Scholarly Research Exchange

Volume 2009 • Article ID 824949 • doi:10.3814/2009/824949

Research Article

Analysis of PfEMP1—*var* Gene Sequences in Different *Plasmodium falciparum* Malarial Parasites

Aarti Ozarkar,[1] D. Prakash,[2] Dileep Deobagkar,[1] and Deepti Deobagkar[1]

[1] *Molecular Biology Research Laboratory, Department of Zoology and Centre for Advanced Studies, University of Pune, Pune 411007, India*
[2] *Malaria Section, National Centre for Cell Science, Pune 411007, India*

Correspondence should be addressed to Deepti Deobagkar, deepti.deobagkar@gmail.com

Received 4 September 2008; Accepted 5 February 2009

Plasmodium falciparum synthesizes *P. falciparum* erythrocyte membrane protein-1 (PfEMP-1), a product of the multicopy *var* gene family, which localizes on the surface of infected erythrocytes. This protein plays an important role in cytoadherence and immune evasion. Comparative analysis of the molecular sequences of the DBLα domain of the *var* gene from different isolates of the parasite reveals variations in the number of cysteines and presence of small conserved motifs like DGEA, RGD, GAG-binding motifs. Phylogenetic analysis while highlighting the extensive diversity leads to clustered them in separate clades far apart from each other. Discriminant factor analysis of physicochemical properties of amino acid sequences revealed that the aliphatic index, isoelectric point, and instability index have more effect in deciding the variance of different isolates sequences. The origin of diverse repertoire of the DBLα domain in the parasites highlights the complexity of host-parasite relationship in the context of parasite survival.

1. Introduction

Malaria has emerged as a major health problem especially in tropical and subtropical regions of the world [1]. *P. falciparum* erythrocyte membrane protein-1 (PfEMP-1) is a product of the multicopy *var* gene family [2]. The gene has 60 copies; however, only one gene is expressed at a time. This variant surface antigen has a major role in evasion of the host immune response and cytoadherence [3–6]. PfEMP1 acts as a ligand for host endothelial epithelial receptor. This results in sequestration of infected erythrocytes in the microvasculature of the brain, placenta, and other organs thereby causing cerebral malaria and severe malaria. PfEMP1 is a virulence factor and plays an important role in pathophysiology of the disease (severe malaria and cerebral malaria) and enhances survival of the parasite. Thus, it has been considered as one of the vaccine targets [7]. PfEMP1 is composed of four domains: an N-terminal segment (NTS), Duffy binding-like (DBL) domains, cysteine-rich interdomain region (CIDR), and C2 domains [8]. DBLα domain is also becoming the target of immunoepidemiological and vaccine production studies to analyze diversity in the parasite population's worldwide [9].

We have reported distinct size polymorphism of DBLα domain of the *var* gene in cultured and clinical isolates of malaria parasites [10]. There are several reports that characterize the extent of sequence diversity of *var* genes in different geographical regions [11–21]. It seems likely that the extent of diversity reported till date indicates only the tip of the *var* gene diversity iceberg [22]. Malaria is prevalent and reemerging in India. Till date there are scanty reports regarding *var* gene diversity from India [21, 23]. Thus in order to understand the divergence in the sequence and structural motifs of *var* genes in Indian and Thailand parasite lines and field isolates from western part of India, we have analyzed sequences from the DBLα domain of the *var* gene whose flanking regions are highly conserved and compared our sequence data with the reported other Indian sequences deposited in GenBank.

2. Materials and Methods

2.1. Parasites Culture. *P. falciparum* parasite lines were cultured as described earlier [10]. Briefly, the erythrocytic stages from India (*FAN5 HS, PUNE-1, FMN-17*), Thailand (*SOHS, MP-14*), and Netherlands (*3D7*) were cultured in

RPMI 1640 medium supplemented with 0.5% Albumax II (Gibco-BRL, Md, USA) [24]. The field isolates from India (*C, D, NS1, NS2, and NS6*) were not cultured and were used directly for isolation of genomic DNA.

2.2. Isolation of Genomic DNA. Parasites were collected by centrifugation from cultures with 20% parasitemia with mature trophozoites stages and washed with PBS. Parasitized RBCs (*p*-RBC) were lysed with saponin (0.15%), and the parasite pellet washed with PBS. DNA was isolated by treatment with proteinase K in the presence of sodium dodecyl sulphate followed by phenol chloroform extraction and subsequent ethanol precipitation [25].

2.3. var Gene Primers and PCR Amplification. DBLα primers (synthesized by Invitrogen, Md, USA) of *var* gene were used for amplification of parasite genomic DNA (forward primer: CGACACCGGCGACATTATAAGAGG (primer1) and reverse primer: TCGCAGGTATTGTGGCACGTAGTC (primer2)). The primers were specific for the two highly conserved sites in DBLα domain and flank the polymorphic segment of DBLα domain. The amino acid sequences were DTGDIIRG and DYVPQYLR (see Figure 1 in Supplementary Material available online at doi:10.1155/2009/824949). PCR was carried out as described earlier [10].

2.4. Cloning and Sequencing. The PCR products were purified using PCR purification system (Roche). The DNA was ligated into pGEM-T easy system (Promega) and used to transform *E. coli* TOP10 cells (Invitrogen). The recombinants were sequenced using T7 and Sp6 primers with BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems 3730 sequencing machines).

2.5. Bioinformatic Analysis. The nucleotide sequences derived from the parasite lines and field isolates were analyzed for sequence similarities by NCBI BLAST (http://www.ncbi.nlm.nih.gov/Blast.cgi). The nucleotide sequences were translated into amino acid sequences using Expasy Translation tool. Pepstat analysis was performed to derive information about the amino acid sequences. By the SIM Alignment Tool, local similarity program (http://ca.expasy.org/tools/sim-prot.html) analysis was carried out. The multiple sequences were aligned using MultAlin program (http://bioinfo.genotoul.fr/multalin/multalin.html). Pairwise sequence alignment was generated using ISHAN-Integrated Software for Homology Analysis (http://physics.unipune.ernet.in/~pbv/ishan.html) [26]. The percentage of amino acid identity between sequences was calculated. An identity matrix analysis was performed, in which each sequence from a parasite was compared to all other sequences from the same isolate and from different isolates. The resultant matrix contained the percent sequence identity of each sequence relative to all other sequences. The mean and range values for the sequence comparison data were calculated. The percent similarity values against the sequences was plotted in MsExcel. The sequences (38) from other Indian isolates (ICGEB-R1, R15, R35, and MRC20) and 3D7 were derived from GenBank database and were also used for comparative analysis of the data. Multiple sequence alignment of the sequences for phylogenetic analysis of the parasite isolates wascarried out using ClustalW (http://www.ebi.ac.uk/clustalw/). Phylogenetic tree was constructed by using WebPhylip. The trees were reconstructed using protpars program, that implements maximum parsimony method. The phylogenetic analysis was performed to determine the relationships between the *var* gene sequences. The T- and B-cell epitopes prediction was carried out, as described earlier [27]. Although the prediction was carried out for few sequences by us, we have studied extensively the T-cell epitopes for all the sequences.

