Genetic Diversity and Population Structure of Plasmodium Falciparum in Nigeria: Insights From Microsatellites Loci Analysis

Fehintola Victoria Ajogbasile
Reedeemer's University College of Natural Sciences  https://orcid.org/0000-0002-3149-3085

Adeyemi T. Kayode
Reedeemer's University College of Natural Sciences

Kazeem O. Akano
Reedeemer's University College of Natural Sciences

Paul E. Olunyi
Reedeemer's University College of Natural Sciences

Jessica N. Uwanibe
Reedeemer's University College of Natural Sciences

Benjamin B. Adegboyega
African Centre for Excellence for the Genomics of Infectious Diseases (ACEGID)

Courage Philip
African Centre of Excellence for the Genomics of Infectious Diseases (ACEGID)

Oluwagboadurami G. John
Covenant University

Philomena J. Eromon
Africa Centre of Excellence for the Genomics of Infectious Diseases (ACEGID)

George Emechebe
Imo State University

Finimo Finimo
Federal Medical Centre Yenogoa

Nnenna Ogbulaofo
National Malaria Elimination Program: Federal Ministry of Health

Nma Jiya
Uthman Dan Fodio University, Sokoto

Uche Okafor
University of Nigeria, Nsukka

Jose Ambe
University of Maiduguri

Robinson D. Wammanda
Ahmadu Bello University

Stephen Oguche
University of Jos

Olugbenga A. Mokuolu
University of Ilorin

Akinbode Arowojolu
University of Ibadan College of Medicine

Onikepe A. Folarin
Reedeemer's University College of Natural Sciences

Christian T. Happi (happic@run.edu.ng)
African Centre of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer's University, Ede, Nigeria  https://orcid.org/0000-0002-3056-6705

Research

Keywords: Malaria, Plasmodium falciparum, Genetic diversity, Microsatellite, Nigeria

DOI: https://doi.org/10.21203/rs.3.rs-90709/v1
Abstract

Background

Malaria remains a public health burden especially in Nigeria. To develop new malaria control and elimination strategies or refine existing ones, understanding parasite population diversity and transmission patterns is crucial.

Methods

In this study, we characterized parasite diversity and structure of Plasmodium falciparum isolates from 633 dried blood spot samples in Nigeria, using 12 microsatellite loci of P. falciparum. These microsatellites were amplified via semi-nested polymerase chain reaction (PCR) and fragments were analyzed using GeneMapper and GENALEX 6.5.

Results

Estimates of parasite diversity such as Mean complexity of infection (range: 1.71-2.66) and Expected heterozygosity (range: 0.76-0.82) were high, while parasite population sub-structuring was low (Analysis of molecular variance= 0.039, Fixation index= 0.038 and Linkage disequilibrium= 0.0219).

Conclusion

We conclude that the high level of genetic diversity and low population structuring in this study suggests that parasite populations circulating in Nigeria are homogenous. This implies that a uniform control strategy will be effective across the six geographical zones of Nigeria. The results obtained can be used as a baseline for parasite diversity and structure, aiding in the formulation of appropriate therapeutic and control strategies in Nigeria.

Background

Although the incidence of malaria infections and malaria-associated mortality has reduced in many African countries [1, 2, 3], transmission continues in endemic regions despite intensified efforts towards prevention, control and eradication [4, 5]. This is due, in part, to the high genetic diversity of Plasmodium falciparum that contributes to increased transmission rate and spread of resistant parasites [6]. Therefore, understanding the extent of genetic diversity, transmission intensity, and parasite population structure in Nigeria - the most malaria burdened country, is essential if the goal of malaria control or elimination is to be achieved.

Molecular techniques play important roles in the analyses of genetic diversity, transmission dynamics, and population structure of P. falciparum field isolates. Early molecular studies focused mostly on the use of polymorphic markers such as merozoite surface protein 1 (msp-1) and merozoite surface protein 2 (msp-2) and glutamate-rich protein (Glurp) to characterise falciparum genetic diversity and structure in Nigeria [7, 8, 9]. These markers were also useful in monitoring drug efficacy with regards to classification of recurrent falciparum parasitaemia as re-infection or recrudescent infection [6, 10, 11]. However, there have been contrasting reports of polymorphisms in msp-1 and msp-2 in earlier studies in Nigeria [6, 12, 13, 14] which is associated with the fact that these antigenic markers are often under intense immune pressure [15, 16, 17]. The genotyping results provided by these markers can therefore potentially lead to a masked and distorted view of the population structure and transmission patterns which may account for observed variations across parasite populations circulating in a given environment [6].

