

## Accepted Manuscript

Genetic variability and population structure of *Plasmodium falciparum* parasite populations from different malaria ecological regions of Kenya

Luicer A. Ingasia, Jelgat Cheruiyot, Sheila Akinyi Okoth, Ben Andagalu, Edwin Kamau

PII: S1567-1348(15)30013-7  
DOI: doi: [10.1016/j.meegid.2015.10.013](https://doi.org/10.1016/j.meegid.2015.10.013)  
Reference: MEEGID 2516

To appear in:

Received date: 2 April 2015  
Revised date: 9 October 2015  
Accepted date: 10 October 2015

Please cite this article as: Ingasia, Luicer A., Cheruiyot, Jelgat, Okoth, Sheila Akinyi, Andagalu, Ben, Kamau, Edwin, Genetic variability and population structure of *Plasmodium falciparum* parasite populations from different malaria ecological regions of Kenya, (2015), doi: [10.1016/j.meegid.2015.10.013](https://doi.org/10.1016/j.meegid.2015.10.013)

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



**Genetic variability and population structure of *Plasmodium falciparum* parasite populations from different malaria ecological regions of Kenya**

**Luicer A. Ingasia<sup>1</sup>, Jelagat Cheruiyot<sup>1</sup>, Sheila Akinyi Okoth<sup>2,3</sup>, Ben Andagalu<sup>1</sup>, Edwin Kamau<sup>1#</sup>**

<sup>1</sup>*Department of Emerging and Infectious Diseases (DEID), United States Army Medical Research Unit-Kenya (USAMRU-K), Kenya Medical Research Institute (KEMRI)/Walter Reed Project (WRP), Kisumu, Kenya*

<sup>2</sup>*Malaria Branch, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, United States,*

<sup>3</sup>*Atlanta Research and Education Foundation/VA Medical Center, Decatur GA, United States*

#Address correspondence to [edwin.kamau.mil@mail.mil](mailto:edwin.kamau.mil@mail.mil)

Keywords: *Plasmodium falciparum*; Kenya; Genetic diversity; Population structure;

**Abstract**

Transmission intensity, movement of human and vector hosts, biogeographical features, and malaria control measures are some of the important factors that determine *Plasmodium falciparum* parasite genetic variability and population structure. Kenya has different malaria ecologies which might require different disease intervention methods. Refined parasite population genetic studies are critical for informing malaria control and elimination strategies. This study describes the genetic diversity and population structure of *P. falciparum* parasites from the different malaria ecological zones in Kenya. Twelve multi-locus microsatellite (MS) loci previously described were genotyped in 225 *P. falciparum* isolates collected between 2012 and 2013 from five sites; three in lowland endemic regions (Kisumu, Kombewa, and Malindi) and two in highland, epidemic regions (Kisii and Kericho). Parasites from the lowland endemic and highland epidemic regions of western Kenya had high genetic diversity compared to coastal lowland endemic region of Kenya [Malindi]. The Kenyan parasites had a mean genetic differentiation index ( $F_{ST}$ ) of 0.072 ( $P=0.011$ ). The multi-locus genetic analysis of the 12 MS revealed all the parasites had unique haplotypes. Significant linkage disequilibrium (LD) was observed in all the five parasite populations. Kisumu had the most significant index of association values (0.16;  $P < 0.0001$ ) whereas Kisii had the least significant index of association values (0.03;  $P < 0.0001$ ). Our data suggest high genetic diversity in Kenyan parasite population with the exception of parasite from Malindi where malaria has been on the decline. The presence of significant LD suggests that there is occurrence of inbreeding in the parasite population. Parasite populations from Kisii showed the strongest evidence for epidemic population structure whereas the rest of the regions showed panmixia. Defining the genetic diversity of the parasites in different ecological regions of Kenya after introduction of the artemether-lumefantrine is

important in refining the spread of drug resistant strains and malaria transmission for more effective control and eventual elimination of malaria in Kenya.

ACCEPTED MANUSCRIPT

## 1. Introduction

*Plasmodium falciparum* genetic variability, population structure and gene flow across regions are critical in malaria control and elimination strategies since they impact parasite genotypic and phenotypic characteristics. These characteristics include clonal diversity (Anderson et al., 2000), response to host immunity (Gupta and Anderson, 1999; Klein et al., 2008), virulence (Mackinnon and Read, 2004), and dynamics in the emergence and the spread of drug resistance traits (Ariey et al., 2013; Mehlotra et al., 2008; Mita et al., 2011). Differences in parasite genetic diversity and population structure may result from differences in the effective population size, the intensity of malaria transmission and the levels of inbreeding in low and high transmission regions (Anderson et al., 2000).

Neutral microsatellite markers sampled across the genome have been used to analyze parasites from different parts of the world. Parasites from regions with low transmission intensity have strong linkage disequilibrium (LD), low diversity, and extensive population differentiation (are highly structured). For instant, parasite populations in Southeast Asia have moderate to high levels of population structure and those in South America are clonal and highly structured (Anderson et al., 2000; Griffing et al., 2011; Mu et al., 2010). In contrast, parasites from high transmission intensity regions such as sub-Saharan Africa (SSA) have weak LD, high diversity, and low levels of differentiation (Anderson et al., 2000). However, exceptions have been reported in parasites from urban populations of Senegal, Niger and the Republic of Djibouti which have been described as having significant population structure, low levels of genetic diversity and significant LD (Bogreau et al., 2006). Significant LD has also been found in regions of high transmission and diversity in Zimbabwe, Republic of Congo (DRC), and Senegal (Anderson et al., 2000; Durand et al., 2003; Leclerc et al., 2002). It is yet to be determined

whether regions of high transmission intensity with low genetic differentiation have parasite populations that are genetically similar or identical with heritable traits.

Evidence suggests that transmission intensity is important in determining parasite genetic variability and population structure (Anderson et al., 2000; Susomboon et al., 2008). Other important factors include the movement of human and vector hosts (Schultz et al., 2010), the biogeographical features that may interrupt or promote the gene flow (Anderson et al., 2000; Joy et al., 2003) and the local malaria control measures (Anthony et al., 2005; Zhong et al., 2007). This underscores the importance of doing refined parasite population genetic studies at the local level, which can clarify the implications of parasite genetics in order to inform local control (Anthony et al., 2005; Schultz et al., 2010) and elimination strategies.

Based on the malaria risk map and the eco-epidemiology of malaria, Kenya is stratified into four malaria ecological regions: The endemic lake and coastal regions (risk class equal to or above 20%); epidemic-prone highland region (risk class 5 to less than 20%); seasonal transmission risk region (risk class less than 5%); and the low-risk region (risk class less than 0.1%). The endemic regions have high, stable malaria transmission whereas epidemic highlands are characterized by unstable transmission (Wanjala et al., 2011; Noor et al., 2009). Due to increased malaria interventions, some regions in Kenya such as the endemic coastal region have experienced significant decline in malaria transmission (Okiro et al., 2007; Okiro et al., 2009; O'Meara et al., 2008). Such changes may not only impact parasite genetic diversity, but also the need to modify control and intervention strategies to reflect the new transmission and disease rates. Genetic diversity and population structure of the parasite populations using multilocus microsatellite markers have mostly focused on the western part of Kenya (Bonizzoni et al., 2009; Gatei et al., 2010; Vardo-Zalik et al., 2012; Zhong et al., 2007). There is a gap in knowledge on

the *Plasmodium* genetic diversity and the population structure across the different malaria transmission regions of Kenya. This information is critical for tracking the spread of drug resistant strains in the parasite population, malaria transmission, control and eventual elimination. The goal of this study was to determine the genetic diversity and population structure of *P. falciparum* malaria parasite populations from endemic lake and coastal regions, and the epidemic-prone highland regions of Kenya, which differ in malaria transmission and disease endemicity.

## 2. Materials and Methods

### 2.1 Study sites and sample collection

The samples used in this study were collected from five different sites between 2012 and 2013 under the approved protocol titled “Epidemiology of malaria drug sensitivity patterns in Kenya.” This protocol was approved by the ethical review committee of the Kenya Medical Research Institute (KEMRI, Protocol # 1330), Nairobi, Kenya and Walter Reed Army Institute of Research (WRAIR, Protocol # 1384) Institutional Review Board (IRB), Silver Spring, MD. The samples were obtained from consenting patients presenting with uncomplicated malaria at the study site clinics, aged between 6 months and 65 years. Eligibility criteria included a history of fever (temperature of  $\geq 37.5^{\circ}\text{C}$ ) within 24 h prior to presentation, mono-infection with *P. falciparum* and a baseline parasitemia of 2,000-200,000 asexual parasites/mL. Persons treated for malaria within the preceding 2 weeks were excluded from the study. Written informed consent was obtained from adult subjects (>18 years of age) or legal guardians for subjects < 18 years of age. The presence of malaria was confirmed by microscopy and rapid diagnostic test (RDT; Parascreen®, Zephyr Biomedicals, Verna Goa, India). Whole blood was collected and aliquots

preserved for analysis as specified in the study protocols. The study subjects were treated with oral AL (Coartem) administered over three consecutive days, a standard of care for *P. falciparum* malaria in Kenya. On enrollment, the patients demographics, (including age and gender), place of birth, place of residence, occupation and travel history in the last 2 months was captured on a clinical data sheet. The study samples were collected from patients residing near the clinical collection locations (district hospitals) in Kisumu, Kombewa, Malindi, Kisii and Kericho. These sites were selected because they capture diverse malaria ecological zones of Kenya. Kisumu and Kombewa are in the endemic lake region and Malindi is in the coastal lowland region, both with stable high malaria transmission whereas Kisii and Kericho are epidemic-prone highland regions with unstable malaria transmission (Fig. 1).

