Gene xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Gene

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

Population genetic structure of the *Plasmodium vivax* circumsporozoite protein (*Pvcsp*) in Sri Lanka

Sajani Dias [a], Thilan Wickramarachchi [a,1], Imeshi Sahabandu [a], Ananias A. Escalante [b], Preethi V. Udagama [a,*]

[a] Department of Zoology, Faculty of Science, University of Colombo, Colombo 03, Sri Lanka
[b] Center for Evolutionary Medicine & Informatics, The Biodesing Institute, Arizona State University, Tempe, AZ, USA

ARTICLE INFO

Article history:
Accepted 7 January 2013
Available online xxxx

Keywords:
Sri Lanka
Plasmodium vivax
Circumsporozoite Protein
VK210 variant
Genetic diversity
Phylogeographic analysis

ABSTRACT

Molecular methods elucidate evolutionary and ecological processes in parasites, where interaction between hosts and parasites enlighten the evolution of parasite lifestyles and host defenses. Population genetics of *Plasmodium vivax* parasites accurately describe transmission dynamics of the parasites and evaluation of malaria control measures. As a first generation vaccine candidate against malaria, the Circumsporozoite Protein (CSP) has demonstrated significant potential in *P. falciparum*. Extensive polymorphism hinders the development of a potent malaria vaccine. Hence, the genetic diversity of *Pvcsp* was investigated for the first time in 60 Sri Lankan clinical isolates by obtaining the nucleotide sequence of the central repeat (CR) domain and examining the polymorphism of the peptide repeat motifs (PRMs), the genetic diversity indices and phylogenetic relationships. PCR amplicons determined size polymorphism of 610, 700 and 710 bp in *Pvcsp* of Sri Lanka where all amino acid sequences obtained were of the VK210 variant, consisting variable repeats of 4 different PRMs. The two most abundant PRMs of the CR domain, GDRADGQPA and GDRAAGQPA consisted ~2-4 repeats, while GNRAAGQPA was unique to the island. Though, different nucleotide sequences termed repeat allotypes (RATs) were observed for each PRM, these were synonymous contributing to a less polymorphic CR domain. The genetic diversity of *Pvcsp* in Sri Lanka was due to the number of repetitive peptide repeat motifs, point mutations, and intragenic recombination. The 19 amino acid haplotypes defined were exclusive to Sri Lanka, whereas the 194 *Pvcsp* sequences of global isolates generated 57 more distinct a.a. haplotypes of the VK210 variant. Strikingly, the CR domain of both VK210 and VK247 variants was under purifying selection interpreting the scarcity of CSP non-synonymous polymorphisms. Insights to the distribution of RATs in the CR region with geographic clustering of the *P. vivax* VK210 variant were revealed. The cladogram reiterated this unique geographic clustering of local (VK210) and global isolates (VK210 and VK247), which was further validated by the elevated fixation index values of the VK210 variant.

1. Introduction

Understanding the genetic diversity and population structure of *P. vivax* parasites is essential to accurately describe the transmission dynamics of vivax malaria and to the program evaluation of malaria control measures (Zhong et al., 2011). Population genetics data are crucial, for example, to predict how fast phenotypes of interest, such as novel antigenic variants, particular relapsing patterns or drug resistance, arise and spread in natural populations (Zilversmit and Hartl, 2005). Hence, for a clear quantitative definition of genetic diversity, different genetic markers and parasite typing systems are used. The well-characterized polymorphic regions in both pre-erythrocytic and erythrocytic genes have been extensively used to analyze genetic diversity patterns in *P. vivax* populations both locally and globally (Thakur et al., 2008; Wickramarachchi et al., 2010; Patil et al., 2010; Dias et al., 2011a, b; Premaratne et al., 2011). A further branch of research using molecular markers is the study of Molecular Ecology. It has important practical applications in the field of evolution and control of drug resistance (Criscione et al., 2005).

One such genetic marker is the sporozoite's major surface protein, the circumsporozoite protein (CSP), which is multifunctional in nature. It mediates sporozoite development and the initial interactions between the sporozoite and its two hosts, the mosquito mid gut and the mammalian liver (Sinnis and Nardin, 2002; Coppi et al., 2011). A

Abbreviations: API, Annual parasite incidence; AZB, Azerbaijan; CR, Central repeat; CSP, Circumsporozoite protein; DBP, Duffy binding protein; DNA, Deoxyribonucleic acid; dNTP, Deoxyribonucleotide triphosphate; F_{ST}, Fixation index; LD, Linkage disequilibrium; ML, Maximum likelihood; PCR, Polymerase chain reaction; PRM, Peptide repeat motif; PvCSP, *P. vivax* Circumsporozoite protein; RAT, Repeat allotype; RFLP, Restriction fragment length polymorphism; UV, Ultra violet.

* Corresponding author at: Department of Zoology, Faculty of Science, University of Colombo, No 94, Cumaratunga Munidasa Mawatha, Colombo 03, Sri Lanka. Tel.: +94 11 2503399; fax: +94 11 2503148.
E-mail address: dappvr@yahoo.com (P.V. Udagama).
[1] Current Affiliation: Microtech Biological (Pvt) Ltd, No 211/1C, Palanwatta Road, Rathmaldeniya, Pannipitiya, Sri Lanka.

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.01.003

notable feature of this protein is that it consists of tandem arrays of relatively short amino acid motifs (Brito and Ferreira, 2011). The structure of CSP is highly conserved among *Plasmodium* species infecting rodents, primates, and humans. CSPs comprise of a central repeat (CR) region that is diverse across *Plasmodium* species, and flanking the repeats are two conserved domains: region I, a 5' amino acid (a.a) sequence at the N terminus of the repeats, and a known cell-adhesive motif C-terminal to the repeats termed the type I thrombospondin repeat (Coppi et al., 2011).

The genotyping of *P. vivax* infections based on the *csp* gene in various geographical areas were evaluated by the use of various tools; i) combination of Polymerase Chain Reaction and Restriction Fragment Length Polymorphism (PCR/RFLP) (Kho et al., 1999; Imwong et al., 2005; Kim et al., 2006; Manamperi et al., 2008; Storti-Melo et al., 2009; Zakeri et al., 2010) and ii) sequencing (Mann et al., 1994; Kho et al., 1999; Lim et al., 2001; Kim et al., 2002; Cui et al., 2003; Leclerc et al., 2004; Zakeri et al., 2006; Patil et al., 2010; Souza-Neiras et al., 2010).

P. vivax CSP (PvCSP) displays two major types of peptide repeat motifs (PRMs), each consisting of nine amino acids, GDRA[D/A]GQPA and ANGAGNQPG, defining variants known as VK210 and VK247, respectively (Rosenberg et al., 1989). Both VK210 and VK247 variants are globally distributed, but geographic biases have been described (Cochrane et al., 1990; Qari et al., 1993a, b; Leclerc et al., 2004; Zakeri et al., 2006; Patil et al., 2010). A third type of repeat unit (APGANQ[E/G]GAA), identical to that described for *P. simiovale*, characterizes the so called *P. vivax*-like parasites (Qari et al., 1993a). Insertions and deletions in the central repeat domain, from either sexual recombination during meiosis or intrahelical strand-slippage events during mitotic DNA replication (McConkey et al., 1990), generate novel CSP variants that may be positively selected if the mutant parasites evade host immunity.