Microsatellites have been suggested to be better alternatives to msp-1, msp-2 and GLURP due to their abundance, putative neutrality and higher levels of polymorphisms [18]. This molecular technique remains one of the most efficient and reliable methods for analyzing the genetic diversity data of falciparum populations for epidemiological and drug efficacy purposes within countries and across continents [19]. In past studies of using microsatellite analyses, it was observed that parasites from areas of low malaria transmission [19] (<1% infection) have less genetic diversity but more population structure and greater linkage disequilibrium (i.e. more non-random association among alleles across multiple loci) [4, 19, 20, 21]. Contrary, in regions of high malaria transmission, individuals are more likely to be infected by more than one P. falciparum parasite thereby resulting in an increase in the rate of recombination and subsequently, high diverse population with low linkage disequilibrium [18, 19, 22]. Although, some studies report a deviation from the norm whereby high levels of heterozygosity (a measure of genetic diversity) is observed in several low transmission countries [18, 23, 24]. This suggests that a high level of heterozygosity may reflect past human demographic processes as opposed to recent epidemiological factors [25].

The objective of this study was to investigate the genetic diversity of circulating Plasmodium falciparum parasites and their population structures in Nigerian children 6-96months old with uncomplicated infections, treated with artemisinin-based combination therapies (ACTs).

Methods

Study site

Filter papers containing dried blood spot (DBS) obtained from 633 children from nine Nigerian States covering all six geographical zones i.e., North-east: Adamawa State (n = 48), North-west: Sokoto State (n = 50), North-central: Kano (n = 100), Plateau (n = 100) and Kwara States (n = 58), South-west: Oyo State (n = 50), South-south: Bayelsa State (n = 45), and South-east: Enugu State (n = 100), Imo State (n = 82) with P falciparum malaria
infections on Day 0 were randomly selected for this study. These States are parts of sentinel sites for the National Malaria Elimination Program of the Federal Ministry of Health in Nigeria for the year 2014-2018.

Study site

Filter papers containing dried blood spot (DBS) obtained from 633 children from nine Nigerian States covering all six geographical zones i.e., North-east: Adamawa State (n = 48), North-west: Sokoto State (n = 50), North-central: Kano (n = 100), Plateau (n = 100) and Kwara States (n = 58), South-west: Oyo State (n = 50), South-south: Bayelsa State (n = 45), and South-east: Enugu State (n = 100), Imo State (n = 82) with P. falciparum malaria infections on Day 0 were randomly selected for this study. These States are parts of sentinel sites for the National Malaria Elimination Program of the Federal Ministry of Health in Nigeria for the year 2014-2018.

Sample collection

Two to three drops of finger-pricked blood samples were blotted on 3mm Whatman filter paper (Whatman International Limited, Maidstone, United Kingdom) before treatment initiation (Day 0). The blood samples impregnated on to filter papers were allowed to air-dry properly at room temperature, and dry blood spots (DBS) were kept in airtight envelopes with silica gel until analysed.

DNA extraction

DNA was extracted from DBS for parasite genetic diversity and population structure studies as previously described [26]. DNeasy Blood and Tissue extraction kit (Qiagen, Germany) was used to extract parasite DNA from DBS following the manufacturer's protocol.

*Plasmodium falciparum* genotyping by microsatellite loci analysis

Semi-nested PCR amplification of 12 *P. falciparum* microsatellite loci was done using a previously described protocol [17]. The 12 microsatellites loci were Poly A, PfG377, TA81, ARA2, TA87, TA40, TA42, 2490, TA1, TA60, TA109 and PfPk2 [27]. FAM, YAK YELLOW, and ATTO550N-labeled PCR products for the different loci amplified were pooled together (GeneScan™ 500, Applied Biosystems, Foster City, CA) for electrophoresis on the ABI 3500XL Genetic Analyzer at the African Centre of Excellence for the Genomics of Infectious Diseases (ACEGID), Redeemer's University Ede, Osun State, Nigeria. Peakscanner (Applied Biosystems) and GeneMarker (Softgenetics) software were used for normalization across runs and automatic determination of allele length and peak heights in samples containing multiple alleles per locus. Minor alleles were scored when the minor peaks were ≥ 20% the height of the predominant allele in the isolate and with a relative fluorescent unit of at least 100.

Data analysis

Measures of parasite genetic diversity

Complexity of infection (COI)

The COI was computed as the number of alleles per microsatellite loci divided by the number of amplified samples per microsatellite loci. The mean COI per State was calculated as the average of all COI values per microsatellite loci.

Parasite allelic frequency

The allele frequencies, per locus were calculated using GENALEX 6.5 [28]. This frequency was calculated for each State involved in the study.