### 2.2 Sample size calculation

Sample size was calculated using the Raosoft® sample size calculator formula (Raosoft) taking into consideration the incidence of malaria reduced to 15% as reported by Kenya malaria indicator survey 2010.

$$x = Z(c/100)2r(100-r)$$

$$E = \text{Sqrt} [(N - n) x/n (N-1)] = 4.63$$

Confidence level at 95%

$$n = N x / ((N-1) E^2 + x) = 271$$

The samples were collected from a study primarily designed to provide a description of the epidemiology of malaria and drug sensitivity pattern in Kenya. However, while considering that that each study site was able to enroll a minimum of 1 subject and a maximum of 3 subjects per week, a total of between 48 and 144 subjects per site per year were expected. A total of 909



samples were collected between 2012 and 2013 from the study sites Kisumu, Kisii, Kericho Malindi and Kombewa). Therefore, an average of 48 samples per site was randomly selected.

### *2.3 Plasmodium falciparum genotyping of microsatellite loci*

Genomic DNA was prepared from whole blood samples using the Qiagen DNA Mini extraction protocol (Qiagen, Valencia, CA), as recommended by the manufacturer. The samples were genotyped using a hemi-nested PCR strategy to amplify twelve microsatellite loci distributed across the *P. falciparum* genome as previously described by Anderson et al. (1999). The loci assayed were Poly $\alpha$  (Chromosome 4), TA42 (Chr5), TA81 (Chr5), TA1 (Chr6), TA109 (Chr6), TA87 (Chr6), TA40 (Chr10), 2490 (Chr10), ARAII (Chr11), pfG377 (Chr12), PfPk2 (12), and TA60 (Chr13). GENEMAPPER software Version 4.1 (Applied Biosystems, Foster City, CA) was used to automate measurement of allele length and to quantify peak heights. Microsatellite genotyping was used to identify infections containing >1 parasite clone in one or more loci. Isolates for which alleles were successfully amplified in at least 10 loci out of 12 were included in the analyses while those isolates in which alleles were amplified in fewer than 10 loci were excluded.

### *2.4 Assessment of Multiple Infections*

Multiple infections (MIs) were defined as those specimens in which one or more of the twelve loci showed multiple alleles. Data points (alleles) from samples that amplified poorly for particular loci (maximum peak height of 200 fluorescent units) were discarded. Multiple alleles per locus were scored if fluorescence peaks corresponding to minor alleles were  $\geq 33\%$  the height of that of the predominant allele in the isolate. Allele scoring was determined for each parasite clone where the greatest number of alleles amplified at any locus was considered the multiplicity of infection for that sample.

### 2.5 Estimation of genetic diversity and population structure

Genetic diversity was measured by the number of alleles per locus and expected heterozygosity on the basis of allele frequency data in each population. The allele frequencies were determined using the predominant allele at each locus within each infection. The predominant allele at each locus was defined as the highest peak height in electropherogram traces. The expected heterozygosity ( $H_e$ ) was measured at each locus in each of the five parasite locations as  $(1 - \sum p_i^2)$ , where  $p_i$  is the frequency of the  $i^{\text{th}}$  allele. In addition, the variance in allele size and the number of alleles ( $N_a$ ) at each locus was determined in the five populations. The level of genetic diversity, estimated effective population size and number of effective alleles ( $N_e$ ) in the different populations was compared using Kruskal-Wallis H test. The population subdivision was measured using Wright's  $F$  statistics and the genetic distance between parasite populations using Nei's method (1978) while the genetic relationships among individual parasite haplotypes was determined by counting the proportion of alleles shared among the 12-locus haplotypes ( $P_s$ ).

### 2.6 Constructing Multi-locus Microsatellite haplotypes and Linkage Disequilibrium

The median-joining network were constructed in NETWORK version 4.6.1.1 (<http://www.fluxus-engineering.com/sharenet.htm>) to visualize the relationships among *P. falciparum* isolates from different malaria transmission regions in Kenya. The predominant allele at each locus was used to construct the infection haplotypes. In addition, all the observed alleles with peak heights greater than 33% of the height of the predominant allele in each locus were used to construct haplotypes by permutation. This imposed additional recombination of the data in the construction of the infection haplotypes. We conducted analyses of both the complete data set with all the alleles and an abridged data set in which predominant alleles were used. A

permutation procedure was used to test the null hypothesis of random association among loci for each parasite population (Haubold et al., 1998; Souza et al., 1992). The program LIAN, version 3 (Haubold and Hudson, 2000), was used to compute the number of alleles shared (D) between all pairwise comparisons of complete 12-locus haplotypes and to measure the variance of this distance measure (VD). To investigate if the observed data differed from random expectations, the observed VD was compared with the distribution of VD values in 100,000 simulated data sets in which alleles at each locus were randomly reshuffled among genotypes. The index of association ( $I_A$ ) was used to measure the strength of LD. The standardized  $I_A$  statistic ( $I_{AS}$ ) was calculated as  $((V_D/V_e - 1)/(1 - r))$ , where  $r$  is the number of loci, a measure of haplotype-wide linkage and the 95% confidence limits determined by Monte Carlo simulations (LMC) (Hudson, 1994).

### 3. Results

#### 3.1 Genetic diversity

Complete allelic data was obtained from 188 of 225 (83.6%) malaria infected blood samples collected from the five sites. These included 140 samples with full genotypes from the twelve loci and 48 samples with allelic data from eleven loci. The number of samples analyzed per site ranged from 23 to 46, with Malindi having the lowest and Kericho having the highest (Table 1). Multiply-infected parasite samples were determined by the number of alleles per locus; parasites with more than one allele at any of the loci were considered to be polyclonal. All the samples analyzed in this study were polyclonal. Parasites were either analyzed by use of predominant allele at each locus or by combination of possible permutations for all alleles present.

Table 2 shows the genetic diversity of parasites from all the five regions in Kenya. Parasites were highly polymorphic with a mean number of alleles of 13.02. The Poly $\alpha$  locus was the most polymorphic with a mean of 18.40 alleles whereas 2490 and PfG377 were the least polymorphic loci with means of 8.20 and 8.60 alleles, respectively. Allelic diversity at each locus was summarized as the  $H_e$  from the distribution of allele frequencies while factoring in sample sizes of different populations. Parasites from Kenya showed high genetic diversity with the mean unbiased  $H_e$  of 0.81. The Poly $\alpha$  locus had the highest  $H_e$  (0.92) whereas PfG377 had the lowest (0.63). Table 1 shows the genetic diversity of isolates from each of the five sites. Levels of genetic diversity showed heterogeneity among the different malaria transmission regions in Kenya. Kisumu parasites showed the highest diversity whereas those from Malindi were the least diverse; the mean number of alleles in all the loci for Kisumu parasites was 14.75 whereas Malindi was 10.33. The difference in numbers of alleles between the different sites did not reach statistical significance. TA1 was the most polymorphic locus with 24 alleles in Kombewa parasites whereas TA42 in Malindi parasites and 2490 in Kisii parasite isolates were the least polymorphic with 5 alleles each. From the number of distinct alleles observed at each locus, the  $H_e$  estimated per locus is shown in Table 1. Kisumu parasites had the highest mean  $H_e$  (0.86) whereas Kisii had the lowest (0.79). Analysis of each the locus across the five sites showed Poly $\alpha$  had the highest  $H_e$  mean (0.92) whereas TA42 had the lowest (0.63). From regional and locus-specific analyses, the highest  $H_e$  was observed in TA1 among Kombewa parasites (0.95) whereas TA42 was lowest in Malindi parasites (0.37). Most alleles were distributed widely in all the five parasite populations, with Kisumu having the largest number of private alleles (5) compared to Malindi which had the least (1). Shannon diversity index (I) is a quantification measure of abundance and evenness of species present, which gives more information about the

composition of a population. Kisumu had the highest I value (2.3) compared to Malindi, which had the lowest (2), further confirming higher genetic diversity in Kisumu compared to Malindi.

