Positive diversifying selection on *Plasmodium vivax* RON2 protein

JIANXIA TANG[1,2], YANG DAI[2], HONGWEI ZHANG[3], RICHARD L. CULLETON[4], YAOBO LIU[2], SONG ZHAO[2], XIAOTING WANG[2], XIAOHONG GUAN[1], OSAMU KANEKO[5] *and* YINCHANG ZHU[1,2*]

[1] *School of Basic Medical Science, Nanjing Medical University, Nanjing, Jiangsu, 210029, P. R. China*
[2] *Key Laboratory on Technology for Parasitic Disease Prevention and Control, Ministry of Health, Jiangsu Institute of Parasitic Diseases (JIPD), Meiyuan Yangxiang 117, Wuxi, Jiangsu, 214064, P. R. China*
[3] *Department of Sanitary Inspection, Henan Center of Disease Control and Prevention, Zhengzhou, Henan, 450003, P. R. China*
[4] *Malaria Unit, Institute of Tropical Medicine (NEKKEN), Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan*
[5] *Department of Protozoology, Institute of Tropical Medicine (NEKKEN) and the Global Center of Excellence Program, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan*

(*Received 18 August 2011; revised 26 October 2011; accepted 9 December 2011; first published online 10 February 2012*)

SUMMARY

Plasmodium rhoptry neck protein 2 (RON2), which is released from the neck portion of the merozoite rhoptries and interacts with the microneme protein Apical Membrane Antigen 1 (AMA1), plays a crucial role in erythrocyte invasion. In this study, we sequenced the *Plasmodium vivax* RON2 gene from 19 *P. vivax* isolates collected in central China in order to establish whether this protein is under positive diversifying selection, which may occur as a result of protective host immune pressure†. In comparison with the *P. vivax* Sal-1 reference line, we found 10 amino acid substitutions dispersed throughout the open reading frame as well as indels caused by polymorphism in a repeat unit (21–23 repeats of (Q/E/K/N/H)(G/D)G (H/L/Y/P)G) in the second tandem repeat region located at amino acid positions 541–650. A McDonald-Kreitman test with RON2 sequences from the primate malaria parasite *Plasmodium knowlesi*, detected significant departure from neutrality in the *Pv*RON2 3′ region (nucleotide positions 2668–6609). These results suggest that the *Pv*RON2 gene has evolved under positive diversifying selection.

Key words: *Plasmodium vivax*, RON2, positive diversifying selection, cell invasion.

INTRODUCTION

Malaria caused by *Plasmodium vivax* is responsible for 100–300 million clinical cases per year worldwide (Price *et al*. 2007; Greenwood *et al*. 2008). Although often considered 'benign' (at least in comparison to the disease caused by *Plasmodium falciparum*), severe *P. vivax* cases have been periodically reported (Price *et al*. 2007). Thus efficient control tools are required against not only *P. falciparum* but also for *P. vivax*.

Erythrocyte invasion is an essential step in the life cycle of malaria parasites. The stages of the *Plasmodium* parasite that invade cells contain distinctive organelles at their apical end, termed micronemes, rhoptries, and dense granules, which release proteins during invasion. Recently, in another apicomplexan parasite *Toxoplasma gondii*, a microneme protein Apical Membrane Antigen 1 (*Tg*AMA1) was found to form a complex with 3 rhoptry neck (RON) proteins: *Tg*RON2, *Tg*RON4 and *Tg*RON5 during cell invasion (Alexander *et al*. 2005; Lebrun *et al*. 2005; Boothroyd and Dubremetz, 2008). Orthologues of these genes have been identified in the genomes of malaria parasite species and some have been characterized for their expression and complex formation. The *P. falciparum* orthologue of RON4, *Pf*RON4, has been shown to associate with *Pf*AMA1 (Alexander *et al*. 2006) and to be localized at the moving junction (Baum *et al*. 2008). Complex formation between *Pf*RON2 and *Pf*AMA1 was also shown in *P. falciparum* (Cao *et al*. 2009). Protein expression at the apical end of the merozoite was confirmed for *Pf*RON5 (Curtidor *et al*. 2011). When complex formation between RON2 and AMA1 was inhibited in *P. falciparum*, the parasite could not invade erythrocytes, suggesting that the complex formed by these proteins is essential for a successful invasion (Richard *et al*. 2010). Although the amino acid sequences of the RON2 and AMA1 proteins have diverged between species, the interaction

* Corresponding author: Meiyuan Yangxiang 117, Wuxi, Jiangsu Province 214064, People's Republic of China. Tel: +86 510 8551 7721. Fax: +86 510 8551 0263. E-mail: yczhu2006@yahoo.com.cn

† Sequence data from this study have been deposited to the GenBank™/EMBL/DDBJ databases under Accession numbers AB676030 – AB676067.

doi:10.1017/S0031182011002447

between them has been maintained in both *P. falciparum* and *T. gondii*, which suggests functional conservation (Lamarque *et al*. 2011).

