Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular & Biochemical Parasitology 138 (2004) 37–49
MOLECULAR & BIOCHEMICAL PARASITOLOGY

A *SICAvar* switching event in *Plasmodium knowlesi* is associated with the DNA rearrangement of conserved 3′ non-coding sequences☆

Vladimir Corredor[a,b], Esmeralda V.S. Meyer[a], Stacey Lapp[a], Claudia Corredor-Medina[a], Curtis S. Huber[c], Andrew G. Evans[a], John W. Barnwell[c], Mary R. Galinski[a,d,*]

[a] *Emory Vaccine Center, Yerkes National Primate Research Center, Emory University, 954 Gatewood Road, Atlanta, GA 30329, USA*
[b] *Departamento de Ciencias Fisiologicas, Facultad de Medicina, Universidad Nacional de Colombia, Bogota, Columbia 30341, USA*
[c] *Division of Parasitic Diseases, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Chamblee, GA 30341, USA*
[d] *Department of Medicine, Division of Infectious Diseases, Emory University, 954 Gatewood Road, Atlanta, GA 30329, USA*

Received 27 October 2003; received in revised form 23 April 2004; accepted 9 May 2004

Abstract

Plasmodium knowlesi variant antigens are expressed at the surface of infected erythrocytes and are encoded by the Schizont Infected Cell Agglutination variant antigen (*SICAvar*) multigene family. The 3′ region of the *SICAvar* gene locus encoding the 205 kDa variant antigen expressed in the Pk1(B+)1+ parasites was found to be altered compared to the Pk1(A+) parental clone. Here we report that this alteration is the result of a DNA rearrangement and that the original and altered *205 SICAvar* alleles appear to encode bona fide variant antigens. Importantly, *205A* and *205B SICAvar* RNA sequences are detectable in similar apparent quantities as determined by quantitative real-time reverse-transcription polymerase chain reaction (qRT-PCR) amplification experiments. However, expression of the 205 kDa SICA protein at the surface of the infected erythrocyte is not characteristic of the Pk1(A+) parasites and the *205 SICAvar* transcript has not been detected in Pk1(A+) parasites by northern blot analysis. Furthermore, we report that many distinct *SICAvar* transcripts were detected in *P. knowlesi* Pk1(B+)1+ cDNA library hybridization screens. Of special interest, in light of these data, distinctive differences at the 3′ end of the *205A* and *205B* alleles are observed, which may be of functional importance. An analysis of the 3′ untranslated region (UTR) of *SICAvar* genes in more than 100 sequences revealed a surprising common sequence pattern characterized by blocks of imperfect, GT-rich, heptad repeated motifs (Block I), followed by A and T rich homopolymers (Block II) and in a large number of genes, GC-rich segments (Block III). We show that this region undergoes extensive recombination and that the preferential stability of the *205 SICAvar* transcript in Pk1(B+)1+ parasites may be associated with the presence of its specific Block III sequences. We speculate that the conserved yet polymorphic *SICAvar* 3′UTR sequences, and comparable regions in *P. falciparum var* genes, function in the stage-specific and developmentally regulated post-transcriptional gene silencing (PTGS) of variant antigen transcripts.

Keywords: *Plasmodium*; Malaria; Antigenic variation; *var* genes; *SICAvar* genes; Erythrocytes

1. Introduction

Antigenic variation in malaria and the presence of variant antigens at the surface of infected erythrocytes were originally studied in the simian malaria parasite species, *Plasmodium knowlesi* [1–6] and subsequently in the human malaria, *P. falciparum* [7,8]. Discoveries ensued in the mid 1990s leading to the identification of variant antigen gene families encoding the *P. falciparum* erythrocyte membrane protein I (PfEMP1) [9–11] and the related *P.*

Abbreviations: SICA, schizont-infected cell agglutination; UTR, untranslated region; TM, transmembrane; CRDs, cysteine rich domains; PFGE, pulsed-field gel electrophoresis; gDNA, genomic DNA; nt, nucleotide; aa, amino acid; ORFs, open reading frames; RT-PCR, reverse transcription-polymerase chain reaction; qRT-PCR, quantitative reverse transcription-polymerase chain reaction

☆ *Note:* Nucleotide sequence data reported in this paper are available in the Genbank™ database under accession numbers AY642228–AY642266.

* Corresponding author. Tel.: +1 404 727 7214; fax: +1 404 727 8199.
E-mail address: galinski@rmy.emory.edu (M.R. Galinski).

0166-6851/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2004.05.017

knowlesi schizont infected cell agglutination (SICA) antigens [12]. These gene families, named *var* and *SICAvar*, respectively, contain 50–100 members and encode proteins between ∼180 and 400 kDa that are comprised of a predominant variable extracellular region, a transmembrane (TM) domain, and a cytoplasmic tail. The large extracellular regions contain a series of distinct yet related cysteine rich domains (CRDs).

How malaria parasites control the expression of these antigens is important to understand since they are associated with virulence. In *P. falciparum*, this is partly due to the adhesive nature of its CRDs, which cause infected red blood cells to adhere to receptors on endothelial and other cells (reviewed in [13]). *P. knowlesi* infected red blood cells, in contrast, have evolved with less robust adhesive properties and unique sequestration characteristics [14]. Yet, *P. knowlesi* SICA[+] parasites are associated with extreme virulence in rhesus monkeys [4].

To date, the molecular mechanisms used by *Plasmodium* to govern the expression of the variant antigens have been elusive. *P. falciparum var* genes are distributed on all 14 chromosomes and transcription of multiple loci can take place in situ at both internal and telomeric chromosomal sites [15–20]. The mechanism(s) whereby a specific transcript among many becomes expressed as protein at the surface of infected cells is at present unknown. A mutually exclusive transcriptional control mechanism has been put forth to explain the detection of many *var* gene transcripts by reverse-transcription polymerase chain reaction (RT-PCR) amplification in early ring stages; yet only one in trophozoites [19]. Nuclear run-on experiments, showing the specific hybridization of one of five *var* gene Duffy binding-like-α (DBL-α) domains to transcripts generated from trophozoites selected for adhesion to CD36 or CSA, garnered support for this theory [19]. Nevertheless, *var* gene transcripts representing multiple *var* genes have been detected in adhesion-selected parasite populations as well as from micromanipulated single-cell ring and trophozoite-stage parasites ([18,19,21–23], and reviewed in [24]). As in *P. knowlesi* [12], northern blot analyses of *P. falciparum var* genes [25,26] have shown that *var* gene transcripts typically accumulate in the ring-stage parasites and become diminished in trophozoites, when the protein is expressed. Transcriptional control mechanisms could possibly explain the presence of predominant transcripts observed in some experiments, whether detected by northern blot analyses or RT-PCR (e.g. see [21,25,26–28]); however, the basis for such mechanisms as the sole factor regulating the expression of *var* genes remains preliminary. Post-transcriptional gene silencing (PTGS) mechanisms could also be considered to explain much of the currently available data [29].

A unique attribute of *P. knowlesi* compared to *P. falciparum* experimental systems is that stable isogenic clonal lines expressing distinct variant antigens can be derived in vivo, where specific host immune responses and the presence of the spleen are required for switching events ([4–6] and unpublished data). We originally set out to determine what genetic changes may have occurred when the Pk1(A+) clonal population, expressing SICA proteins of 210 and 190 kDa switched its phenotype in vivo leading to the generation of the Pk1(B+)1+ clonal population that alternatively expresses SICA proteins of 205 and 200 kDa [6]. We have shown that the *205 SICAvar* gene expressing the 205 kDa SICA protein has a 10-exon structure and that a 3′ alteration of this locus is associated with this switched phenotype [12]. Here we show that this is due to a rearrangement beginning in intron IX and includes ∼4 kb of coding and non-coding sequences; i.e. including regions comparable to the single intron and exon II characteristic of *P. falciparum var* genes ([9–11], and reviewed in [29]). Quantitative investigations of *SICAvar* transcripts representing the original and altered *SICAvar* alleles, and comparative studies of 108 sequences related to the 3′ regions of *SICAvar* genes, suggest that *SICAvar* 3′UTR sequences may be functionally important for post-transcriptional regulation of this variant antigen gene family.

2. Materials and methods

2.1. Parasites, total RNA and genomic DNA

P. knowlesi infected red blood cells were propagated in rhesus monkeys at the Yerkes National Primate Research Center, Emory University, and processed to obtain stage-specific parasites essentially as described [4,12]. Blood was drawn when parasitemias reached at least 5%, at approximately 3 p.m., when it is routinely observed that most if not all schizonts have ruptured and the newly released merozoites have reinvaded new host red blood cells. Purified ring-stage infected erythrocytes are generally processed immediately for the purification of ring-stage RNA or cultured first for a 2 h period. Middle-stage trophozoite RNA is prepared from the development of these parasites after ∼13 h in culture. Stage-specific RNA stocks generated by these procedures, starting with highly synchronous infected red blood cells from an infected animal, reliably represent the distinct specific stages, while reconstituted cryopreserved ring-infected red blood cells generally do not mature as well in a highly synchronous fashion. Total RNA was purified from each infected red blood cell population using Trizol (Gibco) reagents and standard procedures recommended by the manufacturer. Genomic (g) DNA was prepared as described [20] from parasites matured in culture to the schizont stage.

