

RESEARCH ARTICLE

Open Access

# Long term persistence of clonal malaria parasite *Plasmodium falciparum* lineages in the Colombian Pacific region

Diego F Echeverry<sup>1,2\*</sup>, Shalini Nair<sup>3</sup>, Lyda Osorio<sup>2,4</sup>, Sanjay Menon<sup>5</sup>, Claribel Murillo<sup>2</sup> and Tim JC Anderson<sup>3</sup>

## Abstract

**Background:** Resistance to chloroquine and antifolate drugs has evolved independently in South America, suggesting that genotype - phenotype studies aimed at understanding the genetic basis of resistance to these and other drugs should be conducted in this continent. This research was conducted to better understand the population structure of Colombian *Plasmodium falciparum* in preparation for such studies.

**Results:** A set of 384 SNPs were genotyped in blood spot DNA samples from 447 *P. falciparum* infected subjects collected over a ten year period from four provinces of the Colombian Pacific coast to evaluate clonality, population structure and linkage disequilibrium (LD). Most infections (81%) contained a single predominant clone. These clustered into 136 multilocus genotypes (MLGs), with 32% of MLGs recovered from multiple (2 – 28) independent subjects. We observed extremely low genotypic richness ( $R = 0.42$ ) and long persistence of MLGs through time (median = 537 days, range = 1 – 2,997 days). There was a high probability (>5%) of sampling parasites from the same MLG in different subjects within 28 days, suggesting caution is needed when using genotyping methods to assess treatment success in clinical drug trials. Panmixia was rejected as four well differentiated subpopulations ( $F_{ST} = 0.084 - 0.279$ ) were identified. These occurred sympatrically but varied in frequency within the four provinces. Linkage disequilibrium (LD) decayed more rapidly ( $r^2 = 0.17$  for markers <10 kb apart) than observed previously in South American samples.

**Conclusions:** We conclude that Colombian populations have several advantages for association studies, because multiple clone infections are uncommon and LD decays over the scale of one or a few genes. However, the extensive population structure and low genotype richness will need to be accounted for when designing and analyzing association studies.

**Keywords:** *Plasmodium falciparum*, Colombia, Clonality, Relatedness, Persistence, Genotypic richness, Population structure, SNPs, Linkage disequilibrium, Association studies

## Background

Studies of parasite genetic structure are of practical importance in low transmission malaria regions for a number of reasons. First, long term persistence of identical multilocus genotypes (MLGs) may generate an upward bias in PCR-corrected estimates of treatment failure rates and genetic data can be used to estimate the size of this

bias. Second, the longer range of linkage disequilibrium (LD) expected in such populations can simplify genetic mapping since genomic regions influencing traits of interest may be detected with a lower density of markers. At the same time, however, population structure may introduce bias into such analyses. Therefore, detailed evaluations of both LD and population structure are needed to design appropriate association studies in such regions. Third, while low malaria transmission areas are responsible for a small proportion of the world's malaria cases, such regions appear to play a disproportionate role in evolution of drug resistance. For example, resistance to chloroquine (CQ) has evolved twice in South America,

\* Correspondence: [dechever@purdue.edu](mailto:dechever@purdue.edu)

<sup>1</sup>Department of Entomology, Purdue University, 901 West State Street, West Lafayette, IN 47907, USA

<sup>2</sup>International Center for Medical Research and Training, CIDEIM, Carrera 125 # 19-225 Av. La María, Cali, Colombia

Full list of author information is available at the end of the article

with additional origins in Southeast Asia and Papua New Guinea [1]. However, there is no evidence that CQ resistance has arisen in Africa, where ~90% of the world malaria cases occur, and a similar story also holds for resistance to pyrimethamine [2]. Finally, effective control measures have reduced malaria transmission in many hyperendemic regions of sub-Saharan Africa. Over time we expect that malaria parasite population structure in these regions may become more similar to that currently observed in Southeast Asia and South America.

Malaria is endemic in 20 countries in Central and South America, with Brazil and Colombia accounting for ~65% of all reported cases [3]. Colombia has approximately 100,000 malaria cases per year and >60% of the Colombian population is at risk for malaria. Malaria transmission is unstable with all age groups affected, and nearly all malaria cases are symptomatic [4]. Using genotyping of the antigenic genes merozoite surface protein (*msp*) 1 and 2, parasite populations from Quibdó city (Chocó, Colombia) were shown to be genetically depauperate with low levels of multiclonal infections [5]. In fact, a single haplotype accounted for more than 60% of all the 390 samples studied. A subset of 56 of these samples were further studied with five microsatellites confirming very limited variation: 13 different haplotypes were found among these 56 isolates [5]. This low genetic diversity is consistent with other molecular study performed in Panguí (Chocó) [6], Turbo and Zaragoza (Antioquia, Colombia) [7] and with studies performed in South America, but contrasts with similar studies in Africa [8-11].

Most work on South American *P. falciparum* has employed antigen or microsatellite markers. As data from these markers are difficult to compare between laboratories, we sought to use methods that are more portable. Here, we examined the population genetics and population structure of *P. falciparum* infections using 384 SNPs in *P. falciparum* sampled from the Pacific region of Colombia between 1993–2007. We used these data to: (1) test methods for genome wide SNP typing using limited parasite DNA from dried bloodspots; (2) determine the number of parasite populations present within the region sampled; (3) measure the persistence of identical multilocus genotypes in time and space, and (4) examine patterns of LD across the genome. Our central goals are to determine how parasite population structure impacts interpretation of clinical drug trials and the design of genetic association analyses in this region of Colombia.

## Methods

### Parasite samples

Microscopically confirmed *P. falciparum* samples included in this study came from five cities of the Colombian Pacific region: Tadó and Quibdó in the province of Chocó,

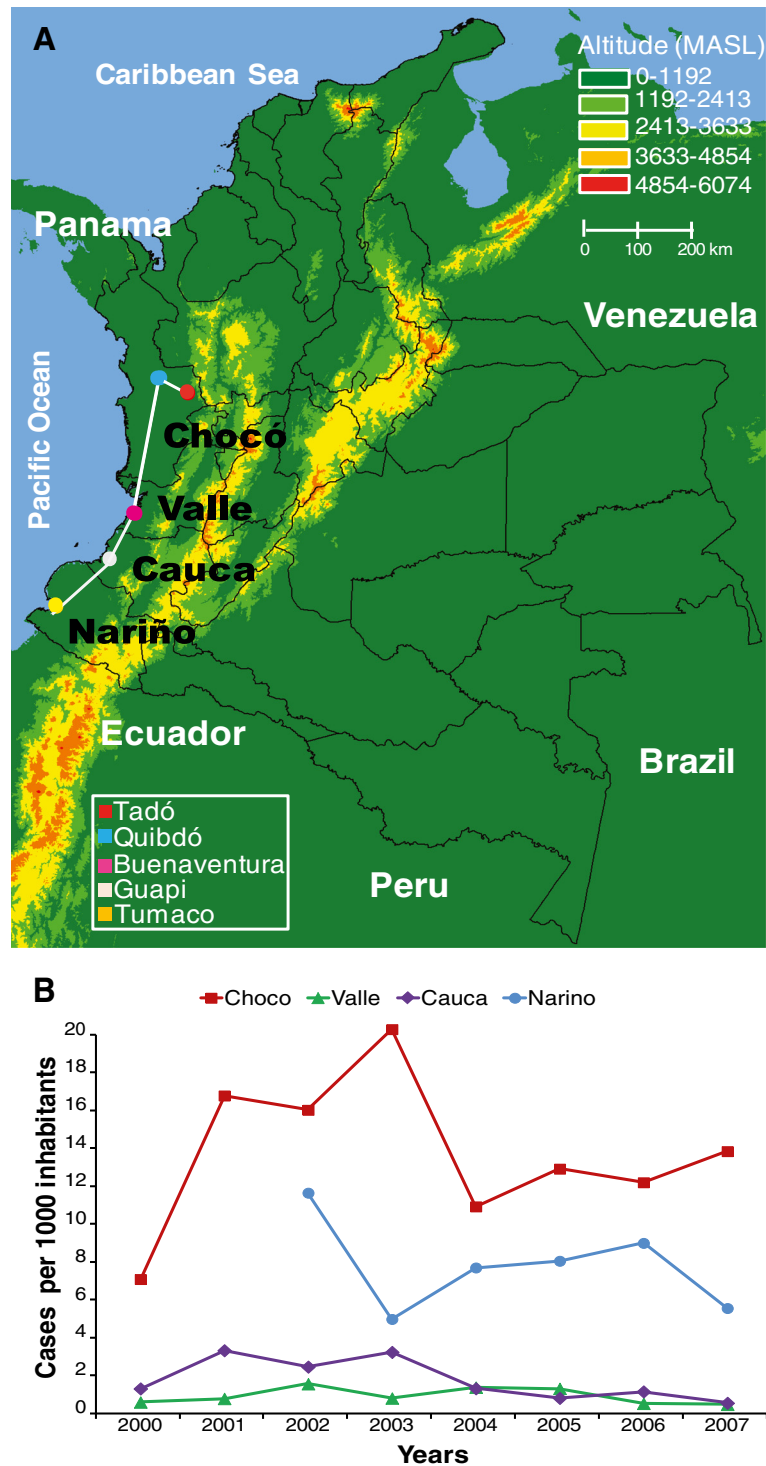
Buenaventura in Valle, Guapi in Cauca, and Tumaco in Nariño (Figure 1A, Table 1). Collectively, these four provinces account for up to 75% of *P. falciparum* cases reported in Colombia [4]. Strong differences in the transmission of falciparum malaria are observed among these locations (Figure 1B and Table 1). Finger-prick blood spot samples were collected on filter paper (Whatman 3 MM, Whatman International, Maidstone, England) from 447 Colombian *P. falciparum* samples. Samples were obtained from primary infections of subjects with uncomplicated malaria who took part in studies conducted by CIDEIM from 1993 to 2007. Ninety five percent were obtained from fresh blood and 5% from samples previously adapted to culture. Informed consent was obtained from all the subjects enrolled, as approved by CIDEIM Institutional Review Board (IRB). We included nine *P. falciparum* reference strains: 3D7 (collected in the Netherlands), 7G8 (Brazil), Dd2 (Indochina/Laos), FCB (Southeast Asia), FCC2 (China), HB3 (Honduras), K1 (Thailand), Santa Lucia (Guatemala) and V1/S (Vietnam) from the Malaria Research and Reference Reagent Resource Center - MR4 (<http://www.mr4.org>), as controls for the genotyping methodology.

