Journal of Infection and Public Health 12 (2019) 576–584

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: http://www.elsevier.com/locate/jiph

Genetic diversity and allelic variation in MSP3α gene of paired clinical *Plasmodium vivax* isolates from Delhi, India

Asha Kaul[a], Prerna Bali[a], Shadab Anwar[b], Ajay K. Sharma[c], Birendra K. Gupta[c], Om P. Singh[a], Tridibes Adak[a], Mohammad Sohail[a,c,*]

[a] Division of Molecular Parasitology, National Institute of Malaria Research, Sector-8, Dwarka, New-Delhi, India
[b] School of Life Sciences, Jawaharlal Nehru University, New Delhi, India
[c] Research Laboratory, University Department of Zoology, Vinoba Bhave University, Hazaribag, Jharkhand, India

ARTICLE INFO

Article history:
Received 26 February 2018
Received in revised form 22 October 2018
Accepted 31 January 2019

Keywords:
Plasmodium vivax
MSP-3α
Polymorphism
Malaria
Sequence diversity
Restriction pattern
Paired isolates

ABSTRACT

Background: Plasmodium vivax malaria accounts for 80% of the malaria cases in Delhi, India. The gene merozoite surface protein 3 alpha (MSP3α) is highly polymorphic and has been used as marker in many *P. vivax* population studies.
Methods: MSP3α has been used to assess the genetic diversity of *P. vivax* samples from Delhi (India) having more than one malaria episode (s) i.e. clinically identified relapse cases using PCR-RFLP and sequencing.
Results: Three major genotypes 2.0 kb (A), 1.4 kb (B) and 1.2 kb (C) were amplified from 72 isolates with frequencies of 72.2%, 19.44% and 9.72% respectively. One sample out of 72 showed mixed infection having both A and B type genotypes. 82.05% patients showed same genotype while only 17.94% patients showed different genotypes after subsequent malaria episodes. 18 different genotypes with Alu I and 35 with Hha I were identified among 72 samples analyzed by restriction fragment length polymorphism (RFLP). 18 Pvmsp3α nucleotide sequences were analyzed and it did not reveal any distinct intragenic differences within sequences of the same type, however, allelic diversity among the three types ($\Pi = 0.029703$) was observed. Phylogenetic analysis showed allelic family types A, B and C were not clustered but distributed in different branches. The results indicate that the *P. vivax* parasite population is highly diverse in Delhi, India. A large number of amino acid substitutions were found at the locus of the isolates when compared with the Belem Strain ($\Pi = 0.030528$). The substantial sequence diversity is largely restricted to certain domains of encoded protein. Analysis of synonymous and nonsynonymous substitutions suggested that different selection forces were operating on different regions of the protein molecule.
Conclusion: We propose that genotyping of the *Pv*MSP-3α gene as one of the molecular tools for differentiating relapse from new infection in epidemiological settings. The analyses of sequence polymorphism in PvMSP-3α gene enable it as potential candidate for inclusion in a *P.vivax* vaccine research.

Introduction

The global burden of malaria due to *Plasmodium vivax* is about 70–80 million cases annually. Malaria in India contributes around 2–2.5 million cases every year. *P.vivax* is the most widely distributed plasmodium species and is very prevalent in Delhi, India. It causes about 60–65% of all malaria infections in India [1–3]. The prevalence of *P. vivax* population may increase again due to its ability to evade the host immune system which in turn would lead to the emergence of a highly diverse parasite population. Mostly genetic studies have been done on *Plasmodium falciparum* using polymorphic markers MSP1, MSP2 and GLURP [4,5]. The focus has now been shifted to *P. vivax*, with a number of genes encoding MSPs, CSP, AMA I and GAM I being identified [6,7]. Merozoites are one of the prime vaccine targets, and antibodies against merozoite surface molecules have shown to block parasite invasion in erythrocytes in vitro [8]. *Plasmodium* merozoite surface proteins are targets of naturally acquired and vaccine-induced immunity against malaria and PvMSP3α a member of MSP3 gene family has been recently validated as a genetic marker to study polymorphism from diverse geographic regions [9–13]. MSP3-α is a single copy gene that is highly polymorphic used for *P. vivax* genotyping, [9,13]. It encodes

* Corresponding author at: Research Laboratory, University Department of Zoology, Vinoba Bhave University, Hazaribag, Jharkhand, India.
E-mail addresses: soh.khan@hotmail.com, soh.khan@vbu.ac.in (M. Sohail).

https://doi.org/10.1016/j.jiph.2019.01.068

a merozoite surface protein with an alanine-rich central domain predicted to form a coiled-coil tertiary structure [14].

In contrast to *P.falciparum*, the management and control of *P.vivax* is perplexed and complicated by its ability to cause relapse infections, a phenomenon that has intrigued parasitologists for more than century but the reason and mechanisms of which remain enigmatic. Relapse is the result of the activation of quiescent liver-stage developmental forms, known as "hypnozoites" that remain dormant within hepatocytes for varying intervals before spontaneously dividing and developing into schizonts and subsequently releasing invasive merozoites into the bloodstream to infect red blood cell. Relapse is an important aspect of *P.vivax* life cycle bearing upon chemotherapy and its assessment. To date, the underlying mechanism of influencing relapses and the patterns of relapses in vivax infection are remain unclear. However, limited experimental and clinical studies have proposed that vivax malaria infections acquired in different geographic and climatic regions have demonstrated striking differences in their pattern of relapse. In general, strains from subtropical or tropical zones are associated with early primary infections followed by frequent relapses in short intervals, whereas strains from temperate zones are associated with primary infections that tend to delayed, with fewer relapses [15–17]. Mixed relapse patterns are observed in some area [1,18]. In addition, some investigators have reported that a relapse pattern are strain specific, reproducible in different hosts and appears to be largely independent of infected host's immune response [19]. Thus, it is suggested that the timing of relapses is strain-dependent and determined by the genetic makeup of the individual sporozoites, leading to the concept of tachysporozoites (strains with short latent periods) and bradysporozoites (strains with long latent periods) [16,20]. However, relapse biology remains poorly understood at the molecular level with contrast findings on genetic relatedness among paired primary infection and relapse isolates. Earlier molecular studies indicate that parasites associated with primary infection and relapses are not genetically different [21,22] whereas recent investigation by Imwong et al. [23] and Chen et al. [24] have demonstrated that parasite isolates associated with relapse of infection rarely have the same genotype as the parasites that caused the primary infection. The objectives of the present study were to distinguish between new infection or relapse (isolates with different genotypes) and treatment failure (isolates with same genotypes) using PCR-RFLP. The protocol was tested with blood samples from patients with multiple malaria episodes. In this study, further sequencing of number of isolates was done to determine the genetic diversity of reemerging *P. vivax* in the same patient.

