Parasite Evolutionary Genomics

Malaria Parasite Genomics and Evolution of Drug Resistance

Development of Nuclear DNA Markers for Evolutionary Studies in *Plasmodium falciparum*

Recent researches in evolutionary genetics have revealed that estimation of genetic diversity is strongly dependent on the genetic markers used, thus making appropriate evolutionary inference at species and gene levels difficult. Considering these facts, we have used published whole-genome sequence information to develop nuclear DNA markers in the human malaria parasite, *P. falciparum* that would help in understanding the precise roles of demography and natural selection in the evolution of *P. falciparum*. For designing the putatively neutral DNA fragments (that bear no or very weak signals of past selection events) which would help in deciphering the demographic history of *P. falciparum*, we scanned the whole genome of *P. falciparum* available in the public domain and isolated introns in every kind of gene (known, putative and hypothetical). Specifically, we considered introns of 450–850 bp long and designed primers in the exons flanking these introns (exon priming intron crossing, or EPIC fragments). We could only predict 170 introns that could be considered as partially putatively neutral (Fig. 19). Further, we have divided the whole *pfcrt* gene into three different fragments for amplification by PCR and five internal fragments (within the three main fragments) for sequencing purposes. The development of nuclear DNA markers has far-reaching significance in malaria research.

Comparative Evolutionary Genetic Insights into *Plasmodium falciparum* Functional Genes with Reference to *P. vivax* genome

Complex and rapidly evolving behaviours of the two human malaria parasites, *P. falciparum* and *P. vivax* have always been mysterious to the evolutionary biologists as the former is the most virulent and the later is most prevalent malaria parasite species across the globe. With the availability of whole genome sequence data, it is now feasible to pinpoint genomic similarities and differences between the parasites with the comparative evolutionary genetic approaches, and thus, define new measures for malaria control. We herewith utilized available genome information of these two species and compared functional genes of *P.*

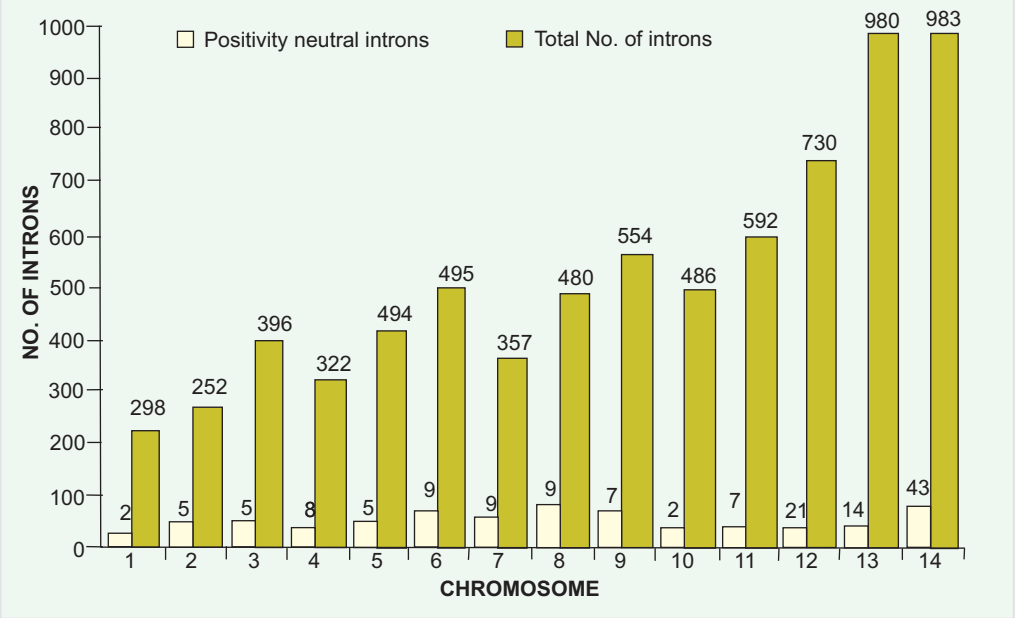

Fig. 19: Distribution of total number of introns and selected putatively neutral introns in *Plasmodium falciparum* chromosomes

