MALARIA JOURNAL

RESEARCH   **Open Access**

Population genetic structure of *Plasmodium falciparum* across a region of diverse endemicity in West Africa

Victor A Mobegi[1,2], Kovana M Loua[3], Ambroise D Ahouidi[4], Judith Satoguina[2], Davis C Nwakanma[2], Alfred Amambua-Ngwa[2] and David J Conway[1,2*]

Abstract

Background: Malaria parasite population genetic structure varies among areas of differing endemicity, but this has not been systematically studied across *Plasmodium falciparum* populations in Africa where most infections occur.

Methods: Ten polymorphic *P. falciparum* microsatellite loci were genotyped in 268 infections from eight locations in four West African countries (Republic of Guinea, Guinea Bissau, The Gambia and Senegal), spanning a highly endemic forested region in the south to a low endemic Sahelian region in the north. Analysis was performed on proportions of mixed genotype infections, genotypic diversity among isolates, multilocus standardized index of association, and inter-population differentiation.

Results: Each location had similar levels of pairwise genotypic diversity among isolates, although there were many more mixed parasite genotype infections in the south. Apart from a few isolates that were virtually identical, the multilocus index of association was not significant in any population. Genetic differentiation between populations was low (most pairwise F_{ST} values < 0.03), and an overall test for isolation by distance was not significant.

Conclusions: Although proportions of mixed genotype infections varied with endemicity as expected, population genetic structure was similar across the diverse sites. Very substantial reduction in transmission would be needed to cause fragmented or epidemic sub-structure in this region.

Keywords: *Plasmodium falciparum*, Microsatellite, Population structure, Transmission intensity

Background

Plasmodium falciparum causes an annual burden of hundreds of millions of episodes of clinical malaria, and between approximately 0.5 and 1.5 million deaths, mostly in endemic populations of sub-Saharan Africa [1-3]. Transmission intensity of malaria parasites is highly variable temporally and geographically [4,5], and this variation plays a role in determining parasite population genetics and evolution locally. When human malaria infections contain a mixture of multiple haploid parasite clones, genetically different gametocytes may be taken into a mosquito blood meal, leading to heterozygous diploid parasites and meiotic reassortment and recombination as the haploid products are formed [6]. However, when infections are uncommon and mostly occur as single genotypes, there would be relatively more inbreeding of the parasites [7], leading to likely changes in population genetic structure during malaria control and elimination.

Microsatellite surveys of *P. falciparum* from endemic populations have clearly shown that infections are less genotypically mixed in areas of low transmission, which does apparently reduce the effective recombination rate and allow linkage disequilibrium to persist locally [8]. Such a pattern occurs in many populations in South America and Southeast Asia, and a wide spectrum of population structures can be seen among different sites within individual countries such as Brazil [9], Malaysia [10], Thailand [8,11,12], Philippines [13], and Papua New Guinea [8,14]. The degree of isolation and genetic

* Correspondence: david.conway@lshtm.ac.uk
[1]Department of Pathogen Molecular Biology, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK
[2]Medical Research Council Unit, Fajara, Banjul, The Gambia
Full list of author information is available at the end of the article

differentiation of local subpopulations within countries also varies in these areas, due to differences in the history of discrete transmission foci as well as some fragmentation in the distribution of current endemicity [8-14].

In contrast to the situation in such regions, most of the global *P. falciparum* infections occur in sub-Saharan Africa, and there is continuous endemicity throughout most of this region [1,5]. A number of independent studies indicate that local *P. falciparum* populations in Africa are genetically very diverse, with a high effective recombination rate due to many mixed genotype infections [6,8,15]. Furthermore, local populations are not strongly isolated from each other due to frequent migration of humans within sub-regions of the continent [8,16,17]. Exceptions are populations on the very edge of the endemic distribution such as in Djibouti [18], or on remote islands such as the Comoros [19], that show a fragmented structure as expected. However, individual studies have tended to focus on only one or a few sites, and parasite population structures in Africa have not been compared very systematically. It is particularly important to do so at this time, as substantial reductions in malaria in parts of Africa have recently occurred [20]. Such changes may have started to affect parasite population genetics, an understanding of which is needed for considering potential elimination or sustained endemic control.
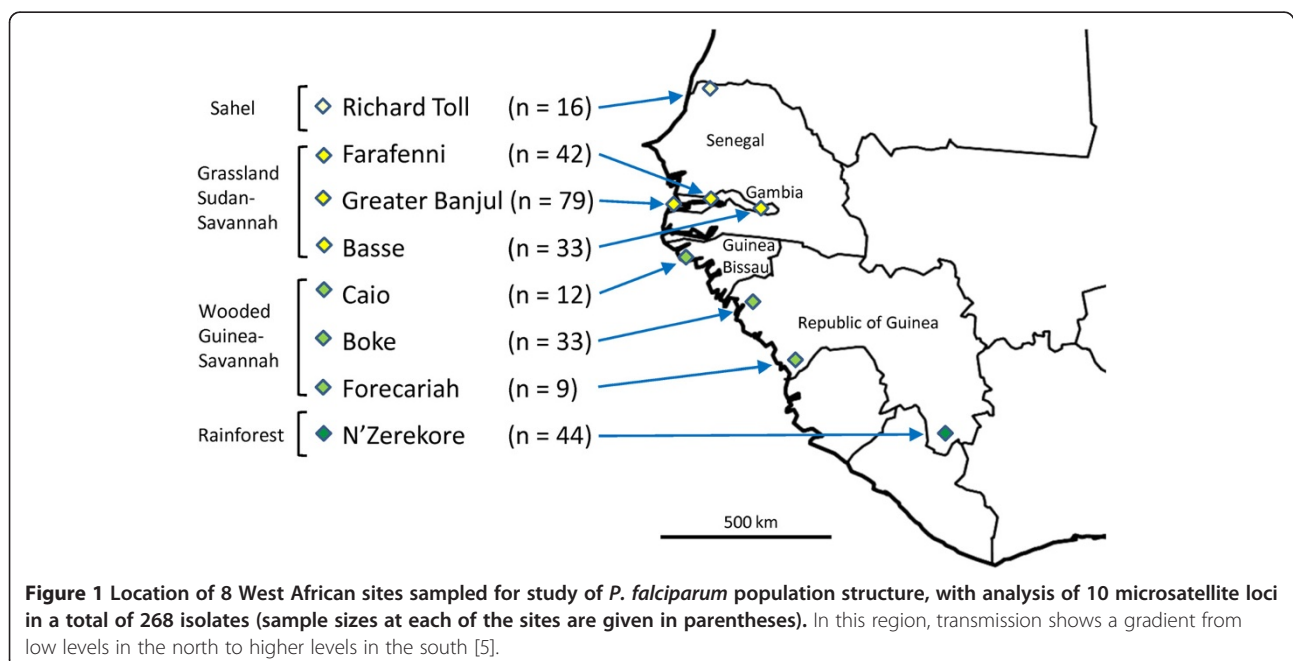
Here, the genetic structure of *P. falciparum* populations was compared among eight locations sampled in four contiguous countries in West Africa (Senegal, Gambia, Guinea Bissau, and Republic of Guinea). These sites vary in endemicity on a north–south gradient of seasonal rainfall which correlates with transmission [5], with ecology ranging from Sahel in the north to dense savannah and rainforest in the south. In the lower endemic northern areas there has recently been more malaria control, particularly in The Gambia and Senegal where significant decline in malaria has been observed during the past decade [21-23]. A set of ten highly polymorphic microsatellite loci was used to determine proportions of mixed genotype infections and pairwise differences among genotypic profiles of isolates, and to assess multi-locus allelic associations within populations and the degree of genetic differentiation among populations. The findings have implications for understanding the epidemiology of malaria in this region, and the potential for more effective control.

