Molecular Microbiology (2003) doi:10.1046/j.1365-2958.2003.03814.x

Evidence for the importance of genetic structuring to the structural and functional specialization of the *Plasmodium falciparum var* gene family

Susan M. Kraemer[1,2] and Joseph D. Smith[1,2]*
[1]*Seattle Biomedical Research Institute, 4 Nickerson Street, Seattle, WA 98109, USA.*
[2]*Department of Pathobiology, University of Washington, Seattle, WA 98195, USA.*

Summary

The *var* gene family encodes *Plasmodium falciparum* erythrocyte membrane 1 (PfEMP1) proteins that act as virulence factors responsible for both antigenic variation and cytoadherence of infected erythrocytes. These proteins orchestrate infected erythrocyte sequestration from blood circulation and contribute to adhesion-based complications of *P. falciparum* malaria infections. For this study, we analysed the genetic organization and strain structure of *var* genes and present evidence for three separately evolving groups that have, in part, functionally diverged and differ between subtelomeric and central chromosomal locations. Our analyses suggest that a recombination hierarchy limits reassortment between groups and may explain why some *var* genes are unusually conserved between parasite strains. This recombination hierarchy, coupled with binding and immune selection, shapes the variant antigen repertoire and has structural, functional and evolutionary consequences for the PfEMP1 protein family that are directly relevant to malaria pathogenesis.

Introduction

Because of their clonally variant and cytoadherent properties, members of the *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) family play a central role in malaria pathogenesis (Miller *et al*., 2002). The PfEMP1 proteins, encoded by *var* genes (Baruch *et al*., 1995; Smith *et al*., 1995; Su *et al*., 1995), are expressed at the infected erythrocyte surface and allow parasitized erythrocytes to sequester from blood circulation to avoid spleen-dependent killing mechanisms. Sequestration is a *P. falciparum* virulence determinant. Although most infections do not cause severe disease, disease severity is increased when infected erythrocytes accumulate in vital organs, such as the brain or the placenta (Miller *et al*., 2002). The surface location of PfEMP1 proteins exposes them to intense antibody responses (Bull and Marsh, 2002). By transcriptional switching between *var* genes, parasites evade immunity and extend opportunities to be transmitted to mosquitoes.

Var genes have a two-exon structure with an ≈170 bp to 1.2 kb intron. The first exon is highly polymorphic, encodes the extracellular binding region and transmembrane domain and varies in size between 3.5 and 9.0 kb (Gardner *et al*., 2002). The second exon is more conserved, encodes an acidic cytoplasmic tail and is between 1.0 and 1.5 kb. There is an enormous diversity of *var* genes in the parasite population. Whereas each parasite genotype encodes approximately 60 different *var* loci (Thompson *et al*., 1997; Gardner *et al*., 2002), gene repertoires differ extensively between parasite strains (Peterson *et al*., 1995; Kyes *et al*., 1997; Ward *et al*., 1999 Fowler *et al*., 2002). However, a small number of exceptional *var* genes have been described that are unusually conserved between parasite genotypes (Fried and Duffy, 2002; Rowe *et al*., 2002; Salanti *et al*., 2002; 2003; Winter *et al*., 2003) Strain-transcendent *var* may have important roles in malaria pathogenesis, as both serological and epidemiological investigations suggest that a limited subset of particularly virulent proteins is responsible for disease in pregnant mothers and severe childhood malaria (Gupta *et al*., 1994; 1999; Fried and Duffy, 1996; Fried *et al*., 1998; Bull *et al*., 1999; Nielsen *et al*., 2002).

It has been suggested that *var* loci could become relatively fixed in the parasite population if a gene has an unusual flanking sequence or gene orientation that limits reassortment with other *var* family members. This concept has important implications for variant antigen diversification and parasite evasion of host immunity, and may be directly relevant to malaria pathogenesis if caused by 'conserved' variants, but has not been investigated directly. Besides the 3D7 parasite strain that was sequenced completely for the Malaria Genome Project (Gardner *et al*., 2002), there is only sketchy information about variant antigen repertoires from different parasite strains. In general, a greater understanding of the *var*

Accepted 8 September, 2003. *For correspondence. E-mail joe.smith@sbri.org; Tel. (+1) 206 284 8846, ext. 384; Fax (+1) 206 284 0313.

gene family conservation and diversity, the factors regulating variant antigen gene diversification and the expression of particular *var* during disease could provide critical insights into malaria pathogenesis and aid vaccine development against PfEMP1 proteins.

PfEMP1-binding regions contain multiple receptor-like domains called Duffy binding-like (DBL) domains and cysteine-rich interdomain regions (CIDR). These adhesion domains can be grouped by sequence similarity (Smith *et al*., 2000) into seven types of DBL domains (α, α_1, β, γ, δ, ε and x) and four types of CIDR domains (α, α_1, β and γ). Adhesion domain sequence criteria have been used as a basis for dissecting PfEMP1 protein forms and may relate to binding function. For example, the 59 PfEMP1 proteins encoded by the 3D7 parasite strain have been arbitrarily assigned one of 16 different type designations based upon their domain organization (Gardner *et al*., 2002). In a genome-wide binding analysis, the majority of the 3D7 PfEMP1 proteins encoded CD36 binding capacity in the semi-conserved protein head structure, whereas a subset did not (Robinson *et al*., 2003). CD36 binding head structures were designated type 1 head structures and contained a DBL-α domain and a CD36-binding CIDR-α type of domain. In contrast, type 2 head structures had a DBL-α or DBL-α_1 subtype domain plus a non-CD36 binding CIDR-α_1 or CIDR-γ domain, whereas atypical head structures lacked a CIDR domain altogether and were predicted not to bind CD36 (Fig. 1). Thus, PfEMP1 proteins are categorized by type of domain organization and distinguished further by type of protein head structure (the first two adhesion domains), the latter of which has been related to CD36 binding function. The functional divergence in CIDR domain binding to CD36 is clinically significant in that non-CD36 binding variants are responsible for pregnancy malaria (Fried and Duffy, 1996; Beeson *et al*., 2000; Flick *et al*., 2001).

For this study, we analysed how CD36 binding and non-binding *var* evolve in the highly recombinogenic *var* gene family. We describe a unique subtelomeric *var* gene architecture that is entirely composed of PfEMP1 proteins that are not predicted to bind CD36. The implications of these findings for *var* gene evolution and investigation of adhesion-based pathogenesis are discussed.

Results

A distinctive subtelomeric var *gene architecture is associated with PfEMP1 proteins predicted not to bind CD36*

To study the relationship between PfEMP1 protein function and gene context, results from a genome-wide CIDR–CD36 binding analysis (Robinson *et al*., 2003) were related to the 3D7 parasite genomic map. Interestingly, although *var* genes encoding CD36 binding capacity were present in central and subtelomeric chromosomal regions, nearly all PfEMP1 proteins predicted not to bind CD36 localized to the telomere ends and were transcribed towards the telomere (Fig. 1A).

From the genome analysis and other recent studies, the *var* upstream non-coding sequences have been shown to cluster into three distinct sequence families, UpsA, UpsB and UpsC. These three sequence groups specifically differ between central and subtelomeric *var* genes and have distinct chromosomal orientations (Voss *et al*., 2000; Gardner *et al*., 2002; Vazquez-Macias *et al*., 2002). The non-CD36 binding group described above associated with UpsA.

To understand better how the genomic compartmentalization of CD36 binding and non-CD36 binding PfEMP1 adhesion groups occurred, the genetic sequence contexts of 3D7 *var* genes were studied here. As *var* promoter elements exist at least 2000 bp upstream of the coding sequence (Voss *et al*., 2003), 2000 bp of 5′ and 3′ flanking sequence were collected from 3D7 *var* genes for phylogenetic and bootstrap analysis. This analysis confirmed that the majority of 3D7 *var* upstream elements grouped into one of three major clades named UpsA, UpsB and UpsC (Fig. 1B), with a few exceptions described below. As reported earlier (Gardner *et al*., 2002), all UpsA and most UpsB sequences were associated with subtelomeric *var* genes. UpsC sequences were associated with central *var* genes (Fig. 1).