2.6. Statistical Analysis. The physicochemical properties of an amino acid sequence were determined using Protparam. The data was analyzed using multivariate statistics, as they offer the advantage of taking into account all the variables in a single analysis, thus making it possible to assess variation in the molecular weight, pI, amino acid index, aliphatic index, hydropathic index, and instability index of an amino acid sequence. The sequences were grouped among four groups, namely, Indian lab adapted, Indian field, ICGEB sequences [23], and MRC-20 sequences [21]. Discriminant factor analysis (DFA) was done to find out variables which were most useful for discriminating between the sequences of different groups. We performed Pillai's trace statistics to find out whether the clusters were significantly different from each other [28].

3. Results

The genomic DNA from parasites amplifies multiple bands in the size range of 350–700 base pair (bp) when amplified with the primer1 and primer2 (see Figure 2 in Supplementary Material available online at doi:10.1155/2009/824949) [10]. The PCR products were cloned, and 328 plasmids were screened. The Indian isolate *FAN5 HS* showed maximum number of clones whereas in Thai isolate *SOHS* yielded maximum number of clones. A total of 189 nucleotide sequences from Indian and Thai parasite lines and field isolates were obtained. The nucleotide sequences have been deposited in GenBank (accession numbers: *DQ 364441–364452*, *DQ 408208–408223*, *EF 143927–EF 143965*, *EU 333323–EU 333374*). SIM analysis of amino acid sequences predicted 99 unique sequences—Indian lab adapted *(FAN5 HS-12, PUNE-1-5, FMN17-4*), Thai lab adapted (*SOHS*-55, *MP-14*-7), and field isolates (*C-8, D-4, NS1-20, NS2-14, NS6-6*). The nucleotide sequences when submitted for NCBI-blast showed more than 50% homology to sequences from different parasite lines and isolates in the GenBank.

The nucleotide sequences of DBLα were translated into amino acid sequence and analyzed by PepStats analysis. It provides information about various features of the amino acid sequence. Majority of the sequences had small, polar, and charged residues (see Table 1). The amino acid sequence was analyzed using Multalin (see Figures 1 and 2). The analysis of the aligned DBLα sequences indicates presence of

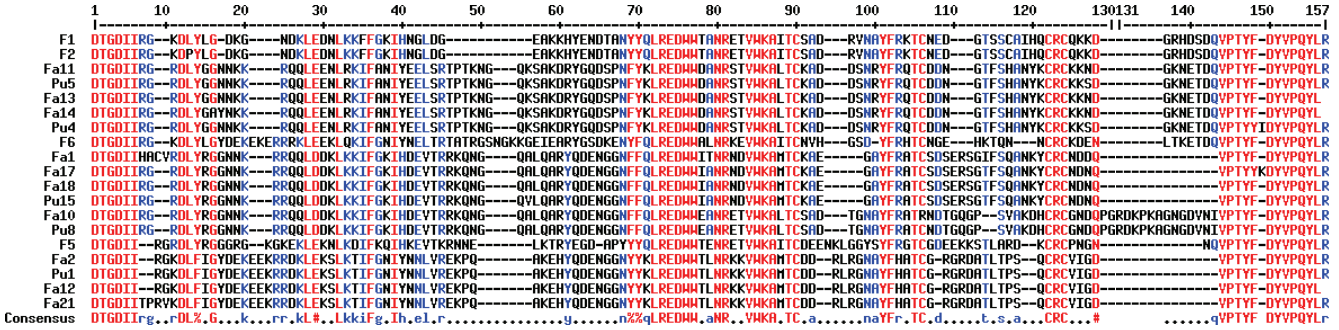

Figure 1: MuItiple sequence alignment of sequences derived from Indian lab adapted isolates. Multiple alignments of sequences derived from the DBLα domain of the parasite lines from India (FAN5 HS = Fa, PUNE-1 = Pu, and FMN-17 = F). Dots indicate gaps necessary for alignment. The Mulltalin program was used to align amino acid sequences, with the program default parameters. DGEA motif was observed in FMN-17 isolate (F1 and F2) and GAG binding motifs were observed in all the sequences. Positions of limited variability 2 (PoLV2 = LREDWW) and CRC are conserved in all sequences.

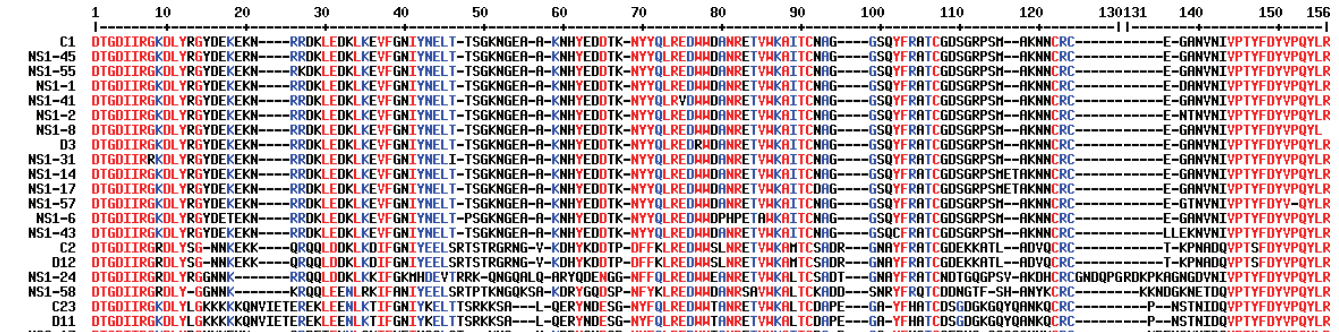

Figure 2: MuItiple sequence Alignment of DBLα domain from field isolates from India. Multiple alignments of sequences derived from the DBLα domain of the Indian field isolates (C, D, NS1, NS2, and NS6). Dots indicate gaps necessary for alignment. The Mulltalin program was used to align amino acid sequences with the program default parameters. Positions of limited variability 2 (PoLV2 = LREDWW) and CRC are not conserved in all sequences.