Materials and methods

Patients and clinical data

A total of 72 isolates from 33 patients (28 patients (paired) suffered from two, 4 patients (triplicate) suffered from three and 1 patient (quadruplet) suffered from four malarial episodes) with an evidence of vivax infection on microscopic examination of outpatient of either sex (Table 1) who attended local Malaria Clinic at 2-Nanak Enclave, National Institute of Malaria Research, New-Delhi, India were included in this study. Stringent diagnostic criteria were used to diagnose vivax infection with our trained technical staff. Data collected included basic demographic information, pattern and onset of infection and use of medications. Clinical records were used to verify patients' data and the study protocol was carried out in accordance with the National Institute of Malaria Research, Delhi, human ethical guideline as reflected in the guidelines of the medical ethics committee, Ministry of Health, India and informed consent was obtained from each patient. Blood specimens were collected from all age groups during different transmission periods of the year from positive cases of *P.vivax* malaria, who had undergone clinical investigation and confirmed on the basis of clinical symptoms and a parasite blood film was checked after staining with JSB stain [25]. All patients were found to be only infected with *P.vivax*.

DNA isolation

Blood samples were collected from patients with vivax infection as dried blood spots on filter papers (Whatmann 3). Genomic DNA was extracted using DNA isolation kit (QIAmp DNA mini Kit: Qiagen, Krefeld, Germany), following the manufacturer's instructions.

Genotyping of PvMSP-3alpha

The genetic polymorphism of PvMSP-3α gene was carried out as described by Bruce et al. [9]. The MSP-3α gene was amplified by a nested PCR. In a 20 μl PCR reaction volume, 2 μl of template DNA was used in primary reaction and 1 μl of the primary product was used in nested PCR. The PCR parameters were as follows: an initial denaturation step at 95 °C for 3 min preceded the cycles of denaturation at 94 °C for 30 s, annealing at 56 °C for primary and 57 °C for nested reactions for 30 s and an extension at 68 °C for 2.5 min for 30 cycles in an automated thermocycler (Gene Amp 9700, Applied Biosystem, Foster, LA, USA) The PCR product was then subjected to electrophoresis on a 0.8% agarose gel, stained with ethidium bromide and observed under UV light.

Restriction fragment length analysis of PvMsp-3alpha

For restriction digestion analysis, 4 μl of each PCR product was digested individually with 5 units of Alu I or Hha I restriction enzymes (NEB) in a 20 μl reaction volume at 37 °C for 4–5 h. The digested products were visualized under UV illumination after electrophoresis on 1.8% agarose gels.

Sequencing of the Pv MSP-3alpha

18 PCR products; 11 out of 52 Type A, 3 out of 14 Type B and 4 out of 7 Type C representing 2.0 kb, 1.4 kb and 1.2 kb respectively were sequenced using Big Dye terminator kit.v3.1 and the DNA analyzer 3720XL DNA analyzer (Applied Biosystems, Foster, USA). The sequences were then analyzed using Mega 4.0 software package [26].

Analysis of DNA sequencing

Nucleotide sequence data of Type A, Type B and Type C alleles obtained in this study were deposited with GenBank accession numbers HQ328853 to HQ328855 and MK072739 to MK072747 respectively. The allele sequences were aligned and compared with 14 additional previously published PvMSP-3α sequences from various geographic localities with Gen Bank accession no. AF 4991946, 4991951, 4991955, 4991957 through 4991962, AF 093584, AY 833013, AY 266088, AY 266091 and AJ 864941. All aligned nucleotide sequences were analyzed using Clustal W. An analysis of the prevalence of synonymous (dS) and non-synonymous substitutions (dN) in PvMSP-3α was undertaken in order to infer the type of selective pressure, if any, operating on protein. Z tests were performed using Mega 4.0 over the complete sequence alignment [27].

Table 1
Clinical data of *P.vivax* positive patients and their recurrent samples showing different genotypes and RFLP pattern.