Parasite allelic diversity

The expected heterozygosity (He), which represents the probability of being infected by two parasites with different alleles at a given locus, was calculated using the formula:

\[
He = \frac{(n/n-1) (1-\Sigma p^2)}{\frac{1}{2}} \text{- Equation 1}
\]

Where n is the number of isolates analysed, and p represents the frequency of each different allele at a locus [29]. He values range from 0 to 1. Values closer to 0 indicate little or no allelic diversity while values closer to 1 indicate high allelic diversity.

Measures of parasite population differentiation

Analysis of molecular variance (AMOVA)

Inter- and intra-population variance was determined with analysis of molecular variance (AMOVA, i.e., ΦPT). ΦPT value of zero (0) is considered indicative of no genetic differentiation among populations.

Fixation index (Fst)
The population divergence was measured by calculating the fixation index (Fst) for all pairs of parasite population in each State. The software, GENALEX 6.5 was used to compute the Fst value.

Principal component analysis (PCA)

Principal component analysis (PCA) was performed with the online program, ClustVis [30] across all nine States and separately for each State.

Linkage disequilibrium (LD)

Linkage disequilibrium (LD) was calculated for all nine States and separately for each State using the standardized index of association, \(\phi^2_A\) (LIAN version 3.5 web interface) [31] and the majority allele at each locus in each infection. This index was calculated as

\[
\phi^2_A = \frac{1}{n} - 1 \left( \frac{\text{VD}}{\text{VE}} - 1 \right) \quad \text{Equation 2}
\]

Where VE is the expected variance of the \(n\)th number of loci for which two individuals differ. VD is the observed variance. Randomization test, previously described [32], was done to determine whether the ratio of VD/VE was significantly higher than 1.

Results

Demographics and baseline characteristics

Overall, 329 (51.97%) were male and the mean age of all children included in the study was 48.4±15.8 months. Also, mean enrollment body temperature was 37.5±2.5°C. Overall geometric asexual parasitemia was 16,219/L\(^{-1}\) (range: 2003-198200).

Parasite genetic diversity

Complexity of infection (COI) and Mean complexity of Infection (mCOI)

The mCOI recorded from the nine Nigerian States ranged from 1.71 in Sokoto (lowest) to 2.66 in Oyo (highest) (Table 1). The COI obtained using the TA42 locus across all States was low (<1.28) while at loci such as Poly A and ARA 2, high COI values (≥2.16) were recorded (Table 1). Although mCOI observed in most (4 of 5) States in the Northern region were < 2.0 and in most (3 of 4) States in the Southern region were >2.0, there was no significant difference in mCOI values obtained in these regions (p>0.05).

Table 1: The COI and mCOI values of parasites in nine Nigerian States using 12 Microsatellites loci

| Loci  | Adamawa (n=48) | Bayelsa (n=45) | Enugu (n=100) | Ibadan (n=50) | Imo (n=82) | Kano (n=100) | Kwara (n=56) | Sokoto (n=46) | Plateau (n=100) |
|-------|----------------|---------------|---------------|---------------|------------|--------------|--------------|---------------|----------------|
| Poly A| 2.63           | 3.02          | 3.14          | 2.36          | 2.28       | 2.25         | 2.34         | 2.16          | 3.19           |
| PfPK2 | 1.92           | 3.2           | 2.4           | 1.8           | 1.68       | 1.83         | 1.75         | 1.58          | 2.53           |
| Ta81  | 1.88           | 1.78          | 4.17          | 2.9           | 1.69       | 1.64         | 2.1          | 1.81          | 2.52           |
| ARA2  | 5.05           | 4.57          | 6.81          | 4.14          | 4.89       | 2.41         | 4.43         | 3.7           | 6.45           |
| TA40  | 1.5            | 1.25          | 2.37          | 2.38          | 1.24       | 1.17         | 1.2          | 1.13          | 1.99           |
| TA87  | 1.78           | 1.38          | 2.25          | 8.9           | 1.48       | 1.49         | 1.55         | 1.68          | 2.46           |
| 2490  | 1.21           | 1.27          | 1.78          | 1.17          | 1.2        | 2.02         | 1.36         | 1.48          | 1.64           |
| TA1   | 1.7            | 1.84          | 3.01          | 2.04          | 1.74       | 2.35         | 1.86         | 1.56          | 2.29           |
| TA42  | 1.1            | 1.2           | 1.28          | 1.11          | 1.17       | 1.12         | 1.03         | 1.13          |                |
| PFG377| 1.44           | 1.52          | 1.7           | 1.63          | 1.3        | 1.47         | 1.33         | 1.37          | 1.36           |
| TA109 | 1.95           | 1.93          | 1.96          | 2             | 1.45       | 1.93         | 1.41         | 1.52          | 2.11           |
| TA60  | 1.45           | 1.48          | 1.92          | 1.33          | 1.39       | 1.62         | 1.62         | 1.44          | 2.02           |
| mCOI  | 1.97           | 2.04          | 2.73          | 2.66          | 1.79       | 1.78         | 1.84         | 1.71          | 2.47           |