universally conserved blocks in the sequences (Supplementary Figure 1). Comparison of sequences showed that there is marked conservation in the positions of certain residues in DBLα domain. The conserved blocks were interspersed with variable blocks, and these sequences varied extensively in both sequence and length. The conserved block LREDWW was conserved throughout the lab adapted parasite lines (Indian and Thai), whereas it was semiconserved in field isolates of India. The sequences showed presence of conserved large bulky aromatic amino acids such as phenylalanine (F), tryptophan (W), and tyrosine(Y) and both hydrophobic as well as hydrophilic residues. It is interesting to note that a few sequences are shared and show similarity with other parasite lines while most sequences show extensive variation.

Table 1: Comparative summary of Pepstat analysis.

Sr. no.	Property (Residues)	*FAN5 HS*-1	*MP14-7*	*C23*	ICGEB25	MRC20-70
		Number (mole %)				
1	Tiny (A + C + G + S + T)	32 (25%)	32 (26%)	36 (27%)	27 (23%)	26 (21%)
2	Small (A + B + C + D + G + N + P + S + T + V)	61 (48%)	61 (49%)	63 (46%)	54 (45%)	53 (43%)
3	Aliphatic (I + L + V)	17 (13%)	16 (13%)	19 (14%)	18 (15%)	18 (15%)
4	Aromatic (F + H + W + Y)	18 (14%)	17 (14%)	17 (12%)	17 (14%)	16 (13%)
5	Nonpolar (A + C + F + G + I + L + M+ P + V + W + Y)	58 (45%)	55 (45%)	59 (43%)	53 (44%)	54 (44%)
6	Polar (D + E + H + K + N + Q + R + S + T + Z)	71 (55%)	68 (55%)	77 (57%)	66 (55%)	68 (55%)
7	Charged (B + D + E + H + K + R + Z)	42 (32%)	42 (34%)	44 (32%)	38 (31%)	45 (37%)

Although the DBLα domain has an average of 4 cysteine residues, the number varied from 2 to 9 cysteines. The parasite line from Thailand (*S61*) showed presence of 2 cysteines and was shorter in length (120 amino acids). The cysteines at CRC position were conserved throughout the Indian parasite lines whereasin field isolates and parasite lines from Thailand, it was semiconserved. The cysteine position in the homology block VWKAiTC is conserved in all sequences whereas in homology block YFraTC , it is semiconserved in parasite lines from Thailand. In parasite line from India, *FAN5 HS*, a single sequence had 3 cysteines and the cysteine in YFraTC block was absent.

Two of the *FAN5 HS* sequences showed extra long sequences of 276 and 250 amino acids with additional 156 and 148 amino acids at 5′ end, respectively. Similar case was noted in a field isolate NS6. A single sequence of the NS6 field isolate has showed presence of an extra long sequence of length of 165 amino acids with additional 35 amino acids at the 5′ end of the sequence. There was a duplication of 6 amino acid residues (DTGDII) at the 5′ end of the sequence. The *FAN5 HS* sequences showed presence of conserved blocks in DBLα domain namely GACAPYRRLH and CTLARSFADIGDI while NS6 sequence showed presence of CTLARSFADIGDI motif (Supplementary Figure 1).

The presence of small motifs (GAG-glycosaminoglycan binding, DGEA motifs, RGD motifs) was observed both in semiconserved and highly variable regions (see Table 2). The number of GAG binding motifs varied in all the sequences. The DGEA motif was observed only in the Indian parasite line, *FMN-17* (F1, F2) while RGD motif was observed only in parasite line from Thailand, namely, SOHS (S64).

Pairwise sequence comparisons were carried out to determine average sequence identity. It can be seen that although most of the sequences varied in their degrees of identity between 19%–62%, a few sequences were 99% identical to each other (see Table 3). There is a low sequence identity among the var sequences of Indian lab isolate, *FAN5 HS*. The sequences showed variation and were separated in different peaks (see Figure 3). The sequences analyzed in this study were grouped close together than other reported Indian sequences, whereas all Indian isolates showed similarity to each other when compared to 3D7. A given parasite line appears to contain sequences which cluster in distinctly different groups showing more similarity to sequences from other isolates and differed extensively from each other.

Table 2: Summary of motifs present.

Sr. No.	Motifs present	Isolates name
1	DGEA motif	*FMN-17-1* and *FMN-17-2*
2	RGD motif	*SOHS*-S64
3	LDV motif	Few 3D7 and ICGEB sequences
4	GAG—binding motifs	Majority of parasite lines of India, Thailand, and field isolates sequences

The phylogenetic unrooted tree was constructed using amino acid sequences from parasite lines of India and Thailand and field isolates (see Figure 4). These sequences were compared to 3D7 sequences and other previously reported Indian sequences available from GenBank database. The sequences of the individual isolates were clustered markedly in separate clades apart from each other. The parasite lines from India, *FAN5 HS* and *PUNE-1,* were closely related to each other whereas *FMN-17* was far away from them. The parasite line from Thailand and field isolates was clustered in distinctly separate groups. A few field isolates were close together to parasite lines from India and Thailand. However when compared with 3D7 sequences, the 3D7 sequences fall in separate clusters apart from parasite lines from India and Thailand and field isolates.

