THE ROYAL SOCIETY

Received 4 September 2001
Accepted 30 October 2001
Published online 11 February 2002

A large gene family for putative variant antigens shared by human and rodent malaria parasites[1]

Christoph S. Janssen*, Michael P. Barrett, C. Michael R. Turner and R. Stephen Phillips

Division of Infection and Immunity, IBLS, University of Glasgow, Glasgow G12 8QQ, UK

A major mechanism whereby malaria parasites evade the host immune response to give chronic infections in patients' blood for months, or even years, is antigenic variation. In order to generate variant antigens, parasites require large multigene families. Although several gene families involved in these phenomena have been identified in the human malaria *Plasmodium falciparum*, to date no variant antigen gene families have been identified in malaria species that will infect widely used rodent laboratory hosts. Here we present, for the first time, to our knowledge, a large multigene family conserved in both rodent and human malarias, which is a strong candidate as a major variant antigen gene family. In each of four species of *Plasmodium*, three rodent malarias and the human pathogen *P. vivax*, homologues of the gene family were found to have a conserved three-exon structure. In the rodent malaria *P. chabaudi,* transcription of members of the gene family was developmentally regulated with maximum expression in late trophozoite stages, which is the developmental stage known to express variant antigen proteins.

Keywords: malaria; antigenic variation; multigene families; variant antigen transcription; rodent malaria

1. INTRODUCTION

Malaria, a disease caused by the protozoan parasites of the genus *Plasmodium*, remains the most important parasitic infection of man, causing 300–500 million clinical cases and 2.7–3 million deaths each year (Phillips 2001). The parasites persist in people by employing the immune evasion mechanism of antigenic variation, which extends both the disease and opportunities for transmission to new hosts via mosquito vectors. Antigenic variation is achieved by malaria parasites through the exclusive expression of a single variant member of a large antigen gene family by any single parasite at a given time. Within an infection, parasites switch expression among members of this variant multigene family, thus enabling parts of the parasite population to evade specific immune responses developed to previously expressed antigens of the same family (Phillips *et al.* 1997).

The first indication that asexual blood-stage malaria parasites could undergo antigenic variation was the work of Cox (1962) with the rodent parasite *P. berghei*, while Brown and Brown demonstrated that repeated antigenic variation could occur during the course of a chronic infection with the simian malaria *P. knowlesi* (Brown & Brown 1965). Subsequently, antigenic variation was shown to occur in two further monkey malarias, *P. cynomolgi bastianelli* (Voller & Rossan 1969) and *P. fragile* (Handunnetti *et al.* 1987), in the murine malaria *P. chabaudi* AS strain (McLean *et al.* 1982; Phillips *et al.* 1997) and in *P. falciparum* in the squirrel monkey (Hommel *et al.* 1983) and *in vitro* (Biggs *et al.* 1992). The most studied variant antigen in *P. falciparum* is known as erythrocyte membrane protein 1 (PfEMP1) (Leech *et al.* 1984) and the encoding genes belong to a large diverse family called *var* genes (Smith *et al.* 1995; Su *et al.* 1995; Baruch *et al.* 1997). Two further gene families, *stevor* and *rif*, have also been implicated in antigenic variation in *P. falciparum* (Cheng *et al.* 1998; Kyes *et al.* 1999). Very recently, a surface variant antigen of *P. vivax* was found to be encoded by another multigene family, the *vir* gene family (del Portillo *et al.* 2001).

Antigenic variation clearly poses a particular problem in the context of vaccines targeted at the asexual blood stages of the parasite. Variant surface antigens are highly immunogenic and are targets of the protective immune response (Saul 1999). A better understanding of the impact of antigenic variation on the dynamics of the immune response to blood-stage parasites will facilitate vaccine development. However, experimental investigations of antigenic variation in the human malarias are limited by the lack of suitable laboratory hosts for the human parasites. Experimental immunology and genetics require that parasite–host interactions be amenable to manipulation under controlled laboratory conditions and hence rodent malarias provide the experimental models of choice.

We identified a major gene family in the rodent malaria parasite *P. chabaudi* during analysis of data derived from a genome survey sequencing project (Janssen *et al.* 2001). We have also identified homologues of this family in the rodent malarias *P. yoelii* and *P. berghei*, and, importantly, in the human malaria *P. vivax*. Here, we present data on the genomic structure of these genes, their relationship to each other and their transcription pattern during the asexual erythrocytic cycle in *P. chabaudi*. Recently, data were published on the *P. vivax* gene family *vir*, which is homologous to the above rodent malaria gene families, giving strong evidence that it encodes immunovariant proteins that have a functional role in antigenic variation (del

*Author for correspondence (c.janssen@bio.gla.ac.uk).