Number of alleles, Allelic frequency and diversity
The mean number of alleles (Na) observed ranged from 9.75 (Adamawa State) to 15.67 (Enugu State) and the mean number of effective alleles (Ne) observed ranged from 5.4 (Oyo State) to 8.2 (Enugu State) (Table 2). The allelic frequencies at each of the 12 loci in parasite populations obtained in each of the nine States are presented in Table 2. The highest allelic frequencies in all nine States were observed in microsatellites loci TA40 (ranged from 200-220), TA42 (ranged from 182-183), PfPK2 (ranged from 161-170), TA81 (ranged from 160-169), TA109 (ranged from 153-174), TA1 (ranged from 153-161), Poly A (ranged from 103-159), PfG377 (103), TA87 (ranged from 90-106), 2490 (ranged from 80-83), ARA2 (ranged from 52-55), and TA60 (ranged from 52-55).

Table 2: Number of different and effective alleles parasite populations according to study location

| State     | Adamawa (n=48) | Bayelsa (n=45) | Enugu (n=100) | Ibadan (n=50) | Imo (n=82) | Kano (n=100) | Kwara (n=56) | Sokoto (n=46) | Plateau (n=100) |
|-----------|---------------|---------------|--------------|-------------|-----------|-------------|------------|-------------|---------------|
| Locus     | Na  | Ne  | Na  | Ne  | Na  | Ne  | Na  | Ne  | Na  | Ne  | Na  | Ne  | Na  | Ne  | Na  | Ne  | Na  | Ne  |
| PolyA     | 24  | 17.6| 20  | 8.3 | 31  | 28.6| 21  | 16.6| 25  | 12.6| 33  | 19.6| 23  | 17.6| 21  | 16.6| 21  | 13.6|
| PfPK2     | 12  | 7.4 | 14  | 6.4 | 13  | 6.4 | 11  | 5.8 | 11  | 5.8 | 13  | 6.4 | 10  | 5.8 | 10  | 6.4 | 10  | 5.8 |
| TA81      | 11  | 6.9 | 15  | 6.5 | 19  | 6.5 | 15  | 6.5 | 13  | 6.5 | 16  | 6.5 | 10  | 6.5 | 12  | 6.5 | 10  | 6.5 |
| ARA2      | 9   | 7.8 | 11  | 5.8 | 17  | 8.2 | 14  | 7.8 | 16  | 7.8 | 19  | 8.2 | 15  | 7.8 | 15  | 7.8 | 14  | 7.8 |
| TA87      | 11  | 8.3 | 14  | 7.4 | 22  | 11.4| 14  | 8.4 | 24  | 12.4| 13  | 8.4 | 18  | 12.4| 16  | 11.4| 24  | 14.4|
| TA40      | 5   | 3.1 | 8   | 3.8 | 10  | 5.6 | 6   | 3.8 | 9   | 5.6 | 10  | 5.6 | 4   | 2.8 | 10  | 5.6 | 6   | 3.8 |
| TA42      | 6   | 1.7 | 4   | 1.9 | 11  | 5.9 | 9   | 2.9 | 2   | 1.3 | 11  | 5.9 | 5   | 2.3 | 11  | 5.9 | 4   | 1.9 |
| 2490      | 5   | 2.6 | 6   | 3.8 | 7   | 3.8 | 3   | 2.8 | 10  | 4.2 | 8   | 3.8 | 10  | 4.2 | 8   | 3.8 | 10  | 4.2 |
| TA1       | 10  | 5.8 | 13  | 5.8 | 24  | 14.8| 7   | 3.8 | 13  | 4.2 | 19  | 11.8| 12  | 5.8 | 15  | 5.8 | 13  | 5.8 |
| PFPG377   | 5   | 3   | 5   | 3.8 | 8   | 3.8 | 9   | 3.8 | 6   | 3.8 | 7   | 3.8 | 8   | 3.8 | 6   | 3.8 | 8   | 3.8 |
| TA109     | 10  | 4.8 | 11  | 5.8 | 14  | 5.8 | 10  | 5.8 | 14  | 5.8 | 14  | 5.8 | 10  | 5.8 | 10  | 5.8 | 10  | 5.8 |
| TA60      | 9   | 5.1 | 6   | 3.8 | 10  | 5.8 | 11  | 4.2 | 8   | 3.8 | 9   | 4.2 | 9   | 3.8 | 11  | 4.2 | 8   | 3.8 |
| Mean      | 9.8 | 6.2 | 10.6| 5.4 | 15.7| 8.2 | 10.7| 5.4 | 12.5| 6.2 | 14.2| 7.1 | 11.3| 6.1 | 14.7| 6.9 | 10.8| 6.1 |
| SE        | 1.5 | 1.2 | 1.4 | 1.1 | 2.1 | 1.6 | 1.2 | 1.1 | 1.8 | 1.1 | 1.2 | 1.5 | 1.1 | 1.8 | 1.2 | 1.5 | 1.1 |

