.** 2009. New *PfATP6* mutations found in *Plasmodium falciparum* isolates from Vietnam. Antimocrob. Agents Chemother. **53**:4570–4571.
4. **Bwijo, B., A. Kaneko, M. Takechi, I. L. Zungu, Y. Moriyama, J. K. Lum, T. Tsukahara, T. Mita, N. Takahashi, Y. Bergqvist, A. Bjorkman, and T. Kobayakawa.** 2003. High prevalence of quintuple mutant *dhps/dhfr* genes in *Plasmodium falciparum* infections seven years after introduction of

sulfadoxine and pyrimethamine as first line treatment in Malawi. Acta Trop. **85:**363–373.
5. **Cojean, S., V. Hubert, J. Le Bras, and R. Durand.** 2006. Resistance to dihydroartemisinin. Emerg. Infect. Dis. **12:**1798–1799.
6. **Dahlström, S., M. I. Veiga, P. Ferreira, A. Mårtensson, A. Kaneko, B. Andersson, A. Björkman, and J. P. Gil.** 2008. Diversity of the sarco/endoplasmic reticulum Ca^{2+}-ATPase orthologue of *Plasmodium falciparum* (PfATP6). Infect. Genet. Evol. **8:**340–345.
7. **Dondorp, A. M., F. Nosten, P. Yi, D. Das, A. P. Phyo, J. Tarning, K. M. Lwin, F. Ariey, W. Hanpithakpong, S. J. Lee, P. Ringwald, K. Silamut, M. Imwong, K. Chotivanich, P. Lim, T. Herdman, S. S. An, S. Yeung, P. Singhasivanon, N. P. Day, N. Lindegardh, D. Socheat, and N. J. White.** 2009. Artemisinin resistance in *Plasmodium falciparum* malaria. N. Engl. J. Med. **361:**455–467.
8. **Eastman, R. T., and D. A. Fidock.** 2009. Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria. Nat. Rev. Microbiol. **7:**864–874.
9. **Eckstein-Ludwig, U., R. J. Webb, I. D. Van Goethem, J. M. East, A. G. Lee, M. Kimura, P. M. O'Neill, P. G. Bray, S. A. Ward, and S. Krishna.** 2003. Artemisinins target the SERCA of *Plasmodium falciparum*. Nature **424:**957–961.
10. **Excoffier, L., G. Laval, and S. Schneider.** 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. **1:**47–50.
11. **Ferreira, I. D., A. Martinelli, L. A. Rodrigues, E. L. do Carmo, V. E. do Rosário, M. M. Póvoa, and P. Cravo.** 2008. *Plasmodium falciparum* from Pará state (Brazil) shows satisfactory in vitro response to artemisinin derivatives and absence of the S769N mutation in the *SERCA*-type PfATPase6. Trop. Med. Int. Health **13:**199–207.
12. **Ibrahim, M. L., N. Khim, H. H. Adam, F. Ariey, and J.-B. Duchemin.** 2009. Polymorphism of PfATPase in Niger: detection of three new point mutations. Malar. J. **8:**28.
13. **Jambou, R., E. Legrand, M. Niang, N. Khim, P. Lim, B. Volney, M. T. Ekala, C. Bouchier, P. Esterre, T. Fandeur, and O. Mercereau-Puijalon.** 2005. Resistance of *Plasmodium falciparum* field isolates to *in-vitro* artemether and point mutations of the *SERCA*-type PfATPase6. Lancet **366:**1960–1963.
14. **Jambou, R., A. Martinelli, J. Pinto, S. Gribaldo, E. Legrand, M. Niang, N. Kim, L. Pharath, B. Volnay, M. T. Ekala, C. Bouchier, T. Fandeur, P. Berzosa, A. Benito, I. D. Ferreira, C. Ferreira, P. P. Vieira, M. G. Alecrim, O. Mercereau-Puijalon, and P. Cravo.** 2010. Geographic structuring of the *Plasmodium falciparum* sarco (endo) plasmic reticulum Ca^{2+}ATPase (Pf*SERCA*) gene diversity. PLoS One **5:**e9424.
15. **Krishna, S.** 2006. Re-evaluation of how artemisinins work in light of emerging evidence of *in vitro* resistance. Trend Mol. Med. **12:**200–205.
16. **Legrand, E., B. Volney, J. B. Meynard, P. Esterr, and O. Mercereau-Puijalon.** 2007. Resistance to dihydroartemisinin. Emerg. Infect. Dis. **13:**808–809.
17. **Marma, A. S., T. Mita, H. Eto, T. Tsukahara, S. Sarker, and H. Endo.** 2010. High prevalence of sulfadoxine/pyrimethamine resistance alleles in *Plasmodium falciparum* parasites from Bangladesh. Parasitol. Int. **59:**178–182.
18. **Menegon, M., A. R. Sannella, G. Majori, and C. Severini.** 2008. Detection of novel point mutations in the *Plasmodium falciparum* ATPase6 candidate gene for resistance to artemisinins. Parasitol. Int. **57:**233–235.
19. **Mu, J., P. Awadalla, J. Duan, K. M. McGee, J. Keebler, K. Seydel, G. A. T. McVean, and X.-Z. Su.** 2007. Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. Nat. Genet. **39:**126–130.
20. **Mugittu, K., B. Genton, H. Mshindal, and H. P. Beck.** 2006. Molecular monitoring of *Plasmodium falciparum* resistance to artemisinin in Tanzania. Malar. J. **5:**126.
21. **Nei, M.** 1987. Molecular evolutionary genetics. Columbia University Press, New York, NY.
22. **Noedl, H., Y. Se, K. Schaecher, B. L. Smith, D. Socheat, and M. M. Fukuda.** 2008. Evidence of artemisinin-resistant malaria in western Cambodia. N. Engl. J. Med. **359:**2619–2620.
23. **Nosten, F., and N. J. White.** 2007. Artemisinin-based combination treatment of falciparum malaria. Am. J. Trop. Med. Hyg. **77**(Suppl. 6)**:**181–192.
24. **Randrianarivelojosia, M., J. L. Harisoa, L. P. Rabarijaona, L. A. Raharimalala, L. Ranaivo, V. Pietra, J. B. Duchemin, F. Rakotomanana, V. Robert, P. Mauclere, and F. Ariey.** 2002. *In vitro* sensitivity of *Plasmodium falciparum* to amodiaquine compared with other major antimalarials in Madagascar. Parassitologia **44:**141–147.
25. **Rozas, J., J. C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas.** 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics **19:**2496–2497.
26. **Tahar, R., P. Ringwald, and L. K. Basco.** 2009. Molecular epidemiology of malaria in Cameroon. XXVIII. In vitro activity of dihydroartemisinin against clinical isolates of *Plasmodium falciparum* and sequence analysis of the *P. falciparum* ATPase 6 gene. Am. J. Trop. Med. Hyg. **81:**13–18.
27. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24:**1596–1599.
28. **Tanabe, K., N. Sakihama, T. Hattori, L. Ranford-Cartwright, I. Goldman, A. A. Escalante, and A. A. Lal.** 2004. Genetic distance in housekeeping genes between *Plasmodium falciparum* and *Plasmodium reichenowi* and within *P. falciparum*. J. Mol. Evol. **59:**687–694.
29. **Tanabe, K., N. Sakihama, D. Walliker, H. Babiker, A. A. Abdel-Muhsin, B. Bakote'e, H. Ohmae, N. Arisue, T. Horii, I. Rooth, A. Färnert, A. Björkman, and L. Ranford-Cartwright.** 2007. Allelic dimorphism-associated restriction of recombination in *Plasmodium falciparum mspl*. Gene **397:**153–160.
30. **Tanabe, K., T. Mita, T. Jombart, A. Eriksson, S. Horibe, N. Palacpac, L. Ranford-Cartwright, H. Sawai, N. Sakihama, H. Ohmae, M. Nakamura, M. U. Ferreira, A. A. Escalante, F. Prugnolle, A. Björkman, A. Färnert, A. Kaneko, T. Horii, A. Manica, H. Kishino, and F. Balloux.** 2010. *Plasmodium falciparum* accompanied the human expansion out of Africa. Curr. Biol. **70:**1–7.
31. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24:**4876–4882.
32. **Uhlemann, A. C., A. Cameron, U. Eckstein-Ludwig, J. Fischbarg, P. Iserovich, F. A. Zuniga, M. East, A. Lee, L. Brady, R. K. Haynes, and S. A. Krishna.** 2005. A single amino acid residue can determine the sensitivity of *SERCA*s to artemisinins. Nat. Struct. Mol. Biol. **12:**628–629.
33. **Valderramos, S. G., D. Scanfeld, A. C. Uhlemann, D. A. Fidock, and S. Krishna.** 2010. Investigations into the role of the *Plasmodium falciparum* SERCA (PfATP6) L263E mutation in artemisinin action and resistance. Antimicrob. Agents Chemother. **54:**3842–3852.
34. **World Health Organization.** 2006. WHO guidelines for the treatment of malaria. World Health Organization, Geneva, Switzerland.
35. **Zhang, G., Y. Guan, B. Zheng, S. Wu, and L. Tang.** 2008. No *PfATPase6* S769N mutation found in *Plasmodium falciparum* isolates from China. Malar. J. **7:**122.