### DNA extraction and whole genome amplification (WGA)

For each sample, four or six mm punches of the blood spot were used for DNA recovery and purification. We used a three step process to prepare DNA for SNP genotyping: (1) DNA was recovered from blood spots using the GenSolve kit (GenVault Corporation, Carlsbad, CA); (2) the QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA) was used to purify recovered DNA and concentrated using a speed vacuum (20 min at 45°C) to achieve a final volume ~10 µL, and (3) the illustra GenomiPhi V2 DNA amplification kit (GE Healthcare, Piscataway, NJ) was used to amplify 1 µL of gDNA. The final volume of DNA was 60 µL (eluted in TE buffer) and DNA was quantified using a NanoDrop 1,000 spectrophotometer (Thermo Fisher Scientific Inc, Wilmington, DL).

### GoldenGate SNP genotyping

A custom GoldenGate<sup>®</sup> genotyping assay was designed using 384 SNPs obtained from coding genes. These included 126 synonymous and 258 non-synonymous SNPs (Additional file 1). The SNPs were selected using the query system on PlasmoDB ([www.plasmodb.org](http://www.plasmodb.org)) [15]. We selected SNPs that were polymorphic (with a minor allele observed in at least two Central and South American samples), situated >50 kb from telomeres to exclude antigenic genes, that were assigned a score of >0.6 using the Illumina Design Tool (ILLUMINA<sup>®</sup> ADT, Illumina, San Diego, CA). The SNPs were genotyped using the GoldenGate<sup>®</sup> genotyping assay [16], following Illumina protocols with 50 ng of starting DNA in 5 µL



**Figure 1 Sampling of *Plasmodium falciparum* populations and malaria epidemiology in the Colombian Pacific region. A)** Map of Colombia - South America showing four malaria endemic provinces located in the Pacific region: Chocó, Valle, Cauca and Nariño, and the cities where samples were collected. About 90% of the population in this region located between the Pacific Ocean and the Western branch of the Andean Mountains are Afro-Colombians. Lines represent distance between sample collection sites: Quibdó-Tadó: ~66 km; Quibdó-Buenaventura: ~253 km; Buenaventura-Guapi: ~149 km and Guapi-Tumaco: ~109 km. **B)** Malaria endemicity in the four states defined by the annual parasite incidence (API) between 2000 – 2007. Data was not available from 2000 – 2001 in Nariño. For complementary information see Table 1.

**Table 1 Summary of *Plasmodium falciparum* samples, epidemiology data and malaria vectors in the Colombian Pacific region**

State City	Location	Inhabitants (% per state) <sup>a</sup>	Multiclonal infection/ primary cases	Number of monoclonal samples	Competent vectors in the state <sup>b</sup>	Malaria falciparum cases per state per year <sup>c</sup>	Number of monoclonal samples collected per state per year <sup>d</sup>	
<b>Chocó</b> Quibdó Tadó	5°41'19"N- 76°39'33"W	388476 109121(28%)	25/102 (25%)	77	<u><i>An. darlingi</i></u> *	2000: 2616	1997: 2	
	5°16'03"N- 76°33'57"W	15962(4%)				<u><i>An. nuneztovari</i></u>	2001: 6383	2000: 6
						<i>An.pseudopuntipennis</i> *	2002: 6089	2001: 13
						<i>An. neivai</i> *	2003: 7745	2002: 1
						<i>An. puntimacula</i> *	2004: 3863	2004: 14
							2005: 4441	2005: 6
							2006: 4320	2006: 13
							2007: 5025	2007: 22
<b>Valle</b> Buenaventura	3°25'52"N- 77°03'17"W	4161125 324207 (8%)	8/57 (14%)	49	<u><i>An. albimanus</i></u> , <u><i>An. nuneztovari</i></u> <i>An.pseudopuntipennis</i> * <i>An. neivai</i> * <i>An. puntimacula</i> *	2000: 985	1993: 4	
						2001: 1767	1994: 1	
						2002: 5003	1999: 4	
						2003: 1859	2000: 1	
						2004: 4369	2002: 1	
						2005: 4115	2004: 13	
						2006: 745	2005: 15	
		2007: 643	2006: 10					
<b>Cauca</b> Guapi	2°34'17"N- 77°53'15"W	1268937 28649 (2%)	11/78 (14%)	67	<u><i>An. albimanus</i></u> <u><i>An. nuneztovari</i></u> <i>An. neivai</i> * <i>An. puntimacula</i> *	2000: 1209		
						2001: 3770		
						2002: 2670	1999: 1	
						2003: 3657	2000: 1	
						2004: 1177	2003: 65	
						2005: 466		
						2006: 948		
		2007: 198						
<b>Nariño</b> Tumaco	1°48'34"N- 78°45'46"W	1541956 161490 (11%)	31/163 (19%)	132	<u><i>An. albimanus</i></u> <i>An. neivai</i> * <i>An. puntimacula</i> *	2000: ND	1999: 1	
						2001: ND	2000: 2	
						2002: 17417	2001: 10	
						2003: 7085	2002: 12	
						2004: 6250	2003: 56	
						2005: 9992	2004: 2	
						2006: 9245	2005: 23	
						2007: 7994	2006: 1	
			2007: 25					
<b>Total</b>			<b>75/400 (18.75%)</b>	<b>325</b>				

<sup>a</sup>Information according to the 2005 Colombian National Census [12]; <sup>b</sup>Reports are based on field studies, underlined species are considered primary vectors in Colombia [13,14]; <sup>c</sup>Information provided by SIMIGILA – Instituto Nacional de Salud (INS) of Colombia; <sup>d</sup>Include only the 325 monoclonal samples analyzed; ND: Not Determined; \*Complexes of species.

of water for each sample. Clustering was done using the BeadStudio package (Illumina, San Diego, CA). We defined clusters of parasites with alternative bases (wild-type or mutant) at each SNP: parasites falling in between these two extremes were assumed to be mixed infections.

#### Clonality assessment, relatedness and persistence

We calculated the proportion of alleles shared (*ps*) between all pairwise comparisons of all single clone parasites following the procedure of Anderson et al. [17,18]. The genetic distance matrix was estimated using

ARLEQUIN v 3.1 software [19], a UPGMA phenogram was constructed using the metric  $1 - p_s$  with PHYLIP v 3.69 [20], and plotted using FIGTREE software v 1.3.1 [21]. Parasite samples identical at all genotyped SNPs (i.e. clones) were identified by inspecting the terminal branches of the phenogram. Parasites with unique haplotypes at all tested SNPs are referred as multilocus genotypes (MLGs). We calculated genotypic richness (R), a measure of the proportion of unique genotypes in the population sample as  $R = G/n$ , where  $G$  is the total number of MLGs found in  $n$  samples [22]. We computed the probability of sampling the same MLG from different infected patients separated by different time intervals within each province. This analysis was done using the clonal sub-range fix option included in the GenClone 2.0 software [23,24], with time intervals relevant for *in vivo* antimalarial efficacy studies.