Methods

Plasmodium falciparum DNA samples

Genotyping of *P. falciparum* was performed on DNA extracted from blood samples from 268 infected individuals, sampled from eight sites in four West African countries between 2005 and 2009 (Figure 1). All subjects or their guardians gave written informed consent to provide a blood sample for studies of malaria that included genotyping of malaria parasites. Protocols were reviewed and approved by the Gambia Government and MRC joint Ethics Committee and the Ethics Committee of the London School of Hygiene and Topical Medicine, and investigators adhered to current guidelines on Good Clinical Practice. For five of the sites, in The Gambia (Basse, Farafenni and Greater Banjul area), Senegal (Richard Toll), and Guinea Bissau (Caio), DNA was

Figure 1 Location of 8 West African sites sampled for study of *P. falciparum* population structure, with analysis of 10 microsatellite loci in a total of 268 isolates (sample sizes at each of the sites are given in parentheses). In this region, transmission shows a gradient from low levels in the north to higher levels in the south [5].

previously extracted from blood samples positive for *P. falciparum* by slide examination, and those also positive by PCR with sufficient material for genotyping were analyzed here. The samples from Basse (n = 33), Farafenni (n = 42) and Caio (n = 12) were from community surveys conducted in January and February 2008 [24], shortly after the end of the 2007 malaria transmission season (most cases occur between August and November each year). Samples from the Greater Banjul area (n = 79) were from clinical cases positive for *P. falciparum* by microscopy presenting to four local health facilities (Royal Victoria Teaching Hospital in Banjul, the MRC clinic in Fajara, Jammeh Foundation for Peace Hospital in Serekunda, and Brikama Health Centre) during the malaria seasons of 2005–2009 [25]. Samples from Richard Toll in northern Senegal (n = 16) were collected from the Centre de Santé de Richard Toll during the malaria season of 2005 and were determined to be *P. falciparum* positive in a survey for parasite detection and drug resistance. In the Republic of Guinea, 105 filter-paper blood samples were collected from patients presenting with malaria between December 2009 and February 2010 who were positive for malaria parasites by rapid diagnostic test (RDT) and slide examination at the Government Health Centres in Boke, N'Zerekore and Forecariah. Genomic DNA was extracted in The Gambia using a Corbett Robotics X-Tractor Gene™ robot (Corbett Robotics Pty Ltd, Australia), and *Plasmodium* species identification was performed on the DNA samples using genus-and species-specific primers [26], identifying 102 *P. falciparum*-positive samples from Guinea, of which 86 had sufficient DNA for multi-locus microsatellite analysis here (44 from N'Zerekore, 33 from Boke, and 9 from Forecariah).

Microsatellite genotyping of *Plasmodium falciparum*

Plasmodium falciparum DNA from each individual sample was genotyped for 10 polymorphic microsatellite marker loci, chosen from a set of 12 previously described [27] (two loci that gave less efficient PCR amplification were not included). These loci consist of specific sequences flanking (TAA)$_n$ repeats in the haploid *P. falciparum* genome [28], and were PCR-amplified using a heminested protocol [27] with slight changes in PCR reaction mix constituent concentrations and primer dye labels: 6-FAM for locus *TA1* on chromosome 6 (Chr 6), *TA81* (Chr 5) and *ARA2* (Chr 11); VIC for *TA42* (Chr 5) and *PfPK2* (Chr 12); PET for *POLYα* (Chr 4), *TA87* (Chr 6) and *TA60* (Chr 13); NED for *TA109* (Chr 6) and *Pfg377* (Chr 12).

PCR products for the different loci amplified from each individual DNA sample were pooled into two groups of 5 loci together with GeneScan™ 500 LIZ internal size standard (Applied Biosystems, Foster City,

CA) for electrophoresis on an ABI 3130XL Genetic Analyzer. One pool comprised loci *TA42*, *TA109*, *POLYα*, *TA60* and *TA81* while the other comprised *TA1*, *TA87*, *ARA2*, *Pfg377* and *PfPK2* for a given individual sample. GeneMapper software version 4 (Applied Biosystems, UK) was used for scoring the allele sizes and for quantifying peak heights for samples containing multiple alleles per locus. Multiple alleles per locus were scored if electrophoretic peaks corresponding to minor alleles were ≥20% the height of that for the predominant allele in the isolate.