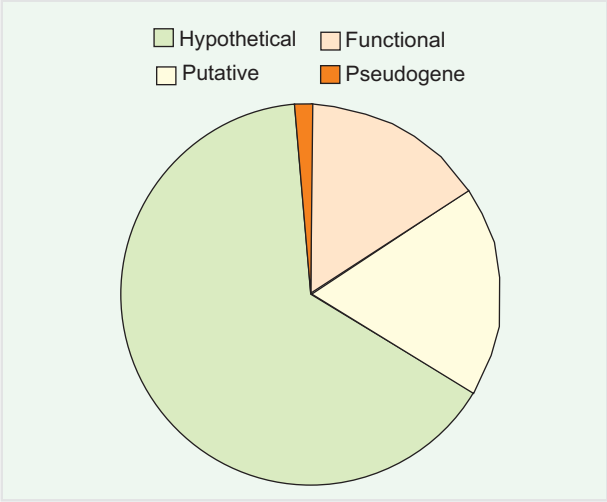

Fig. 20: Total genes in *Plasmodium falciparum* genome. Note that the putative and functional categories (together considered as functional) of genes have been utilized in the present study

falciparum with partially-assembled whole genome sequences of *P. vivax*. About 82% of total functional genes of *P. falciparum* were found to be conserved in *P. vivax* and rest 18% to be unique to *P. falciparum*. Although both types of genes were distributed across all 14 chromosomes of *P. falciparum*, the distribution was slightly biased towards two separate chromosomes for each category (Fig. 20). About a half of the conserved genes was intron-less, whereas almost all unique genes have introns. However, number of introns was comparatively higher (usually >2) in the intron-possessing conserved genes than in the unique genes (mostly <2). Statistically significant positive correlations between total intron length and gene lengths were detected in 11 chromosomes for unique genes, whereas only in three chromosomes for conserved genes. Three most conserved genes (Actin, Elongation factor alpha 1 and Ribosomal protein L 10 putative) between *P. falciparum* and *P. vivax* were found to be highly conserved in four other species of *Plasmodium* (except Actin gene in *P. chabaudi*) and were mostly intron-less. Phylogenetic trees were constructed separately for each of the three genes; in two genes (Actin and Elongation factor alpha 1) different *Plasmodium* species were placed in almost similar positions, whereas Ribosomal protein L 10 putative show different relationships between *Plasmodium* species. Three unique gene families in three *Plasmodium* species (*P. falciparum*, *P. vivax* and *P. knowlesi*) were studied in detail for total intron length and correlations between intron lengths and gene lengths, which corroborate findings on the overall patterns of whole unique genes of *P. falciparum*. The results are discussed in terms of chromosome and intron evolution in *Plasmodium* in general, relevance of introns in differential functions of *P. falciparum* genes and genetic similarities and differences between *P. falciparum* and *P. vivax* and its implications in malaria, in particular.

Fine-scale Genetic Characterization of *Plasmodium falciparum* Chromosome 7 Encompassing the antigenic *var* and the Drug-Resistant *pfcrt* Genes

The fact that malaria is still an uncontrolled disease is reflected by the genetic organization of the parasite genome. Efforts to curb malaria should begin with proper understanding of the mechanism by which the parasites evade human immune system and evolve resistance to different antimalarial drugs.