Discriminant factor analysis of physicochemical properties of amino acid sequences derived from different parasite

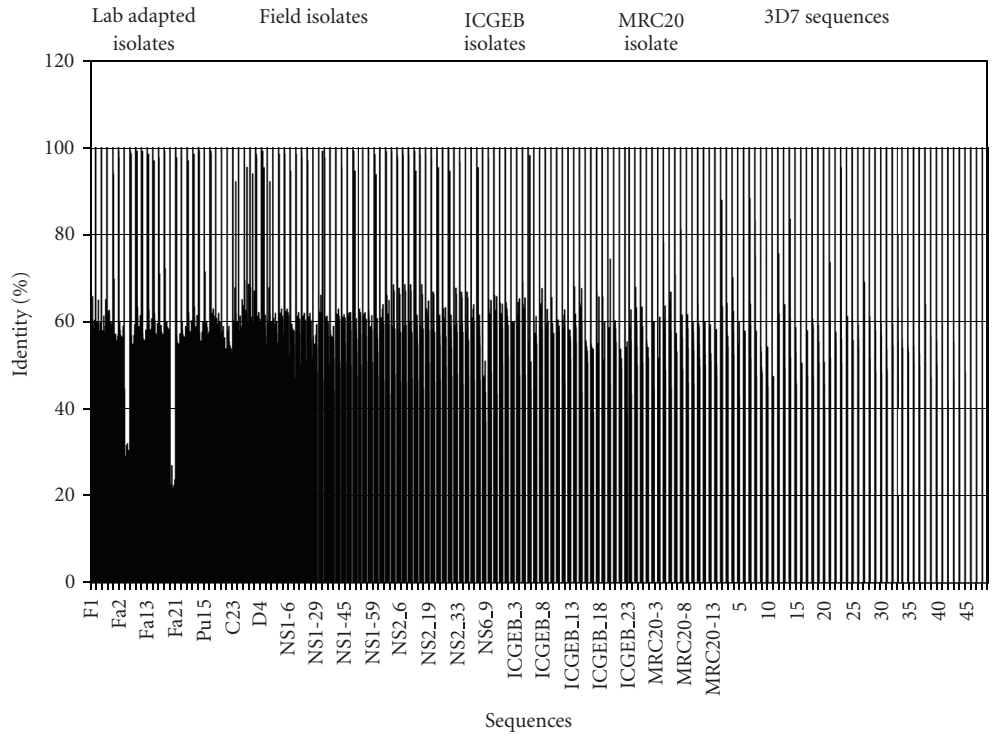

Figure 3: Pairwise sequence alignment graph of Indian isolates and 3D7. Pairwise sequence alignment was performed using ISHAN software and the matrix was plotted using MsExcel. x-axis shows isolates sequences whereas y-axis shows percent identity.

lines revealed that the four groups, that is, Indian lab adapted, Indian field, ICGEB isolates, and MRC-20, do not form significantly different clusters (see Figure 5(a)). The aliphatic index, isoelectric point, and instability index have more effect in deciding the variance of different isolates sequences, where as molecular weight of the protein, its amino acid index and hydropathicity index are not significant in deciding the variation (see Figure 5(b)).

The T-cell epitope prediction revealed that majority of the sequences showed weak binding motifs; however few sequences (11) showed strong binding motifs (see Table 4). The ability to bind T cells seems to show differences in the various genes. The sequences showed presence of B-cell epitopes in all sequences.

4. Discussion

The diversity in the *var* gene in malarial parasite has been reported. In the present study, we have critically analyzed *var* gene sequences in terms of motifs, amino acid, physicochemical properties, and constructed phylogenetic trees from *var* gene sequences derived from parasite lines from India and Thailand and also field isolates from Western part of India. In this analysis, the variation in the DBLα domain in sequences derived from Indian and Thai laboratory adapted parasite lines and field isolates was compared to sequences reported in GenBank from other Indian isolates and 3D7. Extensive polymorphism and variations were observed in DBLα domain. The multicopy DBLα domain seems to be highly diversified while conserving the salient features of the domain. Thus we have observed a highly dynamic and variable picture of *var* gene organization. Polymorphism within the hypervariable region leads to length variation and sequence variability. The *var* gene diversity at genomic level among parasite isolates both within and among endemic areas has been reported earlier [11–21]. The parasite line from India (*FAN5 HS*) and field isolate (*NS6*) both show presence of extra long sequences where DTGDII motif is repeated. It is likely that events such as duplication and DNA recombination may be possible. Thus gene recombination and duplication have been an important mechanism for generating diversity in *var* genes [12, 13, 29–31]. It has been shown that children with symptomatic infections had a greater repertoire of variant-specific antibodies [16]. Additionally, the *var* genes of *P. falciparum* have been reported to undergo constant changes due to frequent recombinationor rearrangements that generate a vast repertoire of *var* genes in nature.

Natural *Plasmodium falciparum* populations are genetically diverse, to an extent that within some geographic regions nearly all isolates contain unique parasite genotypes with regard to polymorphic single copy genes [11]. The repertoire is so immense that it raises questions about the possible molecular genetic mechanism instrumental in creating such high variability while conserving the important functionally conserved regions. The possible implications of this in immune evasion and persistence of parasites in host are of great importance to pathogenicity and virulence. It was reported that global *var* gene repertoire was immense even among geographically close isolates. High degree of similarity was observed among sequences from 3D7 [14]. In

Table 3: Pairwise comparison from *var* genes from *Plasmodium falciparum* Indian and Thai parasite lines and Indian field isolates.

Sr. no.	Isolates name	Range (%)	Average identity (%)
	Indian parasite lines		
1	FAN5 HS/FAN5 HS	19–99	59
2	PUNE-1/PUNE-1	46–98	72
3	FMN-17/FMN-17	49–99	73
4	FAN5 HS/PUNE-1	19–99	59
5	FAN5 HS/FMN-17	47–56	51
6	PUNE-1/FMN-17	48–57	52
	Thai parasite lines		
1	SOHS/SOHS	43–99	71
2	MP-14/MP-14	53–99	76
3	MP-14/SOHS	49–99	74
	Indian field Isolates		
1	C/C	48–97	72
2	D/D	52–59	55
3	NS1/NS1	45–99	72
4	NS2/NS2	53–99	76
5	NS6/NS6	41–99	70
6	C/D	48–99	73
7	C/NS1	44–99	71
8	C/NS2	48–62	55
9	C/NS6	39–61	50
10	D/NS1	50–98	74
11	D/NS2	46–60	53
12	D/NS6	42–60	51
13	NS1/NS2	45–60	52
14	NS1/NS6	38–99	68
15	NS2/NS6	40–99	69

Table 4: Summary of strong binding T-cell epitopes in Indian isolates.