Molecules located on the merozoite surface are potential targets of host immune responses. Diversifying selection has been suggested to have been involved in the evolution of the microneme protein AMA1, suggesting that the genetic diversity of this gene has been maintained at the population level in order to evade host immunity (Polley and Conway, 2001). Due to the fundamental role of RON2-AMA1 complex formation in host cell invasion, RON2 is, potentially, a target of invasion inhibitory antibodies and, therefore, for the development of anti-vaccines and/or drugs. Previously, an excess of non-synonymous substitutions over synonymous substitutions was detected in the gene encoding *Pf*RON2 following the analysis of 6 culture-adapted parasite lines, suggesting that *Pf*RON2 is subject to positive diversifying selection (Cao *et al*. 2009). Similar to *Pf*RON2, *Pv*RON2 was also reported to be expressed in late schizonts and is localized at the rhoptry neck in the merozoite (Arévalo-Pinzón *et al*. 2011), but the degree of polymorphism of this protein within parasite populations has not been previously determined. In this study, we evaluated the diversity of the entire coding region of *Pv*RON2 using parasites isolated from China in order to measure the degree of the polymorphism and to identify signatures of a positive diversifying selection on this protein.

MATERIALS AND METHODS

Ethics statement

The study was approved by the Institutional Review Board (IRB00004221) of Jiangsu Institute of Parasitic Diseases, Wuxi, China. Questionnaire surveys, physical examination and laboratory work were conducted after the purpose of the study had been explained to participants, who were given the right to withdraw from the study at any time without consequences. Written informed consent was obtained from each participant.

Bioinformatics methods

A BLASTP search was performed against *P. vivax* (Sal-I line) and *Plasmodium knowlesi* (H strain) database in PlasmoDB using the *Pf*RON2 amino acid sequence as a query, and RON2 orthologues, *Pv*RON2 (PVX_117880) and *Pk*RON2 (PKH_125430), were identified (Carlton *et al*. 2008; Pain *et al*. 2008; Aurrecoechea *et al*. 2009). The transmembrane (TM) domains were predicted by TopPred 0.01 (von Heijne, 1992), SOSUI v1.11 (Hirokawa *et al*. 1998), HMMTOP v2.0 (Tusnády and Simon, 2001), TMpred (Hofmann and Stoffel, 1993), and TMHMM v2.0 (Krogh *et al*. 2001) web-based algorithms. Repeat sequences and domains were predicted by XSTREAM 1.73 (Newman and Cooper, 2007).

Source of Plasmodium vivax

P. vivax-infected blood samples were collected from 13 symptomatic patients who had acquired the infections in Henan province in 2009 (HN02, HN19, HN29, HN40, HN44, HN49, HN80, HN89, HN90, HN114, HN133, HN159, and HN160; only the 3′ region was anlaysed for HN89), 6 patients from Anhui province in 2007 and 2008 (BB0703, BB0704, BB0706, GS0702, HY0701, and P4; only the 5′ region was analysed for P4), and 2 patients from Jiangsu in 2009 (JS10 and JS17; the 5′ region was analysed for JS10 and the 3′ region was anlaysed for JS17). Infections were diagnosed by examination of Giemsa's solution-stained thick blood films. Finger-prick blood samples obtained from patients were preserved by spotting onto Whatman 3MM filter papers, naturally dried and kept at −20 °C until DNA extraction, while venous blood samples obtained from patients were preserved with EDTA anticoagulant and kept in liquid nitrogen prior to DNA extraction.

