Review Article

Revisiting the multigene families: *Plasmodium var* and *vir* genes

Vineeta Singh¹, Purva Gupta¹ & Veena Pande²

¹National Institute of Malaria Research (ICMR), Dwarka, New Delhi; ²Department of Biotechnology, Kumaun University, Nainital, Uttarakhand, India

ABSTRACT

Malaria is an infectious disease that is widespread in tropical and subtropical regions. The malaria parasite is able to skip the host immunity and thus maintains not only persistent but also repeated infections. There are a number of multigene families in *Plasmodium* that code for the variant antigens and are targets for protective immunity. In this article, we summarize the virulence genes of *P. falciparum* (*var* genes) and *P. vivax* (*vir* genes) which play key roles in disease pathogenesis by evading elimination by the host immune system. These genes occurring within the parasite population are mostly present in the subtelomeric regions of the chromosome.

Key words Multigene family; *Plasmodium*; *var* and *vir* genes

Malaria parasites

Malaria, one of the most widespread infectious diseases in humans is caused by the apicomplexan *Plasmodium falciparum*, *P. vivax*, *P. malariae*, *P. ovale* and *P. knowlesi*. Most severe form of malaria is caused by *P. falciparum* where the infected red blood cells (RBCs) have distinct antigenic properties generally with cytoadherent phenotypes¹. *Plasmodium vivax* is not fatal, though it is responsible for considerable morbidity in those populations where it is endemic through hypnozoites (the dormant parasite stages in liver) causing relapse within weeks to months after primary episode². The diverse *Plasmodium* genome poses a formidable challenge till date to the scientists community worldwide.

Diverse *Plasmodium* genome

The completion of *P. falciparum* and *P. vivax* genome project has revealed many contrasting features between the two genomes along with several similarities seen at the gene function level distributed among 14 chromosomes³⁻⁴. The genome of *P. vivax* is 26.8 megabase (Mb) distinctly larger than *P. falciparum* genome of 23.3 Mb. Both the species are almost similar in terms of the number of genes where *P. vivax* has 5433 and *P. falciparum* has 5403 genes. The average gene length in *P. vivax* is 2164 bp whereas in *P. falciparum* the gene length is 2283 bp. The average intron length in *P. vivax* is 192 bp and in *P. falciparum* it is 179 bp though the average intergenic length between them is ~1994 and 1745 bp, respectively. The difference in the average intron and intergenic regions could be the reason for the larger genome size of *P. vivax* as the coding regions of both the species are considerably comparable. The major difference in the two genomes is the A+T nucleotide content found to be 57.7% in *P. vivax* and 80.6% in *P. falciparum*⁴. The distribution of A+T rich regions varies in both the genomes by being restricted to the subtelomeric areas of the chromosomes in *P. vivax* and evenly distributed in *P. falciparum* similar to the multigene families encoding variant surface antigens (VSAs) following the same pattern of distribution in *P. vivax* and *P. falciparum*.

Antigenic variation in *Plasmodium*

Disease pathogenesis in malaria is related to the ability of the parasitized RBCs to escape immune response and establish chronic infections⁵⁻⁶. The infection of RBCs by *Plasmodium* results in progressive and dramatic structural and biochemical modifications of the RBCs that can worsen into life-threatening complications of malaria⁷. The phenomenon where parasite exhibits variable antigens on the surface of infected erythrocytes enables the parasite to endure and escape the host immune response known as antigenic variation. The symmetry of interactions between different antigenic variants can be explained by various mathematical models of antigenic variation in malaria⁸.

The expression of VSAs has been linked to disease pathogenesis besides the survival of parasite by establishing long-lasting chronic infections⁵. Most species of *Plasmodium* have several multicopy gene families situated at the subtelomeric ends of chromosomes coding for
VSAs that are exported to the surface of the infected host erythrocyte to help the parasite evade the immune system. These multigene families are also highly diverged between species, and undergo high levels of recombination, generating further diversity.

**Multigene families of Plasmodium**

Several *Plasmodium* species contain multigene families on the telomeric and subtelomeric regions of their chromosomes which code for VSAs. The largest multigene family recognized so far is the *Plasmodium* interspersed repeats (pir) which includes repetitive interspersed family (rif) in *P. falciparum*, var in *P. vivax*, pir in *P. knowlesi* and the cir/bir/yir family in three rodent malaria parasites *P. chabaudi*/*P. berghei*/*P. yoelli*, respectively. Since, the completion and annotation of the genome sequence of *P. falciparum*, three multigene families have been identified, i.e. the var genes encoding *P. falciparum* erythrocyte membrane protein 1 (PfEMP1), the rif encoding the RIFIN proteins and the subtelomeric variant open reading frame (stevor) genes coding for the STEVOR proteins. The multigene family in *P. vivax* is variant interspersed repeats (vir) which encodes for VIR proteins. These multigene families are responsible for the antigenic variation of the parasite in natural infections which enables it to escape the host immune response for its survival resulting in prolonged chronic infections. Theoretically, these genes can be considered as effective vaccine candidates which could have a great impact on malaria morbidity and mortality subsequently.