### 3.2 Genetic differentiation

AMOVA analyses is performed using the mean of squares to compute the evolutionary distance between the haplotypes and describes the genetic distance into within and among populations. The AMOVA analysis revealed that most (93%) of the variance in allele frequencies was within the parasite populations whereas 7% of the variation was explained by differences amongst parasite populations from the different malaria transmission regions in Kenya (Table 3a). Tables 3b, c, and d show AMOVA analysis between specific sites. The largest within-haplotype genetic distance was observed between Kisumu and Malindi parasite populations (96%). In addition, Kisumu and Malindi had the smallest genetic difference among the parasite populations (4%) (Table 3b). The smallest variance in allele frequencies among individuals was observed in parasite populations from Kisii and Kericho (90%); these two sites also had the largest difference in genetic variation among parasite populations (10%) (Table 3c).

Genetic distance evaluates the degree of genetic differentiation achieved during the speciation process or any other stage of evolution divergence. Nei's genetic distance estimates the mean number of substitutions that occurred since two populations shared their last common ancestor. The pairwise population matrix values for the Nei's genetic distance ranged between 0.27 - 0.90 with Malindi vs. Kombewa having the least distance at 0.27 (parasites are more closely related) and Kisumu vs. Kisii having the largest value at 0.90 (Table 4).

Wright's fixation index is a measure of population differentiation caused by genetic structure. It measures the amount of genetic variation explained by population structure and provides insights into the evolutionary process that influences the structure of genetic variation

within and among populations. Wright's F-statistics revealed a mean genetic differentiation index ( $F_{ST}$ ) of 0.07 among the parasite populations for all the 12 MS loci (Table 2). TA60 locus had the largest  $F_{ST}$  (0.15) whereas Poly $\alpha$  had the lowest (0.03). The majority (8 of 12) of markers showed moderate genetic differentiation. The pairwise population values for the fixation index ranged between 0.027 - 0.067 with Kombewa vs. Malindi having the lowest  $F_{ST}$  whereas Kisii vs. Kericho had the largest (Table 5). Interpretation of the  $F_{ST}$  values at each locus was based on three categories defined earlier as no differentiation (0), low genetic differentiation (0 - 0.05), moderate differentiation (0.05 - 0.15) and great differentiation (0.15 - 0.25) (Balloux et al., 2002; Weir et al., 1984).

Multi-locus genetic analysis of the 12 MS showed that there were no matching haplotypes; a total of 188 unique haplotypes were generated (Fig. 2). The highest multi-locus combination was seen in two samples from Kisumu which had identical alleles at 9 of 12 loci while six samples (2 from Kisii, 2 from Malindi and 2 from Kericho) had identical alleles at 8 of 12 loci. A distance-based approach was used to determine the relationships among individual parasite haplotypes. Principal coordinate analysis revealed strong clustering of haplotypes by location; parasite isolates from Kisii were distinct from the rest of the regions in Kenya included in the study (Fig. 3). Furthermore, the network joining tree showed Kisii parasites grouping separately from other isolates (Fig. 4).

### 3.3 Linkage Disequilibrium

Analysis of multilocus index of association was performed using the Monte Carlo method (100,000 random re-sampling) to test the significance of LD both for the complete data set and after removal of the isolates which matched at 8 or more loci. Significant LD was observed in all the five populations from Kenya for the complete data set. Kisumu had the most significant

index of association values (0.16;  $p < 0.0001$ ) while Kisii had the least significant index of association values (0.03;  $p < 0.0001$ ) (Table 6).

#### 4. Discussion

This study showed Kenyan *P. falciparum* parasites have significantly high number of alleles and high expected heterozygosity. Further, parasites in the endemic lake lowland and the epidemic highland regions of western Kenya have comparable genetic diversity. These results corroborate those of previous studies (Bonizzoni et al., 2009; Gatei et al., 2010; Vardo-Zalik et al., 2012; Zhong et al., 2007). Interestingly, in the highlands, the  $H_e$  increased from 0.66-0.69 in 2005 (Zhong et al., 2007) to 0.79-0.80 in 2012-2013 (the current study). This coincides with the reported increase in malaria transmission in the highlands (Wanjala et al., 2011). Parasites from Malindi, the lowland coastal region of Kenya had the lowest parasite genetic diversity which coincides with reports of reduction of malaria infections and transmission in this region (Kalayjian et al., 2013; Okiro et al., 2007; Okiro et al., 2009; O'Meara et al., 2008).

Previous studies conducted in Kombewa using samples collected in 2005 by Zhong *et al.* (2007) and Vardo-Zalik *et al.* (2012) showed that parasites from this region had  $H_e$  values of 0.75 and 0.74, respectively. By 2008 in the same parasite population, the  $H_e$  had a slightly increased to 0.80 with a mean number of alleles of 13.50 (Vardo-Zalik et al., 2012). Our data show that about five years later, the genetic diversity of parasites in Kombewa remained unchanged with a  $H_e$  of 0.82 and mean number of alleles of 13.42. These findings are in-line with our recent data which showed parasites from western Kenya (collected after the introduction of artemether-lumefantrine (AL) in 2006) were significantly more diverse compared to those collected before the introduction of AL ( $p < 0.0001$ ) (Chebon et al., 2015: manuscript submitted

for publication). There are two hypotheses for this observation: First, the data may imply that transmission intensity in this region has remained unchanged in the last five years or has not dropped below a threshold value (parasite population bottleneck) that would result in reduction of genetic diversity. As a result, the genetic diversity in the parasite population has remained unchanged over the study period. Second, it may indicate that AL is not exerting as much pressure to the parasite population compared to chloroquine and SP. The genetic data in this and the previous studies (Vardo-Zalik et al., 2012) suggest that malaria transmission has remained unchanged over time, which suggests there is need for new or improved malaria interventions if elimination is to be attained. The current malaria control initiatives in Kombewa include insecticide treated bed net (ITN) usage, residual sprays and AL for case management of malaria illness.

In the last decade, there has been a sharp decline in hospital admissions for malaria infections and transmission in the coastal region of Kenya (Kalayjian et al., 2013; Okiro et al., 2007; Okiro et al., 2009; O'Meara et al., 2008). This decline is attributed to the expanded malaria control efforts including the ITN usage, and the adoption of AL as the first-line treatment for uncomplicated malaria (Amin et al 2007; Fegan et al., 2007; Gikandi et al., 2008). To the best of our knowledge, this study presents for the first time genetic evidence indicating a change in malaria transmission levels in the coastal region of Kenya. To conclusively verify our findings however, it is important to perform genetic analysis of parasite samples collected longitudinally from the coastal regions of Kenya, preferably before AL was adopted as the first-line treatment.

A statistically significant population structure was observed in parasites collected from the different ecological regions ( $F_{ST}$  0.0072,  $P = 0.011$ ). Previous studies reported low genetic differentiation in Kenyan parasite populations (Gatei et al., 2010; Vardo-Zalik et al., 2012;



Zhong et al., 2007). The observed moderate genetic differentiation between highland and lowland parasite populations is consistent with the notion of parasite introduction by human travel from the malaria-endemic lowland to the epidemic highland region (Bonizzoni et al., 2009; Vardo-Zalik et al., 2012). This indicates that there may be limited barriers to gene flow. This would possibly impact parasite migration through natural selection and genetic drift. In addition to human travel, population structure of malaria parasite populations may be affected by parasite life history, site-specific epidemiology, and malaria control measures such as ITNs and indoor residual spray.

Significant LD has implication on the spread of multi-locus drug resistance haplotypes with high levels of inbreeding, which increases their dispersal (Schultz et al., 2010). Significant LD was found in both highland and lowland populations in the present study. None of the significant LD involved loci located on the same chromosome, which suggested that these loci are only statistically linked and not physically linked. Of interest, parasite population from Kisumu was present with a substantially stronger index of association ( $IAS = 0.16$ ) than those reported in other African countries ( $IAS < 0.02$ ) with comparable malaria prevalence (Anderson et al., 2000). Studies conducted in Zimbabwe and the DRC, regions of high malaria transmission, described the parasite population as having high genetic diversity and significant LD (Anderson et al., 2000; Durand et al., 2003). This data suggests the occurrence of inbreeding even in areas with intense and perennial malaria transmission. The data obtained in our study is consistent with the trends observed in Zimbabwe and the DRC.