Microsatellite data analysis

For each locus in each individual isolate, the predominant allele (where multiple alleles were detected) or the only allele (where only a single allele was scored) was counted for population genetic analyses. Summary indices including numbers of alleles, allelic diversity within each population, and allele frequencies per locus per population were calculated using FSTAT Version 2.9.3.2 (updated from [29]). Allelic diversity was calculated for each of the 10 microsatellite loci based on the allele frequencies, using the formula for 'expected heterozygosity' $H_e = [n/(n-1)][1 - \sum p^2]$, where n is the number of isolates analyzed and p represents the frequency of each different allele at a locus. H_e has a potential range from 0 (no allele diversity) to 1 (all sampled alleles are different).

For isolates that were fully genotyped at all 10 loci, analysis of multi-locus genotypic profiles of the isolate was performed, considering the majority allele at each locus in the case of mixed genotype infections. Pairwise comparisons among all isolates within each population were performed using Splitstree [30] to generate and graphically visualize a matrix of similarity among isolates based on numbers of identical or mismatched loci. To test for existence of multilocus linkage disequilibrium, the standardized index of association (I_A^S) was calculated, based on genotypic profiles of the majority allele at each locus in each infection, using the LIAN version 3.5 Web interface [31]. This index is calculated as $I_A^S = 1/L-1(V_D/V_E-1)$, where L is the number of loci used, and the variance (V_D) in pairwise numbers of mismatched alleles between isolates was compared with that expected under linkage equilibrium (V_E) by simulation of data by 10000 iterations under the null hypothesis of $V_D = V_E$ by Monte Carlo random sampling [31].

The level of genetic differentiation between the populations was calculated using fixation indices (F_{ST}) computed with FSTAT, and a test for isolation by distance was performed from the pairwise genetic distances and geographical distances between populations, using a Mantel test of matrix correlation (Genepop version 4.0.10 web interface, Isolde program) [32].

Table 1 Allelic diversity (expected heterozygosity, H_e) at 10 microsatellite loci in 8 local populations of *Plasmodium falciparum* in West Africa

| Locus | Populations sampled (n = numbers of individual isolates) | | | | | | | |
| | Guinea | | | Guinea Bissau | Gambia | | | Senegal |
	N'Zerekore (n = 44)	Boke (n = 33)	Forecariah (n = 9)	Caio (n = 12)	Basse (n = 33)	Farafenni (n = 42)	Greater Banjul (n = 79)	Richard Toll (n = 16)
TA1	0.85	0.88	0.89	0.91	0.88	0.85	0.86	0.91
TAA87	0.85	0.79	0.81	0.82	0.86	0.85	0.85	0.74
ARA2	0.88	0.88	0.72	0.85	0.84	0.82	0.82	0.71
Pfg377	0.64	0.57	0.56	0.64	0.63	0.52	0.53	0.69
PfPK2	0.87	0.85	0.86	0.89	0.88	0.89	0.85	0.93
Polya	0.85	0.90	0.83	0.96	0.78	0.89	0.89	0.86
TAA60	0.80	0.83	0.89	0.68	0.82	0.79	0.82	0.75
TAA81	0.85	0.83	0.89	0.74	0.84	0.80	0.79	0.76
TAA109	0.83	0.86	0.86	0.85	0.88	0.86	0.87	0.82
TA42	0.34	0.54	0.39	0.68	0.16	0.22	0.25	0.00
Mean	0.78	0.79	0.77	0.80	0.76	0.75	0.75	0.72

Results

All of the ten microsatellite loci were highly polymorphic in all eight sites from the four countries sampled, with the overall numbers of different alleles per locus ranging from 7 (for locus *TA42*) to 24 (for locus *Polya*). For each of the 268 isolates, the genotype profile is listed in Additional file 1, and allele frequencies at each of the ten loci in each of the eight sites are given in Additional file 2. All isolates were genotyped successfully for at least eight of the ten loci, and 186 (69.4%) isolates had results for all loci (most missing data were from locus *TA42* which was genotyped in 76.9% of isolates) (Additional file 1). Allelic diversity at each locus was summarized as the expected heterozygosity (H_e) from the distribution of allele frequencies, and was similar across all the sites, with averaged values across the loci between 0.72 (for Richard Toll) and 0.80 (for Caio)

(Table 1; $P > 0.1$ for each Wilcoxon Signed Ranks test on pairs of sites).

The number of genotypes detected in an isolate was defined as the maximum number of alleles scored at any of the individual loci. The numbers of parasite genotypes detected per isolate differed substantially across populations (Kruskal-Wallis test, $P < 0.001$), and tended to be higher at sites in the south (Figure 1 and Table 2). In particular, the three sites in the Republic of Guinea had significantly higher numbers (means of 4.2 in Forecariah, 4.1 in Boke, and 3.7 in N'Zerekore) than each of the other sites, in Guinea-Bissau (mean of 2.6 in Caio), Senegal (2.2 in Richard Toll) and The Gambia (2.4 in Basse, 2.1 in Farafenni and 1.7 in the Greater Banjul area) (Table 2; $P < 0.001$ for each Mann–Whitney test). Remaining pairwise comparisons among sites were not significant, except that the numbers of genotypes per

Table 2 Multiple genotype infections assessed by typing 10 *P. falciparum* microsatellite loci in 268 isolates from 8 West African locations

| Country | Location | Number of isolates | Number of isolates with given no. of detected genotypes | | | | | Mean number of genotypes per isolate [a] | Proportion of locus scores with >1 allele [b] |
			1	2	3	4	>4		
Guinea	N'Zerekore	44	1	6	8	14	15	3.7	0.58
Guinea	Boke	33	0	3	5	14	11	4.1	0.54
Guinea	Forecariah	9	0	0	2	3	4	4.2	0.51
Guinea Bissau	Caio	12	0	6	5	1	0	2.6	0.30
Gambia	Greater Banjul	79	34	32	13	0	0	1.7	0.14
Gambia	Basse	33	1	20	11	1	0	2.4	0.35
Gambia	Farafenni	42	6	26	8	2	0	2.1	0.33
Senegal	Richard Toll	16	3	9	3	0	1	2.2	0.26