Within the UpsA clade defined by Gardner *et al*. (2002), most sequences were closely related to each other with the exceptions of PFL0030c and the PFE1640w pseudovar (also referred to as 3D7 chr5var). The latter two sequences, although they associated with the UpsA clade, were statistically distinct from other UpsA sequences by bootstrap analysis (Fig. 1B; data not shown). Of the two sequences, the PFL0030c was more divergent whereas PFE1640w had a limited amount of sequence identity to the UpsA consensus that was primarily confined to the 100–200 bp upstream of the predicted protein start (data not shown; *Supplementary material*, Fig. S1). In addition, none of the other UpsA sequences contained the two different repetitive motifs of 150 bp and 60 bp characteristic of PFE1640w (Vazquez-Macias *et al*., 2002). Because of their distinctiveness, the PFL0030c and PFE1640w upstream sequences were designated UpsD and UpsE (Fig. 1B).

The UpsB clade also contained some sequences that did not group tightly with the main cluster (Fig. 1B). These divergent sequences had relatively good nucleotide identity to the UpsB consensus residues in the first 600 bp upstream (data not shown). For this reason, there was no firm support to assign them to a distinct upstream sequence type. Thus, at least five different types of *var*

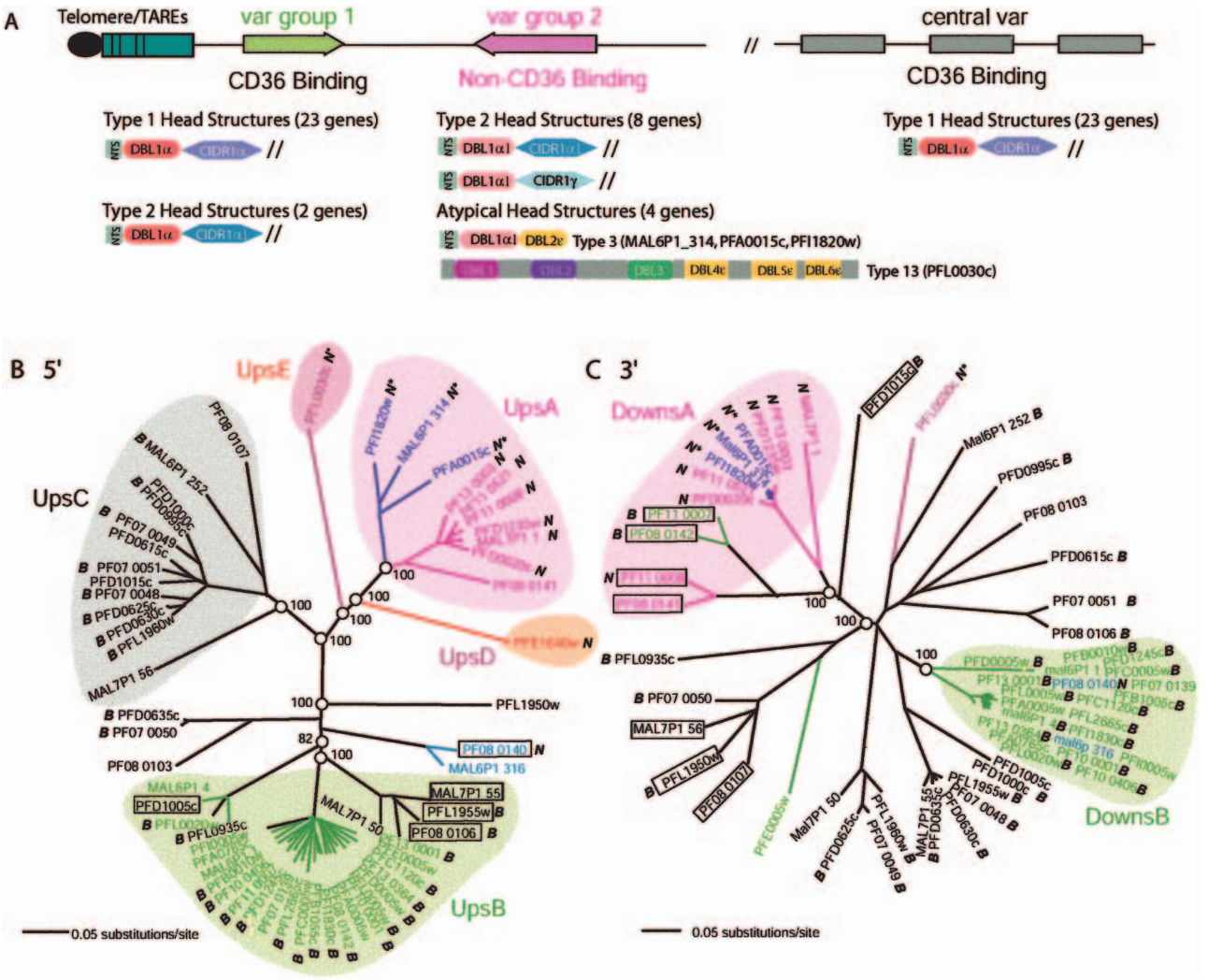

Fig. 1. Structural and functional specialization of *var* architectural groups.
A. Chromosome schematic of the two subtelomeric and central *var* groups. For simplicity, the tandemly arranged subtelomeric *var* genes are depicted as though there were no intervening genes. However, it is not uncommon for one or two *rif* genes to be between these *var* genes (Gardner *et al*., 2002). Nearly all group 1 and central *var* genes would be predicted to bind CD36, whereas group 2 *var* genes are predicted not to bind CD36. Listed are the number of genes in each group and type of protein head structure. Type 1 head structures contain a DBL1-α domain and a CD36 binding CIDR1-α domain. Type 2 head structures contain either a DBL1-α or a DBL1-α_1 subtype domain and a non-CD36 binding CIDR1-α_1 subtype or CIDR-γ domain. Atypical head structures lack a CIDR domain. Phylogenetic analysis of 2000 bp of 5′ (B) and 3′ (C) flanking sequences from 3D7 var genes (Gardner *et al*., 2002). Circles indicate distinct clades with the corresponding bootstrap value. Genes are coloured according to chromosomal location and types of 5′ flanking sequences. Green and blue sequences are located in the subtelomeric region and oriented towards the centromere (subtelomeric *var* group 1), black sequences are in central chromosomal locations, and orange, burgundy, purple and pink sequences are located in the subtelomeric regions and oriented towards the telomere (subtelomeric *var* group 2). Sequences are shaded according to grouping. Five types of *var* upstream sequences were designated UpsA to UpsE and two types of downstream sequences DownsA and Downs B. For the boxed sequences, there were regions of overlap between the 2000 bp analysed and the 5′, 3′ or coding sequence of adjacent *var* genes. The distribution of sequences on the tree did not change when the overlapping regions were excluded from the phylogenetic comparison (data not shown). N, CIDR1 domain does not bind CD36; B, CIDR1 domain binds CD36; N*, protein head structures do not contain CIDR domains and are not predicted to bind CD36 (Robinson *et al*., 2003).

upstream sequences (UpsA to UpsE) could be distinguished (Fig. 1B).

By comparison, the downstream regions of the two subtelomeric *var* genes grouped into two distinct families, DownsA and the extremely conserved DownsB (Fig. 1C). The downstream region of central *var* genes, however, was not conserved (Fig. 1C; Gardner *et al*., 2002). In some notable exceptions, the PFE1640w downstream was not analysed because the gene is truncated in the first exon and ends in a telomere repeat. The PFL0030c sequence that has a distinct upstream sequence (UpsE) also groups separately from the DownsA cluster.

Based on these analyses, three broad categories of *var* gene architectures were defined as subtelomeric *var*

groups 1 (flanked by UpsB, DownsB), subtelomeric *var* group 2 (flanked by UpsA, UpsD or UpsE and DownsA) and central *var* genes (flanked by UpsC) (Fig. 1).