Portillo *et al.* 2001). We therefore propose that the genes we present here are most probably members of a very large variant antigen gene family. Most significantly, this is, to our knowledge, the first time that such genes have been described in both human and laboratory model malarias.

2. MATERIAL AND METHODS

The parasites used in this study were an AS strain of *P. chabaudi*, originally provided by Professor D. Walliker (University of Edinburgh) as a cloned and mosquito-transmitted line (see Phillips *et al.* (1997) for details). Inbred BALBc or outbred ICR mice kept in a reverse light : dark cycle were infected with *P. chabaudi* and bloodstream parasites were recovered, as described (Janssen *et al.* 2001).

Partial sequence data for members of a major *P. chabaudi* gene family were obtained from analyses of a genome survey sequencing project that we carried out previously and have described elsewhere (Janssen *et al.* 2001). Sequence information of some members of this gene family, which we propose to name *cir* (‘*fam3*’ in the previous publication), was extended using rapid amplification of cDNA ends (RACE) to give a full-length mRNA sequence. Briefly, a PCR-based method was employed using Clontech’s SMART RACE technology according to the manufacturer’s instructions. Obtained products were cloned into plasmid vectors and sequenced using dye-terminator chemistry (Bowman *et al.* 1999). Sequences with identical overlapping ends were assembled into a putative full-length mRNA sequence. These sequences were then verified by cloning and sequencing of PCR products obtained from amplification of full-length *cir* cDNA using primers targeting the untranslated regions at both 3′ and 5′ ends.

To obtain an estimate of the relative abundance of *cir* in the genome, Southern blots of genomic DNA, digested with various restriction enzymes, were probed with ^{32}P-dATP-labelled full-length *cir* cDNA. Full-length genomic sequences of *cir* were obtained by cloning and sequencing of the PCR products of genomic DNA amplification with primers targeting either end of the full mRNA sequence.

In order to profile the transcription patterns of members of this gene family during parasite development in the red blood cells, Northern blot analysis using RNA isolated from five points in the synchronous 24 h asexual erythrocytic cycle was carried out. A synchronous parasite clone was passaged through one mouse before being used to infect mice held in reverse light cycle conditions. Parasites were collected at early ring, ring, late ring, late trophozoite and schizont stages of asexual development. The total RNA was isolated from the parasites using an acid guanidinium thiocyanate-phenol method (Chomczynski & Sacchi 1987) and digested with DNase. The RNA was Northern blotted by standard procedures (Sambrook *et al.* 1989). Equal loading of gel lanes and blot transfer efficiency were verified by methylene blue staining of membranes after blotting. Hybridization probes were generated by PCR of *cir* cDNA and radiolabelled with ^{32}P-dATP using the random priming method (Feinberg & Vogelstein 1983). Hybridizations and washes were carried out as described previously (Janssen *et al.* 2001). Autoradiography was carried out using standard methods (Sambrook *et al.* 1989).

Genomic (GSS) and expressed sequence tag (EST) data for *P. berghei* ANKA clone 15cy1 (Carlton & Dame 2000) were obtained from the University of Florida, http://parasite.vetmed.ufl.edu/berg.htm. Preliminary sequence data from the *P. yoelii* genome (strain 17X NL, clone 1.1) were obtained from The Institute for Genomic Research website (www.tigr.org). This sequencing program is carried out in collaboration with the Naval Medical Research Centre and is supported by the US Department of Defence. The sequence of the *P. vivax*, IVD10 YAC clone was produced by the Pathogen Sequencing Group at the Sanger Centre and can be obtained from ftp://ftp.sanger.ac.uk/pub/pathogens/vivax/.

Similarity searching of public databases that contain nucleic acid and protein sequences with the *cir* sequence was carried out using the BLAST algorithm (Altschul *et al.* 1997). The *P. yoelii* genome project, *P. vivax* YAC and the *P. berghei* GSS and EST databases were queried locally with *cir*, using the FASTA algorithm (Pearson 2000). Sequences considered to be homologous by virtue of a highly significant similarity ($E < 10^{-8}$) were analysed to identify protein coding regions and other features within the ARTEMIS package (Rutherford *et al.* 2000). EST and mRNA sequences were aligned with homologous genomic sequence data by the DIALIGN program (Morgenstern *et al.* 1996; Morgenstern 1999), using the option of translating the compared ‘nucleic acid segments’ to ‘peptide segments’. *In silico* analyses of the predicted consensus protein sequence were carried out using PIX at the Human Genome Mapping Project Resource Centre (http://www.hgmp.mrc.ac.uk/).

Phylogenetic analysis was performed on the *cir* homologue dataset using the maximum-likelihood method (Strimmer & von Haeseler 1997) applied to pairwise sequence distances calculated using quartet puzzling, which automatically assigns estimations of support to each internal branch (Strimmer & von Haeseler 1996). Trees were drawn using the program TREEVIEW (Page 1996).