#### Population structure assessment

This analysis was conducted using one representative of each MLG after exclusion of multiple infections. We used the Bayesian model-based clustering algorithm implemented in STRUCTURE v 2.1 [25,26] to investigate genetic structure. We used an admixture model with correlated allele frequencies [26], with a Markov Chain Monte Carlo (MCMC) of 10,000 'burn-in' steps followed by 100,000 interactions and 20 replicate runs at each  $K$  value (1 to 16). The optimal  $K$  partition was estimated following the STRUCTURE directions and the methodology proposed by Evanno et al. [27].

#### Estimation of linkage disequilibrium

HAPLOVIEW software was used to compute linkage disequilibrium (LD) [28]. We calculated LD between each pairwise combination of linked markers and plotted the relationship between physical distances and LD. To measure the extent of LD, we used the correlation coefficient ( $r^2$ ). Analyses were completed for both in the complete set of unique haplotypes and for the subpopulations identified by STRUCTURE software. We estimated the level of LD between unlinked markers (SNPs in different chromosomes for each subpopulation) in order to estimate the background LD caused by small sample size in the populations, or relatedness, as was performed by Van Tyne et al. [29].

## Results

#### Efficient SNP genotyping using Goldengate<sup>®</sup> assay from blood spots

Parasitemias of genotyped samples ranged from <0.01% (including two samples with 80 parasites/ $\mu$ L) to 1.09% of infected red blood cells. Following multiple displacement amplification of DNA extracted from blood spots [30], we obtained an average of  $\sim$ 12  $\mu$ g of DNA per sample

with an OD<sub>260</sub>:OD<sub>280</sub>  $\sim$ 1.8 although most of the amplified DNA came from human origin. All 17 samples adapted to culture were successfully genotyped.

A total of 384 SNPs distributed across the 14 chromosomes of the parasite were genotyped (Figure 2, Additional file 1), with an average of one SNP per 60 kb. Forty three SNPs (11%) were rejected due to poor quality (poorly resolved or scored in <90% samples) and 34 (9%) were non variable SNPs (minor allele frequency - MAF <5%), leaving 307 (80%) informative SNPs with MAF >5% (Additional file 1). We genotyped 447 primary infections and nine reference strains. Thirty two samples (7%) with parasitemias between 150 and 18,200 parasites/ $\mu$ L had incomplete SNP data (>10% of missing data) and 15 samples were excluded due to conflict with identification. A total of 400/447 (89.5%) Colombian Pacific coast primary infections were included in the statistical analyses (Table 1).

The reference strains DD2, HB3, 7G8 and Santa Lucia were evaluated twice in different plates. The validation rate (concordance level) between the same reference strains in different plates to the total SNP calls was high, 99.1, 99.4, 99.7 and 96%, respectively (Additional file 2). We also compared concordance between GoldenGate SNP calls and those in the genome sequence data available at PlasmoDB. Several discrepancies were found. For FCB, DD2, 3D7, HB3 and 7G8, we observed high concordance (between 95.1 to 99.4%), while for K1, V1/S, FCC2 and Santa Lucia concordance was lower (88.1 to 92.6%) (Additional file 2). As the GoldenGate<sup>®</sup> genotyping is a robust technique, SNP discordances may be explained by low sequencing coverage (1.25X) of these samples [29].

#### Clonality assessment reveals strong relatedness and long term persistence

Polyclonal infections (>1 clone of *P. falciparum* per sample) were defined as the samples with >10 heterozygous SNP calls and were found in 75/400 (19%) samples (Table 1). For those samples, the mean number of heterozygous SNPs was 27.7 (range 11 – 66). We observed 25, 14, 14, and 19% multiple infections in Chocó, Valle, Cauca and Nariño provinces respectively. There was a strong positive correlation between carriage of multiple genotypes and transmission intensity ( $r^2 = 0.99$ ) (Figure 3).

A subset of 325 monoclonal samples from Chocó, Valle, Cauca and Nariño were retained for further analysis (Table 1). These samples had 250 informative SNPs, as several SNPs were not informative (MAF <5%) after exclusion of polyclonal infections. The 325 monoclonal parasites comprised 136 unique MLGs, with 44 (32%) represented by MLGs infecting more than one patient (range 2 – 28) (Figure 4A). MLG 036 comprises two culture-adapted samples from Quibdó and Tadó that were indistinguishable from the Dd2 reference strain

from Southeast Asia. Contamination during *in vitro* adaptation (W2, the Dd2 progenitor, is cultured at CIDEIM since 2000) or DNA manipulation may explain this observation.

The sampling period (03/15/1999 to 07/09/2007) covers >3000 days. Using the whole dataset, we observed 45 MLGs that persisted for more than one day (Figure 5A). We examined the clonal persistence in time and space of the 15 most common MLGs (range of 5 – 28 clones per MLG) (Figure 5B). Eight MLGs (n = 67 parasites) were restricted to a single province, while seven MLGs (n = 92 parasites) were recovered from more than one province. A surprisingly long persistence was found with a median of 537.5 (range 1 – 2997) days. Thus, our data suggest that parasite clones persist through the Pacific region for up to eight years (Figures 5A and 5B). This is equivalent to 48 parasite generations assuming six generations per year [31].

In the four locations, the probability of sampling the same MLG decayed slowly and remained between 0.03 and 0.11 at time intervals <500 days (Figure 5C). In Valle, probabilities of sampling identical MLGs were closed to 30% for patients sampled within the first 21 days.

#### Non panmitic population in *P. falciparum* samples from Colombian Pacific region

From the STRUCTURE input file, the mean Var [LnP(D)] showed an unimodal distribution reaching a plateau at  $K=4$  (Figure 4B) and a relatively constant  $\alpha$  value after  $K=4$ , stabilizing at <0.1, suggesting four different subpopulations (Figure 4C). This concurs with Evanno's methodology [27] using the  $K$  vs.  $\Delta K$  (Additional file 3). These findings reject panmixia in the Colombian parasites analyzed. The four subpopulations exhibit significant differentiation with  $F_{ST}$  values ranging from 0.084 - 0.279 (Additional file 4). The subpopulation Col-1 is represented by 82 isolates in 46 MLGs, and were highly prevalent in the North Pacific region (Chocó province) with a frequency >80% decreasing dramatically to the South; in contrast, subpopulation Col-4 with 87 isolates in 21 MLGs was highly prevalent in Cauca (>70%) and its frequency was lower to the Pacific North. Subpopulations Col-2 with 86 isolates in 55 MLGs and Col-3 (70 isolates in 14 MLGs) showed a more heterogeneous distribution, comparable between Valle and Nariño, although Col-2 and Col-3 had the higher frequency (>40%), respectively (Figure 4D); this suggests a high chance of parasite admixture in these provinces. As expected, the lowest differentiation between the provinces was found between Valle - Nariño ( $F_{ST}=0.023$ ) while the highest was for Cauca-Chocó ( $F_{ST}=0.117$ ) (Additional file 4). The UPGMA tree based on allele sharing distances showed strong support for the four clusters (Figure 4A).

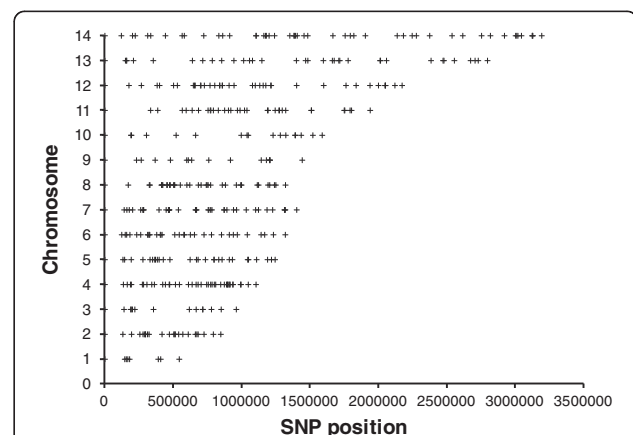
#### Extent of linkage disequilibrium

We computed LD for the entire population and for the four subpopulations of parasites identified by STRUCTURE (Figure 6). In the entire population of unique haplotypes (n = 136), the mean  $r^2$  was 0.16 for markers spaced <10 kb apart and decayed to background levels within ~240 kb. We observed striking differences in the decay of LD values in the four subpopulations (Figure 6), decaying slower in subpopulation Col-3 and Col-4. Within subpopulations Col-3 (n = 14) and Col-4 (n = 21) mean  $r^2$  was ~0.5 at inter-marker distances <20 kb, complete pairwise LD ( $r^2 = 1$ ) was observed for many marker pairs at distances up to ~1200 kb, the mean  $r^2$  decayed to background levels in ~500 kb. In addition, Col-3 and Col-4 showed high background levels of LD ( $r^2 = 0.11$  and 0.08 respectively) between markers on different chromosomes.