DNA isolation, polymerase chain reaction (PCR) amplification and sequencing

DNA was extracted from blood spotted on filter papers and venous blood samples using a QIAmp DNA micro-kit (QIAGEN, Germany) following the manufacturer's protocol. Extracted DNA samples were stored at −30 °C until use. The oligonucleotide primers were designed based on the nucleotide (nt) sequence of *Pv*RON2 (Sal-I line, PVX_117880) in the PlasmoDB. Two independent DNA fragments were PCR-amplified; the first one contains 5′ region encoding N-terminal amino acid positions (aa) 1–924 with the primers PvRON2-F0 (CTCTCTGGATTCACATCTCTCCGTTC) and PvRON2-R3 (CTTTACGTTCAGTCCTTCTTCTGAC) and the second one contains 3′ region encoding C-terminal aa 841–2203 with the primers PvRON2-F2 (GGATACACGCCTATCATCGTAAAGTATGAC) and PvRON2-R2 (GCACACCTGGACGAAGCTTTCTTCC). After excluding primer sequences, 2 DNA fragments share 224 nt sites. Oligonucleotides used in this study are summarized in Table 1.

PCR amplification was performed with KOD -Plus-DNA polymerase (Toyobo, Japan) in a 20 µl final reaction. Amplification conditions for the first fragment was as follows: 1 cycle of 2 min at 94 °C, followed by 40 cycles of 15 sec at 92 °C, 15 sec at 55 °C and 3 min at 68 °C, with a final 10 min extension step at 68 °C. The second fragment was

https://doi.org/10.1017/S0031182011002447 Published online by Cambridge University Press

Table 1. Oligonucleotides used in this study

Name	Sequence
For PCR amplification	
PvRON2-F0	CTCTCTGGATTCACATCTCTCCGTTC
PvRON2-R3	CTTTACGTTCAGTCCTTCTTCTGAC
PvRON2-F2	GGATACACGCCTATCATCGTAAAGTATGAC
PvRON2-R2	GCACACCTGGACGAAGCTTTCTTCC
For sequencing	
PvRON2 V-F0	GACCGCAAAAGCAAAGTAAAATG
PvRON2 V-F1	GACTCTTTACCGCACACCAA
PvRON2 V-R1	CCGTACGTGTCCTCGTCTTC
PvRON2C-F1	AGAGATGCAACCGCTAAGTG
PvRON2C-R1	CCATTGCGCTGAGGATAGTG
PvRON2C-F2	AAGAGCCTTTGATATGTTGC
PvRON2C-R2	GCTTATGCAGCTCTGCCACT

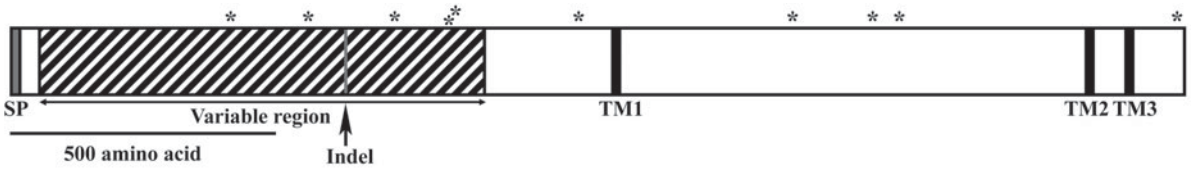

Fig. 1. Amino acid polymorphic sites of *Pv*RON2. SP and TM indicate putative signal peptide (amino acid positions 1–17) and predicted transmembrane regions, respectively. The light grey bar indicates indel. Asterisks indicate polymorphic sites.

amplified as above except 45 cycles for the second step and 4 min 30 sec for the extension time were used. PCR products were visualized on a 1% agarose gel under UV transillumination. The specific band was cut out from the gel and purified using a Wizard SV Gel and PCR Clean-Up system (Promega, Madison, WI, USA). After the addition of an A cohesive end using *Taq* DNA polymerase (Fermentas), PCR products were cloned into pGEM-T easy vector (Promega). Plasmids prepared using a Wizard *Plus* SV Minipreps DNA purification system (Promega) were sequenced using an ABI PRISM®310 genetic analyzer (Applied Biosystems, Foster City, CA). Six clones were selected for each DNA fragment and final sequences were determined from at least 3 identical clones to avoid potential error during PCR amplification and sequencing. Sequence data have been deposited to the GenBank™/EMBL/DDBJ databases.