Acta Tropica 114 (2010) 171–176

Contents lists available at ScienceDirect

Acta Tropica

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

The wheat germ cell-free protein synthesis system: A key tool for novel malaria vaccine candidate discovery

Takafumi Tsuboi[a,b,*], Satoru Takeo[a], Thangavelu U. Arumugam[a], Hitoshi Otsuki[c], Motomi Torii[c]

[a] Cell-free Science and Technology Research Center, Ehime University, Matsuyama, Ehime 790-8577, Japan
[b] Venture Business Laboratory, Ehime University, Matsuyama, Ehime 790-8577, Japan
[c] Department of Molecular Parasitology, Ehime University Graduate School of Medicine, Toon, Ehime 791-0295, Japan

ARTICLE INFO

Article history:
Received 24 August 2009
Received in revised form 13 October 2009
Accepted 28 October 2009
Available online 11 November 2009

Keywords:
Malaria
Plasmodium falciparum
Post-genome
Vaccine candidate discovery
Wheat germ cell-free protein synthesis

ABSTRACT

Malaria kills more than a million people a year, causes malady in about three hundred million people and poses risk to approximately 40% of the world's population living in malarious countries. This disease is re-emerging mainly due to the development of drug-resistant parasites and insecticide-resistant mosquitoes. Therefore, we are now forced to resort to remedy through vaccination. Until now, not even a single licensed malaria vaccine has been developed despite intensive efforts. Even the efficacy of RTS,S, the most advanced and promising vaccine candidate in the pipeline of malaria vaccine development, was only around 50% based on a number of clinical trials. These facts urge malaria researchers to urgently enrich this pipeline, as much as possible, with potential vaccine candidates. With the availability of malaria genome database, the enrichment of this pipeline is possible if we could now employ an efficient protein expression technology to decode the malaria genomic data, without any codon optimization, into quality recombinant proteins. Then, these synthesized recombinant proteins can be characterized and screened for discovering novel potential vaccine targets. The wheat germ cell-free protein synthesis system will be a promising tool to this end. This review highlights the recent successes in synthesizing quality malaria proteins using this tool.

1. Introduction

Malaria kills more than a million people a year, causes malady in about three hundred million people and poses risk to around 3.3 billion people (WHO, 2008). Despite the parasite's complex life cycle, high level of antigenic diversity and mechanism of immune evasion, naturally acquired immunity to malaria does develop after repeated exposure over a period of several years and this immunity confers protection, against symptomatic disease, high-density parasitemia and death (Doolan et al., 2009; Genton, 2008; Gupta et al., 1999; Marsh and Kinyanjui, 2006), through protective antibodies (Cohen et al., 1961; McGregor, 1964). In addition to the naturally acquired immunities, sterile, long-lasting protective immunity has been convincingly proved in many studies, such as vaccination with radiation-attenuated sporozoites (Clyde, 1975; Hoffman et al., 2002; Nussenzweig et al., 1967), and inoculation of infective sporozoites to human volunteers under a prophylactic regimen of chloroquine (Roestenberg et al., 2009). These facts strongly support the reasoning of incorporating malaria vaccines as one of the components of malaria control measures.

2. Malaria vaccine development

The efficacy of malaria control through current interventions that employs drugs and insecticides may not be sustained too long since they rely on too few compounds (Genton, 2008). In fact, the disease is re-emerging mainly due to the emergence of drug-resistant parasites (Greenwood and Mutabingwa, 2002). Very recently, parasites have developed resistance even to the hitherto-promising artemisinin (Dondorp et al., 2009). Therefore, we have to develop and employ malaria vaccines as one of the essential components towards the malaria eradication (Greenwood, 2009). The fact that, in spite of intensive efforts, not even a single licensed malaria vaccine has been developed, urges malaria research community to employ efficient post-genomic approaches (Richie and Saul, 2002). Malaria vaccines are generally divided into three groups based on stages of the parasite life cycle targeted by the vaccine. They are pre-erythrocytic vaccines, asexual blood-stage vaccines, and transmission-blocking vaccines. Fig. 1 gives a quick description about vaccine categories, the stages they target, along with malaria vaccine candidates in clinical trials. Detailed descriptions about

* Corresponding author at: Ehime University, Cell-Free Science and Technology Research Center, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan.
Tel.: +81 89 927 8277; fax: +81 89 927 9941.
E-mail address: tsuboi@ccr.ehime-u.ac.jp (T. Tsuboi).

0001-706X/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.actatropica.2009.10.024

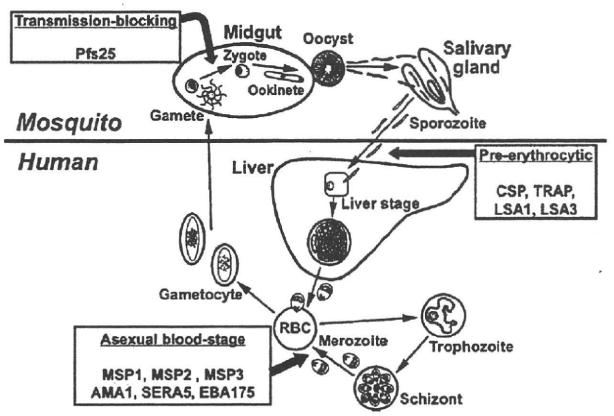


Fig. 1. Malaria vaccines in clinical trials mapped on the parasite life cycle. Malaria vaccines are categorized into three groups based on the target stages of the parasite life cycle. They are pre-erythrocytic, asexual blood-stage, and transmission-blocking vaccines. RBC: red blood cell, CSP: circumsporozoite protein, TRAP: thrombospondin-related adhesion protein, LSA1: liver stage antigen 1, LSA3: liver stage antigen 3, MSP1: merozoite surface protein 1, MSP2: merozoite surface protein 2, MSP3: merozoite surface protein 3, AMA1: Apical membrane antigen 1, SERA5: Serine repeat antigen 5, and EBA175: erythrocyte binding antigen 175. Detailed descriptions about these vaccines were given in other published reviews (Genton, 2008; Richards and Beeson, 2009).

these vaccines are given in other published reviews (Genton, 2008; Richards and Beeson, 2009).

2.1. Pre-erythrocytic vaccines

Pre-erythrocytic vaccines have been designed to prevent entry of sporozoites into hepatocytes and the development of liver stage parasites (Greenwood et al., 2008; Waters, 2006).

2.2. Asexual blood-stage vaccines

Asexual blood-stage vaccines are designed to reduce merozoite invasion, multiplication and growth in order to protect against clinical symptoms and particularly severe disease. It is widely understood that this vaccine induces antibodies that may have roles in prevention of merozoite invasion, clearance of infected erythrocytes, prevention of adhesion and sequestration of parasitized erythrocytes in the vasculature. It is also possible that, in addition to preventing clinical illness, an effective blood-stage vaccine may also contribute to malaria eradication by reducing the efficiency of the transmission of parasites from human host to mosquito by interrupting the blood-stage life cycle in the human body (Duffy, 2007; Genton, 2008; Moll et al., 2007; Richards and Beeson, 2009).

2.3. Transmission-blocking vaccines

Transmission-blocking vaccines are aimed at interrupting the parasite life cycle in the mosquito blood meal. These vaccines elicit antibodies against antigens that are expressed by the sexual stages of the parasite and, thus, stop their subsequent development in the mosquito midgut (Carter, 2001; Tsuboi et al., 2003). These transmission-blocking vaccines, if used in combination with pre-erythrocytic or asexual blood-stage vaccines, might play a key role in finally breaking the transmission of parasites, leading to eradication of the diseases (Targett and Greenwood, 2008).

It is an accepted view that an effective malaria vaccine need to target several stages of parasite and several components of the different stages of parasite (Lasonder et al., 2002) and it must induce protective immune responses equivalent to, or better than, those provided by naturally acquired immunity or immunization with attenuated whole parasite (Gardner et al., 2002). In order to accelerate the discovery of such vaccine candidates, it is indispensable now to establish and exploit two things. One is the optimal recombinant protein synthesis system for synthesizing malaria proteins on a whole-proteome scale and other is the efficient post-genomic high-throughput approaches for sifting potential vaccine candidates out from this whole malaria proteome.

3. Post-genome approaches for novel malaria vaccine candidate discovery research

As described in Fig. 1, decades of research in the pre-genomic era have identified only a handful of vaccine candidates. With the recent completion of the genome projects of human malaria parasites, *Plasmodium falciparum* (Gardner et al., 2002), *P. vivax* (Carlton et al., 2008), primate malaria parasite, *P. knowlesi* (Pain et al., 2008), and rodent malaria parasite, *P. yoelii* (Carlton et al., 2002), we are now in the post-genome era. However, to tangibly reap the benefits from these genomic data, it is indispensable to thoroughly analyze these data using at least two post-genomic high-throughput approaches. One is the functional approach (otherwise called as reverse vaccinology (Flower, 2008)), and the other is the immunoscreening approach.

3.1. Functional approach

In the functional approach, initially putative vaccine candidates are selected from the genome database based on either patent or latent functional criteria. Then these putative candidates are expressed *in vitro* using effective and efficient protein synthesis system to obtain quality proteins for further downstream vaccine candidate assessment studies (Hall et al., 2005). For example, the genes that are essential for the parasite's survival revealed by gene knockout studies (Cowman and Crabb, 2006), or the genes with signatures of strong immune selective pressure revealed by polymorphisms and diversity studies (Mu et al., 2007), or the genes involved in host cell invasion, or the genes whose products are localized on the cell surface or in the apical organelles of the sporozoite, merozoite, and ookinete could be putative vaccine candidates.

3.2. Immunoscreening approach

In the immunoscreening approach, initially putative vaccine candidates are selected based on their immuno-reactivity with the protective antibodies that are elicited in humans, after natural or experimental infection or after vaccination with attenuated organisms (Hoffman et al., 2002; VanBuskirk et al., 2009). Here extensive set of quality recombinant malaria proteins are synthesized, in small scale, using protein synthesis system and are screened intensively, using immunoassays such as enzyme-linked immunosorbent assay, protein microarray, with a large number of human serum samples obtained from non-immune and immune individuals (Doolan et al., 2008). Then, these immuno-reactive putative vaccine candidates that correlate with protection will be synthesized, in large scale, for further downstream vaccine research.

4. What is the optimal recombinant protein synthesis system for malaria proteins?

Whatever may be the approach employed for novel vaccine candidate discovery research, we do need to have an effective and

efficient recombinant protein synthesis system not only for the initial synthesis, characterization, and downstream assessment of putative vaccine candidates, but also for the mass production of vaccine antigens for vaccination purpose in the later stage. Therefore, the choice of the recombinant protein synthesis method is the most crucial factor. One of the main obstacles in malaria vaccine discovery research is the lack of an optimal protein expression system that can decode A/T-rich, low-complexity-region-containing (i.e., repeated stretches of amino acid sequences) malaria genes into high quality (i.e., properly folded) recombinant proteins that are amenable for malaria vaccine research (Gardner et al., 2002). These obstacles pose major limitation to express *P. falciparum* genes not only in conventional *E. coli* cell-based systems (Aguiar et al., 2004; Mehlin et al., 2006; Vedadi et al., 2007), but also in eukaryotic cell-based expression systems such as yeast, baculovirus, or Chinese hamster ovary cell (Tsuboi et al., 2008).