2.2. Pulsed-field gel electrophoresis (PFGE)

Chromosome blocks containing schizont infected red blood cells were prepared for pulsed-field gel electrophoresis (PFGE) experiments as described [30] and fractionated on a BioRad CHEF-DRIII system (120 s/24 h; 240 s/36 h pulses in 1× TAE at 14 °C using 0.8% PFG-grade agarose at 4.5 V/cm

and a 106° angle) and processed for Southern blots using standard procedures.

2.3. Pk1(B+)1+ cDNA libraries, hybridization screening and Southern blots

Pk1(B+)1+ cDNA expression libraries constructed with either random hexamer or oligo dT-primed middle trophozoite to early schizont-stage RNA in Uni-ZAP, or Lambda ZAPII and Lambda ZAP Express vectors (Stratagene) described previously [12] were screened by hybridization with radiolabelled *SICAvar* probes #24 and 538-6 [12], representing the *205 SICAvar* exons VIII and IX sequences, under stringent conditions essentially as described [31]. Approximately 650,000 phage plaques were screened, and the DNA inserts of 22 positive *SICAvar* clones were characterized and partially sequenced to verify their basic characteristics.

Southern blot procedures were performed as described [31], with hybridizations at 65 °C and washes in 6× SSC/0.5% SDS, 2× SSC/0.5% SDS, and 0.2× SSC/0.1% SDS. Probes were: (1) *205* locus-specific: oligomer GSP3/7, (2) 150 bp exon V PCR product (SV19 × SV20); (3) ~500 bp 5′UTR PCR product (SV15 × SV16); (4) exon X: 385 bp exon X PCR product (SV23 × cyto5). All PCR probes were generated from Pk1(B+)1+ gDNA templates.

2.4. Oligomers

Oligonucleotides [PkTM1-2 (exon IX), GSP3 and GSP7 complementary sequences (exon IX), cyto5 (exon X), SV5 (*205B* 3′ non-coding), SV6 (*205A* 3′ non-coding), SV15 (5′UTR), SV16 (5′UTR), SV19 (exon V), SV20 (exon V), SV23 (exon X), SV26 (*205B* 3′ non-coding), exon VIII qRT-PCR forward (exon VIII) and exon VIII qRT-PCR reverse (exon IX)] (see Supporting Data) were synthesized at the Emory University Microchemical and Proteomics Facility.

2.5. Reverse transcription-polymerase chain reactions, quantitative real-time reverse transcription-polymerase chain reactions (qRT-PCR), and cloning

Reverse transcription-polymerase chain reactions were performed starting with 1 ug of total RNA, Thermoscript Reverse Transcriptase (Invitrogen), and oligo dT, SV6 or SV5 primers (Fig. 1). The RNAs were quantified by multiple spectrophotometric readings and shown by gel analysis to be of high quality. The PCR Master kit (Roche) and PCR purification columns (Qiagen) were utilized, as well as the High Fidelity System (Roche) to amplify products for cloning into the pCR 2.1 TOPO system (Invitrogen). The PCR conditions used for the GSP3 + cyto5 amplifications were as follows: 94 °C 2 min, 80 °C 2 min, followed by 35 cycles of 94 °C 1 min, 57 °C 1 min, and 72 °C 2 min, and a final extension of 72 °C 7 min. Each RT-PCR amplification was performed in 50 and 10 μl (20%) of the resulting amplified material was loaded per lane and electrophoresed in agarose gels to visualize the products.

Quantitative real-time reverse transcription-polymerase chain reactions were initiated with 2 ug total RNA from Pk1(A+) and Pk1(B+)1+ parasites at both ring and trophozoite stages. cDNA was generated with TaqMan® Gold RT-PCR Kit (Applied Biosystems) priming with both oligo dT and the Block III oligomers SV5-205B or SV6-205A. cDNA products were amplified in duplicate with 0.9 μM of each of the exon VIII–IX primers noted below and analyzed in real-time with an iCycler (Bio-Rad), using the Quantitect™ SYBR® Green PCR kit (Qiagen). The PCR conditions were as follows: 95 °C 15 min, followed by 40 cycles of 95 °C 30 s, 59.2 °C 30 s, and 72 °C 1 min with a final extension of 72 °C 7 min.

Exon VIII–IX primer pairs were designed using Primer Express software (Applied Biosystems, Version 1.0) to amplify a segment of the *205 SICAvar* transcript across the exon VIII–IX junction, producing an amplicon of

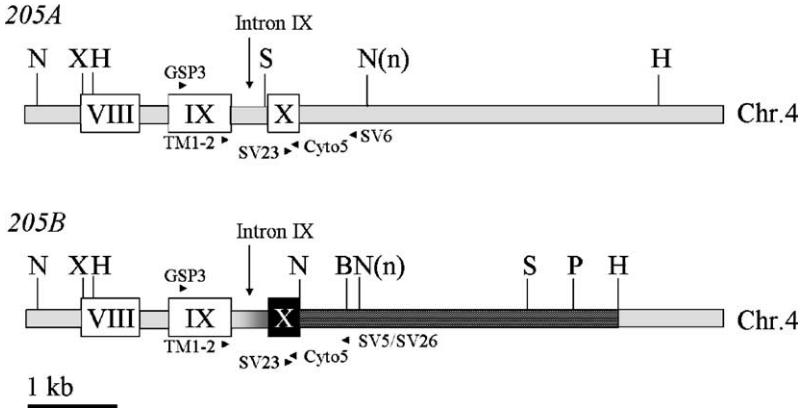

Fig. 1. Schematic highlighting the original (*205A*) and rearranged (*205B*) 3′ portions of the *205 SICAvar* alleles. Conserved and polymorphic restriction sites (N: *Nci*I, X: *Xba*I, H: *Hin*dIII, S: *Ssp*I, B: *Bss*HII, P: *Pvu*II) confirmed by Southern blot analyses are depicted ([12], and data not shown). Oligomers used as hybridization probes or in PCR amplifications are noted by directional arrows. The intron IX, exon X and downstream regions are depicted with different shading patterns to emphasize confirmed differences.

103 bp (exon VIII qRT-PCR forward, 5′-ATGGAATACCATCCTGGGATAC-3′ and exon IX qRT-PCR reverse, 5′-GTTCAACTTCTTTCTGCTTATTTCCC-3′). The reverse primer corresponds to the GSP7 oligomer used in the traditional RT-PCRs and shown previously to be specific for the *205 SICAvar* locus [12]. The forward primer was shown by BLAST analyses of the *P. knowlesi* genome database to be specific for this *SICAvar* locus and only one other *SICAvar* gene. Plasmid controls for real-time RT-PCR experiments were generated by cloning the products of traditional RT-PCR amplifications from Pk1(A+) and Pk1(B+)1+ parasite RNA samples into the pCR®II-TOPO® vector (Invitrogen) and transforming into TOP10 cells. Positive clones were screened and sequenced to confirm fidelity and orientation of the inserts. The plasmid DNA was linearized and in vitro transcribed to RNA (RiboMax Large Scale RNA Production System-SP6, Promega). A standard curve was established from 1×10^9 to 1×10^2 copies.

2.6. DNA sequencing, BLAST and sequence analyses

The BigDye Terminator Cycle Sequencing v2.0 Ready Reaction kit and a 3100 Genetic Analyzer (ABI Prism) were used for all sequencing reactions. The sequence of *SICAvar* cDNAs detected by hybridization of gene expression libraries were also sequenced by standard manual sequencing methods as described [12] and analyzed using MacVector programs (Oxford Molecular). *205A* and *205B* exon X and 3′ non-coding *SICAvar* sequences were queried against Pk1(A+) contigs produced by the Sanger Institute and present at the *Plasmodium* Genome Resource (http://plasmodb.org/PlasmoDB.shtml) [32] as of 3 October 2002. Default parameters used were as described in WashU-BLASTN 2.0. Sequences were aligned using CLUSTAL X [33] and the Bioedit Sequence Alignment Editor (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). The Tandem Repeats Finder algorithm [34] was used to identify repeats using various alignment parameters, minimum score and maximum period sizes. Gene trees were built using the Neighbor-Joining algorithm [35] from distances computed using the proportional (p) number of differences between pairwise sequence comparisons. The minimum number of recombination events (Rm) [36] was estimated using DnaSP 3.53 [37].

All sequence alignments referred to in the text are presented as Supporting Data on-line [site with PDF file to be referenced].

3. Results

3.1. The 205 SICAvar DNA rearrangement involves intron IX and conserved downstream sequences

Previous Southern blot analyses revealed a DNA alteration in the 3′ portion of the *205 SICAvar* gene, which was associated with the switch in expression of the 210/190 kDa SICA phenotype in Pk1(A+) parasites to the 205/200 kDa SICA phenotype in Pk1(B+)1+ parasites [12]. To further investigate the altered region, a 7 kb gDNA clone derived from Pk1(B+)1+ parasites and containing exon VIII through exon X and ∼5 kb of downstream sequences [12] was physically mapped and characterized. An ∼4.0 kb *Nci*I gDNA fragment derived from Pk1(A+) parasites and related PCR products corresponding to exons VIII–X and the immediate downstream region were also cloned and characterized. An additional 300 bp of Pk1(A+) sequence was also identified in the *P. knowlesi* genome sequence database under development (http://plasmodb.org/PlasmoDB.shtml), which is based on gDNA from the Pk1(A+) parasite clone [6]. Fig. 1 summarizes the prominent differences identified in the Pk1(A+) and Pk1(B+)1+ parasites within the *205* locus (i.e. the *205A* and *205B* alleles) and downstream sequences.