Patient code	Date	Age/ sex	Wt. (Kg.)	Temp (°C)	Allele size (Kb)	AluI pattern (size in bp)	Geno type Alu I	HhaI pattern (size in bp)	Geno type Hha I
PV-1	11-03-05	42/ M	55	99	1.4	550,300,230,200,150	1	1.1,400,200	1
	08-06-05				1.4	550,300,230,200,150	1	1.1,400,200	1
	25-04-06				1.4	550,300,230,200,150	1	1.1,400,200	1
PV-2	11-04-05	13 /M	38		2	550,500,430,230,200,150	2	1.1, 400,250, 200,150	2
	18-05-05			100	2	550,500,430,230,200,150	2	1.1, 400,250, 200,150	2
PV-3	21-03-05	32/ M	48		2	550,500,430,230,200,150	2	1.1,500,450,400,250,200	3
	06-06-05			104	2	550,500,430,230,200,150	2	1.1,500,450,400,250,200	3
PV-4	22-03-05	42 /M	54		2	550,500,430,230,200,150	2	1.1,350,250,220,150	4
	08-06-05			102	1.4	550,300,250,200,150	3	1.1,450	5
	30-09-05			102	2	550,500,430,230,200,150	2	1.1,500,250,150	6
PV-5	05-05-05	24/ M	56		2	550,500,430,230,200,150	2	1.1,450,250,220	6
	10-06-05				1.2	550,230,200,150	4	1.1,200	7
PV-6	23-06-05	23 /M	59		2	550,500,430,230,200,150	2	1.1,350,250,220,150	4
	01-08-05			104	2	550,500,430,230,200,150	2	1.1,350,250,220,150	4
PV-7	20-06-05	17 /M	45	99	1.4	550,300,250,200,150	3	1.1,350,150	8
	09-08-05			98.5	1.4	550,300,250,200,150	3	1.1,350,150	8
PV-8	25-10-05	21/ M	48	98.4	2	550,500,430,250,200,150	5	1.1,450,350,150	9
	20-12-05				2	550,500,430,250,200,150	5	1.1,450,350,150	9
PV-9	25-05-05	32/ M	66	98.4	2	550,500,430,250,200,150	5	1.1,300,250,200,150	10
	05-07-05			98.4	2	550,500,430,250,200,150	5	1.1,300,250,200,150	10
	19-08-05			99.1	2	550,500,430,250,200,150	5	1.1,300,250,200,150	10
PV-10	04-07-05	19/ M	50		2	550,500,350,250,200,150	6	1.1,450,350,150	9
	22-08-05			99	1.4	550,300,250,200,150	3	1.1,200,150	11
PV-11	15-07-05	36/ M	50		1.2	550,300,250,150	7	1.1,150	12
	29-08-05			99	1.2	550,300,250,150	7	1.1,150	12
PV-12	11-07-05	38/ M	65	105	2	550,500,430,300,250,200,150	8	1.1,450,250,150	13
	30-08-05			98.4	2	550,500,430,300,250,200,150	8	1.1,450,250,150	13
PV-13	12-07-05	18/ F	30	98.4	2	550,500,430,250,200,150	5	1.1,500,250,150	6
	06-09-05			99	2	550,500,430,250,200,150	5	1.1,500,250,150	6
PV-14	12-07-05	34/ F	60	98.6	2	550,500,430,250,200,150	5	1.1,500,250,150	6
	06-09-05			98.4	2	550,500,430,250,200,150	5	1.1,500,250,150	6
PV-15	01-09-05	45 /F	53	99	2	550,500,430,250,200,150	5	1.1,350,220,150	14
	20-10-05			98.4	2	550,500,430,250,200,150	5	1.1,350,220,150	14
PV-16	05-09-05	35/ M	57	99.4	2	550,500,430,250,200,150	5	1.1,300,250,200,150	10
	14-11-05			100.2	2	550,500,430,250,200,150	5	1.1,300,250,200,150	10
PV-17	01-09-05	23/ M	50	98.4	2	550,500,430,250,200,150	5	1.1,450,400,150	15
	07-04-06			102	2	550,500,430,250,200,150	5	1.1,450,400,150	15
PV-18	30-01-06	26/ M	54	98.2	2	550,500,430,250,200,150	5	1.1,450,250,220	6
	03-04-06			99.5	2	550,500,430,250,200,150	5	1.1,450,250,220	6
	08-05-06			98.9	2	550,500,430,250,200,150	5	1.1,450,250,220	6
PV-19	01-09-05	20/ F	41	99	1.2	550,300,250,150	7	1.1,200	7
	09-05-06			98.9	1.2	550,300,250,150	7	1.1,200	7
PV-20	01-09-05	50 /F	50		2	550,500,430,250,200,150	5	1.1,550,250,220	16
	24-05-06			101.7	2	550,500,430,250,200,150	5	1.1,550,250,220	16
PV-21	18-08-05	42 /M	70	103	1.4	550,430,200,150	9	1.1, 400	17
	30-05-06			98.9	1.4	550,430,200,150	9	1.1, 400	17
PV-22	12-09-05	22/M	55	99.1	2	550,500,430,250,200,150	5	1.1, 400,300,220,150	18
	22-06-06			99.1	2	550,500,430,250,200,150	5	1.1, 400,300,220,150	18
PV-23	24-08-05	45/ F	55	99	2	550,500,350,250,200,150	6	1.1,550.250,220	19
	23-06-06			98.2	2	550,500,350,250,200,150	6	1.1,550.250,220	19
PV-24	26-07-05	18 /M	60		1.2	550,300,250,150	7	1.1,200	7
	23-06-06			98.6	1.2	550,300,250,150	7	1.1,200	7
PV-25	08-05-06	30 /M	61	97.7	2	550,500,430,250,200,150	5	1.1,500,350,150	20
	11-07-06			98.4	2	550,500,430,250,200,150	5	1.1,500,350,150	20
PV-26	03-07-06	17 /F	43	98.4	1.4	550,300,250,200,150	3	1.1,220,150	21
	07-08-06			99.5	1.4	550,300,250,200,150	3	1.1,220,150	21
PV-27	03-07-06	35 /M	49	103	2	550,500,430,250,200,150	5	1.1,500,400	22
	21-08-06			104.1	2	550,500,430,250,200,150	5	1.1,500,400	22
PV-28	31-08-05	42 /M	70		2	550,500,430,230,200,150	2	1.1,400,250,220	23
	21-08-06			98.6	2	550,500,430,230,200,150	2	1.1,400,250,220	23
PV-29	04-07-06	22 /M	56	98.5	2	550,500,350,250,200,150	6	1.1,450,250,150	13
	04-12-06				1.4	550,300,250,200,150	3	1.1,200,150	11
PV-30	27-06-07	59 /M	70	99.2	2	550,500,300,250,200,150	10	1.1,300,250,220,200,150	24
	09-10-07			101	2	550,500,300,250,200,150	10	1.1,300,250,220,200,150	24
PV-31	22-05-08	31/M	56	98.6	2	550,500,350,300,200,150	11	1.1,350,250,220,150	4
	18-07-08			100.2	2	550,500,350,300,200,150	11	1.1,350,250,220,150	4
	05-09-08			98.6	2, 1.4	550,250,200,150	12	1.1,500,450,250,220	25
	20-10-08			98.4	1.4	550,300,250,200,150	3	1.1,220,150	21
PV-32	09-06-08	19/ F	40	98.4	2	550,500,430,230,200,150	2	1.1,450,300,220	26
	29-07-08			99	2	550,500,430,230,200,150	2	1.1,450,300,220	26
PV-33	25-08-08	35 /M	57	104.1	2	550,500,430,250,200,150	5	1.1,400,250,220,150	27
	15-10-08			99.7	2	550,500,430,250,200,150	5	1.1,400,250,220,150	27

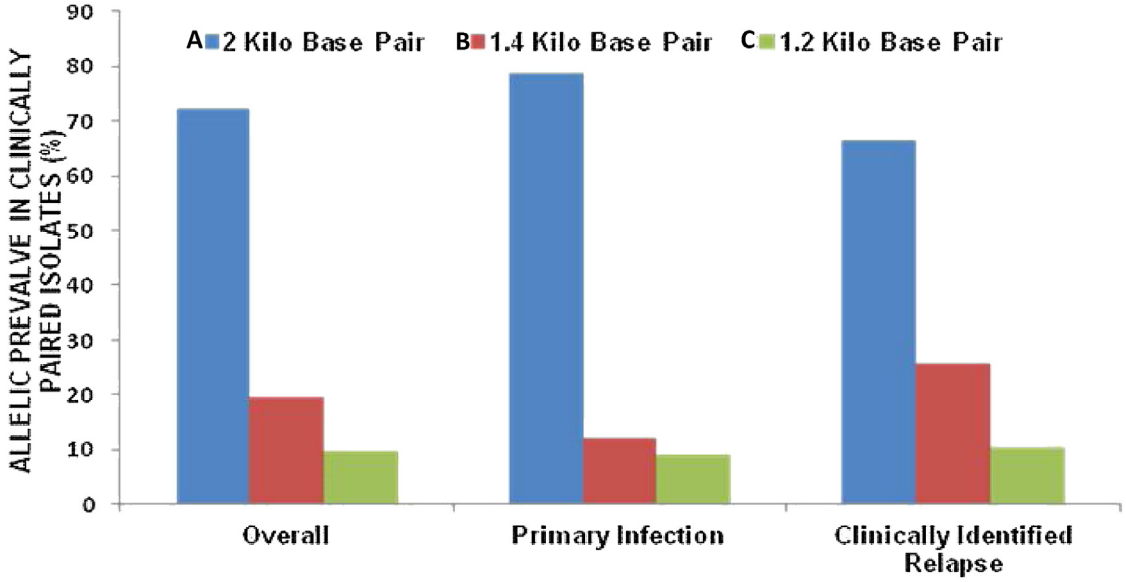

Fig. 1. Frequency distribution of three genotypes (A, B and C) in *P. vivax* clinical isolates during primary and subsequent malaria episodes.

