Analysis of *Plasmodium falciparum var* Genes Expressed in Children from Papua New Guinea

Nicole Falk,[1,a] Mirjam Kaestli,[1,a,b] Weihong Qi,[1,b] Michael Ott,[2] Kay Baea,[3] Alfred Cortés,[3,b] and Hans-Peter Beck[1]

[1]Swiss Tropical Institute and [2]F. Hoffmann–La Roche AG, Basel, Switzerland; [3]Papua New Guinea Institute of Medical Research, Madang, Papua New Guinea

Background. The variable antigen *P. falciparum* erythrocyte membrane protein–1 (PfEMP1) is a major virulence factor in malaria. A large number of *var* genes encode PfEMP1, and we hypothesized that a restricted PfEMP1 repertoire determines clinical disease presentation. We conducted a case-control study in Papua New Guinea and analyzed transcribed *var* genes in naturally infected children.

Methods. *var* messenger RNA was isolated from 78 children with asymptomatic, mild, or severe malaria. We prepared complementary DNA from the upstream region into the DBL1α domain and picked, on average, 20 clones for sequencing.

Results. Twenty-fiv percent of centrally located *var* genes were shared between children, whereas only 5% of subtelomeric genes were shared, indicating lower diversity in the former group. Linkage between group B or C *var* upstream sequences and DBL1α groups was not observed, which impeded prediction by DBL1α analysis. A higher proportion of *var* group A sequences was detected in symptomatic malaria, and a subgroup of frequently encountered *var* genes with complex head structure seems to be associated with severe malaria. A subset of *var* group C genes was frequently expressed in older children with asymptomatic high levels of parasitemia.

Conclusion. Despite this vast diversity, restricted disease-associated *var* genes were identifie and might be used for innovative interventions based on PfEMP1.

Antigenic variation is an important evasion mechanism associated with parasite sequestration and virulence in *Plasmodium falciparum* malaria. The adherence of parasitized red blood cells to endothelial cells and to uninfected erythrocytes (rosetting) is characteristic of *P. falciparum* infection, leading to microvascular obstructions [1–3]. This is mediated by the variant surface antigen *P. falciparum* erythrocyte membrane protein–1 (PfEMP1), which is encoded by 1 of approximately 60 *var* genes [4]. *var* genes are classifie into 3 major groups (i.e., *var* group A, B, and C) and 2 intermediate groups (i.e., BA and BC) in accordance with their chromosomal location and their 5′ upstream region [5]. *var* group A and B genes are subtelomerically located, in contrast *var* group C genes are arranged in the center of chromosomes. Differences in transcriptional regulation between *var* group B and C have been described elsewhere [6] and might indicate differences in their function.

PfEMP1 is structured into several domains with an extracellular part composed of a variable N-terminal segment (NTS), various Duffy binding–like domains (DBL), and cysteine-rich interdomain regions [5]. These domains have been associated with different binding specificitie for host receptors [7–10]. The extracellular part is followed by a transmembrane domain and the conserved intracellular acidic terminal segment that anchors the protein to the cytoskeleton [9, 11, 12]. PfEMP1 enables the parasite to avoid splenic clearance by sequestration, but consequently it is presented to the immune system, eliciting an immune response. Therefore, the parasite undergoes antigenic variation by switching expression to alternative PfEMP1s to subvert

Received 24 September 2008; accepted 20 February 2009; electronically published 23 June 2009.

Potential conflicts of interest: none reported.

Financial support: Swiss National Science Foundation (grant 3100A0–104043/1 to H.P.B.); Optimus Foundation (grant to H.P.B.).

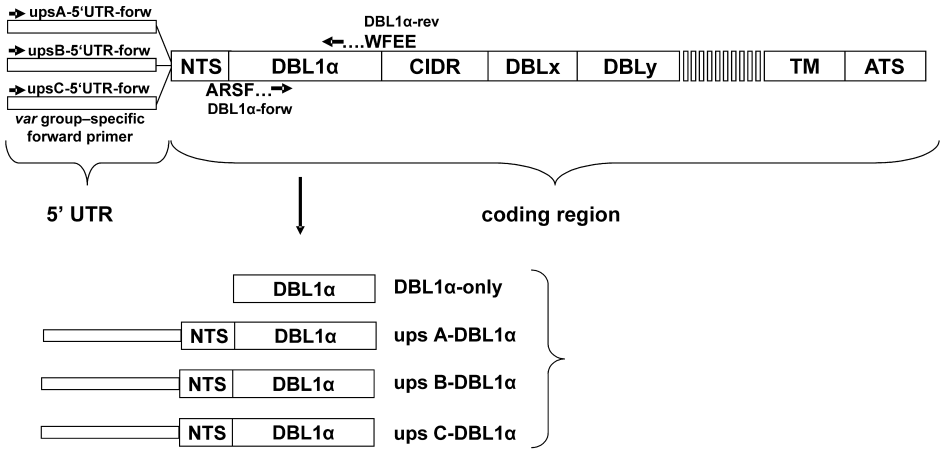
[a] These authors contributed equally.

[b] Present affiliations: Menzies School of Health Research, Emerging and Infectious Diseases Division, Charles Darwin University, Darwin, Australia (M.K.); Functional Genomics Center Zurich, Uni/ETH Zurich, Zurich, Switzerland (W.Q.); Institució Catalana de Recerca i Estudis Avançats and Institute for Research in Biomedicine, Cell, and Developmental Biology, Barcelona, Spain (A.C.).

Reprints or correspondence: Hans-Peter Beck, Swiss Tropical Institute, Socinstrasse 57, CH-4051 Basel, Switzerland (hans-peter.beck@unibas.ch).

The Journal of Infectious Diseases 2009; 200:347–56

0022-1899/2009/20003-0005$15.00
DOI: 10.1086/600071



primers	primer sequence	length of product, bp
upsA - 5'UTR - forw	5'-AACTTACCATAAATTATCATCAAA -3'	1172-1528
upsB - 5'UTR - forw	5'-ATGTAATTGTTGTTTTTTTTTTTGTTAGAATATTTAAA -3'	
upsC - 5'UTR - forw	5'-CACATATA(A/G)TACGACTAAGAAACA -3'	
DBL1α - forw	5'-GCACGAAGTTTTGCAGATAT(A/T)GG -3'	365 - 485
DBL1α - rev	5'-AA(A/G)TCTTC(T/G)GCCCATTCCTCGAACCA-3'	

Figure 1. *var* gene structure and polymerase chain reaction products generated with *var* group–specific 5′ UTR forward or DBL1α forward and DBL1α reverse primer. ATS, acidic terminal segment; CIDR, cysteine-rich interdomain region; DBL, Duffy binding–like domain; NTS, N-terminal segment; TM, transmembrane domain.

the immune response. This switch often changes the adhesive properties [13].