3. RESULTS

Members of the *cir* gene family were previously identified as short *P. chabaudi* GSS fragments (‘*fam3*’) (Janssen *et al.* 2001). However, the quantity of sequence information originally available proved insufficient to identify related genes in other organisms and it was not possible to assign functions to these genes.

Initial Northern hybridization revealed a *cir* gene transcription product of *ca*.1.8 kb in length in late trophozoites (Janssen *et al.* 2001). Extension of one of the original GSS *cir* sequences using RACE generated two distinct full-length mRNA sequences, *cir*1 and *cir*2 (accession numbers AJ315472, AJ315473). The length of complete transcripts was 1587 nucleotides (nt) for *cir*1 and 1599 nt for *cir*2. Open reading frames (ORFs) generated by conceptual translation of the mRNA sequences were 918 nt (306 amino acids (aa)) and 936 nt (312 aa) for *cir*1 and *cir*2, respectively. Comparison of the mRNA sequences with sequences of corresponding genomic clones of *cir*1 and *cir*2 (accession numbers AJ315474, AJ315475) revealed both genes to have a three exon/two intron structure. A short first exon (30 nt) is separated from the main exon (812–830 nt) by an intron of 137 nt. The final exon (89 nt) follows an 101 nt intron. The untranslated regions (UTRs) are *ca*. 306 nt at the 5′ end and 109 nt at the 3′ end.

Southern blot hybridizations (figure 1*a*) revealed that the *P. chabaudi* genome contains an extensive family of sequences related to *cir*. Although Southern hybridizations did not allow for a direct measurement of copy numbers

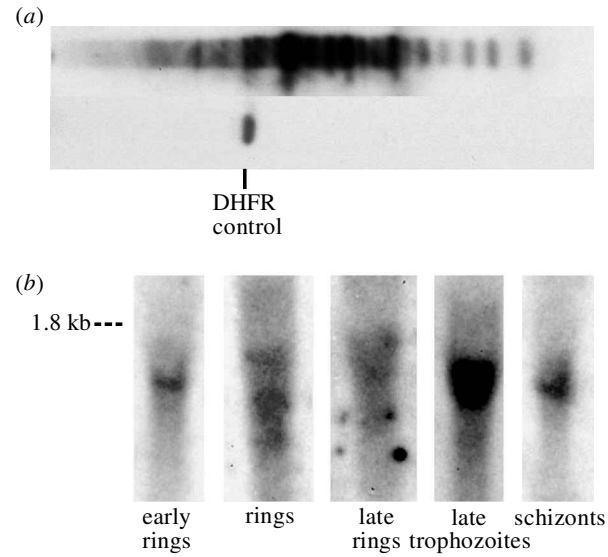

Figure 1. (*a*) *P. chabaudi* AS strain genomic DNA was digested with *Nsi*I. Dihydrofolate reductase (DHFR) was used as a control for digest completion. The final wash of the hybridization was at 60 °C with 0.5 × SSPE. (*b*) A synchronous clone of *P. chabaudi* AS strain parasites was passaged through one mouse before being used to infect mice held in reverse light cycle conditions. Parasites were collected at the indicated stages of asexual development and total RNA was extracted by standard methods. Equal loading of gel lanes and blot transfer efficiency were verified by methylene blue staining of membranes after blotting. The final washes of the hybridizations were at 55 °C with 1 × SSPE.

of *cir* in the whole genome, by extrapolation of the representation of *cir* members in the original GSS (Janssen *et al.* 2001) an estimate can be made of approximately 600 family members.

Northern blot analysis using RNA isolated from five points in the synchronous asexual erythrocytic cycle revealed that *cir* transcription peaks at the late trophozoite stage (figure 1*b*). Transcription levels during other stages of the life cycle are very low, at the limit of detection by Northern hybridization. During progression of the parasites' erythrocytic cycle from early ring to late ring/early trophozoite stage, changes in number and size of hybridizing RNA bands are evident. Ring-stage parasite RNA appears to contain three distinct *cir*-like sequences (figure 1*b*). This pattern is similar, though less pronounced, in late ring/early trophozoite parasites, while only one major transcription product appears to hybridize with *cir* in late trophozoites. Schizonts show greatly reduced levels of *cir* transcript, apparently of the same size as that evident in late trophozoites.

FASTA (Pearson 2000) and BLAST (Altschul *et al.* 1997) algorithm-based searches for *cir*-related sequences in other organisms revealed highly similar gene families in *P. berghei*, *P. yoelii* and *P. vivax*. At the amino-acid level, identities of 30–50% were found between *cir* and *P. berghei* and *P. yoelii* homologues (figure 2). We propose to name the *cir* homologues in *P. berghei* '*bir*', and '*yir*' in *P. yoelii*. Importantly, *bir* sequences were identified in the *P. berghei* EST database, showing that these genes are also transcribed. The *P. vivax* homologues of *cir* genes are the *vir* multigene family members described by del Portillo *et al.* (2001). Identity levels between *cir* and *vir* range between 20 and 30% at the amino-acid level. No evidence could be found for sequences sharing significant similarity to *cir* in the *P. falciparum* genome, or among the non-redundant sequences of GenBank.

