Published in final edited form as:

Infect Genet Evol. 2013 October ; 19: 164–175. doi:10.1016/j.meegid.2013.07.009.

Microsatellite analysis of chloroquine resistance associated alleles and neutral loci reveal genetic structure of Indian *Plasmodium falciparum*

Prashant K. Mallick¹, Patrick L. Sutton², Ruchi Singh³, Om P. Singh¹, Aditya P. Dash⁴, Ashok K. Singh⁵, Jane M. Carlton², and Virendra K. Bhasin^{5,*}

¹National Institute of Malaria Research, Indian Council of Medical Research, Dwarka Sector-8, Delhi-110077, India

²Center for Genomics and Systems Biology, Department of Biology, New York University, New York, NY 10003, USA

³National Institute of Pathology, New Delhi-110029, India

⁴W.H.O SEARO, India

⁵Department of Zoology, University of Delhi-110007, India

Abstract

Efforts to control malignant malaria caused by *Plasmodium falciparum* are hampered by the parasite's acquisition of resistance to antimalarial drugs, *e.g.*, chloroquine. This necessitates evaluating the spread of chloroquine resistance in any malaria-endemic area. India displays highly variable malaria epidemiology and also shares porous international borders with malaria-endemic Southeast Asian countries having multi-drug resistant malaria. Malaria epidemiology in India is believed to be affected by two major factors: high genetic diversity and evolving drug resistance in *P. falciparum*. How transmission intensity of malaria can influence the genetic structure of chloroquine-resistant *P. falciparum* population in India is unknown. Here, genetic diversity within and among *P. falciparum* populations is analyzed with respect to their prevalence and chloroquine resistance observed in 13 different locations in India. Microsatellites developed for *P. falciparum*, including three putatively neutral and seven microsatellites thought to be under a hitchhiking effect due to chloroquine selection were used. Genetic hitchhiking is observed in five of seven microsatellites flanking the gene responsible for chloroquine resistance. Genetic admixture analysis and F-statistics detected genetically distinct groups in accordance with transmission intensity of different locations and the probable use of chloroquine. A large genetic break between the chloroquine-resistant parasite of the Northeast-East-Island group and Southwest group ($F_{ST} = 0.253$, $P < 0.001$) suggests a long period of isolation or a possibility of different origin between them. A pattern of significant isolation by distance was observed in low transmission areas ($r = 0.49$, $P = 0.003$, $N = 83$, Mantel test). An unanticipated pattern of spread of hitchhiking suggests genetic structure for Indian *P. falciparum* population. Overall, the study suggests that transmission intensity can be an efficient driver for genetic differentiation at both neutral and adaptive loci across India.

© 2013 Elsevier B.V. All rights reserved.

*Corresponding author: virendrabhasin@gmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Plasmodium falciparum; India; Chloroquine resistance; Hitchhiking; Transmission area; Gene-flow

1. Introduction

Positive natural selection contributes to the fixation of beneficial alleles in a population (Kim and Stephan, 2002). During this evolutionary process, it is not uncommon to observe a parallel increase in the frequency of specific neutral alleles, which are linked to these beneficial alleles by virtue of proximity. The carriage of neutral alleles during positive natural selection is called genetic hitchhiking (Smith and Haigh, 1974) and it typically results in a reduction of neutral variation in the population through a phenomenon called selective sweep (Anderson, 2004). The pattern of genetic hitchhiking largely depends on two factors: (i) the frequency of random mating within a population, and (ii) the presence of geographic structure (genetic isolation by distance) in the population. The geographic structure of a population may delay the propagation of beneficial mutations by limited migration of parasites across geographic space, or the spread of beneficial mutations over a certain geographic distance may result in weaker hitchhiking (Barton, 2000). Thus, understanding the patterns of selective sweep in a population may reveal hidden geographic structure within the population (Kim and Maruki, 2011).

In *Plasmodium falciparum*, several studies have investigated genetic hitchhiking around mutated genes implicated in resistance to antimalarials like pyrimethamine (Nair et al., 2003; Roper et al., 2004), sulphadoxine (Alam et al., 2011; Vinayak et al., 2010) and chloroquine (Chen et al., 2005; Mehlotra et al., 2005; Wootton et al., 2002). Other studies have established that geographic structure might affect the extent of a selective sweep associated with drug resistance (McCollum et al., 2007; Mu et al., 2005; Pearce et al., 2005). Studies on chloroquine resistance (CQR) found that: (i) multiple foci exist for the origin of CQR, one in Asia spreading to Africa, two in South America, one in Papua New Guinea (PNG) and one in the Philippines; (ii) Southeast Asia and Africa share the resistant CVIET haplotype (amino acids at codon 72 to 76) in the *Plasmodium falciparum chloroquine resistance transporter* (*pfcr*) gene, suggesting the role of human migration in conjunction with the spread of chloroquine-resistant parasites (Wootton et al., 2002); and (iii) resistant SVMNT haplotype characterizes the South American type, which is prevalent throughout South America and the Pacific region (Mehlotra et al., 2005). Studies evaluating genetic variation within microsatellite loci flanking the *pfcr* gene reported invasion of the PNG type SVMNT throughout India (Lumb et al., 2012; Mixson-Hayden et al., 2010). Invasion of the same resistant SVMNT into Pakistan has also been reported (Rawasia et al., 2012).

Analyzing the genetic variation within neutral microsatellite loci provides an effective means to determine parasite population structure and can unravel potential links between the epidemiological pattern and genetic structure of *P. falciparum* (Abdel-Muhsin et al., 2004; Anderson et al., 2000). Malaria exhibits extreme epidemiological diversity throughout India in terms of prevalence and distribution of *Plasmodium* species, transmission intensity, vector distribution, and ecology of vector breeding sites. For example, the presence of local parasite population structure has been suggested to be a factor of exposure-related immunity, identifiable through seroprevalence frequencies to various *P. falciparum* surface antigens. However, the evidence for this remains somewhat controversial, as populations located in close proximity were reported to be more related (Biswas et al., 2008), in contrast identical genetic structures were also found between two distant localities (*i.e.*, eastern and northeastern India) (Joshi et al., 2007). Previous work has shown an association between

genetic diversity of the *pfcr*t-haplotypes and variable transmission intensities throughout India (Mallick et al., 2012). However, there is limited information available on how the genetic diversity of *P. falciparum* is structured within the Indian population, and it has yet to be established whether variation in malaria epidemiological factors across different geographical locations provides a basis for *P. falciparum* evolutionary plasticity and local adaptation in the country. In the absence of a good understanding of *P. falciparum* population structure in India, the gene-flow of the resistant SVMNT across India has not yet been explored, nor has the spread of neutral alleles around the *pfcr*t been investigated.

This study aims to assess the genetic diversity and population structure of *P. falciparum* in several Indian populations which over time have been subject to different levels of chloroquine pressure. Studying the correlation between transmission intensity and genetic structure, and analyzing the spread of hitchhiking may facilitate a deeper understanding of how drug utilization shapes the genetic structure of *P. falciparum* in India. This study uses genome-wide putative neutral microsatellite loci and *pfcr*t-flanking microsatellite loci assumed to be under hitchhiking effect due to chloroquine selection, in combination with robust population statistics, to illustrate the significant geographical structure of *P. falciparum* populations in India and also to investigate the spread of CQR associated with genetic hitchhiking.

2. Materials and methods

2.1. Ethical Considerations

The Ethics Committee of National Institute of Malaria Research (NIMR) approved this study. All individuals or the parents/guardians of children gave written informed consent before inclusion in the study.

2.2. Selection of *P. falciparum* isolates

All isolates were collected from microscopically positive *P. falciparum*-infected patients between 2002 and 2006. Genomic DNA samples (N = 287) isolated from patients infected with *P. falciparum* were selected from an earlier study in which they had been genotyped for *pfcr*t-haplotypes (Mallick et al., 2012). Additionally, genomic DNA was isolated from blood samples collected in Car Nicobar (CN) (N = 12) and Madhya Pradesh (MP) (N = 17) during years 2002 and 2005, respectively. DNA was extracted from these 29 samples using Qiagen® Blood Mini Kit (Hilden, Germany) and *pfcr*t-haplotypes were determined by Sanger sequencing. The distribution of *pfcr*t-haplotypes at different geographic locations is illustrated in Fig. 1. All samples were screened for single-clone infections by genotyping the single-copy antigenic genes merozoite surface protein-1 (*m*sp-1) and *m*sp-2 as described (Snounou et al., 1999) and further underwent microsatellite genotyping at three genome-wide loci at chromosome 5 (TAA60), chromosome 13 (TAA80) and chromosome 11 (ARA2) as described (Anderson et al., 1999). Samples having more than one allele at any of the three microsatellite loci were excluded from further investigation. Finally, a total of 188 single-clone *P. falciparum* isolates were selected for further evaluation : Assam (ASM) (N = 13), Orissa (ORS) (N = 25), West Bengal (WBG) (N = 13), Jharkhand (JHK) (N = 17), Chhattisgarh (CHG) (N = 12), Car Nicobar (CN) (N = 7), Maharashtra (MAH) (N = 9), Tamil Nadu (TN) (N = 18), Goa (GOA) (N = 15), Gujarat (GUJ) (N = 15), Rajasthan (RAJ) (N = 21), Uttar Pradesh (UP) (N = 14), Madhya Pradesh (MP) (N = 9). Based on the prevalence of falciparum malaria and the transmission intensity described in Mallick *et al.*, 2012, these study sites were classified as high transmission areas (HTA: ASM, ORS, JHK, WBG, CHG, MAH and CN) and low transmission areas (LTA: GOA, GUJ, RAJ, TN, MP and UP). Genomic DNA of eight *P. falciparum* isolates obtained from MR4 were included to represent varied geographic origins; Southeast Asian: INDO (MRA819G), V1/S

(MRA820G), Dd2 (MRA150G); South American: HB3 (MRA155G), Brazil 7G8 (MRA152G); African: T2/C6 (MRA818G), D6 (MRA398); and wild 3D7 (MRA102G).