An alignment comprising exon VIII through exon X and ∼0.6 kb of the 3′UTR sequences (Alignment #1 (Aln-1-nt); see Supporting Data) shows complete identity until the second half of intron IX, which has 5 indels and 46 nucleotide (nt) differences out of 414 bp. Exon X displays 30 substitutions, and all but two of the resulting amino acid (aa) differences encoded are conservative changes (Alignment #2 (Aln-2-aa); see Supporting Data). Both sequences encode open reading frames (ORFs) and appear to be *bonafide SICAvar* alleles. However, the 205 kDa SICA protein is only detected in the Pk1(B+)1+ parasites [4,6,12]. As further detailed below, the 3′ non-coding sequences are conserved yet polymorphic. GT-rich imperfect heptad repeats are present immediately adjacent to the stop codon; the *205A* allele contains 28 copies (375 bp) of the GGGTTCACA(G/T)TTTA consensus motif and the *205B* allele has eight copies (118 bp) of the GGGTT(T/C)A sequence. Interestingly, both sequences include the GGGTT(T/C)A telomeric motif described for *Plasmodium* (see [38,39]).

3.2. The 205A SICAvar 3′ sequences are deleted or mutated in the Pk1(B+)1+ genome

To further investigate the nature of the rearrangement involving the *205A* and *205B SICAvar* 3′ coding and non-coding regions, and attempt to identify the donor and recipient sequences involved in the event that resulted in the generation of the *205B* allele, several probes were hybridized to a series of gDNA restriction enzyme digests including *Nci*I digestions of *P. knowlesi* gDNAs as shown in Fig. 2A. As demonstrated previously [12], the exon IX locus-specific oligomer, GSP3/7 hybridizes as a single-copy sequence and is diagnostic of an alteration in this region. In contrast, exon X and 3′UTR oligomers (SV6-205A and SV26-205B), derived from the *205A* and *205B* alleles, respectively, hybridize to multiple bands (the conserved exon X sequence detects a large number of *SICAvar* gene family members and the 3′UTR oligomers detect a more limited number, with each oligomer detecting a different subset of the related 3′UTR

sequences). Importantly, these data show that the 4.0 kb *Nci*I fragment that hybridizes with the SV6-205A oligomer and corresponds to the *205A* allele (see Fig. 1) is not present in the Pk1(B+)1+ gDNA digests, and no other apparent alternative band is detected. This suggests that the 3′ *205A SICAvar* sequence that was replaced in the rearrangement that generated the *205B SICAvar* allele was either deleted, as we speculate, or alternatively was mutated at the location corresponding to the SV6-205A sequence. To the contrary, the SV26-205B oligomer hybridizes to the same subset of bands in all three DNA samples, as expected since the *Nci*I sites at the end of exon X of the *205B* allele and 1.5 kb downstream (see Fig. 1) would be similarly positioned before and after this sequence was transferred to the *205A* locus. These experiments also show the multicopy, semi-conserved nature of sequences in the 3′UTR region. In contrast, the 5′UTR of this *SICAvar* locus hybridizes as a single copy (Fig. 2B).

Given the multicopy nature of the exon X and 3′UTR sequences, and in consequence the inability to generate allele-specific probes, the identification of the *205B* 3′ donor sequences in Pk1(A+) parasites has so far been elusive. However, *205 SICAvar* locus-specific sequences (5′UTR 500 bp sequence, GSP7/GSP3 oligomers, and a 150 bp exon V probe) have enabled the initial chromosomal mapping of this locus to chromosome IV. In contrast, a conserved exon

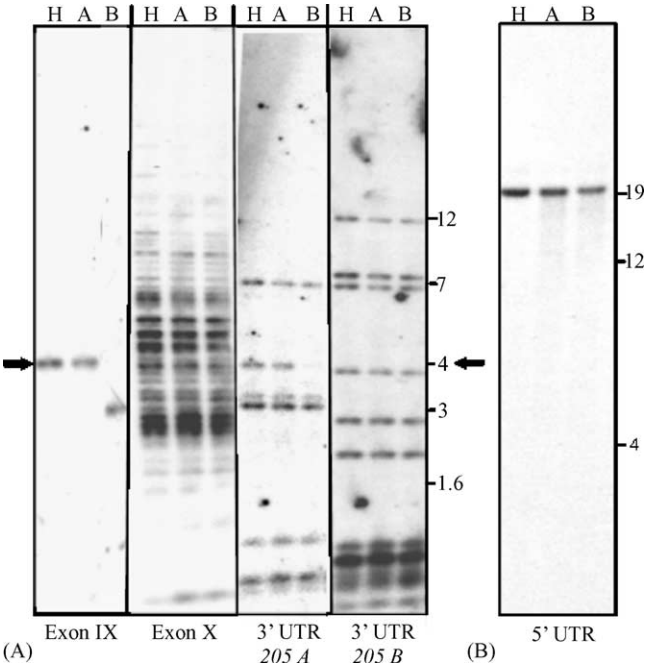

Fig. 2. Southern blot hybridizations of *P. knowlesi* H strain (H), Pk1(A+) (A) and Pk1(B+)1+ (B) gDNAs digested with *Nci*I (2A) or *Hin*dIII (2B) restriction enzyme. (A) shows the hybridization of the same membrane with: exon IX locus-specific oligomer GSP3, conserved exon X probe (385bp), and 3′UTR oligomers SV6-205A and SV26-205B. Arrows indicate the position in the third panel (3′UTR-205A) where the 4.0 kb fragment is present in the H and A gDNAs but missing in the B gDNA. (B) shows the hybridization pattern of a 5′UTR locus-specific probe (∼500 bp) to H, A and B gDNA digests.

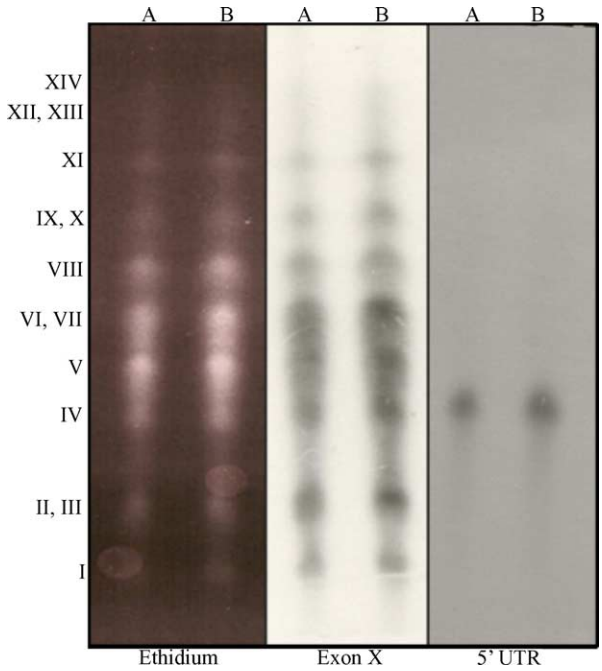

Fig. 3. The *205 SICAvar* locus maps to chromosome IV. An ethidium bromide stained PFG chromosomal fractionation and hybridization with exon X (385 bp) and 5′UTR locus-specific (∼500 bp) probes are shown. The positions of Pk1(A+) and Pk1(B+)1+ chromosomes (A and B, respectively) are numbered.

X probe hybridizes to all fourteen chromosomes, which are apparent using differing PFGE conditions (Fig. 3 and data not shown).

Together, these results indicate that upon switching from the Pk1(A+) to Pk1(B+)1+ SICA phenotype, and expression of the 205 kDa SICA protein in the Pk1(B+)1+ parasites, a new allele was generated at the *205 SICAvar* locus. This involved the transfer of ∼4.0 kb of sequences from an as yet unidentified chromosomal location to the 3′ portion of the *205* locus on chromosome IV, and the apparent accompanying loss of the original 3′ sequence present at this location.