Because of antigenic variation, PfEMP1 is often not considered a feasible malaria vaccine candidate. Recent studies, however, have shown that structural similarities exist in the molecule's head structure [14], and only a few PfEMP1 variants have been associated with certain clinical presentations, such as pregnancy-associated malaria [15], in which variants bind to chondroitin sulfate A. Other PfEMP1 molecules have been associated with severe disease [16–19], for example, PfEMP1 molecules with DBL1-like domains lacking 1 or 2 of the cysteines characteristic of group A and B/A *var* genes [20,21]. Up-regulation of group A and B *var* genes in children with severe malaria was observed in Tanzania but not in Kenya [21, 22]. In Papua New Guinea, up-regulation of *var* group B genes was evident in patients with severe malaria, but up-regulation of *var* group A was not observed [23]. Despite inconsistencies, these results support the notion that structures or a limited number of disease-related *var* genes exist and that their identificatio would enable the development of interventions against severe disease.

To study the association of expressed *var* genes and clinical presentation of malaria, we analyzed the *var* transcripts of parasites recovered from children with severe, mild, or asymptomatic malaria in a case-control study in Papua New Guinea. *var* messenger RNA (mRNA) was reverse transcribed and amplifie with polymerase chain reaction (PCR), followed by cloning and sequencing. Sequences were classifie on the basis of amino acid motifs and the number of cysteine residues in the DBL1α domain, as proposed elsewhere [21]. This is one of few studies addressing *var* expression ex vivo, which adds to understanding of the clinical relevance of PfEMP1. We assessed diversity using bioinformatic tools, but this assessment also highlighted the limitations of the current approach with regard to the identific tion of specifi subsets of expressed *var* genes.

MATERIALS AND METHODS

Study samples. Samples were collected during a case-control study in Madang, Papua New Guinea, as described elsewhere [23]. Ethical clearance for this study was obtained from the Papua New Guinea Medical Research Advisory Committee. The blood samples analyzed were obtained from the following groups: (1) 15 children with severe malaria, including 4 cases with cerebral malaria as define by the World Health Organization [24]; (2) 29 children with mild cases of malaria and no further symptoms of another disease; and (3) 34 age-matched children who were parasitemic but asymptomatic.

Isolation of* var *transcripts and complementary DNA (cDNA) synthesis. Blood samples were treated with ethylenediaminetetraacetic acid and blood pellets were used for *var* mRNA isolation and reverse transcription, as described else-

Table 1. Number of *var* Transcripts Detected, According to Clinical Status of Subjects with Malaria

Sequence, clinical status	Total *var* sequences, no.	Different *var* sequences, no.	Mean no./ child (range)
5′UTR-DBL1α			
upsA			
All	133	19	0.4 (0–3)
Asymptomatic	4	1	0 (0–1)
Uncomplicated	107	13	0.6 (0–3)
Severe	22	5	0.4 (0–2)
upsB			
All	607	78	4.3 (0–13)
Asymptomatic	227	28	3.2 (0–11)
Uncomplicated	252	28	4.8 (0–13)
Severe	128	22	5.1 (0–9)
upsC			
All	603	51	2.4 (0–6)
Asymptomatic	256	21	2.3 (0–6)
Uncomplicated	237	22	2.3 (0–6)
Severe	110	8	2.8 (0–5)
Total	1343	148	7.1 (0–14)
DBL1α domain only	Total DBL1α-only sequences, no.	Different DBL1α-only sequences, no.	
Asymptomatic	482	148	6.3 (1–14)
Uncomplicated	280	141	6.7 (2–14)
Severe	163	75	6.1 (1–9)
Total	925	364	6.4 (1–14)

where [25]. In brief, parasite RNA was immediately extracted using TRIzol (Invitrogen) in accordance with the manufacturer's instructions. TRIzol extraction was repeated and RNA was treated twice with RQ1 DNase (Promega). Full-length *var* transcripts were obtained by incubation of RNA with biotinylated oligonucleotides complementary to exon 2 and captured by use of streptavidin-linked Dynal beads (Dynal, UK). Reverse transcription was performed on captured hybrids using Sensiscript reverse transciptase (Qiagen). Sample aliquots without reverse transcriptase served as negative controls.

Amplificatio of DBL1α domain and* var *group-specifi PCR. Degenerated DBL1α-5′ and DBL1α-3′ primers were used to amplify DBL1α-domains, as described elsewhere [25] (hereafter referred to as "DBL1α-only"). Upstream sequences were amplifie using *var* group A-, B-, and C-specifi forward primers (figu e 1) and DBL1α-3′ reverse primers, resulting in products of approximately 1.4 kb (hereafter referred to as "5′UTR-DBL1α"). PCR conditions were as follows: 5 min at 94°C, followed by 35 cycles of 30 s, 95°C; 60 s, 52°C; and 90 s, 68°C. PCR products were purifie using the High Pure PCR Purifi ation Kit (Roche).

Cloning and sequencing. PCR products were processed for sequencing as described elsewhere [25]. In brief, fragments were ligated into pGEM-T (Promega) or pET vectors for TOPO cloning (Invitrogen), in accordance with the manufacturer's instructions, and transfected into *Escherichia coli* SURE cells (Stratagene). If possible, a mean of 20 clones (range, 2–40) were picked per transfection and sequenced using the forward primer T7 (pGEM-T) or M13 (TOPO).

Sequence analysis. Sequences were checked and edited using MT Navigator (ABI; version 1.02b3), DNASTAR (version 4), NCBI BLAST, and ClustalW.

In general, a consensus sequence was represented by several sequences from the same transfection. Sequences were considered identical if they differed by <5 nucleotides. Nucleotide sequences were translated using ExPASY translation tool and checked for identity using a 95% identity cutoff.