Comparative genomic analysis revealed that intron/exon structure and splice junction sequences are highly conserved among *cir* homologues in the rodent malaria species (figure 3). All of the genes analysed have a very short first exon, a long second exon accounting for over 80% of the coding region and a short third exon. Both *cir* genes had identical splice junction sequences as follows: end 1st exon ctg/gtat, 1st intron/2nd exon junction tag/tgtg, end 2nd exon aag/gtaa and 2nd intron/3rd exon junction tag/tatt. The four *yir* sequences analysed in detail (accession numbers AJ320478–AJ320481) had splice junctions largely identical to those of *cir*: end 1st exon rtg/gtat, 1st intron/2nd exon junction tag/tgtr, end 2nd exon aag/gtaa 2nd intron/3rd exon junction tag/tatt. *Bir* genes were more difficult to analyse completely, since genomic information for *P. berghei* is more limited. The two complete *bir* genomic sequences analysed (accession number AJ320482) had splice junctions identical to those of *yir*, with the exception of the 1st exon end, which was gtg/gtay. Two *bir* genomic sequence fragments had splice junctions for the first two exons identical to those of the other *bir* sequences. *Vir* genes, as annotated by del Portillo *et al.* (2001), generally have the same intron/exon structure as their rodent parasite homologues. Sequences of *vir* splice junctions are also very similar to their rodent malaria counterparts, but show greater diversity among different members of the family. The *in silico* PIX analysis at the HGMP Resource Centre revealed a putative transmembrane domain spanning at least 20 amino acids at the very end of the second exon in all rodent *cir* homologue sequences. The third exon, which is highly conserved, is predicted as a cytoplasmic domain. No N-terminus consensus signal-peptide sequences were detected in any of the analysed sequences.

Phylogenetic analysis was carried out on all described *cir* homologue genes from the rodent malarias and three *vir* genes chosen on the basis of their similarity to *cir* (figure 4). The resultant tree shows the rodent malaria genes clustering together, as might be predicted from other studies (Brooks & McLennan 1992). The *vir* genes form the outgroup and cluster separately. Notably, the variation among the *vir* genes appears greater than distances among homologues in the other species.

4. DISCUSSION

We have identified a variant multigene family, which is highly conserved in three malaria species that infect rodents and is also found in the human parasite *P. vivax*. This gene family has been shown to have a substantial number of variant members in each species studied; estimates of gene copy numbers of up to 600 per genome in *P. chabaudi* and *P. vivax* seem probable. The elucidation of the biological role of these genes is still in its infancy; however, the significance of the gene family to the parasites is apparent from the number of times it is represented in the genome. The high level of sequence diversity found within