Of the 59 *var* genes present in the 3D7 genome, 36 mapped to subtelomeric regions and 23 to central regions (Gardner *et al*., 2002). Twenty-five of the subtelomeric *var* genes were transcribed towards the centromere (*var* group 1), and 11 were transcribed towards the telomere (*var* group 2) (Fig. 1A). Altogether, 22 of the 28 telomere ends had a *var* group 1 gene as the telomere-proximal *var* gene (Gardner *et al*., 2002). A subset of these, eight chromosome ends, had the *var* 1/*var* 2 group arrangement illustrated in Fig. 1A. For simplicity, this tandem arrangement is depicted as though there were no intervening genes. However, it is not uncommon for one or two *rif* genes to be between the two *var* genes (Gardner *et al*., 2002). A ninth chromosome end, chromosome 12, had this arrangement except that there were two *var* group 1 genes next to the *var* group 2 gene PFL0030c (Gardner *et al*., 2002). The remaining *var* group 2 genes, MAL7P1.1 and PF11-0521, were not flanked by a *var* group 1 gene. Because telomere-proximal genes are frequently deleted when *P. falciparum* parasites are adapted to *in vitro* cultivation, it is not clear whether single *var* group 2 genes are natural arrangements or if a *var* group 1 gene was deleted from these chromosome ends.

The subtelomeric *var* group 2 was distinguished from other *var* groups by encoding proteins that were predicted not to bind CD36. All the *var* group 2 proteins had either a type 2 protein head structure (seven genes plus the PFE1640w pseudovar) or atypical head structures (four genes) that lacked a CIDR domain (Fig. 1A). Of the four PfEMP1 proteins that lack a CIDR domain, three had a binding region domain organization referred to as 'type 3', and the fourth (PFL0030c) had a domain organization referred to as 'type 13' (Fig. 1A) (Gardner *et al*., 2002). For convenience, the three 3D7 *var* genes (PFA0015c, MAL6p1.314 and PFI1820w) and others related to them are referred to collectively as 'type 3-like' *var* in this paper.

Remarkably, in the whole-genome analysis (Robinson *et al*., 2003), only two PfEMP1 proteins that were not predicted to bind CD36 did not localize with the group 2 *vars*. The two exceptions, PF08-0140 and MAL6P1.316 (blue sequences in Fig. 1B), had a slightly different type 2 head structure from *var* group 2 proteins. Whereas they had CIDR1-α_1 subtype domains that did not bind CD36, their DBL1 domain was the 'α type' rather than the 'DBL1-α_1' subtype (Fig. 1A) (Robinson *et al*., 2003). Thus, proteins in the subtelomeric *var* group 2 were structurally and functionally distinct from central and subtelomeric *var* group 1 and may be selected for different adherence properties.

The subtelomeric var *group 2 contains several unusually conserved, 'strain-transcendent'* var *genes*

To assess the extent of *var* gene overlap between different parasite strains, most previous investigations have compared small DBL1 domain fragments amplified by degenerate primers with semi-conserved motifs. These studies have documented an enormous diversity of *var* gene sequences (Taylor *et al*., 2000a; Fowler *et al*., 2002; Tami *et al*., 2003) and a limited number of unusually conserved *var* fragments. To date, two 3D7 *var* group 2 loci, the PFE1640w and the PFL0030c (or *var2csa*), have been shown to be unusually conserved between parasite strains (Salanti *et al*., 2003; Winter *et al*., 2003). The FCR3-CSA *var* gene (or *var1csa*), which is chimeric to the PFE1640w pseudovar and may share a common ancestral origin, is a third highly conserved *var* group 2 locus (Gardner *et al*., 2002; Rowe *et al*., 2002; Salanti *et al*., 2002; Vazquez-Macias *et al*., 2002). To test the idea that the unique gene orientation of the subtelomeric *var* group 2 would minimize recombination and cause other genes in this group to be more conserved between parasite strains, primers based upon the 3D7 UpsA, UpsE and DownsA sequences were used for amplification.

For this comparison, three parasite strains (IT4/25/5, Dd2 and MC) collected at different times and from distant geographical locations were investigated, and several complete *var* genes totalling nearly 100 kb of new *var* sequence were cloned and sequenced. To facilitate this analysis, long polyymerase chain reactions (PCRs) were optimized. The longest *var* product that was cloned successfully was ≈10 kb, but products of 6 kb and greater were routinely amplified and cloned. Larger *P. falciparum* DNA fragments are notorious for undergoing recombination in bacteria. Recombination did not occur for larger *var* gene fragments as confirmed by multiple, independent shorter PCRs repeated on genomic DNA for each newly cloned *var* (data not shown). Overall, three different types of subtelomeric *var* group 2 gene products were studied; (i) the 'type 3' *var*; (ii) the PFL0030c-like *var* genes; and (iii) *var* genes encoding type 2 head structures (Fig. 1A).

To clone and sequence 'type 3' *var* genes from the Malayan Camp and It4/25/5 strains, type 3-tailored upstream and downstream primers were used to amplify ≈6 kb PCR products that were sequenced rapidly using oligonucleotides based upon the 3D7 'type 3' *var* genes. A comparison of the genes amplified with those from the 3D7 strain showed that at least 800 bp of 5′ flanking sequence and 400 bp of downstream sequence were highly conserved (Fig. 2A, *Supplementary material*, Fig. S1). In addition, each of the genes had a conserved and unusually short intron, ≈150–170 bp (Fig. 2A; data not shown). As the intron region has been implicated in *var* gene silencing (Deitsch *et al*., 2001), possible implica-

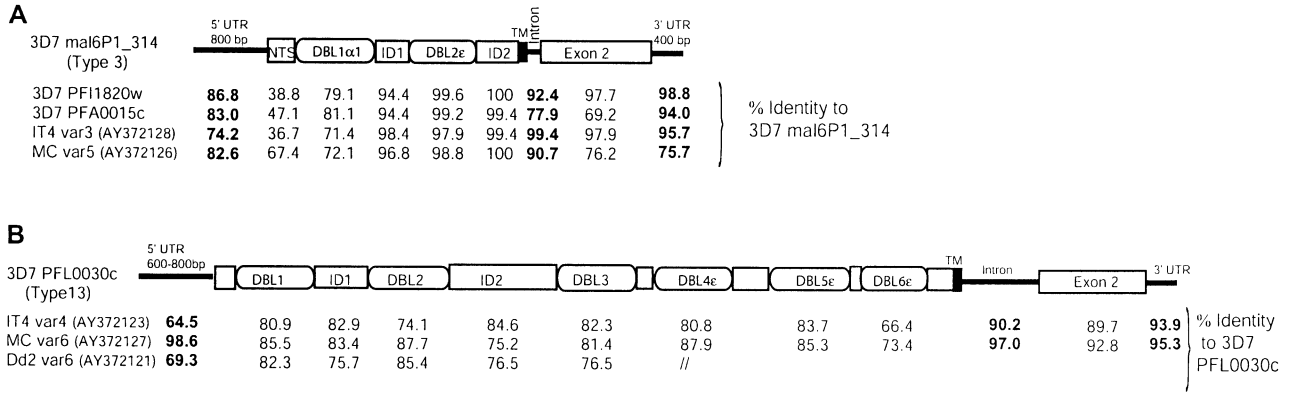

Fig. 2. Comparison of 'type 3' and PFL0030c-like *var* genes from different parasite genotypes. Percentage nucleotide identity (bold type) or amino acid identity (normal type) is shown for (A) type 3 *var* and (B) PFL0030c-like *var*. Accession numbers of genes amplified for this paper are listed in parentheses.

tions on gene regulation or switching exist. The degree of coding region similarity has not been reported for 'type 3' *var*. A comparison of the five sequences (Fig. 2A) showed a remarkable degree of sequence conservation both within and between parasite strains. These proteins were nearly identical over much of their predicted binding region and only diverged at the amino-terminus and the intracytoplasmic region (Fig. 2A).