Results

MSP-3α genotyping of the 72 blood samples collected from 33 patients who suffered from multiple malaria episodes revealed either the same or different monoclonal genotypes except in one patient, who suffered from four malaria episodes with a monoclonal genotype (2.0 kb) in the first, second and fourth malaria episode while in third episode a polyclonal genotype (both type 2.0 kb and 1.4 kb) was observed. Based on the size of the PCR products, three major types A (2.0 kb), B (1.4 kb) and C (1.2 kb) were identified. In 72 samples, 72.2% of Type A the most predominant followed by 19.44% of type B and 9.72% of type C were observed. During the first malaria episode, 78% showed type A, 12% type B and 9% type C was observed while 66% of type A, 25% of type B and 10% of type C was observed in samples collected during the subsequent malarial episodes (Fig. 1)

PCR – RFLP

Digestion with Hha I and Alu I yielded fragment sizes that were highly polymorphic among samples. The combination of RFLP patterns allowed 53 parasite genotypes to be distinguished. The PCR–RFLP patterns of all samples with Alu I showed 9 alleles ranging from 150 to 550 base pairs (bp) and Hha I showed 11 alleles ranging from 150 bp to 1.1 kb. In this study, the RFLP patterns of all samples showed size conservation of largest fragments i.e. 550 bp with Alu I and 1.1 Kb with Hha I, not included for distinguishing different alleles while the smaller fragments showing variation in size were incorporated for RFLP analysis. The sum of RFLP fragment sizes was not equal to the size of uncut product in some of the samples. Type A showed 7 and 11, type B showed 6 and 5 and type C showed 4 and 2 allelic variants when digested with Alu I and Hha I respectively. All samples gave multiple positive results on analysis at different time interval ranging from 33 to 355 days. Out of the 39 subsequent samples collected from 33 patients suffering from malaria, 32 (82.05%) had identical and 7 samples (17.94%) had different genotypes compared to primary samples. Out of the 32 samples, 21 had identical RFLP pattern with both Alu I and Hha I and 9 had identical RFLP pattern with Alu I but different with Hha I while 2 samples had different RFLP pattern with both Alu I and Hha I. 7 samples which had different genotypes showed different RFLP pattern with both Alu I and Hha I. Digestion with Hha I revealed 22 different genotypes in type A and 6 in type B and 2 in type C. No allelic group was prevalent in comparison to each other. Alu I digestion revealed 7 different genotypes (A1 to A7) present in 52 samples of type A and 3 genotypes in 14 samples of type B and 2 genotypes in 7 samples of type C. Genotypes A1 and A2 were the most abundant with apparition frequency of 13/52 (25%) and 29/52 (55.76%) of the infections.

Frequency distribution of alleles

Digestion with Alu I showed 9 alleles ranging from 150 to 550 base pairs (bp) and Hha I showed 11 alleles ranging from 150 bp to 1.1 kb. After exclusion of the conserved alleles i.e. 550 bp with Alu I and 1.1 kb with Hha I digestion, the frequency of 150 bp and 250 bp were found to be higher i.e. 93% and 48% with Alu I and Hha I. respectively among remaining alleles (Fig.2).

Analysis of MSP-3α gene sequences

11 sequences (6 primary and 5 of recurrent malaria episodes) of type A, 3 (1 primary and 2 of subsequent episodes) of type B and 4 (3 primary and 1 of subsequent malarial episode) of type C were found to be identical within the allelic class at DNA level. The sequenced region corresponds to the central alanine-rich domain of PvMSP-3α sequence of Belem strain (AF 093584) used as a reference in our study [14]. We studied the phylogenetic relationships among the parasite isolates using coding region of Pvmsp3α gene. Fig.3 shows that all alleles of Type B and C have been originated from single Type A. The types B and C were distributed in different branches shown in phylogenetic tree. Sequence comparisons indicate that type A was found similar to Belem; type B similar to Belem, Chesson and Salvador whereas type C was found to be similar to Chesson and Salvador. Phylogenetic tree comparing sequence types A, B and C with other 14 sequences retrieved from GenBank. Cladogram was built using the entire sequence alignment. The cladogram showed that Venezuela and Thailand although from different geographical regions shared extensive sequence similarity showing a high bootstrap value. The cladogram clearly showed similarity between type A allele and Bangladesh sequence (99%). Type B allele was found to be closer to N. Korea and Sri Lanka (93%) and type C was found to be closer to Salvador (98%), Papua New Guinea and Thailand (Fig.4). The deduced peptide sequences of the PvMSP-3α showed