The subpopulation Col-2 (n = 55) exhibited the most rapid decay in LD, and patterns of decay for this subpopulation mirrored those observed in the whole population (Figure 6).

#### Discussion

Successful genotyping of 307 SNPs in 400 *P. falciparum* infections from four provinces in the Colombian Pacific region over a ten year period allows detailed description of parasite population genetics in this region. These data reveal (a) parasite MLGs persisting for up to eight years (median of 538 days), (b) stratification of parasites into four subpopulations that occur sympatrically within sampling locations, (c) LD decays by half in <10 kb, but varies between subpopulations. We discuss the advantages of the SNP genotyping method used, and the implications of our findings for design of association studies and evolution of drug resistance in low endemic malaria areas.



**Figure 2 Distribution of Single Nucleotide Polymorphisms (SNPs) used for genotyping.** Location of 384 SNPs through the 14 chromosomes of *P. falciparum*. There were 6, 22, 12, 39, 27, 30, 36, 34, 14, 15, 31, 37, 33, and 48 SNPs on chromosomes 1 – 14, respectively.

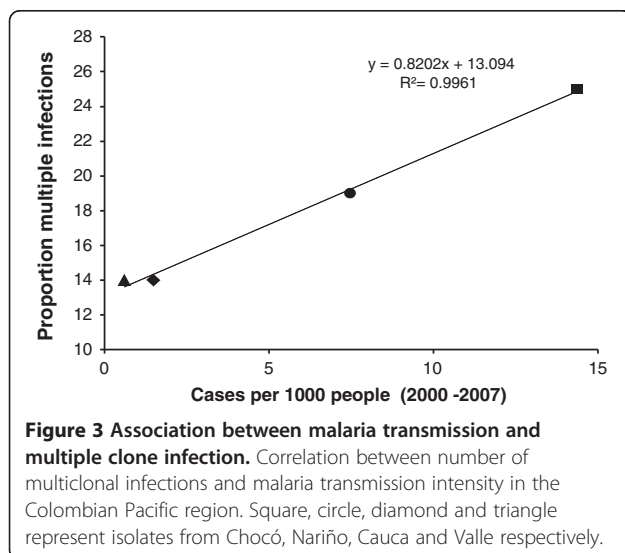
### Effective genotyping using dried blood spots

In the present study, we were able to rapidly score multiple markers from finger prick blood spots with high reproducibility. The SNP markers selected provide a set of 250 informative SNPs for further genetic studies at the local/regional level. The major advantage of SNPs over microsatellites is that they are more abundant, mutationally stable, located in genes and “portable”; in other words, they are easily scored and comparable between studies.

### Clonality and persistence of MLGs in time and space

Genetic variability studies show a direct relationship between degree of parasite endemicity and genetic variation [32,33]. In high endemic malaria areas, such as sub Saharan Africa, multiclonal *P. falciparum* infections, high genetic diversity, and low LD are common. In contrast, in low endemic areas, such as South American countries, malaria patients are expected to have infections caused by a single clone with limited genetic diversity and more extensive LD [8]. We found that 19% of *P. falciparum* infections were multiclonal in samples collected from the Colombian Pacific region. This is consistent with previous studies in low endemic malaria areas and contrasts with studies in Africa where the percentage of multiclonal infections can reach up to 90% [34]. This correlation between multiclonal infections and transmission intensity was confirmed here within low endemic areas (Figure 3), suggesting that this metric provides an indirect measure of transmission intensity [35,36].

A decade ago, Anderson et al., using 12 microsatellites, stated that *P. falciparum* from South America had the lower level of genetic diversity worldwide, with 30 Colombian samples (collected in Antioquia province),



**Figure 3 Association between malaria transmission and multiple clone infection.** Correlation between number of multiclonal infections and malaria transmission intensity in the Colombian Pacific region. Square, circle, diamond and triangle represent isolates from Chocó, Nariño, Cauca and Valle respectively.

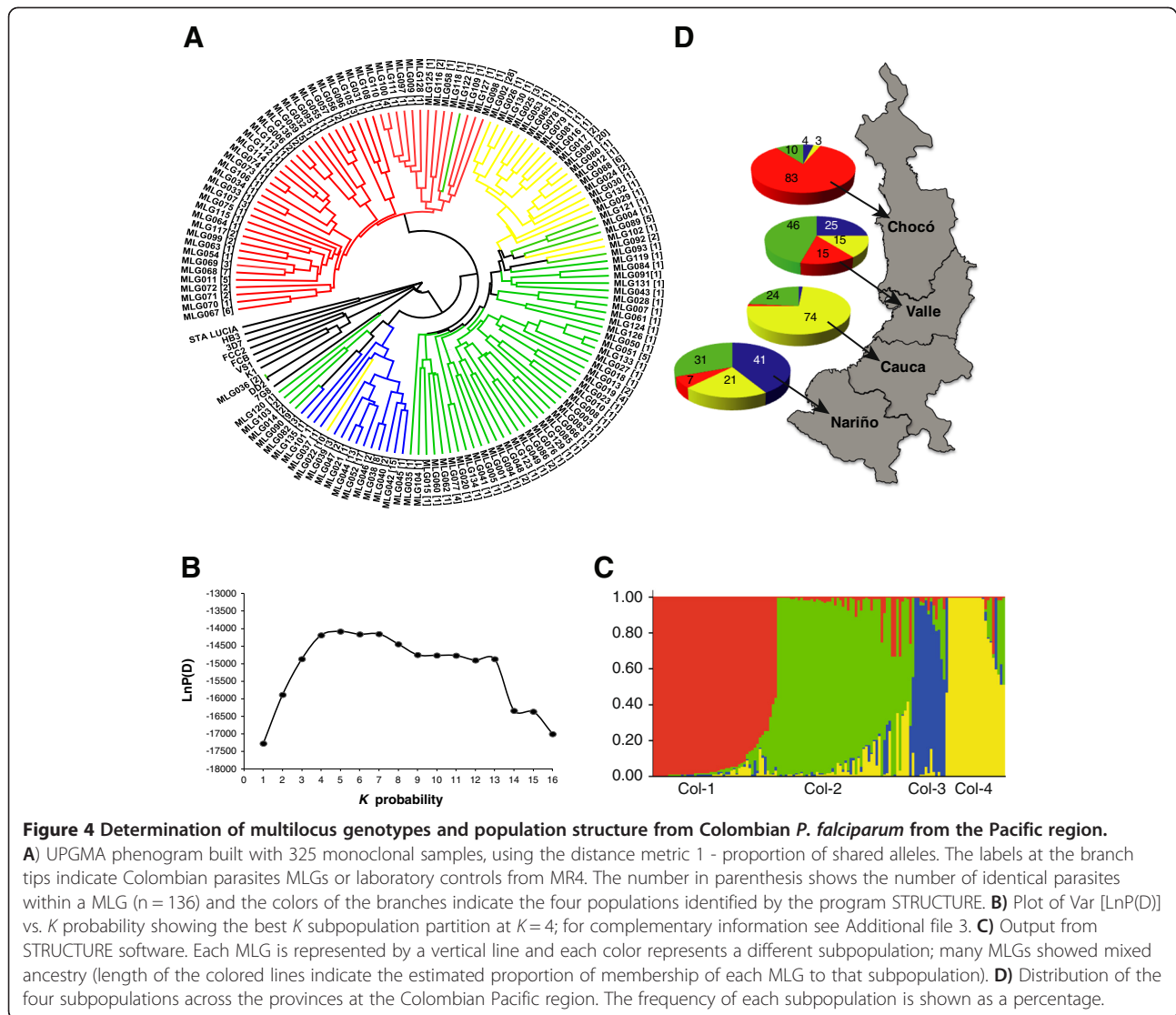
showing the lowest diversity [8]. Another study using 56 samples from Chocó and five polymorphic microsatellites, also suggested low diversity [5].