Sequences were classifie on the basis of the number of cysteine residues and specifi amino acid motifs (positions of limited variability [PoLV]) within DBL1α domains, as proposed by Bull et al. [21]. DBL1α sequences were assigned to 1 of 6 groups and their distribution was analyzed with regard to *var* groups A, B, and C and the clinical status of the child from whom the sequences were isolated.

Multiple sequence alignments were performed in ClustalX (version 1.83) and unrooted, minimum evolution phylogenetic trees (based on amino acids and p-distance) were created with 1000 bootstrap replicates by use of Molecular Evolutionary

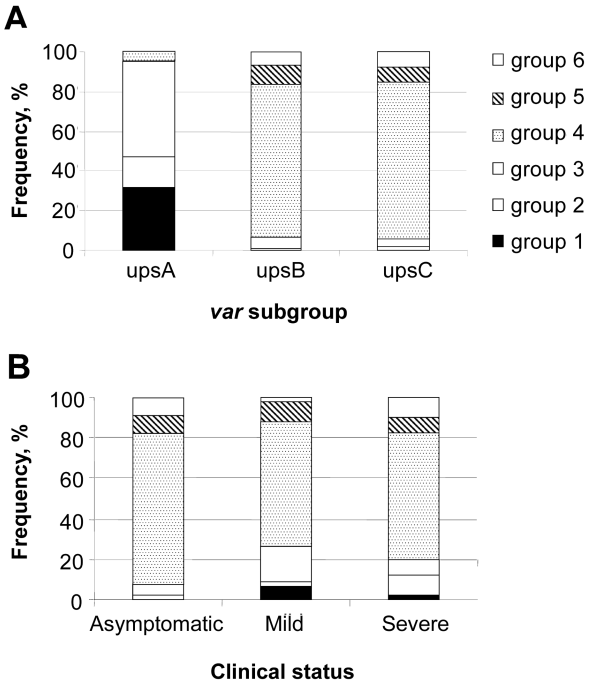


Figure 2. Grouping of different DBL1α sequences into 6 sequence groups, according to the classification model of Bull et al. [31]. *A,* Distribution of DBL1α sequence groups according to *var* subgroup (groups A, B, and C); *B,* Distribution of DBL1α sequence groups according to clinical status.

Genetics Analysis software (MEGA 3.1). Multidimensional scaling analysis (MDS) and analysis of similarities (ANOSIM) were performed using Primer (version 6.1.9; Primer-E). One-way ANOSIM, a nonparametric permutation procedure (999 permutations), tests the null hypothesis that there is no difference between 2 groups and is based on a Poisson-corrected distance matrix of amino acid sequences, as is MDS, a nonmetric multivariate ordination method.

All sequences have been submitted to GenBank (accession numbers are EU787517-EU787985).

Statistical analysis. Intercooled Stata (version 8.2; Stata) was used to perform Fisher's exact and Mann-Whitney *U* tests. All tests were 2-tailed, and differences were considered significant if P was <.05.

Species richness estimation. For sequence richness determination, sequences from Papua New Guinea were compared to a sequence set from Tanzania (J.P. Mugasa et al., unpublished data), to sequence data from Mali [19], and a global sequence collection [26]. DBL1α sequences from all samples were compared to each other by use of BLASTCLUST [27]. Sequences with 95% identity were assigned the same sequence type. The number of sequence types depended on how many samples were taken into account. Perl scripts were developed to simulate the sequential inclusion of all samples in all possible combinations. Accumulation curves were generated by plotting the increase in recovery of new sequences as a function of sampling effort.

RESULTS

Sequence data on 5′UTR-DBL1α was available for 24 of 34 asymptomatic children, 29 of 29 children with mild malaria, and 14 of 15 children with severe malaria. DBL1α-only sequences were available for 28 of 34 asymptomatic children, 23 of 29 children with mild malaria, and 14 of 15 children with severe malaria (table 1). Failure to obtain sequence data from all samples was mainly the result of failure to amplify cDNA for this particular stretch of the sequence. DBL1α-only sequences were primarily used to analyze the diversity of *var* transcripts, whereas phylogenetic analyses were performed on 5′UTR-DBL1α sequences.

var *group distribution in clinical subgroups.* Using msp2 genotyping [28], a mean multiplicity of infection of 1.4 infecting strains per child was determined, as described elsewhere [23], with no association between the number of infecting strains and the clinical outcome. We found a mean of 6.4 different DBL1α-only sequences per child, with no statistically significan differences observed between asymptomatic children, children with mild malaria, and children with severe malaria, even when adjusting for the number of infecting parasite strains. *var* group B and C transcripts were found in equal numbers in the 3 clinical subgroups, but only 1 *var* group A sequence was observed in 1 child with asymptomatic malaria (table 1).

Of 109 group C 5′UTR-DBL1α sequences, 27 (24.8%) were detected in more than 1 child, but only 11 (4.8%) of 231 group B sequences were shared ($P < .001$, by Fisher's exact test). Most *var* group B or C sequences found in >1 child were found in either asymptomatic children or children with mild malaria; in these 2 groups, only 19 group A sequences were detected, of which none was shared.

Distribution of motifs in DBL1α sequences. The sequences from block D (ARSFADIGDI) to block H (WFEEW) (fi ure 1) [9] of all DBL1α sequences were grouped according to the number of cysteines in this region and PoLV [21]. Figure 2*A* shows the distribution of sequence groups 1 to 6 according to *var* group A, B, or C transcripts. The majority of *var* group B and C transcripts contained sequences of group 4–6, with group 4 being the dominant. *var* group A contained group 1–3 sequences almost exclusively.