2.3. Microsatellite typing

All single-clone infections were genotyped for seven microsatellite loci flanking the *pfprt*, B5M97 (−24kb), B5M77 (−20kb), 2E10 (−5kb), 9B12 (+1kb), PE12a (+6kb), 2H4 (+22kb), and PE14F (+106kb), as previously described (Wootton et al., 2002). All primary and nested PCR's were performed in a volume of 20 µl with 0.75 U of Taq DNA polymerase, 2.5 mM magnesium chloride, 0.25 mM each dNTP and 0.1 mM of each primers. The primary PCR product was diluted 1:100 for the following nested PCR. For all nested PCR reactions the 5' end of the primers were modified with one of the following fluorophores: FAM, HEX or TET. Following PCR amplification the length variation was measured commercially at The Centre for Genomic Application, New Delhi, India. Genotypes were analyzed using ABI Peak Scanner v1.0 software. To check on scoring error two reference isolates (Dd2 and 7G8) were amplified and used to calibrate allele sizes.

2.4. Microsatellite analysis

2.4.1. Genetic Diversity—The genetic variation in terms of expected heterozygosity (H_e) for each microsatellite locus was calculated using the formula $H_e = [n/(n-1)] [1 - \sum p_i^2]$, where n is number of *P. falciparum* isolates genotyped for a particular locus and p_i is the frequency of the i th allele. The sampling variance of H_e was calculated as $[2(n-1)/n^3] (2(n-2) [\sum p_i^3 - (\sum p_i^2)^2])$ (Nash et al., 2005). Significant differences between the allelic diversity and H_e of any two groups was assessed using a Mann-Whitney test in GraphPad Prism 5 (GraphPad Software, Inc). Allele diversity was measured in terms of the mean number of distinct alleles (allelic richness) and the mean number of private alleles (private allelic richness) using the ADZE program (Szpiech et al., 2008). The program uses a rarefaction method (Kalinowski, 2004) to correct the sample size factor across whole population.

2.4.2. Linkage Disequilibrium—LIAN version 3.5 (Haubold and Hudson, 2000) was used to infer linkage disequilibrium (LD) in neutral and flanking microsatellite loci of various *pfprt*-haplotypes. The standardized index of association (I_A^S) was calculated for the entire multi-locus haplotype and also within each unique multi-locus haplotype obtained from Arlequin software v3.0 to account any bias resulting from the presence of dominant haplotypes. The null hypothesis of complete linkage equilibrium ($I_A^S = 0$) was tested by Monte-Carlo simulations using 10,000 random permutations of the data. Due to the small sample size of some geographic regions, the LD was calculated by considering *pfprt*-haplotypes, rather than by location.

2.4.3. Population structure—STRUCTURE 2.3.1 software (Pritchard et al., 2000) was used to assign isolates to subpopulations (K) according to allele frequencies at each locus. Here, the software was used to test how microsatellite haplotypes clustered according to the geographic origins of isolates. For the admixture model, burn-in period was set at 50,000 steps, followed by 100,000 Markov chain iterations. Five independent runs were performed for each predefined K -value ($K = 1-8$). These analyses were done for two different combinations of microsatellite loci; (1) for the neutral loci TAA60, TAA80, ARA2; and (2) for the *pfprt*-flanking loci B5M97, B5M77, 2E10, 9B12, PE12a, 2H4, PE14F. STRUCTURE software assigns probabilities of relatedness for each individual haplotype (or sample) into genetically distinct clusters; these probabilities are then used to infer the membership to specific clusters (or subpopulations) (Rosenberg et al., 2003).

2.4.4. Population Differentiation—Population differentiation among all regions was measured by pairwise F_{ST} values (Weir and Cockerham, 1984). The F_{ST} values and its

significance was calculated using 10,000 permutation tests in Arlequin (Excoffier et al., 2005). N_m values calculated as $N_m = 1/2[(1/F_{ST}) - 1]$, were estimated to account for the degree of migration between the subpopulations. An analysis of molecular variance (AMOVA) (Excoffier et al., 1992) was used to determine statistical differences between the clusters obtained by STRUCTURE and also in different geographic region (Northern, Northeastern, Eastern, Western, Central, Southern and Island). AMOVA analysis was used to determine the level at which level genetic variation is distributed (*i.e.*, within population or between populations). The significance of covariance components associated with the different possible level of genetic structure was tested using nonparametric methods (10,000 permutations). A Mantel test (Mantel, 1967) was performed to assess the correlation between geographical and genetic distances (isolation by distance-pattern) of all populations, using Arlequin (Excoffier et al., 2005). Significance between genetic and geographical distances was determined using a Mantel test with 10,000 permutations.

2.4.5. Genetic relationship—A median joining network was constructed using unique multi-locus haplotypes at *pfcr*-flanking microsatellite loci in NETWORK version 4.5.1.0 (http://www.fluxus-engineering.com/sharenet_rm.htm), to illustrate the genetic relationship between the various *pfcr*-haplotypes found among different regions of India and worldwide isolates.

3. Results

The length polymorphism in two single-copy genes (*msh-1* and *msh-2*) identified 213 single-clone infections out of 316 isolates. A further round of genotyping at three genome-wide microsatellite loci (TAA60, TAA80, and ARA2) identified 25 additional multi-clone infections, leaving a final 188 single-clone infections. Of these, 169 isolates successfully amplified for all seven microsatellite loci flanking the *pfcr*.

3.1. Allele diversity in microsatellite markers

Genetic diversity was computed using two sets of loci: (i) three putatively neutral loci and (ii) seven loci flanking *pfcr* that are thought to be under chloroquine selective pressure. All microsatellite loci were polymorphic and the number of alleles per locus ranged from five (at PE12a) to 20 (at 2H4). The level of allelic diversity was extensive between geographic locations (Supplementary Table 1). The mean number of alleles per locus (N_A) at neutral loci indicated higher diversity in high transmission areas (HTA) ranging from 3.7 to 6.0, in comparison with low transmission areas (LTA), which ranged from 2.7 to 4.7 alleles per locus. A similar trend was observed in the *pfcr*-flanking microsatellite loci, where HTA ranged from 2.4 to 4.6, in comparison with LTA, which ranged from 1.4 to 2.7 alleles per locus. There was a significant reduction ($P=0.0107$, Mann-Whitney test) in N_A at resistant *pfcr*-flanking loci when compared to the neutral loci.

Since observed number of alleles per locus depends upon sample size, the sample size was standardized to the smallest sample size across populations (CN; $N = 7$) using ADZE. The observed allelic richness " A_R " (mean number of distinct alleles) and private allelic richness " P_R " (number of alleles private to the population) are presented in Table 1. Even after standardizing the data, the results were consistent, indicating that HTA have higher A_R than LTA in both neutral and *pfcr*-flanking loci. The only exception were isolates from Madhya Pradesh (from LTA), which showed intermediate values for all measures of diversity. Higher allelic diversity was observed within the neutral loci, when compared to the *pfcr*-flanking loci at all the geographic locations. Furthermore, significant reduction ($P<0.0001$, Mann-Whitney test) was observed in allelic diversity amongst the resistant *pfcr*-haplotypes in comparison to wild CVMNK.

3.2. Expected Heterozygosity and trend of selective sweep around *pfcr*-haplotypes

Genetic diversity in terms of expected heterozygosity (H_e) was calculated for each locus, while patterns of a selective sweep were tested by only considering the *pfcr*-flanking loci. The mean number of distinct alleles (A_R) and mean H_e for the wild-type were significantly higher ($A_R = 6.0 \pm 0.54$; $H_e = 0.81 \pm 0.07$), compared to resistant CVIET ($A_R = 2.92 \pm 0.33$; $H_e = 0.53 \pm 0.06$; $P < 0.01$, Mann-Whitney test) and resistant SVMNT ($A_R = 2.10 \pm 0.21$; $H_e = 0.32 \pm 0.04$; $P < 0.004$, Mann-Whitney test). This observation is consistent with the fact that these resistant haplotypes are under drug pressure as indicated by the decline in values of A_R and H_e . A reduction in allelic variation was observed among the microsatellite loci between -24kb upstream to +22kb downstream of the *pfcr*. The “valley” of reduced variation for SVMNT was much deeper and broader than CVIET (Fig. 2). When reduction of heterozygosity was assessed at different locations, it provided a selective sweep around -20kb to +6kb in all sites under high transmission conditions, whereas sites exposed to low transmission conditions showed a wider selective sweep around -24kb to +22kb. The exception in LTA was Madhya Pradesh, where a considerably higher amount H_e was observed. The immediate proximal loci to resistant *pfcr*-haplotypes showed modest variation in HTA. The expected heterozygosity of SVMNT (mean $H_e = 0.24 \pm 0.04$, $N = 82$) from LTA was significantly lower in comparison of SVMNT (mean $H_e = 0.45 \pm 0.06$, $N = 40$) from HTA. The regional distribution of expected heterozygosity (H_e) is presented in Table 2.

To investigate the possibility of clonal expansion in prevalence of *pfcr*-haplotypes, A_R and H_e were measured within three neutral loci on chromosomes 5, 11, and 13. In all geographic regions, the mean H_e was significantly higher for the three neutral loci in comparison to the seven *pfcr*-flanking loci ($P < 0.009$, Mann Whitney). The mean H_e at neutral loci was comparable between wild and resistant *pfcr*-haplotype (Table 2). The reduction in genetic variation at both *pfcr*-flanking and neutral loci observed in GOA, TN and UP could be an indication of clonal expansion.