Studies on PvCSP in Sri Lanka are scanty, where Mendis et al. (1992), investigated the anti-circumsporozoite protein antibodies against a peptide consisting the GDRADGQPA and GDRAAGQPA repeat motifs in *P. vivax* patients from Kataragama, an endemic area in Sri Lanka. Though an age related prevalence pattern was observed, the pattern was not an age acquired response, rather due to differences in inoculation rates in different age groups, and the antibody prevalence was less than 20%, suggesting that the peptide was less immunogenic in these individuals. A pilot study by Manamperi et al. (2008), demonstrated size polymorphism of *Pvcsp* by a nested polymerase chain reaction (PCR) at the central repeat domain using 12 blood smear slides of *P. vivax* patients from Kataragama. Four size variants ranging from ~600-750 bp were observed.

Hence, we investigated in depth the genetic diversity of the *Pvcsp* gene in Sri Lanka, where low transmission and unstable malaria prevails, by tests of diversity, natural selection and recombination. Further, both local and global *P. vivax* isolates were analyzed for patterns of sequence variation in the *csp* gene by examining the polymorphism of the peptide repeat motifs (PRMs). Relationships of the *Pvcsp* worldwide isolates were traced as reconstructing the historical geography of isolates and identifying different genetic subdivisions within species are major objectives of phylogeographical analysis (Avise, 2000).

2. Materials and methods

2.1. P. vivax clinical isolates

Ethical clearance was obtained for this study from the Ethical Review Committee of the Faculty of Medicine, University of Colombo (EC/10/131). Previously collected blood samples prior to anti malarial therapy from December, 1998 – March, 2000 were used to extract *P. vivax* parasite DNA (Wickramarachchi et al., 2010). These isolates were collected from two malaria endemic areas, Anuradhapura (8°22'N, 80°20'E; N = 20) and Kataragama (6°25'N, 81°20'E; N = 32) and from a malaria non-endemic area, Colombo (7°55'N, 79°50'E; N = 30). Patients from Colombo were infected during visits to regions of the island with *P. vivax* transmission, including the two endemic areas (Fonseka and Mendis, 1987; Dias et al., 2011a). During 1995–2000 the annual parasite incidence (API) per 1000 due to *P. vivax* was 0–1.25, 20–40 and 80–160 in Colombo, Anuradhapura and Kataragama, respectively (Briët et al., 2003). Parasitemia ranged from 0.0001 to 0.025%.

Single clone infections confirmed by genotyping at the polymorphic *Pvmsp-3α* locus using a combined nested polymerase chain reaction (PCR) and restriction fragment length polymorphism (RFLP) (Wickramarachchi et al., 2010), were used for this study.

2.2. Amplification and sequencing of the P. vivax circumsporozoite protein

The CR domain of *P. vivax csp* was amplified by nested PCR essentially as described by Imwong et al. (2005). A fragment of ~680-bp was amplified in a total volume of 20 μl using 1 μl of template DNA in the primary reaction and 0.4 μl of the primary PCR product in the nested reaction. Both reactions contained 0.4 units of Taq polymerase (Promega, USA), 0.25 μM of each primer, 0.125 mM dNTPs with 0.8 mM $MgCl_2$. Amplified products were visualized by UV on a 0.8% agarose gel containing 0.25 mg/ml of ethidium bromide and purified using the QIAquick PCR purification kit (Qiagen, USA). Sequencing of the PCR products were done on both strands using VCS-NF and VCS-NR primers (Imwong et al., 2005) at Macrogen, Korea. Some isolates underwent four-fold sequence coverage to confirm rare polymorphisms.

2.3. Data analysis

The consensus nucleotide sequences were generated by Seqman II (DNAStar, Madison, Wisconsin, USA) software and aligned manually. Different nucleotide sequences encoding the same peptide repeat motif, termed repeat allotypes (RATs) were recorded according to Patil et al. (2010).

Measures of genetic polymorphism, *i.e.* polymorphic sites (S), nucleotide diversity (π) and haplotype diversity (h) were calculated using DnaSP version 5.1 (Librado and Rozas, 2009). Natural selection was determined by the difference between the non-synonymous and synonymous substitutions (Dn − Ds) estimated by modified Nei and Gojobori's method (MEGA 5, Escalante et al., 2004; Tamura et al., 2011) and the statistical analyses of recombination and linkage disequilibrium (LD) were tested by DnaSP 5.1 (Librado and Rozas, 2009).

Using all available *Pvcsp* global sequences (VK210 and VK247 strains) from the GeneBank and sequences from Sri Lanka generated from the current study, the overall similarities were determined by using phylogenetic methods; i) the construction of a phylogenetic tree using the Maximum Likelihood (ML) method with the General Time Reversible substitution model using a discrete Gamma distribution (+G) with 5 discrete gamma rate categories (MEGA 5, Tamura et al., 2011) and ii) calculating the genetic differentiation of the VK210 strain between geographic regions *via* F_{ST} values using DnaSP 5.1 (Librado and Rozas, 2009). The best substitution model used for constructing the phylogenetic tree was selected by the DNA/Protein model in MEGA 5. Of the 24 different nucleotide substitution models, model with the lowest BIC (Bayesian Information Criterion) score was considered to describe the substitution pattern the best. Further, the phylogenetic tree was drawn without taking recombination into account, as the study was based on a single *P. vivax* taxon (Arenas and Posada, 2010). It is worth noting that whereas tandem repeats should not be used for establishing ancestry (Escalante et al., 1995), in the context of this investigation, these methods were used to obtain a representation of the similarities among RAT structures.

Please cite this article as: Dias, S., et al., Population genetic structure of the *Plasmodium vivax* circumsporozoite protein (*Pvcsp*) in Sri Lanka, Gene (2013), http://dx.doi.org/10.1016/j.gene.2013.01.003

3. Results

3.1. Genotyping of Pvcsp in Sri Lanka

Amplification of *Pvcsp* in 82 Sri Lankan isolates at the CR region resulted in 3 approximate different product sizes of 610, 700 and 710 bp (Fig. 1). The 710 bp product (N = 34; 41%) was significantly ($P<0.01$) more common than the 700 bp (N = 32; 39 %) and the 610 bp (N = 16; 20%) products, in all three study areas.