Group 1–3 sequences have been previously shown more likely to be associated with severe disease [19–21, 29–31]. Although the majority of the *var* transcripts observed in this study belonged to *var* group 4–6 (figu e 3*B*), group 1–3 sequences were significantl less common in asymptomatic children than in children with clinical malaria ($P = .007$, by Fisher's exact test) (figu e 2*B*). A detailed analysis of transcripts and asso-

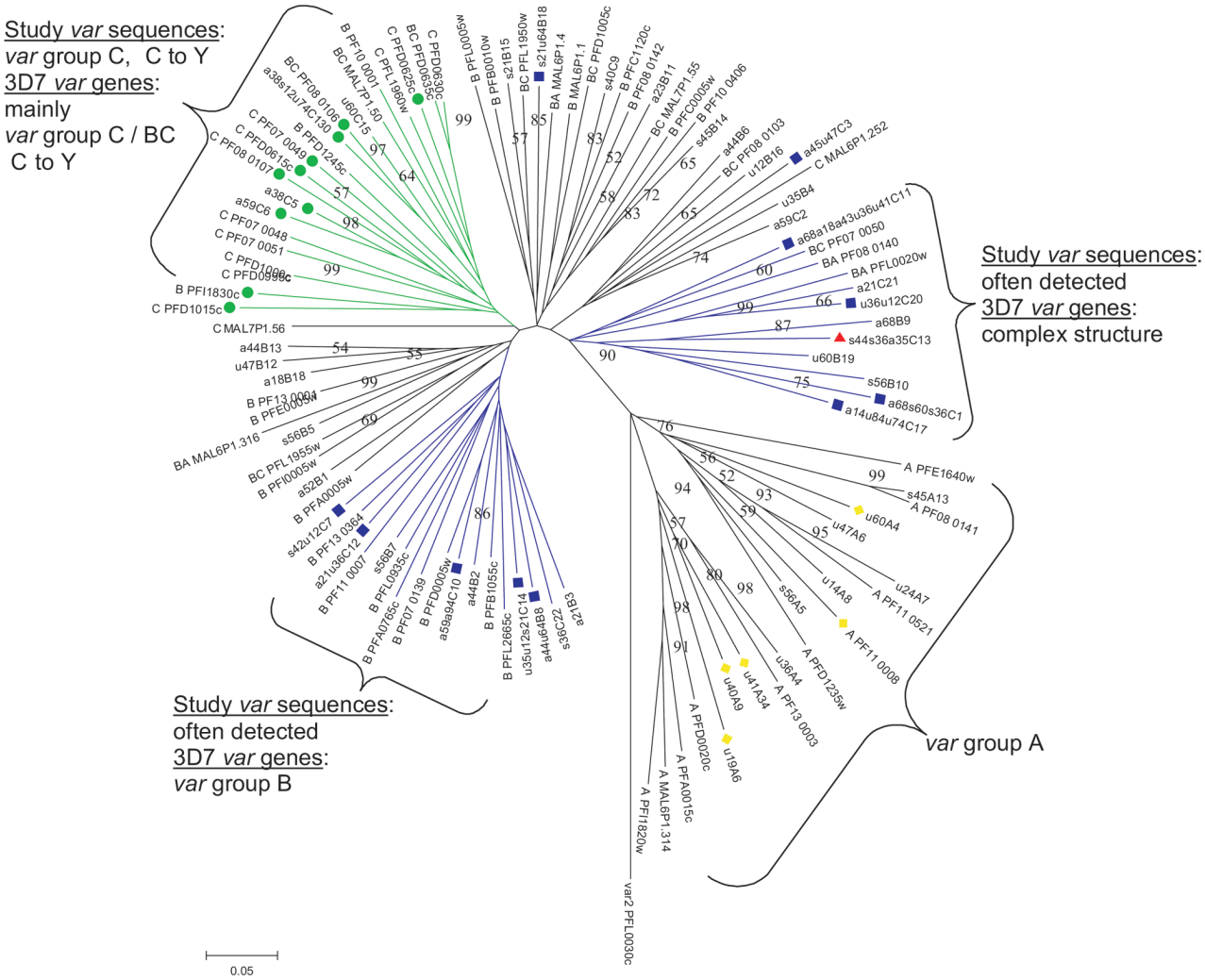


Figure 3. Unrooted minimum evolution, consensus radial tree (1000 bootstrap replicates; bootstraps above are 50% indicated) of the *var* N-terminal segment (NTS)–DBL1α amino acid stretch of study *var* transcripts, with 3D7 *var* genes used as a reference (study sequences are labeled as follows: *blue squares, var* sequences detected in several children; *green circles,* sequences with the Y motif in block B of DBL1α; *red triangle,* an identical NTS-DBL1α sequence found in 2 children with cerebral malaria; and *yellow diamonds, var* sequences of rosetting isolates). The labeling of the study sequences is as follows: (1) the small letters "a," "u," or "s" indicate asymptomatic, uncomplicated (i.e., mild), or severe malaria; (2) the first number refers to the age of the child in months; (3) the capital letters "A," "B," or "C" indicate the *var* group; (4) the number following the *var* group indicator corresponds to the sequence number. Successive sets of small letters and numbers indicate the same sequence found in several children.

ciations with sequence length and PoLV motifs can be found in the appendix, which appears only in the online version of the *Journal*

Upstream region of DBL1α-only sequences. The importance of upstream regions for classificatio of *var* genes has been recognized, but this information is difficul to determine. To test whether comparison of DBL1α sequences with the 3D7 genome might predict the upstream region in fiel samples, we used BLAST on DBL1α domains of UTR-DBL1α sequences for which we knew the upstream sequence, to compare with the 3D7 genome to determine the *var* group in silico.

DBL1α domains in linkage with *var* group A sequences were correctly determined in all but 1 case (17 of 18 sequences), whereas 21 (28.4%) of 74 group B sequences were incorrectly allocated to *var* group C, and 30 (62.5%) of 48 confir ed group C sequences were wrongly assigned to *var* group B. Overall, sequence similarity analysis comparing DBL1α with 3D7 led to incorrect assignments for 51 (36.4%) of 140 sequences.

Phylogenetic analysis of NTS-DBL1α sequences. A distance matrix–based, radial phylogenetic tree of NTS-DBL1α sequences with the *var* transcripts obtained and with 3D7 *var* genes as reference indicated 7 main clusters (figu e 3). Several clusters had an increased proportion of *var* sequences with specifi characteristics as described below.