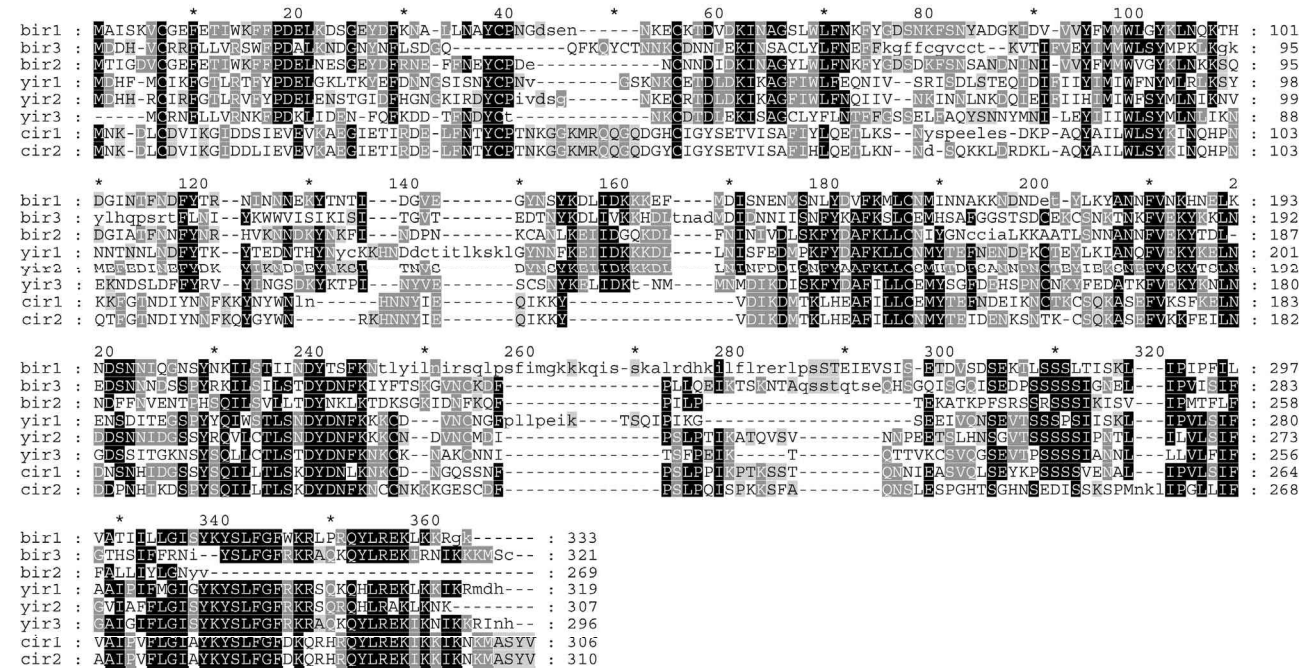

Figure 2. DIALIGN2-generated alignment of translations of exons from *cir* sequences. Shading highlights conserved amino-acid residues. See § 2 for details. Accession numbers of sequences: *cir*1 & *cir*2 AJ315474, AJ315475; *yir* AJ320478–AJ320480; *bir* AJ320482.

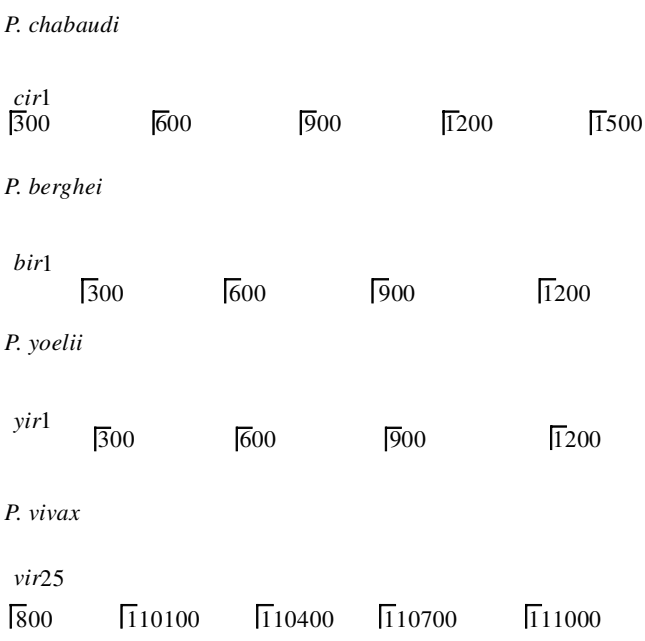

Figure 3. Exons of annotations of *cir* and homologues on contigs from *P. chabaudi*, *P. berghei*, *P. yoelii* and *P. vivax*. White boxes represent exons; numbers show nt position on contig. Accession numbers of sequences: *vir*25 AL360354; *cir*1 AJ315474; *yir*1 AJ320478; *bir*1 AJ320482.

the family is consistent with, and may indicate, immune selection and consequent antigenic variation. Significantly, the developmentally-regulated expression of *cir* in *P. chabaudi* is consistent with the expression of variant antigen types on infected erythrocyte surfaces during the course of an infection (McLean *et al.* 1982; Brannan *et al.* 1994).

The scope for investigations on the biological role of *cir*, the significance it may have for host–parasite interactions

