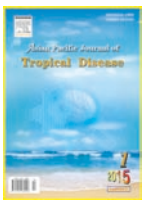
HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Disease

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



Original article doi: 10.1016/S2222-1808(15)60862-3

Characterization of sequence diversity in *Plasmodium falciparum* SERA5 from Indian isolates

Rahul C.N[1], Shiva Krishna K[1], Meera M[2], Sandhya Phadke[3], Vidya Rajesh[1*]

[1]*Department of Biological Sciences, Birla Institute of Technology and Science, Pilani, Hyderabad Campus, Andhra Pradesh, India*

[2]*Sir Ronald Ross Institute of Tropical and Communicable Diseases, Hyderabad, Andhra Pradesh, India*

[3]*Ehrlich Laboratories, Chennai, Tamil Nadu, India*

ARTICLE INFO

Article history:
Received 4 May 2015
Received in revised form 11 May 2015
Accepted 23 May 2015
Available online 1 June 2015

KeyWords:
Plasmodium falciparum
Serine repeat antigen-5
Genetic diversity
Haplotype
Antigenicity index
India

ABSTRACT

Objective: To characterize the sequence diversity of blood-stage *Plasmodium falciparum* serine repeat antigen-5 (PfSERA5) which is lacking in a malaria-endemic country like India.
Methods: In this study, parasitic DNA was obtained from field isolates collected from various geographic regions. Subsequently, PfSERA5 gene sequence was PCR amplified and DNA sequenced.
Results: We reported the existence of unique repeat polymorphisms and novel haplotypes for both the octamer repeat (OR) and serine repeat (SR) regions of the N-terminal fragment of PfSERA5 from Indian isolates. Several isolates from India were identical to low-frequency African haplotypes. Unique finding of our study was an Indian isolate showing deletion in a perfectly conserved 14 mer sequence within octamer repeat. Indian haplotypes reported in this study were found to be distributed into the three earlier classified allelic clusters of FCR3, K1 and Honduras showcasing broad diversity as compared to worldwide haplotypes.
Conclusions: This study is the first report on genetic diversity of PfSERA5 antigen from India. Further evaluation of these haplotypes by serotyping would provide useful information for investigating variant-specific immunity and aid in malaria vaccine research.

1. Introduction

High global incidence of malaria still remains a persistent problem and costly affair for mankind. Among the countries, India contributes 60% of malaria cases from South East Asia to this global incidence[1,2]. Emergence of drug resistance against frontline drugs like artemesinin has raised concern and propelled a need to develop effective vaccine for malaria[3,4]. High genetic diversity has profound implication in vaccine research and its understanding is important to guide rational vaccine design[5]. Among the known vaccine candidates, *Plasmodium falciparum* (*P. falciparum*) serine repeat antigen 5 (PfSERA5) is a potential blood stage target. This highly expressed 120 kDa antigen is processed in late blood stage into a 47 kDa N-terminal, 50 kDa central and 18 kDa C-terminal fragments. The N-terminal 47 kDa fragment of protein contributes to its vaccine potential and shows repeat polymorphism. It has also been well characterized to show immunological role representing distinct epitopes recognized by antibodies, with enough potential to inhibit parasite growth *in vitro*. Hence, PfSERA5 is currently considered for vaccine design[6-8]. India is an important study centre for complex malaria. Although, genetic diversity of the octamer repeat (OR) and serine repeat regions (SR) of PfSERA5 has been reported from nine countries recently[9,10], information from a large malaria-endemic centre like India is lacking. This genetic polymorphism report from Indian field isolates adds knowledge to the existing repertoire of antigenic variations for PfSERA5.

2. Materials and methods

Blood samples (n = 38) were collected from hospitals located in different parts of India representing different malaria-endemic settings like Tamil Nadu, Andhra Pradesh, Orissa, and Uttar Pradesh. Institutional ethical guidelines were adhered for sample collection from these regions. Malaria-positive blood samples were subjected to standard DNA extraction procedure[11].