First, the *var* group A cluster was the only cluster clearly separated, and transcripts from rosetting parasites were found

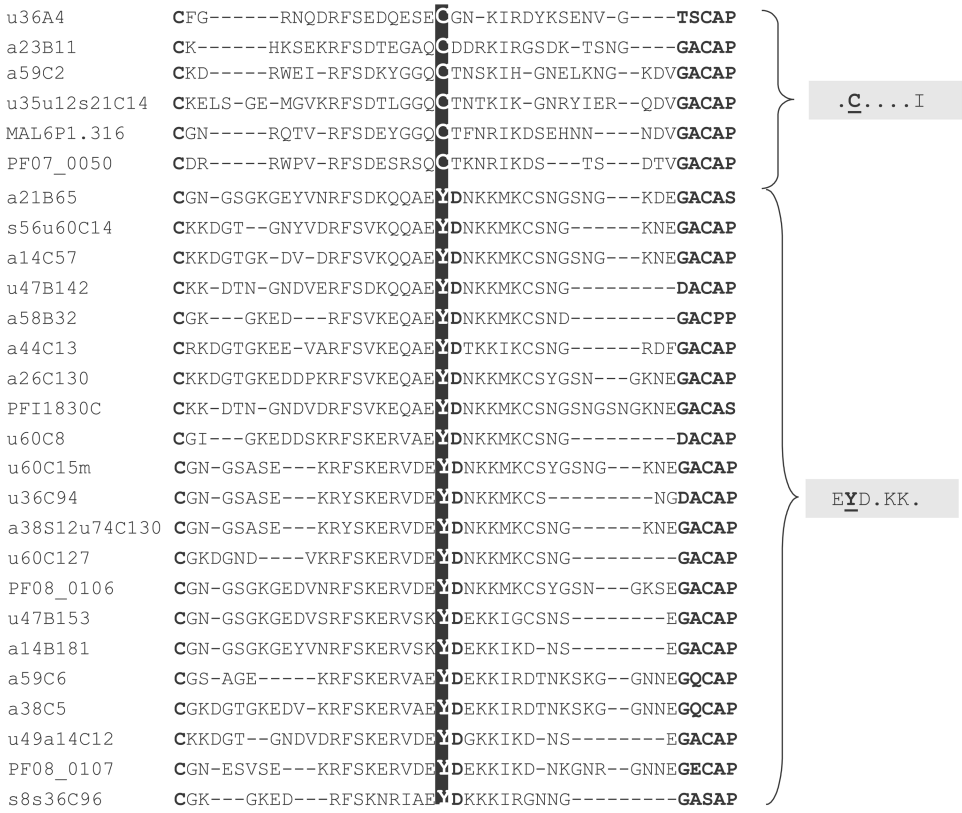


Figure 4. Multiple sequence alignment of DBL1α block A from cysteine 1 to cysteine 3. Selection of study and 3D7 *var* sequences with amino acid differences; differences between groups are indicated by bold type.

in this group. Second, clusters with commonly expressed *var* genes that were found in several children were mainly grouped into 2 clusters (figu e 3, blue clusters). One cluster comprised group B and C transcripts and three 3D7 *var* genes of the intermediate groups BA or BC, representing *var* genes with complex domain structures, including DBL2β-C2. In 2 of 4 children with cerebral malaria, the group C sequence s44s36a35C13 was observed to be the most abundant after cloning and sequencing. The second cluster also contained a mixture of group B and C sequences, which exclusively clustered with 3D7 *var* group B sequences.

Third, *var* group C cluster comprised mainly *var* group C or BC sequences (figu e 3, green cluster). Two closely related sequences (a59C6 and a38C5) originated from 2 children with asymptomatic malaria but very high parasitemia of 108,000 and 44,000 parasites/μL, respectively, and low hemoglobin levels. These *var* group C sequences were the dominant *var* transcripts (7 of 31 and 10 of 15 sequenced clones) in these children, and no *var* group A or B transcript was amplified confir ing previous quantitative PCR data [24]. These sequences have a distinct DBL1α block A with a cysteine-to-tyrosine substitution (figu e 4). Subsequently, this substitution was found in 19 different *var* transcripts, of which 14 were *var* group C genes. Fourteen of these sequences had a cysteine 7 aa further downstream instead. These 19 sequences were derived from 18 children (median age, 46 months; 4 children had severe malaria, 5 had mild malaria, and 9 were asymptomatic) who had a parasitemia level 7 times higher (median parasitemia level, 64,000 parasites/μL) than that of children not expressing this *var* variant (median parasitemia level, 8,920 parasites/μL) ($P = .067$, by use of the Mann-Whitney U test).

Clustering of* var *transcripts of severe and asymptomatic malaria cases. Although there was no obvious clustering of full-length NTS-DBL1α sequences of severe malaria cases, an MDS plot showed statistically significan clustering of the DBL1α stretch from block E to F [9] (figu e 5 and fi ure A4, which appears only in the online appendix) of dominant *var* transcripts from children with severe malaria and children with asymptomatic malaria (ANOSIM test, $P = .002$). In particular, 7 of 12 *var* sequences from children with severe malaria were clearly grouped apart (figu e 5), a result that is also evident in a distance tree of this sequence stretch that contains all *var* transcripts (figu e A4). This distinct cluster contained *var* transcripts of parasites from 14 children, 7 of whom had severe malaria ($P = .006$, by use of Fisher's exact test). *var* transcripts expressed in 3 of 4 children with cerebral malaria were also grouped in this cluster ($P = .025$, by use of Fisher's exact test).