3.3. *SICAvar 3′UTR sequences are conserved and reveal distinctive patterns*

To better understand the apparent conserved structures of the *SICAvar* 3′UTR and neighboring downstream sequences we carried out comparative analyses of over 100 related sequences derived by PCR and RT-PCR in our laboratory and a comparable number identified in the *P. knowlesi* genome database (http://plasmodb.org/PlasmoDB.shtml). BLAST searches, using *205B* exon X as the query sequence, yielded 108 contigs with conserved sequences (*E*-values ranging from 1.9e−85 to 0.45). A series of alignments were generated that show the distinctive characteristics of the conserved exon X and 3′ non-coding sequences (see Supporting Data). Alignment 3-89 (Aln-3-89) displays 89 distinct sequences containing exon X, and Alignment 4-72 (Aln-4-72)

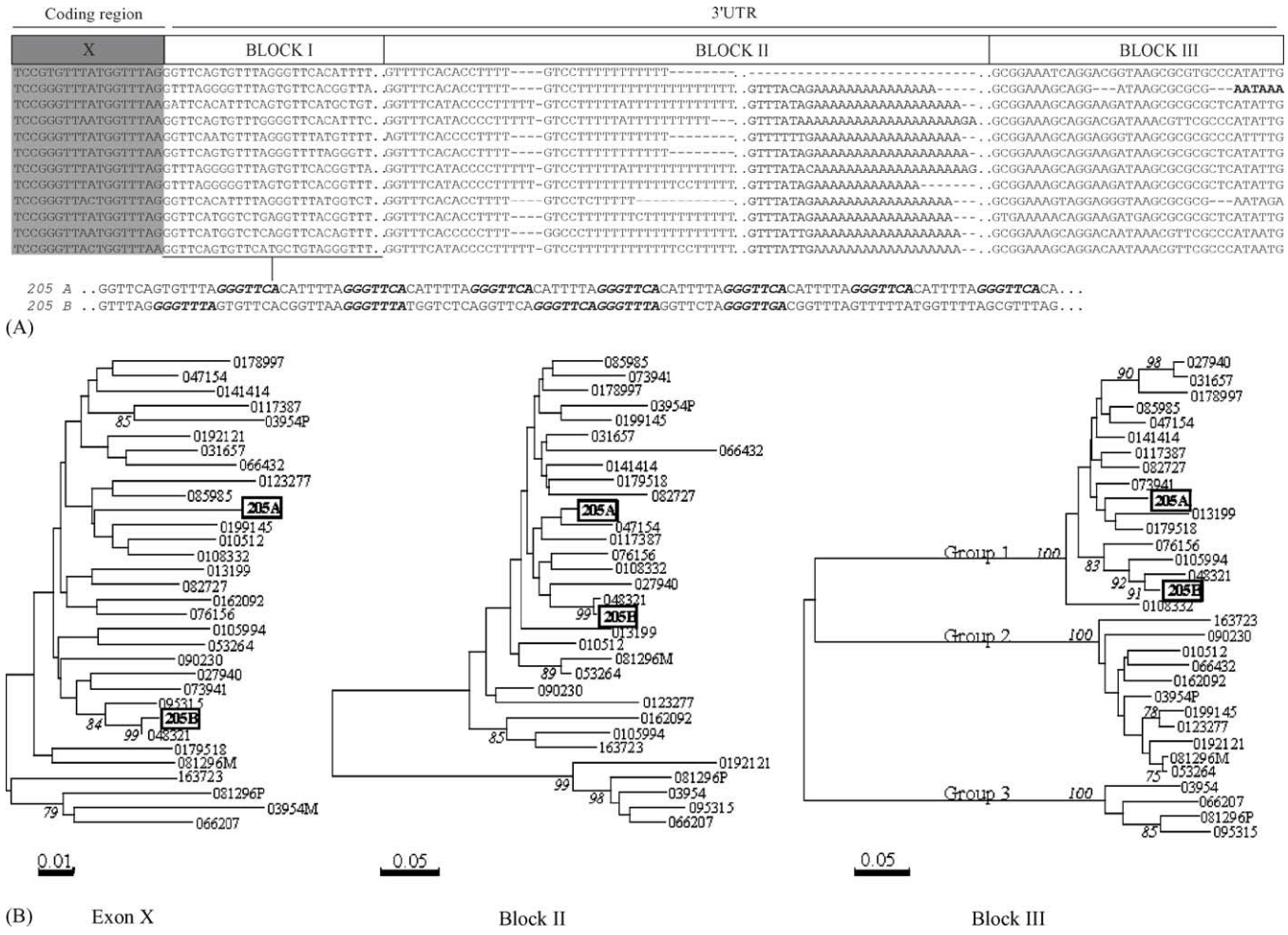

Fig. 4. (A) Schematic representation of the conserved *SICAvar* exon X coding and distinctive semi-conserved Blocks I–III regions that comprise the 3′UTRs. The schematic shows limited segments of these regions from 12 representative sequences identified in BLAST searches. The upper two sequences represent the *205A* and *205B SICAvar* sequences, respectively. The Block I repeated motifs are expanded to show their patterns and the presence of the *Plasmodium* telomeric motif; which is bolded. The putative AATAAA polyadenylation motif unique to the *205B* Block III sequence is shown. To view the complete alignments, see Supporting Data. (B) Neighbor-Joining cladograms displaying the different clustering of Exon X, Block II and Block III sequences. The 32 contigs used in the analysis are noted by contig numbers from the *P. knowlesi* database (as of 3 October 2002) (http://plasmodb.org/PlasmoDB.shtml). The *205A* and *205B* sequences are boxed. Bootstrap values >75%, shown at the bifurcation nodes, correspond to 1000 bootstrap replicas.

highlights the immediate 3′UTR sequences available from 72 contigs. This region, which we have called Block I, ranges in size from 44 to 375 bp and is comprised of imperfect GT-rich heptad repeats, similar to those noted above in the *205A* and *205B* alleles. Alignment 5-66 (Aln-5-66) shows exon X contigs with >500 bp of 3′ non-coding sequences, and reveals a second distinct region (Block II) that varies in size from ∼180 to 1300 bp, with multiple A and T homopolymer tracts and numerous duplications and indels, some of which are common to certain groups of sequences, which suggests that extensive recombination has taken place. The consensus GGTTTCACACC(T)n^1GTCC(T)n^2 (n^1 = 4–8; n^2 = 21) can also be noted in the 5′ part of Block II. In this group of sequences (Aln-5-66), Block II is flanked by a region of ∼100 bp with interspersed GC-rich tracts, which we have designated Block III. Southern blot and BLAST analyses show that Block III sequences are divergent and present in multiple copies (see Aln-6-51). Fig. 4A shows a schematic that summarizes the fundamental findings relating to the conservation of exon X and the 3′UTR sequence patterns we have designated Blocks I–III. Of potential importance, the *205B* allele has a putative polyadenylation sequence (AATAAA) 454 bp downstream from its stop codon, which is not present in the *205A* allele. When comparing 57 related sequences in our analysis only two other sequences were found to have AATAAA at a similar position, which is at the beginning of the Block III region. These findings led us to similarly examine *P. falciparum* 3′UTR *var* gene alignments reported recently [40]. We found that despite an A/T content of ∼80%, the AATAAA motif is in fact also only rarely observed in these sequences. Downstream from Block III, the sequences become unique. However, groups of sequences

downstream of Block III can be discerned that display imperfect repeats present in high copy number (e.g. the consensus sequence CCGGGTTAAAAGTAAAATTGGT), which has, indeed, also made it difficult to readily map the sequences downstream of the 205 alleles as well as other *SICAvar* genes in this large multigene family (unpublished data).

3.4. Evidence of extensive recombination in the 3′ region of SICAvar genes

Thirty-two sequences containing exon X and 3′UTRs with well-defined Blocks II and III were aligned and gene trees built to assess the stability of linkage groups (Fig. 4B, and see Aln-7-32 in Supporting Data); due to its repeated nature, Block I was not included in the analysis. Block III defines three groups of sequences with bootstrap values of 100%: Group 1, 17 sequences; Group 2, 11 sequences; and Group 3, 4 sequences. In this gene tree, the *205A* and *205B* alleles both belong to Group 1. In contrast, when trees are built using exon X or Block II sequences, bootstrap values diminish and different clusters are obtained, revealing the effect of recombination. In certain cases sequences belonging to different groups in the Block III tree now cluster with high bootstrap values (e.g. in the exon X tree: the node grouping 095315, 048321 and the *205B* sequences; and in the Block II tree: the node grouping 0162092, 0105994 and 163723; or 0192121 and Group 3 sequences). Moreover, clusters with high bootstrap values within Groups 1 and 2 are broken in the exon X and Block II trees (e.g. the nodes including *205B*, 048321 and 0105994; 031657, 027940 and 0178997, as well as 0123277 and 0199145 sequences). However, Group 3 sequences tend to remain clustered. Together, these data suggest that recombination has occurred at different frequencies between and within groups. To better appreciate the effect of recombination within each group of sequences, we estimated the minimum number of recombination events (Rm) [36] for each of the three groups defined by Block III (Aln-8-G1, Aln-9-G2, Aln-10-G3). A minimum of 22 recombination sites for Group 1, 24 for Group 2, and 11 for Group 3 were determined. Given the topology of the trees recovered (i.e. Group 3 tends to associate with members of its own group), these data support the interpretation that recombination tends to occur at a higher frequency within members of the same group.