All States had a mean allelic diversity (He) value of 0.79. Independently, each State had mean He values ≥ 0.76 with the highest value of 0.82 observed in Enugu State and lowest He value of 0.76 observed in Oyo State (Table 3). Kruskal-Wallis test further showed no significant difference between the mean He observed across the nine States (p>0.05).

Table 3: Allelic diversity (He) of microsatellite loci from parasite populations in the nine states
Parasite population differentiation

Analysis of molecular variance (AMOVA)

Comparisons of parasite populations using AMOVA showed that genetic differentiation amongst the nine States was low with $\Phi_{PT} = 0.039$ with a $p$-value $< 0.05$ which suggests that only 3.9% of genetic variance exists among all States, and to a large extent (96.1%), parasite populations are similar across all nine States.

Fixation index (Fst)

The fixation index between parasite populations (Fst) is 0.038; that is, the genetic diversity between the nine States constituted 3.8% of the total genetic variance ($p < 0.05$) which essentially suggests that all nine States are not significantly genetically diverse from each other.

Principal component analysis (PCA)

Based on the PCA plots, low diversity existed among the States. However, plots showing each State independently revealed within-population diversity especially in Bayelsa, Imo and Kano States (Figure 2 - Panels B, E and F, respectively).

Results obtained from analyses showed no significant index of association in all parasite populations considered as the obtained LD value was 0.0179 (Table 4). Although the highest LD value of 0.0715 was obtained in Adamawa State, and the lowest LD value of 0.0037 was obtained in Kwara State, there was no significant difference ($p > 0.05$). This is indicative of similar parasite structuring in all nine States.

Table 4: Linkage disequilibrium analysis for *P. falciparum* populations obtained in each state
Population | $V_D$ | $V_E$ | $\hat{F}_A$
--- | --- | --- | ---
Adamawa | 3.5435 | 1.9832 | 0.0715
Bayelsa | 3.1642 | 2.0411 | 0.05
Enugu | 1.8239 | 1.5728 | 0.0145
Ibadan | 2.2501 | 1.9772 | 0.0125
Imo | 2.3986 | 1.855 | 0.0266
Kano | 2.2776 | 1.7312 | 0.0287
Kwara | 1.6931 | 1.7651 | 0.0037
Sokoto | 2.0717 | 1.6981 | 0.02
Plateau | 1.8405 | 1.6313 | 0.0117
**ALL** | **1.8031** | **1.507** | **0.0179**

$V_D$: the observed variance.

$V_E$: the expected variance of $n$ - the number of loci for which two individuals differ

$\hat{F}_A$: Linkage disequilibrium

**Discussion**

Nigeria remains the country with the highest global malaria burden. Hence, molecular studies on *P. falciparum* diversity and population structure become essential in monitoring the impact of different intervention strategies in the control of malaria transmission. This study employed the use of 12 microsatellites to evaluate *P. falciparum* genetic diversity and population structure in nine Nigerian States. Although microsatellites are better alternatives to polymorphic markers such as *msp*-1, *msp*-2, and *Glurp*, there are only a few reports of its use in studies conducted in Nigeria.

Our analysis of the microsatellite data generated in this study revealed high parasite diversity across all states. For instance, the mCOI (measure of parasite diversity) in all nine States was high (ranging from 1.71-2.66). Although, higher mCOI values (4.38-5.4) have been reported in earlier studies conducted prior to the introduction of artemisinin combination therapies (ACTs) [10, 11, 33], mCOI values obtained in this study suggest a steady decline in parasite diversity 13 years post-adoption of ACTs in Nigerian children. This may largely be attributed to the adoption and deployment of ACTs in Nigeria. In addition, other concurrent interventions such as broader distribution of long-lasting insecticide treated net (LLIN), may be a contributing factor [26]. Another measure of parasite diversity is the number of effective alleles (Ne) detected per microsatellite locus. It is expected that the number of Ne detected per locus is likely to be high in areas with high malaria endemicity and vice versa [5, 19]. The observed mean Ne in parasites obtained from all States (5.4 - 8.2) were comparable to those reported in other high-endemic regions of Sub-Saharan Africa [4, 5, 34]. The distribution of observed mean Ne in the Northern and Southern States were similar (p>0.05). This is equally expected as malaria endemicity continues to be high throughout Nigeria. Estimated allelic diversity (computed as expected heterozygosity-He), was high in all States with values ranging from 0.774 to 0.825. This suggests that parasites from these States exhibited high heterozygosity, which depicts high parasite transmission [35]. Similar high He values have been reported in other parts of Nigeria (Ekiti State: 0.79 and Lagos State: 0.65) and other countries with high levels of malaria transmission [5, 6, 18, 36, 37].