Species richness determination. To estimate the diversity

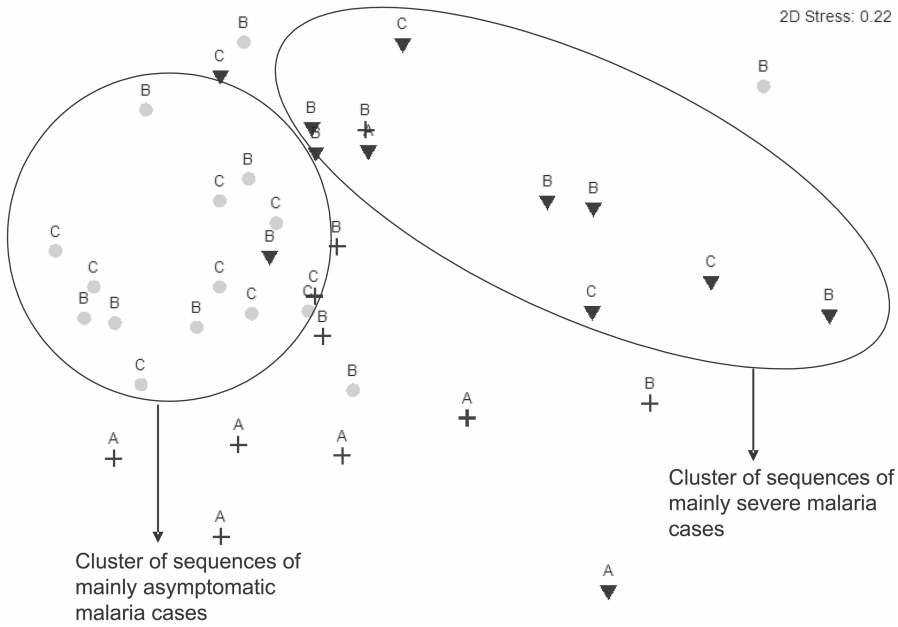


Figure 5. Multidimensional scaling analysis reflecting a distance matrix of *var* sequences between semiconserved DBL1α blocks E and F from children with severe *(black triangles),* mild *(crosses),* and asymptomatic malaria *(gray circles).* Letters A, B, and C indicate the *var* groups of these sequences.

of *var* DBL1α sequences, we simulated species accumulation curves based on four data sets (figu e 6). For each data set, the number of DBL1α sequence types was plotted against the number of samples studied, and empirical plots were well fitte by a linear function. The different nonparametric estimators of species richness implemented in Eco-Tools (http://www.eco-tools.net) were applied to all data sets, but none was stabilized before reaching the full number of samples.

In all cases, curves did not reach a plateau, which indicated that the diversity of DBL1α sequences is vast, and more sampling efforts are needed to completely capture sequence diversity. The slopes of the curves varied among data sets (fi ure 6), reflectin different sequence diversity. Other factors also contributed to the observed differences, such as the source of sequences (cDNA or genomic DNA) or the number of clones sequenced per sample.

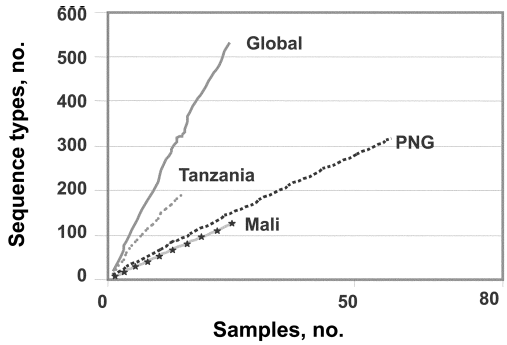
DISCUSSION

Little information is available on *var* gene expression in naturally infected malaria cases [19–23, 32]. To describe the expression of *var* genes, we conducted a case control study in Papua New Guinea and analyzed *var* transcripts of parasites isolated from 78 children with asymptomatic, mild, and severe malaria.

There was not a statistically significan difference between asymptomatic, mild, and severe malaria cases with respect to the number of different *var* transcripts of group B or C detected by cloning and sequencing, which is in accord with the data from Mali [19]. We previously reported a statistically signifi ant quantitative shift from *var* group B to group C transcripts in the same symptomatic and asymptomatic malaria cases, demonstrated by real-time PCR [23]. However, quantitative analysis based on cloning and sequencing of PCR products introduces bias through primers, amplificatio plateaus, and cloning and cannot be compared directly with quantitative PCR [33]. A combination of quantitative and qualitative information about *var* transcripts provides the most meaningful data.

An increase in the number of *var* group A transcripts in children with symptomatic malaria was observed, compared with the result for children with asymptomatic malaria. This is in agreement with the results of a study by Bull et al. [21] and subsequent studies [19, 29–31], which showed that DBL1α sequences of cys2 type (groups 1–3, mainly *var* group A genes) were primarily found in children with symptomatic malaria. This shift in the distribution of DBL1α groups between clinical presentations was only evident for dominantly expressed sequences.

A total of 370 different DBL1α-only sequences were detected, of which 7% were found in >1 child. Approximately 25% of all *var* group C sequences were detected more than once in several children, whereas only 5% of group B sequences were shared. This indicates a larger diversity of *var* group B than would be expected from the ratio of *var* group B to C genes per genome, which is only 3 in 3D7. This fi ding supports previous studies that reported high recombination rates, especially between group B genes [30, 34–36]. Phylogenetic anal-



Sample regions	PNG	Tanzania	Mali	Global
Reference	this study	unpublished	[19]	[26]
Sample type	patient samples	patient samples	patient samples	cloned isolates
Multiplicity of infection	1.3	>1	2.4	-
No. samples	61	15	26	25
Sequence tags per sample	12 (median)	45 (mean)	16 (median)	20 (median)
Type of sequence tags	cDNA	cDNA	cDNA	genomic DNA
Total no. sequence tags	963	668	277	608
Total no. sequence types	327	193	127	538

Figure 6. Comparison of *P. falciparum* erythrocyte membrane protein 1 DBL1α accumulation curves simulated for different data sets. *Asterisks,* Mali; Tanzania; *dotted line,* Papua New Guinea; global. The value on the *x*-axis is the no. of patient samples or cloned isolates studied. The data set is described in the table below the figure.

yses also suggested frequent genetic exchange between group B and C genes because they did not cluster separately, and we were unable to determine the respective *var* groups of DBL1α sequences by use of BLAST analysis.

Only 15 children with severe malaria (including 4 with cerebral malaria) were admitted to the hospital during the 4-month study period. The low incidence of severe malaria has been attributed to good access to antimalarials, the omnipresence of *P. vivax*, and the genetic background of Papua New Guineans [37]. Therefore, findin the same *var* NTS-DBL1α sequence dominantly transcribed in 2 of 4 children with cerebral malaria was surprising. In a distance tree, this sequence clustered with other frequently detected *var* transcripts and with 3 genes of complex structure of the 3D7 *var* group BA/BC genes. These 3D7 *var* genes contain additional domains such as DBL2β-C2, which was previously found to be associated with cerebral malaria and binding to the intercellular adhesion molecule–1 [38–40]. Another of these genes, PFL0020w, was previously found to be the second most transcribed *var* gene in an artificia *P. falciparum* 3D7 challenge infection in a non-immune human host [41]. PFL002w has a DBL1α sequence (DQ519151) identical to that of a *var* transcript from parasites isolated from heart tissue obtained from 2 children with fatal malaria in Malawi [32]. Thus, special attention should be paid to *var* genes within this cluster, as they might be strongly associated with the morbidity of malaria.