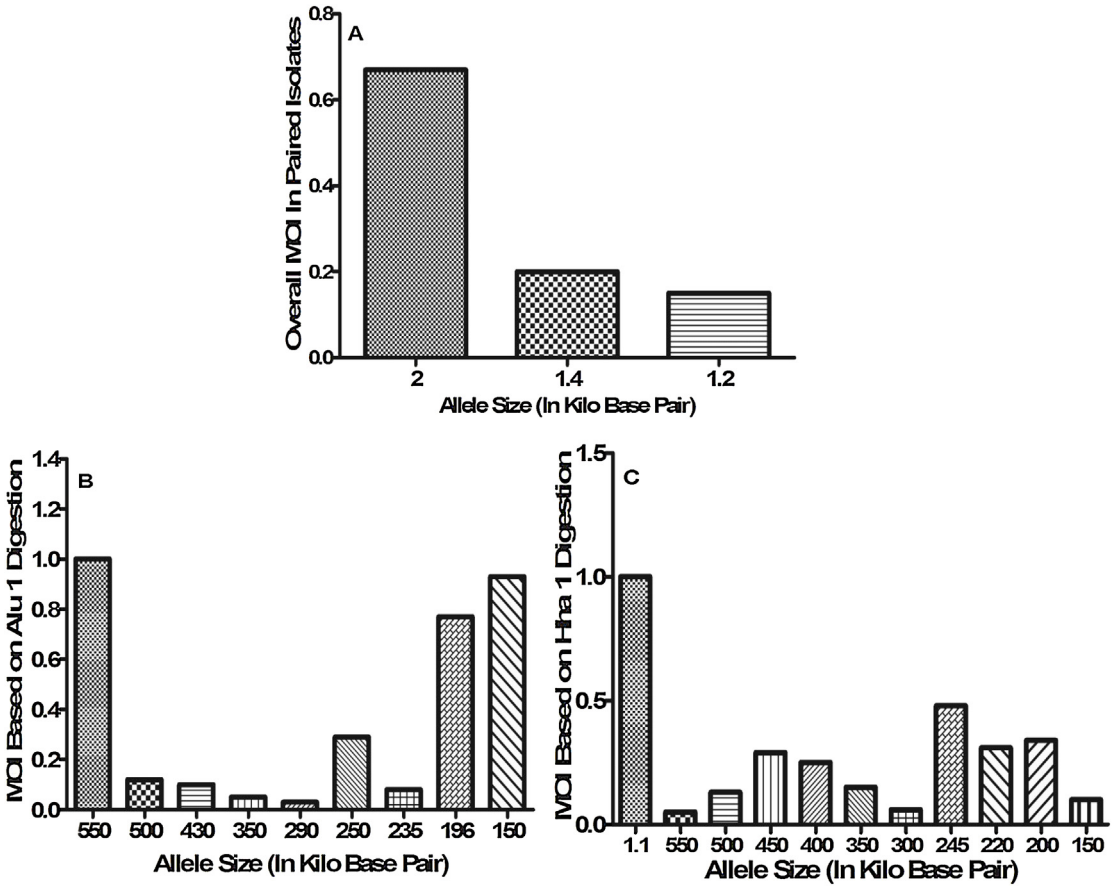

Fig. 2. (A) Overall multiplicity of infection in paired isolates. Frequency distribution of alleles digested with (B) Alu I and (C) Hha I at *Pvmsp-3α* among *P. vivax* clinical isolates.

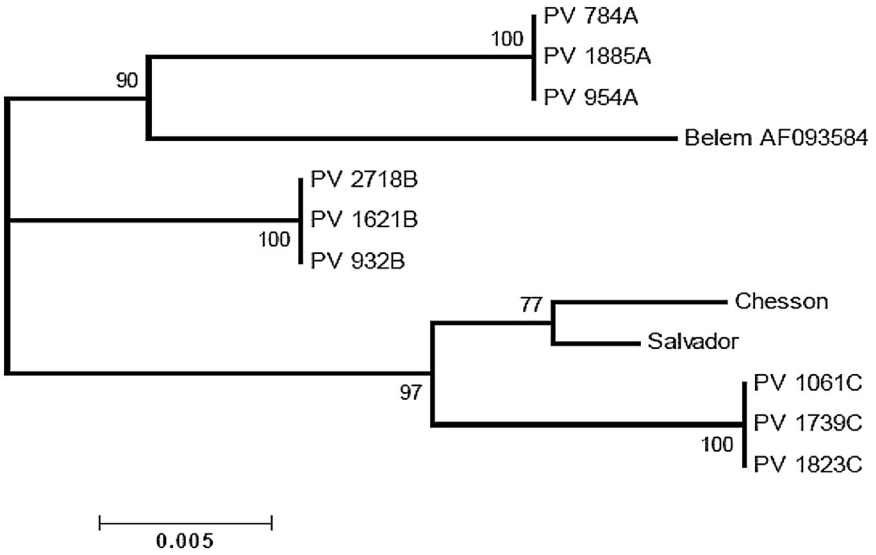

Fig. 3. Comparison of phylogenetic relationships among 9 sequences (3 of each A, B and C) of *Pv*msp3α gene from Indian clinical isolates with that of Belem, Chesson and Salvador strain.

a large number of amino acid substitutions. When compared to Belem, the sequences of types A, B and C showed 21, 20 and 28 nucleotide substitutions respectively. Out of these 6, 8 and 9 were synonymous and 15, 12 and 19 were found to be non-synonymous mutations in A, B, and C types respectively. As a result of these non-synonymous substitutions at second codon positions, most of the polymorphic sites were dimorphic or trimorphic. In major, distinct blocks shared by different alleles discriminated in three fragments at different nucleotide (nt) positions have been distinguished. Amino acid sequence LIETELAI in type A and MFETEIAI in type B at nt positions 16–36 were found instead of SIETEIAV in Belem while type C was having SIETEIAI amino acid motif similar to Chesson and Salvador. KAKEASD at nt positions 238–258 was observed only in type A in place of NAKDASE in Belem, Ches-

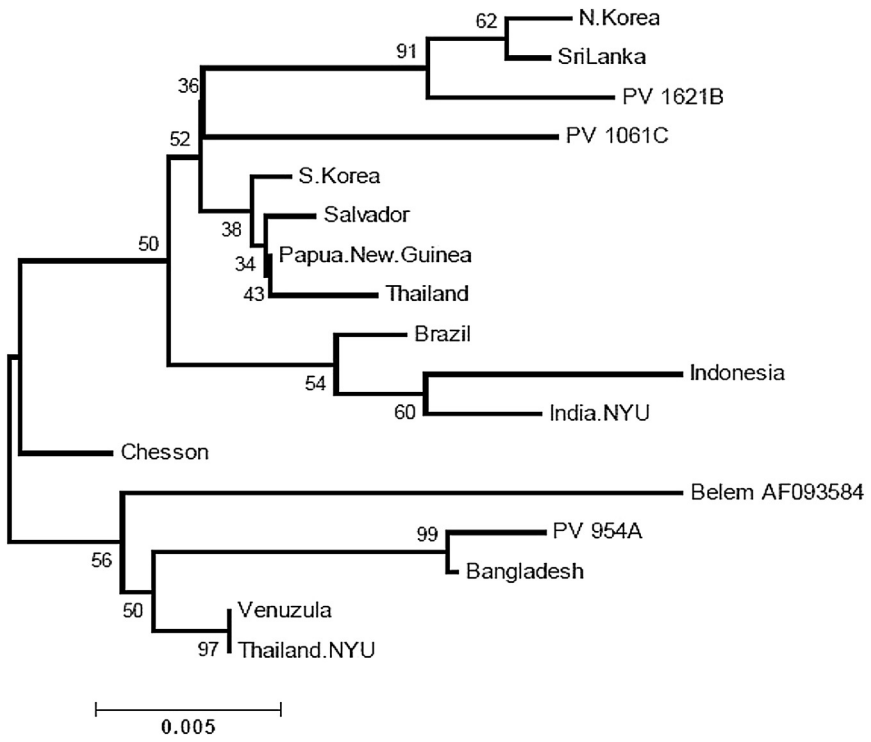

Fig. 4. Phylogenetic tree of *Pv*MSP-3α sequences constructed using the Neighbor- joining method. Tree was constructed using 14 available sequences and 3 representative sequences of classes A, B and C of Indian clinical isolates.

son and Salvador. At another site, the amino acid sequences were found to be dimorphic with either LSKLEE (Belem and Chesson) or MSELEK (Salvador). At nt positions 364–384 type A possess LSKLEE similar to Belem and Chesson while types B and C possess MSELEK similar to Salvador. At nt positions 505–528, types A and B possess KEATAAKL similar to amino acid motif of Belem. On the other hand, type C possess TAANVVKL which is similar to Chesson and Salvador having TAANVVKD but with a difference in the motif at the last amino acid from D to L. Amino acid motif AKKAE in Belem, Chesson and Salvador is replaced by SKKAK found solely in Type C isolates at nt positions 703–717, and another motif NVARS was found in all three types (A, B and C) at nt positions 775–789 in place of DVARA in Belem. Moreover, non-synonymous mutations were observed solely in type A at nt positions 217 (E→K), 499 (A→T), 733 (A→V), in type B at nt position 64 (N→Y) and in type C at nt positions 541 (A→V) and 601 (A→R). In addition, Single Nucleotide Polymorphism (SNPs) were observed at nt position 604 (A→E) both in type A and C and at nt position 100, (N→K) at 121 (K→E) and at 802 (T→A) in all the three types (A, B and C) when compared with Belem (Fig.5).