3.2.1. Unique microsatellite haplotypes among *pfcr*-haplotypes—Analysis of unique flanking microsatellite haplotypes associated with various *pfcr*-haplotypes can provide inferences about the dispersal of CQR across India. Five *pfcr*-flanking microsatellite loci (-24kb, -20kb, -5kb, +1kb and +6kb) were used to construct the unique 5-locus haplotypes. The two farthest microsatellite loci downstream 2H4 (+22kb) and PE14F (+106kb) were excluded due to high H_e , which indicate non-selective properties. A total of 54 unique 5-locus haplotypes were observed amongst all the *pfcr*-haplotypes (Supplementary Table 2). Among these 5-locus haplotypes, 16 were associated with wild CVMNK ($N = 16$), 18 with resistant CVIET ($N = 31$), and 20 with resistant SVMNT ($N = 122$). Two 5-locus haplotypes were not restricted to a specific *pfcr*-haplotype: (i) CK3 was shared between wild and resistant CVIET and (ii) haplotype CK17 was shared between wild and resistant SVMNT. None of the 5-locus haplotypes were shared between two resistant *pfcr*-haplotypes. Six different 5-locus haplotypes were observed in eight worldwide isolates (Supplementary Table 2).

3.3. Linkage disequilibrium

Favorable mutations leading to a genetic sweep generally cause a reduction in genetic variation and increase linkage disequilibrium (LD) around the target of selection (Anderson, 2004). To have an assessment of selection, LD was measured as a standardized association index I_A^S between multiple loci flanking *pfcr*-haplotypes. There was no significant LD ($I_A^S = -0.0219$, $P = 0.687$, Monte Carlo simulation) around wild CVMNK, whereas significant LD was observed around resistant haplotypes. The I_A^S values for CVIET and SVMNT were 0.13 ($P < 0.0001$, Monte Carlo simulation) and 0.14 ($P < 0.0001$, Monte Carlo simulation)

respectively. No significant LD was observed at neutral loci in wild CVMNK ($I_A^S = -0.0311$, $P=0.670$, Monte Carlo simulation) and CVIET ($I_A^S = 0.005$, $P=0.378$, Monte Carlo simulation). However, I_A^S value for SVMNT was 0.09 ($P<0.0001$, Monte Carlo simulation) in neutral loci indicates clonal expansion at LTA. Multi-locus LD with seven and five *pfcr*-flanking microsatellite loci and in neutral loci is illustrated in Supplementary Table 3.

3.4. Population structure

The algorithm underlying STRUCTURE detects the number of subpopulations (K) represented in the entire population. A K=2 had the highest clusteredness value ($k=0.58$) over five simulations for the three neutral loci, while a K=3 had highest clusteredness value ($k=0.85$) over five simulations for the seven *pfcr*-flanking loci.

The clustering pattern of the neutral loci obtained with K=2 (Supplementary Fig. 1a) indicates discrete clustering of *P. falciparum* genotypes into high transmission and low transmission areas with limited admixture in both regions. The clusters were significantly associated with transmission intensity (high versus low) ($P<0.0001$, Fisher exact), i.e., Neutral cluster1 was associated with HTA and Neutral cluster2 associated with LTA (Table 3). The population structure of *P. falciparum* according to ancestry at K=2 is shown in Supplementary Fig. 1b. Neutral cluster1 (red) consists of 78 isolates, 70 (89.74%) of which were collected in HTA, while Neutral cluster2 (green) cluster consists of 91 isolates, 75 (82.41%) of which were collected in LTA.

The clustering pattern of the seven *pfcr*-flanking microsatellite loci shows a much higher level of admixture in HTA; while in LTA, Flanking cluster1 (red) of K=3 is the predominant cluster, with a complete absence of Flanking cluster2 (green) of K=3 (Supplementary Fig. 2a). Flanking cluster1 was significantly associated with LTA ($P<0.001$, Fisher exact) and SVMNT ($P<0.001$, Fisher exact). Flanking cluster2 was associated with CVIET ($P<0.001$, Fisher exact) and observed only in Eastern, Northeastern, and CN Island, which was not unexpected due to the proximity to South East Asia (SEA). However, Flanking cluster3 (blue) of K=3 was significantly associated with the wild-type, as well as with the HTA ($P<0.001$, Fisher exact) (Supplementary Fig. 2b). Both the SVMNT and CVIET haplotype appeared in Flanking cluster3, and seem to have evolved from the wild-type through some historical event. By superimposing the two analyses, it is clear that Flanking cluster2 and 3 were significantly associated with Neutral cluster1 ($P<0.001$, Fisher exact), while Flanking cluster1 was significantly associated with Neutral cluster2 ($P<0.001$, Fisher exact).

3.5. Population differentiation

Using the three neutral loci, pairwise F_{ST} values were calculated among all subpopulations. No population differentiation occurred between ORS-CN ($F_{ST} = 0.01$, $P=0.363$) or between ASM-MAH ($F_{ST} = 0.05$, $P=0.069$). Population differentiation occurred between ASM-CHG ($F_{ST} = 0.07$, $P<0.04$), and up to $F_{ST} = 0.44$ between CHG-TN ($P<0.0001$) (Supplementary Table 4). It seems ASM has a limited genetic similarity to CHG, while CHG and TN were most genetically separated. When pairwise F_{ST} values were calculated with seven *pfcr*-flanking loci, it ranged from 0.06 between RAJ-TN ($P<0.03$), and upto 0.46 between CN-GUJ/TN ($P<0.0001$) (Supplementary Table 4), indicating RAJ population has limited genetic similarity to TN and while CN-GUJ/TN has strongest genetic differentiation between them. While, no population differentiation occurred between GOA-TN ($F_{ST} = 0.02$, $P=0.241$) or between ORS-CHG ($F_{ST} = 0.06$, $P=0.067$).

N_m values (Slatkin, 1987) estimate gene-flow at neutral loci between different geographic regions to indicate the level of migration between parasite populations. The N_m values were fairly erratic for non-significant F_{ST} values (i.e when $P > 0.05$), albeit the overall values

indicate high gene-flow between populations. The Nm values for significant F_{ST} values at neutral loci ranged from $Nm=0.64$ between CHG-TN, to $Nm=6.45$ between ASM-CHG, indicating low migration between CHG-TN and high migration between ASM-CHG. However, Nm estimates for *pfcr*-flanking loci indicated the spread of CQR between the subpopulation. Migration of *pfcr*-flanking loci was shown to be highest between RAJ-TN ($Nm=7.54$) and the lowest was between CN-GUJ/TN ($Nm=0.59$). The Nm values at neutral and *pfcr*-flanking loci are illustrated in Supplementary Table 5.

3.5.1. Interpreting F_{ST} and Nm values—The significant pairwise F_{ST} and Nm values were interpreted to generate a pattern of gene-flow about CQR across India. Pairwise F_{ST} values were categorically represented into a range of genetic differentiation; 0.0 to 0.05 indicates minor genetic differentiation, 0.05 to 0.15 indicates moderate genetic differentiation, 0.15 to 0.25 indicates high genetic differentiation, and values above 0.25 indicate strong genetic differentiation among the subpopulations (Hartl and Clark, 1997; Wright, 1978). The estimated level of migration (Nm) was interpreted with the criteria that $Nm > 1.0$ indicates enough gene-flow between subpopulations and $Nm > 4.0$ is an indicative of panmixia (high gene-flow) (Hartl, 1988). Using these categories, the pairwise F_{ST} and Nm values identified four distinct groups of populations in data: first group, Northeast-East-Island (ASM, WBG, ORS, JHK, CHG and CN); second, Northwest (U.P, RAJ and GUJ); third Southwest (GOA and TN); and fourth Central (MAH, M.P).

No population differentiation occurred between subpopulations comprising the Northeast-East-Island group ($F_{ST} = 0.05$, $P > 0.05$), except WBG, which showed moderate levels of differentiation. Additionally, this group showed low to high levels of differentiation with other groups (Supplementary Table 6a). The Southwest group showed high genetic differentiation with the Northwest and Central groups. A strong genetic differentiation was observed between the Northeast-East-Island group and the Southwest group. All four groups showed significant genetic differentiation ($F_{ST} = 0.04$, $P < 0.005$) and high gene-flow between them, i.e., the pairwise F_{ST} ranges from 0.041 to 0.25 and Nm ranges from 1.47 to 11.54 (Supplementary Table 6a, b). While analyzing the genetic differentiation and gene-flow at *pfcr*-flanking loci, the central group appears to be genetically related with all the groups, implying an interaction with all groups of populations, which may be explained by its geographical location. Such evidence of strong genetic differentiation between different groups of subpopulations and high gene-flow within groups permits the development of a hypothetical route of gene-flow about CQR among Indian *P. falciparum* (Fig. 3).

AMOVA based on neutral loci showed that variation among the subpopulations constituted 17.51% of the total variation when no grouping of subpopulations was inferred, and 7.70% variation among populations when subpopulations were grouped by geographical region. Analysis of the two clusters detected with STRUCTURE resulted in 17.26% variation between clusters. Furthermore, the new groups defined by Nm and F_{ST} values show 13.17% variation among the four groups and 7.14% of variation among subpopulations within groups. Separate AMOVAs for each new group defined by Nm and F_{ST} values was also performed, in which the variation between subpopulations ranged from 4.51% to 13.58% (Table 4).