From the 82 *P. vivax* isolates, 60 (Anuradhapura, N = 17; Kataragama, N = 29; Colombo, N = 14) *Pvcsp* nucleotide sequences were successfully obtained consisting of an amino acid region corresponding to that reported by Leclerc et al. (2004) (GeneBank accession numbers JQ362595 - JQ362654). All of these 60 *P. vivax* isolates were of the VK210 variant consisting of variable repeats of 4 different PRMs, GDRADGQPA (PRM1), GDRAAGQPA (PRM2), GNRAAGQPA (PRM3) and GNGAGGQAA (PRM4) (Fig. 2). Among the 60 PvCSP amino acid sequences, three notable sequences with varying lengths were observed (137 a.a., 128 a.a., and 119 a.a. sequence). The 128 a.a. sequence (N = 57; C2- C17) was significantly more common than both the 137 (N = 1; C1) and the 119 (N = 2; C18 & C19) a.a. sequences in the study population. The Sri Lankan isolates corroborated the variations in the peptide repeat motifs GDRADGQPA (PRM1) and GDRAAGQPA (PRM2) of other global isolates, with different alternations of non-synonymous codons GCT or GAT, respectively, coding for Alanine (A) or Aspartic acid (D) (Leclerc et al., 2004; Zakeri et al., 2006; Patil et al., 2010). Strikingly, the PRM, GNRAAGQPA (PRM3), observed in this study is unique to the island and is reported for the first time (Fig. 2).

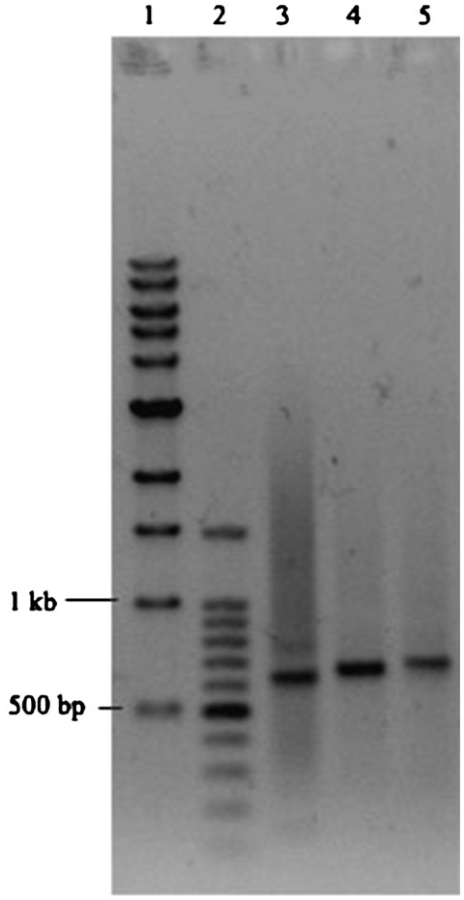

Fig. 1. Size polymorphism of *Pvcsp* of Sri Lankan clinical isolates (WxH = 6 x 10 cm). Lanes 1–5 represents respectively, the 1 kb DNA ladder, 100 bp DNA ladder, the 610 bp, 700 bp and 710 bp amplicons of *Pvcsp*.

Different nucleotide sequences, with synonymous substitutions encoding the same PRM are termed repeat allotypes (RATs) (Rich et al., 1997). Four each from the eight and five RATs identified by Patil et al. (2010) for GDRADGQPA and GDRAAGQPA, respectively were detected from our study, where GGAGACAGAGCAGATGGACA**G**CCAGCA of the former a.a. sequence differed only by a single nucleotide polymorphism (Table 1). Except for a single isolate, where the nucleotide sequence of GNGAGGQAA comprised a single nucleotide polymorphism (GGAAATGGTGCAGGTGGACA**T**GCAGCA), the rest of the Sri Lanka isolates coincided with the nucleotide sequences observed by Patil et al. (2010) (Table 1). Conversely, the polymorphisms of the PRMs giving rise to PRM1-4 are types of RATs generated by non-synonymous substitutions.

The 19 amino acid haplotypes defined from the Sri Lankan population were exclusive to the island. Of these, 52% (N = 31) was of the C4 a.a. haplotype. Except for the C10, C19 and C16 a.a. haplotypes consisting of 6, 5 and 2 isolates, respectively, the rest of the a.a. haplotypes represented a single isolate each (Supplementary Table S1).

The 194 world wide isolates of VK210 variant obtained for this study from Azerbaijan, Brazil, China, Gabon, Iran, Korea, the Philippines, and the Solomon Islands, defined 57 a.a. haplotypes (Supplementary Table S1 and Fig. 2). Apart from the three main peptide repeat motifs GDRA(A/D)GQPA and GNGAGGQAA, 10 other peptide repeat motifs (PRM) were observed. The PRM, GNGAGGQPA, was observed in only a single isolate each from Korea and China, while GDGAAGQPA was observed in 12 Korean and 05 Chinese isolates consisting of 2–5 repeat units and a single repeat unit consisting in an isolate from Azerbaijan towards the 5' end (Mann et al., 1994; Kho et al., 1999; Kim et al., 2002; Leclerc et al., 2004). Three other PRMs were identified; i) **R**DRAAGQPA in a single isolate each from Iran and Azerbaijan, ii) GDRA**P**GQPA unique to 2 Iranian isolates and iii) G**N**RADGQPA unique to 09 Brazilian isolates. Two non-repeat peptide motifs of GDRPAGQPA and ANGAGGQAA were observed from a single isolate from Gabon, and YRASGQPAA from a single isolate from Azerbaijan. Except for 28 Brazilian isolates manifesting 1–3 repeat units of GDRAAGQ**A**A, all the local and global a.a. sequences included the GDRAAGQ**P**A post repeat unit at the 3' end, identical to that of the VK210 variant (Fig. 2).

3.2. Genetic diversity and phylogenetic relationship of Pvcsp in isolates of Sri Lanka and geographically different countries

As the VK210 isolates acquired from different geographic regions of the world consisted of different sequence lengths, polymorphic sites and diversity was calculated accordingly from the nucleotide sequence range of 1–411 bp of the central repeat region. Genetic diversity indices were calculated for local parasite populations (N = 60) and for those of Brazil (N = 52), Iran (N = 113) and Korea (N = 12).

The 60 *Pvcsp* sequences of Sri Lankan isolates revealed 41 polymorphic positions with 10 singleton sites, 25 parsimony informative sites with two variants and 6 parsimony informative sites with three variants giving rise to 47 point mutations (Supplementary Table S1). The Brazilian isolates revealed the majority of polymorphic sites (71) and mutations (76) compared to Iran (43, 46) and Korea (31, 32). Elevated nucleotide diversity per site (π) was recorded from the Korean isolates ($\pi = 0.047 \pm 0.008$S.D.) followed by isolates from Brazil, Sri Lanka and Iran. A majority of a.a. haplotypes (N = 23) were defined by those isolates of Brazil where an elevated haplotype diversity (Hd) of 0.955 ± 0.016 S.D. was also recorded, whereas the 19 a.a. haplotypes of Sri Lanka recorded a Hd of 0.903 ± 0.033 S.D. However, though the 5 Korean a.a. haplotypes generated a high Hd (0.894 ± 0.078 S.D.), the 17 Iranian a.a. haplotypes generated a lower Hd of 0.748 ± 0.031 S.D.