Discussion

It has been suggested that since the largest size class i.e. type A occurs at the highest frequency in all sampled natural populations, the deletions found in smaller size classes may result in a loss of fitness for those parasites which carry them [12,13,28] that is in concord with our results in which type A (72.2%) was the most predominant followed by B and C (19.44% and 9.72%) respectively. The frequencies of the three Pvmsp3-alpha types were consistent with those found in Papua New Guinea and Thailand [9,13,29]. The MSP-3α gene seems to be a good candidate for studying the genetic diversity of *P.vivax* populations. In evaluating genetic diversity and dynamics of the parasite population in asymptomatic patients in Papua Guinea, Bruce and others [9,13] identified 24 alleles of the MSP-3α gene in *P.vivax* infected isolates based on PCR-RFLP analysis using two restriction enzymes, whereas in our study, 20 alleles were identified suggesting that *P.vivax* population was equally diverse. Many factors contribute towards the genetic diversity of *P.vivax* population, one of them being relapse and gametocytemia which favor the cross-fertilization and meiotic recombination of distinct parasites within the vector and the other being drug resistance which may affect the duration of infection and enhance genetic recombination in mosquitoes [13]. Recurrent samples from the same patient were taken which could have been either relapse or newly acquired infection. Samples taken from the same patient at different time intervals were analyzed enabling us to assess the turnover of paired samples of *P.vivax* parasite populations collected over 33 and 355 day intervals. In our study, subsequent samples of 7 patients collected over 45–150 days intervals showed different genotypes. The rate of turnover depends upon the rate of acquisition of new infection and rate of clearance of existing infections. Clearance rates depend on immunity that is species and most probably genotype specific [30]. It was difficult to determine the epidemiologic and clinical significance of each allelic type of the Pvmsp3α locus and the types didn't clearly correlate with the different incubation period of patients (short and long) which is well known in reemerging *P.vivax* malaria. In our study, 82% of subsequent samples showed same genotype indicating treatment failure/relapse and 18% showed different genetic composition indicating new infection/relapse with a different parasite. Although, it has been concluded that parasites associated with primary infection and those with relapse are usually of a similar genetic composition, evidence for novel genotypes in *P. vivax* associated with relapses was obtained for 24% of the patients (1 of 6 patients in Canada [31], 2 of 10 patients in Brazil [32] and 2 of 5 patients in Thailand) [22,9] inferred re-infection with the same genotype rather than persistence of an infection in subsequent samples from same patient who showed a consistent RFLP pattern was in accordance with our results where almost out of 82% uncut recurrent samples, 77% and 53% samples when digested with Alu I and Hha I respectively from same patients showed consistent

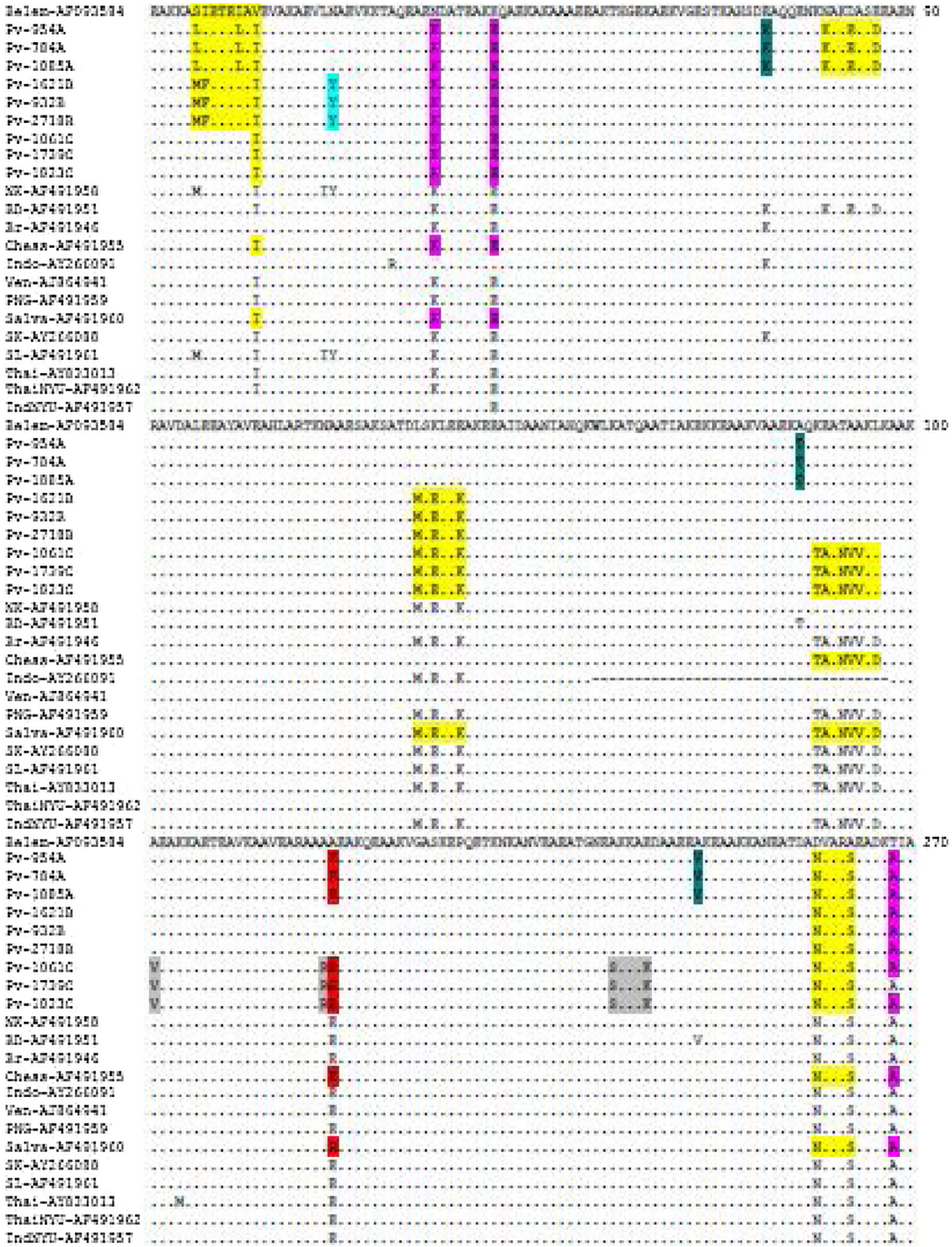

Fig. 5. Alignment of deduced amino acid sequences of *Pv*MSP-3α of the three major classes A, B and C with the 14 sequences available from GenBank database. The alignment of the sequences is based on that of Belem strain. Residues highly conserved in the sequences are shown by dots and polymorphic regions are highlighted.