AMOVA based on *pfcr*-flanking loci showed that variation among the subpopulations constituted 18.48% of the total variation when no grouping of subpopulations was inferred, and 13.06% variation among populations when subpopulations were grouped for seven different geographical regions. When the same analysis was used on three clusters detected with STRUCTURE, high variation (36.77%) was detected between clusters. Next, an AMOVA was used to determine the amount of variation for the four groups defined by interpretation of Nm and F_{ST} . A total of 15.32% variation was identified between these four

groups with 6.38% of variation between subpopulations within the group, and 78.30% of variation within subpopulations. Separate AMOVAs of each of the four groups was also performed, in which the variation between subpopulations ranged from 0.30% to 7.57% (Table 5). AMOVA results indicate that almost all the genetic variation was localized within groups or clusters resulting in genetic differentiation between different groups or clusters. There was little genetic differentiation found among subpopulations within the same geographic region ($F_{SC}=0.08$, $P<0.0001$); however, substantial genetic differentiation was found among the four geographic regions ($F_{CT}=0.13$, $P<0.0001$). The significant differences found between geographic regions lends support to the overall differences observed when all populations are combined ($F_{ST}=0.20$, $P<0.0001$), and support the division of the populations into four groups by geographic region, as done by pairwise F_{ST} and Nm values.

In the Mantel test, a significant positive correlation between $F_{ST}/(1-F_{ST})$ (Rousset's Distance Measurement) and the natural log of geographic distance was detected across all the populations for both neutral (Mantel test, $P=0.03$, $r=0.31$, $N=169$) and *pfcr*-flanking loci ($P=0.001$, $r=0.39$, $N=169$) (Fig. 4); however, different levels of genetic diversity at HTA and LTA may have skewed this data. To offset this potential bias, the analysis was repeated for each transmission area, separately. The division of populations resulted in non-significant Mantel probabilities and a decrease in the correlation coefficient for *pfcr*-flanking loci (in HTA; $r=0.26$, $P=0.10$ and in LTA; $r=0.28$, $P=0.18$). In the case of the neutral loci, LTA show a significant correlation coefficient ($r=0.50$, $P=0.03$, $N=83$); however, a decrease in the correlation coefficient ($r=0.15$, $P=0.23$, $N=86$) and a non-significant Mantel probability was observed in HTA.

3.6. Genetic relationship among individual multi-locus haplotypes

A median joining network permits visualization of probable mutation paths undergone by the *pfcr*-haplotypes (CVMNK, CVIET and SVMNT), which in turn helps to explain the evolutionary relationship between the haplotypes observed in the sample set. NETWORK analysis of multi-locus microsatellite haplotypes around various *pfcr*-haplotypes (Supplementary Table 2) generated two clusters of evolutionarily divergent clones. However, these two clusters are not separated by a high frequency of ancestral divisions (nodes). The 5-locus haplotype ST1 was found to be associated with the cluster predominately identified by SVMNT (yellow square box in Fig. 5). The ST1 appears to share a close relationship to many other 5-locus haplotypes (ST5, 7, 10, 16 and 18), which are found in different locations across India. The network indicates that the cluster of parasites with multi-locus haplotypes flanking SVMNT are of recent origin, while the cluster associated with the CVIET haplotype is more ancestral, indicated by an increased number of mutations between various 5-locus haplotypes associated with CVIET. The wild-type appears more closely related to the CVIET than that of the SVMNT. While the 5-locus haplotypes of SVMNT which are closely related with CVMNK were mostly from HTA. The cluster of CVIET haplotype shows close genetic relationship with Southeast Asian haplotype (CT6).

4. Discussion

Results from this study provide important information about the affect of chloroquine usage and transmission intensity on the population structure of Indian *P. falciparum* parasites.

4.1. Parasite diversity across Indian *P. falciparum* populations

P. falciparum populations from high transmission areas (HTA) generally shows higher genetic variation than those from low transmission areas (LTA) (Anderson et al., 2000) and

maintain multiclonal infections (*i.e.*, multiple genotypes per infected person) in such areas (Paul et al., 1998). We observed that genetic diversity within both neutral microsatellite loci and *pfcr*t-flanking microsatellite loci varies according to transmission intensity. The extensive microsatellite diversity observed within HTA corroborates earlier studies (Garg et al., 2007; Joshi, 2003), which demonstrated extensive polymorphism in surface antigens akin to high levels of multiple infections in HTA. Higher values for many measures of genetic diversity (A_r , P_R and H_e) within HTA indicate high rates of recombination and less inbreeding in parasite populations (Conway et al., 1999). In this study, a reduction in all measures of genetic variation was observed at flanking loci of resistant *pfcr*t-haplotypes, when compared with flanking loci of wild-type (Table 1, 2). However, such reduction was not observed at neutral loci in resistant *pfcr*t-haplotypes, suggesting that there may be selection around the *pfcr*t.

4.2. Trend of chloroquine selection

The significant reduction in genetic diversity (H_e) within the flanking loci of the resistant *pfcr*t-haplotypes relative to wild-type indicates chloroquine selective pressure on the *pfcr*t. The reduced width of the selective sweep around the resistant CVIET compared to the resistant SVMNT (Fig. 2), is possibly due to the higher recombination rates associated with HTA, while inbreeding and reduced recombination rates are associated with LTA (Anderson et al., 2000; Conway et al., 1999; Nash et al., 2005). An approximate two-fold higher H_e value for the flanking SVMNT haplotypes was observed in HTA, when compared to H_e value of flanking loci of the SVMNT haplotype from LTA (Table 2) reiterates higher recombination events in populations at HTA. Variable strength of selection at different geographic regions cannot be ignored, *i.e.*, weak selection strength at HTA in comparison to LTA. Another reason may be the order of colonization of different resistant haplotypes at various regions. The CVIET was possibly introduced earlier than SVMNT in India and may have evolved multiple times due to recombination events adding genetic diversity to the population under HTA epidemiology. In a selective sweep, reduction of heterozygosity was also accompanied by large regions of elevated linkage disequilibrium around resistant haplotypes. The observed LD value supports above narrower selection valley as it got reduced with addition of a distant loci at upstream (+106kb) onwards. The decline in LD is slightly lower for CVIET ($I_A^S=0.08$) in HTA than for SVMNT ($I_A^S=0.10$) in LTA, consent with a weaker selection event at HTA, discussed above. The difference in relative fitness of *pfcr*t-haplotypes within different epidemiological conditions at the time of sampling may have contributed to this differential pattern of selective sweep. However, the data presented here concurs with a recent report explained molecular basis of CQR differs with changing transmission intensities in India (Lumb et al., 2012). Hence, a different pattern of genotypes had developed in various regions (broadly HTA and LTA), which translated into the detection of geographical structure among CQ-resistant Indian *P. falciparum*.

4.3. Inference of population structure among Indian *P. falciparum*

The STRUCTURE analysis of three neutral loci yields two discrete clusters associated with transmission intensity; Neutral cluster1 was significantly comprised of isolates from HTA and Neutral cluster2 was significantly comprised of isolates from LTA (Table 3). A minimal admixture observed between HTA and LTA supports earlier reports of the association between the genetic structure of *P. falciparum* and patterns of transmission (Anderson et al., 2000). The STRUCTURE results also corroborate the differential measurements of genetic diversity at HTA and LTA. Further, pairwise F_{ST} estimates substantial levels of genetic differentiation between populations from HTA and LTA, in concordance with population differentiation indicated by STRUCTURE (Supplementary Table 4). The observed value of $Nm \gg 1$ between all of the subpopulations does not support genetic drift (Slatkin, 1987) as a cause of observed genetic differentiation.

Low population differentiation was detected between *P. falciparum* subpopulations within HTA except WBG, which has moderate differentiation when compared with other regions. Corroborating earlier reports of genetically similar surface antigens (*msp1* and *msp2*) between Assam and Eastern India (Joshi et al., 2007) indicate a pattern of current and/or historical migration. CN and Eastern India were also found to be genetically similar, indicating that there is considerable parasite gene-flow between mainland and island parasite populations, which supports previously reported data based on surface antigen (*ama-1*) (Garg et al., 2007). In LTA, MP showed low genetic differentiation with RAJ, GOA and TN ($F_{ST}=0.05$, $P>0.094$), RAJ-GUJ-UP and GOA-TN showed high genetic differentiation between them. The Southwest (GOA-TN) part of India has the most genetically isolated population from rest of India, except for MP. The three-level hierarchical analysis of genetic differentiation (AMOVA) supported the division of Indian *P. falciparum* population to four geographical groups, as maximum variations observed were within subpopulations rather than among subpopulations of a particular geographical group.

Geographical distance alone does not adequately explain the divergent parasite populations across India, especially given that populations of ASM-MAH, MAH-CN, and ORS-CN are greater than 1200 km apart and showed no genetic differentiation between them. However, the analysis of the Mantle test revealed a significant isolation by distance (IBD) at LTA only, *i.e.*, genetic differentiation depends on geographical distances between studied regions of LTA. No such correlation between genetic distances and geographic distances in HTA was observed. This again supports the division of Indian *P. falciparum* population to four geographical groups. Further, Mantel test involving all populations excluding the Southwest group revealed no correlation between genetic distances and geographic distances. It is likely that the genetic differentiation between the Southwest region and all other geographically distinct groups is responsible for the significant IBD observations.