Sequence alignment and analysis

The *Pv*RON2 open reading frame nt sequences were aligned with that retrieved from the genome database (Sal-I) using the CLUSTAL W program (Thompson *et al*. 1994) with manual correction; Nucleotide diversity (π) and its standard error (S.E.) were computed using the Jukes and Cantor method in MEGA 4.0 after removing insertion/deletions (indels). The mean number of synonymous substitutions per synonymous site (d_S) and non-synonymous substitution per non-synonymous site (d_N), and their standard errors were computed using the Nei and Gojobori method (Nei and Gojobori, 1986) with Jukes and Cantor correction, implemented in MEGA 4.0. The statistical difference

Table 2. Prediction of *PvRON2 transmembrane* domains using different algorithm

(TM and aa indicate transmembrane domain and amino acid position in *Pv*RON2 (Sal-I). TMHMM and HMMTOP use a hidden Markov model (HMM) with special architecture developed to search transmembrane topology. SOSUI use 4 physicochemical parameters: the hydropathicity index of Kyte and Doolittle, an amphiphilicity index, an index of amino acid charges, and the length of each sequence. TMpred was trained with a database of naturally occuring transmembrane proteins. For the analysis by TMpred, only the strongly preferred model is shown. TopPred considers the abundance of positively charged residues in the part of the sequence on the cytoplasmic side of the membrane, in addition to the hydrophobicity.)

Algorithm	TM1	TM2	TM3
TMHMM	not predicted	not predicted	aa 2087–2109
HMMTOP	not predicted	not predicted	not predicted
SOSUI	not predicted	not predicted	aa 2080–2102
TMpred	aa 1127–1143	aa 2016–2034	aa 2088–2109
TopPred	aa 1123–1143	aa 2015–2035	aa 2082–2102

https://doi.org/10.1017/S0031182011002447 Published online by Cambridge University Press

Table 3. Nucleotide and amino acid polymorphism of *Pv*RON2

Region	Nucleotide position (amino acid position)[a]	Type	Codon (amino acid) and frequency			
			Sal-I type		non-Sal-I type	
5′ region	1240–1242 (414)	Nsyn	GTA (Val)	1	GGA (Gly)	19
	1675–1677 (559)	Nsyn	CAT (His)	18	TAT (Tyr)	2
	2146–2148 (716)	Nsyn	AGT (Ser)	19	GGT (Gly)	1
	2455–2457 (819)	Nsyn	TAT (Tyr)	1	TCT (Ser)	19
	2476–2478 (826)	Nsyn	GGC (Gly)	1	GAC (Asp)	19
3′ region	3181–3183 (1061)	Nsyn	GCC (Ala)	19	GTC (Val)	1
	3442–3444 (1148)	Syn	GTG (Val)	19	GTA (Val)	1
	4390–4392 (1464)	Nsyn	CCG (Pro)	18	TCG (Ser)	2
	4840–4842 (1614)	Nsyn	GCA (Ala)	16	ACA (Thr)	4
	4987–4989 (1663)	Nsyn	ATC (Ile)	1	AGC (Ser)	19
	6553–6555 (2185)	Nsyn	ATT (Ile)	1	ATG (Met)	19

[a] Nucleotide and amino acid numbering are after Sal-I line. Syn and Nsyn indicate synonymous substitution and non-synonymous substitution, respectively.

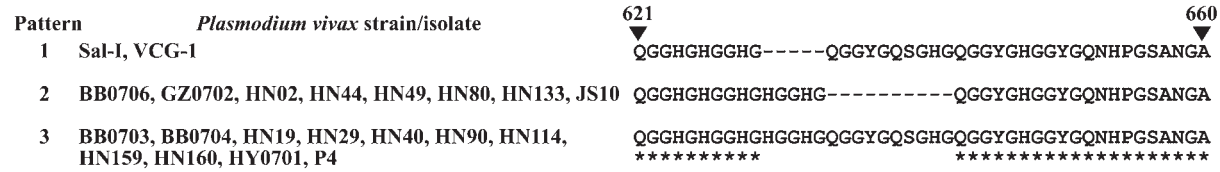

Fig. 2. Three patterns of repeat polymorphism of *Pv*RON2 second tandem repeat region.

between d_S and d_N was tested using a Fishers' exact test. Positive selection was also evaluated using the McDonald-Kreitman test with *Pk*RON2 nt sequence (McDonald and Kreitman, 1991).