Sr. No.	Isolates name	Sequence
1	PUNE-1-4	ETDQVPTYY
2	C2, D12	CSADRGNAY
3	NS1-24	CSADTGNAY
4	NS1-66, NS6-9, NS6-12, NS6-15	YQDESGNYY
5	ICGEB-5	SDRCGHNYN
6	ICGEB-16	LTDAKAKSY, CSAEENDRY
7	MRC20-1	LTDPKAKEY

three different geographical parasite lines, highly divergent *var* gene sequences were observed [20]. Vast amount of global *var* gene diversity was reported whereas limited amount of diversity was observed in PNG isolates [21]. Limited diversity was observed in India as the isolates were collected from an area having low malaria transmission rate. However, we observed considerable diversity among clones, even within a single isolate. It was reported that ^{var}COMMON type gene family is found in few Kenyan and Indian ICGEB isolate R35 [32]. However, our sequences derived from parasite lines and field isolates data analysis do not show ^{var}COMMON type gene.

It was reported that DBLα domain comprises of 400 amino acids and contains 16–18 conserved cysteine residues [8, 12, 33]. The number of cysteine residues is important for classification of DBLα domain and severity of disease [8, 16]. Cysteine 8 was conserved in all sequences. The position of cysteine at CRC motif was conserved in all Sudanese *var* genes. We also observed the similar conserved cysteines (CRCs) in the Indian isolates. Variants with unusual number of cysteines may form a subset with altered antigenic and adhesive properties. Sequences predominantly expressed in patients with severe malaria could be subgrouped on the basis of number of cysteine residues. These sequences were commonly found in the *var* gene repertoire of parasites from patients with mild malaria [34]. This suggests that positions of cysteines are important for DBLα structure, possibly because they are involved in disulphide bonding or other aspects of conformation or folding. The cysteine disulphide bonds may be important for stabilizing the surface-located variable loops and protuberances which appear to play a critical role in folding or structural conformation of

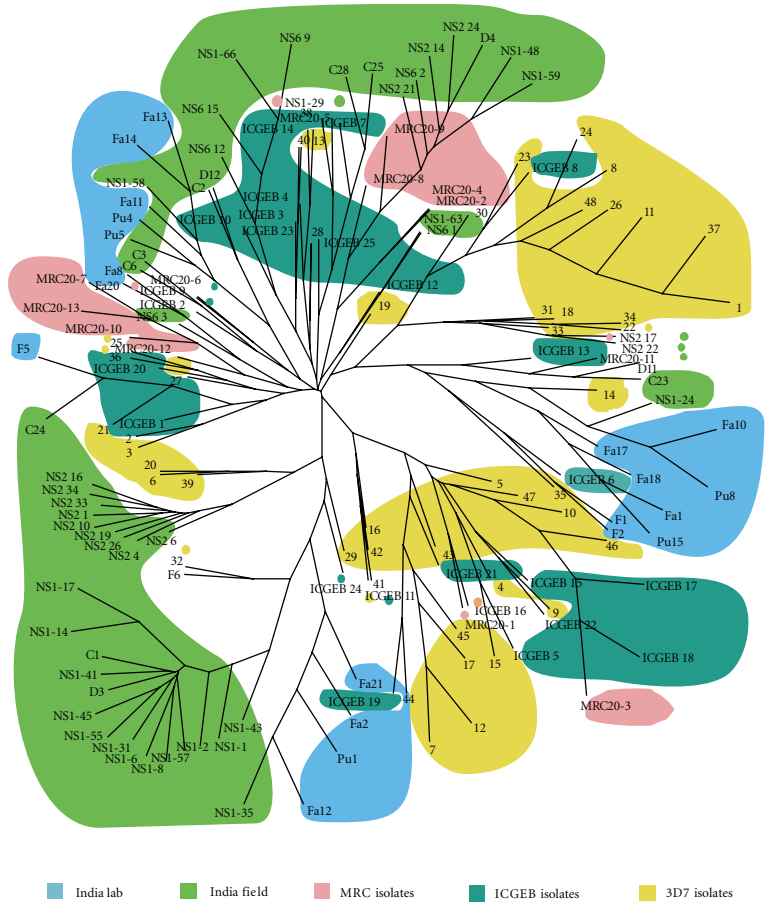

Figure 4: Phylogenetic tree of sequences of parasite lines from India lab adapted and field isolates and ICGEB and MRC20 sequences and 3D7. Unrooted phylogram was constructed by using WebPhylip—protpars program that implements maximum parsimony method. Yellow block represents 3D7 DBLα sequences, light blue block represents Indian lab adapted parasite lines (FAN5 HS, Pune-1, FMN-17), green block represents Indian field isolates (C, D, NS1, NS2, NS6), dark blue represents ICGEB sequences, and pink block represents MRC-20 sequences. The sequences of the isolates are clustered in separate groups apart from each other.

whole PfEMP1 protein. Presence of conserved cysteines and hydrophobic amino acid residues suggests that they are structurally important.

The *var* genes in *Plasmodium falciparum* genome are classified into 17 different protein architectural types based on domain compositions [35]. The genome is divided into 3 major subgroups (A, B, C) and two intermediate groups (B/A, B/C) on the basis of 5′ upstream (Ups) sequence and chromosomal location [36, 37]. Group A and B/A genes have DBLα 1 type domain whereas groups B, C, B/C, or B/A have DBLα domain. Based on the number of conserved cysteine residues in the sequences , *var* genes are grouped as DBLα and DBLα1 domains. DBLα domain has 4 conserved cysteine residues whereas DBLα1 has 2 cysteine residues [38]. However, we too report that majority of the sequences have DBLα domain, that is, 4 cysteine residues, and all the isolates are from severe malaria cases.

A number of charged potentially exposed amino acid residues are also present in these segments. Divergent segments contain multiple hydrophilic residues suggesting that they are likely to be exposed and may serve as epitopes for agglutinating antibodies. Presence of common segments even in divergent regions of DBLα domains and the hydrophilic nature of the sequences suggest that it is likely that they are both exposed and form epitopes for antibody recognition. The presence of conserved cysteines and tryptophan (W) residues is amongst the most highly conserved residues suggesting that DBL domains share conserved three-dimensional structures. We have shown that there is variation in T-cell and B-cell epitopes and also in their 3D structures. The difference in number and position of cysteine residues suggests a rearrangement of disulfide bonds leading to different three-dimensional structures [39].