**The var gene family:** The best characterised multigene family in *P. falciparum* is the var gene family which encodes PfEMP1, a protein that is exported to the ‘knob like’ binding structures on the surface of infected erythrocytes with a key role in antigenic variation and cytoadherence. A total of 60 var genes are known to be present in single haploid genome of *P. falciparum* though only a single PfEMP1 type is expressed on the surface of the parasitized RBC stably inherited through successive cell cycles or a different gene can be expressed by switching during the course of infection. The completion of the genome project revealed 35 of the var genes in the subtelomeric regions and the remaining genes in clusters away from the ends of the chromosomes. Var gene repertoires have been extensively sequenced and worked on in 3D7 *P. falciparum* reference clone. The var genes are large (6–13 kb) and comprise of two exons where the first exon codes for the variable extracellular domain and transmembrane region, and the second exon for a highly conserved cytoplasmic or acidic terminal segment (ATS). The var genes present in the central regions of chromosomes can either be present singly or in groups which are arranged in tandem arrays with three to seven var genes. Each chromosome end contains one to three var genes followed by a group of rif, stevor and other gene families (Fig. 1). In the *P. falciparum* infection, cytoadhesion of the parasitized blood cells to the endothelial cells is due to the interactions between PfEMP1 encoded by the var gene family and defined host receptors on endothelial cells like cluster of differentiation-36 (CD36), intercellular adhesion molecule-1 (ICAM-1), and chondroitin sulfate A (CSA). Multiple receptor-like domains, the duffy binding-like (DBL) domain and the cysteine rich inter domain region (CIDR) are present in the binding regions of PfEMP1. Although, there is a substantial variation in the sequence of the PfEMP1 proteins, the adhesion domains can be grouped on the basis of the sequence similarity and the domain design can offer insights in the binding function. There are many different DBL-γ domains binding to CSA and several CIDR-α type domains which bind to CD36.

Var gene ancient sequence fragments termed homology blocks (HBs) recombine at very high rate and expression rates of some HBs are associated more strongly with the severe disease phenotypes than the expression rates of DBL-α types. More specifically, expression profiles of HB differ significantly for severe and mild disease and also for rosetting vs impaired consciousness associated severe disease.

The total number of functionally characterized PfEMP1 domains are still very limited and it is clear that the binding functionality of PfEMP1 can not be completely known by sequence analysis alone.

Based on sequence similarities the 5′ promoter regions of the var genes are divided into five distinct groups, i.e. Upstream—UpsA, UpsB, UpsC, UpsD and UpsE. In

![Fig. 1: Schematic representation of *P. falciparum* var, rif and stevor genes on the chromosome 3 (the gene sizes are not to scale).](image-url)
the 3D7 genome, there are nine UpsA var genes, 22 UpsB var genes, 13 UpsC var genes, one UpsD var pseudogene and one UpsE var gene. The remaining 13 var genes comprise of UpsBC as their promoter sequences, i.e. they are phylogenetically between UpsB and UpsC\(^{16}\). UpsD is grouped with UpsA and UpsC var genes located centrally in the chromosome, UpsB var genes are either central or subtelomeric and transcribed away from the telomere arranged in tandem arrays with other UpsB or UpsC var genes whereas UpsA and UpsE var genes are subtelomeric and transcribed towards the telomere in opposite direction to the UpsB genes\(^{17}\). In the group, UpsA var genes are more closely related with each other and all encode non-CD36-binding type CIDR domains whereas, UpsB and UpsC genes have CD36-binding CIDR domains. The UpsD and UpsE-linked var genes, viz. var1CSA and var2CSA and the UpsA-linked Type 3 var are exceptionally conserved in parasite isolates though the functional importance of the various promoters remains unclear\(^{15, 18–19}\). Var gene transcription is tightly controlled by a special transcription mechanism that silences all the other var genes except the one that is being expressed per genome and this process could be controlled epigenetically\(^{20}\). This frequent switching of the var genes ensures parasite survival against the host immune system and is one of the major causes of the severity of \(P. falciparum\) malaria indicating var genes as an important research area for vaccine development. Gene recombination between var paralogs is the major mechanism in generation of the extreme diversity in the variant antigen repertoire\(^{21}\).