The genotypic richness (R) [22] was 0.42 for all the monoclonal samples included in this study, the lowest reported in comparison with other studies from similar malaria eco-epidemiological features such as Venezuela, Peru, Brazil, Cambodia and Thailand with R values of 0.60 – 0.98 [9,11,17,18,37-39]; however this measure is strongly influenced by sampling intensity and hence comparisons between countries may be biased [35]. Our study confirmed low genotype richness in *P. falciparum* from Colombia, with a third of the MLGs infecting  $\geq 2$  patients (Figure 4A) and long persistence (Figure 5A) in cities separated up to more than 500 km (Figure 1). Our results contrast with studies from neighboring countries including Venezuela, Brazil and Peru, where genotype richness is markedly higher and the number of polyclonal infections has increased (more than double) from 2003 to 2007 [9,38,40].

### Implications for drug efficacy studies

PCR genotyping of parasite infections before and after treatment is widely used to differentiate between reinfection and recrudescence and to adjust measures of treatment failure rates in antimalarial drug efficacy studies [41,42]. However when parasite populations are highly inbred, there is high probability of patients being reinfected with the same parasite genotype [43]. To evaluate this probability in Colombian samples, we examined the probability of sampling identical genotypes. In Valle province, this probability was the highest, during the interval of 15–21 days (~30%), followed by Chocó (29–42 days) and Nariño (22–28 days) with a probabilities close to 15% and Cauca (43–63 days) ~10%. The overall population mean probabilities were between 3 - 11% for up to 500 days (Figure 5C). These results suggest that PCR evaluation should be used with care in Colombia, because there is a strong probability of misclassifying some new parasite infections as recrudescences, thereby overestimating treatment failure rates (Figure 5C).

Colombia implemented Artemisinin Combination Therapies (ACTs) at the end of 2006. Three drug efficacy studies were performed with these compounds in Antioquia and Chocó provinces, showing 99 to 100% of efficacy [44-47]. One subject from the rural area of Tadó (in Chocó) presented parasitemia and fever at day 28 post treatment with Coartem®; further genetic analyses of the *msh1* gene suggested a recrudescence [47]. However, it is unclear whether this was a true treatment failure, or a case of reinfection with the same genotype. The use of even more polymorphic markers is not going to overcome the limitations of PCR in this scenario. Three alternative approaches



could be used to aid interpretation of such studies: i) the use of statistical approaches designed to account for this bias, ii) definition of primary efficacy in terms of parasite clearance rates and iii) the use of “malaria-free locations” for malaria patients during post treatment surveillance [43,48].

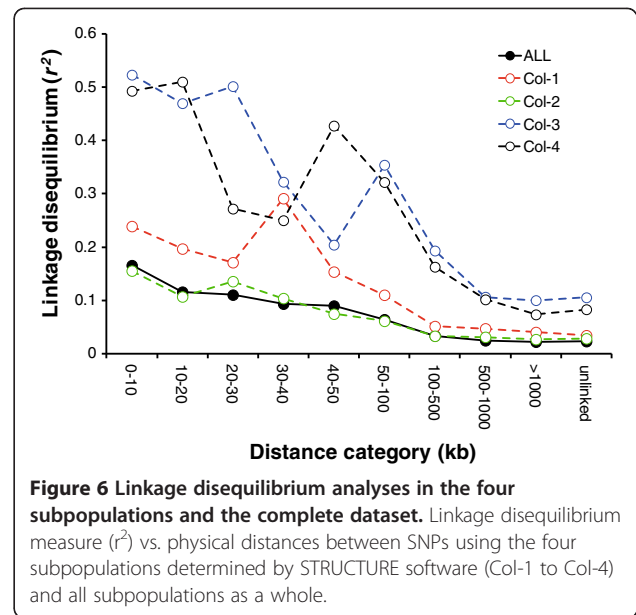
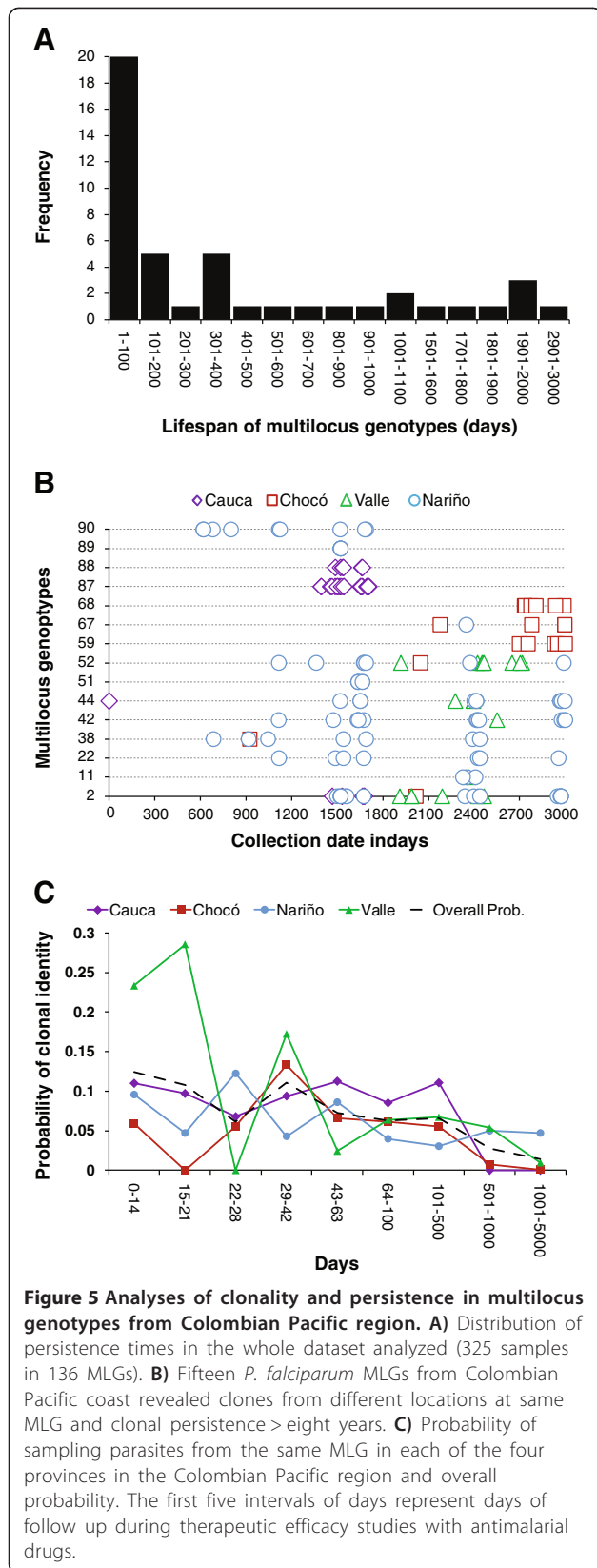
#### Strong genetic structure in the Colombian Pacific

Both allele sharing methods and Bayesian clustering define four subpopulations and mixed ancestry in the area of study (Figure 4), suggesting that our population structure results are robust. This is in line with previous studies performed in Brazil and Peru, where three to five subpopulations were revealed [9,10,39]. The presence of parasite subpopulations in the Colombian Pacific coast may be partially explained by the bottleneck in *Plasmodium* populations, approximately 9,000 cases in 1960, produced

by the implementation of malaria control strategies [49] and subsequent focal reemergence of parasites with different genetic backgrounds.

The coexistence of different sub-populations within locations is consistent with limited genetic exchange, together with extensive migration among locations. *Plasmodium falciparum* genetic interchange in Colombia was suggested recently for parasites through the Andean mountains and North and South of the Pacific region [50]. Identification of identical MLGs in different sampling locations also demonstrates migration of parasite genotypes without breakdown due to recombination. An epidemiological study performed in Quibdó with 670 *P. falciparum* infected patients, revealed that 66% of the cases are from the urban and rural area of the city, while the 33 and 1% are from neighboring municipalities and provinces, respectively [5].





Local adaptation to different vectors may play a role in the parasite population structure. For example, coadaptation between vector and parasite has been suggested for mosquitoes and *Plasmodium vivax* in Mexico [51], where subpopulations of parasites differentially infected *Anopheles albimanus* and *An. pseudopuntipennis*. Three primary and three secondary vectors are found in the Pacific region (Table 1), and at least five different ecological sub-regions had been identified (<http://www.eoearth.org/>). *Anopheles* populations in Colombia vary locally in their vectorial competence, breeding habitats, and feeding preferences [13]. One possible explanation is that parasite population structure in the Pacific region is also shaped by geographic restriction of compatible vectors. For example, parasites from Col-1 subpopulation may be adapted to *An. darlingi* in Chocó, since this vector had not been registered in the other provinces of the Colombian Pacific region [13,14]. Further experimental investigation is necessary to test this hypothesis. The presence of *An. darlingi* in Chocó, could explain the higher number of multiclinal infections (25%), as this species is considered the most effective malaria vector in Latin America [13,14].