Sequences consisting of position of limited variability (PoLVs 1–4) and the number of cysteine residues have been suggested as signature sequences of the DBLα domain [16]. Each PoLV was four amino acids long and situated adjacent to conserved amino acid residues at the boundary of previously defined islands of homology. In the sequences from parasite lines from India, PoLV2 (LREDW) is conserved

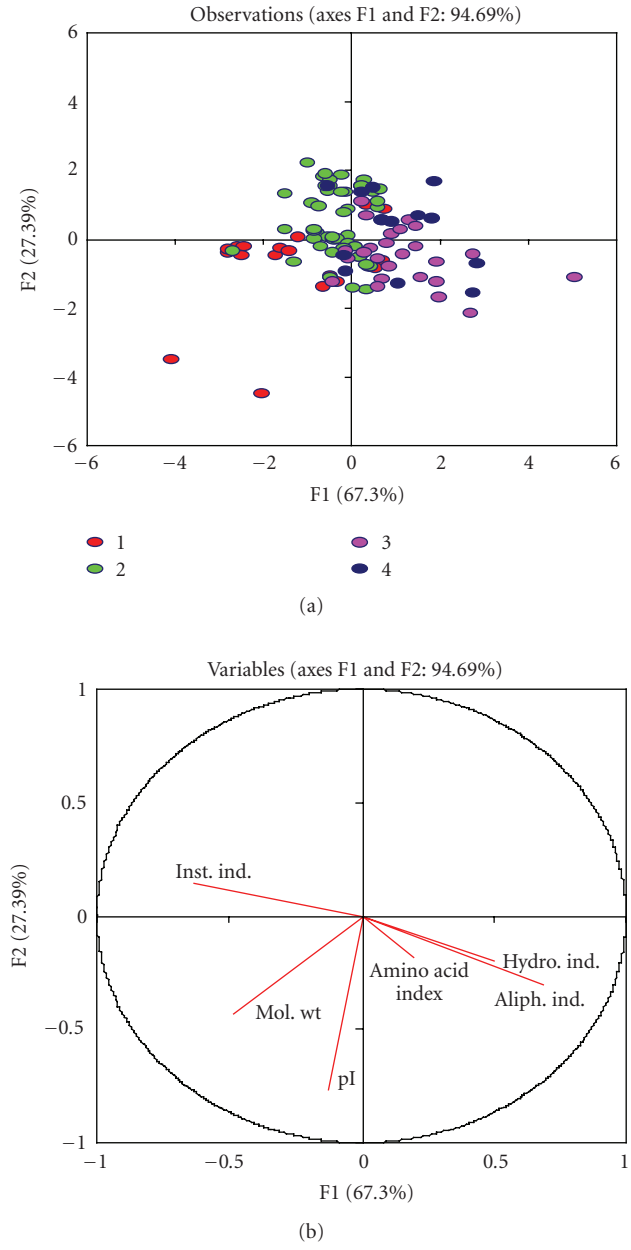

Figure 5: Discriminant factor analysis of physicochemical properties of an amino acid sequence. (a) Clusters of different groups, (b) variables which discriminate between the clusters. Group 1 = Indian lab adapted isolates (red), 2 = Indian field isolates (green), 3 = ICGEB isolates (pink), and 4 = MRC-20 (blue). Aliphatic index, isoelectric point and instability index have more effect in deciding the variance of different isolates, whereas molecular weight, amino acid index, and hydropathicity index are not significant.

whereas in parasite lines from Thailand and field isolates all PoLV 1–4 are different. Thus the sequences derived from parasite lines and field isolates differed from each other, and variation was also observed within sequences of a single parasite line.

The glycosaminoglycan (GAG) binding motifs are clusters of positively charged amino acid residues. These small amino acid motifs are linked to the severe states of malaria [22]. We observed them in sequences obtained from parasite lines from India and Thailand and field isolates. These motifs are responsible for involvement of the protein in rosette formation of iRBC. It was observed that similar motifs could not be linked to severity of the disease [15]. It has been suggested that intravenous injection of glycosaminoglycan disrupts rosettes and releases already sequestered parasites into circulation. Thus it can act as candidate for adjunct therapy in severe malaria [40]. Heparin sulfate on the surfaces of uninfected RBC may act as receptors for rosetting and that GAG-binding motifs of PfEMP1 mediate this binding [39]. DGEA mediates integrin-collagen interactions in platelets. It has been suggested that it is important in cell adhesion molecules and can therefore be a target for therapeutic intervention [41]. The RGD motif was observed in parasite line from Thailand *SOHS* and in 3D7 sequences. The LDV motif was observed in 3D7 and other Indian isolates (ICGEB-R1, R15, R35), but it was not present in any of our sequences. The RGD and LDV motifs are associated with protein-protein interaction and cell attachment [42]. They are cell adhesion motifs involved in ligand-receptor interactions involving the integrin family [43]. Thus it can act as a potential therapeutic agent for the treatment of disease.

The T- and B-cell epitopes in different sequences of DBLα domain of parasites show extensive variation. The epitopes varied in its location and amino acid composition. Due to such variation, it is difficult to make an effective vaccine. Thus the conserved epitopes are required for an effective vaccine design against malaria.

The evolutionary relationships of conserved cysteine rich motifs in adhesive molecules of malaria parasites have been studied [44]. The report suggests that rapid divergence originated from multiple gene duplication events. There is a sequential pattern of repeated duplication and diversification. The unrooted phylograms showed that the genes show diversity even among the lab adapted and field isolates from India. The sequences in the genes of Indian and Thai isolates are quite diverse from that of 3D7. This interesting pattern demonstrates the variations in the sequence organization of the *var* gene domains and raises interesting possibility about role of mutations and recombination in the generation of this diversity. The role of host-induced immune pressure remains to be elucidated.

