

RESEARCH

Open Access



Genetic polymorphism of the extracellular region in surface associated interspersed 1.1 gene of *Plasmodium falciparum* field isolates from Thailand

Natpasit Chaianantakul^{1*} , Tippawan Sungkapong¹, Jirapinya Changpad¹, Keawalin Thongma¹, Sasiwimon Sim-ut¹ and Morakot Kaewthamasorn²

Abstract

Background: A novel variable surface antigens (VSAs), Surface-associated interspersed proteins (SUFIRINs), is a protein that is modified on the surface of infected red blood cell (iRBC). Modified proteins on the iRBC surface cause severe malaria, which can lead to death throughout the life cycle of a malaria parasite. Previous study suggested that SURFIN_{1.1} is an immunogenic membrane-associated protein which was encoded by using the *surf*_{1.1} gene expressed during the trophozoite and schizont stages. This study aimed to identify the regions of SURFIN_{1.1} and investigate the genetic diversity of the extracellular region of the *surf*_{1.1} gene.

Methods: A total of 32 blood samples from falciparum malaria cases that were diagnosed in Si Sa Ket Province, Thailand were collected. *Plasmodium* genomic DNA was extracted, and the extracellular region of *surf*_{1.1} gene was amplified using the polymerase chain reaction (PCR). A sequence analysis was then performed to obtain the number of haplotypes (H), the haplotype diversity (Hd), and the segregating sites (S), while the average number of nucleotide differences between two sequences (Pi); in addition, neutrality testing, Tajima's *D* test, Fu and Li's *D** and *F** statistics was also performed.

Results: From a total of 32 patient-isolated samples, 31 DNA sequences were obtained and analysed for *surf*_{1.1} gene extracellular region polymorphism. Researchers observed six distinct haplotypes in the current research area. Haplotype frequencies were 61.3%, 16.2%, and 12.9% for H1, H2, and H3, respectively. The remaining haplotype (H4-H6) frequency was 3.2% for each haplotype. Hd was 0.598 ± 0.089 with the Pi of 0.00381, and S was 15. The most common amino acid polymorphic site was E251Q; other sites included N48D, I49V, E228D, E235S, L265F, K267T, E276Q, and S288F. Fu and Li's *D** test value was -1.24255 , Fu and Li's *F** test value was -1.10175 , indicating a tendency toward negative balancing selection acting on the *surf*_{1.1} N-terminal region. The most polymorphic region was variable 2 (Var2) while cysteine-rich domain (CRD) was conserved in both the amino acid and nucleotide extracellular region of *surf*_{1.1} gene.

Conclusions: The Thai *surf*_{1.1} N-terminal region was well-conserved with only a few polymorphic sites remaining. In this study, the data regarding current bearing on the polymorphism of extracellular region of *surf*_{1.1} gene were

*Correspondence: natpasitc@nu.ac.th

¹ Department of Medical Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok 65000, Thailand
Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

reported, which might impact the biological roles of *P. falciparum*. In addition, may possibly serve as a suitable candidate for future development of SURFIN-based vaccines regarding malaria control.

Highlights

- The regions of SURFIN_{1.1} were identified: SURFIN_{1.1} is comprised of extracellular, transmembrane (TM), and intracellular regions.
- Nucleotide and amino acid sequences of the extracellular region of *P. falciparum* SURFIN_{1.1} from a total of 31 field isolates were obtained and analyzed for genetic polymorphism: six different haplotypes were identified.
- The extracellular region of the SURFIN_{1.1} among field isolates was conserved, especially in the cysteine-rich domain (CRD) sub-region.
- High polymorphism was shown in the variable region 2 (Var2), followed by N-terminal (N-ter) and variable region 1 (Var1), respectively.
- The findings presented herein may enable the discovery and development of a novel SURFIN-based vaccine for prevention and control of malaria.

Keywords: Malaria, *Plasmodium falciparum*, Gene polymorphism, Variable surface antigens, SURFIN, *Surf* gene

Background

Malaria caused by *Plasmodium* parasitic protozoa is one of the most serious tropical diseases faced by humans and other mammals. The disease is transmitted by the *Anopheles* spp. mosquito. There are five species of the most common strains of human malaria, including *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, *Plasmodium ovale*, and *Plasmodium knowlesi*. The last species (*P. knowlesi*) is capable of transmitting between macaques and humans [1, 2]. Malaria is a global public health concern. In 2018, there were 228 million malaria cases globally, with more than 4,00,000 cases of morbidity. This disease has been especially impactful amongst children under the age of five in Africa [3]. In 2016, there were 19,079 malaria cases in Thailand resulting in 33 deaths [4]. Even though malaria cases have gradually reduced globally, the mortality amongst the number of cases remains high because of the complications involved in the form of cerebral malaria [5]. In Thailand, *P. falciparum* and *P. vivax* are the most prevalent malaria species.

Cerebral malaria is the most serious form of *P. falciparum* infection, which is a consequence of sequestration (1) binding between parasite proteins on infected red blood cell (iRBC) to the endothelial cell of the microvascular system; (2) rosette formation between iRBC and normal RBC [5]. *Plasmodium falciparum* erythrocyte membrane protein-1 (PfEMP-1) is one of the major parasite-derived surface proteins located on the iRBC membrane [6]. PfEMP-1 binds to the receptors on vascular epithelial cells and becomes part of the host cell immune evasion [5, 6]. Repetitive interspersed family proteins (RIFINs) are also a binding function of iRBC to

microvascular endothelial cells, which are also included in the rosette formation [7]. The PfEMP-1, RIFINs proteins, and the new parasite-derived surface protein, also known as surface-associated interspersed proteins (SURFINs), have been researched and characterized [6].

SURFINs are polymorphic proteins expressed on the iRBC membrane [6]. These proteins are encoded by using surface-associated interspersed genes (*surf* genes), which include *surf*_{1.1}, *surf*_{1.2}, *surf*_{1.3}, *surf*_{4.1}, *surf*_{4.2}, *surf*_{8.1}, *surf*_{8.2}, *surf*_{8.3}, *surf*_{13.1} and *surf*_{14.1}. The *surf* genes exhibit a differential expression pattern during the different erythrocytic stages of the parasite cycle [8, 9]. The most studied and well-characterized members of SURFINs are SURFIN_{4.1} and SURFIN_{4.2}, which are expressed by *surf*_{4.1} and *surf*_{4.2}, respectively [9]. SURFIN_{4.1} was found in the parasitophorous vacuole (PV), but not on the iRBC membrane. SURFIN_{4.2} is accumulated in PV and found at the knobs of iRBC with PfEMP-1 [6]. Consequently, SURFIN_{4.2} might be a function in the binding of endothelial cells as well as the immune response of host cells because of the polymorphic antigen properties of this protein [6, 10].

SURFIN_{4.2} is comprised of extracellular, transmembrane (TM), and intracellular regions [6]. The extracellular region of SURFIN_{4.2} is of importance for host immune activation, antigenic variable and antigenic polymorphism on the iRBC surface [11]. Therefore, the study of SURFIN_{4.2} protein and *surf*_{4.2} gene polymorphism is designed to assist researchers in determining the benefits of potential drugs and vaccines that were developed based on research of a well-conserved region of surface antigen. There has been much research done on the SURFIN_{4.2} and the remaining nine other SURFINs. These proteins can be used to identify the expression on

the surface of iRBC that might be the new target of the anti-malarial drug or new antigen for the malaria vaccine [8, 11, 12]. Among these SURFINs, SURFIN_{1.1} is one of the more interesting surface antigens that are expressed in the late stages of the parasite cycle.

A previous study showed that the *surf*_{1.1} gene expressed SURFIN_{1.1} protein during the trophozoite and schizont stages of malaria parasites [9]. The cytoadherence mechanisms of *P. falciparum* take place in these stages. Therefore, the SURFIN_{1.1} protein might function as a ligand for the binding of iRBC to vascular endothelial cells [6]. To date, no studies have examined on this protein. However, this protein is predicted to be a highly immunogenic membrane-associated protein in *P. falciparum* [13]. Therefore, the objectives of this study are to identify the regions of SURFIN_{1.1} in *P. falciparum* and investigate the extracellular region of the *surf*_{1.1} gene polymorphism in the field isolates. The findings from this study were presented in the context of the elucidation of gene polymorphism and potential targets for the development of malaria vaccines required to control the malaria disease, one of the most serious and devastating mosquito-borne diseases impacting humans.

Methods

Parasite isolates and DNA extraction

The research population consisted of a total 32 blood samples (isolate ID A1–A32) that were collected from symptomatic malaria-infected patients at the Kantaralak Hospital, Si Sa Ket, Thailand (2016). All clinical isolates were reported as single-species infections of *P. falciparum* as determined by light microscopic examination of Giemsa-stained blood smears. This study received ethical approval from the Naresuan University (IRB No. P10091/63).

Genomic DNA was extracted from all the isolates using a commercially available DNA extraction kit (QIAGEN, Germany) following the manufacturer's instructions. DNA was extracted from 200 µL of whole blood in a final elution volume of 200 µL. DNA samples were kept at – 20 °C before use.

PCR analysis of parasite species

Confirmation of the microscopic detection of *P. falciparum* and other potential co-infected species that might be in the samples were achieved using a nested PCR amplification assay based on the SSU rRNA gene [2, 14, 15].