The model of metapopulation structure in *P. falciparum* suggests the potential for spreading of drug resistance alleles [52,53]. Parasites from Colombia may follow this model as they show no panmixia and inbreeding. This fact highlights the need to closely monitor the efficacy of ACTs in Colombia and neighboring countries, since emergence of drug resistance to different antimalarials occurred and disseminated rapidly in this region [2]. Artemisinin resistance has been confirmed in Southeast Asia [54], an area with similar low transmission conditions as South

America. Finally, the presence of *P. falciparum* subpopulations in the Colombian Pacific region could explain the different patterns of drug susceptibility (*in vivo* and *in vitro*), as the magnitude of resistance to amodiaquine, sulphadoxine-pyrimethamine, and mefloquine varies between the South and North of this region [50,55].

#### Implications for association studies

Association studies require considerable investment of time and resources. Therefore, it is critical to first demonstrate that the traits of interest have a genetic basis and to quantify the heritability in order to calculate appropriate sample sizes [17,18]. Colombian *P. falciparum* populations are well suited to study the heritability of a trait of interest as the estimation of this parameter is achievable when identical clones in populations have been identified.

Colombian *P. falciparum* samples represent a challenge for association studies owing to strong population structure, and presence of many identical or closely related genotypes. In this study there were 136 unique genotypes among 400 parasites sampled. Hence, only a third of parasites sampled would be informative for association analyses. Both population stratification and cryptic relatedness can generate spurious associations [56]. On the other hand, the low numbers of multiple clone infections simplifies detection of genotype/phenotype associations. Both sampling and statistical approaches can minimize bias in this situation. A two phase sampling strategy provides one possible approach to minimize cost and effort, while maximizing study power. In phase one, preliminary genotyping of the parasite population using 96–384 SNPs can rapidly identify identical clones and multiple clone infections. In phase two, a single representative of each clone can be genotyped using higher densities of SNPs using Illumina sequencing [57] or microarray based approaches [58] and characterized for the trait of interest. From the statistical standpoint, powerful mixed model approaches developed by plant geneticists allow for effective control of both stratification and cryptic relatedness [59]. This methodology was recently used to establish the role of the *PF10\_0355* membrane protein in low susceptibility to arylaminoalcohols antimalarials [29]. These approaches must be considered in areas of low transmission for *P. falciparum* association studies.

Offspring with limited to zero recombination are expected to occur in South America [60]. Subpopulations Col-3 and Col-4 exhibit the lower genotypic richness ( $R$ ) with  $R \sim 0.21$ , leading to higher LD in comparison with the other populations. Subpopulations Col-1 and Col-2 both with  $R$  of  $\sim 0.59$ , are likely older and have more likelihood of experiencing recombination. In our samples we estimated persistence up to 48 generations, which reflects

transmission over many generations of segments of ancestral haplotypes comprising linked markers.

Despite evidence for high levels of inbreeding, the extent of LD observed was lower than observed in other studies of South American populations. We observed mean  $r^2$  value of 0.16 between markers spaced <10 kb apart in the whole data set (Figure 6), while Neafsey et al. 2008 reported mean  $r^2$  of 0.7 for markers spaced <10 kb apart for samples from Brazil [61]. Strong artifactual LD can be generated by combining subpopulations with differing allele frequencies in a single population sample. We therefore expected that LD would be elevated in the total population relative to the individual subpopulations. In fact, we observed the opposite (Figure 6).

The extent of LD varies among the four subpopulations. Mean  $r^2$  for intermarker distances <10 kb range from 0.15 for Col-2 to 0.52 in Col-3 (Figure 6). Linkage disequilibrium in Col-3 and Col-4 also decays to background levels ( $r^2$  between markers on different chromosomes) at  $\sim 500$  kb, more gradually than for Col-1 and Col-2 (Figure 6). Mixing with other parasite populations could also explain the rapid decay in LD in the Col-2 subpopulation. The Col-2 subpopulation is dominant in Buenaventura (Valle State) the most important Colombian port in the Pacific Ocean. The extensive movement of people through this port may increase the chance of parasite admixture.

Several factors may contribute to the elevated LD observed in Col-3 and Col-4. First, these subpopulations are small ( $n = 14$  and  $21$  for Col-3 and Col-4 respectively) so the extended LD may be an artifact of low sample size. Three observations are consistent with this. First, background levels of LD (between unlinked markers on different chromosomes) are much higher in Col-3 and Col-4 than in Col-1 and Col-2. Second, random resampling of 14 haplotypes from the total sample of unique genotypes ( $n = 136$ ) increased values of  $r^2$  by an average of 0.06 in each distance category. Third, relatedness or recent admixture may contribute to elevated LD in Col-3 and Col-4. These subpopulations show lower expected heterozygosity ( $H = 0.25$  for Col-3 and  $H = 0.21$  for Col-4) compared with Col-1 ( $H = 0.27$ ) and Col-2 ( $H = 0.34$ ), suggesting that they may contain closely related parasites. Finally, genotypic richness is lower in Col-3 and Col-4 ( $R = 0.19$  and  $0.25$  respectively) compared with Col-1 and Col-2 ( $R = 0.57$  and  $0.61$  respectively). We suggest that observations of differences in LD decay within parasite subpopulations should be viewed with caution unless artifactual effects of sample size and relatedness can be clearly rejected.

A feature of the LD information is important for association mapping. Only 2.4% of markers situated within 10 kb show  $r^2 \geq 0.8$  in the whole dataset. Hence, even in this low transmission region, genome sequencing or efficient genotyping of tagging SNPs will be needed to avoid

false negative associations. On the other hand, the rapid decay in LD should enable localization of causative SNPs to genome regions containing 1–5 genes.

## Conclusion

Our study shows the impact of low genotypic richness, persistence of MLGs and population structure of *P. falciparum* on the establishment, distribution and propagation of MLGs in low endemic malaria areas. These features have important implications for the design of ACT clinical efficacy trials and genotype/phenotype association studies. SNP surveys such as these, using moderate numbers of markers, will be critical for maximizing the power, and minimizing bias, in association studies in similar endemic areas in South America using genome sequencing [57] or high resolution microarray methods [58].

## Additional files

### Additional file 1: Single nucleotide polymorphisms (SNPs) used for genotyping Colombian *P. falciparum* samples from the Pacific region.

Detailed information of the 384 coding SNPs used for the GoldenGate<sup>®</sup> genotyping assay including SNP location, synonymous or non-synonymous SNP status, gene information and summary of the performance of each SNP during the genotyping assay are shown.

**Additional file 2: Validation of *P. falciparum* genotyped SNPs using reference strains (controls) performing comparison between them and against PlasmoDB SNPs data.** Bead Studio package output comparisons between technical replicates for reference strains (Dd2, HB3, 7G8 and Santa Lucia) and comparisons between genotyped reference samples (n = 9) and SNP alleles from PlasmoDB. Alleles expected for each SNP are shown in the SNP/OPA column. Alleles expected for the *P. falciparum* reference genome (3D7 strain) and the major and minor allele are also shown.

### Additional file 3: *K* value determination using the Evanno's approach. $\Delta K$ vs. *K* probability plot following Evanno's approach.

The analysis suggest the best *K* value at *K* = 4, suggesting four subpopulations of *P. falciparum* parasites from the Colombian Pacific region.

### Additional file 4: Pairwise fixation indexes in the Colombian Pacific coast *P. falciparum* samples.

A)  $F_{st}$  values between subpopulations identified by STRUCTURE software and B)  $F_{st}$  values between provinces. The  $F_{st}$  values were computed using the GENALEX software [62].