3.5. SICAvar transcripts are produced representing the 205A and 205B alleles

Experiments were conducted to establish if the *205A* and *205B* alleles are both transcribed and to begin to assess the functional relevance of the conserved nature of the *SICAvar* 3′UTR sequences. First, PCRs were performed with gDNA from Pk1(A+) and Pk1(B+)1+ parasites using the forward exon IX locus-specific primer, GSP3, or conserved exon IX transmembrane primer, TM1-2, paired with an oligo dT primer or the reverse 3′UTR oligomers SV6-205A and SV5-205B (Fig. 1). PCRs that combined GSP3 and the oligo dT primer amplified each allele. Sequencing reactions show that oligo dT priming occurs predominantly if not exclusively at Block II A homopolymeric tracts (data not shown). PCRs that combined GSP3 and *205A* and *205B* 3′UTR primers specifically amplified the corresponding regions of each allele only from the respective gDNAs. In contrast, when the conserved TM1-2 primer was combined with these 3′UTR primers products representing many different *SICAvar* genes were obtained (not shown).

Specific RT-PCR experiments were thus initiated using either oligo dT or the SV6-205A and SV5-205B primers to initiate the reverse transcription reaction to test for the presence of transcripts representing each allele (Fig. 5). Total RNA samples from Pk1(A+) and Pk1(B+)1+ ring- and trophozoite-stage parasites were used as templates. Subsequent PCRs were performed with GSP3 as the forward primer and cyto5, a conserved exon X oligomer, as the reverse primer. Fig. 5A shows that *205* locus-specific RT-PCR products of the expected size (∼800 bp) were generated from both the Pk1(A+) and Pk1(B+)1+ RNA templates, from ring and trophozoite-stage samples. These products were sequenced and shown to correspond to the 3′ regions of the respective alleles. These 3′ amplification data confirmed that both alleles are completely transcribed, regardless of the fact that the 205 kDa protein has only been detected in the Pk1(B+)1+ parasites harboring the *205B* allele. Additional RT-PCR experiments, amplifying upstream sequences, corroborate this finding (not shown) and confirmed that the *205A* and *205B* alleles have identical upstream coding sequences, thus limiting the differences so far detected to the 3′ region.

Of special interest, Fig. 5A shows that RT-PCR products of similar intensities were consistently obtained when the RT reaction was initiated from the Block II region (i.e., from the A homopolymeric tracts), regardless of whether the starting RNA sample was from the Pk1(A+) or Pk1(B+)1+ parasites, or from the ring or trophozoite stages of development. These data were intriguing because they are not consistent with our northern blot data ([12] and unpublished data) which show a large hybridization signal on Pk1(B+)1+ ring RNA and a greatly diminished signal on Pk1(B+)1+ trophozoite RNA, and no signal representing the *205 SICAvar* locus on Pk1(A+) RNAs from either stage. In stark contrast, however, reactions initiated with the Block III primers consistently in a series of experiments resulted in dramatic differences between the Pk1(A+) and Pk1(B+)1+ samples, as well as between the two stages of development. These data are consistent with the northern blot data ([12], and unpublished data), with weak RT-PCR signals generated when a northern signal was not obtained or was weak, and an intense RT-PCR signal generated when the northern blot signal was intense (that is, for the Pk1(B+)1 ring-stage RNA sample).

Quantitative real-time RT-PCR experiments were subsequently performed to verify these data (Fig. 5B and data not shown). Reverse transcription reactions were again initiated with either oligo dT (Block II) or SV6-205A and SV5-205B

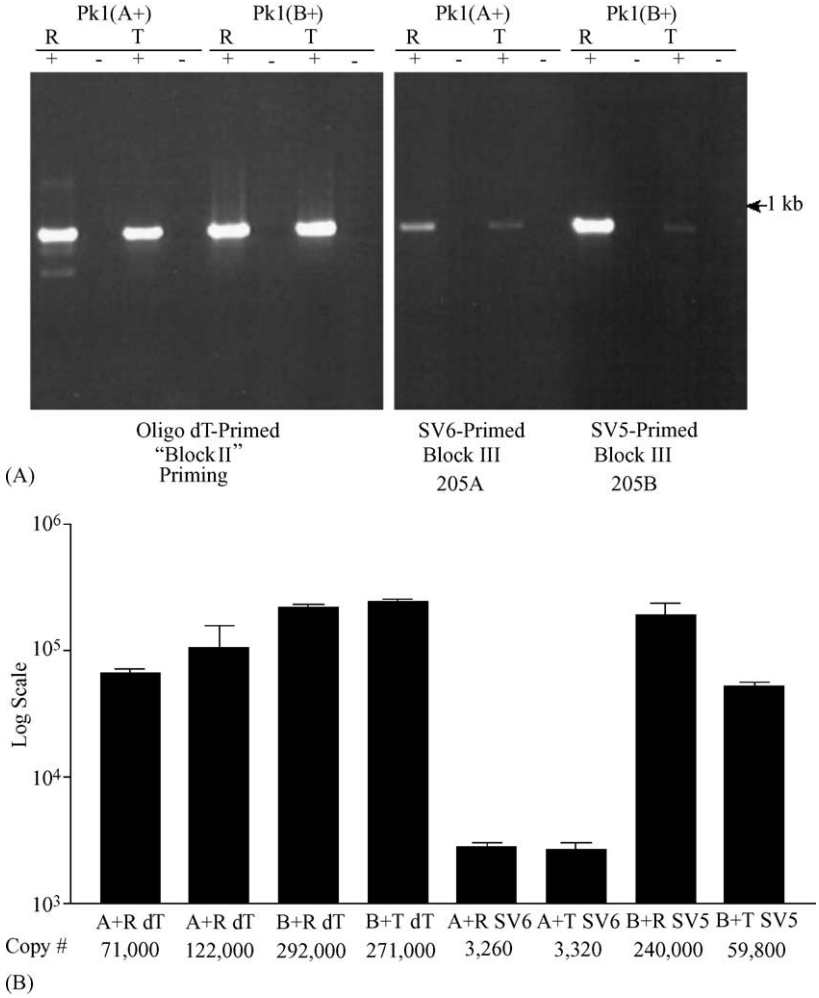

Fig. 5. *205A* and *205B SICAvar* RNAs are detected in apparent similar quantities and Block III sequences may be associated with transcript stability. (A) RT-PCR: Pk1(A+) and Pk1(B+)1+ total RNA samples from the ring (R) or trophozoite (T) stages were reverse transcribed with either an oligo dT primer that was confirmed to target Block II hompolymeric tracts or SV6-205A and SV5-205B Block III primers, as noted. The plus and minus samples denote those with or without reverse transcriptase in the reactions. The reverse transcribed cDNA products were then PCR amplified with GSP3 (exon IX) and cyto5 (exon X) oligonucleotides (Fig. 1). The ~800 bp products generated are presented. (B) qRT-PCR: Pk1(A+) [A+] and Pk1(B+)1+ [B+] total RNA samples from ring (R) or trophozoite (T) stages were reverse transcribed with either oligo dT or SV6-205A and SV5-205B Block III primers, as noted. The reverse transcribed cDNA products were then PCR-amplified in duplicate with exon VIII–IX primer pairs (see Section 2). The results are displayed as a bar graph with standard errors depicted and the resulting copy #s displayed using a log scale. The actual copy #s are also noted in the corresponding horizontal table.

(Block III) primers, and the resulting cDNAs were amplified with gene-specific primers straddling the exon VIII–IX junction, which would result in a 103 bp product. As shown in Fig. 5B, which represents one of a series of three such experiments, the relative amounts of product obtained in the qRT-PCR experiment was consistent and comparable to those obtained in the traditional RT-PCR experiments (Fig. 5A and data not shown). The products generated by Block II primers show no more than a 2–4-fold difference in copy number, with each of the Pk1(A+) products being slightly less than the Pk1(B+)1+ products, consistent with the results visualized in Fig. 5A. Also consistent with the original RT-PCR data, Block II primers produced a 21- and 36-fold greater copy number than the Block III primers on the Pk1(A+) ring and trophozoite-stage samples, respectively, while similar apparent amounts of product as the Block III primers on the Pk1(B+)1+ ring stage sample. The Block II primers also produced a 4.5-fold greater copy number than the Block III primers on the Pk1(B+)1+ trophozoite sample. Given the overall consistency between the RT-PCR and qRT-PCR data, one might question the relatively weak signal obtained in the traditional RT-PCR experiment from the Pk1(B+)1+ trophozoite sample. However, this may simply be due to the fact that this RT-PCR was performed on RNA from a different trophozoite preparation that is predicted to have been slightly more mature; and thus would be predicted to have less full-length *SICAvar* transcript. Importantly in this regard, although there is great reduction in observed *SICAvar* transcripts between the ring and trophozoite stage of development ([12] and unpublished data] the actual point in development when transcription of the *SICAvar* genes ceases or becomes diminished has not yet been defined.