Although parasite diversity was high, further analysis of microsatellite data generated revealed low parasite population differentiation. Analysis of molecular variance (AMOVA) and genetic differentiation index ($F_{ST}$) values obtained were 0.039 and 0.038 respectively, which is low [36, 38]. This implies that about 96% of genetic variations observed among parasites were within populations. The principal component analysis (PCA) of all nine States further confirmed genetic similarities amongst parasite populations as similar clustering patterns consistent with low levels of genetic differentiation were observed. Linkage disequilibrium (LD) values for each parasite population ranged from 0.0037 in Kwara to 0.0715 in Adamawa. The overall association index was 0.0219, which is weaker than those typically reported in regions with low transmission [21, 23]. Studies have associated low LD values such as those reported in this study, to high levels of malaria transmission; which leads to increased cross-breeding and meiotic recombination that results in LD breakdown [5, 6, 19, 39]. Although the LD values obtained in this study remain low, there is a need to continually monitor parasite populations within Nigeria to detect new variants that may inform adaptation against interventions currently employed. The perceived lack of genetic differentiation or sub-structuring between States as evidenced by results obtained from AMOVA, Fst, PCA and LD analysis, is probably as a result of immense human migration between these populations as part of the usual socioeconomic activities and indiscriminate vector migration within the country [6, 40, 41, 42].

**Conclusion**
This study represents the first use of 12 polymorphic microsatellite loci to characterize parasite diversity and structure in Nigeria across regions representing all the six geographical zones of the country. The high level of genetic diversity and low population structuring in this study suggests that parasite populations circulating in Nigeria are homogenous. This implies that a uniform control strategy will be effective across the six geographical zones. The results obtained can be used as a baseline for parasite diversity and structure, aiding in the formulation of appropriate therapeutic and control strategies in Nigeria.

**Abbreviations**

PCR: Polymerase chain reaction  
MSP1: Merozoite surface protein 1  
MSP2: Merozoite surface protein 2  
GLURP: Glutamate-rich protein  
DBS: Dried blood spot  
COI: Complexity of infection  
mCOI: Mean complexity of infection  
Na: Number of alleles  
Ne: Number of effective alleles  
He: Expected heterozygosity  
AMOVA: Analysis of molecular variance  
FST: Fixation index  
PCA: Principal component analysis  
LD: Linkage disequilibrium  
ACTs: Artemisinin combination therapies  
LLIN: Long lasting insecticide net

**Declarations**

**Acknowledgement**

The authors thank all the patients, their parents or guardians for volunteering to participate in the study. We also acknowledge the principal investigators (PI) in each of the nine sentinel locations considered in this study. The National Malaria Elimination Program of the Federal Ministry of Health in Nigeria for the year 2014-2018.

**Funding**

This work was supported by grants from African Centre of Excellence for Genomics of Infectious Diseases (ACEGID), World Bank (ACE019) and The National Institute of Health (grants U01HG007480, U54HG007480). The U.S President’s Malaria Initiative (USPMI) funded the primary drug efficacy study from which samples were obtained for the current study.

**Author information**

Fehintola V. Ajogbasile and Adeyemi T. Kayode contributed equally to this work.

**Affiliations**

**African Centre of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer’s University, Ede, Nigeria.**

Fehintola V. Ajogbasile, Adeyemi T. Kayode, Kazeem O. Akano, Paul E. Oluniyi, Jessica N. Uwanibe, Benjamin B. Adegboyega, Courage Philip, Philomena J. Eromon, Onikepe A. Folarin Christian T. Happi

**Department of Biological Sciences, Faculty of Natural Sciences, Redeemer’s University, Ede, Nigeria**
Author contributions
F.V.A., A.T.K., B.B.A., C.P., J.N.U., O.G.J., N.O., U.O., S.O., J.A., R.D.W., N.J., F.F., G.E., and O.M carried out the experiments and acquired data; C.T.H. conceived the study; F.V.A., A.T.K., K.O.A. and P.E.O performed data analysis and interpretation; F.V.A and A.T.K wrote the paper; K.A.O., P.E.O., J.N.U., P.J.E., A.S., O.A.F and C.T.H reviewed and revised the manuscript; P.J.E., A.S., O.A.F and C.T.H. supervised the study.