As shown recently (Salanti *et al*., 2003), PFL0030c-like *var* genes are also highly conserved between parasite strains. In addition, the 3D7 strain also contains a second, PFL0030c-like pseudogene at the end of chromosome 13 (MAL13P1-354) that is transcriptionally orientated towards the telomere and has a nearly identical upstream sequence to PFL0030c. To clone and sequence PFL0030c-like *var* genes from other parasite strains, these ≈10 kb genes were amplified in two separate PCRs using gene-specific primers positioned at approximately −800 and +500 bp relative to the coding sequence and internal primers in the third DBL domain (Fig. 2). The initial PCR products overlapped at the primers. To confirm their connection, a third PCR product from DBL domains 2–4, which straddled these PCR products, was cloned and sequenced. For Dd2, only the 5′ half of the gene was cloned, but complete genes were cloned and sequenced from Malayan Camp and It4/25/5. As was done with the 'type 3' *var*, oligos based upon the 3D7 *var* sequence were used to sequence the new genes.

Although it is known that the coding sequences of PFL0030c-like *var* are conserved, this analysis demonstrated significant conservation of the gene flanking sequence (Fig. 2B). The upstream regions were nearly identical except that the It4/25/5 and Dd2 strains had an insertion of ≈250 bp, accounting for the lower sequence identity values for It4 var4 and Dd2 var6 (Fig. 2B; data not shown). PFL0030c-like *var* have a truncated N-terminal segment at the beginning of the protein and lack a conventional head structure because they do not have a CIDR domain (Fig. 1A; data not shown). The first three DBL domains do not cluster with other DBL types but form unique clades of their own (Fig. 4B). In addition, the inter-domain regions after DBL domains 1 and 2 were not present in other *var* genes, but both had several conserved cysteine residues (data not shown). These distinctive features suggest that the PFL0030c-like *var* do not recombine frequently with other *var* genes.

Lastly, to test whether UpsA sequences were associated with type 2 head structures in non-3D7 parasite genotypes, an UpsA primer was paired with a degenerate DBL-β primer. In large PfEMP1 proteins, DBL-βc2 tandem domains are the most common domains after either type 1 or type 2 head structures (Smith *et al*., 2000; Gardner *et al*., 2002). Using the UpsA primer, each of five randomly cloned *var* genes from Malayan Camp, Dd2 and It4/25/5 had DBL1 and CIDR1 domains characteristic of type 2 head structures (Fig. 4B and C). By comparison, when an UpsC primer was paired with the same degenerate DBL-β primer, only PfEMP1 proteins with type 1 head structures were amplified from IT4/25/5 genomic DNA (Kraemer *et al*., 2003; data not shown).

The downstream gene regions of the new genes with type 2 head structures were cloned using gene-specific DBL-β primers paired with a DownsA primer (Table 1). These *var* were more diverse than the 'type 3 *var*' and the PFL0030c-like homologues (Figs 2 and 3). To aid their sequencing, 'universal' DBL type-specific primers developed for PCR amplification (Kraemer *et al*., 2003) were adapted to sequencing reactions. When adhesion domains were not repeated in the same cloned product, several contigs were produced that could be rapidly closed by 'walking'. The new proteins included two examples of a DBL-γc2 domain pairing, where previously the c2 domain has only been associated with the DBL-β type (Smith *et al*., 2000; Gardner *et al*., 2002).

Table 1. Oligos used in cloning *var* genes from different genotypes.

Oligo name	Sequence	Fragment cloned
PFL0030c-like *var*		
PFL0030c 5p 5.1	GTCTGTGAATGCAATGACAG	Gene flanking
PFL0030c DBL3 3.1	CGTGATATAATTGCTGTACC	5′ to DBL3
PFL0030c DBL3 5.1	GGTACAGCAATTATATCACG	DBL3 to gene
PFL0030c 3p 3.1	GTAAGATGTAACAAGATATTAC	flanking 3′
PFL0030c DBL2 5.2	GTAAGTCGTGTAAGGAAAGTG	DBL2 to DBL4
PFL0030c DBL4 3.2	GCAACTATTTGTAATGTTTCC	
Type 3 *var*		
Type3 5p 5.1	GGATAAGTGATRACATAATRT	Gene flanking 5′
Var gr2 3p 3.1 (DownsA)	CAAATAATCAMATGTGTCAAAYAR	to gene flanking 3′
Var containing type 2 head structures		
Var gr2 5p 5.1 (UpsA)	ATKTATTAYATTTGTTGTAGGTGA	Gene flanking 5′ to DBLβ
degenDBLβ 3.2	AATCKTTGDGG RATRTARTC	

This analysis demonstrated that subtelomeric group 2-like *vars* were broadly distributed between parasite genotypes and had related features. Some group 2 *vars*, such as the PFL0030c-like *var* and 'type 3' *var*, were extremely conserved (Fig. 2). Others were less strictly conserved between parasite genotypes (Fig. 3) but had type 2 head structures rather than type 1 head structures.

Evidence for three separately evolving var *groups across parasite genotypes*

The unusual interstrain conservation of the *var* group 2-like genes is consistent with the hypothesis that gene location, orientation and non-coding flanking sequence limit *var* recombinational possibilities. To test further the idea that a recombination hierarchy exists between *var* groups, additional sequence comparisons were performed. Initially, the 10 new *var* genes amplified for this study were compared with *var* group 2 genes from the 3D7 strain using phylogenetic criteria. As mentioned above, the 'type 3' *var* and PFL0030c-like *var* were highly related between strains (Figs 2 and 4). However, even the more divergent *var* genes encoding type 2 head structures were similar between parasite strains, as evidenced by two approaches. First, for the domains comprising the protein's head structure and cytoplasmic regions, each of the subtelomeric *var* group 2-like sequences tended to group together in the NTS, DBL, CIDR and exon 2

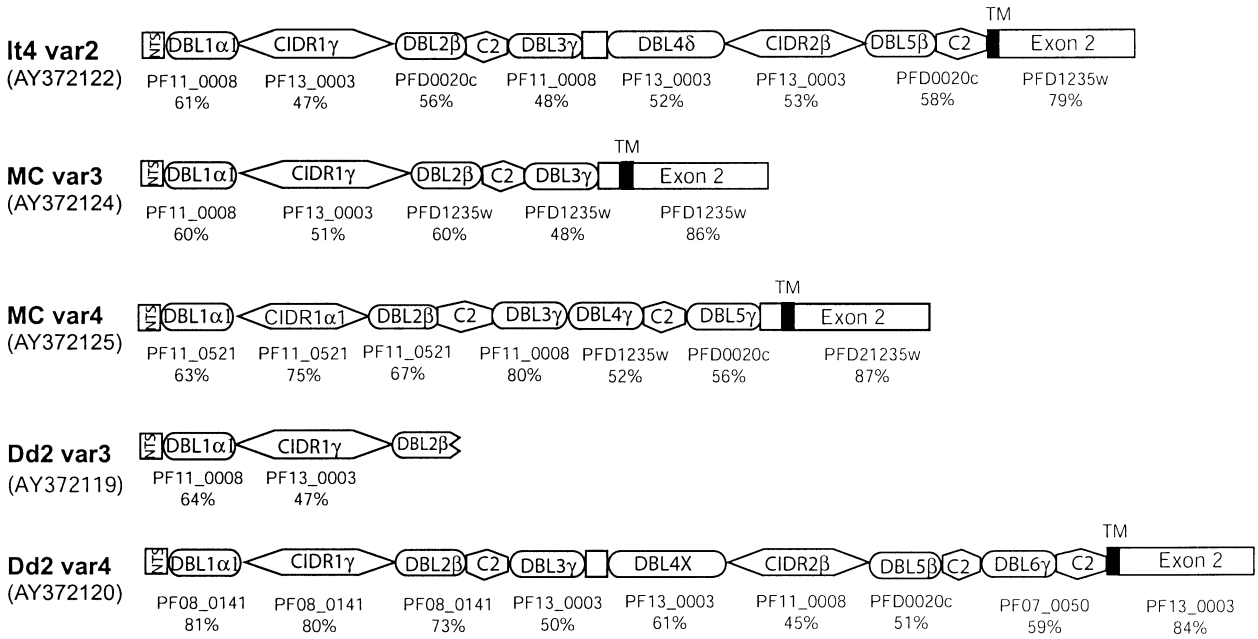

Fig. 3. Comparison of *var* group 2-like proteins that contain type 2 head structures. PfEMP1 sequences amplified from different parasite genotypes are compared with PfEMP1 proteins from the 3D7 parasite strain. Percentage amino acid identity to the 3D7 sequence with greatest BLAST identity is listed. Accession numbers are listed in parentheses.