The Sri Lankan population demonstrated a significant purifying selection at the central repeat region of PvCSP ($Dn - Ds = -0.044 \pm 0.014$ S.D.; Z test $P<0.05$), where excess of synonymous substitutions per site ($Ds = 0.059 \pm 0.014$ S.D.) were observed compared to non-

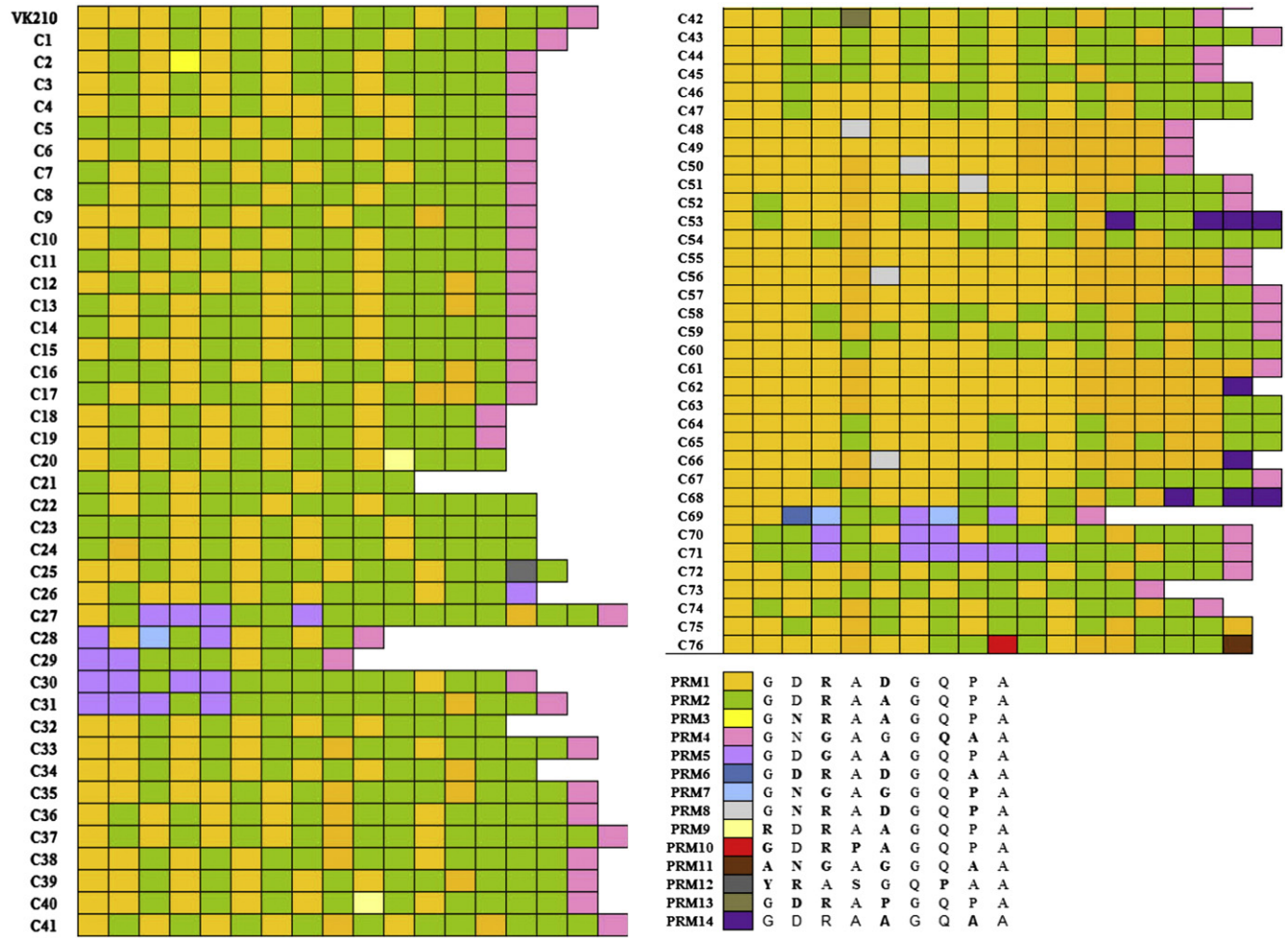

Fig. 2. Distribution of the peptide repeat motif amino acid sequences of the central repeat domain (CR) of PvCSP in geographically diverse malaria endemic areas (WxH = 16.9 x 24 cm). C1-C76 denotes the a.a. haplotypes. VK210 denotes the reference sequence of PvCSP (Accession No. M28745; Leclerc et al., 2004); C1-C19 from Sri Lanka, C20 –C26 from Azerbaijan, C27-C31 from South Korea, C32-C45 from Iran, C46-C68 from Brazil, C69-C71 from China, C72-C73 from the Philippines, C74-C75 from the Solomon Islands, C76 from Gabon. PRM (Peptide Repeat Motif) 3 was unique to Sri Lanka.

synonymous substitutions (Dn = 0.015 ± 0.004 S.D.). Similarly, a purifying selection was observed by each of the parasite populations from Brazil (Dn − Ds = −0.107 ± 0.023 S.D.; Z test $P<0.05$), Iran (Dn − Ds = −0.044 ± 0.024 S.D.; Z test $P<0.05$) and Korea (Dn − Ds = −0.058 ± 0.029 S.D.; Z test $P<0.05$).

Similarly, the VK247 variant from both Western Colombia (N = 26; Dn-Ds = −0.041 ± 0.017 S.D.; Z test P<0.05) and Iran (N = 23; Dn-Ds = −0.082 ± 0.042 S.D.; P<0.05) observed a purifying selection at the CR domain.

Table 1
Peptide repeat motifs and the different nucleotide sequences of the repeat allotypes (RATs) of the central repeat domain of *P. vivax* CSP in the Sri Lankan isolates.

Peptide Repeat Motif	Nucleotide Sequences of Repeat Allotype (RAT's)
GDRADGQPA	1. GGAGACAGAGCAGATGGACAGCCAGCA
	2. GGTGATAGAGCAGATGGACAGCCAGCA
	3. GGCGATAGAGCAGATGGACAGCCAGCA
	4. GGAGACAGAGCAGATGGACA**G**GCCAGCA
GDRAAGQPA	1. GGTGATAGAGCAGCTGGACAACCAGCA
	2. GGAGATAGAGCAGCTGGACAACCAGCA
	3. GGAGATAGAGCAGCTGGACAGCCAGCA
	4. GGTGATAGAGCAGCTGGACAGCCAGCA
GNRAAGQPA	1. GGAAATAGAGCAGCTGGACAGCCAGCA
GNGAGGQAA	1. GGAAATGGTGCAGGTGGACAGGCAGCA
	2. GGAAATGGTGCAGGTGGACA**T**GCAGCA
	3. GGAAATGG**A**GCAGGTGGACAGGCAGCA

Note: bold and underlined nucleotide positions were polymorphic but depicted the same peptide repeat motif.