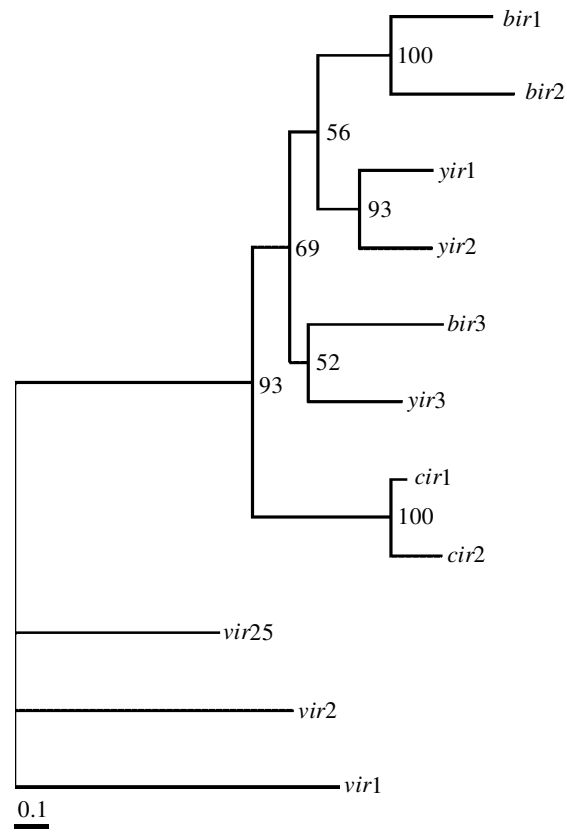

Figure 4. The phylogenetic tree was derived by the maximum-likelihood method applied to pairwise sequence distances calculated using quartet puzzling (see http://www.tree-puzzle.del). Branch lengths are drawn to scale: bar = 0.1 amino-acid replacements per site. Numbers indicate the support values for each branch. Accession numbers of sequences: *cir*1 & *cir*2 AJ315474, AJ315475; *yir* AJ320478–AJ320481; *bir* AJ320482; *vir* AL360354.

and its relevance to human disease has been greatly enhanced by the discovery of homologues in three other malaria species. The degree of conservation at the predicted amino-acid level among sequences from rodent malarias is high in the context of genes that are variant members of gene families. Variation among genes examined within a single genome was on a par with variation in predicted protein sequences among genomes of the three rodent malaria species (figure 2). The *P. vivax vir* genes showed greater sequence diversity. The analysed *vir* genes were chosen on the basis of their highest similarity to the rodent malaria genes, which placed constraints on their potential variability. Nevertheless, the *vir* genes showed greater diversity among themselves than any of the rodent malaria genes (figure 4). This observation is reiterated by a remarkable conservation of splice junction sequences among the genes in the rodent parasite species, which does not appear to hold true to the same extent for the *P. vivax vir* genes. The tree analysis results presented here mirror the phylogenetic relationships among the species as inferred from small subunit rRNA (Siddall & Barta 1992) and circumsporozoite protein (CSP) (Ayala *et al.* 1998) sequences.

On individual chromosomes, *cir* is very probably arrayed in sub-telomeric regions. Telomeric repeat motifs (Pace *et al.* 1987; Dore *et al.* 1990) mapped to most of the larger contigs containing *bir* and *yir* gene. Likewise, *vir* genes have been shown to be arranged in sub-telomeric arrays (del Portillo *et al.* 2001).

Investigations into transcription of *cir* revealed slight variations of transcript size when hybridizing *cir* probes to Northern blots of RNA derived from parasites isolated from separate experimental infections. This variation was observed to a greater extent in *cir*-hybridizing bands of RNA derived from a synchronous infection sampled during the asexual erythrocytic cycle (figure 1*b*). It is probable that transcript size variation in different parasite populations reflects a predominance of expression of a single variant member of the gene family. In this context it is noteworthy, however, that the variant *cir* mRNA transcripts, which were cloned and sequenced, came from an RNA preparation derived from a single infection experiment. A possible explanation for this result may derive from the fact that the sensitivity of the PCR-based methods used to produce these clones is such that even low-abundance transcripts may be isolated. Considering the premise that *cir* may well be involved in antigenic variation, this result would not be unexpected since it has been shown that major and minor variant antigen types do coexist in a parasite population (Brannan *et al.* 1994), reflecting the relatively high *in vivo* switching rates. Processing of pre-mRNA *cir* transcripts may provide an explanation for the multiple-transcript patterns observed in ring and late ring stages. The two-intron structure indicated by the genomic/mRNA sequence comparison may be reflected by the triple band hybridization pattern observed in these developmental stages, showing the various splicing stages of the pre-mRNA. Significantly, the same three-band pattern could be obtained using variant gene-specific reverse transcription-polymerase chain reaction (RT-PCR) (results not shown), while the genomic DNA control reaction of the same conditions only yields a single amplicon of the same size as the largest RT-PCR product. The observed *cir*/RNA hybridization patterns thus shed some light on a potential control mechanism possibly utilized in *cir* expression. Transcript abundance is low during the beginning and end of the developmental cycle, transcription being evidently most active during the middle of the cycle. Enhanced transcript stability and a slow down in transcription as parasites reach the late trophozoite stage would account for the peak in detectable *cir* transcripts, while lowering the amount of pre-processed mRNA detectable. A halt to *cir* transcription and a significant decrease in the half-life of *cir* mRNA during schizont development would explain the dramatic reduction in detectable products. Although the parasites used in this study were undergoing synchronous development at the population level, slight variations in developmental state among individual parasites are expected, with an overlap of *ca*. 2 h between the appearance of the first-ring stages and rupture of the last schizonts at the end of the cycle. It is therefore probable that *cir* transcripts detected in early-ring parasite preparations are actually carried over from some late schizonts.