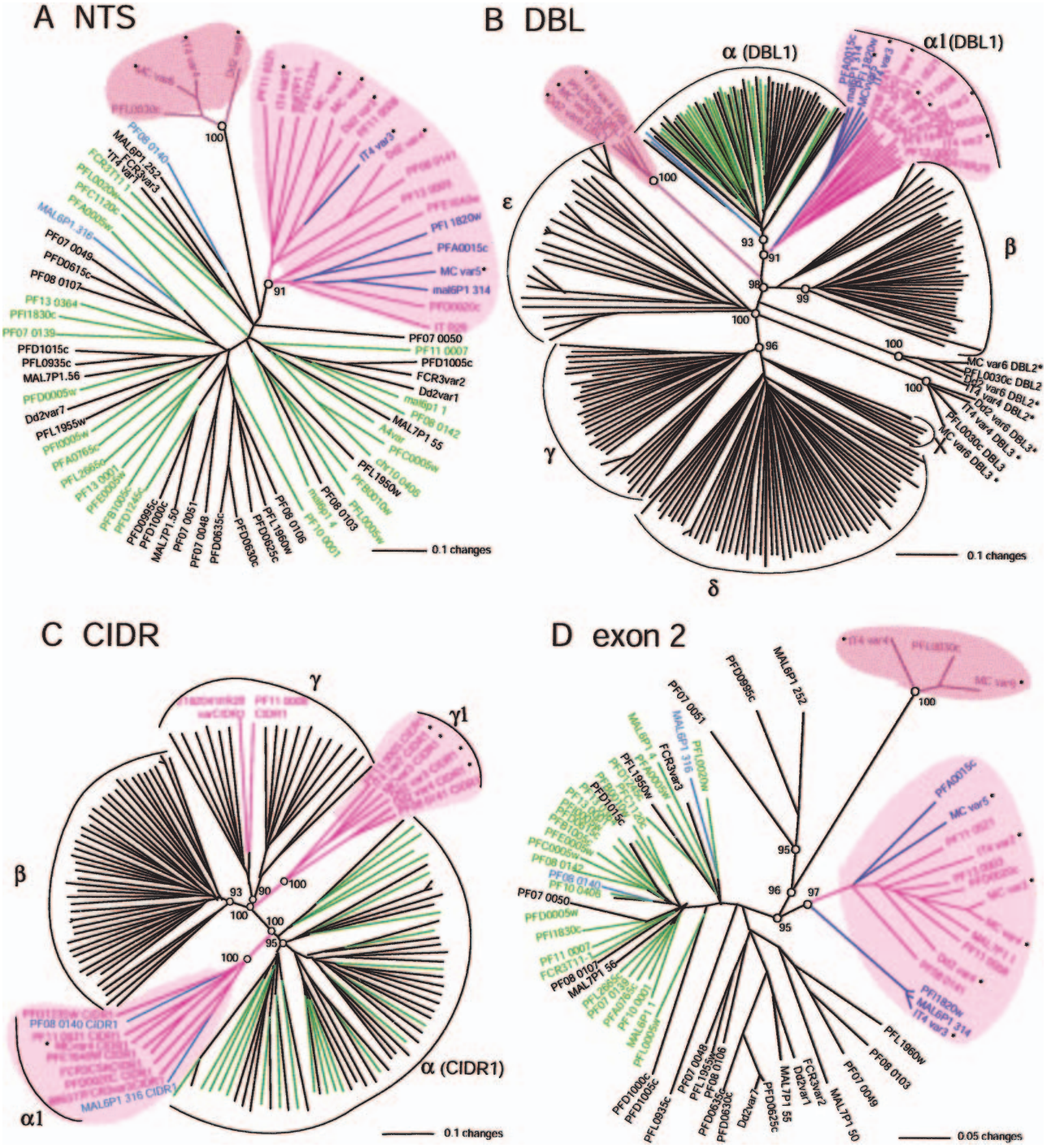

Fig. 4. Sequence comparison of the subtelomeric *var* group 2 genes with other *var* genes. Phylogenetic trees were generated based upon domains in PfEMP1 proteins. The sequences compared included the 59 3D7 PfEMP1 proteins (Gardner *et al.*, 2002), eight PfEMP1 proteins from different parasite genotypes that had previously been mapped to chromosome locations (FCR3T11-1, A4var, Dd2var1, FCR3var2, FCR3var3, Dd2var7a, Dd2var7b, ItR29) and the 10 new *var* genes amplified for this study (indicated by asterisks). Amino acid sequences are compared. Genes and gene groups are coloured as in Fig. 1. For simplicity, only the DBL1 and CIDR1 head structure domains were coloured in the DBL and CIDR trees. In addition, in the DBL and CIDR trees, most of the gene names were removed, except for the subtelomeric *var* group 2, to make the figure legible.

(cytoplasmic region) trees (Fig. 4A–D). In addition, nearly all the predicted DBL and CIDR adhesive domains in the new genes were most similar to 3D7 subtelomeric *var* group 2 genes by BLAST analysis (Fig. 3). The only exception was the Dd2 var4 DBL6γc2.

The similarity between *var* group 2 genes also included the non-coding flanking sequence. When the upstream, downstream and intron sequences were compared across parasite strains, groups of sequences that were defined based upon the 3D7 genome remained intact (*Supplementary material*, Fig. S1; data not shown). As *var* 5′ regions have promoter activity and co-operate with the intron to silence *var* genes (Deitsch *et al*., 1999; 2001; Calderwood *et al*., 2003; Voss *et al*., 2003), these unique features of *var* group 2 sequences could have regulatory implications.

To extend this analysis to non-*var* group 2 genes, several previously mapped *var* genes from non-3D7 parasite genotypes were added to the phylogenetic analyses. Based on chromosomal location, FCR3T11-1 and A4var are members of subtelomeric *var* group 1 (Hernandez-Rivas *et al*., 1997; Horrocks *et al*., 2002); Dd2var1, FCR3var2, FCR3var3, Dd2var7a and Dd2var7b are central *var* (Su *et al*., 1995; Deitsch *et al*., 1999); and ItR29 is a member of subtelomeric *var* group 2 (Horrocks *et al*., 2002). For each of these *var*, the non-coding flanking sequences grouped according to prediction (*Supplementary material*, Fig. S1). Moreover, the coding sequence of all central and subtelomeric *var* group 1 genes had type 1 head structures and grouped together in the NTS and exon 2 analyses, whereas the R29 *var* had a type 2 head structure and grouped with other *var* group 2 genes in the NTS tree (Fig. 4; data not shown). Therefore, this analysis showed that central and subtelomeric *var* groups are surrounded by specific and distinct flanking sequences and have diverged in characteristic patterns in their protein coding sequence. These data suggest that the different groups may be recombining and evolving separately from each other.

Discussion

The availability of a complete *P. falciparum* genomic sequence creates dramatic new opportunities to investigate the parasite's biology. For this study, we analysed the gene structure and chromosomal organization of *var* genes and used this information to gain insight into how the gene family evolves and causes disease. In common with other organisms, *P. falciparum* variant genes are concentrated at the telomere ends (Fischer *et al*., 1997; Thompson *et al*., 1997; Gardner *et al*., 2002). However, unlike many other variant gene families that function only as clonally variant targets of immunity, *var* genes also encode cytoadherent activity. Thus, the *var* gene repertoire is shaped by selective forces acting to both maintain binding capability and escape immunity (Roberts *et al*., 1993). At the genomic level, these varied pressures have been manifested in an apparent genetic compartmentalization of PfEMP1 binding phenotypes to particular *var* gene sequence contexts.