The regions in our study are connected by a vast network of roads, with extensive human migration and transit of goods among the different regions, especially in western Kenya. Human travel from the lowlands to the highlands of western Kenya has been shown to play an important

role in introducing parasites from one region to the other (Bonizzoni et al., 2009; Shanks et al., 2005). An isolation-by-distance model could not explain the observed structure; the parasite population separated by the largest distance (Kombewa versus Malindi which is separated by a distance of ~ 1000 kilometers) had the lowest genetic differentiation whereas parasites separated by shorter distances showed stronger genetic differentiation (Kisii versus Kericho or Kisumu, which are bordering each other and the sample collection sites are separated by ~ 100 kilometers). The study by Schultz *et al.* (2010) showed similar results, whereby genetic differentiation in Papua New Guinea parasite populations from different regions did not correlate with geographic distance. Vector densities and transmission intensity in the highlands are generally much lower than in the adjacent lowland areas (Minakawa et al., 2006). Kisii and Kericho are both highland regions located in western Kenya (elevation > 1,500 m), with comparable amounts of total annual rainfall (2,243 vs. 1718 mm, respectively) and mean annual temperature (21.0 °C vs. 18 °C, respectively) (Wanjala et al., 2011). Our data suggest genetic flow between these two regions is lower compared to the other regions. This is intriguing given that Kisii and Kericho share a border and there is as much human travel between these two regions as there is in other surrounding regions that show evidence of parasite genetic flow (Gatei et al., 2010; Vardo-Zalik et al., 2012; Zhong et al., 2007). To further elucidate our findings, a larger number of samples must be analyzed. It would also be important to capture the GPS coordinates of where the samples (patient) originate from.

Parasite populations from Kisii showed the strongest evidence for epidemic population structure (Fig. 3). Kisii is characterized by ecosystems that have epidemic hotspots with localized micro-epidemics (Wanjala et al., 2011). These localized micro-epidemics might be driven by small-scale fishing farming whereby abandonment of these farms has resulted in

increased numbers of malaria vectors (Howard and Omlin, 2008). Epidemic population structures have also been reported in South American and Southeast Asian populations of *P. falciparum* (Anderson et al., 2000). This might explain the unique parasite structure which could lead to parasite isolation and local selfing. Thus, a small increase in the abundance of vectors may lead to a significant malaria outbreak in the highlands.

## 5. Conclusion

This study shows the genetic diversity in Kenyan *P. falciparum* parasites remain high, with the lake lowland endemic and epidemic highland regions of western Kenya having comparable genetic diversity. However, the parasite genetic diversity is lower in Kenyan coastal lowland endemic region which can be attributed to reduction in malaria disease burden and transmission. Further, for the first time, we have shown the existence of Kenyan parasite populations with moderate genetic differentiation and significant LD. The data presented here suggest that there is panmixia which leads to sufficient genetic flow between the different regions of Kenya, likely driven by both human travel and local transmission. Interestingly, an isolation-by-distance model could not explain the observed structure because the genetic differentiation between different parasite populations did not correlate with the distance. Parasites from some of the highland regions seemed to be slightly isolated, which has led to stronger genetic differentiation. This information is important in helping guide the malaria control and elimination agenda as well as malaria vaccine development.

## Funding

This work was supported by the Armed Forces Health Surveillance Center, Division of Global Emerging Infections Surveillance and Response System Operations.

#### **Author Contributions Statement**

L.A.I executed the project, ran the assays, analyzed the data and wrote the manuscript. . J.C. ran the assays and analyzed the data. S.A.O. analyzed the data analysis and wrote the manuscript.

B.A. analyzed the data and reviewed the manuscript. E.K. conceived of the project, analyzed the data and wrote the manuscript.

#### **Acknowledgement**

We would like to thank all the patients, clinical and other support staff at the study sites. We would like to thank all our colleagues in the Malaria Drug Resistance Laboratory for their technical and moral support. The opinions and assertions contained herein are private opinions of the authors and are not to be construed as reflecting the views of the US Army Medical Research Unit-Kenya (USAMRU-K), the Department of the Army, the Department of Defense or the U.S. Government.

**Competing financial interests:** The authors declare no competing financial interests.

**Conflict of Interest:** None declared

## References

- Amin, A., Zurovacm D., Kangwana, B.B., Greenfield, J., Otieno, D.N., Akhwale, W.S., Snow, R.W., 2007. The challenges of changing national malaria drug policy to artemisinin-based combinations in Kenya. *Malar J.* 6, 72.
- Anderson, T.J., Haubold, B., Williams, J.T., Estrada-Franco, J.G., Richardson, L., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., Whitworth, J., Velez, I.D., Brockman, A.H., Nosten, F., Ferreira, M.U., Day, K.P., 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Mol. Biol. Evol.* 17, 1467–1482.
- Anthony, T.G., Conway, D.J., Cox-Singh, J., Matusop, A., Ratnam, S., Shamsul, S., Singh, B., 2005. Fragmented population structure of *Plasmodium falciparum* in a region of declining endemicity. *J. Infect. Dis.* 191, 1558-1564.
- Ariey, F., Witkowski, B., Amaratunga, C., et al. A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature.* 2014; 13:50–55.
- Balloux, F. and Lugon-Moulin N., 2002. The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 11, 155-165.
- 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, M.A., Puijalón, O., Rogier, C., 2006. Genetic diversity and structure of African *Plasmodium falciparum* populations in urban and rural areas. *Am. J. Trop. Med. Hyg.* 74, 953-959.
- Bonizzoni, M., Afrane, Y., Baliraine, F. N., Amenya, D.A., Githeko, A.K., Yan G., 2009. Genetic structure of *Plasmodium falciparum* populations between lowland and highland sites and antimalarial drug resistance in Western Kenya. *Infection, Genetics and Evolution* 9, 806-812.
- Durand, P., Michalakakis, Y., Cestier, S., Oury, B., Leclerc, M.C., Tibayrenc, M., Renaud, F., 2003. 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.* 68, 345-349.
- Eyase F.L., Akala H.M., Ingasia L., Cheruiyot A., Omondi A., Okudo C., Juma D., Yeda R., Andagalu B., Wanja E., Kamau E., Schnabel D., Bulimo W., Waters N.C., Walsh D.S., Johnson J.D., 2013. The role of *pfmdr1* and *pfcr1* in changing chloroquine, amodiaquine, mefloquine and lumefantrine susceptibility in western-Kenya *P. falciparum* samples during 2008–2011. *PLoS ONE* 8:e64299
- Fegan, G.W., Noor, A.M., Akhwale, W.S., Cousens, S., Snow, R.W., 2007. Effect of expanded insecticide-treated bednet coverage on child survival in rural Kenya: a longitudinal study. *Lancet* 370, 1035–1039.
- Gatei, W., Kariuki, S., William, H., Kuile, F., Terlouw, D., Howard-Philips, P., Nahlen, B., Gimnig, J., Lindblade, K., Wlaker, K., Hamel, M., Crawford, S., Williamson, J., Slutsker, L., Shi P.Y., 2010. Effects of transmission reduction by insecticide treated bed nets (ITNs) on parasite genetics population structure: I. The genetic diversity of *Plasmodium falciparum* parasites by microsatellite markers in western Kenya. *Malar. J.* 9, 353
- Gething, P., Elyazar, I., Moyes, C., Smith, D., Battle, K., Guerra, C., Patil, A., Tatem, A., Howes, R., Myers, M., George, DB., Horby, P., Wertheim, H., Price, R., Mueller, I., Baird, J., and Hay, S. 2012. A long neglected world malaria map: *Plasmodium vivax* endemicity in 2010. *PLoS Negl Trop Dis.* 6.