Identification the regions of SURFIN_{1.1}

SURFIN_{1.1} (PF3D7_0113100) was analysed by comparing with previously well-studied SURFINs which included SURFIN_{4.1} (PF3D7_0402200), SURFIN_{4.2} (PF3D7_0424400), and SURFIN_{1.3} (PF3D7_0115000).

These SURFINs amino acid sequences were compared with the Clustal Omega program from EMBL-EBI (www.ebi.ac.uk/Tools/msa/clustalo/). The transmembrane (TM) of SURFIN_{1.1} was also determined using a TMHMM Server v.2.0 (www.cbs.dtu.dk/services/TMHMM/).

Amplification of the extracellular region of *surf*_{1.1} gene and sequencing

The extracellular region of *surf*_{1.1} was amplified with forward primer NewF (GTGCTTGTAGAAACCCC) and reverse primer NewR (CCTTTCGAGTTGTTCCAT ATAC) or forward primer NewF2 (GGTGTCTTATA TACGAAAGCG) and the same reverse primer NewR. The amplification was performed in a 50 µL reaction mixture containing a 1 × KOD-Plus-Neo buffer, 0.2 mM dNTPs, 1 mM MgSO₄, 1 U of KOD-Plus-Neo DNA polymerase (Toyobo, Japan), and a 1 µL (~ 20–40 ng) of the genomic DNA template. The Thermal cycler condition includes an initial denaturation at 94 °C for 2 min; 40 cycles of 94 °C for 30 s, 55 °C for 30 s, 68 °C for 1 min; and final extension at 68 °C for 2 min.

The PCR products were analysed using a 1% agarose gel electrophoresis after ethidium bromide staining; the PCR products were then examined under UV transillumination. The PCR products were purified using a QIAquick Gel Extraction Kit (QIAGEN, Germany) and then sequencing the nucleotide sequences with the ABI 3730 DNA analyzer (Applied Biosystems) by Macrogen, Korea.

Data analysis

The nucleotide sequences of extracellular region of *surf*_{1.1} gene were aligned using a MegAlign 15 (DNASTAR, USA). The mean numbers of synonymous substitutions per synonymous sites (d_S) and non-synonymous substitutions per non-synonymous sites (d_N) were computed using the Nei and Gojobori method using the Jukes and Cantor correction [16]. The statistical difference between d_N and d_S of a codon-based test was calculated with a one-tailed Z-test using 500 bootstrap replications in MEGA X [17]. A value of d_N over d_S at the 95% confidence level was considered significantly evident for positive selection. The deduced amino acids were translated from nucleotide sequences in order to investigate sequence diversity. Neutrality tests, based on measures of allele frequencies or heterozygosity within species were performed in DnaSP v6 software with the following analysis: Tajima's D [18] and Fu and Li's Tests [19]. Sliding window plots of nucleotide diversity, Tajima's D , Fu and Li D^* and F^* tests were carried out using a 90 bp of window length and a 3 bp of step size. The secondary structure of SURFIN_{1.1} was predicted using a PSIPRED server [20, 21].

The phylogenetic analysis was constructed in MEGA X [17] by using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model [22]. The final tree was visualized in the same software with 1000 replicates Bootstrap topologies reliability test.

Results

The regions of SURFIN_{1,1} compared with well-known SURFINS

To identify the regions of SURFIN_{1,1}, the SURFINS included SURFIN_{4,2}, SURFIN_{4,1}, and SURFIN_{1,3}, which were used as reference sequences for comparing the

SURFIN_{1,1} amino acid sequences. The amino acid sequences in the extracellular, transmembrane, and some parts of the intracellular region of SURFIN_{4,2}, SURFIN_{4,1}, SURFIN_{1,3}, and SURFIN_{1,1} were aligned (Fig. 1). The regions of each SURFIN from amino acid sequences alignment between these SURFINS were summarized, as shown in Table 1. The amino acid identity between SURFIN_{1,1} and the remaining three other SURFINS was ~ 30%. To determine the TM of SURFIN_{1,1}, the amino acid sequence of SURFIN_{1,1} was analysed by using the TMHMM server. The predicted TM of SURFIN_{1,1} was located at amino acid residues

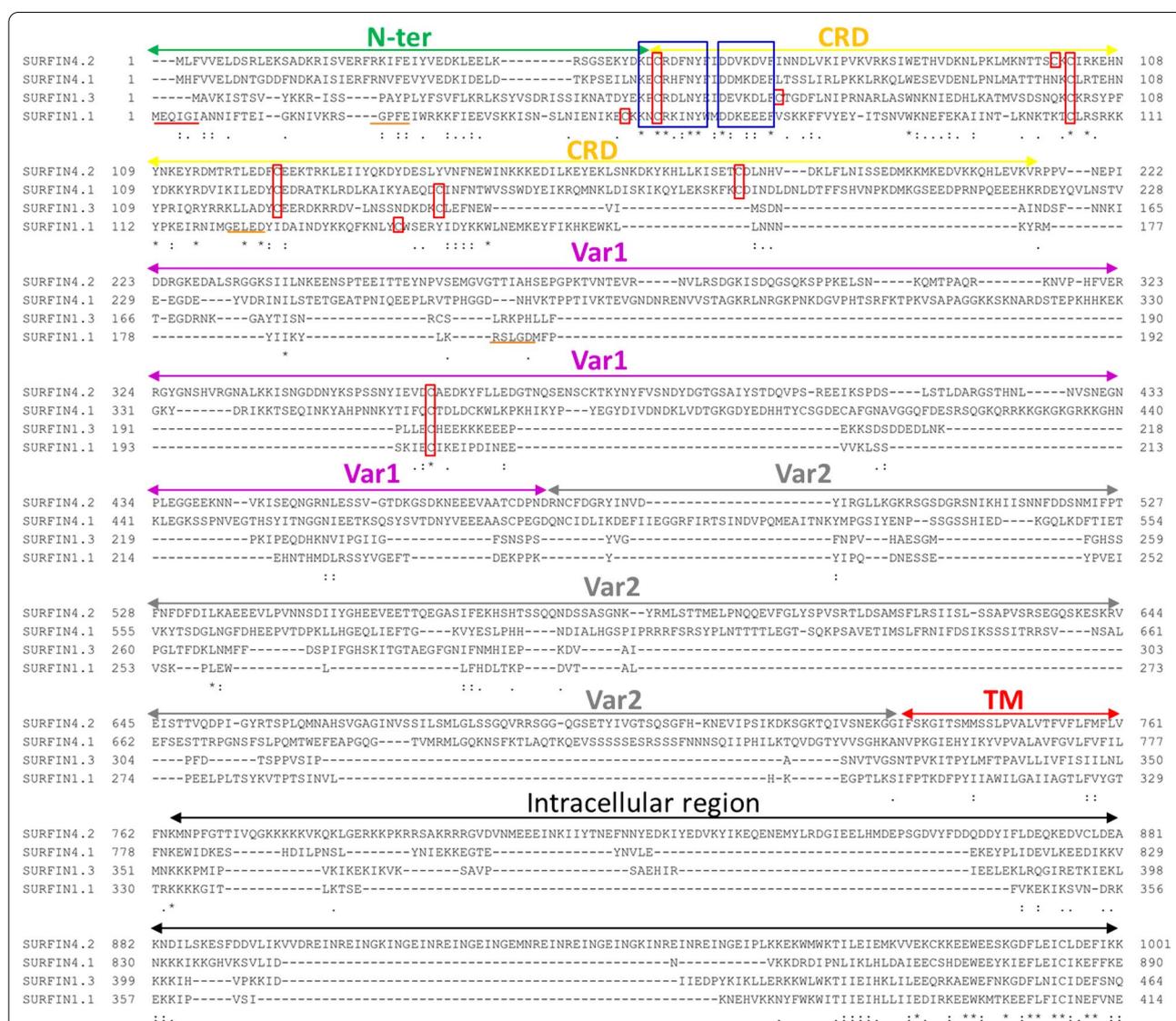


Fig. 1 Amino acid sequences alignment of SURFIN_{4,2}, SURFIN_{4,1}, SURFIN_{1,3}, and SURFIN_{1,1}. The regions of SURFINS included extracellular (N-ter, CRD, Var1, and Var2), transmembrane (TM), and intracellular regions. Cysteine residue was marked in red box. A short conserved polar amino acid sequences were marked in blue box. The extra five amino acids in SURFIN_{1,1} were marked with red underlined. PEXEL and PEXEL-like sequences were marked with brown underlined. * = Identity amino acid, : = Highly conserved amino acid, · = Semi-conserved amino acid

Table 1 The regions of SURFINs and amino acid position

SURFIN	Length (amino acid)	MW (kDa)	Extracellular region				Transmembrane region (TM)	Intracellular region	Identity to SURFIN _{1.1} (%)
			N-ter	CRD	Var1	Var2			
4.2	2380	286	1–50	51–195	196–482	483–733	734–764	765–2380	30.2
4.1	2156	258	1–50	51–198	199–492	493–749	750–780	781–2156	29.1
1.3	1925	235	1–50	51–155	156–241	242–322	323–353	354–1925	31
1.1	1555	192	1–55	56–176	177–236	237–301	302–332	333–1555	–

N-ter N-terminal region, CRD Cysteine-rich domain, Var1 Variable region 1, Var2 Variable region 2

308 to 330 (Additional file 2: Fig. S1). The predicted TM amino acid position of SURFIN_{1.1} by the TMHMM server was consistent with the TM region that was analysed by using an amino acid comparison among SURFINs. The TM region of SURFIN_{1.1} is comprised of hydrophobic amino acids, which are consistent with the PfEMP1 transmembrane domain [23]. In SURFIN_{4.1} and SURFIN_{4.2}, the TM region is needed for the trafficking of these proteins to the endoplasmic reticulum [24] and iRBC surface and Maurer’s cleft [25]. From these results, SURFIN_{1.1} consisting of three regions includes (1) Extracellular region: aa 1–301; (2) Transmembrane region: aa 302–332; and (3) Intracellular region: aa 333–1555 (Fig. 2). The intracellular region of SURFIN_{1.1} showed conserved tryptophan-rich domains (WR1–3) intersected by stretches of higher variability. The extracellular region of SURFIN_{1.1} was divided into four sub-regions based on the amino acid sequence of the SURFIN_{4.2} reference protein. These extracellular sub-regions of SURFIN_{1.1} included N-terminal (N-ter): aa 1–55, cysteine-rich domain (CRD): aa 56–176,

variable region 1 (Var1): aa 177–236, and variable region 2 (Var2): aa 237–301 (Fig. 2).