The DBLα domain of pfEMP1 displays extensive divergence in both sequence and length even among the same parasite line. Despite the extent of sequence diversity in DBLα domains, it is predicted that due to the presence of conserved cysteines and homology blocks between invariant cysteines, the DBL domain may have a common fold. As a result, the receptor-binding pockets may lie in the same region of diverse DBL domains. The evolution of variation in the *var* gene sequences in different geographical parasite lines and field isolates with respect to small motifs seems to play important role for parasite survival as well as to evade immunity.

Acknowledgments

The authors thank Dr. G. C. Mishra, Director of NCCS, Department of Biotechnology (DBT), and GOI for providing facilities; ICMR, UGC, DST-FIST, and Government of India for providing financial assistance. A. Ozarkar acknowledges SRF from Institute of Bioinformatics and Biotechnology (IBB-UPE). They also thank Mr. Mangesh S. Deval (NCCS), for help in maintenance of parasite cultures.

References

[1] M. J. Gardner, N. Hall, E. Fung, et al., "Genome sequence of the human malaria parasite *Plasmodium falciparum*," *Nature*, vol. 419, no. 6906, pp. 498–511, 2002.

[2] N. Kriek, L. Tilley, P. Horrocks, et al., "Characterization of the pathway for transport of the cytoadherence-mediating protein, PfEMP1, to the host cell surface in malaria parasite-infected erythrocytes," *Molecular Microbiology*, vol. 50, no. 4, pp. 1215–1227, 2003.

[3] J. D. Smith, G. Subramanian, B. Gamain, D. I. Baruch, and L. H. Miller, "Classification of adhesive domains in the *Plasmodium falciparum* erythrocyte membrane protein 1 family," *Molecular and Biochemical Parasitology*, vol. 110, no. 2, pp. 293–310, 2000.

[4] J. G. Beeson, S. J. Rogerson, B. M. Cooke, et al., "Adhesion of *Plasmodium falciparum*-infected erythrocytes to hyaluronic acid in placental malaria," *Nature Medicine*, vol. 6, no. 1, pp. 86–90, 2000.

[5] I. W. Sherman, S. Eda, and E. Winograd, "Cytoadherence and sequestration in *Plasmodium falciparum*: defining the ties that bind," *Microbes and Infection*, vol. 5, no. 10, pp. 897–909, 2003.

[6] N. Rasti, M. Wahlgren, and Q. Chen, "Molecular aspects of malaria pathogenesis," *FEMS Immunology and Medical Microbiology*, vol. 41, no. 1, pp. 9–26, 2004.

[7] D. Donati, L. P. Zhang, Q. Chen, et al., "Identification of a polyclonal B-cell activator in *Plasmodium falciparum*," *Infection and Immunity*, vol. 72, no. 9, pp. 5412–5418, 2004.

[8] J. D. Smith, B. Gamain, D. I. Baruch, and S. Kyes, "Decoding the language of *var* genes and *Plasmodium falciparum* sequestration," *Trends in Parasitology*, vol. 17, no. 11, pp. 538–545, 2001.

[9] C. Russell, O. Mercereau-Puijalon, C. Le Scanf, M. Steward, and D. E. Arnot, "Further definition of PfEMP-1 DBL-1α domains mediating rosetting adhesion of *Plasmodium falciparum*," *Molecular and Biochemical Parasitology*, vol. 144, no. 1, pp. 109–113, 2005.

[10] A. D. Ozarkar, D. Prakash, D. N. Deobagkar, and D. D. Deobagkar, "Distinct polymorphism of DBLα domain of the *var* gene in laboratory-cultured and clinical field isolates of *Plasmodium falciparum* malaria parasites," *Current Science*, vol. 93, no. 2, pp. 219–223, 2007.

[11] S. Kyes, H. Taylor, A. Craig, K. Marsh, and C. Newbold, "Genomic representation of *var* gene sequences in *Plasmodium falciparum* field isolates from different geographic regions," *Molecular and Biochemical Parasitology*, vol. 87, no. 2, pp. 235–238, 1997.

[12] C. P. Ward, G. T. Clottey, M. Dorris, D.-D. Ji, and D. E. Arnot, "Analysis of *Plasmodium falciparum* PfEMP-1/*var* genes suggests that recombination rearranges constrained sequences," *Molecular and Biochemical Parasitology*, vol. 102, no. 1, pp. 167–177, 1999.

[13] H. M. Taylor, S. A. Kyes, and C. I. Newbold, "*Var* gene diversity in *Plasmodium falciparum* is generated by frequent recombination events," *Molecular and Biochemical Parasitology*, vol. 110, no. 2, pp. 391–397, 2000.

[14] E. V. Fowler, J. M. Peters, M. L. Gatton, N. Chen, and Q. Cheng, "Genetic diversity of the DBLα region in *Plasmodium falciparum var* genes among Asia-Pacific isolates," *Molecular and Biochemical Parasitology*, vol. 120, no. 1, pp. 117–126, 2002.

[15] N. Ikenoue, S. Kawazu, T. Hatabu, E. A. Villacorte, P. T. Rivera, and S. Kano, "PCR-amplification, sequencing, and comparison of the *var*/PfEMP-1 gene from the blood of patients with falciparum malaria in the Philippines," *The Southeast Asian Journal of Tropical Medicine and Public Health*, vol. 33, supplement 3, pp. 8–13, 2002.

[16] P. C. Bull, M. Berriman, S. Kyes, et al., "*Plasmodium falciparum* variant surface antigen expression patterns during malaria," *PLoS Pathogens*, vol. 1, no. 3, article e26, pp. 202–213, 2005.

[17] L. Albrecht, E. F. Merino, E. H. E. Hoffmann, et al., "Extense variant gene family repertoire overlap in Western Amazon *Plasmodium falciparum* isolates," *Molecular and Biochemical Parasitology*, vol. 150, no. 2, pp. 157–165, 2006.

[18] A. R. Trimnell, S. M. Kraemer, S. Mukherjee, et al., "Global genetic diversity and evolution of *var* genes associated with placental and severe childhood malaria," *Molecular and Biochemical Parasitology*, vol. 148, no. 2, pp. 169–180, 2006.

[19] H. M. Kyriacou, G. N. Stone, R. J. Challis, et al., "Differential *var* gene transcription in *Plasmodium falciparum* isolates from patients with cerebral malaria compared to hyperparasitaemia," *Molecular and Biochemical Parasitology*, vol. 150, no. 2, pp. 211–218, 2006.