In summary, these series of experiments show (1) the presence of RNA derived from the *205 SICAvar* locus in both the Pk1(A+) and Pk1(B+)1+ parasite clones in both the ring and trophozoite developmental stages; (2) that abundant *205A* RNA can be detected in Pk1(A+) parasites despite the fact that the 205 kDa protein has not been detected in this clone; (3) that Block II and III priming yield different albeit consistent quantifiable results; and (4) that Block III primers only generated comparable high copy numbers in samples where a full-length transcript has been observed in northern blot analyses. These data provide initial evidence that (a) the *205A* transcript is produced but silenced; (b) the *205A SICAvar* transcripts undergo rapid processing, with a cleavage or other type of alteration associated with their silencing occurring at or near the boundary between the Block II and III regions; and (c) the observed reduction in *205B* RNA detected upon development from the ring-stage to the trophozoite-stage with Block III primers is not associated with a similar reduction with Block II primers. This implies that processing (silencing) of the *205B* transcript is initiated at the trophozoite stage and may similarly be dependent on a cleavage (or other alteration) at the Block II–III boundary. As discussed below, whether RNA derived from Block II priming solely represents transcription products or RNA synthesized de novo in the context of a silencing mechanism is an important question to be addressed in future experiments. However, in this regard, the presence of similar low levels of product in the Pk1(A+) samples with Block III priming – compared to Block II priming – could support the likelihood that full-length transcription and associated processing is occurring in both stages. Otherwise, we would only expect to see a product generated with Block III primers from the ring stage sample. It should further be noted that possible differences in the efficiency of reverse transcription from the Block III region of the *205A* or *205B SICAvar* UTRs will be taken into consideration in future experimentation. Moreover, we must remain open to the possibility that these results could to some degree reflect subtle changes in the natural integrity of the *SICAvar* transcripts.

3.6. Hybridization of Pk1(B+)1+ cDNA libraries identifies 22 distinct SICAvar gene clones, indicating that while they express distinct proteins, many SICAvar genes are transcribed in cloned parasites

Importantly, in the initial course of these studies, we determined that many distinct *SICAvar* transcripts could be obtained from Pk1(B+)1+ cDNA library hybridization screening, where PCR amplification sensitivity or biases were not factors. Pk1(B+)1+ cDNA expression libraries representing transcripts from late trophozoites and two-nucleated schizonts (see Section 2) were screened with the DNA of the original *SICAvar* clones #24 and 538-6 [12] as probes. These probes include exon VIII and IX sequence from the *205 SICAvar* gene. Twenty-two distinct *SICAvar* gene clones were identified and partly sequenced to show they correspond to other members of this gene family. The majority proved to be partial *SICAvar* cDNAs, all including exon IX sequence, as well as portions of exon VIII and/or exon X or 3′UTR sequence. Two of the clones also contain introns, and thus, like clone #24, may represent pre-mRNA or gDNA remaining in the cDNA preparation used to make this library. Like clone #24, these two particular clones also terminate with A/T rich sequence representing intron IX. This analysis confirmed: (1) that many *SICAvar* transcripts are produced in clonal *P. knowlesi* parasites, even though this and other *P. knowlesi* parasite clones are known to produce two defined variant proteins at the surface of infected erythrocytes ([6,12,41] and unpublished data), and (2) that detection of multiple *SICAvar* transcripts is feasible from material that is not PCR-amplified; thus the detection of numerous distinct transcripts is not simply due to the sensitivity of the detection procedures, picking up low-level and potentially insignificant basal transcription products, as could be the case when RT-PCR methods are employed as a standard to determine whether and to what extent single or multiple variant antigen gene transcripts are produced. Regardless, RT-PCR, cloning and sequencing experiments have shown that many different 3′ *SICAvar* sequences are readily obtained from highly purified ring and trophozoite stage Pk1(A+) and Pk1(B+)1+ RNA samples (data not shown); thus, the production of many full-length *SICAvar* transcripts in *P. knowlesi* clones has been confirmed through multiple methods.

4. Discussion

We have shown in this report that a previously identified mitotic 3′ genomic DNA alteration in the *205 SICAvar* gene associated with the switch in expression of the 210/190 kDa Pk1(A+) SICA antigen specificity to the 205/200 kDa specificity in Pk1(B+)1+ parasites [12] is the result of a DNA rearrangement involving the transfer of ∼4.0 kb of sequence from an unknown location, to the 3′ portion of the *205A SICAvar* locus in chromosome IV. This rearrangement resulted in the generation of a new allelic variant, the *205B* allele, with different intron IX, exon X and 3′UTR sequences. Of special note, the 3′ non-coding sequences involved in the rearrangement have conspicuous conserved sequence patterns, yet are polymorphic and differ widely in size. Moreover, we have determined that both the *205A* and *205B SICAvar* alleles are completely transcribed; that is, RT-PCR products representing upstream as well as the 3′ most regions were readily amplified for each allele from their respective parasite clones and not just the *205B* allele known to be expressed at the surface of Pk1(B+)1+ infected red blood cells. Significantly, real-time RT-PCR experiments initiated with oligo dT (i.e., Block II priming) confirmed that *205A* and *205B* RNA sequences can be detected in essentially similar quantities; with only small 2–4-fold differences noted. These data, discussed further below, are complemented with data demonstrating that many if not all *SICAvar* transcripts are produced by clonal

Pk1(B+)1+ parasites, as determined by the detection of 22 distinct *SICAvar* cDNAs representing the 3′ most exons in a limited hybridization screening of Pk1(B+)1+ cDNA expression libraries. The finding that many *SICAvar* transcripts, importantly, including the *205A* allele, are completely transcribed but without corresponding protein products evident in the respective *P. knowlesi* clones, prompts questions regarding what controls might be present that prevent their translation. Similarly, what aspects of the *205B* allele permit translation?

In this context, of utmost interest are our findings that relate to the 3′UTRs. Analyses of many gene family members indicate that three differentially conserved yet polymorphic regions, which we have called Blocks I and III, are all part of the transcriptional unit. This structure is noteworthy in that Block I is comprised of imperfect repeated motifs that vary considerably in size and include the GGGTT(T/C)A *Plasmodium* telomeric repeat motif (see [38,39]). Block II, which follows, is predominated by numerous alternating A and T homopolymer tracts. While this region can be aligned, no two sequences come close to being identical. Block III, as we have shown, defines groups of related sequences with GC-rich segments that are also polymorphic. Importantly, the 3′UTRs of the *P. falciparum var* genes were also shown recently to have some degree of conservation and perceptible sequence patterns within discernible groups [40]. These investigators speculated that the 3′UTR sequences may be functionally relevant for regulating transcriptional control, in line with the conventional viewpoint to date that transcriptional control mechanisms are of primary importance for governing the selective control of *var* gene expression [19,21–23,25,42–45]. Alternatively, given that multiple *var* gene transcripts have been shown to exist in *P. falciparum* [18,19,21,22], even in single cells [18,23], we have questioned if PTGS mechanisms as proposed here may function at some level in both species [29].

The fact that the *SICAvar* 3′UTRs – but not necessarily the 5′UTRs – have conserved sequence patterns is not what is expected for genes evolving independently from an ancient common ancestor and suggest a possible functional constraint. From this point of view it is important to consider that 3′UTR sequences may regulate transcripts at the level of transport, stability or translational control (reviewed in [46–48]). Hence, for example, 3′UTR sequences may be targets for protein binding to AU-rich elements regulating mRNA stability [49], complementary small RNA inhibition of translation [50,51], or RNAi mediated post transcriptional gene silencing [52,53]. Differential splicing or polyadenylation could also dictate which messages are transcribed or silenced (reviewed in [54,55]).

From these perspectives, two points are of particular interest. First, while RT-PCR amplifications initiated with oligo dT (Block II priming) consistently resulted in the generation of *205A* and *205B SICAvar* products of relatively similar quantities, Block III priming routinely generated a comparable high copy number only from the Pk1(B+)1+ ring-stage RNA sample. Critically, the Block III-primed products shown here mirror what is observed by northern blot analysis. By northern blot analysis, a full-length *205 SICAvar* transcript was only detected in Pk1(B+)1+ and not Pk1(A+) RNA samples, and in the Pk1(B+)1+ sample the transcript was more abundant in the ring-stage sample [12]. These and other corroborating stage-specific northern blot experiments (unpublished data) have suggested that either the *205A* transcript is not produced, is produced in very low amounts, or is produced to a similar extent as the *205B* transcript but is unstable. The currently presented RT-PCR experiments provide evidence that the *205A* transcript is in fact produced, and show that *205A* RNA segments can be amplified to generate products of a comparable level to the corresponding *205B* products. However, at this stage, it remains in question whether all the observed *205A* product represents actual *205A* transcript or RNA produced de novo, through, for example, an RNAi mediated amplification mechanism (reviewed in [48]). The fact that *205A* RNA is identified in both ring-stage and trophozoite-stage RNA samples via the Block III primer would support the premise that transcription of the *205A* allele is occurring in both stages, and, although this needs to be investigated further, perhaps to the same extent as the *205B* allele. Further investigations along these lines, including nuclear run-on assays, should be able to discern to what degree transcriptional control versus post-transcriptional control is relevant in the expression of the *205 SICAvar* locus in the different clones that either express or do not express the 205 kDa SICA protein. Our data also suggest that a cleavage, or other alteration, of the *205A* transcript at the Block II–III boundary may be associated with the silencing of the *205A* transcript in Pk1(A+) parasites as well as the decrease in the full-length *205B* transcript as the Pk1(B+)1+ parasites develop into trophozoites (Fig. 5B). In this context, the *205A* Block III or associated 3′UTR sequences could promote the immediate programmed processing and silencing of the *205A* transcript. Conversely, the *205B* Block III or associated sequences could confer stability to the full-length *205B* transcript for its incorporation into the translation pathway. These preliminary data are indeed intriguing; yet must be taken with caution as future analyses proceed, with closely monitored sampling time points and regular assessments of the integrity of the RNA under investigation.