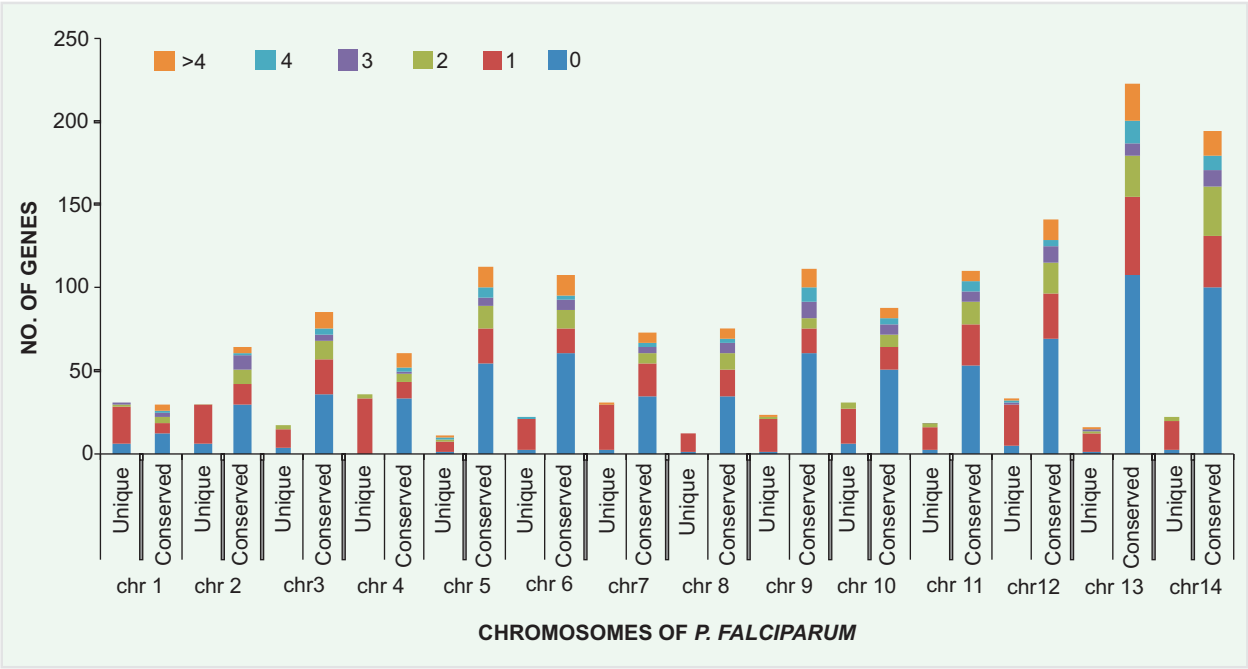

Fig. 21: Distribution of genes with different intron numbers in conserved and unique genes in chromosomes of *Plasmodium falciparum*

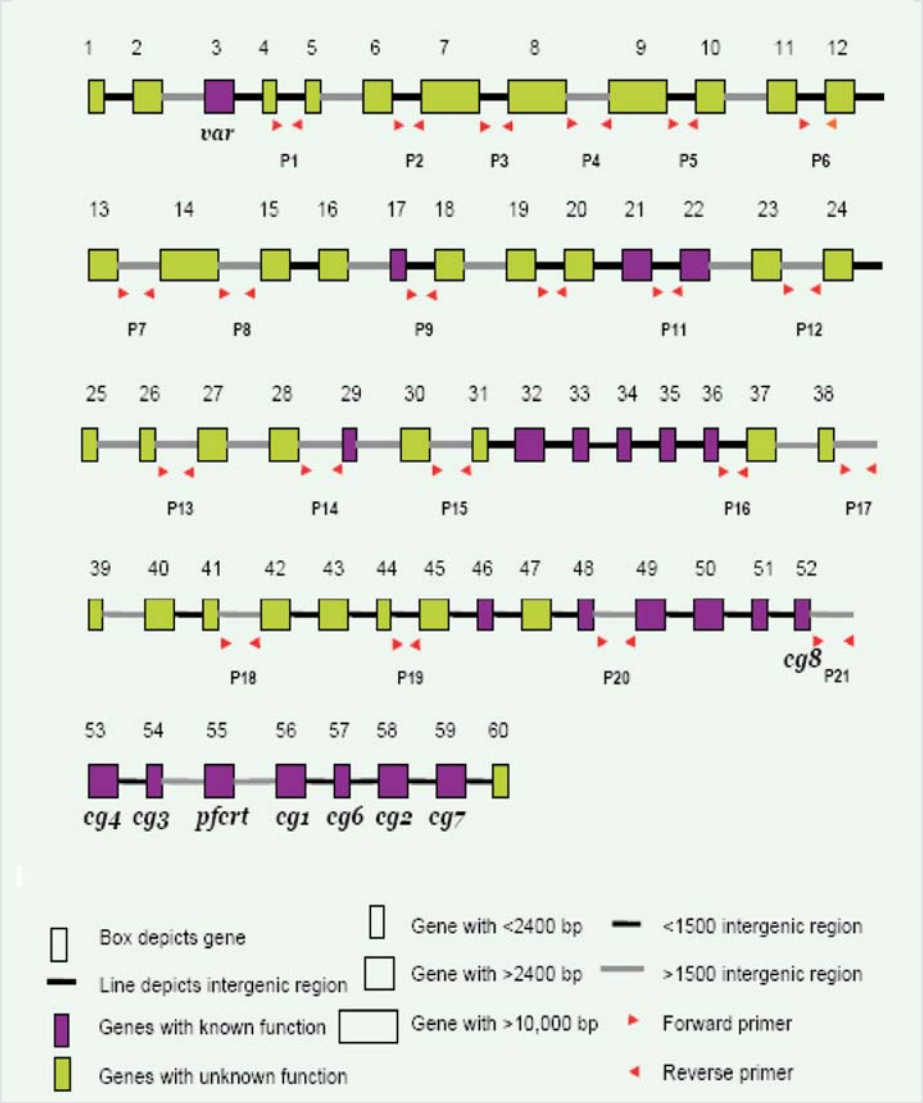

Fig. 22: Fine details of the specified region showing the genes, intergenic regions and locations of the primers designed for amplification of different regions