The minimum number of recombination events between adjacent polymorphic sites (Rm) in the local population was 13, where as those of the VK210 variant from Brazil, Iran and Korea recorded 23, 7 and 4 Rm sites, respectively. A linkage disequilibrium (LD) was maintained across the *Pvcsp* for the entire local population corroborating Patil et al. (2010). Though a similar LD was observed for the Iranian population, the Korean population indicated equilibrium inter-site linkage with increasing nucleotide distance.

The degree of genetic differentiation between populations was estimated by the F_{ST} values, where only a 219 bp region of the 789 bp's of *Pvcsp* were common to all the VK210 worldwide isolates. The Sri Lankan population showed the highest F_{ST} values with Iran ($F_{ST}=0.521$) followed by Brazil ($F_{ST}=0.509$) and Korea ($F_{ST}=0.434$). Similar F_{ST} values were recorded between Iran/Korea ($F_{ST}=0.437$), Brazil/Korea ($F_{ST}=0.57$) and Iran/Brazil ($F_{ST}=0.442$).

The phylogenetic tree drawn using all available VK210 and VK247 worldwide isolates clearly demonstrated two separate clades of the strains and geographic clustering of isolates within each strain (Supplementary Figs. S1). Within the VK247 population, though isolates of Western Colombia and Iran were clearly separated from one another, these had branched off from the same point of origin.

The isolates of *Pvcsp* VK210 variant defined 19 distinct groups (Supplementary Figs. S1). It is worth noting that most of the isolates

were geographically clustered in to groups, which is expected if the pattern of PRM's allow for a quick divergence among populations. Of the Sri Lankan isolates, 2 from Kataragama were exclusively clustered in group 3 while the rest depicted no clustering. Of the global isolates, majority belonged to Group 17, consisting exclusively of 13 Brazilian isolates. Groups 1 and 2 were exclusive for 2 Korean isolates each, while Groups 4 and 5 consisted of the AZB subtypes 2, 3 and 5, 6, respectively. Some Iranian isolates uniquely clustered in groups 6–9 and 13, whereas the Brazilian isolates uniquely clustered in groups 12, 14–19, where Group 16 included 3 sub groups; i) with Brazilian CSP types III, IVa, ii) CSP type IVb and iii) CSP type V. Group 10 included 2 sub groups, one consisting an isolate from China (CH4) and the other with an isolate each from Iran and China (CH5), while group 11 consisted 2 sub groups with one representing the VK210 reference strain and an isolate each from the Solomon islands and the Philippines (Supplementary Figs. S1).

Focusing solely on the RATs, it was observed that their distribution correlate with the geographic clustering of the isolates depicted by the phylogenetic tree drawn (Fig. 2, Supplementary Figs. S1). The main PRM circulating within the Sri Lankan, Korean, Chinese and Iranian isolates is the PRM2, followed by PRM1, where as the PRMs circulating in the Brazilian isolates was vice versa (Fig. 2). The pattern of the two repeats coincides with the distribution of the isolates in the phylogenetic tree, where all isolates of Sri Lanka, Korea, China and Iran was clustered within one part of the phylogenetic tree, whereas the Brazilian isolates were clustered within the other (Fig. 2, Supplementary Figs. S1). Among the isolates belonging to Sri Lanka, Korea, China and Iran, the following was observed on the PRM pattern; i) Sri Lankan isolates of a.a. C1-C19 was highlighted in the phylogenetic tree clustered separately, ii) Korean isolates of a.a. haplotypes C28-C31 were represented in groups 3 and 4, while those isolates of Korea and China belonging to a.a. haplotypes C27, C70 and C71 were in group 12 of the phylogenetic tree, iii) Iranian isolates of a.a. haplotypes C33, 34, 40 and 43 were clustered in groups 8, 9 and 10, respectively. Of the Brazilian isolates: i) a.a. haplotypes C48, 50, 66 and ii) C49, 55, 62 and 63 belonged to groups 18 and 19, iii) a.a. haplotypes C60, 64, 65 and 68 belonged to groups 20 and 21, respectively in the phylogenetic tree (Fig. 2, Supplementary Figs. S1).

4. Discussion

Enhanced understanding of the transmission dynamics and population genetics for *Plasmodium vivax* is crucial in predicting the emergence and spread of novel parasite phenotypes with major public health implications, such as new relapsing patterns, drug resistance and increased virulence (Brito and Ferreira, 2011). For example, the proportion of hypnozoites and the timing and frequency of their activation, differ between temperate and tropical strains of this parasite (Garnham et al., 1975), while, the *Pvcsp* repeat type was found to be associated with transmissibility by defined species of anophelines (Zakeri et al., 2006; Gonzalez-Ceron et al., 2001). Hence, suitable molecular markers are required for these population genetic studies. The single-copy CSP gene encoding the highly immunogenic CSP surface antigen expressed at the parasite's sporozoite stage has proven to be a useful marker for defining population genetics and phylogenetic relationships of *Plasmodium* species, especially of *P. vivax* (Patil et al., 2010; Brito and Ferreira, 2011; Tripathi and Gupta, 2011).

A previous study by Manamperi et al. (2008), analyzing *Pvcsp* in 12 parasite isolates collected from blood smears of *P. vivax* patients from an endemic area in Sri Lanka by a nested PCR, indicated a highly diverse *P. vivax* parasite population in the island. However, the present extensive investigation for the first time demonstrated the genetic diversity of *P. vivax* in Sri Lanka by sequencing the repetitive region of the CSP from 60 isolates collected from two malaria endemic and a non-endemic area in Sri Lanka.

Of the three defined variants of PvCSP, the VK210, VK247 and the *P. vivax* like strains, Machado and Póvoa (2000) suggests that the VK210 variant is the best-adapted variant in the world. This was corroborated by the entire Sri Lankan test population representing the VK210 phenotype together with global isolates from Azerbaijan, Brazil, China, Philippines, India, Iran and Korea (Mann et al., 1994; Leclerc et al., 2004; Kim et al., 2006; Zakeri et al., 2006; Patil et al., 2010). Studies have reported differences in the infectivity of anophelines to the variant genotypes indicating that *Anopheles darlingi* and *A. pseudopunctipennis* were more susceptible to the infection by VK210 (Da Silva et al., 2006). In Sri Lanka, *A. culicifacies* and *A. subpictus* are the two most abundant species of malaria infected anophelines (Konradsen et al., 2000).

Significantly, a study by Storti-Melo et al. (2009) recorded a higher frequency of VK210 single infections among subjects comprising of the FYA/FYB allele coding for the Duffy Binding Protein. A possible explanation that both the CSP and DBP were under strong linkage disequilibrium was opted, where *P. vivax* CSP variants could be tagging functional subset(s) of genetic diversity that modulate the efficiency of erythrocyte invasion towards a specific group of erythrocyte receptors.