Since the rate of peptide growth on ribosomes is 5–10 times slower in eukaryotes than in prokaryotes, this slow rate of peptide growth in eukaryotic protein expression system contributes greatly towards correct conformational folding of multidomain proteins and their solubility (Hartl and Hayer-Hartl, 2002; Netzer and Hartl, 1997). Therefore eukaryotic proteins with multiple domains, such as malaria parasite proteins, when expressed either in *E. coli* cell-free or cell-based system, tend to fold incorrectly, and become insoluble (Netzer and Hartl, 1997). After all, the correct folding and solubility of malaria proteins are indispensable critical factors that directly affect the success in our genome-wide search for potential vaccine antigens. From the correct folding point of view, the eukaryotic-based system is greatly advantageous over *E. coli* based prokaryotic system. The eukaryotic-based systems have been widely used to synthesize eukaryotic multidomain proteins in active forms. However, they all suffer from low throughput and low productivity (Endo and Sawasaki, 2006; Goshima et al., 2008). For example, it is well known that the rabbit reticulocytes lysate system has been employed for the production of quality eukaryotic recombinant proteins. However, the yield of this system is very low, and moreover this system has post-translational modification machinery, i.e., glycosylation (Endo and Sawasaki, 2006). Therefore it is disadvantageous to express malaria genes in rabbit reticulocyte system and other eukaryotic cell-based protein expression systems that posses glycosylation machinery, for these systems can produce inappropriately glycosylated recombinant malaria proteins, resulting in incorrect immune responses (Gowda and Davidson, 1999; Kedees et al., 2002). The fact that both the wheat germ system and the malaria parasite do not have glycosylation machinery (Samuelson et al., 2005) is an additional advantage for expressing malaria proteins without any inadvertent glycosylation in wheat germ system. (Endo and Sawasaki, 2006). Among all eukaryotic protein expression systems, wheat germ cell-free translation system is the most suitable for easy handling and achieving high-throughput, high-solubility, and high-productivity in the synthesis of recombinant proteins (Endo and Sawasaki, 2006; Goshima et al., 2008).

From our experience in expressing around five hundred malaria genes using wheat germ cell-free system, we have found that the wheat germ cell-free system can surmount almost all of the above stumbling blocks in the path of post-genome malaria vaccine candidate discovery and therefore can be used for synthesizing high level of soluble malaria recombinant proteins (an important prerequisite for vaccine candidate assessment and large scale vaccine production), and accelerate malaria vaccine discovery (Tsuboi et al., 2008). Some of the successes that have been published by our collaborators and us have been described in the subsequent sections. Therefore we believe that the wheat germ system is the optimal system for synthesis of malaria proteins.

5. Synthesis of malaria proteins using prokaryotic protein synthesis system

There were several trials to achieve genome-wide expression of *P. falciparum* genes using the conventional *E. coli* based protein synthesis system. Aguiar et al. (2004) using *E. coli* cells, were able to express only 39 out of 292 malaria genes cloned in GST-fusion vector. Mehlin et al. (2006) tried to express 1000 genes encoding relatively small (<450 amino acids) malaria cytosolic proteins, in *E. coli*, for the structural analysis. In that study, only 30% of the genes were expressed and only 6.3% of the proteins were soluble, yielding 0.9–406 mg of protein per liter of culture medium. Vedadi et al. (2007) tried to express 182 malaria genes, for the structural analysis, utilizing genetically engineered *E. coli* strain supplemented with tRNAs that allows reading of high number of A/U codons in malaria mRNA. They achieved marginal improvement in protein solubility, i.e., 38 out of 182 (20.9%) proteins tested were soluble. Using an *E. coli* cell-free rapid expression system, Mu et al. attempted to express 108 polymorphic malaria genes that are considered to be under immune selection pressure to find potential vaccine candidates. Expression of proteins was verified via protein blot using antibodies to the histidine tags incorporated into the C-terminus of the expressed proteins. In their study, they found that, out of 108 genes, 65 genes (i.e., 60.2%) could be expressed (Mu et al., 2007). Recently, Doolan et al. (2008) reported the construction of the *P. falciparum* protein microarrays using *E. coli* cell-free protein synthesis system for high-throughput immunoscreening for vaccine candidate discovery. In their study, they obtained >90% efficiency in their effort to express 250 *P. falciparum* genes using *E. coli* cell-free *in vitro* transcription and translation reactions. And they have attributed their high success rate in this system, at least in part, to the fact that the system is supplemented with rare transfer RNAs to help translate A/T-rich genes. However, in the *E. coli* cell-free system, the rate of protein expression reported by Mu et al. (2007) was lower than the rate reported by Doolan et al. (2008). The above experiments proved that *E. coli* cell-free system is better in productivity and throughput than *E. coli* cell-based system, however, these experiments did not prove that the quality of proteins synthesized in *E. coli* cell-free system is amenable enough for vaccine candidate discovery research.

6. A cursory glance at wheat germ cell-free protein synthesis system

Recently, the wheat germ cell-free protein synthesis system was established for high-throughput synthesis of soluble recombinant proteins (Madin et al., 2000) and is now commercially available as simple protein synthesis kits (CFS Co., Ltd., Matsuyama, Japan). Basically there are two protocols established for practical use in this system (Endo and Sawasaki, 2006). The first one is for small scale, high-throughput or parallel synthesis of proteins. It can be applied for initial testing of gene expression and solubility of proteins. This will facilitate genome-wide biochemical annotation of gene products (Sawasaki et al., 2002). The second protocol is for the large scale production of interested proteins (Endo and Sawasaki, 2006; Kamura et al., 2005; Sawasaki and Endo, 2008). This protocol can produce hundreds of micrograms of protein, and also, offer scope for scale up (Sawasaki and Endo, 2008). These two protocols have already been acknowledged as advantageous for producing high quality eukaryotic proteins for structural and functional genomics studies (Goshima et al., 2008; Vinarov et al., 2004). Taken together, this system is versatile for both small and large scale synthesis of quality eukaryotic proteins.

Table 1
The recent successes in the synthesis of biologically active malaria proteins using wheat germ cell-free expression system[a].

Gene ID[b]	Gene name	Species[c]	Results[d]	References
Enzymes				
PFD0830w	Dihydrofolate reductase–thymidylate synthase	*Pf*	Enzyme activities (DHFR, and TS)	Mudeppa et al. (2007)
PVX_087680	Chitinase	*Pv*	Enzyme activities	Takeo et al. (2009)
PF10_0363	Pyruvate kinase 2	*Pf*	Ab (WB, IFA)	Maeda et al. (2009)
Sporozoite antigens				
PFC0210c	Circumsporozoite protein	*Pf*	Ab (WB, IFA)	Tsuboi et al. (2008)
PFD0215c	P52	*Pf*	Ab (IFA)	VanBuskirk et al. (2009)
PKH_121770	thrombospondin-related adhesion protein	*Pk*	T cell	Jiang et al. (2009)
Merozoite antigens				
PF11_0344	Apical membrane antigen 1	*Pf*	Ab (WB, IFA)	Tsuboi et al. (2008)
PF14_0495	Rhoptry neck protein 2	*Pf*	Ab (IP, WB, IFA, IEM)	Cao et al. (2009)
PY04764	Erythrocyte binding ligand	*Py*	Ab (WB, IFA, IEM)	Otsuki et al. (2009)
PKH_072850	Merozoite surface protein 1	*Pk*	T cell	Jiang et al. (2009)
Ookinete antigen				
PF10_0303	Pfs25	*Pf*	Ab (WB, IFA, TBA)	Tsuboi et al. (2008)

[a] The individual articles may be consulted for further details.
[b] Detailed information is available at the PlasmoDB website. (http://plasmodb.org/plasmo/).
[c] Plasmodium species, Pf: *Plasmodium falciparum*, Pv: *P. vivax*, Py: *P. yoelii*, and Pk: *P. knowlesi*.
[d] Results obtained by the contribution of recombinant proteins synthesized by the wheat germ cell-free system. Ab, antibody; IP, immunoprecipitation; WB, Western blot, IFA, immunofluorescence microscopy, IEM, immunoelectron microscopy. T cell, *in vitro* T cell studies; TBA, transmission-blocking activity.

7. Synthesis of malaria proteins using eukaryotic wheat germ protein synthesis system

Initially, in order to test the suitability of wheat germ cell-free system for high-throughput synthesis of malaria recombinant proteins, we attempted to express 124 genes encoding asexual blood-stage parasite proteins, selected from malaria genome database, PlasmoDB (http://plasmodb.org/plasmo/). Out of 124 genes, 93 of them (i.e., around 75%) yielded soluble protein products. However, the magnitude/extent of solubility among these 93 soluble proteins is on an average 65% and ranges from 26% to 100% (Tsuboi et al., 2008). The average yield of full size recombinant proteins was 1.9 μg per 150 μl of reaction mixture and this amount is sufficient for high-throughput screening of antigens that correlates with protection using hyper-immune serum.

The experimental results published by our collaborators and us (Table 1) confirm that wheat germ cell-free protein synthesis system is very suitable for decoding A/T-rich malaria genes without any codon optimization into biologically active malaria proteins. Firstly, the system was able to synthesize active malaria enzymes that are recalcitrant to expression in other systems, such as *P. falciparum* dihydrofolate reductase–thymidylate synthase (Mudeppa et al., 2007), a chitinase of *P. vivax* (Takeo et al., 2009) and pyruvate kinase type-II isozyme of *P. falciparum* (Maeda et al., 2009). Secondly, the system was able to produce a sufficient amount of good quality malaria proteins, such as repeat-rich circumsporozoite protein (Tsuboi et al., 2008), cysteine-rich P52 (VanBuskirk et al., 2009), thrombospondin-related adhesion protein (Jiang et al., 2009), rhoptry neck protein 2 (Cao et al., 2009), erythrocyte binding ligand (Otsuki et al., 2009), merozoite surface protein 1 (Jiang et al., 2009), apical membrane antigen 1 and highly cysteine-rich Pfs25 (a promising transmission-blocking vaccine candidate) (Tsuboi et al., 2008).