[a] For this analysis the number of genotypes was scored from the locus with the largest number of alleles within the isolate.
[b] For this analysis the proportion of loci with >1 allele within each isolate was scored.

isolate were significantly lower in the Greater Banjul area than in each of the other sites (P < 0.005) except for Richard Toll. In a complementary analysis, the proportion of individual locus scores with more than one allele per isolate was highest in Nzerekore (0.58) and lowest in the Greater Banjul area (0.14) (Table 2 and Additional file 1).
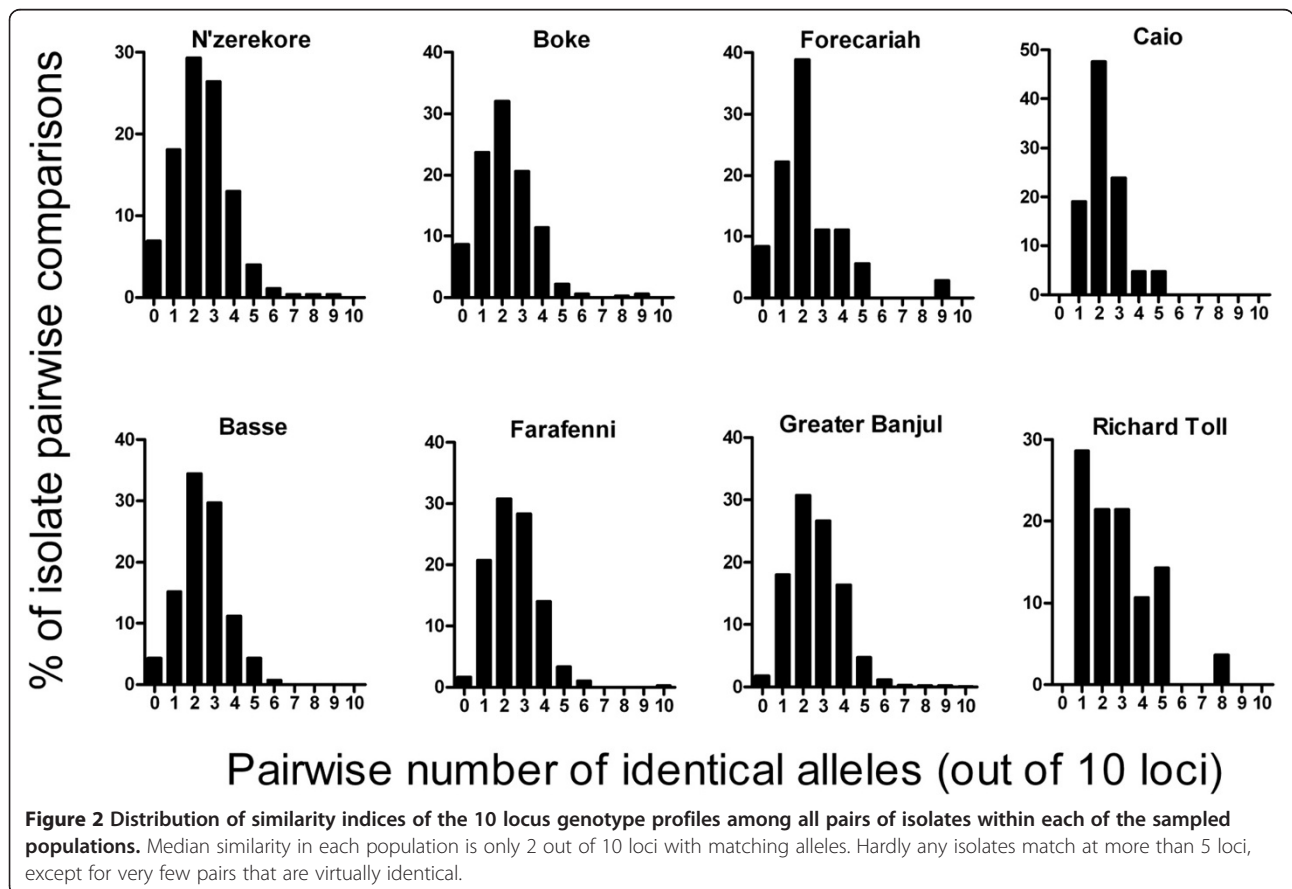
Most isolates had complete genotype data for all loci, and in these the multilocus profiles were examined, considering the majority allele at each of the ten loci. Almost all isolates differed from each other at most loci, and for each population the median number of loci with matching alleles in pairwise comparisons was only two out of 10 (Figure 2). However, a very small number of isolates showed exceptionally high level of identity, with six of the eight populations containing at least one pair of isolates matching at 8 or more loci (matching genotypes were different in each population). For example, in Farafenni a single pair of isolates were identical at all loci while none of the other pairs were identical at more than 6 loci, and in Richard Toll one pair of isolates were identical at 8 loci while none of the rest were identical at more than 5 loci (Figure 2). Performing an initial multilocus index of association (I_A^S) test, significant values were detected in five of the populations (Table 3), but

Table 3 Standardized multi-locus Index of Association (I_A^S) test for *Plasmodium falciparum* isolates genotyped for a complete set of 10 microsatellite loci at 8 West African locations

Country	Location	Including all isolates			Non-identical genotypes [§]		
		n	I_A^S		n	I_A^S	
Guinea	N'Zerekore	24	0.036	**	22	0.019	NS
Guinea	Boke	26	0.032	**	23	0.011	NS
Guinea	Forecariah	9	0.106	***	8	0.041	NS
Guinea Bissau	Caio	7	0.000	NS	7	0.000	NS
Gambia	Basse	24	0.008	NS	24	0.008	NS
Gambia	Farafenni	25	0.007	NS	24	0.000	NS
Gambia	Greater Banjul	63	0.017	**	59	0.009	NS
Senegal	Richard Toll	8	0.150	***	7	0.061	NS

n = number of isolates analysed; [§] to exclude genotypes that were virtually identical one member of each isolate pair matching at 8 or more of the 10 loci was removed (2 isolates in N'Zerekore, 3 in Boke; 1 in Forecariah, 1 in Farafenni, 4 in Greater Banjul and 1 in Richard Toll); NS P > 0.05, not significant; * P < 0.05; ** P < 0.01; *** P < 0.001.