4.4. Spread of hitchhiking in Indian *P. falciparum* population

Collectively, the results (STRUCTURE, F_{ST} , AMOVA and IBD) indicate subdivided populations of *P. falciparum* isolates in India. A deviation from a single random mating population may affect the amount of genetic hitchhiking around a beneficial mutation in a population. Interesting patterns of population differentiation were observed while comparing pairwise F_{ST} between neutral loci and *pfcr*-flanking loci (Supplementary Table 6a) at the four groups of population inferred above; Southwest and Northeast-East-Island show highly differentiated populations at both neutral and *pfcr*-flanking loci indicating heterogeneous population and low gene-flow between them; Central and Northeast-East-Island or Northwest show low genetic differentiation at both neutral and *pfcr*-flanking loci indicating homogeneous population and continuous gene-flow between them; Central and Southwest show moderate differentiation at both neutral and *pfcr*-flanking loci indicating a limited gene-flow between populations; Southwest and Northwest show high genetic differentiation at neutral loci but low genetic differentiation at *pfcr*-flanking loci indicating that evolutionary process (*i.e.*, strength of selection) is faster than rate of migration. In a subdivided population, genetic hitchhiking can introduce population differentiation (large F_{ST}) in an initially homogeneous population (small F_{ST}) (Bierne, 2010; Slatkin and Wiehe, 1998) and the similar pattern is displayed between Northwest and Northeast-East-Island groups (Supplementary Table 6a). The observed gradient of heterozygosity for a selective sweep may be used to infer the geographical movement of the beneficial mutation in different subpopulation, as reduction in expected heterozygosity initially accumulates in subpopulations where beneficial mutation are introduced and subsequently increase with the spread across neighboring subpopulations (Kim and Maruki, 2011). Finally, our data also show a gradual increase of heterozygosity in the flanking loci of SVMNT from the

Southwest to Northeast-East-Island part of India, which suggests a probable introduction of the SVMNT haplotype in southern India and then spread towards other regions.

4.5. Probable migratory route of chloroquine-resistant haplotypes

The STRUCTURE analysis of *pfcr*-flanking loci showed 3 discrete clusters associated with the different *pfcr*-haplotypes (CVIET, SVMNT, and CVMNK). This in turn reflects the earlier observation of varied level of selective sweep associated with regional variation in chloroquine pressure (Lumb et al., 2012). The pairwise F_{ST} estimate at *pfcr*-flanking loci show high genetic differentiation between East-Island (ORS, JHK, CHG, CN) and Northwest (UP, RAJ, GUJ), strong differentiation between East-Island and Southwest (GOA&TN), indicating a limited migration of resistant haplotypes between East-Island with that of Northwest and Southwest. In turn, the Northeast region may have indeed, been a major migratory route of the resistant haplotypes, likely originating from the Southeast Asia region and then spreading into Eastern India and consequently to other parts of India. Meanwhile, the moderate genetic differentiation ($F_{ST}=0.09-0.15$) observed between the Central (MP, MAH) group and all other groups indicate continuous genetic interaction between the resistant parasites in central and other parts of India.

Admittedly, more neutral loci may be needed to make concrete conclusions about geographic structure in Indian *P. falciparum*, but even with these three neutral loci, there is evidence of a genetic structure that is strongly linked with the patterns of malaria transmission. Additionally, resistant haplotypes are also differentially structured (Table 3), probably due to above found geographic structure or varied chloroquine usage in these locations. While determining F -statistics estimations, a small amount of genetic exchange between populations is enough to prevent the accumulation of large genetic differences between them. Thus, the significant strong genetic differentiation found between Southwest and Northeast-East-Island part of India implies limitations in direct dispersal of resistant haplotypes and supports the movement via Central India.

Human migration between malaria-endemic regions plays an important role in the movement of resistant alleles in distant populations (Hume et al., 2003; Lqbal et al., 2002). For example, two studies based on *pfcr*-flanking microsatellite markers (Mixson-Hayden et al., 2010; Rawasia et al., 2012) reported that the resistant SVMNT found in India and Pakistan migrated from PNG. Evidence for this was also found by clusters of neutral microsatellite markers that were reported earlier between PNG (Pacific region) and Thailand (Southeast Asia) (Anderson et al., 2000). In this study, the microsatellite profile of *pfcr*-flanking loci associated with the resistant SVMNT and CVIET resembles the microsatellite profile reported from PNG and SEA, respectively (Mixson-Hayden et al., 2010; Wootton et al., 2002). A recent study (Awasthi et al., 2011) examining the allele frequency of various *pfcr*-haplotypes postulates a probable route of migration for different *pfcr*-haplotypes in India and reported that the resistant SVMNT arrived India from PNG via SEA. Here, the population differentiation, gene-flow and isolation by distance observed between different regions of India stimulate to postulate another probable route for dispersal of resistant SVMNT across India through Sri Lanka or southern India (Fig 3)

Human migration related to labor and tourism had been associated with prevalence of CQR in India, particularly laborers traveling from eastern part of India to the western states (Sethi et al., 1990; Sharma and Sharma, 1988; Sharma, 2000). The study site in TN (Rameswaram Island) has been associated with continuous human migration with Sri Lanka, and being a Hindu holy place, it receives a large number of pilgrims from all over India. Many tourists also stay here before or after visiting Sri Lanka. These human migrations have likely helped to facilitate the spread of chloroquine-resistant parasites between both countries and throughout India (Rajagopalan et al., 1986). This, in turn, may have given chance for

immigration of SVMNT from Sri Lanka to India as a recent study reported ubiquitous appearance of SVMNT in Sri Lanka (Zhang et al., 2011) and we also observed similar frequency of SVMNT from this study site. The predominance of similar mosquito vectors (*Anopheles culicifacies*; sibling B,E) in both locations (Surendran et al., 2000) also supports this hypothesis.

It is now known about human migration between southern India, Sri Lanka and Indonesia for trade and cultural exchange was much greater than hitherto thought (Karafet et al., 2005). Similarly, Pakistan shares international borders with Rajasthan and Gujarat, which reports the fixation of the SVMNT haplotype with similar *pfprt*-flanking loci in both countries (Mallick et al., 2012; Rawasia et al., 2012). On the other hand, countries sharing a border with northeast India, such as Nepal, Bangladesh, and Myanmar, reports higher frequency of CVIET (Banjara et al., 2011; Kawai et al., 2011; Mohapatra et al., 2005). It is worth putting the South Asia-Southeast Asia migration corridor in perspective. The above discussed active routes of human migration lead us to propose immigration of the SVMNT from PNG to India via Sri Lanka or southern India. Finally, the results of median joining NETWORK construction reveal two distinct clusters of resistant CVIET and SVMNT, indicating accumulation of resistant haplotypes on a distinct genetic background. However, SVMNT cluster supports a more recent origin than that of the CVIET cluster, which seems to have undergone many mutational steps over time in HTA epidemiology. It is also possible that multiple selective sweeps have occurred independently in different geographic locations of India due to local adaptation.

Conclusions

The molecular analyses presented here suggest strong genetic subdivision of *P. falciparum* between high and low transmission regions of India, possibly related to contrasting chloroquine utilization in each epidemiological type. Within each regional group, low levels of parasite genetic differentiation were observed across India. The genetic structure determined by the two groups of markers (neutral and *pfprt*-flanking) indicates at least four geographical subpopulations of Indian *P. falciparum* isolates, circulating among each other. The pattern of genetic hitchhiking suggests spread of resistant SVMNT from southern India to other parts of India.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

PKM was supported by funds from the Council of Scientific and Industrial Research, India and grant 5D43TW007884-05 "Promotion of Plasmodium Research and Training in India" from the U.S. National Institutes of Health/Fogarty International Center and by NIMR (ICMR), New Delhi. PLS was supported by an NIH/Fogarty International Center U.S. Global Health Postdoctoral Scientist Fellowship 3D43TW007884-03S1. A set of samples used in the study were collected during a project of therapeutic efficacy of chloroquine supported by a grant USAID SE/05/209279. All laboratory isolates, including MRA819G, MRA820G, MRA150G, MRA155G, MRA152G, MRA818G, MRA398, MRA102G, were kindly provided by the Malaria Reagent and Reference Resource Center (MR4). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Fogarty International Center or the National Institutes of Health or NIMR. We sincerely thank all the staff members of NIMR, Delhi and the respective field units, for their help, cooperation and support during the study. This research article bears the NIMR publication committee approval number 028/2012.

References

Abdel-Muhsin AM, Mackinnon MJ, Ali E, Nassir el KA, Suleiman S, Ahmed S, Walliker D, Babiker HA. Evolution of drug-resistance genes in *Plasmodium falciparum* in an area of seasonal malaria

