Infection, Genetics and Evolution 57 (2018) 64–74

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Genetic structure of two erythrocyte binding antigens of *Plasmodium falciparum* reveals a contrasting pattern of selection

Pramita Chowdhury[a], Srikanta Sen[b], Sumana Datta Kanjilal[c], Sanghamitra Sengupta[a,*]

[a] Department of Biochemistry, University of Calcutta, 35, Ballygunge Circular Road, Kolkata 700 019, West Bengal, India
[b] Mitra Tower, Lake Town, Block-A, Kolkata 700 089, India
[c] Department of Pediatric Medicine, Institute of Post Graduate Medical Education & Research, Kolkata, West Bengal, India

ARTICLE INFO

Keywords:
Plasmodium falciparum
EBA-175
EBA-140
Glycophorin A
Erythrocyte invasion
Host-parasite interaction

ABSTRACT

Erythrocyte binding antigens 175 (EBA-175) and 140 (EBA-140) play key roles in erythrocyte invasion by binding to glycophorin A (GPA) and C (GPC) respectively in human malaria. Since antigenic variation in malaria endemic region is a major barrier to development of effective vaccine, we explore the nature and pattern of sequence diversity of these two vaccine candidates in Kolkata, India. Population genetic parameters based on parasite sequences representing region II of *Pfeba-175* and *Pfeba-140* genes were estimated using DnaSP V.5.10 and MEGA version 6.0. A novel molecular docking approach was implemented to assess the binding affinities of Kolkata *Pfeba-175* variants with GPA. *P. falciparum* Kolkata isolates experienced a recent population expansion as documented by negative Tajima's *D*, Fu & Li's statistics, unimodal mismatch distribution and star-like median-joining network for both loci. Positive selection seemed to play a major role in shaping the diversity of *Pfeba-175* ($d_N/d_S = 2.45$, and McDonald-Kreitman *P*-value = 0.04) with successive accumulation of Q584K/E, E592A and R664S deriving high frequency haplotypes designated here as F2KH3 and F2KH1. *In silico* molecular docking demonstrated that polypeptides encoded by F2KH1 and F2KH3 were capable of engaging the parasite ligand into energetically favorable interaction with GPA. Our data demonstrated emergence of *Pfeba-175* sequences harboring selectively advantageous nonsynonymous substitutions on Pf3D7 sequence background in the Kolkata parasite isolates. A contrasting pattern of Pf3D7-centric expansion of parasite sequences was noted for *Pfeba-140*. Together, this study provides a firm genetic and biological support favoring a dominant role of EBA-175 in erythrocyte invasion.

1. Introduction

Despite the declining prevalence due to improvement in the disease control and management, malaria still remains a major health burden in countries belonging to Africa and South-East Asia (Bhatia et al., 2013). Five species of parasite namely *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae*, and *Plasmodium knowlesi* are capable of infecting human, of which *P. falciparum* is the most virulent strain and associated with majority of the mortality (Snow et al., 2005). According to a recent estimate, 91 countries and territories including India are considered endemic for malaria (WHO World Malaria Report, 2016). Incidence rates of malaria in India are highly heterogeneous due to various factors such as variability in the distribution of Anopheles vector and diversity in the climatic conditions, geographical, socioeconomic and ethnic strata which jointly pose an enormous challenge to malaria control and elimination. As part of the control strategy, effective parasite surveillance is necessary not only to monitor the efficacy of malaria interventions, but to identify any emerging strains which may render a significant risk for disease resurgence. Population genetics approach can reveal important aspects of parasite population flux, migration, substructure and evolution that cannot be defined by traditional surveillance methodologies (Auburn and Barry, 2017). In the present study we derive new insights into important adaptive mechanism of *P. falciparum* through the application of an *in silico* structural biology approach in combination to genetic analysis.