Ethical Declaration
The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the National Health Research Ethics Committee, Federal Ministry of Health (FMoH), Abuja, Nigeria. Informed consent was obtained from parents and legal guardians of participants prior to enrollment in study.

Consent for publication
Not applicable

**Competing interests**

The authors declare that they have no competing interests

**Availability of data and materials**

The data analyzed for this manuscript is available upon request from the corresponding author.

**References**

1. Ceesay SJ, Casals-Pascual C, Erskine J, Anya SE, Duah NO, Fulford AJC, et al. Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis. Lancet. 2008;372: 1545–1554.

2. O'Meara WP, Bejon P, Mwangi TW, Okiro EA, Peshu N, Snow RW, et al. Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. Lancet. 2008;372: 1555–1562.

3. Jaenisch T, Sullivan DJ, Dutta A, Deb S, Ramsan M, Othman MK, et al. Malaria incidence and prevalence on Pemba island before the onset of the successful control intervention on the Zanzibar archipelago. Malar J. 2010;9: 32.

4. Bogreau H, Renaud F, Bouchiba H, Durand P, Assi S-B, Henry M-C, et al. Genetic diversity and structure of African Plasmodium falciparum populations in urban and rural areas. Am J Trop Med Hyg. 2006;74: 953–959.

5. Mulenge FM, Hunja CW, Magiri E, Cullerton R, Kaneko A, Aman RA. Genetic Diversity and Population Structure of Plasmodium falciparum in Lake Victoria Islands, A Region of Intense Transmission. Am J Trop Med Hyg. 2016;95: 1077–1085.

6. Oyebola MK, Idowu ET, Olukosi YA, Iwalokun BA, Aman RA. Genetic Diversity and Population Structure of Plasmodium falciparum in Lake Victoria Islands, A Region of Intense Transmission. Am J Trop Med Hyg. 2016;95: 1077–1085.

7. Ariey F, Chalvet W, Hommel D, Peneau C, Hulin A, Mercereau-Puijalon O, et al. Plasmodium falciparum parasites in French Guiana: limited genetic diversity and high selfing rate. Am J Trop Med Hyg. 1999;61: 978–985.

8. Haddad D, Snounou G, Mattei D, Enamorado IG, Figueroa J, Ståhl S, et al. Limited genetic diversity of Plasmodium falciparum in eld isolates from Honduras. Am J Trop Med Hyg. 1999;60: 30–34.

9. Atroosh WM, Al-Mekhla HM, Mahdy MA, Saif-Ali R, Al-Mekhlafi AM, Surin J. Genetic diversity of Plasmodium falciparum isolates from Pahang, Malaysia based on MSP-1 and MSP-2 genes. Parasit Vectors. 2011;4: 233.

10. Engelbrecht F, Tögel E, Beck HP, Enwezor F, Oettli A, Felger I. Analysis of Plasmodium falciparum infections in a village community in Northern Nigeria: determination of msp2 genotypes and parasite-specific IgG responses. Acta Trop. 2000;74: 63–71.

11. Haddad D, Snounou G, Mattei D, Enamorado IG, Figueroa J, Ståhl S, et al. Limited genetic diversity of Plasmodium falciparum in field isolates from Honduras. Am J Trop Med Hyg. 1999;60: 30–34.

12. Oyedeji SI, Awobode HO, Kun J. Limited Genetic Diversity and Low Multiplicity of Plasmodium falciparum Infections in Children with Severe Malaria in Laa, North-central Nigeria. Journal of Experimental & Clinical Medicine. 2013;5: 143–147.

13. Hughes AL. Positive selection and interallelic recombination at the merozoite surface antigen-1 (MSA-1) locus of Plasmodium falciparum. Mol Biol Evol. 1992;9: 381–393.

14. Hughes MK, Hughes AL. Natural selection on Plasmodium surface proteins. Mol Biochem Parasitol. 1995;71: 99–113.

15. Escalante AA, Lal AA, Ayala FJ. Genetic polymorphism and natural selection in the malaria parasite Plasmodium falciparum. Genetics. 1998;149: 189–202.
23. Pumpaibool T, Arnathau C, Durand P, Kanchanakhan N, Siripoon N, Suegom A, et al. Genetic diversity and population structure of Plasmodium falciparum in Thailand, a low transmission country. Malar J. 2009;8: 155.