Recently, we attempted to express 567 of *P. falciparum* cDNA clones belonging to sporozoite, merozoite, and gametocyte stages in a high-throughput format by the wheat germ cell-free system. Out of 567 genes, 478 of them (i.e., around 84%) yielded soluble protein products (unpublished). Our biochemical, immunocytochemical, and biological analyses have revealed that the recombinant malaria proteins synthesized by this system are of high quality and therefore amenable for the assessment and discovery of potential vaccine targets (Tsuboi et al., 2008). Thus, we indeed believe that the wheat germ cell-free protein synthesis system is a key tool for decoding malaria genome for malaria proteome and vaccine research.

8. Conclusion

From a malaria vaccine perspective, wheat germ cell-free system is the most suitable system to date for easily achieving high-throughput, high-solubility, and high-productivity in the whole-proteome-scale synthesis of malaria proteins and construction of microarrays of malaria proteins that will accelerate post-genomic novel malaria vaccine candidate discovery.

Acknowledgments

We thank Kana Kato, Aya Tamai, and Mai Tasaka for their technical assistance. This work was supported in part by Grants-in-Aid for Scientific Research 19406009 (TT), Scientific Research on Priority Areas 21022034 (TT) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan; and, in part, by a Grant-in-Aid from the Ministry of Health, Labour, and Welfare (H21-Chikyukibo-ippan-005) (TT), Japan.

References

Aguiar, J.C., LaBaer, J., Blair, P.L., Shamailova, V.Y., Koundinya, M., Russell, J.A., Huang, F., Mar, W., Anthony, R.M., Witney, A., Caruana, S.R., Brizuela, L., Sacci Jr., J.B., Hoffman, S.L., Carucci, D.J., 2004. High-throughput generation of *P. falciparum* functional molecules by recombinational cloning. Genome Res. 14, 2076–2082.

Cao, J., Kaneko, O., Thongkukiatkul, A., Tachibana, M., Otsuki, H., Gao, Q., Tsuboi, T., Torii, M., 2009. Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in *Plasmodium falciparum* merozoites. Parasitol. Int. 58, 29–35.

Carlton, J.M., Angiuoli, S.V., Suh, B.B., Kooij, T.W., Pertea, M., Silva, J.C., Ermolaeva, M.D., Allen, J.E., Selengut, J.D., Koo, H.L., Peterson, J.D., Pop, M., Kosack, D.S., Shumway, M.F., Bidwell, S.L., Shallom, S.J., van Aken, S.E., Riedmuller, S.B., Feldblyum, T.V., Cho, J.K., Quackenbush, J., Sedegah, M., Shoaibi, A., Cummings, L.M., Florens, L., Yates, J.R., Raine, J.D., Sinden, R.E., Harris, M.A., Cunningham, D.A., Preiser, P.R., Bergman, L.W., Vaidya, A.B., van Lin, L.H., Janse, C.J., Waters, A.P., Smith, H.O., White, O.R., Salzberg, S.L., Venter, J.C., Fraser, C.M., Hoffman, S.L., Gardner, M.J., Carucci, D.J., 2002. Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. Nature 419, 512–519.

Carlton, J.M., Adams, J.H., Silva, J.C., Bidwell, S.L., Lorenzi, H., Caler, E., Crabtree, J., Angiuoli, S.V., Merino, E.F., Amedeo, P., Cheng, Q., Coulson, R.M., Crabb, B.S., Del Portillo, H.A., Essien, K., Feldblyum, T.V., Fernandez-Becerra, C., Gilson, P.R., Gueye, A.H., Guo, X., Kang'a, S., Kooij, T.W., Korsinczky, M., Meyer, E.V., Nene, V., Paulsen, I., White, O., Ralph, S.A., Ren, Q., Sargeant, T.J., Salzberg, S.L., Stoeckert, C.J., Sullivan, S.A., Yamamoto, M.M., Hoffman, S.L., Wortman, J.R.,