The results from these first steps in the characterization of expression patterns of *cir* genes fit well with their potential role as variant antigens. In the human malaria parasite *P. falciparum,* variant antigen genes belonging to the *rif* gene family (Fernandez *et al.* 1999) show a similar timing of transcriptional peak during the erythrocytic asexual cycle (Kyes *et al.* 1999). The well-described *var* genes of *P. falciparum* (Smith *et al.* 1995; Su *et al.* 1995) appear to undergo more complex transcriptional regulation, with ubiquitous *var* gene transcription during the earlier parts of the asexual developmental cycle and selective transcription of only one dominant variant gene in trophozoites (Chen *et al.* 1998; Scherf *et al.* 1998). Most significantly, immunofluorescent techniques have demonstrated that *P. chabaudi* expresses variant antigen types during the late trophozoite-schizont stages of development in the bloodstream (Brannan *et al.* 1994).

The strongest evidence for *cir* and its homologues coding for variant antigen genes comes from recently presented work demonstrating that *vir* genes appear to code for immunovariant proteins functioning in antigenic variation (del Portillo *et al.* 2001). The relationship of *cir* and its rodent malaria homologues to the human parasite *P. vivax vir* gene family makes them invaluable tools. The results from experimental immunology and genetics with the rodent malaria *cir* genes will enable predictions to be made concerning the interactions of *P. vivax* with its human host, which can be checked by clinical observations for their direct relevance to the human disease.

This work was supported by funding from the Wellcome Trust.

ENDNOTE

[1]Nucleotide sequence data reported in this paper are available in the GenBank, EMBL and DDBJ databases under the accession numbers AJ315472–AJ315475 and AJ320478–AJ320482.

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic. Acids Res.* **25**, 3389–3402.

Ayala, F. J., Escalante, A. A., Lal, A. A. & Rich, S. M. 1998 Evolutionary relationships of human malaria parasites. In *Malaria: parasite biology, pathogenesis, and protection* (ed. I. W. Sherman), pp. 285–300. Washington, DC: American Society for Microbiology Press.

Baruch, D. I., Ma, X. C., Singh, H. B., Bi, X. H., Pasloske, B. L. & Howard, R. J. 1997 Identification of a region of PfEMP1 that mediates adherence of *Plasmodium falciparum* infected erythrocytes to CD36: conserved function with variant sequence. *Blood* **90**, 3766–3775.

Biggs, B. A., Anders, R. F., Dillon, H. E., Davern, K. M., Martin, M., Petersen, C. & Brown, G. V. 1992 Adherence of infected erythrocytes to venular endothelium selects for antigenic variants of *Plasmodium falciparum*. *J. Immunol.* **149**, 2047–2054.

Bowman, S. (and 35 others) 1999 The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*. *Nature* **400**, 532–538.

Brannan, L. R., Turner, C. M. & Phillips, R. S. 1994 Malaria parasites undergo antigenic variation at high rates *in vivo*. *Proc. R. Soc. Lond.* B **256**, 71–75.

Brooks, D. R. & McLennan, D. A. 1992 The evolutionary origin of *Plasmodium falciparum*. *J. Parasitol.* **78**, 564–566.

Brown, K. N. & Brown, I. N. 1965 Immunity to malaria: antigenic variation in chronic infections of *Plasmodium knowlesi*. *Nature* **208**, 1286–1288.

Carlton, J. M. & Dame, J. B. 2000 The *Plasmodium vivax* and *P. berghei* gene sequence tag projects. *Parasitol. Today* **16**, 409.

Chen, Q. J., Fernandez, V., Sundstrom, A., Schlichtherle, M., Datta, S., Hagblom, P. & Wahlgren, M. 1998 Developmental selection of *var* gene expression in *Plasmodium falciparum*. *Nature* **394**, 392–395.

Cheng, Q., Cloonan, N., Fischer, K., Thompson, J., Waine, G., Lanzer, M. & Saul, A. 1998 *stevor* and *rif* are *Plasmodium falciparum* multicopy gene families which potentially encode variant antigens. *Mol. Biochem. Parasitol.* **97**, 161–176.

Chomczynski, P. & Sacchi, N. 1987 Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analyt. Biochem.* **162**, 156–159.

Cox, H. W. 1962 The behaviour of *Plasmodium berghei* strains isolated from relapsed infections of white mice. *J. Parasitol.* **9**, 114–118.

del Portillo, H.A. (and 12 others) 2001 A superfamily of variant genes encoded in the subtelomeric region of *Plasmodium vivax*. *Nature* **410**, 839–842.

Dore, E., Pace, T., Ponzi, M., Picci, L. & Frontali, C. 1990 Organization of subtelomeric repeats in *Plasmodium berghei*. *Mol. Cell Biol.* **10**, 2423–2427.

Feinberg, A. P. & Vogelstein, B. 1983 A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* **132**, 6–13.