24. Branch OH, Sutton PL, Barnes C, Castro JC, Hussin J, Awadalla P, et al. Plasmodium falciparum genetic diversity maintained and amplified over 5 years of a low transmission endemic in the Peruvian Amazon. Mol Biol Evol. 2011;28: 1973–1986.

25. Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGreedy R, Phyol AP, et al. Population genetic correlates of declining transmission in a human pathogen. Mol Ecol. 2013;22: 273–285.

26. Bankole BE, Kayode AT, Nosamiefan IO, Eromon P, Baniecki ML, Daniels RF, et al. Characterization of Plasmodium falciparum structure in Nigeria with malaria SNPs barcode. Malar J. 2018;17: 472.

27. Anderson TJ, Su XZ, Bockarie M, Lagog M, Day KP. Twelve microsatellite markers for characterization of Plasmodium falciparum from finger-prick blood samples. Parasitology. 1999;119 (Pt 2): 113–125.

28. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 2012;28: 2537–2539.

29. Mohammed H, Kassa M, Mekete K, Assefa A, Taye G, Commons RJ. Genetic diversity of the msp-1, msp-2, and glurp genes of Plasmodium falciparum isolates in Northwest Ethiopia. Malar J. 2018;17: 386.

30. Haubold B, Hudson RR. LIAN 3.0: detecting linkage disequilibrium in multilocus data. Linkage Analysis. Bioinformatics. 2000;16: 847–848.

31. Souza V, Nguyen TT, Hudson RR, Piñero D, Lenski RE. Hierarchical analysis of linkage disequilibrium in Rhizobium populations: evidence for sex? Proc Natl Acad Sci U S A. 1992;89: 8389–8393.

32. Happi CT, Gbotosho GO, Folarin OA, Bolaji OM, Sowunmi A, Kyle DE, et al. Association between mutations in plasmodium falciparum chloroquine resistance transporter and p. falciparum multidrug resistance 1 genes and in vivo amodiaquine resistance in p. falciparum malaria–infected children in nigeria. Am J Trop Med Hyg. 2006;75: 155–161.

33. Zhong D, Afrane Y, Githeko A, Yang Z, Cui L, Menge DM, et al. Plasmodium falciparum genetic diversity in western Kenya highlands. Am J Trop Med Hyg. 2007;77: 1043–1050.

34. Conway DJ, Machado RL, Singh B, Dessert P, Mikes ZS, Povoa MM, et al. Extreme geographical xation of variation in the Plasmodium falciparum gamete surface protein gene Pfs48/45 compared with microsatellite loci. Mol Biochem Parasitol. 2001;115: 145–156.

35. Durand P, Michalakis Y, Cestier S, Oury B, Leclerc MC, Tibayrenc M, et al. 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.

36. Balloux F, Lugon-Moulin N. The estimation of population differentiation with microsatellite markers. Mol Ecol. 2002;11: 155–165.

37. Anthony TG, Conway DJ, Cox-Singh J, Matsup A, Ratnam S, Shamsul S, et al. Fragmented population structure of plasmodium falciparum in a region of declining endemicity. J Infect Dis. 2005;191: 1558–1564.

38. Schultz L, Wapling J, Mueller I, Ntsukeni PO, Senn N, Nale J, et al. 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.

39. Lum JK, Kaneko A, Tanabe K, Takahashi N, Björkman A, Kobayakawa T. Malaria dispersal among islands: human mediated Plasmodium falciparum gene flow in Vanuatu, Melanesia. Acta Trop. 2004;90: 181–185.

40. Lum JK, Kaneko A, Taleo G, Amos M, Reiff DM. Genetic diversity and gene flow of humans, Plasmodium falciparum, and Anopheles farauti s.s. of Vanuatu: inferred malaria dispersal and implications for malaria control. Acta Trop. 2007;103: 102–107.

Figures
Figure 1

Geographic representation of sites (highlighted) in Nigeria where samples for in this study were collected. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.
**Figure 2**

Principal Component Analysis (PCA) plot of parasite populations in (A) Adamawa State, (B) Bayelsa State, (C) Enugu State, (D) Oyo State, (E) Imo State, (F) Kano State, (G) Kwara State, (H) Plateau State, (I) Sokoto State (J) Southern region States combined (Bayelsa, Enugu, Ibadan and Imo States), (K) Northern region States combined (Kano, Kwara, Plateau and Sokoto States), and (L) Northern region States and Southern Region States combined 

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- TA60.pdf
- TA42.pdf
- Ta81.pdf
- TA87.pdf
- TA109.pdf
- ARA2.pdf
- PFG377.pdf
- TA40.pdf
- TA1.pdf
- PFPK2.pdf
- 2490.pdf
- PolyA.pdf
• AdditionalFiles.pdf