Gardner, M.J., Galinski, M.R., Barnwell, J.W., Fraser-Liggett, C.M., 2008. Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. Nature 455, 757–763.
Carter, R., 2001. Transmission blocking malaria vaccines. Vaccine 19, 2309–2314.
Clyde, D.F., 1975. Immunization of man against falciparum and vivax malaria by use of attenuated sporozoites. Am. J. Trop. Med. Hyg. 24, 397–401.
Cohen, S., Mc, G.I., Carrington, S., 1961. Gamma-globulin and acquired immunity to human malaria. Nature 192, 733–737.
Cowman, A.F., Crabb, B.S., 2006. Invasion of red blood cells by malaria parasites. Cell 124, 755–766.
Dondorp, A.M., Nosten, F., Yi, P., Das, D., Phyo, A.P., Tarning, J., Lwin, K.M., Ariey, F., Hanpithakpong, W., Lee, S.J., Ringwald, P., Silamut, K., Imwong, M., Chotivanich, K., Lim, P., Herdman, T., An, S.S., Yeung, S., Singhasivanon, P., Day, N.P., Lindegardh, N., Socheat, D., White, N.J., 2009. Artemisinin resistance in *Plasmodium falciparum* malaria. N. Engl. J. Med. 361, 455–467.
Doolan, D.L., Mu, Y., Unal, B., Sundaresh, S., Hirst, S., Valdez, C., Randall, A., Molina, D., Liang, X., Freilich, D.A., Oloo, J.A., Blair, P.L., Aguiar, J.C., Baldi, P., Davies, D.H., Felgner, P.L., 2008. Profiling humoral immune responses to *P. falciparum* infection with protein microarrays. Proteomics 8, 4680–4694.
Doolan, D.L., Dobano, C., Baird, J.K., 2009. Acquired immunity to malaria. Clin. Microbiol. Rev. 22, 13–36.
Duffy, P.E., 2007. *Plasmodium* in the placenta: parasites, parity, protection, prevention and possibly preeclampsia. Parasitology 134, 1877–1881.
Endo, Y., Sawasaki, T., 2006. Cell-free expression systems for eukaryotic protein production. Curr. Opin. Biotechnol. 17, 373–380.
Flower, D., 2008. Bioinformatics for Vaccinology. John Wiley & Sons, Chichester, West Sussex, UK.
Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., Barrell, B., 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419, 498–511.
Genton, B., 2008. Malaria vaccines: a toy for travelers or a tool for eradication? Expert Rev. Vaccines 7, 597–611.
Goshima, N., Kawamura, Y., Fukumoto, A., Miura, A., Honma, R., Satoh, R., Wakamatsu, A., Yamamoto, J., Kimura, K., Nishikawa, T., Andoh, T., Iida, Y., Ishikawa, K., Ito, E., Kagawa, N., Kaminaga, C., Kanehori, K., Kawakami, B., Kenmochi, K., Kimura, R., Kobayashi, M., Kuroita, T., Kuwayama, H., Maruyama, Y., Matsuo, K., Minami, K., Mitsubori, M., Mori, M., Morishita, R., Murase, A., Nishikawa, A., Nishikawa, S., Okamoto, T., Sakagami, N., Sakamoto, Y., Sasaki, Y., Seki, T., Sono, S., Sugiyama, A., Sumiya, T., Takayama, T., Takayama, Y., Takeda, H., Togashi, T., Yahata, K., Yamada, H., Yanagisawa, Y., Endo, Y., Imamoto, F., Kisu, Y., Tanaka, S., Isogai, T., Imai, J., Watanabe, S., Nomura, N., 2008. Human protein factory for converting the transcriptome into an in vitro-expressed proteome. Nat. Methods 5, 1011–1017.
Gowda, D.C., Davidson, E.A., 1999. Protein glycosylation in the malaria parasite. Parasitol. Today 15, 147–152.
Greenwood, B., Mutabingwa, T., 2002. Malaria in 2002. Nature 415, 670–672.
Greenwood, B., 2009. Can malaria be eliminated? Trans. R. Soc. Trop. Med. Hyg. 103 (Suppl. 1), S2–S5.
Greenwood, B.M., Fidock, D.A., Kyle, D.E., Kappe, S.H., Alonso, P.L., Collins, F.H., Duffy, P.E., 2008. Malaria: progress, perils, and prospects for eradication. J. Clin. Invest. 118, 1266–1276.
Gupta, S., Snow, R.W., Donnelly, C.A., Marsh, K., Newbold, C., 1999. Immunity to non-cerebral severe malaria is acquired after one or two infections. Nat. Med. 5, 340–343.
Hall, N., Karras, M., Raine, J.D., Carlton, J.M., Kooij, T.W., Berriman, M., Florens, L., Janssen, C.S., Pain, A., Christophides, G.K., James, K., Rutherford, K., Harris, B., Harris, D., Churcher, C., Quail, M.A., Ormond, D., Doggett, J., Trueman, H.E., Mendoza, J., Bidwell, S.L., Rajandream, M.A., Carucci, D.J., Yates 3rd, J.R., Kafatos, F.C., Janse, C.J., Barrell, B., Turner, C.M., Waters, A.P., Sinden, R.E., 2005. A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. Science 307, 82–86.
Hartl, F.U., Hayer-Hartl, M., 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295, 1852–1858.
Hoffman, S.L., Goh, L.M., Luke, T.C., Schneider, I., Le, T.P., Doolan, D.L., Sacci, J., de la Vega, P., Dowler, M., Paul, C., Gordon, D.M., Stoute, J.A., Church, L.W., Sedegah, M., Heppner, D.G., Ballou, W.R., Richie, T.L., 2002. Protection of humans against malaria by immunization with radiation-attenuated *Plasmodium falciparum* sporozoites. J. Infect. Dis. 185, 1155–1164.
Jiang, G., Shi, M., Conteh, S., Richie, N., Banania, G., Geneshan, H., Valencia, A., Singh, P., Aguiar, J., Limbach, K., Kamrud, K.I., Rayner, J., Smith, J., Bruder, J.T., King, C.R., Tsuboi, T., Takeo, S., Endo, Y., Doolan, D.L., Richie, T.L., Weiss, W.R., 2009. Sterile protection against *Plasmodium knowlesi* in rhesus monkeys from a malaria vaccine: comparison of heterologous prime boost strategies. PLoS ONE 4, e6559.
Kamura, N., Sawasaki, T., Kasahara, Y., Takai, K., Endo, Y., 2005. Selection of 5′-untranslated sequences that enhance initiation of translation in a cell-free protein synthesis system from wheat embryos. Bioorg. Med. Chem. Lett. 15, 5402–5406.
Kedees, M.H., Azzouz, N., Gerold, P., Shams-Eldin, H., Iqbal, J., Eckert, V., Schwarz, R.T., 2002. Plasmodium falciparum: glycosylation status of *Plasmodium falciparum* circumsporozoite protein expressed in the baculovirus system. Exp. Parasitol. 101, 64–68.
Lasonder, E., Ishihama, Y., Andersen, J.S., Vermunt, A.M., Pain, A., Sauerwein, R.W., Eling, W.M., Hall, N., Waters, A.P., Stunnenberg, H.G., Mann, M., 2002. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. Nature 419, 537–542.
Madin, K., Sawasaki, T., Ogasawara, T., Endo, Y., 2000. A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. Proc. Natl. Acad. Sci. U.S.A. 97, 559–564.
Maeda, T., Saito, T., Harb, O.S., Roos, D.S., Takeo, S., Suzuki, H., Tsuboi, T., Takeuchi, T., Asai, T., 2009. Pyruvate kinase type-II isozyme in *Plasmodium falciparum* localizes to the apicoplast. Parasitol. Int. 58, 101–105.
Marsh, K., Kinyanjui, S., 2006. Immune effector mechanisms in malaria. Parasite Immunol. 28, 51–60.
McGregor, I.A., 1964. The passive transfer of human malarial immunity. Am. J. Trop. Med. Hyg. 13 (Suppl.), 237–239.
Mehlin, C., Boni, E., Buckner, F.S., Engel, L., Feist, T., Gelb, M.H., Haji, L., Kim, D., Liu, C., Mueller, N., Myler, P.J., Reddy, J.T., Sampson, J.N., Subramanian, E., Van Voorhis, W.C., Worthey, E., Zucker, F., Hol, W.G., 2006. Heterologous expression of proteins from *Plasmodium falciparum*: results from 1000 genes. Mol. Biochem. Parasitol. 148, 144–160.
Moll, K., Pettersson, F., Vogt, A.M., Jonsson, C., Rasti, N., Ahuja, S., Spangberg, M., Mercereau-Puijalon, O., Arnot, D.E., Wahlgren, M., Chen, Q., 2007. Generation of cross-protective antibodies against *Plasmodium falciparum* sequestration by immunization with an erythrocyte membrane protein 1-duffy binding-like 1 alpha domain. Infect. Immun. 75, 211–219.
Mu, J., Awadalla, P., Duan, J., McGee, K.M., Keebler, J., Seydel, K., McVean, G.A., Su, X.Z., 2007. Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. Nat. Genet. 39, 126–130.
Mudeppa, D.G., Pang, C.K., Tsuboi, T., Endo, Y., Buckner, F.S., Varani, G., Rathod, P.K., 2007. Cell-free production of functional *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase. Mol. Biochem. Parasitol. 151, 216–219.
Netzer, W.J., Hartl, F.U., 1997. Recombination of protein domains facilitated by co-translational folding in eukaryotes. Nature 388, 343–349.
Nussenzweig, R.S., Vanderberg, J., Most, H., Orton, C., 1967. Protective immunity produced by the injection of x-irradiated sporozoites of *Plasmodium berghei*. Nature 216, 160–162.
Otsuki, H., Kaneko, O., Thongkukiatkul, A., Tachibana, M., Iriko, H., Takeo, S., Tsuboi, T., Torii, M., 2009. Single amino acid substitution in *Plasmodium yoelii* erythrocyte ligand determines its localization and controls parasite virulence. Proc. Natl. Acad. Sci. U.S.A. 106, 7167–7172.
Pain, A., Bohme, U., Berry, A.E., Mungall, K., Finn, R.D., Jackson, A.P., Mourier, T., Mistry, J., Pasini, E.M., Aslett, M.A., Balasubrammaniam, S., Borgwardt, K., Brooks, K., Carret, C., Carver, T.J., Cherevach, I., Chillingworth, T., Clark, T.G., Galinski, M.R., Hall, N., Harper, D., Harris, D., Hauser, H., Ivens, A., Janssen, C.S., Keane, T., Larke, N., Lapp, S., Marti, M., Moule, S., Meyer, I.M., Ormond, D., Peters, N., Sanders, M., Sanders, S., Sargeant, T.J., Simmonds, M., Smith, F., Squares, R., Thurston, S., Tivey, A.R., Walker, D., White, B., Zuiderwijk, E., Churcher, C., Quail, M.A., Cowman, A.F., Turner, C.M., Rajandream, M.A., Kocken, C.H., Thomas, A.W., Newbold, C.I., Barrell, B.G., Berriman, M., 2008. The genome of the simian and human malaria parasite *Plasmodium knowlesi*. Nature 455, 799–803.
Richards, J.S., Beeson, J.G., 2009. The future for blood-stage vaccines against malaria. Immunol. Cell Biol. 87, 377–390.
Richie, T.L., Saul, A., 2002. Progress and challenges for malaria vaccines. Nature 415, 694–701.
Roestenberg, M., McCall, M., Hopman, J., Wiersma, J., Luty, A.J., van Gemert, G.J., van de Vegte-Bolmer, M., van Schaijk, B., Teelen, K., Arens, T., Spaarman, L., de Mast, Q., Roeffen, W., Snounou, G., Renia, L., van der Ven, A., Hermsen, C.C., Sauerwein, R., 2009. Protection against a malaria challenge by sporozoite inoculation. N. Engl. J. Med. 361, 468–477.
Samuelson, J., Banerjee, S., Magnelli, P., Cui, J., Kelleher, D.J., Gilmore, R., Robbins, P.W., 2005. The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. Proc. Natl. Acad. Sci. U.S.A. 102, 1548–1553.
Sawasaki, T., Ogasawara, T., Morishita, R., Endo, Y., 2002. A cell-free protein synthesis system for high-throughput proteomics. Proc. Natl. Acad. Sci. U.S.A. 99, 14652–14657.
Sawasaki, T., Endo, Y., 2008. The wheat germ cell-free protein synthesis system. In: Spirin, A.S., Swartz, J.R. (Eds.), Cell-free protein synthesis, methods and protocols. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, pp. 111–139.
Takeo, S., Hisamori, D., Matsuda, S., Vinetz, J., Sattabongkot, J., Tsuboi, T., 2009. Enzymatic characterization of the *Plasmodium vivax* chitinase, a potential malaria transmission-blocking target. Parasitol. Int. 58, 243–248.
Targett, G.A., Greenwood, B.M., 2008. Malaria vaccines and their potential role in the elimination of malaria. Malar. J. 7 (Suppl. 1), S10.
Tsuboi, T., Tachibana, M., Kaneko, O., Torii, M., 2003. Transmission-blocking vaccine of vivax malaria. Parasitol. Int. 52, 1–11.
Tsuboi, T., Takeo, S., Iriko, H., Jin, L., Tsuchimochi, M., Matsuda, S., Han, E.T., Otsuki, H., Kaneko, O., Sattabongkot, J., Udomsangpetch, R., Sawasaki, T., Torii, M., Endo, Y., 2008. Wheat germ cell-free system-based production of malaria proteins for discovery of novel vaccine candidates. Infect. Immun. 76, 1702–1708.
VanBuskirk, K.M., O'Neill, M.T., De La Vega, P., Maier, A.G., Krzych, U., Williams, J., Dowler, M.G., Sacci Jr., J.B., Kangwanrangsan, N., Tsuboi, T., Kneteman, N.M., Heppner Jr., D.G., Murdock, B.A., Mikolajczak, S.A., Aly, A.S., Cowman, A.F., Kappe,

S.H., 2009. Preerythrocytic, live-attenuated *Plasmodium falciparum* vaccine candidates by design. Proc. Natl. Acad. Sci. U.S.A. 106, 13004–13009.
Vedadi, M., Lew, J., Artz, J., Amani, M., Zhao, Y., Dong, A., Wasney, G.A., Gao, M., Hills, T., Brokx, S., Qiu, W., Sharma, S., Diassiti, A., Alam, Z., Melone, M., Mulichak, A., Wernimont, A., Bray, J., Loppnau, P., Plotnikova, O., Newberry, K., Sundararajan, E., Houston, S., Walker, J., Tempel, W., Bochkarev, A., Kozieradzki, I., Edwards, A., Arrowsmith, C., Roos, D., Kain, K., Hui, R., 2007. Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related Apicomplexan organisms. Mol. Biochem. Parasitol. 151, 100–110.
Vinarov, D.A., Lytle, B.L., Peterson, F.C., Tyler, E.M., Volkman, B.F., Markley, J.L., 2004. Cell-free protein production and labeling protocol for NMR-based structural proteomics. Nat. Methods 1, 149–153.
Waters, A., 2006. Malaria: new vaccines for old? Cell 124, 689–693.
WHO, 2008. World Malaria Report 2008. WHO Press, Geneva, Switzerland.