Nucleotide polymorphism of the extracellular region of the surf_{1.1} gene

From a total of 32 patient-isolated samples, 31 DNA sequences (excepted for isolate A18) were obtained and analysed for *surf*_{1.1} gene extracellular region polymorphism. The nucleotide polymorphism among the 31 patients isolated samples and *P. falciparum* 3D7 reference strain showed nucleotide identity at 99–100% with divergence of 0–1% (Additional file 1: Table S1). This result indicated that the extracellular region among field isolates *P. falciparum* was conserved. The nucleotide sequence compared with the *P. falciparum* 3D7 reference strain showed six different haplotypes (H1–H6). Haplotype frequencies were 61.3%, 16.2%, and 12.9% for H1, H2, and H3, respectively (Table 2). The remaining haplotype (H4–H6) frequency was 3.2% for each haplotype. The haplotype (gene) diversity (Hd) was 0.598 ± 0.089 with a nucleotide diversity (Pi) of 0.00381

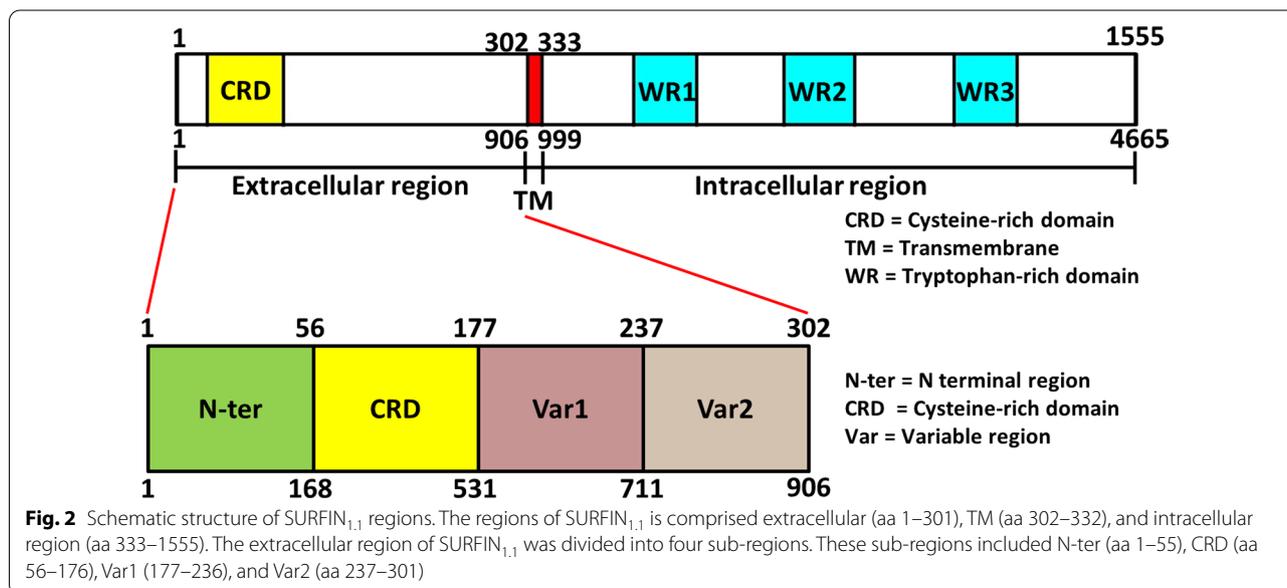


Table 2 *Surf*_{1,1} nucleotide and amino acid mutation patterns of six different haplotypes identified from Thai isolates

Haplotype	Nucleotide mutation	Amino acid mutation	Number of isolates	Frequency (%)
Haplotype 1	A142G	N48D	19	61.3
	A145G	I49V		
	G751C	E251Q		
	T756C	I252*		
	G795T	L265F		
	A800C	K267T		
	G826C	E276Q		
Haplotype 2	A142G	N48D	5	16.2
	A145G	I49V		
	G751C	E251Q		
	T756C	I252*		
	C863T	S288F		
Haplotype 3	A684C	E228D	4	12.9
	C703T	P235S		
	G751C	E251Q		
	T756C	I252*		
	C863T	S288F		
Haplotype 4	T232G	Y78D	1	3.2
	A684C	E228D		
	C703T	P235S		
	G751C	E251Q		
	T756C	I252*		
	C863T	S288F		
Haplotype 5	A142G	N48D	1	3.2
	A145G	I49V		
	G795T	L265F		
	A800C	K267T		
	C863T	S288F		
Haplotype 6	A142G	N48D	1	3.2
	A145G	I49V		
	G751C	E251Q		
	T756C	I252*		
	G795T	L265F		
	G826C	E276Q		
	A831G	L277*		
	G840A	T280*		
	G842A	T280*		
	C843T	S281N		

* Silence mutation

(Table 3). The average number of nucleotide differences (k) was 3.441. A sliding window plot of nucleotide diversity indicates an elevation in Pi values between nucleotide positions 55 to 321 and 595 to 903 (Fig. 3).

A statistical analysis of nucleotide sequences is concluded in Table 3. For the neutrality tests of the extracellular region of *surf*_{1,1}, Tajima's *D* test, Fu and Li's *D** and *F** statistics were performed. Fu and Li's *D** test statistics value was -1.24255 ($P > 0.10$), Fu and Li's *F** test statistics value was -1.10175 ($P > 0.10$). Sliding window plots of Tajima's *D*, Fu and Li's *D** and *F** statistics show a similar tendency of the test values heading in the direction of negative selection at nucleotide positions 145–321 and 664 towards the C-terminal side (Fig. 4). It is notable that, in the sliding window plots of

Table 3 Sequence analysis of *P. falciparum surf*_{1,1} extracellular region from Thai isolates

Analysis	Result	
Number of Haplotypes (H)	6	
Haplotype (gene) diversity (Hd)	0.598	± 0.089
Nucleotide diversity (Pi)	0.00381	
Variance of haplotype diversity	0.00789	
Average number of pairwise nucleotide differences (k)	3.441	
Fu and Li's <i>D*</i> test	-1.24255	$P > 0.10$
Fu and Li's <i>F*</i> test	-1.10175	$P > 0.10$
Tajima's <i>D</i> test	-0.27968	$P > 0.10$
Number of segregating sites (S)	15	
Number of non-synonymous sites (pN)	11	$P > 0.10$
Tajima's <i>D</i> (NonSyn)	0.49830	
Number of synonymous sites (pS)	4	$P < 0.05$
Tajima's <i>D</i> (Syn)	-1.88923	
pN/pS ratio	2.75	
Tajima's <i>D</i> (NonSyn/Syn) ratio	-0.26376	