after removal of individual isolates matching another at 8 or more loci none of the I_A^S values was significant (Table 3). As many isolates had mixed genotypes, the dominant multilocus profile in some could be a

Figure 2 Distribution of similarity indices of the 10 locus genotype profiles among all pairs of isolates within each of the sampled populations. Median similarity in each population is only 2 out of 10 loci with matching alleles. Hardly any isolates match at more than 5 loci, except for very few pairs that are virtually identical.

composite of different genotypes rather than a true haplotype, but previous analysis of other populations shows very similar I_A^S results whether only including single genotype infections or a more broad inclusion of profiles from all infections [8]. Here, there was a sufficient number of single-clone isolates fully genotyped in the Greater Banjul area for a separate analysis (n = 26), which also showed no significant index of association (I_A^S = 0.020, P > 0.05).

Most alleles were distributed widely across different populations, and any 'private' alleles (detected only in one population) were at very low frequencies (Additional file 2). Pairwise comparisons of populations showed that allele frequencies were similar and the F_{ST} values were low, ranging from not significantly greater than zero for many population pairs, through to F_{ST} = 0.065 (P = 0.048) between Richard Toll and Caio (Table 4). The F_{ST} values and geographical distances between all pairs of populations were analysed by a Mantel test of matrix correlation, showing no significant overall evidence for isolation by distance (P = 0.304). However, the northernmost site at Richard Toll tended to be more divergent from the rest of the populations (with F_{ST} values of 0.028 to 0.065; Figure 3 and Table 4).

Discussion

Across a marked gradient in transmission intensity among West African locations sampled here, the parasite populations show remarkably similar population genetic structure. There was a high allelic diversity for the ten *P. falciparum* microsatellite loci in all the eight locations, the overall H_e index for each site ranging from 0.72 to 0.80, the same range as that previously seen in most other endemic African countries [8,16,33]. Although slightly lower diversity has been reported in a low-endemic area

Figure 3 Scatterplot of pairwise genetic divergence between populations (F_{ST} mean for all 10 microsatellite loci) plotted by geographical distance in kilometers between them. Comparisons that include the northernmost population (Richard Toll) are shown with square points, while all others are shown with circular points. An overall test for isolation by distance was not significant (Mantel test of matrix correlation, P = 0.304).

of Dakar in Senegal [18] and a highland area of western Kenya [34], a much lower diversity has only been reported for one African location (H_e of 0.41 in Djibouti which has unstable epidemic transmission) [18].

As expected, the mean numbers of *P. falciparum* genotypes detected per infection varied among the populations studied, with highly mixed infections at each of the sites in Guinea which experience high transmission for much of the year, in contrast with sites having limited seasonal transmission further north in the region. This substantial difference is likely to be due to differences in endemicity as predicted, rather than confounding due to differences in sampling. Samples from five of the sites were from clinical cases presenting to health facilities, and those from the remaining three sites were from community sampling of asymptomatic individuals, but the latter sites were in the middle of the range geographically and in terms of proportions of mixed genotype infections, so this did not cause the observed north–south contrast in values. Review of comparisons elsewhere between asymptomatic and symptomatic infections also shows no consistent overall difference in proportions of mixed genotype infections [35-40].

The multi-locus index of association analysis initially indicated non-random patterns in several populations, but this was shown to be entirely due to only one or a few pairs of virtually identical isolates. In South American and Southeast Asian populations of low endemicity, identical *P. falciparum* isolates are sometimes seen as multi-locus genotypes persist through a number of self-

Table 4 Matrix of genetic differentiation (F_{ST} values below diagonal) and geographical distances (in Km above diagonal) in pairwise comparisons of eight sampled *P. falciparum* populations

	Nzr	Bok	For	Bas	Cai	Far	Gba	Ric
Nzr		696	504	852	932	980	1055	1218
Bok	0.009		213	262	235	323	372	629
For	0.001	0.013		446	439	533	582	825
Bas	0.019	0.021	0.012		264	151	252	382
Cai	0.028	0.032	0.027	0.027		193	174	504
Far	0.007	0.022	0.017	0.000	0.023		104	320
Gba	0.009	0.023	0.005	0.005	0.026	0.000		346
Ric	0.037	0.061	0.040	0.043	0.065	0.028	0.046	

Nzr = N'Zerekore; Bok = Boke; For = Forecariah; Bas = Basse; Cai = Caio; Far = Farafenni; Gba = Greater Banjul area; Ric = Richard Toll.
Pairwise F_{ST} values are the mean for all 10 microsatellite loci, and values significantly above zero (p < 0.01) are underlined. Sample sizes for each location are given in Methods and Table 1.
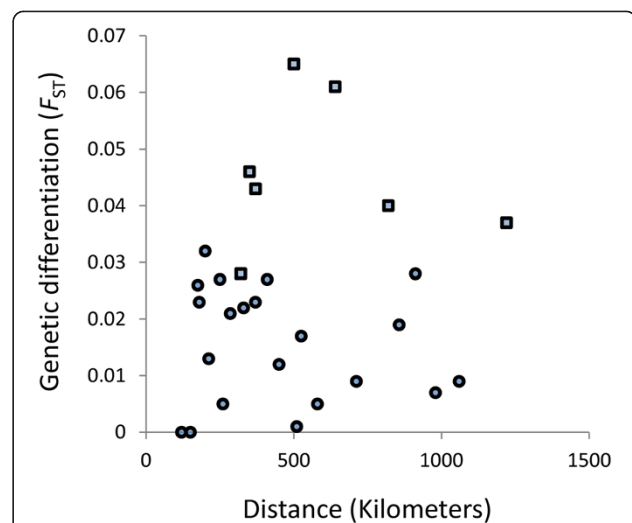
fertilisation and transmission cycles. This has allowed comparisons of identical versus non-identical pairs of clones for estimation of phenotype heritability in a manner analogous to a conventional twin-pair study [41,42]. In Africa, related parasites may occur within isolates but it is rare to see identical genotypes in different infections [43]. It has been previously shown in The Gambia that identical parasite genotypes were more common in pairs of children sleeping in the same house who presented with malaria on the same day [44], and occasionally in children living close to each other within small villages [45]. Such identical genotypes probably result from single mosquitoes infecting more than one individual or from transmission of parasites by different mosquitoes that fed on a single genotype gametocyte carrier.