A species-specific diagnostic PCR protocol was followed to confirm species identity[12]. In confirmed *P. falciparum* samples,

*Corresponding author: Vidya Rajesh, Ph. D, Associate Professor, Department of Biological Sciences, Birla Institute of Technology and Science, Pilani, Hyderabad Campus, Hyderabad, 500078, Andhra Pradesh, India.
Tel: 91 40 66303505
Fax: 91 40 66303998
E-mail: vidya_nrajesh@yahoo.com; vidya@hyderabad.bits-pilani.ac.in
Foundation Project: Supported by a grant from the Department of Science and Technology, New Delhi, India (SR/FT/LS-061).

amplification of Exon II (coding for P47) and Exon IV (coding for P50 and P18) of PfSERA5 was carried out by using specific primers (Table 1). The amplicons were outsourced for DNA sequencing to Life Technologies-Invitrogen, Gurgaon and Bio Serve Technologies Ltd., Hyderabad. Only chromatograms lacking overlapping peaks were included for further analysis to ensure single infection. The sequences were submitted to Genbank under the accession numbers KC236368-KC236390 and KJ397705-KJ397719.

Sequence analysis was done by multiple sequence alignment through Clustal W programme available in BioEdit 7.0.5.3[13]. Phylogenetic tree was generated by neighbour joining method with MEGA5 5.2 software[14]. Statistical analysis of basic population genetic parameters like haplotype diversity (HD)-measure of uniqueness of a particular haplotype in a given population, and nucleotide diversity (π)-average number of pairwise differences per site, were also computed for the sequence data obtained with DnaSP 5[15,16]. Antigenicity plots for distinct haplotypes were generated using online webtool Antigenicity Plot Index[17].

3. Results

3.1. OR region polymorphism from Indian isolates

Analysis of this region in Indian isolates showed repetitions, truncations and variations. Almost all Indian isolates analysed showed the standard number of OR *i.e.* 7 (6 Type I + 1 Type II). However, three isolates showed exceptions with repeat variation in Type Ie and Type-II. They were HF37 (5 repeats of Type Ie), CE145 (1 Type Ie + 3 repeats of Type II) and HF169 (5 repeats of Type Ie + 1 Type II) (Figure 1A). In addition, samples HF105 and CE145 showed a specific glutamine-arginine (G→R) change within type Ia TGESQT (G/R)N (Figure 1A). Truncated OR occurred between 1b and 1c in (23/38) Indian isolates. Out of these, truncated repeat 4/4 -TVGDQAGN, occurred in (21/23) Indian isolates. While two Indian isolates HF170 and HF174 showed 3/5-TVGGQAGN truncated repeat. Truncated repeat 3/5-TVGSPQGS was exceptional to Indian isolate CE145 in the present study (Figure 1A). Phylogenetic

Table 1

Details of primers used for amplification of PfSERA5 gene.

Primer name	Primer sequence	Region targeted	Amplification conditions
Forward: CNRSERA5PF1A	5'gcgtcgacatgaagtcatatatttccttg3'	PfSERA5-Exon II	3 min initial denaturation at 94 °C followed by 36 cycles of 1 min denaturation at 94 °C, 1.3 min annealing at 58.8 °C, and 2 min extension at 72 °C and final 10 min extension at 72 °C
Reverse: CNRSERA5PF2	5'gcgacgcgttactaaaagagcacattgaaa3'		
Forward: CNRSERA5PF3	5'gcggtcgactgctttcaatgtgctcttttag3'	PfSERA5-Exon IV	3 min initial denaturation at 94 °C followed by 36 cycles of 1 min denaturation at 94 °C, 1.3 min annealing at 59 °C, and 2 min extension at 72 °C and final 10 min extension at 72 °C
Reverse: CNRSERA5PF4	5'gcgacgcgttggagagttatgccctattt3'		