RESULTS

Signal peptide, transmembrane and repeat regions of PvRON2 protein sequence

The full-length *Pv*RON2 protein consists of 2203 residues with a putative signal peptide sequence at its N-terminus at aa 1–17. Among 5 web-based tests, 4 algorithms predicted that *Pv*RON2 was a transmembrane protein bearing (potentially) 1–3 TM domains (Fig. 1 and Table 2). The putative TM domain towards the C-terminus (TM3; aa 2087–2109 by TMHMM) was predicted by 4 algorithms; however, other TM domains, TM1 (aa 1123–1143 by TopPred) and TM2 (aa 2015–2035 by TopPred), were only predicted by TMPred and TopPred. Even though we were unable to unambiguously determine the exact number of TM domains, the presence of at least 1 TM suggests that *Pv*RON2 is an intergral membrane protein.

Two tandem repeat (TR) sequences were detected in the *Pv*RON2 amino acid sequence by XSTREAM 1.73. An 8 times repeated 11 amino acid sequence 'G (A/S)(E/D)(G/R)KGYG(P/T)(Y/D)G' was located at aa 258–345, and a 5 amino acid second TR '(Q/E/K/N/H)(G/D)G(H/L/Y/P)G' was located at aa 541–650 (after Sal-I sequence). Based on the second TR number, *Pv*RON2 could be classified into 3 types; type 1 consisting of Sal-I and VCG-1 had 22 repeats, type 2 consisting of 8 China isolates had 21 repeats, and type 3 consisting of 11 China isolates had 23 repeats (Fig. 2).

Polymorphism in the PvRON2 sequence

Ten non-synonymous nt substitutions were found among 20 *Pv*RON2 sequences at nt positions 1241, 1675, 2146, 2456, 2477, 3182, 4390, 4840, 4988, and 6555 (Table 3, Fig. 2). One synonymous substitution was found at nt 3444. An indel polymorphism due to a difference in repeat number was also found at aa 541–650 (after Sal-I sequence) as described above (Fig. 2). Arévalo-Pinzón *et al*. (2011) sequenced *Pv*RON2 cDNA sequence from Colombian VCG-1 isolate and found that there were 2 non-synonymous nucleotide substitutions at nt 1241 (T to G) and 1811 (C to A) (numbers are after Sal-I sequence) and 1 insertion of a nucleotide triplet 'GAA' between nt 1485 and 1486 of Sal-I sequence in comparison with the Sal-I reference line. A substitution at nt 1241 was found in all 19 China isolates in this study; however, a substitution at nt 1811 and an indel both found in the VCG-1 isolate were not found in these 19 Chinese isolates.

Positive diversifying selection on the PvRON2 gene

Nucleotide sequences of *Pv*RON2 from 19 China isolates and the Sal-I isolate were aligned and

https://doi.org/10.1017/S0031182011002447 Published online by Cambridge University Press

Table 4. Nucleotide diversity of *Pv*RON2 gene

Region (nt)	*n*	Indel	Number of sites (base)	*Nd* (S.E.)	*N* (S.E.)	*Sd* (S.E.)	*S* (S.E.)	π (S.E.)	d_N (S.E.)	d_S (S.E.)	d_N/d_S	P ($d_N>d_S$)
5′ region (4–2667)	20	(+)	2589	0·590 (0·282)	1976·20 (8·770)	0·000 (0·000)	612·800 (9·246)	0·00023 (0·00011)	0·00030 (0·00014)	0·00000 (0·00000)	∞	n.s.
3′ region (2668–6609)	20	(−)	3942	0·826 (0·407)	3100·783 13·138	0·100 (0·093)	841·217 (13·059)	0·00024 (0·00011)	0·00027 (0·00013)	0·00012 (0·00011)	2·25	n.s.

n, number of sequences analysed.
nt, nucleotide position.
Sites, nucleotide sites analysed excluding sequence around indels.
π, pairwise nucleotide diversity.
N and *S*, average numbers of non-synonymous and synonymous sites.
d_N, number of non-synonymous substitutions over number of non-synonymous sites.
d_S, number of synonymous substitutions over number of synonymous sites.
S.E., standard error computed using the Nei-Gojobori method with Jukes-Cantor correction using the bootstrap method with 500 replications. The numbers of synonymous (*Sd*) and non-synonymous (*Nd*) differences were calculated by Nei-Gojobori method. Nucleotide positions 1864–1938 were excluded from the analysis for 5′ region to obtain an reliable alignment. Nucleotide (nt) number is after Sal-I reference line. *P* value indicates the d_N is significantly greater than d_S.
n.s. indicates not significant ($P>0·05$).