As was seen here, it has elsewhere been shown that inclusion of a very small number of closely related parasite genotypes can generate a significantly non-random multi-locus index of association, even when there is no linkage disequilibrium among the loci in the population generally [8]. Such occasional occurrence of identical parasites should not lead to a population being considered 'clonal', as such a term is not generally applied to human population genetic structure despite the presence of monozygotic twins. Instead, it has been suggested that the presence of many pairs of genotypically similar isolates may be taken as evidence of an 'epidemic' population structure [8,46], but here there were very few such pairs and none of the populations could be described as having such a structure. It should be noted that if multiple sibling parasite genotypes within the same host were included in a crude analysis, it could give false appearance of linkage disequilibrium [47,48]. Therefore, such a problem is avoided by restricting analysis to one parasite genotype per host [49-51], following the general principle whereby genotypes from closely related family members are not separately counted in studies of human linkage disequilibrium.

Over the sampled range of up to ~ 1200 kilometers between sites in West Africa, this microsatellite analysis found very low levels of genetic differentiation between the local populations of P. falciparum, with most pairwise F_{ST} values being less than 0.03. This is considerably less than the differentiation among local sites within non-African countries that are less endemic, including Malaysia [10], Papua New Guinea [14], the Philippines [13], and Brazil [9], with F_{ST} values exceeding 0.10 between sites separated by similar distances. It is instead consistent with previous sampling from more widely separated African populations, which has shown F_{ST} values of less than 0.05 for a similar set of microsatellite loci [8,16]. Despite the ecological and epidemiological diversity in West Africa, there is likely to be considerable mixing of parasites between different locations, due to frequent movement of

humans in this region [17]. In comparison with South America and Southeast Asia, it will be difficult to identify discrete endemic locations in West Africa where malaria elimination might be achieved and sustained in the face of local migration.

Despite studying only ten polymorphic loci and sampling a very limited number of isolates from some of the sites, marked variation in proportions of multiple clone infections and relatively similar genetic structure of P. falciparum populations has been clearly shown, providing a framework for future genomic-scale studies. Comprehensive analysis of genome sequence variation should allow finer differences in population structure to be detected [52], including variation in patterns across the genome and identification of genes under natural selection, particularly with large sample sizes. Given the high levels of recombination and minimal reproductive isolation of parasite populations in West Africa, we predict that differential signatures of selection in particular populations will be detectable against a background of neutral genomic variation that is more spatially homogeneous.

Conclusions

This analysis of ten microsatellite loci in P. falciparum in West Africa showed there were more mixed parasite genotype infections in highly endemic forested areas in the south than in drier areas with lower transmission in the north, although each location showed similar levels of allelic diversity and pairwise genotypic diversity among isolates. Apart from a few exceptional pairs of isolates that were virtually identical, there was no significant multilocus index of association in any population, genetic differentiation between locations was low, and an overall test for isolation by distance was not significant. Substantial future reduction in transmission would be needed before fragmented or epidemic sub-structure in parasite populations is likely to be seen in this region. Given the high levels of recombination and minimal reproductive isolation of parasite populations in this region, genome-wide studies may detect differential signatures of selection against a background of neutral variation that is relatively spatially homogeneous.

Additional files

Additional file 1: Table S1. Alleles scored in each of the 268 P. falciparum isolates genotyped at 10 microsatellite loci. Microsatellite alleles detected in isolates from the eight sampled populations, with highlighting showing predominant allele calls in the mixed genotypes.

Additional file 2: Table S2. Allele frequencies at 10 microsatellite loci in each of the eight population samples of Plasmodium falciparum. Data on frequency of each allele scored for each microsatellite locus in each of the eight populations sampled.

Author details
[1]Department of Pathogen Molecular Biology, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK. [2]Medical Research Council Unit, Fajara, Banjul, The Gambia. [3]National Institute of Public Health, Conakry, Republic of Guinea. [4]Université Cheikh Anta Diop, Dakar, Senegal.

Authors' contributions
VAM and DJC designed the experiments. VAM performed the experiments. VAM and DJC analysed the data. VAM, KML, ADA, JS, DCN, AA-N, and DJC provided materials and tools. VAM and DJC wrote the paper. All authors read and approved the final manuscript.