It is also worth noting that the *205B* allele has a putative AATAAA polyadenylation signal near the beginning of Block III sequences, which is not present in the *205A* allele, and most other *SICAvar* genes examined. Similarly, and curiously, we have noted that AATAAA motifs are rare in the highly A/T rich *P. falciparum var* gene 3′UTR sequences. Polyadenylation mechanisms may indeed prove to be complex for this gene family. Regardless, it is feasible that polyadenylation characteristics (or lack thereof) may distinguish which transcripts will be translated. Further investigations with analyses of other expressed genes and their transcripts are warranted to test this hypothesis. Moreover, the fact that many transcripts were identified in cDNA library

screens without any PCR amplification step, and without any dominant transcript identified, gives strong support to the hypothesis that many if not all *SICAvar* transcripts are transcribed in clonal parasites, and, importantly, without any apparent preference for more abundant transcription of the locus representing the *SICAvar* genes to be expressed as protein at the surface of the infected host cells. Based on these data, attention needs to be focused on how certain transcripts are ultimately selected for translation.

Our continued investigations of *SICAvar* gene expression will require regular consideration of the potential roles of both transcriptional and PTGS regulatory mechanisms, keeping translation requirements in mind. Foremost, our data do not support a transcriptional control mechanism whereby only one or even a few *SICAvar* transcripts are produced. Clearly, we have detected *SICAvar* transcripts representing many *SICAvar* loci in cloned parasites, which have been produced in their entirety, and *SICAvar* RNA has been identified in both purified ring and mid-stage trophozoites. Nuclear run-on experiments performed with multiple ring and trophozoite-stage samples, with quantification of the nascent *SICAvar* transcripts, would help to clarify the relative contributions of different potential control mechanisms. Scherf et al. have indicated utilizing nuclear run-on assays that a single *P. falciparum var* gene is transcribed in trophozoites [19]. This finding seemed to contradict subsequent data showing the detection by RT-PCR of multiple *P. falciparum var* gene transcripts in a single trophozoite [23]. However, the latter study could reflect the presence of processed *var* transcripts made earlier or RNA made de novo as under consideration here as a possible explanation for the presence of the *205A* and other *SICAvar* RNAs detected in trophozoites. This apparent discrepancy in the literature remains to be clarified. Further, it will be important with regards to understanding transcription of the *SICAvar* genes, to assess if the 5′UTR of other *SICAvar* genes exist as single copy sequences, as shown to be the case for the *205 SICAvar* gene, and whether particular conserved 5′UTR subfamilies exist, as has been shown for the *P. falciparum var* genes [20,43].

The extensive array of repeated motifs that were detected downstream of Block III sequences is reminiscent of the highly recombinogenic subtelomeric structure characteristic of eukaryotic chromosomes [56]. However, a more detailed knowledge of the chromosome structure at the loci harboring *SICAvar* genes awaits a more extensive analysis. The presence of such extensive repeats, as well as long homopolymeric tracts, has also made it problematic to confirm the precise 3′ limit of the observed *SICAvar* rearrangement, as well as the origin of the donor sequence. Importantly, however, our data shows that the original 3′ *205A SICAvar* sequence is not present in the Pk1(B+)1+ genome and appears to be deleted, although this observation could also be due to nucleotide substitutions in this sequence. Future aid of the *P. knowlesi* genome database, as it nears completion, will facilitate the design of potentially unique oligonucleotides to thoroughly map the donor and downstream recipient regions.

The specific recombination event described here also prompts the question whether such events are necessary (and, if so, sufficient) for SICA antigen switching. As a priority, further studies must assess if similar events are associated with other switched phenotypes, generated similarly in vivo. It should also be stressed that the 3′ genetic alteration initially detected in the expressed *205* locus, was not readily apparent, and could have gone unnoticed if the specific diagnostic enzymes and hybridization probes chosen for study at the time had not been analyzed. No other genetic differences have yet been detected between Pk1(A+) and Pk1(B+)1+ parasites. Whether such mitotic recombination events, associated with switch events, could in fact be common in *Plasmodium var* genes in vivo, remains to be clarified.

Conceptually, it is also important to differentiate this recombination event, which is a mitotic event associated with a switch in expression of variant antigens, from meiotic recombination events. As we have demonstrated, the *SICAvar* 3′UTR sequences are indeed subject to numerous ectopic recombination events that have occurred frequently within groups of sequences defined by their 3′ non-coding regions. Similarly, preferential recombination between certain *P. falciparum var* genes has been reported [57]. Recombination during meiosis in mosquito infections has been shown to be frequent in *P. falciparum* [58,59]. However, the relative frequency of mitotic recombination events for either species is unknown. Also, spontaneous or transfection-induced recombination events in the context of investigating the expression and switching of *var* genes in cultured cells have been noted [42,60]. A spontaneous 18 kb deletion that included a *var* gene from within an internal chromosomal *var* gene cluster resulted in the detection by RT-PCR of different *var* gene transcripts in two subcloned populations [42]. Here the inference was that the deletion of one *var* gene locus allowed for the transcription of other *var* genes in the cluster. A subsequent study, however, showed that targeted disruption of a *var* gene did not result in the transcription of a neighboring *var* gene [60]. To date, there are no reports comparable to the event we have described, with an in vivo switch involving a recombinational event that is restricted to the 3′ intron and downstream sequences, with the resulting specific protein expression of the newly generated allele.

Finally, while approximately 108 sequences related to *SICAvar* conserved exon X sequences are present in the *P. knowlesi* genome, it remains to be determined if all of these correspond to full-length *SICAvar* genes, or if *P. knowlesi* has sequences comparable to the *var* C modules in *P. falciparum*, which include the 3′ intron, exon and 3′UTR regions of *var* genes (reviewed in [40]). Whether the final *SICAvar* intron could also be shown to function in the context of the 5′UTR sequence to silence transcription in experimental systems as noted by Deitsch et al. could also be of interest [44]. However, the fact that the *SICAvar 205A* and *205B* transcripts are both clearly transcribed, suggests that the 5′UTR and intron sequences per se are not the

defining factor that silences the expression of the 205 kDa SICA protein in the Pk1(A+) parasites.

In conclusion, this is the first report defining the conserved 3′UTRs of *Plasmodium var*iant antigen genes in detail and in the context of bringing to light the potential functional importance of these sequences in post transcriptional regulation. We hope that continued investigations along these lines will help to discern particular characteristics that can distinguish *SICAvar* (and *var*) gene transcripts that are translated compared to those that are produced but not expressed as protein. Further, the means to study single switching events, such as the one described here, in stable in vivo derived clonal lines expressing well-defined SICA phenotypes, underscores the utility of the *P. knowlesi* system for uncoupling the transcriptional and post-transcriptional control factors that may underpin variant antigen gene expression.

Acknowledgements

The authors acknowledge support from the National Institutes of Allergy and Infectious Diseases, National Institute of Health (AI24710-17, AI35804-07, and AI50409-04A1) and the Yerkes National Primate Research Center base grant #RR-00165 awarded by the Comparative Medicine Program, National Center for Research Resources of the National Institutes of Health. We are also grateful for use of *P. knowlesi* sequence information provided by the Sanger Institute at the *Plasmodium* Genome Resource (http://plasmodb.org/PlasmoDB.shtml) [32], Natalia Kozyr's assistance establishing the quantitative real-time RT-PCR experiments, and Cindy Korir's critical reading of the manuscript and insightful discussions.

Appendix A. Supporting data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molbiopara.2004.05.017.