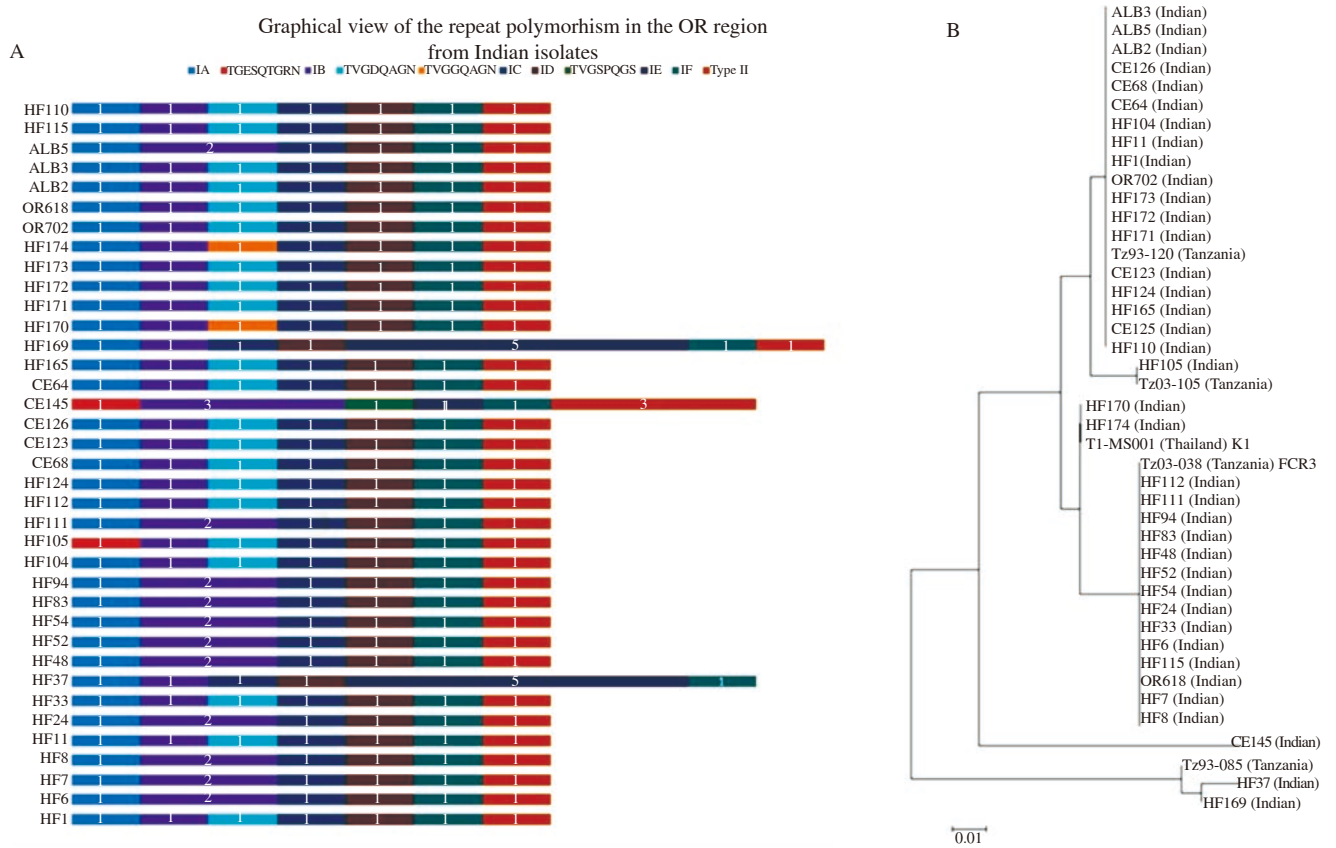


Figure 1. Characterization of sequence diversity in the OR region of P47.