pattern at different time intervals. All field isolates showed a conserved 1.1 kb band after Hha I digestion and a conserved 550 bp band after Alu I digestion which was in accordance with the results of [9] and [12]. Mixed infection was observed only in 1 out of 72 samples which are not in concomitant with the results observed by [9] who found mixed infections at a greater than expected frequency i.e. 6 parasite genotypes in a single host in Papua New Guinea. Moreover the pro9portion of mixed gene infections estimated in Papua New Guinea, India and Thailand ranges from 30% to 65% [13,30,33,34] whereas in Northern isolates from Iran, no mixed genotype infections have been detected by [35] as of yet, suggesting that the difference in multiple infection rates in different malaria endemic regions may reflect the difference in malaria transmission intensity [36]. In our study, the levels of divergence within the allelic classes were found to be much lower/negligible than those between allelic classes, in contrast to the analysis done by [37] who found nine distinct haplotypes among sequences within each of the three size classes depicting higher intra allelic genetic

diversity in the isolates from Venezuelan Amazon. The three major types based on PCR length polymorphism were found to be mixed by [13] suggesting that the deletion mutations may have arisen in multiple parasite lines which is in contrast to our study in which three major types A, B and C had fallen into separate groups. The very low level of sequence diversity within types B and C alleles in our study suggests that all alleles of each given size class have been originated from a single type A allele which is supported by the predicted phylogeny of the conserved C terminus domain [28]. Found many of the closely related sequences were isolates from different geographical origins i.e. clustering of certain South American and Thai isolates was in concordance with our study in which the clustering between isolates from Venezuela and Thailand NYU was observed. The lack of clustering amongst sequences from the same geographical region suggests firstly the presence of an ancient *P. vivax* population, secondly the independent convergence of polymorphic variants in a recent population and thirdly the extensive exchange of parasite populations [38]. Analysis of nucleotide substitutions indicated that different selection forces may have acted on the molecule and shaped the organization of different domains of the gene. Comparison of synonymous and non-synonymous substitutions across the *Pv*MSP-3α gene in our study only revealed dN > dS (not statistically significant, $p < 0.098$) suggesting that this polymorphic domain might be under positive selection which is in accordance to the analysis done by [12]. The positive selection is thought to be maintaining most of the observed polymorphism in malarial antigens, except for *Pf*MSP5 and *Pv*MSP 4 suggesting that they might be exposed to different degrees of diversifying selection such as immune system pressure [38]. New dimorphic sites found in our isolates have not been reported earlier. LIETELAI in type A or MFETEIAI in type B in place of motif SIETEIAV present in Belem and another motif KAKEASD found only in type A isolates while Belem having NAKDASE implying that isolates can have any combination of these two motifs involving recombination between variants. Consistent with other studies [12,13,28] all non-synonymous mutations under positive selection pressure in the conserved C- terminus domain were clustered into two dimorphic sites LSKLEE/MSELEK and KEATAAKL/TAANVVKL, suggesting these conserved motifs are under functional and/or structural constraints among global *P. vivax* populations. Considering these together, this domain of the Pvmsp3α gene may be a candidate for a *P. vivax* vaccine target, particularly as this region is known to be immunogenic and induce merozoite invasion-blocking antibodies. The distribution and nature of polymorphism suggest that there are functional restrictions on mutations in this gene, and have implication for inclusion of *Pv*msp-3α as a candidate in a *P. vivax* vaccine [12].

Conclusion

In conclusion, reinfection with the same genotype rather than persistence of an infection could be inferred in majority of samples from the same patients who showed a consistent genotype and RFLP pattern even at longer time intervals. Therefore, *Pv*MSP-3α genotyping could be used as one of the molecular tools for differentiating relapse from new infection. Moreover, the nucleotide sequences of the *Pv*MSP-3α gene showed SNPs exclusively in each genotype (A, B and C) making them different from each other, thereby increasing the diversity of *Pv*MSP-3α gene even within a limited geographical region. This extensive allelic diversity of *Pv*MSP-3α gene can be used as a molecular marker for distinguishing field isolates. However, in order to design an effective MSP based vaccine, the genetic diversity needs to be taken into consideration as the immune response targeting one type of antigen may not be effective against other forms of the antigens being expressed by *P.vivax* strains, thereby evaluation of the extent of genetic diversity is must before developing an effective vaccine.

Acknowledgements

Authors are thankful to anonymous subjects consented to participate in this study. We thank the supporting staff team, particularly microscopists at Malaria Clinic, at NIMR for their untiring dedication and exceptional skill in examining the blood smears, laboratory technician for their expertise in clinical sample collection is gratefully acknowledged. The authors are also thankful to Indian Council of Medical Research (ICMR), Delhi for financial assistance for the work.