We have initiated such a study and presented here with the results from the *in silico* understanding of a seventh chromosomal region of the malaria parasite, *P. falciparum* encompassing the antigenic *var* genes (coding *pfemp1*) and the drug-resistant gene *pfcrt* located at a specified region of the chromosome 7. We found 60 genes of various functions and lengths, majority (61.67%) of them were performing known functions. Almost all the genes have orthologs in other four species of *Plasmodium*, of which *P. chabaudi* seems to be the closest to *P. falciparum*. However, only two genes were found to be paralogous. Interestingly, the drug-resistant gene, *pfcrt* was found to be surrounded by seven genes coding for several CG proteins out of which six were reported to be responsible for providing drug resistance to *P. vivax*. The intergenic regions, in this specified region were generally large in size, majority (73%) of them were of more than 500 nucleotide bp length. We also designed primers for amplification of 21 non-coding DNA fragments in the whole region for estimating genetic diversity and inferring the evolutionary history of this region of *P. falciparum* genome. (Figs. 21 and 22).

Evolutionary Paradigm of Chloroquine-Resistant Malaria in India

Drug pressure in the field is believed to be responsible for the emergence of drug-resistant *P. falciparum*. Variants of the *P. falciparum* chloroquine resistance transporter (*pfcrt*) gene have been shown to be responsible for conferring resistance to the commonly used drug chloroquine. In particular, an amino acid mutation, K76T, was shown to have a strong positive correlation with the chloroquine resistant varieties of malaria parasites. Global studies have reported highly reduced genetic diversity surrounding K76T in the pfcrt gene, which indicates that the mutation has been a target of positive Darwinian natural selection. However, two recent studies of *P. falciparum* in India found high genetic diversity in the *pfcrt* gene, which, at first sight, do not support the role of natural selection in the evolution of chloroquine resistance in India (Fig. 23).