Manifestation of malaria symptoms due to infection by *P. falciparum* and the burden of morbidity are associated with erythrocyte invasion by blood stage malaria parasite (merozoite). Erythrocyte invasion is an obligate part of parasite life-cycle and involves specific molecular interaction between the receptors on host's red blood cells and various ligands of the merozoites, many of which are attractive vaccine targets (Harvey et al., 2012; Bannister and Mitchell, 2003; Miller et al., 2002). Of the several merozoite proteins implicated in RBC invasion, two

* Corresponding author.
E-mail address: sanghamitrasg@yahoo.com (S. Sengupta).

https://doi.org/10.1016/j.meegid.2017.11.006
Received 24 July 2017; Received in revised form 18 October 2017; Accepted 7 November 2017
Available online 08 November 2017

classes namely Duffy binding-like (DBL) and reticulocyte binding-like protein families (RBL) are referred to as 'alternative-pathway' ligands because the proteins belong to these families have slight variation in their roles and work together with a combination of overlapping functions. This redundancy has primarily evolved to counter erythrocyte receptor polymorphisms. In addition it may help the parasite to circumvent the host's antibody responses through varied expression of these ligands which can functionally compensate each other (Duraisingh et al., 2003; Stubbs et al., 2005; Persson et al., 2008). Briefly, selection pressure and functional constraints exerted by host immune system play a crucial role in parasite survival, malaria pathogenesis vis-à-vis co-evolution of host and parasite proteins (Wright and Rayner, 2014). Copy number variants in the gene responsible for DantuNE blood group variants associated with resistance to severe malaria indicated genetic evolution of host receptor in African population (Leffler et al., 2017).

Till date four functional *P. falciparum* DBL proteins namely erythrocyte binding antigen 175 (EBA-175) (Sim et al., 1990), EBA-140 (Mayer et al., 2001; Thompson et al., 2001; Narum et al., 2002), EBA-181 (Adams et al., 2001) and EBA-1 have been described. These are secreted from micronemes after induction by cytosolic calcium level due to exposure of merozoites to low potassium ion concentrations in blood plasma and responsible for apical reorientation of merozoite (Adams et al., 2001; Mayer et al., 2001; Singh et al., 2010; Harvey et al., 2012). Antibodies against PfEBA-175, one of the highly pursued vaccine candidate, have been demonstrated to inhibit it's binding to host receptor glycophorin A (GPA) and block invasion of merozoites in vitro (Sim et al., 1990). Since, one of the obstacles in the development of effective malaria vaccine is antigenic variation of *P. falciparum* proteins due to their genetic diversity of its antigens; present study embarks on a thorough analysis of genetic architecture of two major EBA proteins namely EBA-175 and EBA-140. Both these proteins are composed of a functionally important cysteine-rich region at their N-terminal, termed as region II (RII), which is composed of two related domains, F1 and F2, responsible for receptor binding (Baum et al., 2003; Maier et al., 2009). EBA-175 engages through multiple glycans of GPA in a sialic acid dependent manner (Sim et al., 1994; Salinas et al., 2014; Pandey et al., 2002). Upon binding to GPA, RII of EBA-175 significantly influences erythrocyte membrane deformability, a process that is essential for successful parasite invasion (Sisquella et al., 2017). Determination of three dimensional structure of the RII of EBA-175 led to the understanding that F2 alone contributes to 75% of the contacts to the glycans on GPA (Tolia et al., 2005). DBL domain of EBA-140 interacts with glycophorin C (GPC) through its O-linked and N-linked sialylated glycans (Camus and Hadley, 1985; Joshua-Tor et al., 2005; Lin et al., 2012; Rydzak et al., 2012). Polymorphisms in the F1 and F2 regions of EBA-140 have been shown to alter abilities of receptor binding specificity in COS cells (Mayer et al., 2002).

Given the importance of EBA-175 and EBA-140 as parasite invasion machinery, we first, explored the level of genetic diversity of RII of *Pfeba-175* and *Pfeba-140* genes from the clinical isolates of *P. falciparum* in Kolkata, West Bengal. Our genetic analysis indicated a scenario of recent population expansion with the spread of few advantageous genetic variants of *Pfeba-175*. In contrast, a *Pfeba-140* gene pool did not show any significant departure from neutrality. Next, using a rationale-driven molecular docking approach, we evaluated the energies and various poses of interaction between EBA proteins with host glycophorins towards ascertaining if the predominant EBA-175 and EBA-140 variants in Kolkata parasite population influence the degree vis-à-vis efficacy of host-parasite interaction during erythrocyte invasion. Results from this two-pronged analysis provide substantial evidence that genetic architectures of *Pfeba-175* and *Pfeba-140* in Kolkata are shaped by different modes of natural selection.