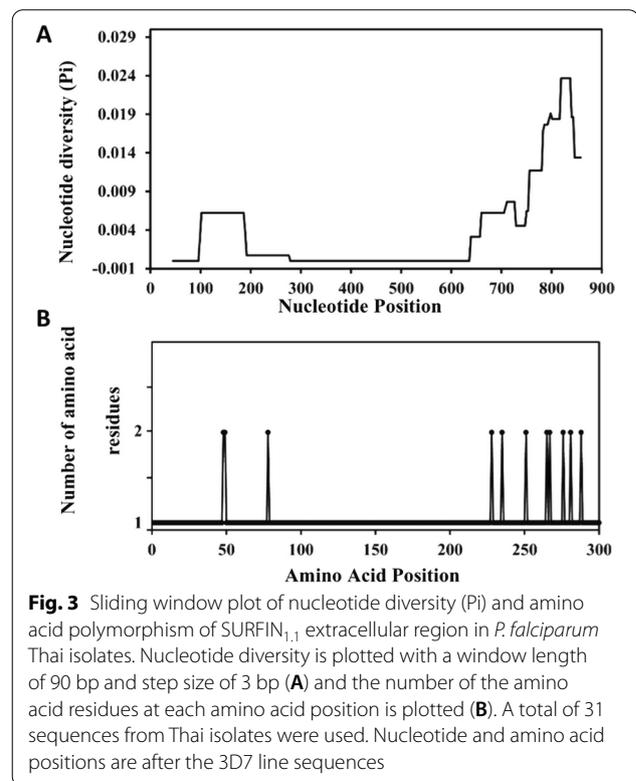
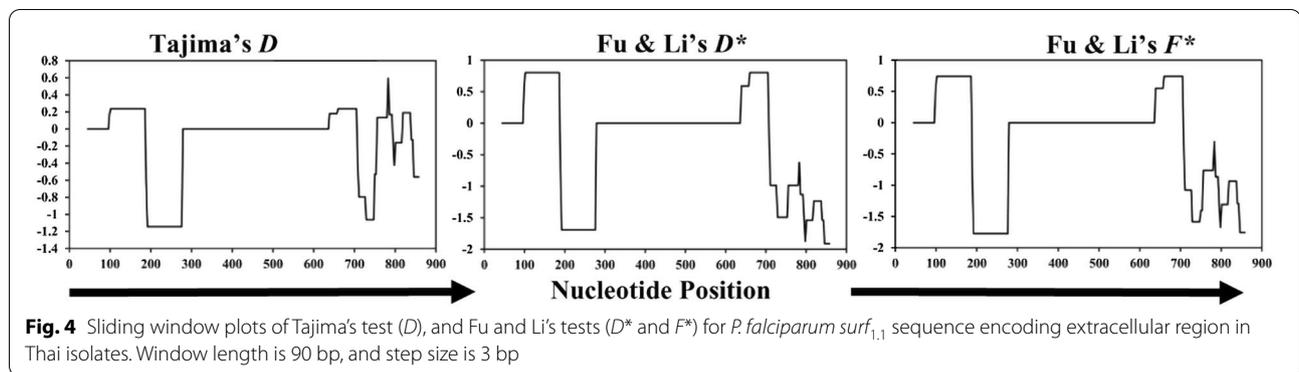


Fig. 3 Sliding window plot of nucleotide diversity (Pi) and amino acid polymorphism of SURFIN_{1,1} extracellular region in *P. falciparum* Thai isolates. Nucleotide diversity is plotted with a window length of 90 bp and step size of 3 bp (A) and the number of the amino acid residues at each amino acid position is plotted (B). A total of 31 sequences from Thai isolates were used. Nucleotide and amino acid positions are after the 3D7 line sequences

Tajima's *D*, Fu and Li's *D** and *F** tests, positive values were observed at nucleotides between position 100 to 200 and 650 to 700. However, statistical tests did not detect any significant value in each window. The number of segregating sites (S) detected in the extracellular



region of $surf_{1.1}$ was 15, and no insertion or deletions were observed in the overall samples sequenced. Tajima's D value was -0.27968 ($P > 0.10$).

The extracellular region of the $surf_{1.1}$ gene was divided into four sub-regions based on amino acid sequence conservation among SURFIN proteins. The extracellular sub-regions of $surf_{1.1}$ included N-ter (nt 1–168), CRD (nt 169–531), Var1 (nt 532–711), and Var2 (nt 712–906). Although the polymorphic sites were distributed across the entire extracellular sub-regions nucleotide sequence, high polymorphism was found in Var2, followed by N-ter and Var1 (Fig. 5A). The most common nucleotide polymorphisms were G751C and T756C (silence mutation) in Var2. Among six different haplotypes, H6 was a highly polymorphic isolate with 10 nucleotide mutation positions. These mutations included 4 silence mutations: T756C, A831G, G840A, and G842A. These silence mutations were observed in Var2. Interestingly, no nucleotide mutation in CRD was observed in H1, H2, H3, H5, and H6. However, there was a T232G mutation in H4.

Amino acid polymorphism of the extracellular region of SURFIN_{1.1} protein

The 31 nucleotide sequences were translated to amino acid sequences and analysed for SURFIN_{1.1} amino acid polymorphism. The amino acid polymorphism amongst 31 patient-isolated samples and *P. falciparum* 3D7 reference strain showed amino acid identity at 97–100% with divergence of 0–3.1% (Additional file 1: Table S2). These results indicated low genetic diversity in the extracellular region of SURFIN_{1.1}. The extracellular sub-regions of SURFIN_{1.1} included N-ter (aa 1–55), CRD (aa 56–176), Var1 (aa 177–236), and Var2 (aa 237–301). Few amino acid polymorphisms have been identified amongst these sub-regions. Amino acid mutations included N48D, I49V, Y78D, E228D, E235S, E251Q, L265F, K267T, E276Q, S281N, and S288F (Table 2). The most common amino acid mutation was E251Q. There were six different haplotypes, with the most frequent haplotype being

H1 (19 isolates, 61.3%). The amino acid mutations in H1 included N48D, I49V, E251Q, L265F, K267T, and E276Q.

Among 11 amino acid mutation positions, 2 mutations were observed in N-ter (aa position 48 and 49), 1 mutation was observed in CRD (aa position 78), 2 mutations were observed in Var1 (aa position 228 and 235), and 6 mutations were observed in Var2 (aa position 251, 265, 267, 276, 281, and 288) (Fig. 5B). Therefore, more amino acid polymorphism accumulated towards the C-terminal of the extracellular region of SURFIN_{1.1}. The amount of polymorphic amino acid sites per length of amino acids in N-ter ($2/55 = 3.6\%$) and Var1 ($2/60 = 3.3\%$) was comparable, whereas a large number of polymorphic sites were observed in Var2 ($6/65 = 9.2\%$) with six missense mutations in this sub-region. Interestingly, there was no amino acid mutation in CRD amongst five haplotypes except for H4 (Fig. 5B, 6, Additional file 3: Fig. S2). In H4, there was a Y78D mutation (Table 2) in CRD, which was found in only 1 isolate (3.2%). The N-ter of SURFIN_{1.1} in patients isolated showed N48D and I49V mutations, which were different from the *P. falciparum* 3D7 reference strain. These results indicated that the CRD was a highly conserved region among these field isolate parasites. The Var2 amino acid residues amongst field isolates showed consensus amino acid residues, including 251Q, 265F, 267T, and 276Q (Fig. 5B, Additional file 3: Fig. S2).

Sequence analysis of the extracellular region of SURFIN_{1.1} from different haplotypes compared with reference strains

The amino acid sequences of the extracellular region from six different haplotypes were compared to global reference strains including 3D7, NF54, 7G8 (Brazil), HB3 (Honduras), Dd2 (Indochina), and D10 (Papua New Guinea) (Fig. 6). In N-ter, N48D and I49V mutations were observed in H1, H2, H5, and H6 ($26/31$ isolates = 83.9%). These amino acid mutations have been reported in the HB3, 7G8, D10, and Dd2 reference strains. Therefore, these mutations might be beneficial for the field *P.*