Repeat arrays in *Plasmodium* antigens (which are exposed to the immune system of the vertebrate host) possess unique characteristics with respect to the number of repeat units, as well as nucleotide and amino acid composition, suggesting that natural selection exerted by the host immune system has shaped features of these arrays (Hughes, 2004). The VK210 variant CSP of *P. vivax*, included ~2–4 local and up to 2–6 global repeats of the two most abundant PRMs, GDRADGQPA and GDRAAGQPA. Similar to *P. falciparum*, these abundant allelic variations with respect to the number of repeat units, suggests that expansion and contraction of these arrays are ongoing (Delhomme and Djian, 2000; Van Rheede et al., 2003; Patil et al., 2010). The mode of evolution of the longer (9-mer) repeats of *P. vivax csp* seems to be quite similar to that hypothesized for the shorter (4-mer) repeats of *P. falciparum csp* (McConkey et al., 1990; Rich et al., 1997) and may involve repeated nonreciprocal, intrahelical recombination events such as strand-slippage during mitotic DNA replication. The number of repetitive peptide repeat motifs is a key factor governing the genetic diversity of *P. vivax csp* of worldwide isolates including those from Sri Lanka (Fig. 2).

Significantly, few nucleotide replacements are observed when RATs, rather than nucleotide sequences, are properly aligned (Rich et al., 1997; Rich et al., 2000). Intragenic recombination, in addition to point mutations, may have generated some of the motifs and RATs observed. The RATs within these CR arrays display a remarkably similar arrangement, arguing for the relatively recent origin from a common ancestor (Patil et al., 2010). Even though a nucleotide diversity of 0.025 was observed for PvCSP in the local population, as a majority of the polymorphisms were synonymous at the amino acid level, the central repetitive domain (CR domain) of PvCSP amino acid haplotypes represented a highly conserved overall structure of the peptide repeat motif arrays. This was common to all geographically distributed isolates. As synonymous substitutions are adaptively neutral, or nearly so, synonymous polymorphisms reflect the time elapsed since the sequences derived from a common ancestral sequence. In tracing the evolutionary history of *P. vivax*, the scarcity of CSP non-synonymous polymorphisms can be interpreted, as it has been for *P. falciparum* (Escalante et al., 1998; Rich et al., 1998; Hartl, 2004), as an indication of an aged species arising from a common ancestor. Using previously published *P. vivax* VK247 CSP isolates from Western Colombia (Hernández-Martínez et al., 2011) and Iran (Zakeri et al., 2006), this study confirmed a similar purifying selection at the CR region, which highlighted the two strains' (VK210 and VK247) evolutionary history. In contrast to our observations, Lim et al. (2005) reported less synonymous segregating sites (only 3 among the 19 a.a haplotypes) in the CR domain from several worldwide *P. vivax* isolates.

Strikingly, in the Sri Lankan isolates, synonymous substitutions of the common repeat sequence occurred frequently in codons 1, 2, and 7, while the mutations at codon 5 were always non-synonymous,

Please cite this article as: Dias, S., et al., Population genetic structure of the *Plasmodium vivax* circumsporozoite protein (*Pvcsp*) in Sri Lanka, Gene (2013), http://dx.doi.org/10.1016/j.gene.2013.01.003

indicating that variation at codon 5 reflected selective pressure. Similar observations were recorded for all the VK210 global isolates.

Extensive variation in repetitive domains of CSP is well defined with linkage disequilibrium and low genetic diversity in areas with low levels of transmission (Anderson et al., 2000). Our study population, together with Brazil and Iran, displayed strong LD arguing for a predominantly clonal mode of reproduction (i.e., high levels of inbreeding and rare homologous meiotic recombination events) in parasites (Patil et al., 2010). Further, in populations with high levels of inbreeding, the "effective" recombination rate will be considerably reduced which has been documented in the local and global population. However, the Korean population displayed a linkage equilibrium across PvCSP, which must have resulted from the low number (N = 12) of samples. Moreover, insertions and deletions in the central repeat domain, from either sexual recombination during meiosis or intrahelical strand-slippage events during mitotic DNA replication (McConkey et al., 1990), generate novel CSP variants that may be positively selected if the mutant parasites evade host immunity. Frequent mitotic recombination coupled with positive selection of new variants may accelerate sequence evolution even where meiotic recombination and out crossing are relatively uncommon in malaria parasite populations (Rich et al., 2000).

Although, 19 different a.a. haplotypes defined from the Sri Lankan population was unique to the island, the sequences were quite similar in the three local test areas except for 2–3 a.a. changes. Moreover, the low F_{ST} values among the three areas (Anuradhapura/Kataragama $F_{ST} = -0.018$, Anuradhapura/Colombo $F_{ST} = 0.12$ and Colombo/Kataragama $F_{ST} = 0.06$) confirmed a gene flow of the parasites among the three populations and hence appear to be a single population. This signifies the inbreeding of the parasites to form new variants, which was further illustrated in the phylogenetic tree, where isolates of the three local study areas were distributed without clustering area wise. Nonetheless, the worldwide isolates showed a strict geographic clustering in both the VK210 and VK 247 variants as reported previously (Leclerc et al., 2004; Zakeri et al., 2006). This was further corroborated by the current study, where the F_{ST} values calculated for VK210 isolates among different geographic populations showed strong population structure, in addition to previous studies (Cornejo and Escalante, 2006; Orjuela-Sánchez et al., 2010). Further, the markedly distinct distribution pattern of Peptide Repeat Motifs 1 and 2 among the Asian and the South American isolates, corroborate a recent study by Chenet et al. (2012). This scenario is reiterated with the geographic structure of the central repeat region of *Pvcsp*, as illustrated by the clustering obtained in the phylogenetic tree.

This study for the first time demonstrated the genetic diversity of the central repeat domain of the *P. vivax* Circumsporozite protein in Sri Lanka, where polymorphism was due to point mutations, insertions and intragenic recombination. This had resulted in 19 amino acid haplotypes and a single PRM unique to the island.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2013.01.003.

Acknowledgments

This work was supported by the Asia Pacific Malaria Elimination Network (grant no.107-14). The patients who generously donated blood for this study, the assistance of the Malaria Research Unit, Department of Parasitology, Faculty of Medicine, University of Colombo, in sample collection (through collaboration on PVU's grant no. F/3008-1 from IFS, Sweden) and the support of the Department of Zoology, Faculty of Science, University of Colombo are acknowledged.

References

Anderson, T.J.C., Haubold, B., Williams, J.T., Estrada-Franco, J.G., Richardson, L., Mollinedo, R., et al., 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol. Biol. Evol. 17, 1467–1482.

Arenas, M., Posada, D., 2010. The effect of recombination on the reconstruction of ancestral sequences. Genetics 184 (4), 1133–1139.

Avise, J.C., 2000. Phylogeaography: the History and formation of species. Harvard University Press, Cambridge, Massachusetts.

Briët, O.J.T., Gunawardena, D.M., van der Hoek, W., Amerasinghe, F.P., 2003. Sri Lanka malaria maps. Malar. J. 2 (22) (http://www.malariajournal.com/content/2/1/22).

Brito, C.F.A., Ferreira, M.U., 2011. Molecular markers and genetic diversity of *Plasmodium vivax*. Mem. Inst. Oswaldo Cruz 106 (1), 12–26.