- Gikandi, P.W., Noor, A.M., Gitonga, C.W., Ajanga, A.A., Snow, R.W., 2008. Access and barriers to measures targeted to prevent malaria in pregnancy in rural Kenya. *Trop. Med. Int. Health* 13, 208–217.
- Griffing, S.M., Mixson-Hayden, T., Sridaran, S., Alam, M. T., McCollum, A.M., Cabezas, C., Marquino, Q.W., Barnwell, J.W., De Oliveira, A.M., Lucas, C., Arrospe, N., Escalante, A.A., Bacon, D.J., Udhayakumar, V., 2011. South American *Plasmodium falciparum* after the malaria eradication era: clonal population expansion and survival of the fittest hybrids. *PLoS ONE* 6:e23486.
- Gupta, S., and Anderson, R.M., 1999. Population structure of pathogens: the role of immune selection. *Parasitol Today* 12, 497–501.
- Hudson, R.R., 1994. Analytical results concerning linkage disequilibrium in models with genetic transformation and conjugation. *J. Evol. Biol.* 7, 535-548.
- Haubold, B., and Hudson R., 2000. LIAN version 3: A program for detecting linkage disequilibrium in multilocus data. *Bioinformatics*.
- Haubold, B., Travisano, M., Rainey, B.P., and Hudson, R.R., 1998. Detecting linkage disequilibrium in bacterial populations. *Genetics* 150, 1341–1348.
- Howard, A.F., and Omlin, F.X., 2008. Abandoning small-scale fish farming in western Kenya leads to higher malaria vector abundance. *Acta Trop.* 105, 67-73.
- Joy, D.A., Feng, X., Mu, J., Furuya, T., Chotivanich, K., Krettli, A.U., Ho, M., Wang, A., White, N.J., Suh, E., Beerli, P., Su, X.Z., 2003. Early origin and recent expansion of *Plasmodium falciparum*. *Science* 300, 318-321.
- Kalayjian, B.C., Malhotra, I., Mungai, P.L., Holding, P.A., King, C.H., 2013. Marked decline in malaria prevalence among pregnant women and their offspring from 1996 to 2010 on the South Kenyan coast. *Am. J. Trop. Med. Hyg.* 89, 1129–1134.
- Klein E.Y., Smith, D.L., Boni, M.F., Laxminarayan, R., 2008. Clinically immune hosts as a refuge for drug-sensitive malaria parasites. *Malar J* 7, 67.
- Leclerc M.C., Durand, P., de Meeus, T., Robert, V., Renaud, F., 2002. Genetic diversity and population structure of *Plasmodium falciparum* isolates from Dakar, Senegal, investigated from microsatellite and antigen determinant loci. *Microbes Infect.* 4, 685-692.
- Mackinnon, M.J., Read, A.F., 2004. Virulence in malaria: an evolutionary viewpoint. *Phil. Trans. R. Soc. Lond. B.* 359, 965–986.
- Marsh, K., Snow, R.W., 1999. Malaria transmission and morbidity. *Parasitologia* 41, 241–246.
- Mehlotra, R.K., Mattera, G., Bockarie, M.J., Maguire, J.D., Baird, J.K., Sharma, Y.D., Alifrangis, M., Dorsey, G., Rosenthal, P.J., Fryauff, D.J., Kazura, J.W., Stoneking, M., Zimmerman, P.A., 2008. Discordant patterns of genetic variation at two chloroquine resistance loci in worldwide populations of the malaria parasite *Plasmodium falciparum*. *Antimicrob. Agents Chemother.* 52, 2212-2222.
- Minakawa, N., Omukunda E., Zhou G., Githeko A., Yan g., 2006. Malaria vector productivity in relation to the highland environment in Kenya. *Am. J. Trop. Med. Hyg.* 75, 448–453
- Mita, T., Venkatesan, M., Ohashi, J., Culleton, R., Takahashi, N., Tsukahara, T., Ndounga, M., Dysoley, L., Endo, H., Hombhanje, F., Ferreira, M.U., Plowe, C.V., Tanabe, K., 2011. Limited geographical origin and global spread of sulfadoxine-resistant dhps alleles in *Plasmodium falciparum* populations. *J. Infect. Dis.* 204, 1980-1988.
- Mu J., Myers R.A., Jiang H., Liu S., Ricklefs S., Waisberg M., Chotivanich K., Wilairatana P., Krudsood S., White N.J., Udomsangpetch R., Cui L., Ho M., Ou F., Li H., Song J., Li G.,

- Wang X., Seila S., Sokunthea S., Socheat D., Sturdevant D.E., Porcella S.F., Fairhurst R.M., Wellems T.E., Awadalla P., Su X.Z., 2010. *Plasmodium falciparum* genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. *Nat. Genet.* 42, 268–271.
- Ndenga, B., Githeko, A., Omukunda, E., Munyekenye, G., Atieli, H., Wamai, P., Mbogo, C., Minakawa, N., Zhou, G., Yan, G., 2006. Population dynamics of malaria vectors in western Kenya highlands. *J. Med. Entomol.* 43, 200–206.
- Noor, A.M., Gething, P.W., Alegana, V.A., Patil, A.P., Hay, S.I., Muchiri, E., Juma, E., Snow, R.W., 2009. The risks of malaria infection in Kenya in 2009. *BMC Infect Dis.* 9, 180.
- Okiro, E.A., Hay, S.I., Gikandi, P.W., Sharif, S.K., Noor, A.M., Peshu, N., Marsh, K., Snow, R.W., 2007. The decline in paediatric malaria admissions on the coast of Kenya. *Malar J.* 6, 151.
- Okiro, E.A., Alegana, V.A., Noor, A.M., Mutheu, J.J., Juma, E., Snow, R.W., 2009. Malaria paediatric hospitalization between 1999 and 2008 across Kenya. *BMC Med.* 7, 75.
- Okiro, E.A., Alegana, V.A., Noor, A.M., Snow, R.W., 2010. Changing malaria intervention coverage, transmission and hospitalization in Kenya. *Malar J.* 9, 285.
- O’Meara, W.P., Bejon, P., Mwangi, T.W., Okiro, E.A., Peshu, N., Snow, R.W., Newton C.R., Marsh, K., 2008. Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. *Lancet* 372, 1555–1562.
- Razakandrainibe, F.G., Durand, P., Koella, J.C., De Meeus, T., Rousset, F., Ayala, F.J., Renaud, F., 2005. “Clonal” population structure of the malaria agent *Plasmodium falciparum* in high-infection regions. *Proc Natl Acad Sci USA.* 102, 17388–17393.
- Schultz, L., Wapling, J., Mueller, I., Ntsuke, P.O., Senn, N., Nale, J., Kiniboro, B., Buckee, C.O., Livingstone, T., Siba, P.M., Reeder, J.C., Barry, A.E. 2010. 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.* 9, 336.
- Shanks, G.D., Biomndo, K., Guyatt, H.L., Snow, R.W., 2005. Travel as a risk factor for uncomplicated *Plasmodium falciparum* malaria in the highlands of western Kenya. *Trans. Roy. Soc. Trop. Med. Hyg.* 99, 71–74.
- Snow, R.W., Omumbo, J.A., Lowe, B., 1997. Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. *Lancet* 349, 1650–1654.
- Souza, V., Nguyen, T.T., Hudson, R.R., Pinero, D., Lenski, R.E., 1992. Hierarchical analysis of linkage disequilibrium in *Rhizobium* populations: evidence for sex? *Proc. Natl. Acad. Sci. USA* 89:8389–8393.
- Susomboon, P., Iwagami, M., Tangpukdee, N., Krusood, S., Looareesuwan, S., Kano, S., 2008. Differences in genetic population structures of *Plasmodium falciparum* isolates from patients along Thai-Myanmar border with severe or uncomplicated malaria. *Malar J.* 7, 212.
- Vardo-Zalik, A.M., Zhou, G., Zhong, D., Afrane, Y., Githeko, A.K., Yan, G., 2012. Alterations in *Plasmodium falciparum* genetic structure two years after increased malaria control efforts in Western Kenya. *Am. J. Trop. Med. Hyg.* 88, 29–36.
- Wanjala, C.L., Waitumbi, J., Zhou, G., Githeko, A.K., 2011. Identification of malaria transmission and epidemic hotspots in the western Kenya highlands: its application to malaria epidemic prediction. *Parasite Vectors* 4, 81.

- Weir, B.S., Cockerham, C.C., 1984. Estimating F-Statistics for the analysis of parasite structure. *Evolution* 38, 1358-1370
- Zhong, D., Afrane, Y., Githeko, A., Yang, Z., Cui, L., Menge, D.M., Temu, E.A., Yan, G., 2007. *Plasmodium falciparum* genetic diversity in western Kenya highlands. *Am. J. Trop. Med. Hyg.* 77, 1043–1050

ACCEPTED MANUSCRIPT



**Figure caption**

**Fig. 1.** Spatial distribution of *P. falciparum* malaria in 2012 stratified by endemicity class in Kenya. The study sites are showed on the map with the different ecological zones.

**Fig. 2.** Summary of multilocus matches by locus for increasing combinations of the 12 loci for Kenyan *P. falciparum* isolates. A total of 188 unique parasites genotypes were constructed using the 12 multilocus haplotype combinations. The x-axis shows the number of combined genotypes (loci) and the y-axis shows the number of parasites carrying the combined genotype.

**Fig. 3.** The graphical presentation of the first two coordinates for 188 individual infections genotypes with the twelve MS. The eigen vectors and eigen values were used to plot the infection haplotypes of the parasite from different ecological regions in Kenya. Color codes shows subgroups infection partitioned by the different malaria endemic regions in Kenya.