evaluated for the signature of positive selection. Nucleotide diversity (π) and its standard error of 5′ region (nt 3–2667) and 3′ region (nt 2668–6609) were $0·00023 \pm 0·00010$ and $0·00024 \pm 0·00011$, respectively (Table 4). A significant excess of non-synonymous substitutions over synonymous substitutions was not detected for both regions (Hughes and Nei, 1988). To further evaluate the signature of positive selection on *Pv*RON2, we performed a McDonald-Kreitman test using the *P. knowlesi* RON2 homologue. For the McDonald-Kreitman test, nucleotide sequences around indels and repeat regions were excluded from the analysis. The *Pv*RON2 sequences used for the analysis were nt 4–663, 676–771, 1027–1065, 1072–1206, 1360–1482, 1489–1623, 1954–2472, and 2482–6609 (Sal-I line). The *Pk*RON2 sequences used for the analysis were nt 4–663, 679–774, 997–1035, 1045–1179, 1303–1425, 1435–1569, 1768–2286, and 2302–6426. For the 5′ region, *Pv*RON2 nt 4–2747 and *Pk*RON2 nt 4–2567 were used. For the 3′ region, *Pv*RON2 nt 2558–6609 and *Pk*RON2 nt 2488–6426 were used. A significant excess of non-synonymous substitutions over synonymous substitutions was detected among *P. vivax* sequences by comparing the fixed differences between *P. vivax* and *P. knowlesi* for the 3′ region ($P < 0·02$), suggesting that positive diversifying selection has acted on this region (Table 5). This analysis did not detect positive selection on the 5′ region.

DISCUSSION

The present study describes the polymorphism in the nucleotide sequences of *Pv*RON2, a homologue of *Pf*RON2, in samples collected from *P. vivax* malaria patients in China. A McDonald-Kreitman test detected positive selection on the 3′ region. Thus, we propose that *Pv*RON2 polymorphism has evolved under positive diversifying selection, possibly mediated by host immune pressure. A significant excess of d_N over d_S was also detected for *Pf*RON2, strengthening the case for RON2 as a potential target of the host immunity (Cao *et al*. 2009). However, because the Z-test used to evaluate the difference between d_N and d_S required a larger number of nucleotide differences than the detected one used for the data set presented here, the detected significant excess of d_N over d_S on the 5′ region of *Pf*RON2 might be a false positive, thus further studies are required for *Pf*RON2.

Three TM domains were predicted by different bio-informatics tools, TM3 was identified by 4 algorithms, but TM1 and TM2 were only supported by 2 algorithms. This difference may be due to the algorithm used in the webware; algorithms that basically rely on hydrophobicity tend to predict TM1 and TM2 in addition to T3, but algorithms that put more weight on amino acid characteristics,

https://doi.org/10.1017/S0031182011002447 Published online by Cambridge University Press

Table 5. The McDonald-Kreitman test of selection for *PvRON2 gene*

Region	*n*	Number of sites(base)	Fixed difference between *P. vivax* and *P. knowlesi*		Polymorphic sites within *P. vivax*		*P*
			Syn	Nsyn	Syn	Nsyn	
5′ region	20	1893	158	187	0	2	n.s.
3′ region	20	3942	314	158	1	5	$P < 0{\cdot}02$

n, number of *Plasmodium vivax* sequences analysed.
nt, nucleotide position.
Sites, nucleotide sites analysed. Nucleotide positions 4–2667 and 2668–6609 of *Pv*RON2 were used for the 5′ and 3′ regions, respectively after excluding sequence around indels and repeat region to obtain a reliable alignment. Nucleotide sequences used for the analysis were nucleotide positions 4–663, 676–771, 1027–1065, 1072–1206, 1360–1482, 1489–1623, 1954–2472, and 2482–6609 for *Pv*RON2 (after Sal-I reference line) and nucleotide positions 4–663, 679–774, 997–1035, 1045–1179, 1303–1425, 1435–1569, 1768–2286, and 2302–6426 for *P. knowlesi* RON2.
Syn and Nsyn indicate synonymous substitution and non-synonymous substitution, respectively. Fisher's exact test (one-tailed) was used.
P value indicates that Nsyn is significantly greater than Syn.
n.s. indicates not significant ($P > 0{\cdot}05$).