The vir gene family: The similar multigene superfamily in \(P. vivax\) is vir gene family, which like the var genes in \(P. falciparum\) might have a key role in antigenic variation. The annotation of the vir gene repertoire from \(P. vivax\) Sal-I genome, has revealed a total of 346 var genes including fragments, and pseudogenes\(^{5, 22}\). These genes range from 156–3434 bp in size and each gene has 1–5 exons where the first exon lacks the signal peptide sequences. The second exon is highly polymorphic containing the predicted transmembrane domain and conserved cysteine residues and the third exon is uniform in size encoding for the putative cytosolic domain (Fig. 2). The vir gene region between the second and the third exon is well-conserved. The vir gene family corresponds to 10% of the coding sequences comprising of 12 subfamilies christened from A to L varying in their extent of allele polymorphism and the remaining 82 genes could not be clustered. Unlike the var genes, the vir genes are not found in the internal regions of the chromosomes but are exclusively subtelomeric in location. VIR proteins encoded by the vir genes in natural infections are not known fully but it is speculated that they have a role to play in spleen-specific cytoadherence and in the chronicity of the disease\(^{23}\). It has also been suggested that the vir genes may have different functions in immune evasion as not all the proteins are presented on the infected RBCs\(^{24}\). Unlike the var genes of \(P. falciparum\), vir genes are abundantly expressed at a given time and a number of subfamilies are transcribed by individual parasites\(^{25}\).

Vir genes are the largest subtelomeric multigene family in \(P. vivax\) which might be responsible for bringing \(P. vivax\) malaria towards severity as \(P. vivax\)-infected erythrocytes also have some ability to cause cytoadherence\(^{25}\). The vir genes are extensively more diverse than other \(Plasmodium\) pir families such as cir (\(P. chabaudi\) 135 members) and bir (\(P. berghei\) 245 members) though vir superfamily shares structural characteristics with PfmC-2tm and surfin family found on the infected erythrocytes\(^5\). Many VIR proteins lack Plasmodium export element (PEXEL) motifs showing that these proteins have subcellular localizations other than the surface membranes of infected reticulocytes, however, further investigations are needed to establish the role of VIR proteins as cytoadhesive ligands and the part they play in the severity of the disease in the changing paradigm of pathogenesis.

DISCUSSION

The complete annotation of the \(P. falciparum\) genome showed two other polymorphic multigene families, viz. rif and stevor to be associated with var in the genome\(^{11}\). The rif gene family is located subtelomerically in the chromosomes, very close to the var genes (Fig. 1). It is not known how many RIFIN proteins are expressed on the surface of an RBC at a time but it is found that they are transcribed by immature trophozoites only for a short period of time\(^{26}\). More than 200 rif genes are found per haploid \(P. falciparum\) genome which shows that the rep-
ertificate of rif genes is much larger than that of var genes. Nevertheless, evidence has been found that different RIFINs are being expressed on different parasite lines indicating that rif genes might have an additional role in antigenic variation and in immune evasion as well.

Located along the var and rif genes, stevor genes are more conserved in comparison but the hypervariable region of STEVOR has significant sequence diversity in different P. falciparum lines. While many var and rif genes are centrally located, the stevor genes are restricted to the chromosome ends. STEVOR appears on the surface of infected RBCs (iRBCs) after both PfEMP1 and RIFIN demonstrating its significant role in the development of late stage parasites.

It has been shown in the previous studies that STEVOR proteins are clonally variable on the surface of the RBC and are able to change the antigenic properties of the RBCs, thus playing a major role in the antigenic variation of the late asexual parasite stages. Similar to rif genes, the role of stevor genes in cytoadhesion is unknown.

The var gene repertoire is highly diverse in a single parasite and extensive diversity is also seen between isolates causing the prevailing immense global diversity in the var genes. Perhaps this existing diversity is responsible for host immune evasion in the parasite. The var genes encode PfEMP1 responsible for the cytoadhesive properties in P. falciparum, thus providing the selective advantage of preventing their clearance in the spleen.

The PfEMP1 protein (200–400 kDa) remains the major target for naturally acquired antibodies and PfEMP1 variants expression is more often associated in children’s plasma causing severe malaria than the non-severe disease. The chronic disease is closely associated with a shift in PfEMP1 expression and this kind of shift is responsible for expressing PfEMP1 molecules which are less optimal for adhesion. The PfEMP1 domain could serve as an ideal vaccine target though the extensive diversity present in PfEMP1 could be an obstacle but the domains which bind to CD36 and CSA are structurally conserved and could serve as potential vaccine targets aimed at cytoadhesion.

It was observed that disruption of the var2CSA gene causes the parasite to lose its CSA-binding phenotype. There is ongoing research to evaluate var2CSA as a vaccine target against malaria.

The expression of var genes and its correlation with the disease severity can reveal the role of different var genes in the pathogenesis of the disease. The recent understanding of antigenic variation after analyzing the var genes gives more insight into the mechanisms of immune evasion, hence promoting pathogenesis.

Studies indicate the vir genes relation with antigenic variation due to which the parasite survives elimination by the host. Recent sequence analysis of four vir genes from Indian populations revealed high diversity in them in natural infections both within and between the isolates.