Am. J. Trop. Med. Hyg., 83(6), 2010, pp. 1230–1237
doi:10.4269/ajtmh.2010.10-0353

Limited Polymorphism of the *Plasmodium vivax* Merozoite Surface Protein 1 Gene in Isolates from Turkey

Fadile Yildiz Zeyrek, Shin-Ichiro Tachibana, Fehmi Yuksel, Nebiye Doni, Nirianne Palacpac, Nobuko Arisue, Toshihiro Horii, Cevayir Coban, and Kazuyuki Tanabe*

Department of Microbiology, Harran University Medical Faculty, Sanliurfa, Turkey; Laboratory of Malariology, International Research Center of Infectious Diseases, Department of Molecular Protozoology, Research Institute for Microbial Diseases, and Laboratory of Malaria Immunology, World Premier International Immunology Frontier Research Center, Osaka University, Osaka, Japan

Abstract. The 200-kD merozoite surface protein of *Plasmodium vivax* (PvMSP-1) is one of the leading vaccine candidates against *P. vivax* malaria. However, the gene encoding PvMSP-1 (*pvmsp1*) is highly polymorphic and is a major obstacle to effective vaccine development. To further understand polymorphism in *pvmsp1*, we obtained 30 full-length *pvmsp1* sequences from southeastern Turkey. Comparative analysis of sequences from Turkey and other areas showed substantially limited polymorphism. Substitutions were found at 280 and 162 amino acid sites in samples from other regions and those from Turkey, respectively. Eight substitutions were unique to Turkey. In one of them, D/E at position 1706 in the C-terminal 19-kD region, the K/E change at 1709 was the only polymorphism previously known. Limited diversity was also observed in microsatellites. Data suggest a recent population bottleneck in Turkey that may have obscured a signature for balancing selection in the C-terminal 42-kD region, which was otherwise detectable in other areas.

INTRODUCTION

Malaria imposes a huge public health burden in tropical and subtropical countries with one million deaths every year.[1] Of the five human malaria parasites, *Plasmodium vivax* is the most prevalent in Asia, Melanesia, the Middle East, South and Central America, accounting for 70–80 million cases annually.[2] Although *P. vivax* is often regarded as a benign and self-limiting infection, it can lead to debilitating illness and remains a major cause of morbidity in malaria-endemic countries. Emergence of drug-resistant *P. vivax* makes the control of *P. vivax* malaria more difficult.[3] Thus, there is an increasing demand for developing effective vaccines against *P. vivax* malaria with potential targets directed against the asexual blood stages, which are responsible for clinical manifestations of the disease.[4] The 200-kD merozoite surface protein 1 (MSP-1), which is abundantly expressed on the surface of merozoites, is one of the leading asexual blood stage vaccine candidates.[5] MSP-1 is conserved in all *Plasmodium* species[6] and is essential for parasite survival.[7]

Plasmodium falciparum MSP-1 undergoes proteolytic processing, producing four major polypeptides of 83-kD, 30-kD, 38-kD, and 42-kD.[5] Coincident with erythrocyte invasion, the C-terminal 42-kD protein is further cleaved to produce the N-terminal 33-kD and C-terminal 19-kD fragments, with all, except one, processed fragments shed.[8] The C-terminal 19-kD fragment, which contains cysteine-rich epidermal growth factor–like domains, remains anchored to the merozoite membrane and is carried into the invaded erythrocytes. Both the 42-kD and 19-kD polypeptides are considered to be promising vaccine candidates for *P. falciparum* and *P. vivax*.[4,9] However, undoubtedly, the gene encoding MSP-1 (*msp1*) is highly polymorphic[6,10,11] and, thus, presents a major obstacle to effective vaccine development. *Plasmodium vivax msp1* (*pvmsp1*) shows extensive allelic variation and is subject to balancing selection,[6,12] suggesting an involvement of parasite evasion from host immune attack.

According to inter-allelic sequence variation, *pvmsp1* contains six highly polymorphic regions interspersed with conserved blocks.[13] Of note is the limited polymorphism in the C-terminal 19-kD in *pvmsp1* with only one amino acid substitution, K/E at 1709,[14] in contrast to five major amino acid substitutions in *P. falciparum msp1*.[11] With its polymorphic nature, *pvmsp1*, particularly the highly polymorphic poly Q region in block 6,[13] previously referred to as conserved block (CB) 5,[15] has frequently been used as a molecular marker to monitor genetic diversity of *P. vivax* in different populations.[16–23] However, investigations on polymorphism of the whole *pvmsp1* have until now been limited.[13]

In this study, we report polymorphism of *pvmsp1* from isolates in Sanliurfa, southeastern Turkey, where malaria has long been one of the most common infectious diseases and *P. vivax* has continuously been identified as the only *Plasmodium* species. Our previous study has identified a high prevalence of persons with naturally acquired antibodies to PvMSP1 in southeastern Turkey.[24] We obtained 30 full-length *pvmsp1* sequences from the same study area. Comparative analysis of polymorphism in *P. vivax* populations from Sanliurfa and other areas showed substantially limited *pvmsp1* polymorphism in parasite populations in Turkey with some polymorphism unique to this country.

MATERIALS AND METHODS

Parasite isolates and DNA extraction. *Plasmodium vivax* isolates were obtained in two towns, Siverek and Harran, in Sanliurfa Province, southeastern Turkey (Figure 1), where malaria persists throughout the year with high rates during July–November.[24] According to the World Health Organization, in 2006 *P. vivax* transmission was reported in seven provinces of Turkey and 84% of the cases were from Diyarbakir and Sanliurfa.[1] In the past decade, malaria incidence rapidly decreased in this area because of government malaria control efforts that used chloroquine and primaquine.[1] Turkey shows a strong political commitment to the Tashkent Declaration, endorsed in 2005, and malaria surveillance activities have

*Address correspondence to Kazuyuki Tanabe, Laboratory of Malariology, International Research Center of Infectious Diseases, Research Institute for Microbial Diseases, Osaka University, 3-1, Yamada-Oka, Suita, Osaka 565-0871, Japan. E-mail: kztanabe@biken.osaka-u.ac.jp

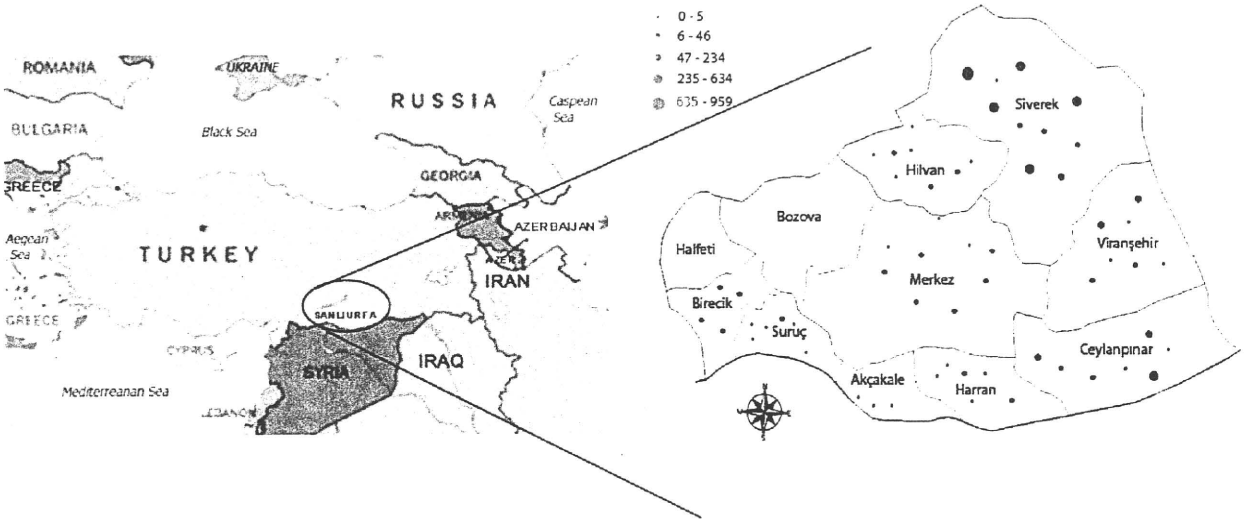


FIGURE 1. Map of Turkey showing Sanliurfa, the study area, where *Plasmodium vivax* isolates were obtained in Siverek and Harran provinces. Circles on the right indicate the number of malaria cases each year during 2001–2009 in provinces in Sanliurfa.

been intensified throughout the country, and priority has been given to provinces in southeastern Anatolia.

A total of 31 blood samples were obtained from patients diagnosed microscopically with *P. vivax* infection at several National Malaria Control Centers within Siverek and Harran: 20 from Siverek during July–December in 2007 and 11 from Harran during September–November in 2008. Although these samples were limited in number (n = 31), they include 32% of total *P. vivax* cases (n = 96) in the study area: 47 in 2007 and 49 in 2008 (Figure 2). Mean age of patients was 21 years (range = 2–55 years) and 61% were male. Giemsa-stained thick blood smears were used to calculate parasitemia (parasites/microliter of blood) as described.[25] Mean ± SD parasite density was 5,057 ± 757 parasites/μL (range = 480–14,720 parasites/μL). The patients were selected randomly from different ethnic and racial groups. All samples were collected after informed consent was obtained from patients or their parents. Sampling authorization was obtained from the Turkish Ministry of Health Sanliurfa Bureau, and ethical approval was obtained from the Research Institute for Microbial Diseases, Osaka University.

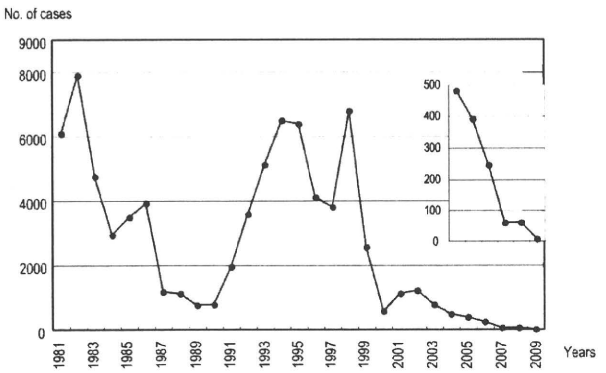


FIGURE 2. Annual incidence of *Plasmodium vivax* cases in Sanliurfa, southeastern Turkey, according to the National Malaria Control Center in Sanliurfa during 1981–2009. The y-axis is scaled for 2004–2009 to show the incidence (**inset** on the right).