References
1. Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI: **The global distribution of clinical episodes of Plasmodium falciparum malaria.** Nature 2005, **434**:214–217.
2. WHO: World Malaria Report. Geneva: World Health Organization; 2010. http://www.who.int/malaria/world_malaria_report_2010/en/.
3. Murray CJ, Rosenfeld LC, Lim SS, Andrews KG, Foreman KJ, Haring D, Fullman N, Naghavi M, Lozano R, Lopez AD: **Global malaria mortality between 1980 and 2010: a systematic analysis.** Lancet 2012, **379**:413–431.
4. Hay S, Rogers D, Toomer J, Snow R: **Annual Plasmodium falciparum entomological inoculation rates (EIR) across Africa: literature survey, Internet access and review.** Trans R Soc Trop Med Hyg 2000, **94**:113–127.
5. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, Kabaria CW, Manh BH, Elyazar IR, Brooker S, Smith DL, Moyeed RA, Snow RW: **A world malaria map: Plasmodium falciparum endemicity in 2007.** PLoS Med 2009, **6**:e1000048.
6. Mu J, Awadalla P, Duan J, McGee KM, Joy DA, McVean GA, Su X-Z: **Recombination hotspots and population structure in Plasmodium falciparum.** PLoS Biol 2005, **3**:e335.
7. Paul REL, Packer MJ, Walmsley M, Lagog M, Ranford-Cartwright LC, Paru R, Day KP: **Mating patterns in malaria parasite populations of Papua New Guinea.** Science 1995, **269**:1709–1711.
8. Anderson TJC, 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.** Mol Biol Evol 2000, **17**:1467–1482.
9. Machado RLD, Póvoa MM, Calvosa VSP, Ferreira MU, Rossit ARB, dos Santos EJM, Conway DJ: **Genetic structure of Plasmodium falciparum populations in the Brazilian Amazon region.** J Infect Dis 2004, **190**:1547–1555.
10. Anthony T, Conway D, Cox-Singh J, Matusop A, Ratnam S, Shamsul S, Singh B: **Fragmented population structure of Plasmodium falciparum in a region of declining endemicity.** J Infect Dis 2005, **191**:1558–1564.
11. Chokejindachai W, Conway DJ: **Case–control approach to identify Plasmodium falciparum polymorphisms associated with severe malaria.** PLoS One 2009, **4**:e5454.
12. Pumpaibool T, Arnathau C, Durand P, Kanchanakhan 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.
13. Iwagami M, Rivera P, Villacorte E, Escueta A, Hatabu T, Kawazu S-i, Hayakawa T, Tanabe K, Kano S: **Genetic diversity and population structure of Plasmodium falciparum in the Philippines.** Malar J 2009, **8**:96.
14. Schultz L, Wapling J, Mueller I, Ntsuke PO, Senn N, Nale J, Kiniboro B, Buckee CO, Tavul L, Siba PM, Reeder JC, Barry AE: **Multilocus haplotypes reveal variable levels of diversity and population structure of Plasmodium falciparum in Papua New Guinea, a region of intense perennial transmission.** Malar J 2010, **9**:336.
15. Conway DJ, Roper C, Oduola AMJ, Arnot DE, Kremsner PG, Grobusch MP, Curtis CF, Greenwood BM: **High recombination rate in natural populations of Plasmodium falciparum.** Proc Natl Acad Sci USA 1999, **96**:4506–4511.
16. Conway DJ, Machado RLD, Singh B, Dessert P, Mikes ZS, Povoa MM, Oduola AMJ, Roper C: **Extreme geographical fixation of variation in the Plasmodium falciparum gamete surface protein gene Pfs48/45 compared with microsatellite loci.** Mol Biochem Parasitol 2001, **115**:145–156.
17. Lynch C, Roper C: **The transit phase of migration: circulation of malaria and its multidrug-resistant forms in Africa.** PLoS Med 2011, **8**:e1001040.
18. Bogreau H, Renaud F, Bouchiba H, Durand P, Assi S-B, Henry M-C, Garnotel E, Pradines B, Fusai T, Wade B, Adehossi E, Parola P, Kamil MA, Puijalon O, Rogier C: **Genetic diversity and structure of African Plasmodium falciparum populations in Urban and Rural areas.** Am J Trop Med Hyg 2006, **74**:953–959.
19. Rebaudet S, Bogreau H, Silaï R, Lepère J-F, Bertaux L, Pradines B, Delmont J, Parola P, Rogier C: **Genetic Structure of Plasmodium falciparum and Elimination of Malaria, Comoros Archipelago.** Emerg Infect Dis 2010, **16**:1686–1694.
20. O'Meara WP, Mangeni JN, Steketee R, Greenwood B: **Changes in the burden of malaria in sub-Saharan Africa.** Lancet Infect Dis 2010, **10**:545–555.
21. Ceesay SJ, Casals-Pascual C, Nwakanma DC, Walther M, Gomez-Escobar N, Fulford AJC, Takem EN, Nogaro S, Bojang KA, Corrah T, Jaye MC, Taal MA, Sonko AA, Conway DJ: **Continued decline of malaria in The Gambia with implications for elimination.** PLoS One 2010, **5**:e12242.
22. Ceesay SJ, Climent C-P, Jamie E, Samuel EA, Nancy OD, Anthony JCF, Sanie SSS, Ismaela A, Samuel D, Omar S, Palmer A, Fofana M, Corrah T, Bojang KA, Whittle HC, Greenwood BM, Conway DJ: **Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis.** Lancet 2008, **372**:1545–1554.
23. Trape JF, Sauvage C, Ndiaye O, Douillot L, Marra A, Diallo A, Cisse B, Greenwood B, Milligan P, Sokhna C, Molez JF: **New malaria-control policies and child mortality in Senegal: reaching Millennium Development Goal 4.** J Infect Dis 2012, **205**:672–679.
24. Satoguina J, Walther B, Drakeley C, Nwakanma D, Oriero E, Correa S, Corran P, Conway D, Walther M: **Comparison of surveillance methods applied to a situation of low malaria prevalence at rural sites in The Gambia and Guinea Bissau.** Malar J 2009, **8**:274.
25. Walther M, Jeffries D, Finney OC, Njie M, Ebonyi A, Deininger S, Lawrence E, Ngwa-Amambua A, Jayasooriya S, Cheeseman IH, Gomez-Escobar N, Okebe J, Conway DJ, Riley EM: **Distinct roles for FOXP3 and FOXP3 CD4 T cells in regulating cellular immunity to uncomplicated and severe Plasmodium falciparum malaria.** PLoS Pathog 2009, **5**:e1000364.
26. Singh B, Bobogare A, Cox-Singh J, Snounou G, Abdullah M, Rahman H: **A genus- and species-specific nested polymerase chain reaction malaria detection assay for epidemiologic studies.** Am J Trop Med Hyg 1999, **60**:687–692.
27. Anderson T, Su X, Bockarie M, Lagog M, Day K: **Twelve microsatellite markers for characterization of Plasmodium falciparum from finger-prick blood samples.** Parasitology 1999, **119**:113–125.
28. Su X-Z, Wellems TE: **Toward a high-resolution Plasmodium falciparum linkage map: Polymorphic markers from hundreds of simple sequence repeats.** Genomics 1996, **33**:430–444.
29. Goudet J: **FSTAT (version 1.2): a computer program to calculate F statistics.** J Hered 1995, **86**:485–486.
30. Huson DH, Bryant D: **Application of phylogenetic networks in evolutionary studies.** Mol Biol Evol 2006, **23**:254–267.
31. Haubold B, Hudson RR: **LIAN 3.0: detecting linkage disequilibrium in multilocus data.** Bioinformatics 2000, **16**:847–849.
32. Rousset F: **Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux.** Mol Ecol Resources 2008, **8**:103–106.