## Abbreviations

SNP: Single nucleotide polymorphism; LD: Linkage disequilibrium; MLG: Multilocus genotype;  $F_{st}$ : Fixation index; CQ: Chloroquine; Msp: Merozoite surface protein; CIDEIM: International Center for Medical Research and Training; MR4: Research and Reference Reagent Resource Center; WGA: Whole genome amplification; UPGMA: Unweighted pair group method with arithmetic mean; MCMC: Markov chain Montecarlo; MAF: Minor allele frequency; ACT: Artemisinin combinatory therapy.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

DFE: Participated in the design of study. Performed genotyping experiments, data analyses and writing of the manuscript. SN: Participated in the supervision of the study, genotyping experiments, data analyses and critically review the manuscript. LO: Participated in the design of the study and writing of the manuscript. SM: participated in genotyping experiments, data analyses and critically review the manuscript. CM: Participated in data

analyses and critically review the manuscript. TJCA: Participated in the design of study and implementation of the genotyping assays. Participate in the supervision of the study, data analyses and writing of the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

This study was supported by an anonymous Swiss Foundation and the Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología COLCIENCIAS (contract ID 198–2007, 2229-405-20319) to DFE, and by NIH R01 AI048071 and AI075145 to TJCA. The molecular work at TBRI was conducted in facilities constructed with support from Research Facilities Improvement Program Grant C06 RR013556 from the National Center for Research Resources, NIH. We would like to acknowledge to Dr. Scott Jackson (Agronomy Department) and Dr. Catherine Hill (Entomology Department) from Purdue University and the malaria team (Gustavo Diaz, Erika Dorado, Madeline Montenegro and Zuleima Pava) at CIDEIM for valuable comments on the manuscript and logistics of the study. We also acknowledge the health authorities, medical staff and malaria control program workers from Chocó, Valle, Cauca and Nariño for their collaboration in establishing the *P. falciparum* collection samples at CIDEIM and to the malaria infected subjects who participate, for making this study possible.

## Author details

<sup>1</sup>Department of Entomology, Purdue University, 901 West State Street, West Lafayette, IN 47907, USA. <sup>2</sup>International Center for Medical Research and Training, CIDEIM, Carrera 125 # 19-225 Av. La María, Cali, Colombia. <sup>3</sup>Department of Genetics, Texas Biomedical Research Institute, 7620 NW Loop 410, San Antonio, Texas, USA. <sup>4</sup>Grupo de Epidemiología y Salud Poblacional, GESPE, School of Public Health, Faculty of Health, Universidad del Valle, Calle 4B # 36-00, Cali, Colombia. <sup>5</sup>Virginia Commonwealth University, Richmond, VA 23220, USA.

Received: 20 September 2012 Accepted: 21 December 2012

Published: 7 January 2013

## References

1. Wootton J, Feng X, Ferdig M, Cooper R, Mu J, Baruch D, Magill A, Su X: **Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*.** *Nature* 2002, **418**(6895):320–323.
2. Mita T, Tanabe K, Kita K: **Spread and evolution of *Plasmodium falciparum* drug resistance.** *Parasitol Int* 2009, **58**(3):201–209.
3. PAHO/WHO: *Status of malaria in the Americas, 2004: A series of data tables.* Washington, D.C, USA: PAHO/WHO; 2004.
4. Instituto Nacional de Salud. Ministerio de la Protección Social: *Protocolo de Malaria.* Bogotá: Instituto Nacional de Salud; 2007.
5. Osorio L, Todd J, Pearce R, Bradley D: **The role of imported cases in the epidemiology of urban *Plasmodium falciparum* malaria in Quibdó, Colombia.** *Trop Med Int Health* 2007, **12**(3):331–341.
6. Gómez D, Chaparro J, Rubiano C, Rojas MO, Wasserman M: **Genetic diversity of *Plasmodium falciparum* field samples from an isolated Colombian village.** *Am J Trop Med Hyg* 2002, **67**(6):611–616.
7. Montoya L, Maestre A, Carmona J, Lopes D, Do Rosario V, Blair S: ***Plasmodium falciparum*: diversity studies of isolates from two Colombian regions with different endemicity.** *Exp Parasitol* 2003, **104**(1–2):14–19.
8. Anderson T, Haubold B, Williams J, Estrada-Franco J, Richardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurwa S, French N, et al: **Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*.** *Mol Biol Evol* 2000, **17**(10):1467–1482.
9. Orjuela-Sánchez P, Da Silva-Nunes M, Da Silva NS, Scopel KK, Gonçalves RM, Malafrente RS, Ferreira MU: **Population dynamics of genetically diverse *Plasmodium falciparum* lineages: community-based prospective study in rural Amazonia.** *Parasitology* 2009, **136**(10):1097–1105.
10. Griffing SM, Mixson-Hayden T, Sridaran S, Alam MT, McCollum AM, Cabezas C, Marquino Quezada W, Barnwell JW, De Oliveira AM, Lucas C, et al: **South American *Plasmodium falciparum* after the malaria eradication era: clonal population expansion and survival of the fittest hybrids.** *PLoS One* 2011, **6**(9):e23486.
11. Machado RL, Pova MM, Calvosa VS, Ferreira MU, Rossit AR, dos Santos EJ, Conway DJ: **Genetic structure of *Plasmodium falciparum* populations in the Brazilian Amazon region.** *J Infect Dis* 2004, **190**(9):1547–1555.