such as TMHMM, HMMTOP and SOSUI, did not detect TM1 and TM2. As this protein is secreted from the rhoptry neck and inserted into the host cell membrane, the predicted TM domains are probably not anchored in the parasite membrane prior to discharge, thus there would be a mechanism to prevent the exposure of these putative TM domains to the environment. They could, for example, be buried in the globular domain before discharge, but be exposed during host cell invasion. Such constraints may make TM1 and TM2 difficult to predict. Performing similar analyses and localization studies of RON2 in *T. gondii*, the following membrane topology of RON2 was proposed; the N-terminal segment before TM2 is located on the cytosolic side, the region between TM2 and TM3 or the segment after TM2 is exposed outside of the cell, and TM1 was actually not a true TM domain but peripherally associated with the host cell membrane by a hydrophobic interaction (Lamarque *et al*. 2011). Based on this model, all polymorphic sites of *Pv*RON2 detected in this study locate in the region that is not exposed to the outside of the erythrocyte. It is not clear how the host immune system may recognize these sites during the asexual stage cycle. However, one possible explanation is that *Pv*RON2 is released into the hepatocyte during sporozoite invasion where it may be targeted by host cellular immunity. RON2 expression in the sporozoite has previously been reported in the rodent malaria parasite *Plasmodium berghei* (Tufet-Bayona *et al*. 2009).

The RON2-AMA1 interaction was found to be strong and species specific (Lamarque *et al*. 2011). No significant binding of a homologous region from the rodent malaria parasite *P. yoelii* RON2 with *Pf*AMA1 was observed, suggesting evolutionary constraints within *Plasmodium* species for the RON2-AMA1 interaction (Srinivasan *et al*. 2011). The region between TM2 and TM3 was proposed to be located on the erythrocyte surface and was shown to bind to AMA1 in both *T. gondii* and *P. falciparum* (Tyler and Boothroyd, 2011; Lamarque *et al*. 2011). Recently, hot spot residues in the binding of *Tg*AMA1-*Tg*RON2 were identified (Tonkin *et al*. 2011), which corresponds to aa 2035 to 2073 of *Pv*RON2. There were no amino acid substitutions in this region among 20 isolates, supporting the conserved and potentially crucial role of this interaction.

ACKNOWLEDGEMENT

We would like to thank Key Lab. on Technology for Parasitic Disease Prevention and Control, Ministry of Health, Jiangsu Provincial and Key Lab. on Molecular Biology of Parasites for providing the necessary equipment and assistance.

FINANCIAL SUPPORT

This work was partly supported by the National Basic Research Program (973 Program) in China (2007CB513100), and the National Natural Science Foundation of China (No. 30700695, and 30872214).

REFERENCES

Alexander, D. L., Arastu-Kapur, S., Dubremetz, J. F. and Boothroyd, J. C. (2006). *Plasmodium falciparum* AMA1 binds a rhoptry neck protein homologous to *Tg*RON4, a component of the moving junction in *Toxoplasma gondii*. *Eukaryotic Cell* **5**, 1169–1173. doi:10.1128/EC.00040–06.

Alexander, D. L., Mital, J., Ward, G. E., Bradley, P. and Boothroyd, J. C. (2005). Identification of the moving junction complex of *Toxoplasma gondii*: a collaboration between distinct secretory organelles. *PLoS Pathogens* **1**, e17.

Arévalo-Pinzón, G., Curtidor, H., Patiño, L. C. and Patarroyo, M. A. (2011). PvRON2, a new *Plasmodium vivax* rhoptry neck antigen. *Malaria Journal* **10**, 60.

Aurrecoechea, C., Brestelli, J., Brunk, B. P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W.,

Miller, J. A., Nayak, V., Pennington, C., Pinney, d.f., Roos, D. S., Ross, C., Stoeckert, C. J., Jr, Treatman, C. and Wang, H. (2009). PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research* **37**, D539–543.

Baum, J., Tonkin, C. J., Paul, A. S., Rug, M., Smith, B. J., Gould, S. B., Richard, D., Pollard, T. D. and Cowman, A. F. (2008). A malaria parasite forming regulates actin polymerization and localizes to the parasite-erythrocyte moving junction during invasion. *Cell Host and Microbe* **3**, 188–198.

Boothroyd, J. C. and Dubremetz, J. F. (2008). Kiss and split: the dual roles of *Toxoplasma* rhoptries. *Nature Reviews Microbiology* **6**, 79–88.