References

[1] Adak T, Sharma VP, Orlov VS. Studies on the Plasmodium vivax relapse pattern in Delhi, India. Am J Trop Med Hyg 1998;59(1):175–9.
[2] Nandy A, Addy M, Maji AK, Bandyopadhyay AK. Monitoring the chloroquine sensitivity of Plasmodium vivax from Calcutta and Orissa, India. Ann Trop Med Parasitol 2003;97(3):215–20.
[3] Maestre A, Sunil S, Ahmad G, Mohmmed A, Echeverri M, Corredor M, et al. Inter-allelic recombination in the Plasmodium vivax merozoite surface protein 1 gene among Indian and Colombian isolates. Malar J 2004;3:4.
[4] Viriyakosol S, Siripoon N, Petcharapirat C, Petcharapirat P, Jarra W, Thaithong S, et al. Genotyping of Plasmodium falciparum isolates by the polymerase chain reaction and potential uses in epidemiological studies. Bull World Health Organ 1995;73(1):85–95.
[5] Snounou G, Zhu X, Siripoon N, Jarra W, Thaithong S, Brown KN, et al. Biased distribution of msp1 and msp2 allelic variants in Plasmodium falciparum populations in Thailand. Trans R Soc Trop Med Hyg 1999;93(4):369–74.
[6] Arnot DE, Barnwell JW, Tam JP, Nussenzweig V, Nussenzweig RS, Enea V. Circumsporozoite protein of Plasmodium vivax: gene cloning and characterization of the immunodominant epitope. Science 1985;230(4727):815–8.
[7] Rosenberg R, Wirtz RA, Lanar DE, Sattabongkot J, Hall T, Waters AP, et al. Circumsporozoite protein heterogeneity in the human malaria parasite Plasmodium vivax. Science 1989;245(4921):973–6.
[8] Berzins K. Merozoite antigen involved in invasion. In: Perlmann P, Troye-Blomberg M, editors. Malaria immunology, 80. Basel, Karge: Chem Immunol; 2002. p. 125–43.
[9] Bruce MC, Galinski MR, Barnwell JW, Snounou G, Day KP. Polymorphism at the merozoite surface protein-3alpha locus of Plasmodium vivax: global and local diversity. Am J Trop Med Hyg 1999;61(4):518–25.
[10] Bruce MC, Donnelly CA, Alpers MP, Galinski MR, Barnwell JW, Walliker D, et al. Cross-species interactions between malaria parasites in humans. Science 2000;287(5454):845–8.
[11] Galinski MR, Ingravallo P, Corredor-Medina C, Al-Khedery B, Povoa M, Barnwell JW. Plasmodium vivax merozoite surface proteins-3beta and-3gamma share structural similarities with P. Vivax merozoite surface protein-3alpha and define a new gene family. Mol Biochem Parasitol 2001;115(1):41–53.
[12] Rayner JC, Corredor V, Feldman D, Ingravallo P, Iderabdullah F, Galinski MR, et al. Extensive polymorphism in the plasmodium vivax merozoite surface coat protein MSP-3alpha is limited to specific domains. Parasitology 2002;125(Pt 5):393–405.
[13] Cui L, Mascorro CN, Fan Q, Rzomp KA, Khuntirat B, Zhou G, et al. Genetic diversity and multiple infections of Plasmodium vivax malaria in Western Thailand. Am J Trop Med Hyg 2003;68(5):613–9.
[14] Galinski MR, Corredor-Medina C, Povoa M, Crosby J, Ingravallo P, Barnwell JW. Plasmodium vivax merozoite surface protein-3 contains coiled-coil motifs in an alanine-rich central domain. Mol Biochem Parasitol 1999;101(1-2):131–47.
[15] Garnham PC. Swellengrebel lecture. Hypnozoites and 'relapses' in Plasmodium vivax and in vivax-like malaria. Trop Geogr Med 1988;40(3):187–95.
[16] Krotoski WA. Discovery of the hypnozoite and a new theory of malarial relapse. Trans R Soc Trop Med Hyg 1985;79(1):1–11.
[17] Contacos PG, Collins WE, Jeffery GM, Krotoski WA, Howard WA. Studies on the characterization of plasmodium vivax strains from Central America. Am J Trop Med Hyg 1972;21(5):707–12.
[18] Adak T, Valecha N, Sharma VP. Plasmodium vivax polymorphism in a clinical drug trial. Clin Diagn Lab Immunol 2001;8(5):891–4.
[19] Cogswell FB. The hypnozoite and relapse in primate malaria. Clin Microbiol Rev 1992;5(1):26–35.
[20] Coatney GR. Relapse in malaria-an enigma. J Parasitol 1976;62:3–9.
[21] Aileen AC, Ian Winchester, Patricia CMadden, Paul Larcey, Ian WHamley, Corrie TI. Synthesis, thermal characterization and rheological properties of a homologous series of polymethacrylate-based side-chain liquid crystal polymers. Polymer 1998;39(5):1197–205.
[22] Khusmith S, Tharavanij S, Bunnag D. Antigenic disparity of Plasmodium vivax causing initial symptoms and causing relapse. Southeast Asian J Trop Med Public Health 1998;29(3):519–24.
[23] Imwong M, Snounou G, Pukrittayakamee S, Tanomsing N, Kim JR, Nandy A, et al. Relapses of Plasmodium vivax infection usually result from activation of heterologous hypnozoites. J Infect Dis 2007;195(7):927–33.
[24] Chen N, Auliff A, Rieckmann K, Gatton M, Cheng Q. Relapses of Plasmodium vivax infection result from clonal hypnozoites activated at predetermined intervals. J Infect Dis 2007;195(7):934–41.

[25] Singh J. J.S.B. stain; a review. Indian J Malariol 1956;10(2):117–29.
[26] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 2007;24(8):1596–9.
[27] Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 1986;3(5):418–26.
[28] Mascorro CN, Zhao K, Khuntirat B, Sattabongkot J, Yan G, Escalante AA, et al. Molecular evolution and intragenic recombination of the merozoite surface protein MSP-3alpha from the malaria parasite Plasmodium vivax in Thailand. Parasitology 2005;131(Pt 1):25–35.
[29] Kim JR, Imwong M, Nandy A, Chotivanich K, Nontprasert A, Tonomsing N, et al. Genetic diversity of Plasmodium vivax in Kolkata, India. Malar J 2006;5:71.
[30] Bruce MC, Galinski MR, Barnwell JW, Donnelly CA, Walmsley M, Alpers MP, et al. Genetic diversity and dynamics of plasmodium falciparum and P. Vivax populations in multiply infected children with asymptomatic malaria infections in Papua New Guinea. Parasitology 2000;121(Pt 3):257–72.
[31] Craig AA, Kain KC. Molecular analysis of strains of Plasmodium vivax from paired primary and relapse infections. J Infect Dis 1996;174(2):373–9.
[32] Kirchgatter K, del Portillo HA. Molecular analysis of Plasmodium vivax relapses using the MSP1 molecule as a genetic marker. J Infect Dis 1998;177(2):511–5.
[33] Kolakovich KA, Ssengoba A, Wojcik K, Tsuboi T, al-Yaman F, Alpers M, et al. Plasmodium vivax: favored gene frequencies of the merozoite surface protein-1 and the multiplicity of infection in a malaria endemic region. Exp Parasitol 1996;83(1):11–9.
[34] Joshi H, Subbarao SK, Adak T, Nanda N, Ghosh SK, Carter R, et al. Genetic structure of Plasmodium vivax isolates in India. Trans R Soc Trop Med Hyg 1997;91(2):231–5.
[35] Zakeri S, Barjesteh H, Djadid ND. Merozoite surface protein-3alpha is a reliable marker for population genetic analysis of Plasmodium vivax. Malar J 2006;5:53.
[36] Babiker HA, Lines J, Hill WG, Walliker D. Population structure of Plasmodium falciparum in villages with different malaria endemicity in east Africa. Am J Trop Med Hyg 1997;56(2):141–7.
[37] Ord R, Polley S, Tami A, Sutherland CJ. High sequence diversity and evidence of balancing selection in the Pvmsp3alpha gene of Plasmodium vivax in the Venezuelan Amazon. Mol Biochem Parasitol 2005;144(1):86–93.
[38] Gomez A, Suarez CF, Martinez P, Saravia C, Patarroyo MA. High polymorphism in Plasmodium vivax merozoite surface protein-5 (MSP5). Parasitology 2006;133(Pt 6):661–72.