**Fig. 4.** Median joining network diagram showing relationship of the parasites from the five different sites. The multilocus microsatellite haplotypes profiles were constructed for each of the parasite generation using the 12 MS. The 188 *P. falciparum* field isolates analyzed formed 188 unique 12-loci MS haplotypes. Each circle in the network represents a unique MS haplotypes with the size of the circle being proportional to the number of isolates showing that particular haplotypes. The grey dots are hypothetical median vectors generated by the software to connect existing haplotypes within the network with maximum parsimony

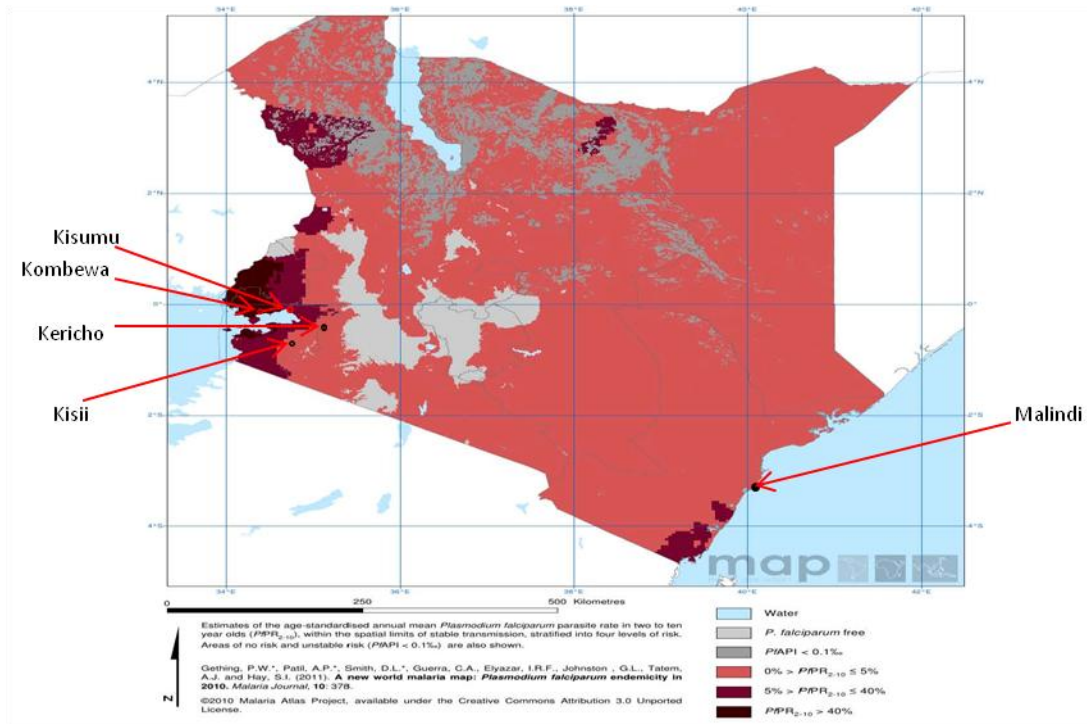


Fig. 1

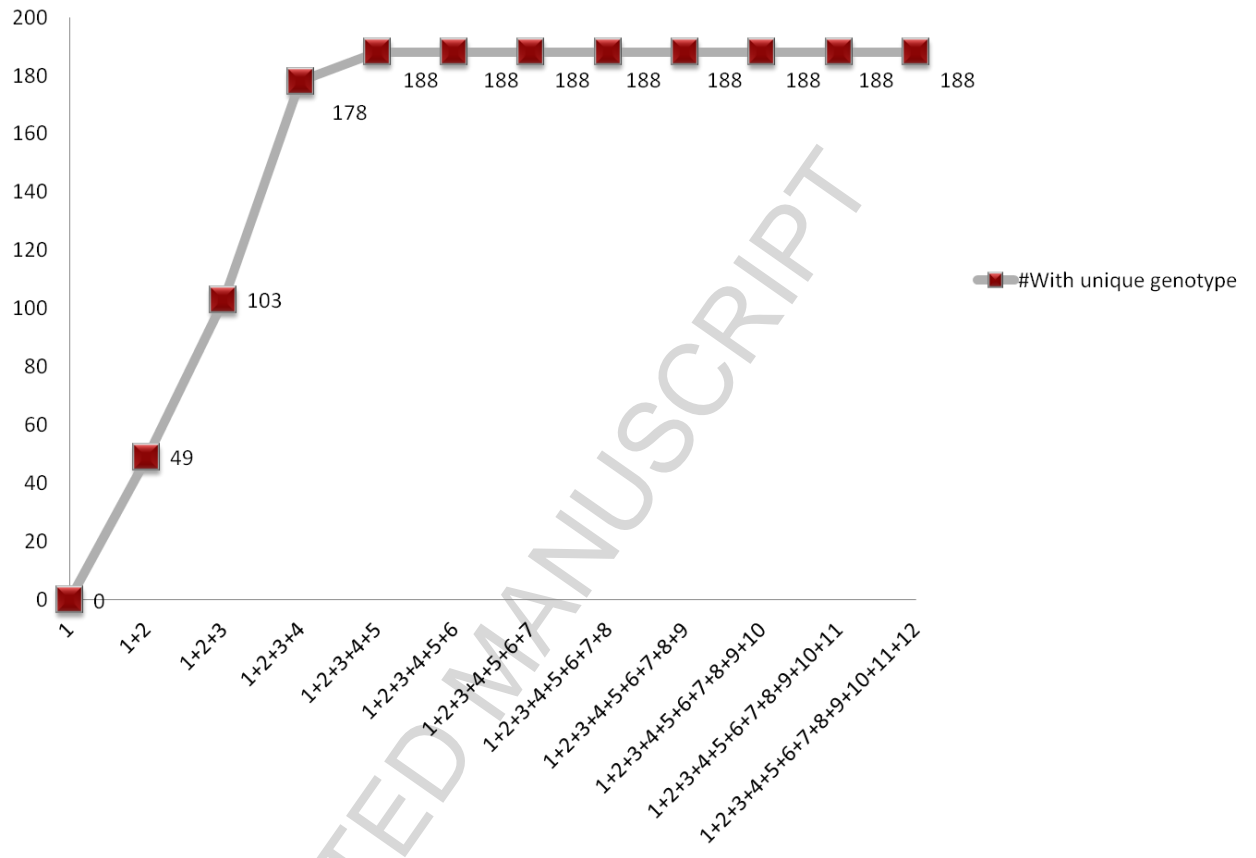


Fig. 2

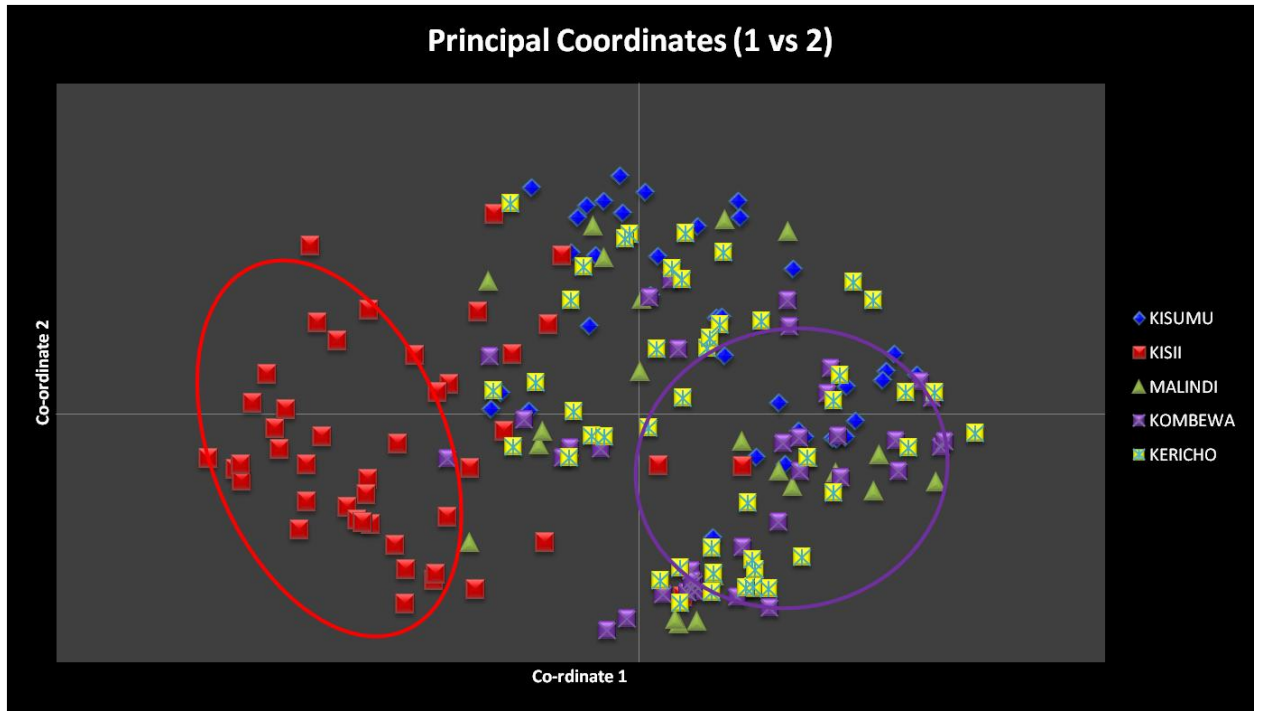


Fig. 3

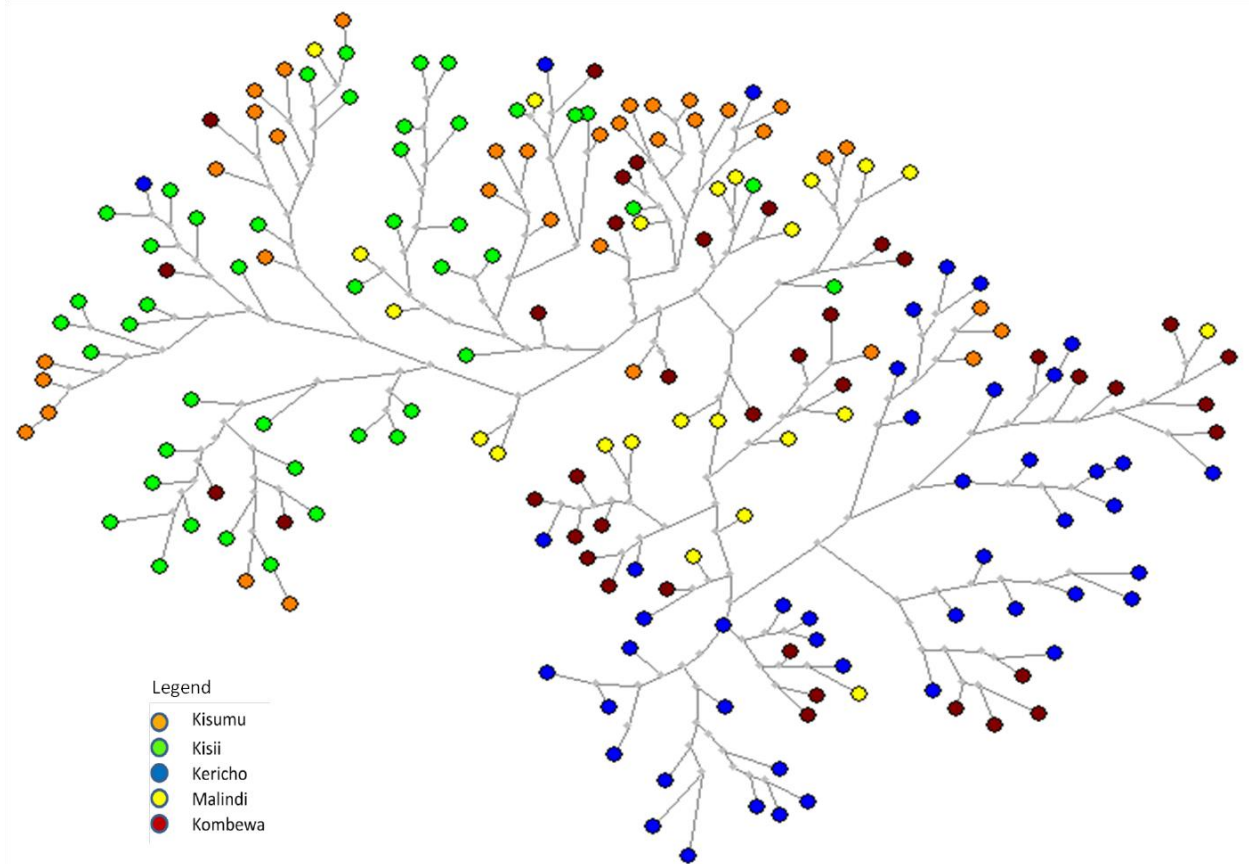


Fig. 4

Table 1: Genetic diversity of *P. falciparum* populations from different endemic regions (western and coastal regions) in Kenya

Locus	KISUMU N=37			KOMBEWA N=37			MALINDI N=23			KISII N=45			KERICHO N=46		
	Na	Ne	$H_e$	Na	Ne	$H_e$	Na	Ne	$H_e$	Na	Ne	$H_e$	Na	Ne	$H_e$
2490	12.0 0	5.26 1	0.8 1	8.00 0	2.74 5	0.6 4	7.00 0	4.17 6	0.7 6	5.00 0	1.95 9	0.4 9	9.00 0	3.68 6	0.7 3
ARA2	13.0 0	8.72 9	0.8 9	15.0 0	10.4 5	0.9 0	10.0 0	5.69 2	0.8 2	11.0 0	5.10 0	0.8 0	11.0 0	7.15 6	0.8 6
PFG3 77	8.00 0	3.41 0	0.7 1	8.00 0	1.95 1	0.4 9	6.00 0	2.35 6	0.5 8	11.0 0	6.60 5	0.8 5	10.0 0	3.35 0	0.7 0
PFPK 2	18.0 0	11.7 0	0.9 2	16.0 0	13.5 5	0.9 3	14.0 0	11.2 6	0.9 1	11.0 0	4.70 9	0.7 9	17.0 0	9.98 0	0.9 0
POLY	19.0 0	12.5 6	0.9 2	19.0 0	14.1 1	0.9 3	14.0 0	9.62 0	0.9 0	21.0 0	15.1 3	0.9 3	19.0 0	12.1 6	0.9 2
TA1	17.0 0	11.7 0	0.9 2	24.0 0	18.2 5	0.9 5	11.0 0	7.45 7	0.8 7	13.0 0	9.13 9	0.8 9	19.0 0	11.5 0	0.9 1
TA10 9	19.0 0	12.7 9	0.9 2	16.0 0	8.83 9	0.8 9	8.00 0	5.81 3	0.8 3	19.0 0	9.98 0	0.9 0	18.0 0	13.0 6	0.9 2
TA40	17.0 0	13.5 4	0.9 3	12.0 0	7.09 6	0.8 6	14.0 0	11.2 6	0.9 1	8.00 0	1.85 6	0.4 6	15.0 0	9.13 9	0.8 9