A: Graphical representation of polymorphism in OR region of N-terminal P47 domain of PfSERA5 gene from Indian isolates; Repeat types: 1A-TGESQTGN variant-TGESQTRN, 1B-TGGGQAGN, Truncated repeat 4/4 -TVGDQAGN, 3/5-TVGGQAGN, 1C-TVGDQAGS, 1D-TGGSPQGS, Truncated 3/5-TVGSPQGS, 1E-TGASPQGS, 1F-TGASQPGS, Type II-SEPSNPVS; B: Phylogenetic analysis of OR region of N-terminal P47 domain of Indian isolates among worldwide isolates.

analysis showed that 36/38 Indian isolates conferred to 5 reported haplotypes while two samples HF37 and CE145 were novel OR haplotypes reported in this study[9] (Figure 1B, Table 2). Among the remainning 36 Indian isolates, 18 showed similarity to haplotype Tz93-120 which is the most prevalent OR haplotype reported worldwide. Fourteen Indian isolates showed similarity to reported haplotype Tz03-38 (FCR3), a prevalent haplotype predominant in Tanzania and Ghana. Two Indian isolates, HF170 and HF174 were similar to Thailand haplotype Ti-MS001 (K1) while isolates HF105 and HF169 were similar to low-frequency haplotypes restricted to African continent Tz03-105 and Tz93-085, respectively.

3.2. SR region polymorphism in Indian isolates

The sequence alignment of SR region of the P47 N-region showing amino acid region from 196-274 amino acids is represented in Figure 2A. The samples were compared with 3D7 standard. Among the Indian isolates, 21 were of K1 type, 14 were of FCR-3 type and 3 were recombinant Honduras type. In our field isolates, Poly-Ser repeat number was in the range of 23-49. Three Indian isolates, HF104, HF33, and HF124 showed minor-frequency single nucleotide polymorphism at amino acid position 217 (S-Serine to N-Aspargine) within the Poly Ser repeat. Phylogenetic analysis of SR haplotypes and their distribution into K1, FCR3 and Honduras is shown in Figure 2b and Table 3, respectively. Analysis of Indian isolates for the central protease domain (P50) and C-terminal region (P18) of PfSERA5 indicated absolute genetic conservation. This signified a vital role of these domains in the parasite life cycle.

3.3. Statistical and computational analysis of polymorphism data

Antigenicity plots of unique Indian OR variants showed differences in pattern of antigenicity in comparison to standard sample isolates

Table 2

Haplotype distribution of the OR region of P47 domain of Indian isolates.

S. No	Indian isolate	N	Worldwide OR haplotype similarity	Prevalence reported
1.	HF124, HF165, CE125, HF110, HF172, OR702, HF1, HF11, HF104, CE64, CE68, CE126, ALB2, ALB3, ALB5, HF171, HF173, CE123	18	Tz93-120	Worldwide
2.	OR618, HF112, HF111, HF94, HF83, HF54, HF52, HF48, HF33, HF24, HF8, HF7, HF6, HF115	14	Tz03-038 (FCR3)	Tanzania, Ghana
3.	HF170, HF174	2	Ti-MS001 (K1)	Thailand
4.	HF105	1	Tz03-105 shows singleton mutation (G→R) in Type 1a OR	Low frequency in Tanzania and Ghana
6.	CE145	1	Novel G→R in Type 1a OR, repeat variation in Type 1e & Type II	India
7.	HF37	1	Novel repeat variation in Type 1e and deletion of Type II	India
8.	HF169	1	Tz93-085 repeat variation in Type 1e	Low frequency in Tanzania