Plasmodium falciparum is known to cause cytoadherence which is one of the prime reasons for severity of the disease.

Has 5403 genes

Has the ability to evade the host immunity by antigenic variation.

---

Snapshot on similarities and dissimilarities between P. falciparum and P. vivax

<table>
<thead>
<tr>
<th>Similarities</th>
<th>P. falciparum</th>
<th>P. vivax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causes severe and complicated malaria</td>
<td></td>
<td>Though it is said to be benign, situation seems to be changing and it is becoming severe</td>
</tr>
<tr>
<td>Resistance to chloroquine well-established</td>
<td></td>
<td>Resistance to chloroquine proven in several parts of the world</td>
</tr>
<tr>
<td>Plasmodium falciparum is known to cause cytoadherence which is one of the prime reasons for severity of the disease</td>
<td>Has 5403 genes</td>
<td>Recently, P. vivax severity is attributed possibly to cytoadherent phenomenon</td>
</tr>
<tr>
<td>Has 5403 genes</td>
<td></td>
<td>Has 5433 genes</td>
</tr>
<tr>
<td>High conservation at functional level exists between two species.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eighty-two percent of the genes in P. falciparum with known functions are conserved in P. vivax</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has the ability to evade the host immunity by antigenic variation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(contd...)
Singh et al: Multigene families of *Plasmodium*

---

**Dissimilarities**

*Plasmodium falciparum* has an African origin and is more recent\(^36\).

*Plasmodium falciparum* is prevalent in Africa and parts of Asia.

<table>
<thead>
<tr>
<th><strong>P. falciparum</strong></th>
<th><strong>P. vivax</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissimilarities</td>
<td></td>
</tr>
<tr>
<td><em>Plasmodium falciparum</em> has an African origin and is more recent(^36).</td>
<td><em>Plasmodium vivax</em> has origin in Asia and is more ancient(^36).</td>
</tr>
<tr>
<td><em>Plasmodium falciparum</em> is prevalent in Africa and parts of Asia.</td>
<td><em>Plasmodium vivax</em> is most commonly distributed in Asia and in south and central America.</td>
</tr>
</tbody>
</table>

Invades all RBCs

Enters host RBCs via multiple pathways.

Gametocytes appear in peripheral blood after clinical symptoms\(^38\).

No dormant stages found.

There are no clinical relapses.

Avoids passage of mature asexual blood stages through the spleen\(^19\).

Continuous *in vitro* culture for *P. falciparum* available.

Isochores not seen.

A+T nucleotide content is 80.6%

A+T regions are restricted to the subtelomeric areas of the chromosome.

Contains multigene families like var, rif and stevor genes.

There are 60 var genes known to be present in *P. falciparum* along with 200 members of rif and 28 members of stevor genes\(^27–28\).

At a time only one var gene is expressed\(^13\).

The var genes are clonally expressed in natural infections.

The var genes have two exons.

There is no grouping of var genes into subfamilies.

Established role of var genes in cytoadherence\(^12\).

The var genes are present in central as well as the subtelomeric regions of the chromosomes.

---

The detailed knowledge of genetic variability among the vir genes will provide an insight in the chronicity of the disease. The most challenging task for malaria researchers is to develop an efficient vaccine. Now that the vir gene sequences are known, DNA microarrays can be designed to look at how the expression of the genes changes. Such experiments will help identify the regulation mechanism of these genes and their functional role in the chronicity of the disease. Furthermore, it will also enable to identify the expression of the genes at different life-cycle stages of the parasite. The challenge to malariologists is to develop innovative ways to tackle this disease. The information about the vir genes should be used to test a suitable approach for developing a vir-based malaria vaccine since it is speculated to play a role in pathogenesis. It remains to be seen whether the emerging *P. falciparum* like characteristics of *P. vivax* associated with VIR proteins are responsible for *P. vivax* shifting from benign to severe. The var and the vir genes repertoire encompass a huge range of evolutionary strategies and constraints, representing an interesting model system to study in depth the host-pathogen evolution.
ACKNOWLEDGEMENTS

We would like to thank the Department of Biotechnology (DBT) and also like to extend thanks to Indian Council of Medical Research (ICMR), for providing us the grant required for carrying out all the research work during which the genesis of this paper took shape.

REFERENCES

7. Calderwood MS, Gannoun-Zaki L, Wellemes TE, Deitsch KW. 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 2003; 278: 34125–32.


**Correspondence to:** Dr Vineeta Singh, Scientist ‘C’, National Institute of Malaria Research (ICMR), Sector–8, Dwarka, New Delhi–110 077, India.

E-mail: vineetas_2000@yahoo.com

**Received:** 8 July 2013 **Accepted in revised form:** 3 April 2014