- transmission in Eastern Sudan. *The Journal of infectious diseases*. 2004; 189:1239–1244. [PubMed: 15031793]
- Alam MT, Vinayak S, Congpuong K, Wongsrichanalai C, Satimai W, Slutsker L, Escalante AA, Barnwell JW, Udhayakumar V. Tracking origins and spread of sulfadoxine-resistant *Plasmodium falciparum* dhps alleles in Thailand. *Antimicrobial agents and chemotherapy*. 2011; 55:155–164. [PubMed: 20956597]
- Anderson TJ. Mapping drug resistance genes in *Plasmodium falciparum* by genome-wide association. *Current drug targets. Infectious disorders*. 2004; 4:65–78. [PubMed: 15032635]
- Anderson TJ, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurwa S, French N, Whitworth J, Velez ID, Brockman AH, Nosten F, Ferreira MU, Day KP. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Molecular biology and evolution*. 2000; 17:1467–1482. [PubMed: 11018154]
- Anderson TJ, Su XZ, Bockarie M, Lagog M, Day KP. Twelve microsatellite markers for characterization of *Plasmodium falciparum* from finger-prick blood samples. *Parasitology*. 1999; 119(Pt 2):113–125. [PubMed: 10466118]
- Awasthi G, Prasad GB, Das A. Population genetic analyses of *Plasmodium falciparum* chloroquine receptor transporter gene haplotypes reveal the evolutionary history of chloroquine-resistant malaria in India. *International journal for parasitology*. 2011; 41:705–709. [PubMed: 21447338]
- Banjara MR, Imwong M, Petmitr S, Sirawaraporn W, Joshi AB, Chavalitshewinkoon-Petmitr P. Factors associated with regional bias of pfcrt (plasmodium falciparum chloroquine resistance transporter) haplotypes in Nepal. *The Southeast Asian journal of tropical medicine and public health*. 2011; 42:1–8. [PubMed: 21323158]
- Barton NH. Genetic hitchhiking. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 2000; 355:1553–1562.
- Bierne N. The distinctive footprints of local hitchhiking in a varied environment and global hitchhiking in a subdivided population. *Evolution; international journal of organic evolution*. 2010; 64:3254–3272.
- Biswas S, Seth RK, Tyagi PK, Sharma SK, Dash AP. Naturally acquired immunity and reduced susceptibility to falciparum malaria in two subpopulations of endemic eastern India. *Scandinavian journal of immunology*. 2008; 67:177–184. [PubMed: 18086262]
- Chen N, Wilson DW, Pasay C, Bell D, Martin LB, Kyle D, Cheng Q. Origin and dissemination of chloroquine-resistant *Plasmodium falciparum* with mutant pfcrt alleles in the Philippines. *Antimicrobial agents and chemotherapy*. 2005; 49:2102–2105. [PubMed: 15855538]
- Conway DJ, Roper C, Oduola AM, Arnot DE, Kremsner PG, Grobusch MP, Curtis CF, Greenwood BM. High recombination rate in natural populations of *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences of the United States of America*. 1999; 96:4506–4511. [PubMed: 10200292]
- Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online*. 2005; 1:47–50. [PubMed: 19325852]
- Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 1992; 131:479–491. [PubMed: 1644282]
- Garg S, Alam MT, Das MK, Dev V, Kumar A, Dash AP, Sharma YD. Sequence diversity and natural selection at domain I of the apical membrane antigen 1 among Indian *Plasmodium falciparum* populations. *Malaria journal*. 2007; 6:154. [PubMed: 18031585]
- Hartl, DL. A primer of population genetics. Sinauer Associates, Inc; Sunderland, Massachusetts: 1988.
- Hartl, DL.; Clark, AG. Principles of Population Genetics. Sinauer Associates; Sunderland, MA: 1997.
- Haubold B, Hudson RR. LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Linkage Analysis. Bioinformatics*. 2000; 16:847–848. [PubMed: 11108709]
- Hume JC, Lyons EJ, Day KP. Human migration, mosquitoes and the evolution of *Plasmodium falciparum*. *Trends in parasitology*. 2003; 19:144–149. [PubMed: 12643998]

- Joshi H. Markers for population genetic analysis of human plasmodia species, *P. falciparum* and *P. vivax*. *Journal of vector borne diseases*. 2003; 40:78–83. [PubMed: 15119076]
- Joshi H, Valecha N, Verma A, Kaul A, Mallick PK, Shalini S, Prajapati SK, Sharma SK, Dev V, Biswas S, Nanda N, Malhotra MS, Subbarao SK, Dash AP. Genetic structure of *Plasmodium falciparum* field isolates in eastern and north-eastern India. *Malaria journal*. 2007; 6:60. [PubMed: 17517129]
- Kalinowski S. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics*. 2004; 5:539–543.
- Karafet TM, Lansing JS, Redd AJ, Reznikova S, Watkins JC, Surata SP, Arthawiguna WA, Mayer L, Bamshad M, Jorde LB, Hammer MF. Balinese Y-chromosome perspective on the peopling of Indonesia: genetic contributions from pre-neolithic hunter-gatherers, Austronesian farmers, and Indian traders. *Human biology*. 2005; 77:93–114. [PubMed: 16114819]
- Kawai A, Arita N, Matsumoto Y, Kawabata M, Chowdhury MS, Saito-Ito A. Efficacy of chloroquine plus primaquine treatment and pfcrt mutation in uncomplicated falciparum malaria patients in Rangamati, Bangladesh. *Parasitology international*. 2011; 60:341–346. [PubMed: 21645634]
- Kim Y, Maruki T. Hitchhiking effect of a beneficial mutation spreading in a subdivided population. *Genetics*. 2011; 189:213–226. [PubMed: 21705748]
- Kim Y, Stephan W. Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics*. 2002; 160:765–777. [PubMed: 11861577]
- Lqbal J, Sher A, Hira PR, Al-Aniezi A. Drug-resistant *Plasmodium falciparum* infection in immigrants and non-immune travellers. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2002; 8:734–738. [PubMed: 12445011]
- Lumb V, Madan R, Das MK, Rawat V, Dev V, Khan W, Khan H, Sharma YD. Differential genetic hitchhiking around mutant pfcrt alleles in the Indian *Plasmodium falciparum* population. *The Journal of antimicrobial chemotherapy*. 2012; 67:600–608. [PubMed: 22207595]
- Mallick PK, Joshi H, Valecha N, Sharma SK, Eapen A, Bhatt RM, Srivastava HC, Sutton PL, Dash AP, Bhasin VK. Mutant pfcrt “SVMNT” haplotype and wild type pfmdr1 “N86” are endemic in *Plasmodium vivax* dominated areas of India under high chloroquine exposure. *Malaria journal*. 2012; 11:16. [PubMed: 22236376]
- Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer research*. 1967; 27:209–220. [PubMed: 6018555]
- McCollum AM, Mueller K, Villegas L, Udhayakumar V, Escalante AA. Common origin and fixation of *Plasmodium falciparum* dhfr and dhps mutations associated with sulfadoxine-pyrimethamine resistance in a low-transmission area in South America. *Antimicrobial agents and chemotherapy*. 2007; 51:2085–2091. [PubMed: 17283199]
- Mehlotra RK, Mattera G, Bhatia K, Reeder JC, Stoneking M, Zimmerman PA. Insight into the early spread of chloroquine-resistant *Plasmodium falciparum* infections in Papua New Guinea. *The Journal of infectious diseases*. 2005; 192:2174–2179. [PubMed: 16288385]
- Mixson-Hayden T, Jain V, McCollum AM, Poe A, Nagpal AC, Dash AP, Stiles JK, Udhayakumar V, Singh N. Evidence of selective sweeps in genes conferring resistance to chloroquine and pyrimethamine in *Plasmodium falciparum* isolates in India. *Antimicrobial agents and chemotherapy*. 2010; 54:997–1006. [PubMed: 20038626]
- Mohapatra PK, Prakash A, Taison K, Negmu K, Gohain AC, Namchoom NS, Wange D, Bhattacharyya DR, Goswami BK, Borgohain BK, Mahanta J. Evaluation of chloroquine (CQ) and sulphadoxine/pyrimethamine (SP) therapy in uncomplicated falciparum malaria in IndoMyanmar border areas. *Tropical medicine & international health : TM & IH*. 2005; 10:478–483. [PubMed: 15860095]
- Mu J, Awadalla P, Duan J, McGee KM, Joy DA, McVean GA, Su XZ. Recombination hotspots and population structure in *Plasmodium falciparum*. *PLoS biology*. 2005; 3:e335. [PubMed: 16144426]
- Nair S, Williams JT, Brockman A, Paiphun L, Mayxay M, Newton PN, Guthmann JP, Smithuis FM, Hien TT, White NJ, Nosten F, Anderson TJ. A selective sweep driven by pyrimethamine treatment

in southeast asian malaria parasites. *Molecular biology and evolution*. 2003; 20:1526–1536. [PubMed: 12832643]

Nash D, Nair S, Mayxay M, Newton PN, Guthmann JP, Nosten F, Anderson TJ. Selection strength and hitchhiking around two anti-malarial resistance genes. *Proceedings. Biological sciences / The Royal Society*. 2005; 272:1153–1161.

Paul RE, Hackford I, Brockman A, Muller-Graf C, Price R, Luxemburger C, White NJ, Nosten F, Day KP. Transmission intensity and *Plasmodium falciparum* diversity on the northwestern border of Thailand. *The American journal of tropical medicine and hygiene*. 1998; 58:195–203. [PubMed: 9502604]

Pearce R, Malisa A, Kachur SP, Barnes K, Sharp B, Roper C. Reduced variation around drug-resistant dhfr alleles in African *Plasmodium falciparum*. *Molecular biology and evolution*. 2005; 22:1834–1844. [PubMed: 15917494]

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000; 155:945–959. [PubMed: 10835412]

Rajagopalan PK, Jambulingam P, Sabesan S, Krishnamoorthy K, Rajendran S, Gunasekaran K, Kumar NP, Prothero RM. Population movement and malaria persistence in Rameswaram Island. *Social science & medicine*. 1986; 22:879–886. [PubMed: 3749960]

Rawasia WF, Sridaran S, Patel JC, Abdallah J, Ghanchi NK, Barnwell JW, Escalante AA, Udhayakumar V, Beg MA. Genetic backgrounds of the *Plasmodium falciparum* chloroquine resistant transporter (pfcrt) alleles in Pakistan. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*. 2012; 12:278–281.

Roper C, Pearce R, Nair S, Sharp B, Nosten F, Anderson T. Intercontinental spread of pyrimethamine-resistant malaria. *Science*. 2004; 305:1124. [PubMed: 15326348]

Rosenberg NA, Li LM, Ward R, Pritchard JK. Informativeness of genetic markers for inference of ancestry. *American journal of human genetics*. 2003; 73:1402–1422. [PubMed: 14631557]

Sethi NK, Choudhri Y, Chuttani CS. Role of migratory population in keeping up endemicity of malaria in metropolitan cities of India. *The Journal of communicable diseases*. 1990; 22:86–91. [PubMed: 2098415]

Sharma RC, Sharma VP. Epidemiological implications of population migration: Part II. Evidence of chloroquine resistant *Plasmodium falciparum* malaria in Kheda district, Gujarat. *Indian journal of malariology*. 1988; 25:117–118. [PubMed: 3077368]

Sharma, V. Status of drug resistance in malaria in India. Narosa Publications; Delhi: 2000.