Chenet, S.M., Tapia, L.L., Escalante, A.A., Durand, S., Lucas, C., Bacon, D.J., 2012. Genetic diversity and population structure of genes encoding vaccine candidate antigens of *Plasmodium vivax*. Malar. J. 14 (11), 68.

Cochrane, A.H., Nardin, E.H., Arruda, M., Maracic, M., Clavijo, P., Collins, W.E., et al., 1990. Widespread reactivity of human sera with variant repeat of the circumsporozoite protein of *Plasmodium vivax*. Am.J.Trop. Med. Hyg. 43, 446–451.

Coppi, A., Natarajan, R., Pradel, G., Bennett, B.L., James, E.R., Roggero, M.A., et al., 2011. The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host. J. Exp. Med. 208 (2), 341–356.

Cornejo, O.E., Escalante, A.A., 2006. The origin and age of *Plasmodium vivax*. Trends Parasitol. 22 (12), 558–563.

Criscione, C.D., Poulin, R., Blouin, M.S., 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Mol. Ecol. 14 (8), 2247–2257.

Cui, L., Mascorro, C.N., Fan, Q., Rzomp, K.A., Khuntirat, B., Zhou, G., et al., 2003. Genetic diversity and multiple infections of *Plasmodium vivax* malaria in Western Thailand. Am.J.Trop. Med. Hyg. 68, 613–619.

Da Silva, A.N.M., Santos, C.C.B., Lacerda, R.N., Machado, R.L.D., Póvoa, M.M., 2006. Susceptibility of *Anopheles aquasalis* and *An. darlingi* to *Plasmodium vivax* VK210 and VK247. Mem. Inst. Oswaldo Cruz 101, 547–550.

Delhomme, B., Djian, P., 2000. Expansion of mouse involucrin by intra-allelic repeat addition. Gene 252, 195–207.

Dias, S., Longacre, S., Escalante, A.A., Udagama-Randeniya, P.V., 2011a. Genetic diversity and recombination at the C-terminal fragment of the merozoite surface protein-1 of *Plasmodium vivax* (PvMSP-1) in Sri Lanka. Infect. Genet. Evol. 11, 145–156.

Dias, S., Somarathna, M., Manamperi, A., Escalante, A.A., Gunasekera, A.M., Udagama, P.V., 2011b. Evaluation of the genetic diversity of domain II of Plasmodium vivax Apical Membrane Antigen 1 (PvAMA-1) and the ensuing strain-specific immune responses in patients from Sri Lanka. Vaccine 29, 7491–7504.

Escalante, A.A., Barrio, E., Ayala, F.J., 1995. Evolutionary origin of human and primate malarias: Evidence from the circumsporozoite protein gene. Mol. Biol. Evol. 12, 616–626.

Escalante, A.A., Lal, A.A., Ayala, F.J., 1998. Genetic polymorphism and natural selection in the Malaria parasite *Plasmodium falciparum*. Genetics 149, 189–202.

Escalante, A.A., Cornejo, O.E., Rojas, A., Udhayakumar, V., Lal, A.A., 2004. Assessing the effect of natural selection in malaria parasites. Trends Parasitol. 20 (8), 388–395.

Fonseka, J., Mendis, K.N., 1987. A metropolitan hospital in a non-endemic area provides a sampling pool for epidemiological studies on vivax malaria in Sri Lanka. Trans. R. Soc. Trop. Med. Hyg. 81, 360–364.

Garnham, P.C., Bray, R.S., Bruce-Chwatt, L.J., Draper, C.C., Killick-Kendrick, R., Sergiev, P.G., et al., 1975. A strain of *Plasmodium vivax* characterized by prolonged incubation: morphological and biological characteristics. Bull. World Health Organ. 52 (1), 21–32.

Gonzalez-Ceron, L., Rodriguez, M.H., Santillan, F., Chavez, B., Nettel, J.A., Hernandez-Avila, J.E., et al., 2001. *Plasmodium vivax*: ookinete destruction and oocyst development arrest are responsible for *Anopheles albimanus* resistance to circumsporozoite phenotype VK247 parasites. Exp. Parasitol. 98 (3), 152–161.

Hartl, D.L., 2004. The origin of malaria: mixed messages from genetic diversity. Nat. Rev. Microbiol. 2, 15–22.

Hernández-Martínez, M.A., Escalante, A.A., Arévalo-Herrera, M., Herrera, S., 2011. Antigenic diversity of the *Plasmodium vivax* circumsporozoite protein in parasite isolates of Western Colombia. Am.J.Trop. Med. Hyg. 84 (2), 51–57.

Hughes, A.L., 2004. The evolution of amino acid repeat arrays in *Plasmodium* and other organisms. J. Mol. Evol. 59, 528–535.

Imwong, M., Pukrittayakamee, S., Grüner, A.C., Rénia, L., Letourneur, F., Looareesuwan, S., et al., 2005. Practical PCR genotyping protocols for *Plasmodium vivax* using *Pvcsp* and *Pvmsp1*. Malar. J. 4, 20.

Kho, W.G., Jang, J.Y., Hong, S.T., Lee, H.W., Lee, W.J., Lee, J.S., 1999. Border malaria characters of re-emerging vivax malaria in the Republic of Korea. Korean J. Parasitol. 37, 71–76.

Kim, T., Kim, Y.J., Song, K., Cha, J.S.S., Kim, Y.K., Suh, Y.K.S.I.B., et al., 2002. The molecular characteristics of circumsporozoite protein gene subtypes from *Plasmodium vivax* isolates in Republic of Korea. Parasitol. Res. 88, 1051–1054.

Kim, J.R., Imwong, M., Nandy, A., Chotivanich, K., Nontprasert, A., Tonomsing, N., et al., 2006. Genetic diversity of *Plasmodium vivax* in Kolkata, India. Malar. J. 5, 71.

Konradsen, F., Amerasinghe, F.P., Van der Hoek, W., Amerasinghe, P.H., 2000. Malaria in Sri Lanka: Current Knowledge on Transmission and Control. International Water Management Institute, Colombo, Sri Lanka.

Leclerc, M.C., Durant, P., Gauthier, C., Patot, S., Billote, N., Menegon, M., et al., 2004. Meager genetic variability of the human malaria agent *Plasmodium vivax*. Proc. Natl. Acad. Sci. USA 101, 14455–14460.

Librado, P., Rozas, J., 2009. DnaSP v5: software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

Lim, C.S., Kim, Y.K., Lee, K.N., Kim, S.H., Hoffman, K.J., Song, K.J., et al., 2001. The analysis of circumsporozoite protein gene sequences from *Plasmodium vivax* isolates in the Republic of Korea. Ann. Trop. Med. Parasitol. 95, 229–235.

Lim, C.S., Tazi, L., Ayala, F.J., 2005. *Plasmodium vivax*: recent world expansion and genetic identity to *Plasmodium simium*. Proc. Natl. Acad. Sci. USA 102, 15523–15528.

Machado, R.L.D., Póvoa, M.M., 2000. Distribution of *Plasmodium vivax* variants (VK210, VK247 and *P. vivax*-like) in three endemic areas of Amazonian Brazil and their correlation with chloroquine treatment. Trans. R. Soc. Trop. Med. Hyg. 94, 377–381.