33. Durand P, Michalakis Y, Cestier S, Oury B, LeClerc MC, Tibayrenc M, Renaud F: Significant linkage disequilibrium and high genetic diversity in a population of Plasmodium falciparum from an area (Republic of The Congo) highly endemic for malaria. Am J Trop Med Hyg 2003, 68:345–349.

34. Zhong D, Afrane Y, Githeko A, Yang Z, Cui L, Menge D, Temu E, Yan G: Plasmodium falciparum genetic diversity in western Kenya highlands. Am J Trop Med Hyg 2007, 77:1043–1050.

35. Amodu OK, Oyedeji SI, Ntoumi F, Orimadegun AE, Gbadegesin RA, Olumese PE, Omotade OO: Complexity of the msp2 locus and the severity of childhood malaria, in south-western Nigeria. Ann Trop Med Parasitol 2008, 102:95–102.

36. Beck HP, Felger I, Huber W, Steiger S, Smith T, Weiss N, Alonso P, Tanner M: Analysis of multiple Plasmodium falciparum infections in Tanzanian children during the phase III trial of the malaria vaccine SPf66. J Infect Dis 1997, 175:921–926.

37. Kyes S, Harding R, Black G, Craig A, Peshu N, Newbold C, Marsh K: Limited spatial clustering of individual Plasmodium falciparum alleles in field isolates from coastal Kenya. Am J Trop Med Hyg 1997, 57:205–215.

38. Roper C, Richardson W, Elhassan IM, Giha H, Hviid L, Satti GMH, Theander TG, Arnot DE: Seasonal changes in the Plasmodium falciparum population in individuals and their relationship to clinical malaria: a longitudinal study in a Sudanese village. Parasitology 1998, 116:501–510.

39. Zwetyenga J, Rogier C, Tall A, Fontenille D, Snounou G, Trape JF, Mercereau-Puijalon O: No influence of age on infection complexity and allelic distribution in Plasmodium falciparum infections in Ndiop, a Senegalese village with seasonal, mesoendemic malaria. Am J Trop Med Hyg 1998, 59:726–735.

40. Engelbrecht F, Felger I, Genton B, Alpers M, Beck H-P: Plasmodium falciparum: malaria morbidity is associated with specific merozoite surface antigen 2 genotypes. Exp Parasitol 1995, 81:90–96.

41. Anderson TJ, Nair S, Nkhoma S, Williams JT, Imwong M, Yi P, Socheat D, Das D, Chotivanich K, Day NP, White NJ, Dondorp AM: High heritability of malaria parasite clearance rate indicates a genetic basis for artemisinin resistance in western Cambodia. J Infect Dis 2010, 201:1326–1330.

42. Anderson TJ, Williams JT, Nair S, Sudimack D, Barends M, Jaidee A, Price RN, Nosten F: Inferred relatedness and heritability in malaria parasites. Proc R Soc Ser B 2010, 277:2531–2540.

43. Nkhoma SC, Nair S, Cheeseman IH, Rohr-Allegrini C, Singlam S, Nosten F, Anderson TJ: Close kinship within multiple-genotype malaria parasite infections. Proc R Soc Ser B 2012, 279:2589–2598.

44. Conway DJ, McBride JS: Genetic evidence for the importance of interrupted feeding by mosquitoes in the transmission of malaria. Trans R Soc Trop Med Hyg 1991, 85:454–456.

45. Conway DJ, McBride JS: Population genetics of Plasmodium falciparum within a malaria hyperendemic area. Parasitology 1991, 103:7–16.

46. Maynard-Smith J, Smith N, O'Rourke M, Spratt B: How clonal are bacteria? Proc Natl Acad Sci USA 1993, 90:4384–4388.

47. Razakandrainibe FG, Durand P, Koella JC, Meeüs TD, Rousset F, Ayala FJ, Renaud F: "Clonal" population structure of the malaria agent Plasmodium falciparum in high-infection regions. Proc Natl Acad Sci USA 2005, 102:17388–17393.

48. Annan Z, Durand P, Ayala FJ, Arnathau C, Awono-Ambene P, Simard F, Razakandrainibe FG, Koella JC, Fontenille D, Renaud F: Population genetic structure of Plasmodium falciparum in the two main African vectors, Anopheles gambiae and Anopheles funestus. Proc Natl Acad Sci USA 2007, 104:7987–7992.

49. Anderson TJC, Paul REL, Donnelly CA, Day KP: Do malaria parasites mate non-randomly in the mosquito midgut? Genet Res 2000, 75:285–296.

50. Mzilahowa T, McCall PJ, Hastings IM: "Sexual" population structure and genetics of the malaria agent P. falciparum. PLoS One 2007, 2:e613.

51. Babiker HA, Ranford-Cartwright LC, Currie D, Charlwood JD, Billingsley P, Teuscher T, Walliker D: Random mating in a natural population of the malaria parasite Plasmodium falciparum. Parasitology 1994, 109:413–421.

52. Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, Maslen G, O'Brien J, Djimde A, Doumbo O, Zongo I, Ouedraogo J-B, Michon P, Mueller I, Siba P, Nzila A, Borrmann S, Kiara SM, Marsh K, Jiang H, Su X-Z, Amaratunga C, Fairhurst R, Socheat D, Nosten F, Imwong M, White NJ, Sanders M, Anastasi E, Alcock D, Drury E, Oyola S, Quail MA, Turner DJ, Ruano-Rubio V, Jyothi D, Amenga-Etego L, Hubbart C, Jeffreys A, Rowlands K, Sutherland C, Roper C, Mangano V, Modiano D, Tan JC, Ferdig MT, Amambua-Ngwa A, Conway DJ, Takala-Harrison S, Plowe CV, Rayner JC, Rockett KA, Clark TG, Newbold CI, Berriman M, MacInnis B, Kwiatkowski DP: Analysis of Plasmodium falciparum diversity in natural infections by deep sequencing. Nature 2012, 487:375–379.