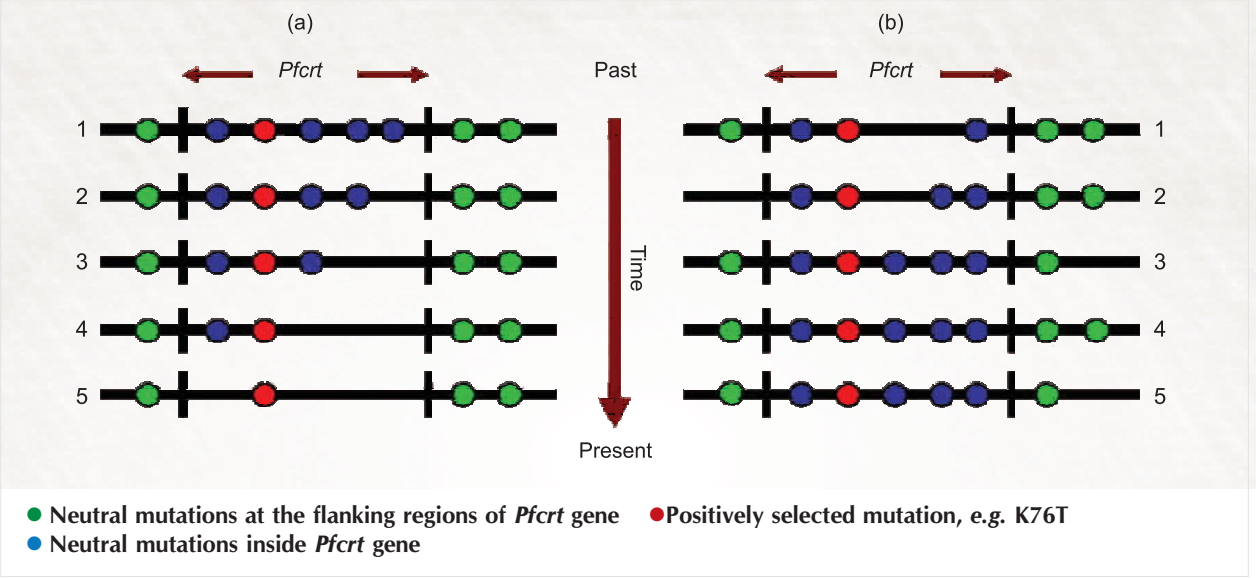

Fig. 23: Evolutionary pattern of the *Pfcrt* gene in chloroquine-resistant *Plasmodium falciparum* populations. (a) Over generations and drug selection pressure, the positively selected mutation (e.g. K76T: red circles) is swept away by natural selection along with the linked neutral genetic variations (blue circles) (1–5). However, variations that are distant from the selected mutation are unaffected by the selective sweep. Thus, neutral variations in the flanking regions of *Pfcrt* (green circles) are maintained (5). (b) An opposite evolutionary trend in which neutral mutations progressively accumulate over time (1–5) in *Pfcrt*. This is especially true in and around the positively selected mutation (*e.g.* K76T) (5).

Human Immunogenic Response to Malaria

Evolutionary Insights into Duffy Gene in Mammalian Taxa with Comparative Genetic Analysis

Evolutionary analyses of genes conserved across taxa are keys to understand the complexity of gene and genome variation. Considering malaria as a devastating infectious disease to humans, and host-parasite interaction mechanism is quite complex, detail evolutionary understanding on human gene responsible for parasite recognition and invasion of human system is the first step to understand the complexity of such interactions. The human duffy gene which is a erythrocyte chemokine receptor has been characterized in detail in this study and compared with eight other different mammalian taxa (*Pan troglodytes, Macaca mulatta, Pongo pygmaeus, Rattus norvegicus, Mus musculus, Monodelphis domestica, Bos taurus* and *Canis familiaris)*. While the genetic architechture of this gene was entirely different across all the nine taxa, a close similarity between human and chimpanzee was evident for several aspects of this gene. Comparisons on other aspects such as ratio of coding and non-coding regions, total gene length number and size of introns and difference of number of nucleotides in human and chimpanzees were also done. Phylogenetic trees were constructed on the basis of exon sizes and in total gene sizes. Most remarkably, human and chimpanzee were only 0.75% different in this gene. The results were discussed on the similarities between human and chimpanzee and gain of introns in human-chimpanzee clade with an inference on the role of evolutionary forces (mainly natural selection) in maintaining such variations across closely-related mammalian taxa. (Fig. 24).

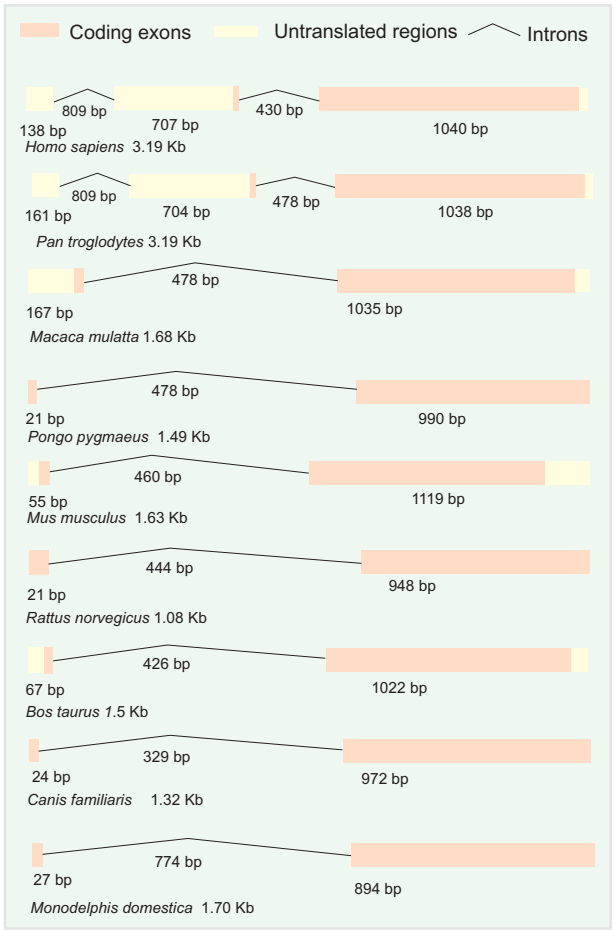

Fig. 24: Detailed characterization of the duffy gene across nine mammalian taxa

Genetic Characterization and Evolutionary Inference of TNF-α with Computational Analyses