In the 3D7 genome, at least three broadly distinct *var* gene groups were defined based upon chromosomal location, gene orientation and 5′ and 3′ flanking sequences (Fig. 1). Subtelomeric *var* group 1 consists of telomeric *var* genes that are transcribed towards the centromere and are flanked by UpsB and DownsB; subtelomeric *var* group 2 contains telomeric *var* genes that are oriented towards the telomere and are flanked by either UpsA, UpsD or UpsE and DownsA; and central *var* are flanked by UpsC and a non-conserved 3′ region (Fig. 1). These features appear to establish a recombination hierarchy that limits reassortment between *var* groups and may create opportunities for the parasite to evolve specialized functions and activities relative to the different groups. A striking example is the subtelomeric *var* group 2 that contains several unusually conserved *var* genes and consists of non-CD36 binding variants. Although CD36 is a key microvasculature receptor used by most parasite isolates (Newbold *et al*., 1997), non-CD36-binding PfEMP1 proteins have important roles in placental sequestration (Fried and Duffy, 1996; Beeson *et al*., 2000; Flick *et al*., 2001) and possibly other vascular sites, such as brain endothelium, where CD36 is poorly or not expressed.

It is surprising that a subtelomeric *var* group would contain so many broadly conserved loci, as this chromosomal region is highly recombinogenic (Lanzer *et al*., 1995; Freitas-Junior *et al*., 2000; Taylor *et al*., 2000b; Figueiredo *et al*., 2002). Recently, the telomere ends of *P. falciparum* chromosomes were shown to cluster together in the nuclear periphery, suggesting a potential mechanism for facilitating *var* gene recombination (Freitas-Junior *et al*., 2000; Figueiredo *et al*., 2002). Our comparative analysis of the subtelomeric *var* group 2 builds on this analysis and suggests a model in which gene orientation, location and flanking sequences are crucial to the strain transcendence of some *var* genes and favour the preferential reassortment of genes within architectural groups (Fig. 5). Conserved subtelomeric *var* group 2 loci, such as PFE1640w (Winter *et al*., 2003) and PFL0030c, have highly distinctive gene flanking sequences (Fig. 1B and C), which may reinforce their withdrawal from general family recombination. This mechanism would promote 'self–self' recombination and interstrain gene conservation. Other group 2 *var* genes encoding type 2 head structures were more diverse between strains. This diversity may occur because there are more of these to recombine with (Fig. 1) and/or the individual genes are under different functional and immune selective pressures from the

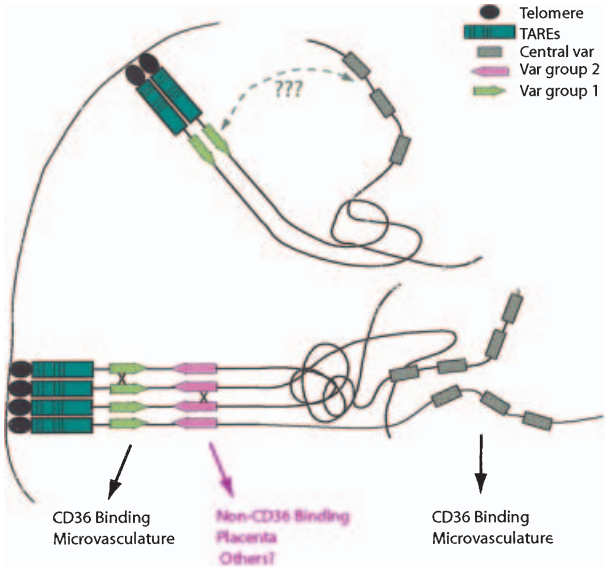

Fig. 5. Idealized model for *var* gene recombination. The telomere ends of *P. falciparum* chromosomes are shown to cluster at the nuclear periphery as has been described by fluorescent *in situ* hybridization analysis (Freitas-Junior *et al*., 2000). *Var* gene recombination (indicated by Xs) is hypothesized to occur preferentially within a *var* architectural group because of gene orientation, location and flanking sequence. Proteins in the subtelomeric *var* group 2 are not predicted to bind CD36 and could be selected for distinct patterns of *in vivo* sequestration.

highly conserved loci. Recently, it was suggested that the mechanism and kinetics of gene silencing are distinct between chromosome central and subtelomeric *var* group 1 promoter elements (Voss *et al*., 2003). As *var* group 2 genes have highly distinctive upstream and intron sequences, it is possible that there are regulatory differences in gene expression, silencing and switching rates between *var* groups that are reinforced through the recombination hierarchy.

Interestingly, where the subtelomeric *var* group 2 genes were highly distinctive, we were unable to detect distinct signature residue features in the coding sequence of extracellular regions of central *var* genes when compared with the subtelomeric *var* group 1 (Fig. 4A–C). However, the two protein groups were partially discriminated by their intracytoplasmic tails or exon 2 sequences (Fig. 4D). It is surprising that central and subtelomeric *var* sequences were not easily distinguished as central chromosomal regions are generally less recombinogenic, and it has been hypothesized that central *var* genes might comprise a relatively conserved parasite reservoir encoding strain-transcendent PfEMP1 functions. More comparative sequence analysis needs to be done on central *var* genes, but one possibility is that limited DNA exchange occurs between central and subtelomeric *var* group 1 sequences (Fig. 5). In addition, both central and subtelomeric *var* group 1 sequences are predicted to bind CD36, so differences in the subtelomeric *var* group 2 may have been exaggerated by functional or immune selection.

It has been hypothesized that immunity directed at variant antigens could cause *P. falciparum* to organize into discrete strains with non-overlapping PfEMP1 repertoires (Gupta *et al*., 1996). A hierarchical model for *var* recombination has additional implications for PfEMP1 antigenic and adhesive diversification that have biological and clinical significance. For instance, multiple adhesion traits can reside in the same PfEMP1 protein and synergize to mediate infected erythrocyte binding and sequestration (Cooke *et al*., 1994; Baruch *et al*., 1996). DNA exchange provides the parasite with opportunities to experiment with new binding combinations. However, by constraining *var* gene reassortment possibilities, proteins may structurally and functionally diverge. One example could be the mutually exclusive binding of infected erythrocytes to CD36 or chondroitin sulphate A (CSA). This adhesion dichotomy has an essential role in pregnancy malaria (Rogerson *et al*., 1995; Fried and Duffy, 1996). Mechanistically, CD36-binding head structures inhibit the ability of downstream DBL-γ domains to bind CSA (Gamain *et al*., 2002). Hence, as infected erythrocytes become functionally specialized for particular organ niches, structural and functional distinctions may arise in PfEMP1 proteins that are reinforced through the recombination hierarchy. The effects of these structural adaptations for PfEMP1 protein vaccine design have been discussed (Gamain *et al*., 2002). For the parasite, it may be advantageous to maintain mutually exclusive adhesion traits in independently evolving *var* groups.

Our results pose interesting questions, such as what are the roles of the three *var* groups in infection and disease and what are the functions of the unusually conserved *var* loci? To address the former, differences in the 5′ flanking regions can be exploited to investigate gene expression during infection. As an example, some conserved subtelomeric *var* group 2 loci that are not predicted to bind CD36 are already being investigated for a role in pregnancy malaria (Buffet *et al*., 1999; Fried and Duffy, 2002; Rowe *et al*., 2002; Salanti *et al*., 2002; 2003; Khattab *et al*., 2003; Winter *et al*., 2003), and others might be identified using such primers. In addition, determining the potential involvement of conserved *var* loci as the ‘common and prevalent’ variants typifying severe childhood malaria infections (Bull *et al*., 1999; Nielsen *et al*., 2002) may provide additional insights into malarial disease.

In conclusion, this model of *var* genetic structuring offers clues to the ‘strain transcendence’ of a minority of *var* genes and could explain the structural and functional specialization of PfEMP1 proteins. This mechanism has important implications for the evolution of the gene family, provides new insights into severe malarial complications

and needs to be considered in PfEMP1 vaccine development and implementation.