An aliquot of venous blood (100–200 μL) was taken by finger prick, spotted onto Whatman® 31ETCHR filter paper (Whatman, Piscataway, NJ), and air-dried. DNA was extracted from filter blots using the EZ1 BioRobot™ (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Sequencing. Full-length *pvmsp1* (5.2–5.3 kb) was amplified by polymerase chain reaction (PCR) using Takara LA Taq (Takara Bio, Otsu, Japan) in a 20-μL reaction mixture as described[26] with primers PVF0 and PVR0.[13] Forty cycles of amplification (20 seconds at 93°C and 5 minutes at 62°C) were preceded by denaturation at 93°C for 1 minute and followed by final elongation at 72°C for 10 minutes. The PCR product was diluted 10-fold, and a 2-μL aliquot was used as template for a second PCR amplification of 20 cycles in a 50-μL reaction mixture using primers PVF0-2 (5′-CGTACATCTTTAAACCCCACACACT-3′) and PVR0. The PCR products were purified by using QIAquick (PCR Product Purification kit; Qiagen). DNA sequencing was performed directly from two independent PCR products, using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) and an ABI 3130 Genetic Analyzer (Applied Biosystems). Sequencing primers were designed to cover specific regions in both directions as described.[13] Four microsatellite loci (MS8, MS9, MS15, and 3.502)[27,28] were also sequenced after PCR amplification by using specific primers shown in Supplementary Table 1. The PCR conditions were the same as for *pvmsp1* and a second amplification was not performed. Contiguous sequences were constructed by using ATGC version 4.01 (Genetyx Corp., Tokyo, Japan). Mixed genotype infections judged from overlapping peak at given positions in an electropherogram were excluded from further analysis. Sequences obtained in this study have been deposited to DNA Database of Japan/European Molecular Biology/GenBank under accession numbers AB564559–AB564588.

Statistical analysis. The *pvmsp1* sequences obtained in this study were analyzed with previously published full-length

sequences (n = 43) from Thailand (n = 20), Brazil (n = 9), South Korea (n = 4), India (n = 1), El Salvador (n = 1), Bangladesh (n = 5), Vanuatu (n = 2), and Côte d'Ivoire (n = 1) (GenBank accession numbers AF435593–AF435599, AF435601–AF435620, AF435622–AF435625, AF435627, AF435629–AF435632, AF435634–AF435639, and DQ220742) (http://www.ncbi.nlm.nih.gov/Genbank/index.html). Between-population comparison of full-length *pvmsp1* was made by using samples from Turkey, Thailand and Brazil. Isolates from Thailand were obtained from Tak Province in 1997–1998, and isolates from Brazil were obtained in Rondonia in 1995 and 1997.[13] Partial sequences of the C-terminal 42-kD region from isolates from India (n = 28)[29] were obtained from GenBank. Partial *pvmsp1* sequences were also obtained from Azerbaijan (n = 36),[19] Iran (n = 191),[21,22,30] Afghanistan (n = 57),[23] Pakistan (n = 33),[30] Myanmar (n = 135),[31] Thailand (n = 33),[32] China (n = 33),[33] and Brazil (n = 78)[16] to analyze polymorphism in the tandem Gln (Q) repeat region in block 6[13] or CB5 previously defined.[15] Frequency distribution of poly Q haplotypes was made for six countries, from which > 20 sequences were available. Sequences were aligned by using CLUSTAL W[34] implemented in MEGA software version 4[35] with manual corrections.

Sequence polymorphism was estimated by using the S, the number of polymorphic nucleotide sites; the number of singleton polymorphic sites; the number of polymorphic amino acid sites; the number of haplotypes and haplotype diversity h; $\theta\pi$, the observed average number of pairwise nucleotide difference per site; and θ_S, the standardized number of polymorphic nucleotide sites (S) per site in the sample expected under neutrality.[36] Tajima's D statistic was estimated for testing departure from neutrality with focus on allele frequency spectrum.[37] We used Tajima's D, which compares $\theta\pi$ and θ_S. Under neutrality, the value is expected to be 0; significantly positive values suggest recent population bottleneck or balancing selection, and negative values suggest population growth or directional selection. We also used Fu and Li's D* and F* tests to test for excess or lack of singleton nucleotides by comparing estimates of θ_S based on the number of singletons versus that derived from S (the D* index) or $\theta\pi$ (the F* index).[38] An excess of intermediate frequency polymorphisms or a lack of rare variants (including singleton nucleotides) results in positive values for D* and F*. All estimates were calculated by using DnaSP software version 4.10[39] and MEGA software.

The mean numbers of synonymous substitutions per synonymous site (dS) and nonsynonymous substitutions per nonsynonymous site (dN) were estimated by using the Nei and Gojobori method[40] with the Jukes and Cantor correction as implemented in MEGA. Standard error was determined by 1,000 bootstrap replications, and dN and dS were compared by using a Z-test of selection implemented in MEGA. If dN was significantly greater than dS, balancing selection appears to be acting. If dS was greater than dN, purifying selection is predicted. The McDonald-Kreitman test[41] was also used to assess a signature for selection, in which the ratio of nonsynonymous and synonymous substitutions was compared between polymorphic (within species) and fixed difference (between closely related species) by using DnaSP. Under neutrality these ratios will be similar, whereas an excess of intraspecific nonsynonymous polymorphisms is suggestive of balancing selection. Ten *msp1* sequences from *P. cynomolgi*,[12] a monkey malaria parasite closely related to *P. vivax*, were used for between-species comparison. Fisher's exact test was used to test for statistical significance. Microsatellite haplotypes (alelles) were determined by counting the number of microsatellite repeats.

RESULTS

Polymorphism of *pvmsp1* in isolates from Turkey. The PCR amplification of the whole *pvmsp1* was successful for 31 *P. vivax* isolates, and 30 full-length sequences were obtained: n = 20 from Siverek and n = 10 from Harran. One isolate from Harran had mixed genotype infection and was not used for further analysis. An alignment of the deduced amino acid sequence of *pvmsp1* from Turkey and other areas (n = 73) showed a number of sequence regions with insertions/deletions and/or tandem repeats of degenerative repeat units, which were scattered throughout the gene (Supplementary Figure 1). Because these sequence regions were not reliably aligned, they were excluded from further analysis, except for a region containing poly Q tandem repeats in block 6, which was used for repeat number polymorphism analysis (see below).

In the sequence region analyzed (4,728 basepairs), there were 543 and 299 polymorphic nucleotide sites in worldwide (n = 73) and Turkey samples, respectively (Table 1). The number of polymorphic sites in Turkey was less than two-thirds of that in Thailand and Brazil. In the 1,576 amino acid sites analyzed, amino acid substitutions were found in 280 and 162 sites in worldwide and Turkey samples, respectively (Supplementary Figure 1). Of the 162 amino acid changes, 8 were newly identified and not observed in other areas. These are N/S at amino acid position 60 (after the Sal-1 sequence), D/N at 931, K/N at 956, S/K at 958, G/T at 959, T/S at 968, P/A at 971, and D/E at 1706. Within Turkey, the number of polymorphic nucleotide and amino acid sites was somewhat lower in Harran than in Siverek. Of note, one of the eight changes that occurred in the C-terminal 19-kD region (D/E substitution at 1706) is a new report because, up to now, only the K/E change at 1709 was the sole polymorphism reported.[13,14] The K/E change was not detected in samples from Turkey.

The number of *pvmsp1* haplotypes was only three in Turkey with haplotype diversity (h) of 0.536, which is much lower than that in Thailand and Brazil (h = 0.974 and 0.944, respectively). Moreover, the three haplotypes found in Turkey were unique to this country. Both $\theta\pi$ and θ_S were also lower in Turkey (56% and 67%, respectively) than worldwide values (Table 1).

Polymorphism in *pvmsp1* poly Q region. The *pvmsp1* contains a highly polymorphic sequence region, which is characterized by the presence (Belem type) or absence (Sal-1 type) of tandem repeats of Gln (Q) residues in amino acid positions 726–748 in the Belem sequence (GenBank accession no. AF435594). Polymorphism of this region in isolates from Turkey was compared with that reported from 8 countries (Table 2). All samples from Turkey had Belem type and no Sal-1 type, whereas both types were present in samples from all other countries analyzed. Haplotype diversity in this sequence region was lowest in Turkey among nine countries for which frequency data are available. In Belem type, the number of Qs (10–31) and the frequency distribution of Q repeat haplotypes greatly varied among seven areas (Supplementary Figure 2). The frequency distribution in Turkey is clearly different from that in Iran and Azerbaijan, and the distribution in Iran and Afghanistan is apparently similar.

TABLE 1

Polymorphism and tests for departure from neutrality in *Plasmodium vivax* merozoite surface protein 1 from Turkey and other areas*

Country	No.	S	Ss	Sa	H	*h*	$\theta\pi$	θS	Tajima D	Fu and Li D	Fu and Li F
Turkey	30	299	0	162	3	0.536 ± 0.077	0.02417 ± 0.00336	0.01596 ± 0.00092	1.7642 ($0.05 < P < 0.1$)	1.9747 ($P < 0.02$)	2.2545 ($P < 0.02$)
Siverek	20	299	0	162	3	0.426 ± 0.122	0.01720 ± 0.00512	0.01783 ± 0.00103	−0.2955 ($P > 0.1$)	1.8184 ($P < 0.02$)	1.3738 ($P > 0.1$)
Harran	10	224	0	119	2	0.533 ± 0.095	0.02610 ± 0.00448	0.01675 ± 0.00112	2.5404 ($P < 0.01$)	1.6706 ($P < 0.02$)	2.1353 ($P < 0.02$)
Thailand	20	495	24	259	16	0.974 ± 0.025	0.03955 ± 0.00309	0.02951 ± 0.00133	0.7668 ($P > 0.1$)	1.2955 ($0.05 < P < 0.1$)	1.3254 ($P > 0.1$)
Brazil	9	449	66	233	7	0.944 ± 0.070	0.04263 ± 0.00381	0.03494 ± 0.00165	0.7356 ($P > 0.1$)	1.0137 ($P > 0.1$)	1.0640 ($P > 0.1$)
Worldwide	73	543	26	280	36	0.918 ± 0.024	0.04326 ± 0.00115	0.02363 ± 0.00101	NA	NA	NA

*S = no. polymorphic nucleotide sites; Ss = no. singleton polymorphic nucleotide sites; Sa = no. polymorphic amino acid sites; H = No. haplotypes; *h* = haplotype diversity; $\theta\pi$ = observed average number of pairwise nucleotide difference per site ± SD; θS = standardized number of polymorphic nucleotide sites per site ± SD; NA = not available.