TNF-α is an important human cytokine that imparts dualism in malaria pathogenecity. At high dosages TNF-α is believed to exhibit pathogenecity against cerebral malaria and at lower dosages TNF-α is protective against severe human malaria. In order to understand the human TNF-α gene closely and to ascertain evolutionary aspect of its dualistic nature on malaria pathogenecity, we first characterized this gene in detail in six different mammalian taxa. The avian taxa, *Gallus gallus* was included in the present study, as TNF-α is not present in birds, therefore, a tandemly placed duplicate of TNF-α (LT-α or TNF-α) was included in this study (Fig. 25). Comparative study was performed on nucleotide length variation, intron and exon size, number variation, differential compositions of coding to the non-coding bases, *etc.* to look for similarities/dissimilarities at the TNF-α gene across all seven taxa. The phylogenetic study revealed the pattern found in other genes, as human, chimpanzee and rhesus monkey were placed in a single clade and rat and mouse in another, with the *G. gallus* in a clearly separate branch. We further focused these three taxa and aligned the amino acid sequences and found less differences between human and chimpanzee but great differences in rhesus monkey from the other two

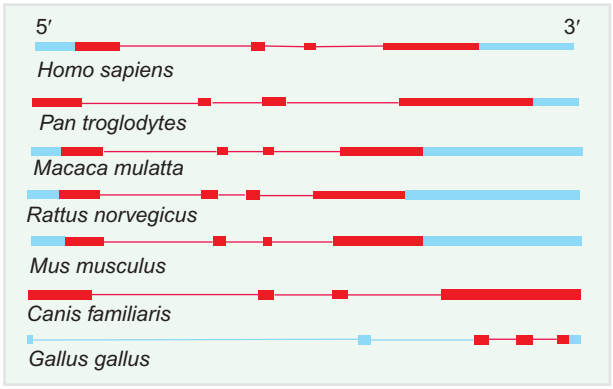

Fig. 25: Fine-scale characterization of TNF-α gene among six mammalian taxa with coding exons (red) and un-translated region (UTR) or non-coding exons (blue). For *G. gallus* information on the TNF-β has been provided. The length of non-coding exons, coding exons and introns are not in scale

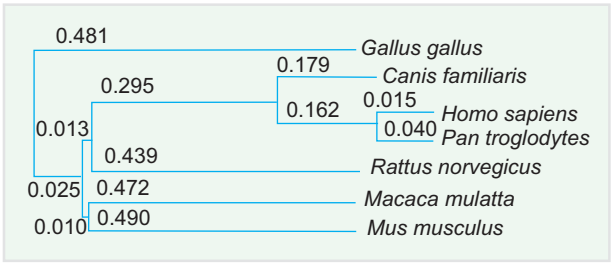

Fig. 26: Phylogenetic positions of seven different mammalian taxa at the human CD36 gene

taxa (Fig. 25). Further, comparison of coding and non-coding nucleotide length variations and coding to non-coding nucleotide ratio between TNF-α and TNF-β among these three mammalian taxa provided a first-hand indication on the role of TNF-α gene, not its duplicate TNF-β in dualistic nature of TNF-α in malaria pathogenecity.

Characterization and Comparative Analysis of Human CD36 Gene

Characterization and comparative analysis of genes of essential functions help understanding the detail composition and evolutionary pathways that the genes have passed through time. With the advent of computational biological tools and availability of whole genome sequences of different organisms, it is now possible to deeply understand the detail characteristics of such genes. We followed such approaches and characterized a human immune-system gene (CD36) responsible for malaria pathogenesis and compared with homologous genes of other six (five mammalian and one avian) taxa. We also studied distribution of CpG Islands across different types of introns (first, small and large) of the CD36 gene separately in each taxa and detected differential distribution patterns. Further, considering all the seven taxa, we constructed a phylogenetic tree on the basis of DNA sequence information of CD36 gene. Number of different copies of this gene was also determined in genome of each taxa and wide variations in copy number across taxa were detected. The detailed study on a human immune-system gene of high importance to malaria pathogenecity provides important information and paves new ways for further research on this gene (Fig. 26).

❑