[20] S. M. Kraemer, S. A. Kyes, G. Aggarwal, et al., "Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates," *BMC Genomics*, vol. 8, article 45, pp. 1–18, 2007.

[21] A. E. Barry, A. Leliwa-Sytek, L. Tavul, et al., "Population genomics of the immune evasion (*var*) genes of *Plasmodium falciparum*," *PLoS Pathogens*, vol. 3, no. 3, article e34, pp. 1–9, 2007.

[22] T. Chookajorn, M. S. Costanzo, D. L. Hartl, and K. W. Deitsch, "Malaria: a peek at the *var* variorum," *Trends in Parasitology*, vol. 23, no. 12, pp. 563–565, 2007.

[23] R. Chattopadhyay, A. Sharma, V. K. Srivastava, et al., "*Plasmodium falciparum* infection elicits both variant-specific and cross-reactive antibodies against variant surface antigens," *Infection and Immunity*, vol. 71, no. 2, pp. 597–604, 2003.

[24] W. Trager and J. B. Jensen, "Human malaria parasites in continuous culture," *Science*, vol. 193, no. 4254, pp. 673–675, 1976.

[25] D. E. Arnot, C. Roper, and R. A. L. Bayoumi, "Digital codes from hypervariable tandemly repeated DNA sequences in the *Plasmodium falciparum* circumsporozoite gene can genetically barcode isolates," *Molecular and Biochemical Parasitology*, vol. 61, no. 1, pp. 15–24, 1993.

[26] P. Shil, N. Dudani, and P. B. Vidyasagar, "ISHAN: sequence homology analysis package," *In Silico Biology*, vol. 6, no. 5, pp. 373–377, 2006.

[27] A. D. Ozarkar, D. Prakash, D. N. Deobagkar, and D. D. Deobagkar, "Prediction of B cell and T cell epitopes of DBLα domain in *Plasmodium falciparum* malaria vaccine candidate *var* gene," *Protein and Peptide Letters*, vol. 14, no. 6, pp. 528–530, 2007.

[28] H. Harris and D. A. Hopkinson, *Handbook of Enzyme Electrophoresis in Human Genetics*, North-Holland, Amsterdam, The Netherlands, 1976.

[29] K. Kirchgatter and H. A. del Portillo, "Association of severe noncerebral *Plasmodium falciparum* malaria in Brazil with expressed PfEMP1 DBL1α sequences lacking cysteine residues," *Molecular Medicine*, vol. 8, no. 1, pp. 16–23, 2002.

[30] L. H. Freitas-Junior, E. Bottius, L. A. Pirrit, et al., "Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*," *Nature*, vol. 407, no. 6807, pp. 1018–1022, 2000.

[31] M. Frank, L. Kirkman, D. Costantini, et al., "Frequent recombination events generate diversity within the multicopy variant gene families of *Plasmodium falciparum*," *International Journal for Parasitology*, vol. 38, no. 10, pp. 1099–1109, 2008.

[32] G. Winter, Q. Chen, K. Flick, P. Kremsner, V. Fernandez, and M. Wahlgren, "The 3D7*var*5.2 (^{var}COMMON) type *var* gene family is commonly expressed in non-placental *Plasmodium falciparum* malaria," *Molecular and Biochemical Parasitology*, vol. 127, no. 2, pp. 179–191, 2003.

[33] J. D. Smith, A. G. Craig, N. Kriek, et al., "Identification of a *Plasmodium falciparum* intercellular adhesion molecule-1 binding domain: a parasite adhesion trait implicated in cerebral malaria," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 4, pp. 1766–1771, 2000.

[34] S. M. Kraemer and J. D. Smith, "A family affair: *var* genes, PfEMP1 binding, and malaria disease," *Current Opinion in Microbiology*, vol. 9, no. 4, pp. 374–380, 2006.

[35] S. M. Kraemer and J. D. Smith, "Evidence for the importance of genetic structuring to the structural and functional specialization of the *Plasmodium falciparum var* gene family," *Molecular Microbiology*, vol. 50, no. 5, pp. 1527–1538, 2003.

[36] T. Lavstsen, A. Salanti, A. T. R. Jensen, D. E. Arnot, and T. G. Theander, "Sub-grouping of *Plasmodium falciparum* 3D7 *var* genes based on sequence analysis of coding and non-coding regions," *Malaria Journal*, vol. 2, no. 1, article 27, pp. 1–14, 2003.

[37] B. A. Robinson, T. L. Welch, and J. D. Smith, "Widespread functional specialization of *Plasmodium falciparum* erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome," *Molecular Microbiology*, vol. 47, no. 5, pp. 1265–1278, 2003.

[38] M. Rottmann, T. Lavstsen, J. P. Mugasa, et al., "Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children," *Infection and Immunity*, vol. 74, no. 7, pp. 3904–3911, 2006.

[39] A. Barragan, V. Fernandez, Q. Chen, A. von Euler, M. Wahlgren, and D. Spillmann, "The Duffy-binding-like domain 1 of *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) is a heparan sulfate ligand that requires 12 mers for binding," *Blood*, vol. 95, no. 11, pp. 3594–3599, 2000.

[40] A. M. Vogt, F. Pettersson, K. Moll, et al., "Release of sequestered malaria parasites upon injection of a glycosaminoglycan," *PLoS Pathogens*, vol. 2, no. 9, article e100, pp. 853–863, 2006.

[41] G. A. Heavner, "Active sequences in cell adhesion molecules: targets for therapeutic intervention," *Drug Discovery Today*, vol. 1, no. 7, pp. 295–304, 1996.

[42] D. I. Baruch, B. L. Pasloske, H. B. Singh, et al., "Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes," *Cell*, vol. 82, no. 1, pp. 77–87, 1995.

[43] M. Ho and N. J. White, "Molecular mechanisms of cytoadherence in malaria," *American Journal of Physiology*, vol. 276, no. 6, pp. C1231–C1242, 1999.

[44] P. Michon, J. R. Stevens, O. Kaneko, and J. H. Adams, "Evolutionary relationships of conserved cysteine-rich motifs in adhesive molecules of malaria parasites," *Molecular Biology and Evolution*, vol. 19, no. 7, pp. 1128–1142, 2002.