2. Materials & methods

2.1. Study area and sample collection

P. falciparum infected malaria patients were enrolled in the year 2012–2014 from Calcutta National Medical College & Hospital, Kolkata, India following WHO guidelines (WHO guidelines, 2010). Approximately, 2 ml of blood sample was collected from each patient in EDTA vial before initiation of any medical interventions. Peripheral blood samples were confirmed *Plasmodium falciparum* infection by both Malaria Parasite Dual Antigen test and Giemsa-staining. The study was approved by the institutional ethical committees of the University of Calcutta and CNMC Hospital. The accompanying persons of all study participants gave their informed consents prior to sample collection.

2.2. DNA isolation, PCR amplification and sequencing analysis of parasite genes

Total genomic DNA was extracted from *P. falciparum* infected malaria patient's blood sample using QIAamp Blood Midi kit (Qiagen, Hilden, Germany) following manufacturer's protocol. For each isolate, F1 (*eba-175:* 825 bp; *eba-140*: 840 bp), F2 (*eba-175:* 852 bp; *eba-140*: 882 bp) and a small linker region (*eba-175*: 105 bp; *eba-140:* 72 bp) of RII segment of *eba-175* and *eba-140* genes were amplified using multiple overlapping primers (Applied Biosystems® GeneAmp® PCR System 9700). Primer designing was accomplished based on *P. falciparum* reference sequence retrieved from PlasmoDB database (http://plasmodb.org): *eba-175* (PF3D7_0731500); *eba-140* (PF3D7_1301600) (Supplementary Table S1). PCR amplification was carried out through the following temperature cycles where initial denaturation for 5 min at 94 °C, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 58–60 °C for 30–45 s; elongation at 72 °C for 1 min and a final extension at 72 °C for 5 min. Amplicons were visualized on 1% agarose gel after ethidium bromide staining and purified by QIAquick Gel Extraction kit (Qiagen, Hilden, Germany). Purified PCR products were subsequently sequenced under the cycling program of initial denaturation at 94 °C for 30 s, 25 cycles of denaturation at 94 °C for 10 s, 5 s at 50 °C and 4 min at 60 °C, followed by holding at 4 °C, in a 9700 thermal cycler. Direct sequencing of amplicons was performed using forward and reverse primers and Big Dye ver. 3.1 dye terminator chemistry. Sequencing was conducted in ABI Prism 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequence identity was validated using NCBI BLAST (http://blast.ncbi.nlm.nih.gov/). Raw sequence files were manually edited to remove ambiguous reads and signal noises and sequences were checked and assembled using BioEdit software (www.mbio.ncsu.edu/BioEdit/bioedit.html). All nucleotide and amino acid sequences were aligned using ClustalW of MEGA v 6.0 with the corresponding sequences of Pf3D7. Sequences from position PF3D7_07_v3:1,358,055 to 1,362,929 and Pf3D7_13_v3:89,319 to 93,352 of PlasmoDB database (http://plasmodb.org/plasmo/) were used to align *Pfeba-175* and *Pfeba-140* genes respectively. The sequence representing the linker region was stitched with that of F1 sequence for each sample before submission. Nucleotide sequences of F1 and F2 of *eba-175* and *eba-140* genes reported in this study were submitted to GenBank (https://www.ncbi.nlm.nih.gov/genbank/).

2.3. Population genetic analysis of parasite sequences

The sequences were aligned with the corresponding prototypic Pf3D7 reference sequence to identify sequence variants. The number of segregating sites (S), haplotypes (H), haplotype diversity (Hd), nucleotide diversity (π, θ) and the average number of pair-wise nucleotide differences, raggedness index (r), minimum number of recombination event (R_m) were estimated using the DnaSP ver. 5.10 (Rozas et al., 2003). Tajima's *D* (Tajima, 1989) statistic was used to test the status of departure from neutrality and measured by the difference between π