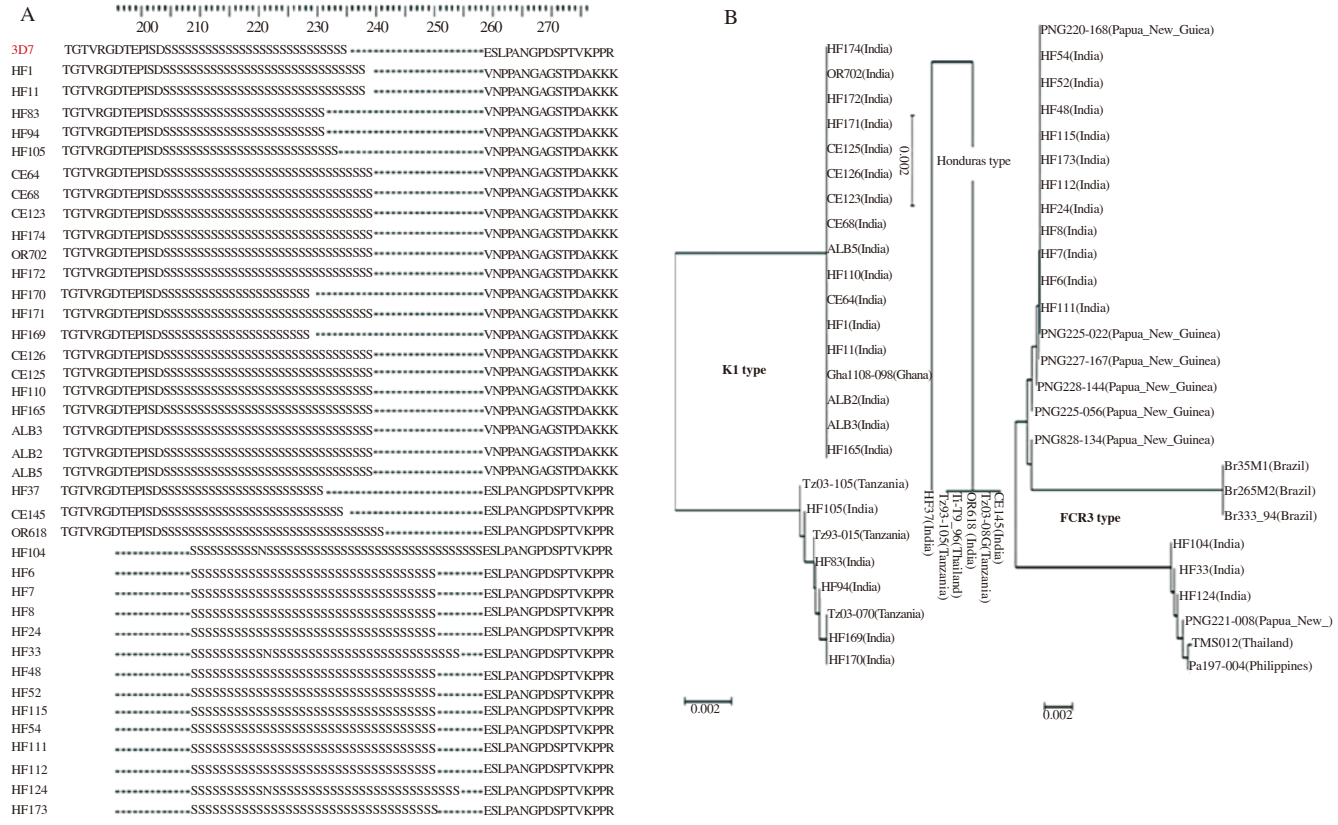


Figure 2. Characterization of sequence diversity of SR region of P47 domain.

A: Multiple sequence alignment of N-terminal P47 SR region of PfSERA5 gene from Indian isolates using standard reference sequence 3D7 strain (highlighted in red box); B: Phylogenetic analysis of SR region of P47 domain of PfSERA5 of Indian isolates.

Table 3
Haplotype distribution of SR region of P47 domain of PfSERA5 in Indian isolates.