References

[1] Brown KN. Antibody induced variation in malaria parasites. Nature 1973;242(5392):49–50.
[2] Brown KN, Brown IN. Immunity to malaria: antigenic variation in chronic infections of *Plasmodium knowlesi*. Nature 1965;208(17):1286–8.
[3] Brown IN, Brown KN, Hills LA. Immunity to malaria: the antibody response to antigenic variation by *Plasmodium knowlesi*. Immunology 1968;14(1):127–38.
[4] Barnwell JW, Howard RJ, Coon HG, Miller LH. Splenic requirement for antigenic variation and expression of the variant antigen on the erythrocyte membrane in cloned *Plasmodium knowlesi* malaria. Infect Immun 1983;40(3):985–94.
[5] Barnwell JW, Howard RJ, Miller LH. Altered expression of *Plasmodium knowlesi* variant antigen on the erythrocyte membrane in splenectomized rhesus monkeys. J Immunol 1982;128(1):224–6.
[6] Howard RJ, Barnwell JW, Kao V. Antigenic variation of *Plasmodium knowlesi* malaria: identification of the variant antigen on infected erythrocytes. Proc Natl Acad Sci USA 1983;80(13):4129–33.
[7] Leech JH, Barnwell JW, Miller LH, Howard RJ. Identification of a strain-specific malarial antigen exposed on the surface of *Plasmodium falciparum*-infected erythrocytes. J Exp Med 1984;159(6):1567–75.
[8] Howard RJ, Barnwell JW, Rock EP, et al. Two approximately 300 kilodalton *Plasmodium falciparum* proteins at the surface membrane of infected erythrocytes. Mol Biochem Parasitol 1988;27(2–3):207–23.
[9] Baruch DI, Pasloske BL, Singh HB, et al. Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. Cell 1995;82(1):77–87.
[10] Su XZ, Heatwole VM, Wertheimer SP, et al. The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. Cell 1995;82(1):89–100.
[11] Smith JD, Chitnis CE, Craig AG, et al. Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell 1995;82(1):101–10.
[12] Al-Khedery B, Barnwell JW, Galinski MR. Antigenic variation in malaria: a 3′ genomic alteration associated with the expression of a *P. knowlesi* variant antigen. Mol Cell 1999;3(2):131–41.
[13] Baruch DI. Adhesive receptors on malaria-parasitized red cells. Baillieres Best Pract Res Clin Haematol 1999;12(4):747–61.
[14] Miller LH, Fremount HN, Luse SA. Deep vascular schizogony of *Plasmodium knowlesi* in *Macaca mulatta*. Distribution in organs and ultrastructure of parasitized red cells. Am J Trop Med Hyg 1971;20(6):816–24.
[15] Fischer K, Horrocks P, Preuss M, et al. Expression of var genes located within polymorphic subtelomeric domains of *Plasmodium falciparum* chromosomes. Mol Cell Biol 1997;17(7):3679–86.
[16] Rubio JP, Thompson JK, Cowman AF. The var genes of *Plasmodium falciparum* are located in the subtelomeric region of most chromosomes. EMBO J 1996;15(15):4069–77.
[17] Hernandez-Rivas R, Mattei D, Sterkers Y, Peterson DS, Wellems TE, Scherf A. Expressed var genes are found in *Plasmodium falciparum* subtelomeric regions. Mol Cell Biol 1997;17(2):604–11.
[18] Chen Q, Fernandez V, Sundstrom A, et al. Developmental selection of var gene expression in *Plasmodium falciparum*. Nature 1998;394(6691):392–5.
[19] Scherf A, Hernandez-Rivas R, Buffet P, et al. Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in *Plasmodium falciparum*. EMBO J 1998;17(18):5418–26.
[20] Gardner MJ, Hall N, Fung E, et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 2002;419(6906):498–511.
[21] Noviyanti R, Brown GV, Wickham ME, Duffy MF, Cowman AF, Reeder JC. Multiple var gene transcripts are expressed in *Plasmodium falciparum* infected erythrocytes selected for adhesion. Mol Biochem Parasitol 2001;114(2):227–37.
[22] Fernandez V, Chen Q, Sundstrom A, Scherf A, Hagblom P, Wahlgren M. Mosaic-like transcription of var genes in single *Plasmodium falciparum* parasites. Mol Biochem Parasitol 2002;121(2):195–203.
[23] Duffy MF, Brown GV, Basuki W, et al. Transcription of multiple var genes by individual, trophozoite-stage *Plasmodium falciparum* cells expressing a chondroitin sulphate A binding phenotype. Mol Microbiol 2002;43(5):1285–93.
[24] Duffy MF, Reeder JC, Brown GV. Regulation of antigenic variation in *Plasmodium falciparum*: censoring freedom of expression? Trends Parasitol 2003;19(3):121–4.

[25] Taylor HM, Kyes SA, Harris D, Kriek N, Newbold CI. A study of *var* gene transcription in vitro using universal *var* gene primers. Mol Biochem Parasitol 2000;105(1):13–23.
[26] Kyes SA, Christodoulou Z, Raza A, et al. A well-conserved *Plasmodium falciparum var* gene shows an unusual stage-specific transcript pattern. Mol Microbiol 2003;48(5):1339–48.
[27] Fried M, Duffy PE. Two DBLgamma subtypes are commonly expressed by placental isolates of *Plasmodium falciparum*. Mol Biochem Parasitol 2002;122(2):201–10.
[28] Peters J, Fowler E, Gatton M, Chen N, Saul A, Cheng Q. High diversity and rapid changeover of expressed *var* genes during the acute phase of *Plasmodium falciparum* infections in human volunteers. Proc Natl Acad Sci USA 2002;99(16):10689–94.
[29] Galinski MR, Corredor V. Variant antigen expression in malaria infections: posttranscriptional gene silencing, virulence and severe pathology. Mol Biochem Parasitol 2004;134(1):17–25.
[30] Corredor V, Enea V. The small ribosomal subunit RNA isoforms in *Plasmodium cynomolgi*. Genetics 1994;136(3):857–65.
[31] Galinski MR, Medina CC, Ingravallo P, Barnwell JW. A reticulocyte-binding protein complex of *Plasmodium vivax* merozoites. Cell 1992;69(7):1213–26.
[32] Bahl A, Brunk B, Crabtree J, et al. PlasmoDB: the *Plasmodium* genome resource. A database integrating experimental and computational data. Nucleic Acids Res 2003;31(1):212–5.
[33] Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25(24):4876–82.
[34] Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 1999;27(2):573–80.
[35] Nei M, Kumar S. Phylogenetic trees. In: Molecular evolution and phylogenetics. New York: Oxford University Press Inc.; 2000. p. 73–85.
[36] Hudson RR, Kaplan NL. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 1985;111(1):147–64.
[37] Rozas J, Rozas R. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 1999;15(2):174–5.
[38] de Bruin D, Lanzer M, Ravetch JV. The polymorphic subtelomeric regions of *Plasmodium falciparum* chromosomes contain arrays of repetitive sequence elements. Proc Natl Acad Sci USA 1994;91(2):619–23.
[39] Figueiredo LM, Freitas-Junior LH, Bottius E, Olivo-Marin JC, Scherf A. A central role for Plasmodium falciparum subtelomeric regions in spatial positioning and telomere length regulation. EMBO J 2002;21(4):815–24.
[40] Mercereau-Puijalon O, Barale JC, Bischoff E. Three multigene families in *Plasmodium* parasites: facts and questions. Int J Parasitol 2002;32(11):1323–44.
[41] Howard RJ. Antigenic variation of bloodstage malaria parasites. Philos Trans R Soc Lond B Biol Sci 1984;307(1131):141–58.
[42] Deitsch KW, del Pinal A, Wellems TE. Intra-cluster recombination and var transcription switches in the antigenic variation of *Plasmodium falciparum*. Mol Biochem Parasitol 1999;101(1-2):107–16.
[43] Voss TS, Thompson JK, Waterkeyn J, et al. Genomic distribution and functional characterisation of two distinct and conserved *Plasmodium falciparum var* gene 5′ flanking sequences. Mol Biochem Parasitol 2000;107(1):103–15.
[44] Deitsch KW, Calderwood MS, Wellems TE. Malaria. Cooperative silencing elements in *var* genes. Nature 2001;412(6850): 875–6.
[45] Winter G, Chen Q, Flick K, Kremsner P, Fernandez V, Wahlgren M. The 3D7var5.2 (var(COMMON)) type var gene family is commonly expressed in non-placental *Plasmodium falciparum* malaria. Mol Biochem Parasitol 2003;127(2):179–91.
[46] Kuersten S, Goodwin EB. The power of the 3′ UTR: translational control and development. Nat Rev Genet 2003;4(8):626–37.
[47] Wilkie GS, Dickson KS, Gray NK. Regulation of mRNA translation by 5′- and 3′-UTR-binding factors. Trends Biochem Sci 2003;28(4):182–8.
[48] Denli AM, Hannon GJ. RNAi: an ever-growing puzzle. Trends Biochem Sci 2003;28(4):196–201.
[49] Shaw G, Kamen R. A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 1986;46(5):659–67.
[50] Olsen PH, Ambros V. The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. Dev Biol 1999;216(2):671–80.
[51] Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell 2001;106(1):23–34.
[52] Metzlaff M, O'Dell M, Cluster PD, Flavell RB. RNA-mediated RNA degradation and chalcone synthase A silencing in petunia. Cell 1997;88(6):845–54.
[53] Metzlaff M, O'Dell M, Hellens R, Flavell RB. Developmentally and transgene regulated nuclear processing of primary transcripts of chalcone synthase A in petunia. Plant J 2000;23(1):63–72.
[54] Legendre M, Gautheret D. Sequence determinants in human polyadenylation site selection. BMC Genomics 2003;4(1):7.
[55] Zarudnaya MI, Kolomiets IM, Potyahaylo AL, Hovorun DM. Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. Nucleic Acids Res 2003;31(5):1375–86.
[56] Mefford HC, Trask BJ. The complex structure and dynamic evolution of human subtelomeres. Nat Rev Genet 2002;3(2):91–102.
[57] Salanti A, Jensen AT, Zornig HD, et al. A sub-family of common and highly conserved *Plasmodium falciparum* var genes. Mol Biochem Parasitol 2002;122(1):111–5.
[58] Freitas-Junior LH, Bottius E, Pirrit LA, et al. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. Nature 2000;407(6807):1018–22.
[59] Taylor HM, Kyes SA, Newbold CI. Var gene diversity in *Plasmodium falciparum* is generated by frequent recombination events. Mol Biochem Parasitol 2000;110(2):391–7.
[60] Horrocks P, Pinches R, Kyes S, et al. Effect of var gene disruption on switching in *Plasmodium falciparum*. Mol Microbiol 2002;45(4):1131–41.