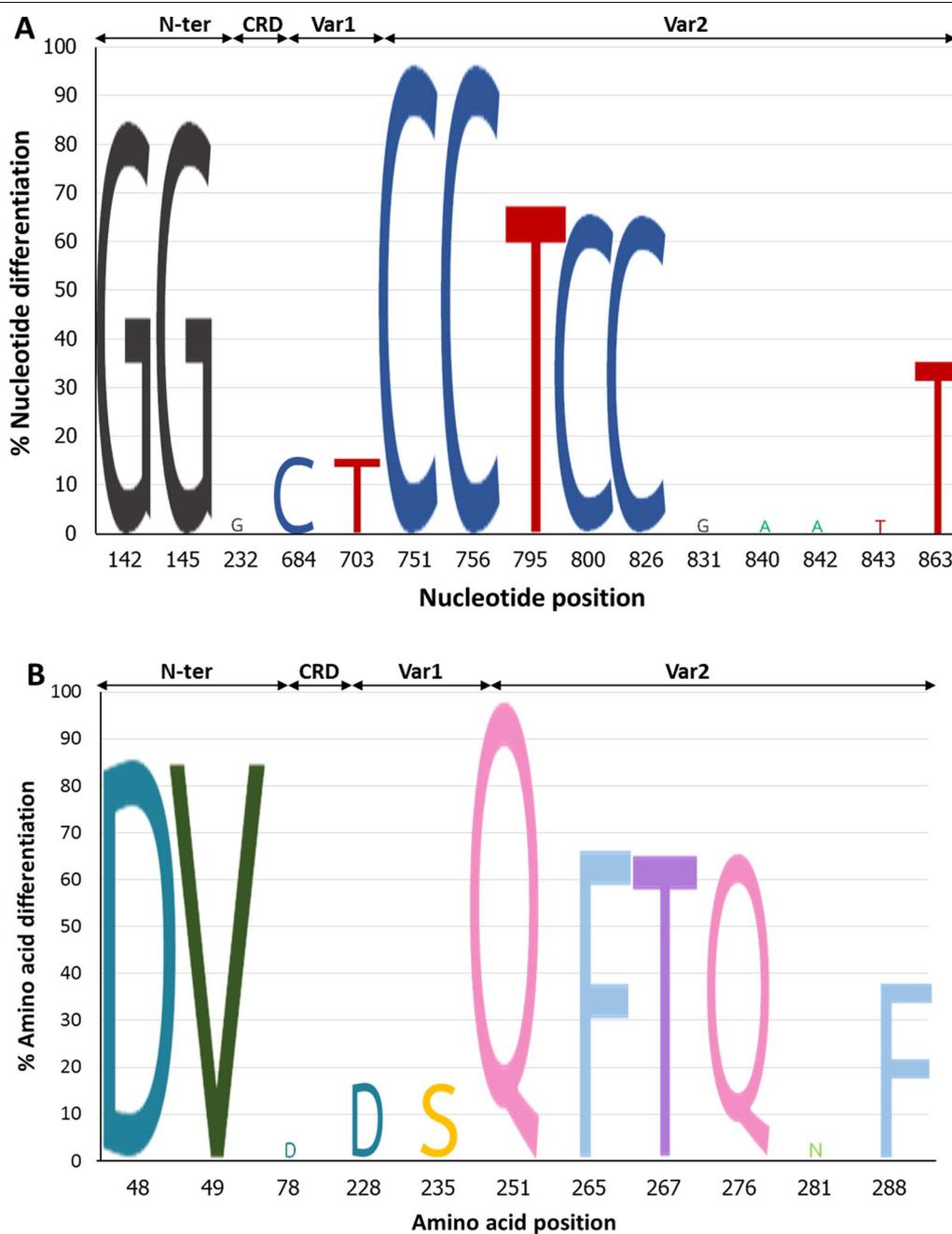
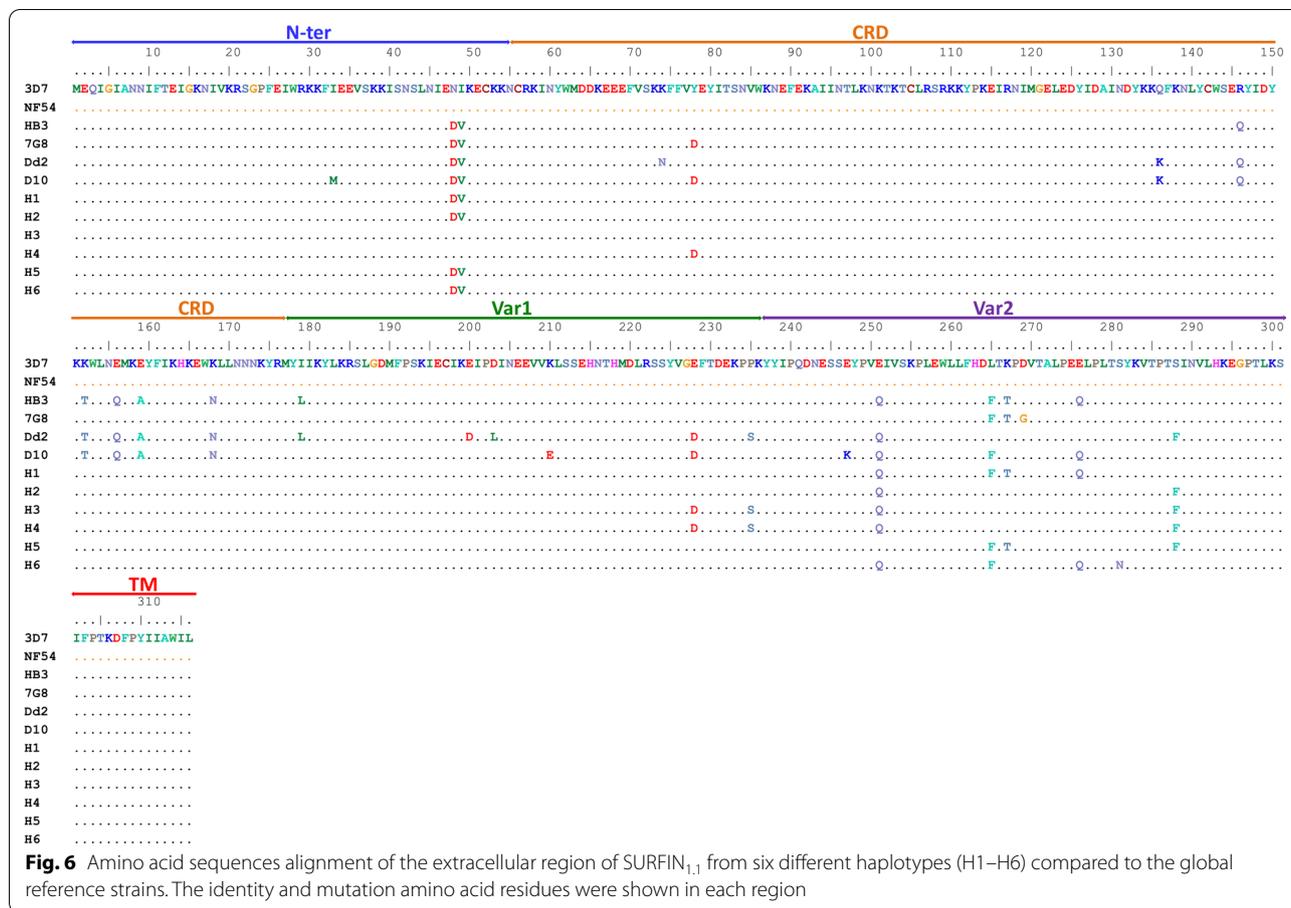


Fig. 5 Nucleotide polymorphism (A) and amino acid polymorphism (B) in the extracellular sub-regions of *surf*_{1.1} included N-ter, CRD, Var1, and Var2 compared to *P. falciparum* 3D7 reference strain. The most polymorphic region was Var2 while CRD was conserved in both the nucleotide and amino acid sequence

falciparum parasite. However, the molecular function of these amino acid residues in SURFIN_{1.1} must be clarified.

In the conserved CRD region, there was Y78D mutation that was found only in H4 (1/31 isolate = 3.2%). Interestingly, this mutation has been reported in 7G8 and D10 strains. This region was highly conserved among

Thai isolates, which were observed in H1, H2, H3, H5, H6 (30/31 isolates = 96.8%). However, other amino acid mutations in CRD have been reported in HB3, Dd2, and D10 strains. In Var1, there were E228D and P235S mutations in H3 and H4 (5/31 isolates = 16.1%). These mutations in Var1 have been reported in Dd2 and D10



strains. For the highly polymorphic region Var2, there were E251Q, L265F, and E276Q mutations in H1 and H6 (20/31 isolates = 64.5%). These mutations in Var2 have been reported in HB3 and D10 strains.

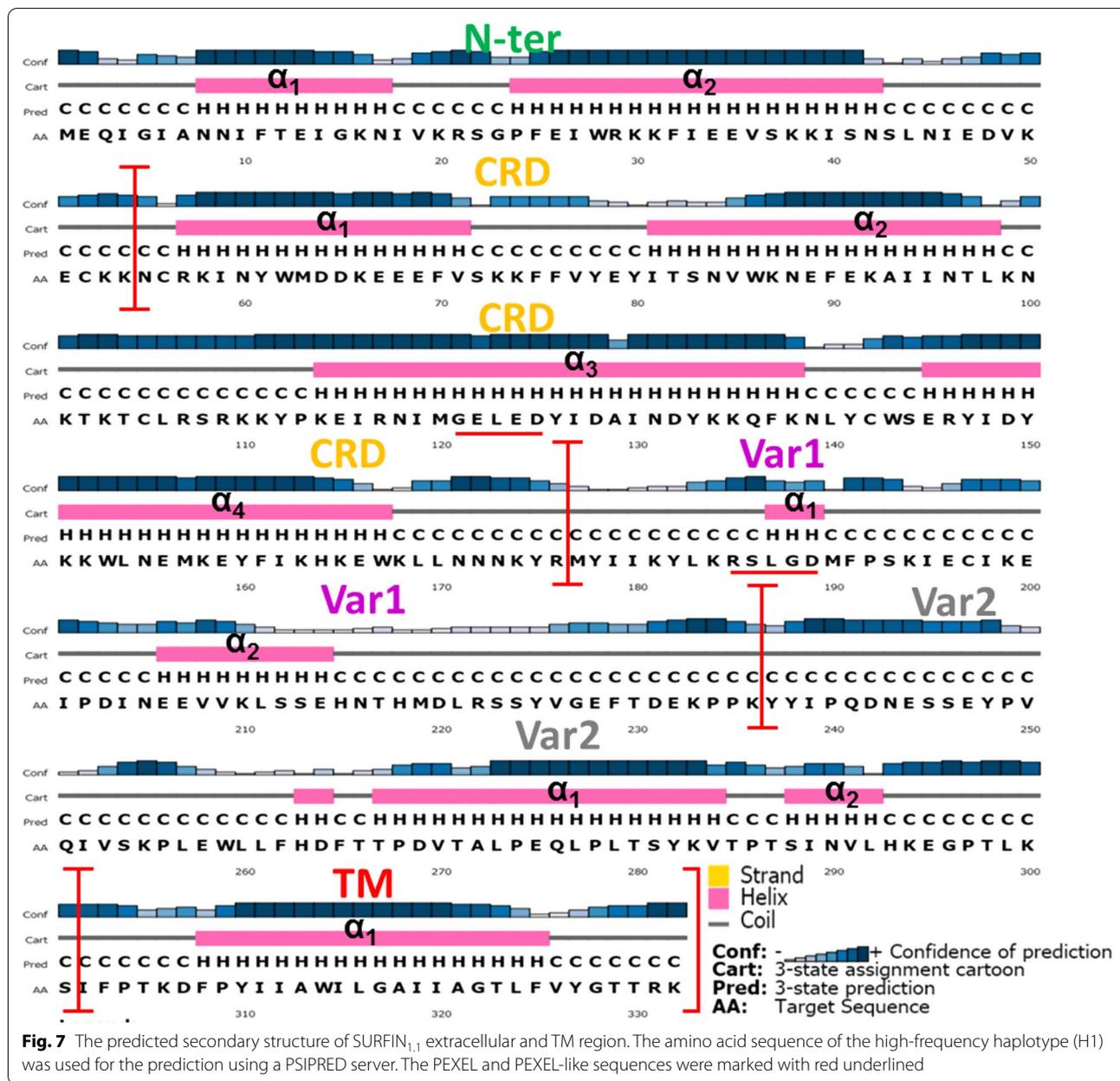
The phylogenetic tree was constructed to describe nucleotide substitution with the reference strains. The results showed that most of the Thai *surf*_{1.1} extracellular region sequences including H1 (19 isolates) and H6 (1 isolate) were not related to those reference strains (Additional file 4: Fig. S3). However, H3 (4 isolates), H4 (1 isolate), and H2 (5 isolates) were related to 3D7 and NF54 reference strains. Interestingly, H5 (1 isolate) was firmly related to the 7G8 strain.