Although there was no obvious clustering of NTS-DBL1α sequences in cases of severe malaria, MDS analysis of block E to F of DBL1α showed statistically significan sequence clustering of a large subset derived from children with severe malaria. This cluster contained *var* transcripts from 3 of 4 children with cerebral malaria, suggesting a restricted subset associated with severe malaria.

Our findin that a subset of *var* sequences were associated with severe disease and detected in several children agrees with the results of previous studies [16, 18, 42] and suggests the existence of conserved *var* genes that are readily expressed in less-immune hosts, increasing parasite survival by providing distinct abilities to bind to vital organs.

Two older children with asymptomatic malaria had unusually high parasitemia levels (>40,000 parasites/μL) with no accompanying symptoms. Parasites obtained from these children almost exclusively expressed only 1 *var* group C transcript and previously had shown a strong proportional increase in the number of *var* group C transcripts [23]. These 2 sequences were closely related to each other and the 3D7 *var* gene PF08_0107 that was previously associated with weak IgG recognition in semi-immune children and with high expression in unselected 3D7 parasites [18]. In these *var* sequences and others within a cluster that contained mainly *var* group C sequences, a cysteine was replaced by tyrosine. It remains to be confi med whether this motif is a marker for a distinct *var* type

or whether it is involved in immune recognition or binding. Such replacement of a cysteine providing disulfid bridges with an aromatic tyrosine might lead to changes in the tertiary structure of PfEMP1.

Our findin suggests that *var* group C genes are primarily expressed in older, semi-immune hosts who already developed immunity against those *var* genes that are involved in severe pathogenesis. Unselected cultured parasites showed highly expressed *var* group C genes (H.P.B., unpublished data, as well as other studies [18, 43]) with low switch-off rates [43]. This might explain the exclusive transcription of group C genes in some semi-immune children. It is as yet unclear whether the high levels of parasitemia observed in these children were a result of *var* group C expression. A "hole in the antibody repertoire" [18, 44, 45] due to rare expression of some of these antigens or other strategies to evade protective immune responses could explain this result. The fact that these children were asymptomatic despite the high parasitemia levels suggests that these PfEMP1 molecules did not mediate cytoadherence in vital organs. In this regard, it is noteworthy that, in a previous quantitative study, parasite burden was positively associated with a proportional increase in *var* group C expression, especially in older children with asymptomatic malaria (age >36 months; $n = 14$; Spearman's $\rho = 0.82$; $P < .001$) [23].

In conclusion, we showed that a greater proportion of *var* group A or cys2 DBL1α sequences are present in children with symptomatic malaria than in children with asymptomatic malaria, and we provide evidence that a subset of frequently encountered *var* genes with complex DBL1α structure might be associated with more severe forms of malaria. We also described another subset of *var* group C genes that are frequently expressed in older children with asymptomatic high levels of parasitemia. However, we were unable to clearly identify an association between define *var* gene expression and severe malaria. Apart from technical reasons, such as the small number of severe malaria cases and the challenges inherent in applying a cloning and sequencing approach to fiel samples, this lack of association could also result from the fact that the mixture of parasites found in peripheral blood express many different *var* genes, as previously reported by Montgomery et al. [32]. Further research on *var* gene expression in natural infection and in different settings is urgently needed to understand the dynamics of *var* gene expression and the associated pathogenesis. Only then can the innovative, advanced proteomic studies be facilitated that are needed to identify PfEMP1 molecules that might become targets of new anti-disease interventions.

Acknowledgments

We would like to thank the children of Papua New Guinea and their parents for participating in the study and we acknowledge support from the staff at the Papua New Guinea Institute of Medical Research in Madang. We would like to thank Sylvia Steiger, Eric Huber, Mark Gillingwater, and Sonja Keller-Maerki for assistance in cloning and sequencing. We would like to acknowledge the support in sequencing from Dorothee Foernzler of Hoffmann La Roche, Basel. We thank the Basel Computational Biology Center ([BC]^2) for computing hardware and software support.

References

1. Marsh K, English M, Crawley J, Peshu N. The pathogenesis of severe malaria in African children. Ann Trop Med Parasitol **1996**; 90:395–402.
2. Dondorp AM, Pongponratn E, White NJ. Reduced microcirculatory flo in severe falciparum malaria: pathophysiology and electron-microscopic pathology. Acta Trop **2004**; 89:309–17.
3. Dondorp AM, Ince C, Charunwatthana P, et al. Direct in vivo assessment of microcirculatory dysfunction in severe falciparum malaria. J Infect Dis **2008**; 197:79–84.
4. Gardner MJ, Hall N, Fung E, et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature **2002**; 419:498–511.
5. Lavstsen T, Salanti A, Jensen AT, Arnot DE, Theander TG. Sub-grouping of *Plasmodium falciparum* 3D7 *var* genes based on sequence analysis of coding and non-coding regions. Malar J **2003**; 2:27.
6. Voss TS, Thompson JK, Waterkeyn J, et al. Genomic distribution and functional characterisation of two distinct and conserved *Plasmodium falciparum var* gene 5′ flankin sequences. Mol Biochem Parasitol **2000**; 107:103–15.
7. Baruch DI, Gormely JA, Ma C, Howard RJ, Pasloske BL. *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 U S A **1996**; 93:3497–502.
8. Smith JD, Kyes S, Craig AG, et al. Analysis of adhesive domains from the A4VAR *Plasmodium falciparum* erythrocyte membrane protein-1 identifie a CD36 binding domain. Mol Biochem Parasitol **1998**; 97: 133–48.
9. Smith JD, Subramanian G, Gamain B, Baruch DI, Miller LH. Classificatio of adhesive domains in the *Plasmodium falciparum* erythrocyte membrane protein 1 family. Mol Biochem Parasitol **2000**; 110:293–310.
10. Bir N, Yazdani SS, Avril M, Layez C, Gysin J, Chitnis CE. Immunogenicity of Duffy binding-like domains that bind chondroitin sulfate A and protection against pregnancy-associated malaria. Infect Immun **2006**; 74:5955–63.
11. 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.
12. Robinson BA, Welch TL, Smith JD. Widespread functional specialization of *Plasmodium falciparum* erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome. Mol Microbiol **2003**; 47:1265–78.
13. Roberts DJ, Craig AG, Berendt AR, et al. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature **1992**; 357: 689–92.
14. Klein MM, Gittis AG, Su HP, et al. The cysteine-rich interdomain region from the highly variable *Plasmodium falciparum* erythrocyte membrane protein-1 exhibits a conserved structure. PLoS Pathog **2008**; 4:e1000147.
15. Rowe JA, Kyes SA. The role of *Plasmodium falciparum var* genes in malaria in pregnancy. Mol Microbiol **2004**; 53:1011–9.
16. Bull PC, Kortok M, Kai O, et al. *Plasmodium falciparum*–infected erythrocytes: agglutination by diverse Kenyan plasma is associated with severe disease and young host age. J Infect Dis **2000**; 182:252–9.
17. Nielsen MA, Staalsoe T, Kurtzhals JA, et al. *Plasmodium falciparum* variant surface antigen expression varies between isolates causing severe and nonsevere malaria and is modifie by acquired immunity. J Immunol **2002**; 168:3444–50.
18. Jensen AT, Magistrado P, Sharp S, et al. *Plasmodium falciparum* as-