TABLE 2

Polymorphism of the *Plasmodium vivax* merozoite surface protwein 1 block 6 poly Q repeat region in isolates from Turkey and other areas*

Country	No.	H	*h*	No. Sal-1 type (%)	No. Belem type (%)	References
Turkey	30	3	0.536 ± 0.077	0 (0)	30 (100)	This study
Azerbaijan	36	7	0.614 ± 0.088	9 (25)	27 (75)	15
Iran	191	22	0.848 ± 0.016	80 (42)	111 (58)	41–43
Afghanistan	57	15	0.821 ± 0.039	31 (54)	26 (46)	44
Pakistan	33	9	0.873 ± 0.027	13 (39)	20 (61)	43
Myanmar	135	7	0.794 ± 0.016	111 (82)	24 (18)	20
Thailand	33	9	0.739 ± 0.057	28 (85)	5 (15)	23, 26
China	33	8	0.712 ± 0.059	17 (52)	16 (48)	47
Brazil	78	12	0.859 ± 0.016	25 (32)	53 (68)	3, 26

H = no. haplotypes; *h** = haplotype diversity ± SD.

Departure from neutrality. Three tests were applied to detect departure from neutrality in *pvmsp1*. The Z-test for a difference between dN and dS showed significantly higher dS than dN (Figure 3) in Turkey, Thailand, and Brazil. When the whole sequence (4,728 basepairs) was divided into three regions (the N-terminal, central, and C-terminal regions), dS was significantly higher than dN in the N-terminal and central regions for all three countries, whereas in the C-terminal 42-kD region, dN was significantly higher than dS in Thailand, Brazil, and India but not in Turkey. In the 19-kD C-terminal region, dN > dS was not significant because there were only two substitutions in this region (Supplementary Figure 1).

Using Tajima's D statistics and Fu and Li's statistics (D* test and F* test), we found that Tajima's D value was weakly (but not significantly) positive only in Turkey. Fu and Li's D* and F* values were significantly positive for Turkey but not for Thailand and Brazil. Within Turkey, samples from Harran yielded significantly positive values for the three statistics. If we consider recent rapid reduction in the incidence of malaria in the study areas (Figure 2), these results suggest a recent bottleneck in *P. vivax* populations in Turkey.

The McDonald and Kreitman test showed significantly higher intraspecific nonsynonymous substitutions than synonymous substitutions in parasite populations from Turkey, Thailand, and Brazil (Table 3), suggesting balancing selection acting on *pvmsp1* in all areas. No signature of balancing selection was evident for *P. cynomolgi*, as reported.[6] Sequence regions showing balancing selection in *pvmsp1* were the central and C-terminal 42-kD regions. A population from India also showed significantly higher intraspecific nonsynonymous substitutions than synonymous substitutions in the 42-kD region. Although high intraspecific nonsynonymous substitutions over synonymous substitutions in a locus is also observed when constraints are relaxed,[42] no singleton alleles were found in samples from Turkey (Table 1), making the occurrence of relaxed constraints in *pvmsp1* unlikely.

Polymorphism in microsatellites. The number of microsatellite repeats (alleles) was only three in four loci examined, with $h = 0.55$ (range = 0.476–0.626) (Table 4). These values are considerably lower than *h* values from other areas: 0.72–0.79 in India, Laos, Thailand, and Colombia[28]; 0.86 in Vietnam[43]; 0.79 in Sri Lanka[27]; and 0.80 in Brazil.[44]

DISCUSSION

The present analysis of *pvmsp1* polymorphism showed a remarkably lower diversity in *P. vivax* populations in Turkey

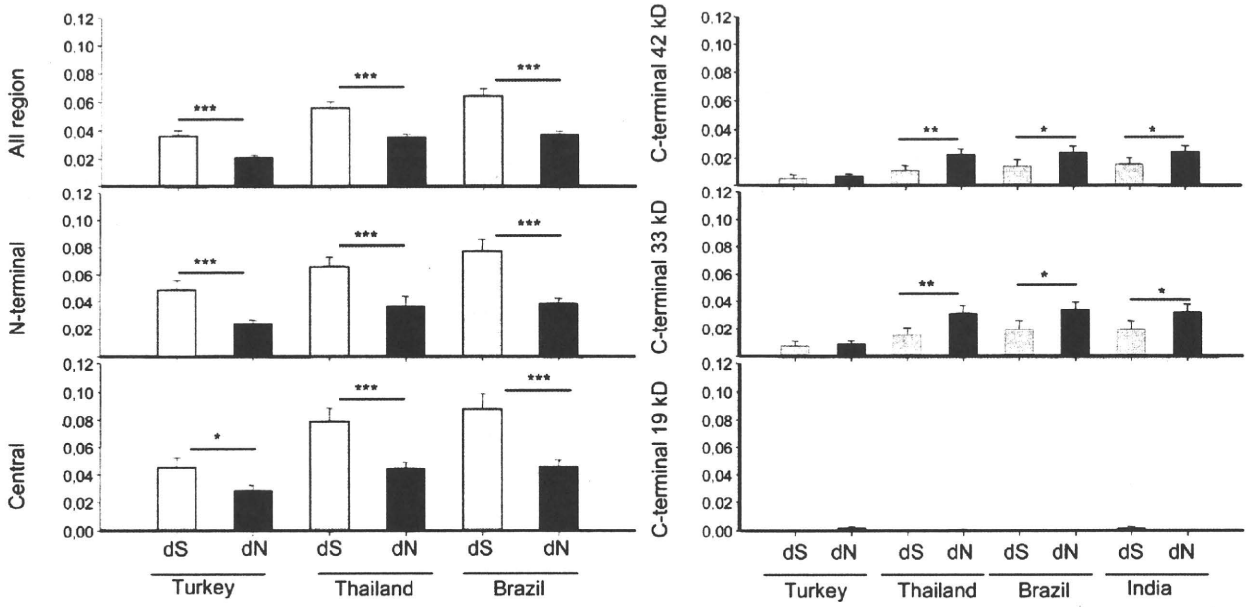


FIGURE 3. Rate of synonymous (dS) and nonsynonymous (dN) substitutions per synonymous and nonsynonymous sites in the *Plasmodium vivax* merozoite surface protein gene (*pvmspl*) from Turkey, Thailand, and Brazil. The *pvmspl* sequence was separated into three regions: the N-terminal, central, and C-terminal regions. The C-terminal 42-kD polypeptide region was further divided into the 33-kD and 19-kD regions. $*P < 0.05$; $**P < 0.01$; and $***P < 0.001$.

than in populations in Thailand and Brazil. Nucleotide diversity was 56–67% of worldwide samples. Nonetheless, there were eight new single nucleotide polymorphisms identified that were unique to Turkey isolates. This finding indicates that additional *pvmspl* sequences from various areas are still likely to be required to fully document polymorphism of the gene. Among the eight single nucleotide polymorphisms, D/E substitutions at 1706 are notable because they reside in the C-terminal 19-kD polypeptide, a candidate vaccine molecule. Given that the amino acid change is not radical, i.e., D and E are acidic residues, it may not cause a dramatic effect on the function of the polypeptide during erythrocyte invasion by the merozoite. However, the substitution may potentially lead to a change of antibody binding because the position is surface exposed.[45]

The observed low diversity of *pvmspl* in Turkey has implications regarding acquisition of immunity against *P. vivax* malaria. It is believed that repeated infections are required for persons in malaria-endemic areas to effectively mount antimalaria protective immunity. This belief largely stems from

TABLE 3

McDonald and Kreitman test results for *Plasmodium vivax* merozoite surface protein 1 from Turkey and other areas*

			WSPD			WSPD			WSPD			WSPD	
			P. vivax	*P. cynomolgi*		*P. vivax*	*P. cynomolgi*		*P. vivax*	*P. cynomolgi*		*P. vivax*	*P. cynomolgi*
pvmspl region	Substitution type	BSFD	Turkey		BSFD	Thailand		BSFD	Brazil		BSFD	India	
Whole gene	Synonymous	173	69	196	155	105	195	156	105	195			
	Nonsynonymous	176	118	191	153	182	182	162	163	184	NA	NA	NA
			$P = 0.005$	$P = 0.77$		$P = 0.007$	$P = 0.72$		$P = 0.017$	$P = 0.53$			
5′ region	Synonymous	62	47	80	51	65	79	52	67	79			
1,851 bp	Nonsynonymous	80	76	97	65	110	88	72	98	90	NA	NA	NA
			$P = 0.369$	$P = 0.784$		$P = 0.244$	$P = 0.579$		$P = 0.820$	$P = 0.413$			
Central	Synonymous	58	20	68	51	37	68	52	34	68			
1,227 bp	Nonsynonymous	39	28	48	33	51	48	34	43	48	NA	NA	NA
			$P = 0.039$	$P = 0.862$		$P = 0.014$	$P = 0.766$		$P = 0.037$	$P = 0.719$			
3′ region	Synonymous	53	2	48	53	3	48	52	4	48	52	7	47
(42 kD)	Nonsynonymous	57	14	46	55	21	46	56	22	46	55	25	46
1,098 bp			$P = 0.017$	$P = 0.682$		$P = 0.002$	$P = 0.778$		$P = 0.005$	$P = 0.679$		$P = 0.013$	$P = 0.784$
33 kD	Synonymous	43	2	39	43	3	39	42	4	39	42	6	39
789 bp	Nonsynonymous	46	13	37	45	20	37	45	22	37	45	25	37
			$P = 0.012$	$P = 0.701$		$P = 0.004$	$P = 0.754$		$P = 0.006$	$P = 0.699$		$P = 0.009$	$P = 0.698$
19 kD	Synonymous	10	0	9	10	0	9	10	0	9	10	1	8
309 bp	Nonsynonymous	11	1	9	10	1	9	11	0	9	10	0	9
			$P = 0.454$	$P = 0.863$		$P = 1.0$	$P = 1.0$		NA	$P = 0.863$		$P = 1.0$	$P = 0.879$

*pvmspl = *P. vivax* merozoite surface protein 1; BSPD = between-species fixed difference; WSPD = within-species polymorphic difference; NA = not available; bp = basepairs. *P* values < 0.05 are shaded in grey.