S. No	Indian isolate	N	Allelic type	Serine repeat number	Worldwide SR haplotype similarity	Prevalence reported
1.	ALB2, ALB3, ALB5, HF110, OR702, HF174, HF172, HF171, HF165, CE125, HF1, HF11, CE64, CE68, CE126, CE123	16	K1	31	Gha1108-098	Ghana
2.	HF105	1	K1	25	Tz03-105	Tanzania, Ghana
3.	HF169, HF170	2	K1	21	Tz03-070	Worldwide
4.	HF83, HF94	2	K1	23	Tz93-015	Tanzania, Ghana
6.	CE145	1	Honduras	27	Tz03-086	Tanzania (single isolate)
7.	HF37	1	Honduras	23	Novel	India
8.	OR618	1	Honduras	33	Ti-T9_96	Thailand (single isolate)
9.	HF104	1	FCR3	49 & S→N at 217th aa	Novel	India
10.	HF33, HF124	2	FCR3	42 & S→N at 217th aa	Novel	India
11.	HF48, HF112, HF6, HF7, HF8, HF24, HF173, HF115, HF111, HF54, HF52	11	FCR3	42	Novel	India

(Figure 3). HD and nucleotide diversity (π) were obtained for both OR and SR regions for sequence results represented in the present dataset. Results indicated higher HD and value for SR region in comparison to the OR region (Table 4).

Table 4
Statistical analysis of genetic polymorphism in PfSERA5-P47 repeat region.

PfSERA5-P47 repeat region	HD	Nucleotide diversity (π)
OR	0.697 + 0.050	0.021 04
SR	0.773 + 0.058	0.332 99

4. Discussion

The present study deals with the polymorphism of nucleotide sequences of PfSERA5 from Indian field isolates. Sequence results obtained for OR region were represented based on the earlier description of repeat types which consist of two repeat types - Type I (subdivided into six subgroups Ia-Ie), and Type II[9]. Three Indian isolates HF37, CE145, HF169 which showed variations in repeat types exclusively for Type 1e and Type II were earlier reported to be restricted to African isolates alone[9]. HF 37 is the only isolate that shows deletion of Type II repeat which disrupts the 14 mer stretch QPGSSEPSNPVSSG (amino acid 59-72), conserved till date in all worldwide isolates and believed to be recognized by MAb43E5[9]. This is the first report of such a deletion in the critical sequence of OR region. Antigenicity plot results in our study differ from an earlier report where the impact of OR sequence variation on immunogenicity was suggested to be insignificant[18]. Significance of these observations on immunogenicity needs to be further evaluated with studies on human sera antibodies.

SR region is a dimorphic locus and is structured into Poly-Ser repeats sandwiched between an upstream 13 amino acid (TGTVRGDTEPISD) and downstream 17 amino acid stretch which shows clustered mutations. Based on these variations, the worldwide haplotypes have been grouped into three allelic clusters FCR3, K1 and Honduras (recombinant type)[10]. Most of the K1 haplotypes reported in the present dataset are similar to low-frequency African haplotypes. The 14 FCR3 type isolates found in India distribute into