TA42	12.0 0	2.99 7	0.6 7	8.00 0	3.16 8	0.6 8	5.00 0	1.60 7	0.3 7	15.0 0	4.63 8	0.7 8	10.0 0	2.82 5	0.6 5
TA60	13.0 0	7.17 6	0.8 6	8.00 0	5.63 2	0.8 2	9.00 0	5.81 3	0.8 3	13.0 0	4.63 8	0.7 8	6.00 0	2.78 4	0.6 4
TA81	13.0 0	6.22 4	0.8 4	11.0 0	7.73 7	0.8 7	13.0 0	9.62 0	0.9 0	14.0 0	9.68 0	0.9 0	8.00 0	2.65 2	0.6 2
TA87	16.0 0	12.3 3	0.9 2	16.0 0	10.1 4	0.9 0	13.0 0	10.8 0	0.9 1	21.0 0	14.3 6	0.9 3	15.0 0	8.60 8	0.8 8
Mean	14.7 5	9.03 6	0.8 6	13.4 2	8.64 2	0.8 2	10.3 3	7.12 0	0.8 0	13.5 0	7.31 9	0.7 9	13.0 8	7.24 0	0.8 0
SE	0.99 3	1.12 3	0.0 3	1.49 4	1.44 4	0.0 4	0.96 5	0.98 5	0.0 5	1.42 5	1.27 5	0.0 5	1.33 4	1.16 4	0.0 4

sample size,  $N_a$  = number of alleles,  $U_{He}$  = unbiased heterozygosity.

The table show the genetic diversity for Isolates collected in 2012-2013 from 5 regions in Kenya (western part; Kisumu, Kisii, Kericho, Kombewa and coastal region; Malindi) \* $N_a$  = number of different alleles,  $N_e$  = number of effective alleles,  $H_e$  = biased expected heterozygosity

ACCEPTED MANUSCRIPT

Table 2: Genetic diversity of Kenyan *Plasmodium falciparum* populations

	Na		Ne		H <sub>E</sub>		F <sub>ST</sub>	Nm
	Mean	SE	Mean	SE	Mean	SE		
<b>2490</b>	8.20	1.16	3.56	0.57	0.68	0.06	0.092	2.46
<b>ARA2</b>	12.00	0.89	7.42	0.98	0.86	0.02	0.069	3.40
<b>PFG377</b>	8.60	0.87	3.53	0.82	0.66	0.06	0.057	4.14
<b>PFPK2</b>	15.20	1.24	10.24	1.50	0.89	0.03	0.048	4.95
<b>POLY</b>	18.40	1.17	12.72	0.94	0.92	0.01	0.033	7.42
<b>TA1</b>	16.80	2.29	11.61	1.84	0.91	0.01	0.051	4.69
<b>TA109</b>	16.00	2.07	10.10	1.34	0.89	0.02	0.048	4.92
<b>TA40</b>	13.20	1.53	8.57	1.99	0.81	0.09	0.115	1.93
<b>TA42</b>	10.00	1.70	3.04	0.48	0.63	0.07	0.045	5.30
<b>TA60</b>	9.80	1.39	5.21	0.73	0.79	0.04	0.153	1.39
<b>TA81</b>	11.80	1.07	7.18	1.30	0.83	0.05	0.119	1.85
<b>TA87</b>	16.20	1.32	11.25	0.98	0.91	0.01	0.041	5.89
<b>Mean</b>	13.02		7.87		0.81		0.072	4.03
<b>SE</b>	0.57		0.53		0.02		0.011	0.53