Fernandez, V., Hommel, M., Chen, Q. J., Hagblom, P. & Wahlgren, M. 1999 Small, clonally variant antigens expressed on the surface of the *Plasmodium falciparum*-infected erythrocyte are encoded by the *rif* gene family and are the target of human immune responses. *J. Exp. Med.* **190**, 1393–1403.

Handunnetti, S. M., Mendis, K. N. & David, P. H. 1987 Antigenic variation of cloned *Plasmodium fragile* in its natural host *Macaca sinica*. Sequential appearance of successive variant antigenic types. *J. Exp. Med.* **165**, 1269–1283.

Hommel, M., David, P. H. & Oligino, L. D. 1983 Surface alterations of erythrocytes in *Plasmodium falciparum* malaria. Antigenic variation, antigenic diversity, and the role of the spleen. *J. Exp. Med.* **157**, 1137–1148.

Janssen, C. S., Barrett, M. P., Lawson, D., Quail, M. A., Harris, D., Bowman, S., Phillips, R. S. & Turner, C. M. 2001 Gene discovery in *Plasmodium chabaudi* by genome survey sequencing. *Mol. Biochem. Parasitol.* **113**, 251–260.

Kyes, S. A., Rowe, J. A., Kriek, N. & Newbold, C. I. 1999 Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with *Plasmodium falciparum*. *Proc. Natl Acad. Sci. USA* **96**, 9333–9338.

Leech, J. H., Barnwell, J. W., Miller, L. H. & Howard, R. J. 1984 Identification of a strain-specific malarial antigen exposed on the surface of *Plasmodium falciparum*-infected erythrocytes. *J. Exp. Med.* **159**, 1567–1575.

McLean, S. A., Pearson, C. D. & Phillips, R. S. 1982 *Plasmodium chabaudi*: antigenic variation during recrudescent parasitaemias in mice. *Exp. Parasitol.* **54**, 296–302.

Morgenstern, B. 1999 DIALIGN 2: improvement of the segment-to-segment approach to multiple sequence alignment. *Bioinformatics* **15**, 211–218.

Morgenstern, B., Dress, A. & Werner, T. 1996 Multiple DNA and protein sequence alignment based on segment-to-segment comparison. *Proc. Natl Acad. Sci. USA* **93**, 12 098–12 103.

Pace, T., Ponzi, M., Dore, E. & Frontali, C. 1987 Telomeric motifs are present in a highly repetitive element in the *Plasmodium berghei* genome. *Mol. Biochem. Parasitol.* **24**, 193–202.

Page, R. D. 1996 TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357–358.

Pearson, W. R. 2000 Flexible sequence similarity searching with the FASTA3 program package. *Meth. Mol. Biol.* **132**, 185–219.

Phillips, R. S. 2001 Current status of malaria and potential for control. *Clin. Microbiol. Rev.* **14**, 208–226.

Phillips, R. S., Brannan, L. R., Balmer, P. & Neuville, P. 1997 Antigenic variation during malaria infection—the contribution from the murine parasite *Plasmodium chabaudi*. *Parasite Immunol.* **19**, 427–434.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A. & Barrell, B. 2000 Artemis: sequence visualization and annotation. *Bioinformatics* **16**, 944–945.

Sambrook, J., Fritsch, E. F. & Maniatis, T. 1989 *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Saul, A. 1999 The role of variant surface antigens on malaria-infected red blood cells. *Parasitol. Today* **15**, 455–457.

Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J. & Lanzer, M. 1998 Antigenic variation in malaria: *in situ* switching, relaxed and mutually exclusive transcription of *var* genes during intra-erythrocytic development in *Plasmodium falciparum*. *EMBO J.* **17**, 5418–5426.

Siddall, M. E. & Barta, J. R. 1992 Phylogeny of *Plasmodium* species: estimation and inference. *J. Parasitol.* **78**, 567–568.

Smith, J. D., Chitnis, C. E., Craig, A. G., Roberts, D. J., Hudson-Taylor, D. E., Peterson, D. S., Pinches, R., Newbold, C. I. & Miller, L. H. 1995 Switches in expression of *Plasmodium falciparum var* genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101–110.

Strimmer, K. & von Haeseler, A. 1996 Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**, 964–969.

Strimmer, K. & von Haeseler, A. 1997 Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proc. Natl Acad. Sci. USA* **94**, 6815–6819.

Su, X. Z., Heatwole, V. M., Wertheimer, S. P., Guinet, F., Herrfeldt, J. A., Peterson, D. S., Ravetch, J. A. & Wellems, T. E. 1995 The large diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* **82**, 89–100.

Voller, A. & Rossan, R. N. 1969 Immunological studies with simian malarias. I. Antigenic variants of *Plasmodium cynomolgi bastianellii*. *Trans. R. Soc. Trop. Med. Hyg.* **63**, 46–56.