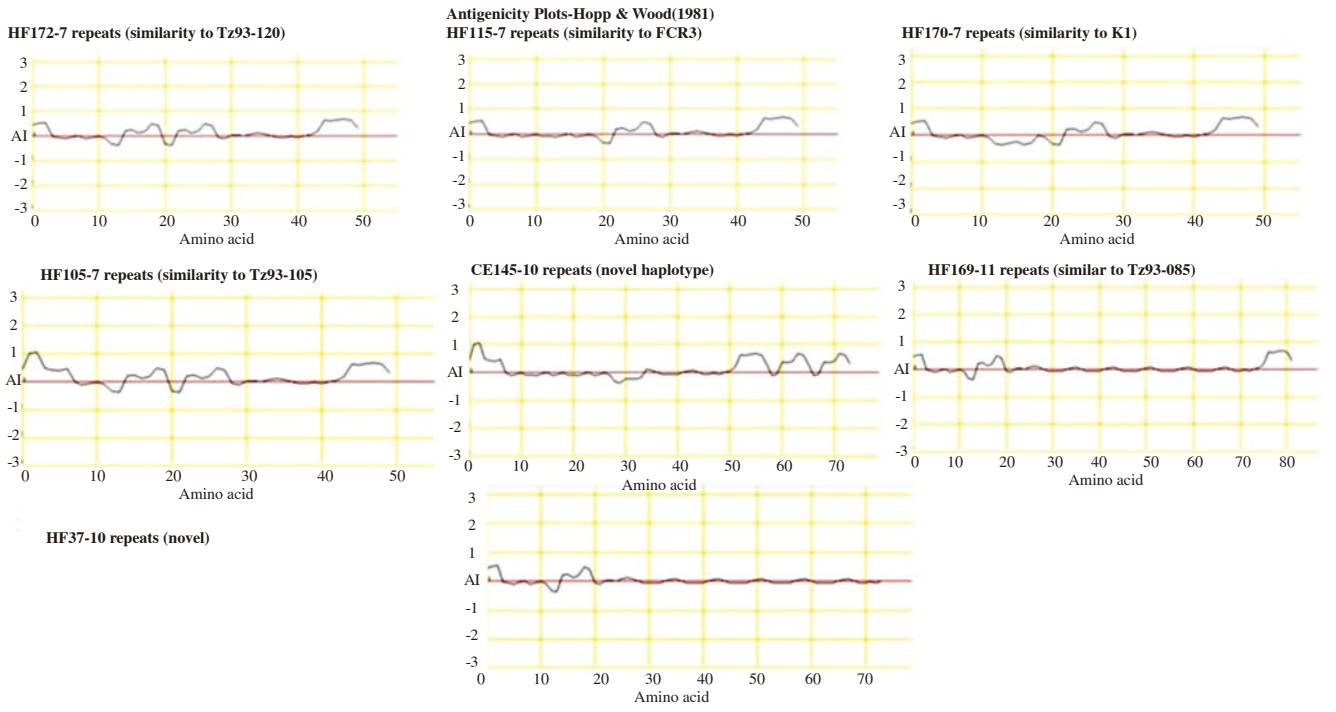


Figure 3. Antigenicity plots for standard (FCR3 & K1) isolates and OR haplotypes from India, displaying the variation of the antigenic index as function of amino acid position.
AI: Antigenic index (woed label of vertical ordinate).

3 novel SR haplotypes unique to these samples. Three Indian isolates CE145, HF37, and OR618 clusters with the Honduras type isolates Tz03-086 (Tanzania), Ti-T9_96 (Thailand) and Tz93-105 (Tanzania) are shown in Figure 2B. The HF37 is a novel Honduras haplotype unreported earlier.

Statistical results of polymorphism data computed for SR and OR regions indicate that genetic diversity in the SR region is higher in comparison to the OR region in Indian field isolates. The overall genetic diversity based on the values for both SR and OR obtained for present dataset indicates maintenance of medium-range genetic diversity in the N-terminal region of PfSERA5 from Indian isolates. This is evident when compared with earlier reports from other countries[9].

Our paper, for the first time, characterizes the sequence diversity in PfSERA5 gene from India. The results add to the antigenic repertoire of PfSERA5 worldwide isolates within PfSERA5 N-terminal region. Several novel OR (2) and SR haplotypes (4) are identified in this study. Predominance of African haplotypes in Indian field isolates suggests probable gene flow from Africa. Further evaluation of these novel mutations by antibody profiling would aid to establish variant-specific immunity.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgments

We would like to thank the University Grant Commission, New Delhi, India for funding the research scholar. We thank Dr. Nalini Rastogi and Dr. Aditya K. Pandey for their help in collection of samples from Allahabad and Orissa respectively. We acknowledge the sequencing support provided by Bio Serve Technologies Ltd., Hyderabad and Life Technologies- Invitrogen, Gurgaon. This work was supported by a grant from the Department of Science and Technology, New Delhi, India (SR/FT/LS-061).