Experimental procedures

Parasites

The A4, Dd2 and Malayan Camp (MC) clonal lines were cultured using standard practices, and genomic DNA was prepared. The A4 clone was originally derived by micromanipulation from *P. falciparum* line IT4/25/5, and the Dd2 strain is a clone of W2-MEF.

PCR amplification and sequencing of var *genes*

PCRs were done using TaKaRa LA *Taq*™ polymerase (Fisher) with 50 ng of genomic DNA, 1× buffer, 0.4 mM dNTPs (each), 2.5 mM $MgCl_2$, 0.5 μM primers and 2.5 U of enzyme run in a DNA Engine Dyad™ Peltier thermal cycler from MJ Research. Oligonucleotide primers are listed in Table 1. PCR conditions were one cycle of 94°C for 1 min followed by 35 cycles of 98°C for 1 min, annealing for 1 min, extension at 62°C for 6–18 min, depending upon fragment size, and a final extension for 10 min at 68°C. Products were cloned into the pCR®4-TOPO vector or pCR®-XL-TOPO from Invitrogen and sequenced to provide at least double coverage. In cases where degenerate DBL-type specific oligos (Kraemer *et al*., 2003) were adapted to sequencing reactions, oligo concentrations were increased 10-fold to account for primer degeneracy. Sequences were assembled using the PHRED/PHRAP/CONSED software suite (Gordon *et al*., 1998).

Sequence analysis

Phylogenetic analyses were done using CLUSTALX for multiple alignments and PAUP*4.0b10 (* phylogenetic analysis using parsimony and other methods) to generate neighbour-joining trees with 1000 bootstrap replicates. For tree generation, the new sequences amplified in the study were compared with 3D7 PfEMP1 sequences and others from non-3D7 parasite genotypes that had previously been mapped to chromosomal locations. 3D7 PfEMP1 sequences are identified by gene names in the figures, and their sequences can be accessed using this identifier at the PlasmoDB (http://plamodb.org/). Gene names and accession numbers for previously published non-3D7 genotypic sequences were: Dd2var (L40608), FCR3T11-1 (U67959), FCR3var2 and FCR3var3 (L40609), MCvar1 (U27338), MCvar2 (U27339), Dd2var7 (L42636), Dd2var7a (AF041422), Dd2var7b (AF041423), ItvarR29 (Y13402) and IT-A4var (3540145). Gene sequences were divided into respective domains using published approaches (Smith *et al*., 2000*).* BLAST analysis of each region was done against the PlasmoDB. Percentage sequence identities were calculated using the algorithm in DNAStar MEGALIGN, version 5.0.

Acknowledgements

The authors thank Siri Nelson, Martin Pentony and Peter Myler for guidance in using CONSED and advice concerning sequencing reactions using degenerate oligos. In addition, we thank Zoe Christodoulou, Alex Rowe and Sue Kyes for sharing their unpublished R29 upstream sequence with us. The authors also thank Leia Smith for a construct containing part of IT4 var2. J.D.S. is supported by an Ellison Medical Foundation New Scholars Award in Global Infectious Disease, the Bill and Melinda Gates Foundation and a National Institutes of Health grant (RO1 AI47953-01A1).

Supplementary material

The following material is available from
http://www.blackwellpublishing.com/products/journals/suppmat/mmi/mmi3814/mmi3814sm.htm

Fig. S1. Sequence comparison of the *var* genes 5′ and 3′ flanking regions.
The flanking sequences of *var* genes amplified from Malayan Camp, It4/25/5, and Dd2 strains were compared to 3D7 *var* genes after trimming to the same length. Circles indicate distinct tree branches with the corresponding bootstrap value. Genes are coloured according to Fig. 1. and sequence groups are shaded.

References

Baruch, D.I., Pasloske, B.L., Singh, H.B., Bi, X., Ma, X.C., Feldman, M., *et al.* (1995) Cloning the *P. falciparum* gene encoding *PfEMP1*, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell* **82:** 77–87.

Baruch, D.I., Gormely, J.A., Ma, C., Howard, R.J., and Pasloske, B.L. (1996) *Plasmodium falciparum* erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. *Proc Natl Acad Sci USA* **93:** 3497–3502.

Beeson, J.G., Rogerson, S.J., Cooke, B.M., Reeder, J.C., Chai, W., Lawson, A.M., *et al.* (2000) Adhesion of *Plasmodium falciparum*-infected erythrocytes to hyaluronic acid in placental malaria. *Nature Med* **6:** 86–90.

Buffet, P.A., Gamain, B., Scheidig, C., Baruch, D., Smith, J.D., Hernandez-Rivas, R., *et al.* (1999) *Plasmodium falciparum* domain mediating adhesion to chondroitin sulfate A: a receptor for human placental infection. *Proc Natl Acad Sci USA* **96:** 12743–12748.

Bull, P.C., and Marsh, K. (2002) The role of antibodies to *Plasmodium falciparum*-infected-erythrocyte surface antigens in naturally acquired immunity to malaria. *Trends Microbiol* **10:** 55–58.

Bull, P.C., Lowe, B.S., Kortok, M., and Marsh, K. (1999) Antibody recognition of *Plasmodium falciparum* erythrocyte surface antigens in Kenya: evidence for rare and prevalent variants. *Infect Immun* **67:** 733–739.

Calderwood, M.S., Gannoun-Azki, L., Wellems, T.E., and Deitsch, K.W. (2003) *Plasmodium falciparum var* genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J Biol Chem* **278:** 34125–34132.

Cooke, B.M., Berendt, A.R., Craig, A.G., MacGregor, J., Newbold, C.I., and Nash, G.B. (1994) Rolling and stationary cytoadhesion of red blood cells parasitized by *Plasmo-*

dium falciparum: separate roles for ICAM-1, CD36 and thrombospondin. *Br J Haematol* **87:** 162–170.

Deitsch, K.W., del Pinal, A., and Wellems, T.E. (1999) Intracluster recombination and var transcription switches in the antigenic variation of *Plasmodium falciparum*. *Mol Biochem Parasitol* **101:** 107–116.

Deitsch, K.W., Calderwood, M.S., and Wellems, T.E. (2001) Malaria. Cooperative silencing elements in var genes. *Nature* **412:** 875–876.

Figueiredo, L.M., Freitas-Junior, L.H., Bottius, E., Olivo-Marin, J.C., and Scherf, A. (2002) A central role for *Plasmodium falciparum* subtelomeric regions in spatial positioning and telomere length regulation. *EMBO J* **21:** 815–824.

Fischer, K., Horrocks, P., Preuss, M., Wiesner, J., Wunsch, S., Camargo, A.A., and Lanzer, M. (1997) Expression of *var* genes located within polymorphic subtelomeric domains of *Plasmodium falciparum* chromosomes. *Mol Cell Biol* **17:** 3679–3686.

Flick, K., Scholander, C., Chen, Q., Fernandez, V., Pouvelle, B., Gysin, J., and Wahlgren, M. (2001) Role of nonimmune IgG bound to PfEMP1 in placental malaria. *Science* **293:** 2098–2100.

Fowler, E.V., Peters, J.M., Gatton, M.L., Chen, N., and Cheng, Q. (2002) Genetic diversity of the DBLalpha region in *Plasmodium falciparum* var genes among Asia-Pacific isolates. *Mol Biochem Parasitol* **120:** 117–126.

Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., *et al.* (2000) Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* **407:** 1018–1022.

Fried, M., and Duffy, P.E. (1996) Adherence of *Plasmodium falciparum* to chondroitin sulfate A in the human placenta. *Science* **272:** 1502–1504.

Fried, M., and Duffy, P.E. (2002) Two DBLgamma subtypes are commonly expressed by placental isolates of *Plasmodium falciparum*. *Mol Biochem Parasitol* **122:** 201–210.

Fried, M., Nosten, F., Brockman, A., Brabin, B.J., and Duffy, P.E. (1998) Maternal antibodies block malaria. *Nature* **395:** 851–852.