Structural characteristics of SURFIN_{1.1} peptide chains

SURFIN_{1.1}, together with SURFIN_{4.2} and SURFIN_{4.1}, are members of the SURFIN multi-gene family. The identity between SURFIN_{1.1} to SURFIN_{4.2} and SURFIN_{4.1} was 30.2% and 29.1%, respectively (Table 1). Amino acid sequence alignment of the extracellular region from the field isolated samples showed a unique amino acid sequence. In the SURFIN_{1.1} amino acid sequence, there was a *Plasmodium* export element (PEXEL) sequence

(R₁₈₅S_LL₁₈₇G_D189) that was located in Var1. The PEXEL-like sequence (G₁₂₁E_LL₁₂₃E_D125) was also observed in CRD (Fig. 7, Additional file 3: Fig. S2). In the N-ter of SURFIN_{1.1}, there were amino acid residues 16-20 (K₁₆N_I18V_K20). Interestingly, the K₁₆, I₁₈, and K₂₀ residues were the same categories with amino acid residues of the PEXEL motif sequence (RxLxE/Q/D).

The amino acid sequence of the high-frequency haplotype (H1) was used for the prediction of SURFIN_{1.1} extracellular and TM region secondary structure. The N-ter of the SURFIN_{1.1} extracellular region showed five extra amino acid residues (MEQIGI) compared to other SURFINs. These amino acid residues were conserved amongst 31 field isolated SURFIN_{1.1} (Additional file 3: Fig. S2). The predicted secondary structure of H1 SURFIN_{1.1} showed a coil structure in the region (Fig. 7). In the conserved CRD region, most field isolates of SURFIN_{1.1} comprised four domains of alpha-helix (α₁₋₄), which might be important for the trafficking of SURFIN_{1.1}. Interestingly, in the 3rd helix (α₃) of CRD, PEXEL-like sequence (G₁₂₁E_LL₁₂₃E_D125) was shown in this region. The PEXEL-like sequence has been shown to be responsible for the trafficking of SURFIN_{4.2} [25]. In the Var1 region, there was a short



helix of PEXEL sequence (R₁₈₅L₁₈₇G₁₈₉) that is necessary for *P. falciparum*-derived proteins exported onto the iRBC surface [26].

Discussion

From previous studies, SURFINS showed structural and sequence similarity with exported iRBC surface proteins including PvSTP1, PkSICAvar, PvVIR, Pf332, and PfEMP1 [6]. However, the presence of SURFINS is different from other iRBC surface proteins. SURFIN_{4,1} is present within PV, around free merozoites as

merozoite-associated material, but is not found on the iRBC surface [9]. For SURFIN_{4,2}, it not only accumulated in the PV, but was also expressed and present in the apex of merozoite [6]. From these studies, the indication is that these SURFINS may be important to malaria parasites because the antibody to SURFIN_{4,2} was found to inhibit merozoite invasion and rosetting formation [6, 11]. Therefore, the function of SURFIN_{1,1} may involve merozoite invasion as the SURFIN_{4,2} [27]. However, the function and molecular basis of SURFIN_{1,1} need to be further clarified.

The regions of SURFIN_{1.1} were identified by comparing with SURFIN_{4.2}, SURFIN_{4.1}, and SURFIN_{1.3}. The extracellular sub-regions of SURFIN_{1.1} include N-ter, CRD, Var1, and Var2. Interestingly, the N-ter of SURFIN_{1.1} contains a stretch of five unique MEQIGI amino acid residues that are missing in other SURFINs. In a previous study, the N-ter of SURFIN_{4.1} was needed for protein translocation across the PV of the parasite [24]. Therefore, these five amino acid residues might be important for SURFIN_{1.1} trafficking. However, further study should be carried out on this stretch of five unique amino acids to verify this hypothesis. The CRD of SURFIN_{1.1} contains four cysteine residues that have been reported in SURFIN_{4.1} [9, 24] and SURFIN_{4.2} [6, 28]. However, the molecular function of these cysteine residues has not been elucidated. Amino acid sequence alignment among SURFIN_{4.2}, SURFIN_{4.1}, SURFIN_{1.3}, and SURFIN_{1.1} revealed a short conserved sequence of positively charged and polar amino acids (K_XCR_{XX}NY) as well as negatively charged amino acids (DD/E_{XX}D/E) between the N-ter and CRD region (Fig. 1). These results were consistent with a previous report that showed the highly-negative charged residues in the N-terminal are required for transportation of PfSBP1 into the iRBC [29]. Therefore, these amino acid residues might be important for the trafficking of SURFIN proteins during the erythrocytic life cycle. The PEXEL motif has been reported for parasite-derived protein trafficking in previous studies [25, 30]. However, the possibilities of these amino acid residues need to be evaluated in the future.

The intracellular region of SURFIN_{1.1} was comprised of conserved WR domains, which were consistent with SURFIN_{4.1} [24], SURFIN_{4.2} [25], and *P. vivax* transmembrane protein: PvSTP1 [6, 31]. The previous report shows that the cytoplasmic region, WR1 of SURFIN_{4.2} is needed for protein trafficking from Maurer's cleft to the iRBC membrane utilizing of co-transportation with PfEMP1 and RIFIN to the iRBC surface [28]. The function of WR domains in SURFIN_{4.2} has been demonstrated, the WR2 of SURFIN_{4.2} bound to F-actin and spectrin of RBC membrane [27]. Therefore, these WR regions might be important for the function of SURFINs during merozoite invasion to normal RBC and/or rosette formation.

Nucleotide polymorphism of the extracellular region of the *surf*_{1.1} gene compared with *P. falciparum* 3D7 reference strain showed six different haplotypes. Among these haplotypes, H1 was the most high-frequency haplotype (19 isolates, 61.3%). The extracellular region of the *surf*_{1.1} gene among field isolates was conserved, especially in the CRD sub-region. High polymorphism was shown in Var2, followed by N-ter and Var1, respectively. This conservation is also observed in its analogous gene *surf*_{4.2} [10] and *Pfcsf* gene [32]; these genes showed low

genetic diversity in the N-ter, CRD, and Var1. Fu and Li's F^* and D^* test statistic and Tajima's D test values, which indicated a probable role in negative balancing selection [18, 19, 32] occurring toward the C-terminal side of the *surf*_{1.1} extracellular region. A significant excess of non-synonymous substitutions (pN) over synonymous substitutions (pS) was detected when the entire sequence of *surf*_{1.1} extracellular region was evaluated ($p < 0.05$). The pN/pS ratio was observed at 2.75, indicating the possible role of positive selection on the extracellular region of the *surf*_{1.1} gene. This finding was consistent with a previous report that suggested an entire sequence of *surf*_{1.1} gene underwent diversifying selection with a pN/pS ratio of 4.33 [13].

Amino acid polymorphism of the extracellular region of SURFIN_{1.1} protein among 31 field isolates showed 11 amino acid polymorphic sites in four sub-regions including N-ter, CRD, Var1, and Var2. Among these regions, the most polymorphic site was the Var2 region. This result was consistent with a previous report in SURFIN_{4.2} that showed Var2 was the most highly polymorphic region [10]. Interestingly, N-ter of SURFIN_{1.1} in the field isolate parasites showed consensus amino acid residues including 48D and 49V. These amino acid residues were different from the *P. falciparum* 3D7 reference strain. These results indicated that the N-ter of SURFIN_{1.1} might be under strong diversifying selection as apical membrane antigen 1 [33]. Relatively low amino acid polymorphism in CRD of SURFIN_{1.1} supports the concept that this region could be an ideal module for a SURFIN_{1.1}-based vaccine. This concept has been reported in the RTS,S malaria vaccine. The development of a regional vaccine based on the conserved region of circumsporozoite protein (PfCSP) was proposed [32]. A consensus amino acid residues in Var2 included 251Q, 265F, 267T, and 276Q. These amino acid residues in Var2 from the field isolate parasites were different from the *P. falciparum* 3D7 reference strain. Therefore, these amino acid residues might have a positive selection for the field *P. falciparum* to overcome the host immune response. However, the molecular function of these amino acid residues in SURFIN_{1.1} must be elucidated further.