References

[1] World Health Organization. World malaria report 2012. Geneva: World Health Organization; 2012. [Online] Available from: http://www.who.int/malaria/publications/world_malaria_report_2012/wmr2012_full_report.pdf [Accessed on 20th April, 2015].

[2] Singh V, Mishra N, Awasthi G, Dash AP, Das A. Why is it important to study malaria epidemiology in India? *Trends Parasitol* 2009; **25**(10): 452-7.

[3] Phyo AP, Nkhoma S , Stepniewska K, Ashley EA, Nair S, McGready R, et al. Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study. *Lancet* 2012; **379**(9830): 1960-6.

[4] Miotto O, Almagro-Garcia J, Manske M, Macinnis B, Campino S, Rockett KA, et al. Multiple populations of artemisinin-resistant *Plasmodium falciparum* in Cambodia. *Nat Genet* 2013; **45**(6): 648-55.

[5] Takala SL, Plowe CV. Genetic diversity and malaria vaccine design, testing and efficacy: preventing and overcoming 'vaccine resistant malaria'. *Parasite Immunol* 2009; **31**(9): 560-73.

[6] Aoki S, Li J, Itagaki S, Okech BA, Egwang TG, Matsuoka H, et al. Serine repeat antigen (SERA5) is predominantly expressed among the SERA multigene family of *Plasmodium falciparum*, and the acquired antibody titers correlate with serum inhibition of the parasite growth. *J Biol Chem* 2002; **277**(49): 47533-40.

[7] Fox BA, Horii T, Bzik DJ. *Plasmodium falciparum*: fine-mapping of an epitope of the serine repeat antigen that is a target of parasite-inhibitory antibodies. *Exp Parasitol* 2002; **101**(1): 69-72.

[8] Yagi M, Bang G, Tougan T, Palacpac NM, Arisue N, Aoshi T, et al. Protective epitopes of the *Plasmodium falciparum* SERA5 malaria vaccine reside in intrinsically unstructured N-terminal repetitive sequences. *PloS One* 2014; **9**(6): e98460.

[9] Tanabe K, Arisue N, Palacpac NM, Yagi M, Tougan T, Honma H, et al. Geographic differentiation of polymorphism in the *Plasmodium falciparum* malaria vaccine candidate gene SERA5. *Vaccine* 2012; **30**(9): 1583-93.

[10] Safitri I, Jalloh A, Tantular IS, Pusarawati S, Win TT, Liu Q, et al. Sequence diversity in the amino-terminal region of the malaria-vaccine candidate serine repeat antigen in natural *Plasmodium falciparum* populations. *Parasitol Int* 2003; **52**(2): 117-31.

[11] Sambrook J, Russell DW. *Molecular cloning: a laboratory manual*. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001, p. 2100.

[12] Das A, Holloway B, Collins WE, Shama VP, Ghosh SK, Sinha S, et al. Species-specific 18S rRNA gene amplification for the detection of *P. falciparum* and *P. vivax* malaria parasites. *Mol Cell Probes* 1995; **9**(3): 161-5.

[13] Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 1999; **41**: 95-8.

[14] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011; **28**(10): 2731-9.

[15] Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 2009; **25**(11): 1451-2.

[16] Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 1989; **123**(3): 585-95.

[17] Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci USA* 1981; **78**(6): 3824-8.

[18] Pratt-Riccio LR, Sallenave-Sales S, de Oliveira-Ferreira J, da Silva BT, Guimarães ML, Santos F, et al. Evaluation of the genetic polymorphism of *Plasmodium falciparum* P126 protein (SERA or SERP) and its influence on naturally acquired specific antibody responses in malaria-infected individuals living in the Brazilian Amazon. *Malar J* 2008; **7**: 144.