Gamain, B., Gratepanche, S., Miller, L.H., and Baruch, D.I. (2002) Molecular basis for the dichotomy in *Plasmodium falciparum* adhesion to CD36 and chondroitin sulfate A. *Proc Natl Acad Sci USA* **99:** 10020–10024.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., *et al.* (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419:** 498–511.

Gordon, D., Abajian, C., and Green, P. (1998) CONSED: a graphical tool for sequence finishing. *Genome Res* **8:** 195–202.

Gupta, S., Hill, A.V., Kwiatkowski, D., Greenwood, A.M., Greenwood, B.M., and Day, K.P. (1994) Parasite virulence and disease patterns in *Plasmodium falciparum* malaria. *Proc Natl Acad Sci USA* **91:** 3715–3719.

Gupta, S., Maiden, M.C., Feavers, I.M., Nee, S., May, R.M., and Anderson, R.M. (1996) The maintenance of strain structure in populations of recombining infectious agents. *Nature Med* **2:** 437–442.

Gupta, S., Snow, R.W., Donnelly, C.A., Marsh, K., and Newbold, C. (1999) Immunity to non-cerebral severe malaria is acquired after one or two infections. *Nature Med* **5:** 340–343.

Hernandez-Rivas, R., Mattei, D., Sterkers, Y., Peterson, D.S., Wellems, T.E., and Scherf, A. (1997) Expressed var genes are found in *Plasmodium falciparum* subtelomeric regions. *Mol Cell Biol* **17:** 604–611.

Horrocks, P., Pinches, R., Kyes, S., Kriek, N., Lee, S., Christodoulou, Z., and Newbold, C.I. (2002) Effect of var gene disruption on switching in *Plasmodium falciparum*. *Mol Microbiol* **45:** 1131–1141.

Khattab, A., Kremsner, P.G., and Klinkert, M.Q. (2003) Common surface-antigen var genes of limited diversity expressed by *Plasmodium falciparum* placental isolates separated by time and space. *J Infect Dis* **187:** 477–483.

Kraemer, S.M., Gupta, L., and Smith, J.D. (2003) New tools to identify var sequence tags and clone full-length genes using type-specific primers to Duffy binding-like domains. *Mol Biochem Parasitol* **129:** 91–102.

Kyes, S., Taylor, H., Craig, A., Marsh, K., and Newbold, C. (1997) Genomic representation of var gene sequences in *Plasmodium falciparum* field isolates from different geographic regions. *Mol Biochem Parasitol* **87:** 235–238.

Lanzer, M., Fischer, K., and Le Blancq, S.M. (1995) Parasitism and chromosome dynamics in protozoan parasites: is there a connection? *Mol Biochem Parasitol* **70:** 1–8.

Miller, L.H., Baruch, D.I., Marsh, K., and Doumbo, O.K. (2002) The pathogenic basis of malaria. *Nature* **415:** 673–679.

Newbold, C., Warn, P., Black, G., Berendt, A., Craig, A., Snow, B., *et al.* (1997) Receptor-specific adhesion and clinical disease in Plasmodium falciparum. *Am J Trop Med Hyg* **57:** 389–398.

Nielsen, M.A., Staalsoe, T., Kurtzhals, J.A., Goka, B.Q., Dodoo, D., Alifrangis, M., *et al.* (2002) *Plasmodium falciparum* variant surface antigen expression varies between isolates causing severe and nonsevere malaria and is modified by acquired immunity. *J Immunol* **168:** 3444–3450.

Peterson, D.S., Miller, L.H., and Wellems, T.E. (1995) Isolation of multiple sequences from the *Plasmodium falciparum* genome that encode conserved domains homologous to those in erythrocyte-binding proteins. *Proc Natl Acad Sci USA* **92:** 7100–7104.

Roberts, D.J., Biggs, B.A., Brown, G., and Newbold, C.I. (1993) Protection, pathogenesis, and phenotypic plasticity in *Plasmodium falciparum* malaria. *Parasitol Today* **9:** 281–286.

Robinson, B.A., Welch, T.L., and Smith, J.D. (2003) Widespread functional specialization of *Plasmodium falciparum* erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome. *Mol Microbiol* **47:** 1265–1278.

Rogerson, S.J., Chaiyaroj, S.C., Ng, K., Reeder, J.C., and Brown, G.V. (1995) Chondroitin sulfate A is a cell surface receptor for *Plasmodium falciparum*-infected erythrocytes. *J Exp Med* **182:** 15–20.

Rowe, J.A., Kyes, S.A., Rogerson, S.J., Babiker, H.A., and Raza, A. (2002) Identification of a conserved *Plasmodium falciparum* var gene implicated in malaria in pregnancy. *J Infect Dis* **185:** 1207–1211.

Salanti, A., Jensen, A.T., Zornig, H.D., Staalsoe, T., Joer-

gensen, L., Nielsen, M.A., *et al.* (2002) A sub-family of common and highly conserved *Plasmodium falciparum* var genes. *Mol Biochem Parasitol* **122:** 111–115.

Salanti, A., Staalsoe, T., Lavstsen, T., Jensen, A.T., Kordai Sowa, M.P., Arnot, D.E., *et al.* (2003) Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering *Plasmodium falciparum* involved in pregnancy-associated malaria. *Mol Microbiol* **49:** 179–191.

Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., *et al.* (1995) Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82:** 101–110.

Smith, J.D., Subramanian, G., Gamain, B., Baruch, D.I., and Miller, L.H. (2000) Classification of adhesive domains in the *Plasmodium falciparum* erythrocyte membrane protein 1 family. *Mol Biochem Parasitol* **110:** 293–310.

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

Tami, A., Ord, R., Targett, G.A., and Sutherland, C.J. (2003) Sympatric *Plasmodium falciparum* isolates from Venezuela have structured var gene repertoires. *Malar J* **2:** 7.

Taylor, H.M., Kyes, S.A., Harris, D., Kriek, N., and Newbold, C.I. (2000a) A study of var gene transcription *in vitro* using universal var gene primers. *Mol Biochem Parasitol* **105:** 13–23.

Taylor, H.M., Kyes, S.A., and Newbold, C.I. (2000b) Var gene diversity in *Plasmodium falciparum* is generated by frequent recombination events. *Mol Biochem Parasitol* **110:** 391–397.

Thompson, J.K., Rubio, J.P., Caruana, S., Brockman, A., Wickham, M.E., and Cowman, A.F. (1997) The chromosomal organization of the *Plasmodium falciparum* var gene family is conserved. *Mol Biochem Parasitol* **87:** 49–60.

Vazquez-Macias, A., Martinez-Cruz, P., Castaneda-Patlan, M.C., Scheidig, C., Gysin, J., Scherf, A., and Hernandez-Rivas, R. (2002) A distinct 5′ flanking var gene region regulates *Plasmodium falciparum* variant erythrocyte surface antigen expression in placental malaria. *Mol Microbiol* **45:** 155–167.

Voss, T.S., Thompson, J.K., Waterkeyn, J., Felger, I., Weiss, N., Cowman, A.F., and Beck, H.P. (2000) Genomic distribution and functional characterization of two distinct and conserved *Plasmodium falciparum* var gene 5′ flanking sequences. *Mol Biochem Parasitol* **107:** 103–115.

Voss, T.S., Kaestli, M., Vogel, D., Bopp, S., and Beck, H.P. (2003) Identification of nuclear proteins that interact differentially with *Plasmodium falciparum* var gene promoters. *Mol Microbiol* **48:** 1593–1607.

Ward, C.P., Clottey, G.T., Dorris, M., Ji, D.D., and Arnot, D.E. (1999) Analysis of *Plasmodium falciparum* PfEMP-1/var genes suggests that recombination rearranges constrained sequences. *Mol Biochem Parasitol* **102:** 167–177.

Winter, G., Chen, Q., Flick, K., Kremsner, P., Fernandez, V., and Wahlgren, M. (2003) The 3D7var5.2 (var (COMMON) type var gene family is commonly expressed in non-placental *Plasmodium falciparum* malaria. *Mol Biochem Parasitol* **127:** 179–191.