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

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. and Fraser-Liggett, C. M. (2008). Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature, London* **455**, 757–763.

Curtidor, H., Patiño, L. C., Arévalo-Pinzón, G., Patarroyo, M. E. and Patarroyo, M. A. (2011). Identification of the *Plasmodium falciparum* rhoptry neck protein 5 (PfRON5). *Gene* **474**, 22–8.

Greenwood, B. M., Fidock, D. A., Kyle, D. E., Kappe, S. H., Alonso, P. L., Collins, F. H. and Duffy, P. E. (2008). Malaria: progress, perils, and prospects for eradication. *Journal of Clinical Investigation* **118**, 1266–1276. doi:10.1172/JCI33996.

von Heijne, G. (1992). Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *Journal of Molecular Biology* **225**, 487–94.

Hirokawa, T., Boon-Chieng, S. and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14**, 378–379.

Hofmann, K. and Stoffel, W. (1993). TMbase – A database of membrane spanning proteins segments. *Biological Chemistry*. Hoppe-Seyler **374**, 166.

Hughes, A. L. and Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class loci reveals overdominant selection. *Nature, London* **335**, 167–170.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* **305**, 567–580.

Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon-Guyot, J., Dubremetz, J. F., Fauquenoy, S., Tomavo, S., Faber, B. W., Kocken, C. H., Thomas, A. W., Boulanger, M. J., Bentley, G. A. and Lebrun, M. (2011). The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathogens* **7**, e1001276. doi:10.1371/journal.ppat.1001276.

Lebrun, M., Michelin, A., El Hajj, H., Poncet, J., Bradley, P. J., Vial, H. and Dubremetz, J. F. (2005). The rhoptry neck protein RON4 re-localizes at the moving junction during *Toxoplasma gondii* invasion. *Cellular Microbiology* **7**, 1823–1833.

McDonald, J. H. and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* **351**, 652–4.

Nei, M. and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**, 418–426.

Newman, A. M. and Cooper, J. B. (2007). XSTREAM: a practical algorithm for identification and architecture modeling of tandem repeats in protein sequences. *BMC Bioinformatics* **8**, 382. doi:10.1186/1471-2105-8-382.

Pain, A., Böhme, 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. and Berriman, M. (2008). The genome of the simian and human malaria parasite *Plasmodium knowlesi*. *Nature, London* **455**, 799–803.

Polley, S. D. and Conway, D. J. (2001). Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene. *Genetics* **158**, 1505–1512.

Price, R. N., Tjitr, E., Guerra, C. A., Yeung, S., White, N. J. and Anstey, N. M. (2007). Vivax malaria: neglected and not benign. *American Journal of Tropical Medicine and Hygiene* **77**, 79–87.

Richard, D., MacRaild, C. A., Riglar, D. T., Chan, J. A., Foley, M., Baum, J., Ralph, S. A., Norton, R. S. and Cowman, A. F. (2010). Interaction between *Plasmodium falciparum* apical membrane antigen 1 and the rhoptry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites. *Journal of Biological Chemistry* **285**, 14815–22.

Srinivasan, P., Beatty, W. L., Diouf, A., Herrera, R., Ambroggio, X., Moch, J. K., Tyler, J. S., Narum, D. L., Pierce, S. K., Boothroyd, J. C., Haynes, J. D. and Miller, L. H. (2011). Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proceedings of the National Academy of Sciences, USA* [Epub ahead of print].

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

Tonkin, M. L., Roques, M., Lamarque, M. H., Pugniére, M., Douguet, D., Crawford, J., Lebrun, M. and Boulanger, M. J. (2011). Host cell invasion by Apicomplexan parasites: Insights from the co-structure of AMA1 with RON2 peptide. *Science* **333**, 463–467.

Tufet-Bayona, M., Janse, C. J., Khan, S. M., Waters, A. P., Sinden, R. E. and Franke-Fayard, B. (2009). Localisation and timing of expression of putative Plasmodium berghei rhoptry proteins in merozoites and sporozoites. *Molecular and Biochemical Parasitology* **166**, 22–31.

Tusnády, G. E. and Simon, I. (2001). The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849–850.

Tyler, J. S. and Boothroyd, J. C. (2011). The C-Terminus of Toxoplasma RON2 provides the crucial link between AMA1 and the host-associated invasion complex. *PLoS Pathogens* 7:e1001282. doi:10.1371/journal.ppat.1001282.