The table show the genetic diversity of Kenyan Isolates collected in 2012-2013 from 5 regions in Kenya \*Na = number of different alleles, Ne = number of effective alleles, He = biased expected heterozygosity, F<sub>ST</sub> = wrights fixation index, Nm = number of migrants



Table 3a: Results of analysis of Molecular Variance for genetic variation in *P. falciparum* populations from Kenya

Source of Variation	Degree of freedom	Sum of squares	Mean Squares	Est. Variance component	% variation	Number of migrants
Among Populations	4	148.50	37.12	0.36	7%	
Among Individual	183	1840.56	10.06	5.03	93%	3.45
Within Individual	188	0.00	0.000	0.00	0%	
Total	375	1989.04		5.39	100%	

Table 3b: Results of analysis of Molecular Variance for genetic variation in *P. falciparum* populations from the low land malaria endemic areas in western Kenya and coastal region (Kisumu and Malindi)

Source of Variation	Degree of freedom	Sum of squares	Mean Squares	Est. Variance component	% variation	Number of migrants
Among Pops	1	21.89	21.85	0.20	4%	
Among Individual	58	602.59	10.39	5.20	96%	6.43
Within Individual	60	0.00	0.000	0.00	0%	
Total	119	624.43		5.40	100%	

Table 3c: Results of analysis of Molecular Variance for genetic variation in *P. falciparum* populations from the highland areas in western Kenya (Kisii and Kericho)

Source of Variation	Degree of freedom	Sum of squares	Mean Squares	Est. Variance component	% variation	Number of migrants
Among Pops	1	61.37	61.37	0.57	10%	

Among Individual	89	873.42	9.81	4.91	90%	2.17
Within Individual	91	0.00	0.00	0.00	0%	
Total	181	934.79		5.47	100%	

Table 3d: Results of analysis of Molecular Variance for genetic variation in *P. falciparum* populations from the low land malaria endemic areas in western Kenya (Kisumu and Kombewa)

Source of Variation	Degree of freedom	Sum of squares	Mean Squares	Est. Variance component	% variation	Number of migrants
Among Pops	1	31.08	31.08	0.28	5%	
Among Individual	72	746.87	10.37	5.19	95%	4.63
Within Individual	74	0.00	0.00	0.00	0%	
Total	147	777.95		5.47	100%	

Table 4: Pair wise population matrix of Nei's Genetic distance

	KISUMU	KISII	MALINDI	KOMBEWA	KERICHO
KISUMU	0.00				
KISII	0.90	0.000			
MALINDI	0.45	0.58	0.00		
KOMBEWA	0.56	0.68	0.27	0.00	
KERICHO	0.64	0.82	0.47	0.40	0.00

The table shows the pair wise Nei's genetic distance for Isolates collected in 2012-2013 from 5 regions in Kenya (western part; Kisumu, Kisii, Kericho, Kombewa and coastal region; Malindi) The Nei's genetic distance values are below the diagonal.

Table 5: Pair wise population matrix of Wrights Fixation index ( $F_{ST}$ )

Pair wise Population  $F_{ST}$  Values

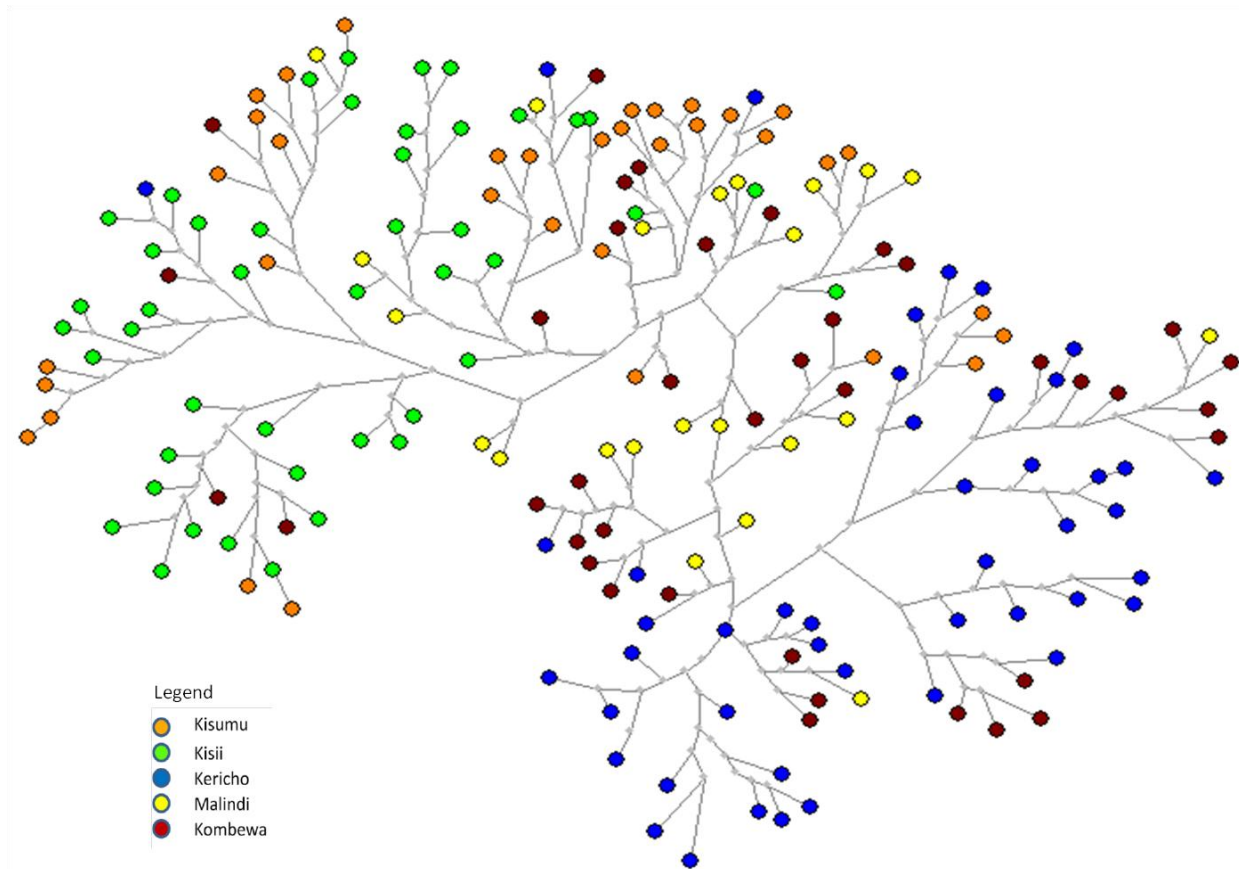
	KISUMU	KISII	MALINDI	KOMBEWA	KERICHO
KISUMU	0.000				
KISII	0.061	0.000			
MALINDI	0.036	0.055	0.000		
KOMBEWA	0.040	0.058	0.027	0.000	
KERICHO	0.047	0.067	0.044	0.037	0.000

P=0.011

Table 6: Results of analysis of Linkage disequilibrium in *P. falciparum*

Source of Variation	Kenyan Parasites	Kisumu	Kericho	Kombewa	Kisii	Malindi
$V_D$	<b>1.41</b>	<b>3.38</b>	<b>2.22</b>	<b>2.03</b>	<b>1.88</b>	<b>2.19</b>
$V_e$	<b>1.08</b>	1.24	1.42	1.46	1.41	1.43
$I_A S$	<b>0.03</b>	0.16	0.05	0.04	0.03	0.05
$P$	<0.01	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001

## Graphical Abstract



The network joining tree showed parasites grouping. Kisii (lime green) parasites separate from other isolates.

## Highlights

- The mean genetic differentiation index ( $F_{ST}$ ) showed a moderate differentiation among the parasite populations (0.072).
- Multi-locus genetic analysis of 12 MS showed that there were no matching haplotypes in Kenya parasite populations.
- Significant linkage disequilibrium (LD) was observed in all the five parasite populations.
- Kisumu had the most significant index of association values (0.16;  $p < 0.0001$ ) whereas Kisii had the least significant index of association values (0.03;  $p < 0.0001$ ).
- Significant LD observed in Kenyan parasites suggests the occurrence of inbreeding in high transmission areas.