sociated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A *var* genes. J Exp Med **2004**; 199:1179–90.
19. Kyriacou HM, Stone GN, Challis RJ, et al. Differential *var* gene transcription in *Plasmodium falciparum* isolates from patients with cerebral malaria compared to hyperparasitaemia. Mol Biochem Parasitol **2006**; 150:211–8.
20. Kirchgatter K, Portillo H del A. Association of severe noncerebral *Plasmodium falciparum* malaria in Brazil with expressed PfEMP1 DBL1 α sequences lacking cysteine residues. Mol Med **2002**; 8:16–23.
21. Bull PC, Berriman M, Kyes S, et al. *Plasmodium falciparum* variant surface antigen expression patterns during malaria. PLoS Pathog **2005**; 1:e26.
22. Rottmann M, Lavstsen T, Mugasa JP, et al. Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children. Infect Immun **2006**; 74: 3904–11.
23. Kaestli M, Cockburn IA, Cortés A, Baea K, Rowe JA, Beck HP. Virulence of malaria is associated with differential expression of *Plasmodium falciparum var* gene subgroups in a case-control study. J Infect Dis **2006**; 193:1567–74.
24. World Health Organization, Communicable Diseases Cluster. Severe *falciparum* malaria. Trans R Soc Trop Med Hyg **2000**; 94(Suppl 1): S1–90.
25. Kaestli M, Cortes A, Lagog M, Ott M, Beck HP. Longitudinal assessment of *Plasmodium falciparum var* gene transcription in naturally infected asymptomatic children in Papua New Guinea. J Infect Dis **2004**; 189: 1942–51.
26. Barry AE, Leliwa-Sytek A, Tavul L, et al. Population genomics of the immune evasion *(var)* genes of *Plasmodium falciparum*. PLoS Pathog **2007**; 3:e34.
27. Altschul SF., Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol **1990**; 215:403–10.
28. Felger I, Tavul L, Beck HP. *Plasmodium falciparum:* a rapid technique for genotyping the merozoite surface protein 2. Exp Parasitol **1993**; 77(3):372–5.
29. Trimnell AR, Kraemer SM, Mukherjee S, et al. Global genetic diversity and evolution of *var* genes associated with placental and severe childhood malaria. Mol Biochem Parasitol **2006**; 148:169–80.
30. Kraemer SM, Kyes SA, Aggarwal G, et al. Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum:* comparisons of geographically diverse isolates. BMC Genomics **2007**; 8:45.
31. Bull PC, Kyes S, Buckee CO, et al. An approach to classifying sequence tags sampled from *Plasmodium falciparum var* genes. Mol Biochem Parasitol **2007**; 154:98–102.
32. Montgomery J, Mphande FA, Berriman M, et al. Differential *var* gene expression in the organs of patients dying of *falciparum* malaria. Mol Microbiol **2007**; 65:959–67.
33. Gölnitz U, Albrecht L, Wunderlich G. *Var* transcription prof ling of *Plasmodium falciparum* 3D7: assignment of cytoadherent phenotypes to dominant transcripts. Malar J **2008**; 7:14.
34. 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**; 407:1018–22.
35. 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**; 105:13–23.
36. Frank M, Kirkman L, Costantini D, et al. Frequent recombination events generate diversity within the multi-copy variant antigen gene families of *Plasmodium falciparum*. Int J Parasitol **2008**; 38:1099–109.
37. Müller I, Bockarie M, Alpers M, Smith T. The epidemiology of malaria in Papua New Guinea. Trends Parasitol **2003**; 19:253–9.
38. Springer AL, Smith LM, Mackay DQ, Nelson SO, Smith JD. Functional interdependence of the DBLβ domain and c2 region for binding of the *Plasmodium falciparum* variant antigen to ICAM-1. Mol Biochem Parasitol **2004**; 137:55–64.
39. Smith JD, Craig AG, Kriek N, et al. Identificatio of a *Plasmodium falciparum* intercellular adhesion molecule-1 binding domain: a parasite adhesion trait implicated in cerebral malaria. Proc Natl Acad Sci U S A **2000**; 97:1766–71.
40. Howell DP, Levin EA, Springer AL, et al. Mapping a common interaction site used by *Plasmodium falciparum* Duffy binding-like domains to bind diverse host receptors. Mol Microbiol **2008**; 67:78–87.
41. 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 U S A **2002**; 99:10689–94.
42. Lindenthal C, Kremsner PG, Klinkert MQ. Commonly recognised *Plasmodium falciparum* parasites cause cerebral malaria. Parasitol Res **2003**; 91:363–8.
43. Frank M, Dzikowski R, Amulic B, Deitsch K. Variable switching rates of malaria virulence genes are associated with chromosomal position. Mol Microbiol **2007**; 64:1486–98.
44. Bull PC, Lowe BS, Kortok M, Molyneux CS, Newbold CI, Marsh K. Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. Nat Med **1998**; 4:358–60.
45. Ofori MF, Dodoo D, Staalsoe T, et al. Malaria-induced acquisition of antibodies to *Plasmodium falciparum* variant surface antigens. Infect Immun **2002**; 70:2982–8.