Manamperi, A., Mahawithanage, S., Fernando, D., Wickremasinghe, R., Bandara, A., Hapuarachchi, C., et al., 2008. Genotyping of *Plasmodium vivax* infections in Sri Lanka using *Pvmsp-3α* and *Pvcs* genes as markers: A preliminary report. Trop. Biomed. 25 (2), 100–106.

Mann, V.H., Huang, T., Cheng, Q., Saul, A., 1994. Sequence variation in the circumsporozoite protein gene of *Plasmodium vivax* appears to be regionally biased. Mol. Biochem. Parasitol. 68, 45–52.

McConkey, G.A., Waters, A.P., McCutchan, T.F., 1990. The generation of genetic diversity in malaria parasites. Annu. Rev. Microbiol. 44, 479–498.

Mendis, C., Del Giudice, G., Gamage-Mendis, A.C., Tougne, C., Pessi, A., Weerasinghe, S., et al., 1992. Anti-circumsporozoite protein antibodies measure age related exposure to malaria in Kataragama, Sri Lanka. Parasite Immunol. 14 (1), 75–86.

Orjuela-Sánchez, P., Karunaweera, N.D., da Silva-Nunes, M., da Silva, N.S., Scopel, K.K.G., Gonçalves, R.M., et al., 2010. Single-nucleotide polymorphism, linkage disequilibrium and geographic structure in the malaria parasite *Plasmodium vivax*: prospects for genome-wide association studies. BMS. Genet. 11, 65.

Patil, A., Orjuela-Sánchez, P., da Silva-Nunes, M., Ferreira, M.U., 2010. Evolutionary dynamics of the immunodominant repeats of the *Plasmodium vivax* malaria vaccine candidate circumsporozoite protein (CSP). Infect. Genet. Evol. 10, 298–303.

Premaratne, P.H., Aravinda, B.R., Escalante, A.A., Udagama, P.V., 2011. Genetic diversity of *Plasmodium vivax* Duffy Binding Protein II (PvDBPII) under unstable transmission and low intensity malaria in Sri Lanka. Infect. Genet. Evol. 11 (6), 1327–1339.

Qari, S.H., Shi, Y.P., Goldman, I.F., Udhayakumar, V., Alpers, M.P., Collins, W.E., et al., 1993a. Identification of *Plasmodium vivax*-like human malaria parasite. Lancet 341, 780–783.

Qari, S.H., Shi, Y.P., Povoa, M.M., Alpers, M.P., Deloron, P., Murphy, G.S., et al., 1993b. Global occurrence of *Plasmodium vivax*-like human malaria parasite. J. Infect. Dis. 168, 1485–1489.

Rich, S.M., Hudson, R.R., Ayala, F.J., 1997. *Plasmodium falciparum* antigenic diversity: evidence of clonal population structure. Proc. Natl. Acad. Sci. U. S. A. 94, 13040–13045.

Rich, S.M., Licht, M.C., Hudson, R.R., Ayala, F.J., 1998. Malaria's Eve: evidence of a recent population bottleneck throughout the world populations of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. U. S. A. 95, 4425–4430.

Rich, S.M., Ferreira, M.U., Ayala, F.J., 2000. The origin of antigenic diversity in *Plasmodium falciparum*. Parasitol. Today 16, 390–396.

Rosenberg, R., Wirtz, R.A., Lanar, D.E., Sattabongkot, J., Hall, T., Waters, A.P., et al., 1989. Circumsporozoite protein heterogeneity in the human malaria parasite *Plasmodium vivax*. Science 245, 973–976.

Sinnis, P., Nardin, E., 2002. Sporozoite antigens: biology and immunology of the circumsporozoite protein and thrombospondin related anonymous protein. In: Perlmann, P., Troye-Blomberg, M. (Eds.), Malaria Immunology. Karger, S., Basel, A.G, Switzerland, pp. 70–96.

Souza-Neiras, W.C., Storti-Melo, L.M., Cassiano, G.C., Couto, V.S.C.A., Couto, A.A.R.A., 2010. *Plasmodium vivax* circumsporozoite genotypes: a limited variation or new subspecies with major biological consequences? Malar. J. 9, 178.

Storti-Melo, L.M., de Souza-Neiras, W.C., Cassiano, G.C., Joazeiro, A.C.P., Fontes, C.J., Bonini-Domingos, C.R., et al., 2009. *Plasmodium vivax* circumsporozoite variants and Duffy blood group genotypes in the Brazilian Amazon region. Trans. R. Soc. Trop. Med. Hyg. 103, 672–678.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. http://dx.doi.org/10.1093/molbev/msr121.

Thakur, A., Alam, M.T., Bora, H., Kaur, P., Sharma, Y.D., 2008. *Plasmodium vivax*: sequence polymorphism and effect of natural selection at apical membrane antigen-1 (PvAMA-1) among Indian population. Gene 419, 35–42.

Tripathi, V., Gupta, D., 2011. Evolutionary analysis of circumsporozoite surface protein and merozoite surface protein-1 (CSP and MSP-1) sequences of malaria parasites. Bioinformation 6 (8), 320–323.

Van Rheede, T., Smolenaars, M.M.W., Madsden, O., de Jong, W.W., 2003. Molecular evolution of the mammalian prion protein. Mol. Biol. Evol. 20, 111–121.

Wickramarachchi, T., Premaratne, P.H., Dias, S., Handunnetti, S.M., Udagama-Randeniya, P.V., 2010. Genetic complexity of Plasmodium vivax infections in Sri Lanka: as reflected at the merozoite surface protein-3 locus. Ann. Trop. Med. Parasit. 104 (2), 95–108.

Zakeri, S., Mehrizi, A.A., Djadid, N.D., Snounou, G., 2006. Circumsporozoite protein gene diversity among temperate and tropical *Plasmodium vivax* isolates from Iran. Trop. Med. Int. Health 11, 729–737.

Zakeri, S., Raeisi, A., Afsharpad, M., Kakar, Q., Ghasemi, F., Atta, H., et al., 2010. Molecular characterization of *Plasmodium vivax* clinical isolates in Pakistan and Iran using PvMSP-1, PvMSP-3 alpha and PvCSP genes as molecular markers. Parasitol. Int. 591, 15–21.

Zhong, D., Bonizzoni, M., Zhou, G., Wang, G., Chen, B., Vardo-Zalik, A., et al., 2011. Genetic diversity of *Plasmodium vivax* malaria in China and Myanmar. Infect. Genet. Evol. 11 (6), 1419–1425.

Zilversmit, M., Hartl, D.L., 2005. Evolutionary history and population genetics of human malaria parasites. In: Sherman, I.W. (Ed.), Molecular approaches to malaria. Press, Washington DC, p. 542.

Please cite this article as: Dias, S., et al., Population genetic structure of the *Plasmodium vivax* circumsporozoite protein (*Pvcsp*) in Sri Lanka, Gene (2013), http://dx.doi.org/10.1016/j.gene.2013.01.003