Slatkin M. Gene flow and the geographic structure of natural populations. *Science*. 1987; 236:787–792. [PubMed: 3576198]

Slatkin M, Wiehe T. Genetic hitch-hiking in a subdivided population. *Genetical research*. 1998; 71:155–160. [PubMed: 9717437]

Smith JM, Haigh J. The hitch-hiking effect of a favourable gene. *Genetical research*. 1974; 23:23–35. [PubMed: 4407212]

Snounou G, Zhu X, Siripoon N, Jarra W, Thaithong S, Brown KN, Viriyakosol S. Biased distribution of msp1 and msp2 allelic variants in *Plasmodium falciparum* populations in Thailand. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1999; 93:369–374. [PubMed: 10674079]

Surendran SN, Abhayawardana TA, De Silva BG, Ramasamy R, Ramasamy MS. Anopheles culicifacies Y-chromosome dimorphism indicates sibling species (B and E) with different malaria vector potential in Sri Lanka. *Medical and veterinary entomology*. 2000; 14:437–440. [PubMed: 11129709]

Szpiech ZA, Jakobsson M, Rosenberg NA. ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*. 2008; 24:2498–2504. [PubMed: 18779233]

Vinayak S, Alam MT, Mixson-Hayden T, McCollum AM, Sem R, Shah NK, Lim P, Muth S, Rogers WO, Fandeur T, Barnwell JW, Escalante AA, Wongsrichanalai C, Ariei F, Meshnick SR, Udhayakumar V. Origin and evolution of sulfadoxine resistant *Plasmodium falciparum*. *PLoS pathogens*. 2010; 6:e1000830. [PubMed: 20360965]

Weir B, Cockerham C. Estimating F-statistics for the analysis of population structure. *Evolution; international journal of organic evolution*. 1984; 38:1358–1370.

- Wootton JC, Feng X, Ferdig MT, Cooper RA, Mu J, Baruch DI, Magill AJ, Su XZ. Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. *Nature*. 2002; 418:320–323. [PubMed: 12124623]
- Wright, S. *Evolution and the Genetics of Population, Variability Within and Among Natural Populations*. The University of Chicago Press; Chicago: 1978.
- Zhang JJ, Senaratne TN, Daniels R, Valim C, Alifrangis M, Amerasinghe P, Konradsen F, Rajakaruna R, Wirth DF, Karunaweera ND. Distribution pattern of *Plasmodium falciparum* chloroquine transporter (pfcrt) gene haplotypes in Sri Lanka 1996–2006. *The American journal of tropical medicine and hygiene*. 2011; 85:811–814. [PubMed: 22049031]

Microsatellite markers revealed a geographic structure in Indian *P. falciparum*.
Four groups of subpopulations observed in Indian *P. falciparum* isolates.
Spread of resistant SVMNT from southern India to other parts of India was noted.
Transmission intensity acts as an efficient driver in spread of CQ-resistant parasite.

1. Gautam Budh Nagar, Uttar Pradesh (2002)

(N=17)

SVMNT- 100%

2. Udaipur, Rajasthan (2003)

(N=26)

SVMNT- 96%, CVMNK- 4%

3. Panna, Madhya Pradesh (2004)

(N=17)

SVMNT- 100%

4. Anand, Gujarat (2005)

(N=33)

SVMNT- 100%

5. Gardchiroli, Maharashtra (2002)

(N=14)

SVMNT- 64%, CVMNK- 36%

6. Panaji, Goa (2004)

(N=26)

SVMNT- 96%, CVIET- 4%

7. Rameswaram, Tamil Nadu (2004)

(N=25)

SVMNT- 100%

13. Kamrup, Assam (2002)

(N=26)

SVMNT- 42%, CVIET-58%

12. Darjeeling, West Bengal (2003)

(N=16)

SVMNT- 75%, CVIET- 25%

11. Simdega, Jharkhand (2006)

(N=34)

CVMNK- 20%, SVMNT- 18%

CVIET- 59%, CVMNT- 3%

10. Sundergarh, Odisha (2003)

(N=54)

CVMNK- 20%, SVMNT- 48%

CVIET- 32%

9. Kanker, Chhattisgarh (2006)

(N=16)

CVMNK- 75%, SVMNT- 25%

8. Car Nicobar (2002)

(N=12)

CVMNK- 12%, SVMNT- 50%

CVIET- 38%

Fig. 1. Distribution of *pfcrT*-haplotypes

Distribution of *pfcrT*-haplotypes in different geographic locations. N; number of isolates; *pfcrT*-haplotypes shown are CVMNK (wild), CVIET (resistant) and SVMNT (resistant).

Fig. 2. Expected heterozygosity (H_e) values for *pfcr1*-flanking microsatellite loci
 The expected heterozygosity (H_e) values for microsatellite loci flanking (-24kb to +22kb) of various *pfcr1*-haplotypes.

Fig. 3. Migration of chloroquine-resistant *pfprt*-haplotypes in India

An illustration of gene-flow in chloroquine-resistant *P. falciparum* across India. The black arrow shows the significant pairwise F_{ST} among the four groups of population outlined in green (Northeast-East-Island), orange (Southwest), pink (Northwest) and blue (Central). The red arrows shows probable gene-flow of various resistant *pfprt*-haplotypes in India. The blue arrows shows another probable route of gene-flow in resistant SVMNT of Indian *P. falciparum* isolates.

Fig. 4. Isolation by distance plot

Analysis of isolation by distance on correlating genetic distance ($F_{ST}/1-F_{ST}$) vs. natural logarithm of pairwise geographic distance. Results involving three neutral loci for total samples are shown in rhombus ($r^2=0.0941$, $P=0.001$) and results comprising seven flanking loci for total samples are shown in triangles ($r^2=0.1519$, $P=0.03$).

Fig. 5. Network analysis of various *pfprt*-haplotypes in Indian *P. falciparum* isolates
Network diagram illustrating two independent clusters related to *pfprt*-haplotype i.e., indicating different genetic background associated with various *pfprt*-haplotypes. SVMNT haplotype in yellow square box and CVIET haplotype in red circle.

Table 1

Allelic diversity in terms of allelic richness and private allelic richness at various geographic regions and *pfcr*-haplotypes.

Area(g=7) ^a	Neutral loci		<i>Pfcr</i> -flanking loci	
	Mean Allelic Richness (± s.d)	Mean Private Allelic richness(± s.d)	Mean Allelic Richness (± s.d)	Mean Private Allelic Richness (± s.d)
Assam	4.32(0.12)	0.85(0.15)	2.30(0.11)	0.25(0.15)
West Bengal	3.46(0.24)	0.36(0.30)	2.59(0.24)	0.10(0.05)
Orissa	4.63(0.23)	0.46(0.09)	2.50(0.31)	0.18(0.08)
Jharkhand	4.40(0.36)	0.44(0.42)	2.90(0.37)	0.40(0.17)
Chhattisgarh	4.04(0.29)	0.37(0.37)	3.20(0.42)	0.61(0.35)
Maharashtra	4.66(0.44)	0.11(0.06)	2.95(0.17)	0.62(0.18)
Car Nicobar	3.66(0.33)	0.26(0.21)	2.00(0.31)	0
Madhya Pradesh	3.28(0.53)	0.19(0.17)	2.02(0.48)	0.26(0.16)
Gujarat	2.51(0.08)	0.18(0.14)	1.15(0.15)	0
Rajasthan	3.29(0.53)	0.08(0.08)	1.31(0.14)	0.04(0.03)
Goa	2.10(0.47)	0.03(0.03)	1.38(0.17)	0.10(0.07)
Tamil Nadu	1.85(0.21)	0	1.17(0.17)	0
Uttar Pradesh	2.5(0.31)	0	1.17(0.18)	0
<i>pfcr</i> -haplotypes(g=16) ^a				
CVMNK	6.33(1.45)	1.48(0.91)	6.00(0.54)	3.18(0.47)
CVIET	6.86(0.36)	1.93(0.18)	2.92(0.33)	0.45(0.28)
SVMNT	5.53(0.49)	1.22(0.43)	2.10(0.21)	0.16(0.05)

^a
g= standardized sample size

Table 2

Expected heterozygosity (H_e) in different geographic regions and *pfcr*-haplotypes.