The amino acid sequences of identified haplotypes H1-H6 were compared with different reference strains including 3D7, NF54, 7G8, HB3, Dd2, and D10. The N48D and I49V mutations in N-ter of SURFIN_{1.1} (found in 26 Thai isolates) have been reported in HB3, 7G8, D10 and Dd2 strains. Therefore, this region of SURFIN_{1.1} could be considered when designing a universal SURFIN_{1.1}-based vaccine. Interestingly, the most common E251Q mutation in Var2 (found in 30 Thai isolates) has been reported in HB3, Dd2, and D10 strains. In a previous report, glutamate and lysine residues in N-ter

were shown to be important for the export of SURFIN_{4.2} to the iRBC [34]. Therefore, these amino acid mutations in the field *P. falciparum* might be important for the trafficking of SURFIN_{1.1}. However, the molecular function of these amino acid residues in SURFIN_{1.1} must be clarified. Amino acid sequence alignment between haplotypes identified from this study and reference strains showed a conserved hydrophobic TM region (Fig. 6) at amino acid positions 302–316. This result was consistent with amino acid sequence alignment among SURFINs (Fig. 1) and a predicted TM by using TMHMM servers (Additional file 2: Fig. S1). From the phylogenetic analysis, the Thai isolates haplotype 1 and 6 (20 isolates) were not related to the reference strains. Therefore, these mutation patterns observed in the SURFIN_{1.1} extracellular region from the field *P. falciparum* might be beneficial for the intra-erythrocytic development of parasites [35].

Most parasite-derived antigens on the surface of iRBC contain a pentameric amino acids sequence (R_XL_XE/Q/D) called PEXEL [36, 37] or vacuolar translocation signal/VTS [26, 38]. These specific amino acid sequences are necessary for the molecular trafficking of parasite-derived proteins to the surface of iRBC. However, the trafficking of SURFIN to the iRBC surface is PEXEL-independent but needs a specific region for transport across PV, Maurer's clefts, and ER [24, 25, 28]. The PEXEL-like sequence was proposed in the molecular trafficking of SURFIN_{4.2} to the iRBC and Maurer's clefts. In SURFIN_{4.2}, the PEXEL-like sequence includes amino acid residues R_{25K}I_{27F}E₂₉, and the PEXEL sequence includes amino acid residues R_{118T}L_{120E}D₁₂₂. Even though these amino acid sequences were not observed in the N-ter of SURFIN_{1.1}, there were amino acid residues K_{16N}I_{18V}K₂₀ in the N-ter. These amino acid residues were the same categories as the amino acid residues of the PEXEL sequence. Therefore, these amino acid residues in N-ter might be important for the molecular trafficking of SURFIN_{1.1}. However, further studies must be engaged to dissect the functional roles of this amino acid sequence in SURFIN_{1.1}.

SURFIN_{1.1} is one of the more highly immunogenic antigens among other SURFINs [13]. Therefore, SURFIN_{1.1} might be a candidate for vaccine development because of its highly conserved variant surface antigens on iRBC. A SURFIN_{1.1}-based vaccine could overcome the antigenic diversity of parasites. The conserved regions include CRD, N-ter, and Var1, which might be beneficial for a conserved epitope vaccine development [8] that could inhibit merozoite invasion or rosette formation. The antigenic diversity of parasite-derived proteins is one of the major challenges for current vaccine candidate development. The diversity of PfEMP1 and PfCSP has been reported and resulted in the escape of parasites from the host immunity [8, 32]. However, whether SURFIN_{1.1}

protein potentially elicits humoral immunity or mediates immune evasion remains to be investigated.

From the predicted secondary structure of SURFIN_{1.1}, amino acid residues MEQIGI formed a coil structure (Fig. 7). In a previous report, the coiled structure promotes peptide penetration through the cell membrane [39, 40]. Therefore, this conserved amino acid sequence might be important for the molecular function of SURFIN_{1.1}. This design might be necessary for the trafficking of SURFIN_{1.1} to iRBC and/or invasion of merozoite to RBC. From a previous study, the PEXEL cleavage and N-acetylated (Ac-xE/Q/D) contains glutamate and glutamine residues; it is recognized by the putative translocon at the PVM [30, 37]. In a previous report, glutamate residue was also found to be critical for the trafficking of REX2 [41, 42]. Therefore, glutamate and glutamine residues observed in N-ter might be important for SURFIN_{1.1}. The structural characteristics of SURFIN_{1.1}, the PEXEL-like and PEXEL sequences were observed in CRD (G_{121E}L_{123E}D₁₂₅) and Var1 (R_{185S}L_{187G}D₁₈₉), respectively. These amino acid residues might be important for SURFIN_{1.1} function and the trafficking of protein to iRBC. These amino acid residues might interact with a putative translocon protein complex that has been reported for protein trafficking [43]. However, this hypothesis must be elucidated by using the mutagenesis analysis of these amino acid residues [44, 45]. The mutation of these amino acid residues might affect the transport of SURFIN_{1.1} and altered parasite-derived surface proteins on iRBC. The predicted secondary structure of TM (Fig. 7) showed a helix structure in this region at amino acid positions 308–325. This result was consistent with SURFINs amino acid sequence alignment (Fig. 1) and TM prediction using TMHMM servers (Fig. S1).

Currently, there are anti-malarial resistant parasites because of the antigenic variation of VSAs and the amino acid mutation of targeted-malarial enzymes. Therefore, developing an effective vaccine and anti-malarial drug is a challenge. The genetic diversity of SURFIN_{1.1} in the Thai isolates presented in this study offers a new conserved surface protein on iRBC that could be a potential option for malaria vaccine development.

Conclusion

In the present work, the regions of SURFIN_{1.1} were identified. The SURFIN_{1.1} is comprised of the extracellular, TM, and intracellular regions. The extracellular region of SURFIN_{1.1} from the Thai isolates was conserved, especially in the N-ter and CRD region. These results suggested that this surface protein might be essential for *P. falciparum* during the erythrocytic life cycle of a parasite. However, further investigations are needed to determine its biological function during the *P. falciparum*

developmental stages. Through understanding this gene and protein polymorphism, it may be possible to identify the geographical distribution, changing patterns, and immunity function of this surface protein. The findings presented herein may enable the discovery and development of a novel SURFIN-based vaccine for the global prevention and control of malaria.

Abbreviations

Aa: Amino acid; CRD: Cysteine-rich domain; ER: Endoplasmic reticulum; nt: Nucleotide; N-ter: N-terminal region; PfEMP-1: *P. falciparum* erythrocyte membrane protein-1; PEXEL: Plasmodium export element; PV: Parasitophorous vacuole; REX: Ring exported proteins; RIFINs: Repetitive interspersed family proteins; SURFIN: Surface-associated interspersed protein; *surf*: Surface-associated interspersed gene; TM: Transmembrane; Var: Variable region; WRD: Tryptophan-rich domain.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12936-021-03876-y>.

Additional file 1: Table S1. Nucleotide identity among 31 field isolates (excepted for isolate A18) and *P. falciparum* 3D7 reference strain. **Table S2.** Amino acid identity among 31 field isolates (excepted for isolate A18) and *P. falciparum* 3D7 reference strain.

Additional file2: Figure S1. The predicted transmembrane region (TM) of SURFIN1.1. The TM was predicted by using TMHMM servers. The predicted TM of SURFIN1.1 was located at amino acid residues 308 to 330.

Additional file3: Figure S2. Amino acid sequences alignment of the extracellular region of SURFIN1.1 of 31 field isolates compared to *P. falciparum* 3D7 reference strain. SURFIN1.1 sequences from 31 patients-isolated were employed in the analyses. Sequence analyses revealed identity and mutation amino acid residues were shown in each region.

Additional file4: Figure S3. Molecular phylogenetic analysis of SURFIN1.1 extracellular region from different haplotypes isolates in Thailand with the global reference strains. The evolution history was conducted in MEGA X using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model. Bootstrap values below 50 are not shown. The reference strains included 3D7, NF54, 7G8, HB3, Dd2, and D10.

Acknowledgements

We express our gratitude to Thidaporn Kaosumran, Vanida Chuenjit, Narumon Chenthana, Siriluk Parimanont, and Tarinee Dansompong for their excellent assistance. We are grateful to all blood donors. The authors are grateful to Mr. Gregory A. Smith (Division of International Affairs and Language Development, Naresuan University) for critical review of the manuscript. This work was supported by Naresuan University (Grant no. R2563C027, 2020).

Authors' contributions

NC conceived the study, participated in its design, data analysis and wrote the manuscript. TS collected the blood samples in the field and designed the study. JC, KT, and SS performed gene amplification and genetic analysis of the *surf*_{1.1} gene. MK participated in data analysis. All authors have read and approved the final manuscript.

Funding

This study was supported by Naresuan University (Grant No. R2563C027, 2020).

Available of data and materials

The data sets used in this study are available from the corresponding author on a reasonable request. All nucleotide sequences used and analysed in

this study have been deposited at DDBJ/ENA/GenBank under the accession number MW767839-767869.

Declarations

Ethics approval and consent to participate

This study was received and approved by Naresuan University Ethics Committee (IRB No. P10091/63). Written informed consents were obtained from participants.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Medical Technology, Faculty of Allied Health Sciences, Naresuan University, Phitsanulok 65000, Thailand. ²Veterinary Parasitology Research Unit, Department of Pathology, Faculty of Veterinary Science, Chulalongkorn University, Bangkok 10330, Thailand.