12. Departamento Administrativo Nacional de Estadística DANE: *Colombia Censo General 2005. Nivel Nacional*. Bogotá DC Colombia: DANE; 2005.
13. Montoya-Lerma J, Solarte YA, Giraldo-Calderón GI, Quiñones ML, Ruiz-López F, Wilkerson RC, González R: **Malaria vector species in Colombia: a review**. *Mem Inst Oswaldo Cruz* 2011, **106**(Suppl 1):223–238.
14. González R, Carrejo N: *Introducción al Estudio Taxonómico de Anopheles de Colombia Claves Taxonómicas y Notas de Distribución, Segunda Edición*. Cali - Colombia: Universidad del Valle Press; 2009.
15. Aurrecochea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS, et al: **PlasmoDB: a functional genomic database for malaria parasites**. *Nucleic Acids Res* 2009, **37**:D539–D543.
16. Shen R, Fan J, Campbell D, Chang W, Chen J, Doucet D, Yeakley J, Bibikova M, Wickham Garcia E, McBride C, et al: **High-throughput SNP genotyping on universal bead arrays**. *Mutat Res* 2005, **573**(1–2):70–82.
17. Anderson TJ, Nair S, Nkhoma S, Williams JT, Imwong M, Yi P, Socheat D, Das D, Chotivanich K, Day NP, et al: **High heritability of malaria parasite clearance rate indicates a genetic basis for artemisinin resistance in western Cambodia**. *J Infect Dis* 2010, **201**(9):1326–1330.
18. Anderson TJ, Williams JT, Nair S, Sudimack D, Barends M, Jaidee A, Price RN, Nosten F: **Inferred relatedness and heritability in malaria parasites**. *Proc Biol Sci* 2010, **277**(1693):2531–2540.
19. Excoffier LGLSS: **Arlequin ver 3.0. An integrated software package for population genetics data analysis**. *Evol Bioinform Online* 2005, **1**:47–50.
20. Felsenstein J: **PHYMLIP (Phylogeny Inference Package) version 3.2**. *Cladistics* 1989, **5**:164–166.
21. Rambaut A, Drummond A: *Fig Tree V1.3.1* [http://tree.bio.ed.ac.uk/software/figtree/]. Edinburgh, United Kingdom: Institute of evolutionary, University of Edinburgh; 2010.
22. Dorken M, Eckert C: **Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae)**. *J Ecol* 2001, **89**(3):339–350.
23. Arnaud-Haond S, Belkhir K: **Genclone: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization**. *Mol Ecol Notes* 2007, **7**:15–17.
24. Harada Y, Kawano S, Iwasa Y: **Probability of clonal identity: inferring the relative success of sexual versus clonal reproduction from spatial genetic patterns**. *J Ecol* 1997, **85**:591–600.
25. Pritchard J, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data**. *Genetics* 2000, **155**(2):945–959.
26. Falush D, Stephens M, Pritchard J: **Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies**. *Genetics* 2003, **164**(4):1567–1587.
27. Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study**. *Mol Ecol* 2005, **14**(8):2611–2620.
28. Barrett J, Fry B, Maller J, Daly M: **Haploview: analysis and visualization of LD and haplotype maps**. *Bioinformatics* 2005, **21**(2):263–265.
29. Van Tyne D, Park DJ, Schaffner SF, Neafsey DE, Angelino E, Cortese JF, Barnes KG, Rosen DM, Lukens AK, Daniels RF, et al: **Identification and functional validation of the novel antimalarial resistance locus PF10\_0355 in *Plasmodium falciparum***. *PLoS Genet* 2011, **7**(4):e1001383.
30. Wang Y, Nair S, Nosten F, Anderson T: **Multiple Displacement Amplification for Malaria Parasite DNA**. *J Parasitol* 2009, **95**(1):253–255.
31. Nair S, Williams JT, Brockman A, Paiphun L, Mayxay M, Newton PN, Guthmann JP, Smithuis FM, Hien TT, White NJ, et al: **A selective sweep driven by pyrimethamine treatment in southeast asian malaria parasites**. *Mol Biol Evol* 2003, **20**(9):1526–1536.
32. Haddad D, Snounou G, Mattei D, Enamorado IG, Figueroa J, Ståhl S, Berzins K: **Limited genetic diversity of *Plasmodium falciparum* in field isolates from Honduras**. *Am J Trop Med Hyg* 1999, **60**(1):30–34.
33. Babiker HA, Ranford-Cartwright LC, Walliker D: **Genetic structure and dynamics of *Plasmodium falciparum* infections in the Kilombero region of Tanzania**. *Trans R Soc Trop Med Hyg* 1999, **93**(Suppl 1):11–14.
34. Bonizzoni M, Afrane Y, Baliraine FN, Amenya DA, Githeko AK, Yan G: **Genetic structure of *Plasmodium falciparum* populations between lowland and highland sites and antimalarial drug resistance in Western Kenya**. *Infect Genet Evol* 2009, **9**(5):806–812.
35. Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGready R, Phyto AP, Nosten F, Anderson TJ: **Population genetic correlates of declining transmission in a human pathogen**. *Mol Ecol* 2012. doi:10.1111/mec.12099.
36. Iwagami M, Rivera P, Villacorte E, Escueta A, Hatabu T, Kawazu S, Hayakawa T, Tanabe K, Kano S: **Genetic diversity and population structure of *Plasmodium falciparum* in the Philippines**. *Malar J* 2009, **8**:96.
37. Pumpaibool T, Arnathau C, Durand P, Kanchanakhon N, Siripoon N, Suegorn A, Sitthi-Amorn C, Renaud F, Harnyuttanakorn P: **Genetic diversity and population structure of *Plasmodium falciparum* in Thailand, a low transmission country**. *Malar J* 2009, **8**:155.
38. Urdaneta L, Lal A, Barnabe C, Oury B, Goldman I, Ayala FJ, Tibayrenc M: **Evidence for clonal propagation in natural isolates of *Plasmodium falciparum* from Venezuela**. *Proc Natl Acad Sci USA* 2001, **98**(12):6725–6729.
39. Branch OH, Sutton PL, Barnes C, Castro JC, Hussin J, Awadalla P, Hijar G: ***Plasmodium falciparum* genetic diversity maintained and amplified over 5 years of a low transmission endemic in the Peruvian Amazon**. *Mol Biol Evol* 2011, **28**(7):1973–1986.
40. Sutton PL, Torres LP, Branch OH: **Sexual recombination is a signature of a persisting malaria epidemic in Peru**. *Malar J* 2011, **10**(1):329.
41. Collins WJ, Greenhouse B, Rosenthal PJ, Dorsey G: **The use of genotyping in antimalarial clinical trials: a systematic review of published studies from 1995–2005**. *Malar J* 2006, **5**:122.
42. Juliano JJ, Taylor SM, Meshnick SR: **Polymerase chain reaction adjustment in antimalarial trials: molecular malarkey?** *J Infect Dis* 2009, **200**(1):5–7.
43. Juliano JJ, Gadalla N, Sutherland CJ, Meshnick SR: **The perils of PCR: can we accurately 'correct' antimalarial trials?** *Trends Parasitol* 2010, **26**(3):119–124.
44. Osorio L, Gonzalez I, Olliaro P, Taylor W: **Artemisinin-based combination therapy for uncomplicated *Plasmodium falciparum* malaria in Colombia**. *Malar J* 2007, **6**:25.
45. Alvarez G, Tobón A, Piñeros J, Ríos A, Blair S: **Dynamics of *Plasmodium falciparum* Parasitemia Regarding Combined Treatment Regimens for Acute Uncomplicated Malaria, Antioquia, Colombia**. *Am J Trop Med Hyg* 2010, **83**(1):90–96.
46. Vásquez A, Sanín F, Alvarez L, Tobón A, Ríos A, Blair S: **Therapeutic efficacy of a regimen of artesunate-mefloquine-primaquine treatment for *Plasmodium falciparum* malaria and treatment effects on gametocytic development**. *Biomedica* 2009, **29**(2):307–319.
47. Rojas-Alvarez DP: *Evaluación de la eficacia terapéutica y la tolerabilidad de las combinaciones fijas de Artesunato/Amodiaquina y Artemeter/Lumefantrina para el tratamiento de la malaria por Plasmodium falciparum no complicada en el departamento del Chocó (Colombia)*. Bogotá: Universidad Nacional; 2010.
48. Juliano JJ, Ariey F, Sem R, Tangpukdee N, Krudsood S, Olson C, Looareesuwan S, Rogers WO, Wongsrichanalai C, Meshnick SR: **Misclassification of drug failure in *Plasmodium falciparum* clinical trials in southeast Asia**. *J Infect Dis* 2009, **200**(4):624–628.
49. Padilla JC, Alvarez G, Montoya R, Chaparro P, Herrera S: **Epidemiology and control of malaria in Colombia**. *Mem Inst Oswaldo Cruz* 2011, **106**(Suppl 1):114–122.
50. Corredor V, Murillo C, Echeverry DF, Benavides J, Pearce RJ, Roper C, Guerra AP, Osorio L: **Origin and dissemination across the Colombian Andes mountain range of sulfadoxine-pyrimethamine resistance in *Plasmodium falciparum***. *Antimicrob Agents Chemother* 2010, **54**(8):3121–3125.
51. Joy D, Gonzalez-Ceron L, Carlton J, Gueye A, Fay M, McCutchan T, Su X: **Local adaptation and vector-mediated population structure in *Plasmodium vivax* malaria**. *Mol Biol Evol* 2008, **25**(6):1245–1252.
52. Ariey F, Duchemin JB, Robert V: **Metapopulation concepts applied to falciparum malaria and their impacts on the emergence and spread of chloroquine resistance**. *Infect Genet Evol* 2003, **2**(3):185–192.
53. Dye C, Williams BG: **Multigenic drug resistance among inbred malaria parasites**. *Proc Biol Sci* 1997, **264**(1378):61–67.
54. Dondorp AM, Nosten F, Yi P, Das D, Phyto AP, Tarning J, Lwin KM, Ariey F, Hanpithakpong W, Lee SJ, et al: **Artemisinin resistance in *Plasmodium falciparum* malaria**. *N Engl J Med* 2009, **361**(5):455–467.
55. Aponte SL, Díaz G, Pava Z, Echeverry DF, Ibarquén D, Ríos M, Murcia LM, Quelal C, Murillo C, Gil P, et al: **Sentinel network for monitoring *in vitro* susceptibility of *Plasmodium falciparum* to antimalarial drugs in Colombia: a proof of concept**. *Mem Inst Oswaldo Cruz* 2011, **106**(Suppl 1):123–129.
56. Balding DJ: **A tutorial on statistical methods for population association studies**. *Nat Rev Genet* 2006, **7**(10):781–791.
57. Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, Maslen G, O'Brien J, Djimde A, Doumbo O, Zongo I, et al: **Analysis of *Plasmodium***

- falciparum* diversity in natural infections by deep sequencing. *Nature* 2012, **487**(7404):375–379.
58. Tan JC, Miller BA, Tan A, Patel JJ, Cheeseman IH, Anderson TJ, Manske M, Maslen G, Kwiatkowski DP, Ferdig MT: **An optimized microarray platform for assaying genomic variation in *Plasmodium falciparum* field populations.** *Genome Biol* 2011, **12**(4):R35.
  59. Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, *et al*: **A unified mixed-model method for association mapping that accounts for multiple levels of relatedness.** *Nat Genet* 2006, **38**(2):203–208.
  60. Mu J, Awadalla P, Duan J, McGee KM, Joy DA, McVean GA, Su XZ: **Recombination hotspots and population structure in *Plasmodium falciparum*.** *PLoS Biol* 2005, **3**(10):e335.
  61. Neafsey DE, Schaffner SF, Volkman SK, Park D, Montgomery P, Milner DA, Lukens A, Rosen D, Daniels R, Houde N, *et al*: **Genome-wide SNP genotyping highlights the role of natural selection in *Plasmodium falciparum* population divergence.** *Genome Biol* 2008, **9**(12):R171.
  62. Peakall R, Smouse P: **GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research.** *Mol Ecol Notes* 2006, **6**(1):288–295.

doi:10.1186/1471-2156-14-2

**Cite this article as:** Echeverry *et al.*: Long term persistence of clonal malaria parasite *Plasmodium falciparum* lineages in the Colombian Pacific region. *BMC Genetics* 2013 **14**:2.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