AREA	N	microsatellite loci flanking <i>pfcr</i> gene										putative neutral microsatellite loci					Total Mean(±s.d)
		BSM97 -24kb	B5M77 -20kb	2E10 -5kb	9B12 +1kb	PE12a +6kb	2H4 +22kb	PE14F +106kb	Mean(±s.d)	TAA81 Chr13	ARA2 Chr11	TAA60 Chr5	Mean(±s.d)				
Assam	13	0.50	0.29	0.51	0.50	0.56	0.83	0.91	0.59 (0.08)	0.81	0.83	0.83	0.82 (0.05)	0.66 (0.08)			
West Bengal	13	0.76	0.56	0.59	0.41	0.38	0.53	0.81	0.58 (0.09)	0.60	0.69	0.76	0.68 (0.12)	0.61 (0.10)			
Orissa	22	0.65	0.18	0.61	0.73	0.51	0.79	0.82	0.61 (0.05)	0.81	0.84	0.89	0.85 (0.32)	0.69 (0.13)			
Jharkhand	15	0.74	0.47	0.48	0.80	0.45	0.66	0.67	0.61 (0.08)	0.89	0.82	0.80	0.83 (0.03)	0.68 (0.07)			
Chhattisgarh	8	0.79	0.64	0.71	0.61	0.43	0.89	0.93	0.71 (0.10)	0.86	0.75	0.79	0.80 (0.07)	0.74 (0.09)			
Maharashtra	8	0.46	0.46	0.64	0.46	0.61	0.86	0.75	0.61 (0.10)	0.89	0.89	0.82	0.87 (0.05)	0.69 (0.09)			
Car Nicobar	7	0.67	0	0.48	0.57	0.57	0.81	0.67	0.54 (0.07)	0.86	0.67	0.81	0.78 (0.07)	0.61 (0.07)			
M. Pradesh	9	0.69	0.50	0.42	0	0	0.58	0.64	0.40 (0.08)	0.83	0.56	0.56	0.65 (0.11)	0.48 (0.09)			
Gujarat	15	0.25	0	0	0	0	0.37	0.53	0.16 (0.05)	0.51	0.47	0.59	0.52 (0.10)	0.27 (0.06)			
Rajasthan	21	0	0.18	0.19	0	0.10	0.70	0.75	0.27 (0.06)	0.77	0.77	0.45	0.67 (0.06)	0.39 (0.06)			
Goa	14	0.36	0.14	0	0	0.14	0.38	0.47	0.22 (0.08)	0.58	0.26	0.14	0.33 (0.11)	0.25 (0.09)			
Tamil Nadu	16	0.33	0	0	0	0	0.13	0.69	0.16 (0.03)	0.50	0.34	0.13	0.32 (0.09)	0.21 (0.05)			
U. Pradesh	8	0.25	0	0	0	0	0.25	0.57	0.15 (0.04)	0.61	0.25	0.46	0.44 (0.14)	0.24 (0.07)			
Total	169																
<hr/>																	
<i>Pfcr</i> -haplotypes																	
		-24Kb	-20Kb	-5Kb	+1Kb	+6Kb	+22Kb	+106Kb		Chr13	Chr11	Chr5					
CVMNK	16	0.85	0.75	0.88	0.86	0.51	0.95	0.87	0.81 (0.07)	0.94	0.76	0.87	0.86 (0.04)	0.83 (0.06)			
CVIET	31	0.63	0.41	0.22	0.68	0.26	0.73	0.76	0.53 (0.06)	0.85	0.82	0.85	0.84 (0.04)	0.62 (0.05)			
SVMNT	122	0.35	0.19	0.16	0.05	0.14	0.59	0.76	0.32 (0.04)	0.75	0.79	0.63	0.72 (0.03)	0.44 (0.04)			
SVMNT <i>LTA</i>	82	0.28	0.14	0.1	0	0.05	0.43	0.71	0.24 (0.04)	0.68	0.72	0.38	0.59 (0.05)	0.38 (0.04)			
SVMNT <i>HTA</i>	40	0.45	0.27	0.23	0.4	0.27	0.76	0.79	0.45 (0.07)	0.87	0.84	0.83	0.84 (0.03)	0.57 (0.06)			

*HTA*₁ in high transmission areas;

*LTA*₁ in low transmission areas; N is the number of single clone isolates; Chr is chromosome number; H_e values <0.45 are shaded grey

Table 3

The estimated membership probability of 13 geographic parasite populations in different clusters of STRUCTURE.

Population	N	Mean inferred cluster of neutral loci			Mean inferred cluster of flanking loci		
		1	2		1	2	3
Assam	13	0.628	0.372		0.483	0.339	0.179
Orissa	22	0.690	0.310		0.341	0.533	0.126
West Bengal	13	0.723	0.277		0.309	0.222	0.468
Jharkhand	15	0.675	0.325		0.154	0.551	0.295
Chhattisgarh	8	0.851	0.149		0.300	0.119	0.581
Maharashtra	8	0.802	0.198		0.522	0.016	0.463
Car Nicobar	7	0.717	0.283		0.323	0.606	0.071
Goa	14	0.186	0.814		0.871	0.030	0.099
Gujarat	15	0.270	0.730		0.949	0.013	0.038
Rajasthan	21	0.323	0.677		0.822	0.020	0.158
Tamil Nadu	16	0.140	0.860		0.946	0.013	0.040
Uttar Pradesh	8	0.257	0.743		0.950	0.013	0.038
Madhya Pradesh	9	0.305	0.695		0.592	0.026	0.382
<i>pfcr</i> -haplotypes							
CVMNK	16	0.751	0.249		0.084	0.277	0.638
CVIET	31	0.712	0.288		0.047	0.825	0.128
SVMNT	122	0.390	0.610		0.802	0.034	0.163

N: number of single-clone isolates, the prominent membership of cluster is shaded grey at different geographic location and *pfcr*-haplotypes.

Table 4

Analysis of Molecular Variance (AMOVA) among Indian *P. falciparum* isolates based on three neutral microsatellite markers.

Population	Source of Variation	Degree of freedom	Sum of Squares	Variance Components	Percentage of Variation	Fixation Indices (P value)
All population	Among populations	12	43.855	0.20801	17.51	$F_{ST}=0.18$ (<0.0001)
	Within subpopulations	156	152.903	0.98015	82.49	
The seven regional groups	Among populations	6	27.508	0.09241	7.70	$F_{CT}=0.08$ (0.04)
	Among subpopulation within populations	6	16.347	0.12756	10.63	$F_{SC}=0.12$ (<0.0001)
	Within subpopulations	156	152.903	0.98015	81.67	$F_{ST}=0.18$ (<0.0001)
The population grouped by transmission areas	Among clusters	1	16.806	0.16702	13.22	$F_{CT}=0.13$ (0.0004)
	Among subpopulation within clusters	11	27.049	0.11657	9.22	$F_{SC}=0.11$ (<0.0001)
	Within clusters	156	152.903	0.98015	77.56	$F_{ST}=0.22$ (<0.0001)
The 2 clusters detected with STRUCTURE	Among clusters	1	21.520	0.22324	17.26	$F_{CT}=0.17$ (<0.0001)
	Among subpopulation within clusters	21	42.469	0.15798	12.21	$F_{SC}=0.15$ (<0.0001)
	Within clusters	147	134.099	0.91224	70.53	$F_{ST}=0.29$ (<0.0001)
Four populations defined by F_{ST} and Nm values	Among populations	3	25.027	0.16194	13.17	$F_{CT}=0.13$ (0.0001)
	Among subpopulation within populations	9	18.828	0.08782	7.14	$F_{SC}=0.08$ (<0.0001)
	Within subpopulations	156	152.903	0.98015	79.69	$F_{ST}=0.20$ (<0.0001)
Northeast-East-Island	Among populations	5	9.615	0.05689	4.51	$F_{ST}=0.05$ (0.002)
	Within subpopulations	72	86.744	1.20478	95.49	
Central	Among populations	1	2.627	0.17706	13.58	$F_{ST}=0.14$ (0.004)
	Within subpopulations	15	16.903	1.12685	86.42	
South Western	Among populations	1	1.255	0.05131	9.50	$F_{ST}=0.10$ (0.05)
	Within subpopulations	28	13.679	0.48852	90.50	
North Western	Among populations	2	5.332	0.13121	13.13	$F_{ST}=0.13$ (0.0003)
	Within subpopulations	41	35.577	0.86774	86.87	

Table 5

Analysis of Molecular Variance (AMOVA) of Indian *P. falciparum* isolates based on seven *pfprt*-flanking microsatellite markers.

Population	Source of Variation	Degree of freedom	Sum of Squares	Variance Components	Percentage of Variation	Fixation Indices (<i>P</i> value)
All population	Among populations	12	68.605	0.33110	18.48	$F_{ST}=0.18$ (<0.0001)
	Within subpopulations	156	227.791	1.46020	81.52	
The seven regional groups	Among populations	6	49.678	0.23798	13.06	$F_{CT}=0.13$ (0.01)
	Among subpopulation within populations	6	18.927	0.12390	6.80	$F_{SC}=0.08$ (0.0002)
	Within subpopulations	156	227.791	1.46020	80.14	$F_{ST}=0.20$ (0.04)
The 3 clusters grouped by transmission areas	Among clusters	1	33.770	0.35903	18.38	$F_{CT}=0.18$ (0.002)
	Among subpopulation within clusters	11	34.835	0.13452	6.89	$F_{SC}=0.08$ (<0.0001)
	Within clusters	156	227.791	1.46020	74.74	$F_{ST}=0.25$ (<0.0001)
The 3 clusters detected with STRUSTRUCTURE	Among clusters	2	79.335	0.77667	36.77	$F_{CT}=0.37$ (<0.0001)
	Among subpopulation within clusters	24	57.053	0.20634	9.77	$F_{SC}=0.15$ (<0.0001)
	Within clusters	143	161.477	1.12921	53.46	$F_{ST}=0.47$ (<0.0001)
Four populations defined by F_{ST} -and Nm values	Among populations	3	41.907	0.28580	15.32	$F_{CT}=0.15$ (0.0006)
	Among subpopulation within populations	9	26.699	0.11898	6.38	$F_{SC}=0.08$ (0.0001)
	Within subpopulations	156	227.791	1.46020	78.30	$F_{ST}=0.22$ (<0.0001)
Northeast-East-Island	Among populations	5	21.531	0.17340	7.57	$F_{ST}=0.08$ (<0.0001)
	Within subpopulations	72	152.418	2.11692	92.43	
Central	Among populations	1	1.792	0.00525	0.30	$F_{ST}=0.003$ (0.41)
	Within subpopulations	15	26.208	1.74722	99.70	
South Western	Among populations	1	0.852	0.01316	1.97	$F_{ST}=0.02$ (0.24)
	Within subpopulations	28	18.348	0.65529	98.03	
North Western	Among populations	2	2.524	0.03725	4.72	$F_{ST}=0.05$ (0.06)
	Within subpopulations	41	30.817	0.75163	95.28	