Received: 15 April 2021 Accepted: 9 August 2021

Published online: 16 August 2021

References

- Collins WE. *Plasmodium knowlesi*: a malaria parasite of monkeys and humans. *Ann Rev Entomol*. 2012;57:107–21.
- Singh B, Sung LK, Matusop A, Radhakrishnan A, Shamsul SS, Cox-Singh J, et al. A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *Lancet*. 2004;363:1017–24.
- WHO. World malaria report 2019. World Health Organization, Geneva; 2019.
- Bureau of Vector Borne Disease. Annual report 2016. Thailand: Department of Disease Control, Ministry of Public Health; 2016.
- Idro R, Marsh K, John CC, Newton CR. Cerebral malaria: mechanisms of brain injury and strategies for improved neurocognitive outcome. *Pediatr Res*. 2010;68:267–74.
- Winter G, Kawai S, Haeggstrom M, Kaneko O, von Euler A, Kawazu S, et al. SURFIN is a polymorphic antigen expressed on *Plasmodium falciparum* merozoites and infected erythrocytes. *J Exp Med*. 2005;201:1853–163.
- Goel S, Palmkvist M, Moll K, Joannin N, Lara P, Akhouri RR, et al. RIFINs are adhesins implicated in severe *Plasmodium falciparum* malaria. *Nat Med*. 2015;21:314–7.
- Chan JA, Fowkes FJ, Beeson JG. Surface antigens of *Plasmodium falciparum*-infected erythrocytes as immune targets and malaria vaccine candidates. *Cell Mol Life Sci*. 2014;71:3633–57.
- Mphande FA, Ribacke U, Kaneko O, Kironde F, Winter G, Wahlgren M. SURFIN4.1, a schizont-merozoite associated protein in the SURFIN family of *Plasmodium falciparum*. *Malar J*. 2008;7:116.
- Kaewthamasorn M, Yahata K, Alexandre JS, Xangsayarath P, Nakazawa S, Torii M, et al. Stable allele frequency distribution of the polymorphic region of SURFIN_{4.2} in *Plasmodium falciparum* isolates from Thailand. *Parasitol Int*. 2012;61:317–23.
- Quintana MDP, Ch'ng JH, Moll K, Zandian A, Nilsson P, Idris ZM, et al. Antibodies in children with malaria to PfEMP1, RIFIN and SURFIN expressed at the *Plasmodium falciparum* parasitized red blood cell surface. *Sci Rep*. 2018;8:3262.
- Wahlgren M, Goel S, Akhouri RR. Variant surface antigens of *Plasmodium falciparum* and their roles in severe malaria. *Nat Rev Microbiol*. 2017;15:479–91.
- Krzyszczonik K, Switnicki M, Kaczanowski S. Analysis of immunogenicity of different protein groups from malaria parasite *Plasmodium falciparum*. *Infect Genet Evol*. 2012;12:1911–6.
- Singh B, Bobogare A, Cox-Singh J, Snounou G, Abdullah MS, Rahman HA. A genus- and species-specific nested polymerase chain reaction

- malaria detection assay for epidemiologic studies. *Am J Trop Med Hyg.* 1999;60:687–92.
15. Snounou G, Singh B. Nested PCR analysis of *Plasmodium* parasites. *Methods Mol Med.* 2002;72:189–203.
 16. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol.* 1986;3:418–26.
 17. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 2018;35:1547–9.
 18. Tajima F. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics.* 1993;135:599–607.
 19. Fu YX, Li WH. Statistical tests of neutrality of mutations. *Genetics.* 1993;133:693–709.
 20. Buchan DWA, Jones DT. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* 2019;47:W402–7.
 21. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol.* 1999;292:195–202.
 22. Hasegawa M, Kishino H, Yano T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol.* 1985;22:160–74.
 23. Baruch DI, Ma XC, Singh HB, Bi X, Pasloske BL, Howard RJ. Identification of a region of PfEMP1 that mediates adherence of *Plasmodium falciparum* infected erythrocytes to CD36: conserved function with variant sequence. *Blood.* 1997;90:3766–75.
 24. Zhu X, Yahata K, Alexandre JS, Tsuboi T, Kaneko O. The N-terminal segment of *Plasmodium falciparum* SURFIN_{4.1} is required for its trafficking to the red blood cell cytosol through the endoplasmic reticulum. *Parasitol Int.* 2013;62:215–29.
 25. Alexandre JSF, Yahata K, Kawai S, Torii M, Kaneko O. PEXEL-independent trafficking of *Plasmodium falciparum* SURFIN_{4.2} to the parasite-infected red blood cell and Maurer's clefts. *Parasitol Int.* 2011;60:313–20.
 26. Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estrano C, et al. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science.* 2004;306:1934–7.
 27. Zhu X, He Y, Liang Y, Kaneko O, Cui L, Cao Y. Tryptophan-rich domains of *Plasmodium falciparum* SURFIN_{4.2} and *Plasmodium vivax* PvSTP2 interact with membrane skeleton of red blood cell. *Malar J.* 2017;16:121.
 28. Kagaya W, Miyazaki S, Yahata K, Ohta N, Kaneko O. The cytoplasmic region of *Plasmodium falciparum* SURFIN_{4.2} is required for transport from Maurer's clefts to the red blood cell surface. *Trop Med Health.* 2015;43:265–72.
 29. Saridakis T, Frohlich KS, Braun-Breton C, Lanzer M. Export of PfsBP1 to the *Plasmodium falciparum* Maurer's clefts. *Traffic.* 2009;10:137–52.
 30. Chang HH, Falick AM, Carlton PM, Sedat JW, DeRisi JL, Marletta MA. N-terminal processing of proteins exported by malaria parasites. *Mol Biochem Parasitol.* 2008;160:107–15.
 31. del Portillo HA, Fernandez-Becerra C, Bowman S, Oliver K, Preuss M, Sanchez CP, et al. A superfamily of variant genes encoded in the subtelomeric region of *Plasmodium vivax*. *Nature.* 2001;410:839–42.
 32. Mohamed NS, Ali Albsheer MM, Abdelbagi H, Siddig EE, Mohamed MA, Ahmed AE, et al. Genetic polymorphism of the N-terminal region in circumsporozoite surface protein of *Plasmodium falciparum* field isolates from Sudan. *Malar J.* 2019;18:333.
 33. Polley SD, Conway DJ. Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene. *Genetics.* 2001;158:1505–12.
 34. Chitama BA, Miyazaki S, Zhu X, Kagaya W, Yahata K, Kaneko O. Multiple charged amino acids of *Plasmodium falciparum* SURFIN_{4.1} N-terminal region are important for efficient export to the red blood cell. *Parasitol Int.* 2019;71:186–93.
 35. Kanoi BN, Nagaoka H, White MT, Morita M, Palacpac NMQ, Ntege EH, et al. Global repertoire of human antibodies against *Plasmodium falciparum* RIFINs, SURFINs, and STEVORs in a malaria exposed population. *Front Immunol.* 2020;11:893.
 36. Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science.* 2004;306:1930–3.
 37. Marti M, Baum J, Rug M, Tilley L, Cowman AF. Signal-mediated export of proteins from the malaria parasite to the host erythrocyte. *J Cell Biol.* 2005;171:587–92.
 38. Tonkin CJ, Pearce JA, McFadden GI, Cowman AF. Protein targeting to destinations of the secretory pathway in the malaria parasite *Plasmodium falciparum*. *Curr Opin Microbiol.* 2006;9:381–7.
 39. Bechara C, Sagan S. Cell-penetrating peptides: 20 years later, where do we stand? *FEBS Lett.* 2013;587:1693–702.
 40. Chaianantakul N, Sungkapong T, Supatip J, Kingsang P, Kamlaithong S, Suwanakitti N. Antimalarial effect of cell penetrating peptides derived from the junctional region of *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase. *Peptides.* 2020;131:170372.
 41. Haese S, Herrmann S, Gruring C, Heiber A, Jansen PW, Langer C, et al. Sequence requirements for the export of the *Plasmodium falciparum* Maurer's clefts protein REX2. *Mol Microbiol.* 2009;71:1003–17.
 42. Dixon MW, Hawthorne PL, Spielmann T, Anderson KL, Trenholme KR, Gardiner DL. Targeting of the ring exported protein 1 to the Maurer's clefts is mediated by a two-phase process. *Traffic.* 2008;9:1316–126.
 43. de Koning-Ward TF, Gilson PR, Boddey JA, Rug M, Smith BJ, Papenfuss AT, et al. A newly discovered protein export machine in malaria parasites. *Nature.* 2009;459:945–9.
 44. Hanssen E, Hawthorne P, Dixon MW, Trenholme KR, McMillan PJ, Spielmann T, et al. Targeted mutagenesis of the ring-exported protein-1 of *Plasmodium falciparum* disrupts the architecture of Maurer's cleft organelles. *Mol Microbiol.* 2008;69:938–53.
 45. Melcher M, Muhle RA, Henrich PP, Kraemer SM, Avril M, Vigan-Womas I, et al. Identification of a role for the PfEMP1 semi-conserved head structure in protein trafficking to the surface of *Plasmodium falciparum* infected red blood cells. *Cell Microbiol.* 2010;12:1446–62.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

