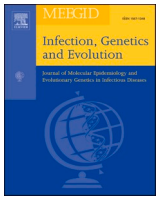
Infection, Genetics and Evolution 105 (2022) 105369

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid



Population genetic analysis of *Plasmodium falciparum* cell-traversal protein for ookinetes and sporozoite among malaria patients from southern Nigeria

Mary A. Oboh [a,*], Clement Isaac [b], Marissa N. Schroeter [a], Olanrewaju B. Morenikeji [c], Alfred Amambua-Ngwa [d], Martin M. Meremikwu [e], Bolaji N. Thomas [a,*]

[a] Department of Biomedical Sciences, Rochester Institute of Technology, Rochester, NY, United States of America
[b] Department of Zoology, Faculty of Life Sciences, Ambrose Ali University, Ekpoma, Nigeria
[c] Division of Biological and Health Sciences, University of Pittsburgh, Bradford, PA 16701, United States of America
[d] Medical Research Council Unit, The Gambia, at the London School of Hygiene and Tropical Medicine, Fajara, the Gambia
[e] University of Calabar Teaching Hospital, Calabar, Cross River State, Nigeria

ARTICLE INFO

Keywords:
Plasmodium falciparum
genetic diversity
polymorphism
phylogenetic relatedness
sub-Saharan Africa
celtos

ABSTRACT

Plasmodium falciparum immune escape mechanisms affect antigens being prioritized for vaccine design. As a result of the multiple surface antigens the parasite exhibits at different life cycle stages, designing a vaccine that would efficiently boost the immune system in clearing infections has been challenging. The *P. falciparum* cell-traversal protein for ookinetes and sporozoite (*Pfceltos*) is instrumental for ookinete traversal of the mosquito midgut and sporozoites invasion of the human liver cells. *Pfceltos* elicits both humoral and cellular immune response but has been reported with multiple single nucleotide polymorphisms in global isolates. A cross-sectional survey, conducted in southern Nigeria, between January–March 2021 recruited 283 individuals. Of this, 166 demonstrated *P. falciparum* infections (86 from Cross River and 80 from Edo), 48 (55.8%) while only 36 (45%) were amplified for *Pfceltos* gene from both sites respectively. Fifty amplified samples were sequenced and analysed for their diversity, polymorphisms and population structure of the gene. The number of segregating sites in Edo State was higher (34) than that of Cross River State. Though nucleotide diversity was higher for Edo compared to Cross River State ($\theta w = 0.02505$; $\pi = 0.03993$ *versus* $\theta w = 0.00930$; $\pi = 0.01033$ respectively), the reverse was the case for haplotype diversity (0.757 versus 0.890 for Edo and Cross River respectively). Of the twelve haplotypes observed from both states, only two (KASLPVEK and NAFLSFEK) were shared, with haplotype prevalence higher in Edo (16% and 36%) than Cross River (8% and 4%). The Tajima's D test was positive for both states, with Fst value showing a strong genetic differentiation (Fst = 0.25599), indicating the occurrence of balancing selection favoring haplotype circulation at a low frequency. The shared haplotypes, low Hst and Fst values presents a challenge to predict the extent of gene flow. High LD values present a grim public health consequence should a *Pfceltos*-conjugated vaccine be considered for prophylaxis in Nigeria.

1. Introduction

Malaria remains a major global health problem, particularly in sub-Saharan Africa (sSA), where *Plasmodium falciparum* is the most prevalent species. Despite the substantial effort directed towards malaria control, infection from *P. falciparum* still remains a huge public health problem. An estimated 241 million malaria cases were reported in 2020 globally, with the World Health Organization (WHO) Africa – WHO Afro region, contributing 95% of these and Nigeria accounting for 27%. Similarly, 627,000 malaria deaths were observed in 2020 and 95% of these deaths from sSA, with Nigeria contributing 27% of these deaths (WHO, 2021). These reports indicate sSA is disproportionately affected by malaria burden.

Control interventions have been fraught with the development of resistance by the parasite and vector to both artemisinin-based treatment and insecticides respectively (Asare et al., 2017; Ndiaye et al., 2005; Ngassa-Mbenda and Das, 2016; *PLoS Med.*, 2017; Fadel et al., 2019; Opondo et al., 2019; Weedall et al., 2019). Recent reports have

* Corresponding authors at: Department of Biomedical Sciences, Rochester Institute of Technology, 153 Lomb Memorial Drive, Rochester, NY 14623, United States of America.
E-mail addresses: maochst@rit.edu (M.A. Oboh), bntsbi@rit.edu (B.N. Thomas).

https://doi.org/10.1016/j.meegid.2022.105369
Received 3 July 2022; Received in revised form 12 September 2022; Accepted 20 September 2022
Available online 28 September 2022

shown reduced susceptibility of *P. falciparum* clinical isolates to artemisinin-combination therapies (ACTs) used in Uganda (Rasmussen et al., 2017; Asua et al., 2021; Tumwebaze et al., 2021; Balikagala et al., 2021; Fukuda et al., 2021) and Rwanda (Uwimana et al., 2020; Straimer et al., 2022; Van Loon et al., 2022). In addition, control of insect vector has also seen its own challenge with several insecticides showing reduced potency against local mosquito species (Olusegun-Joseph et al., 2020; Pwalia et al., 2019). These and the moderate efficacy of the only available vaccine warrants continuous search for additional drugs and vaccine to increase the toolkit for malaria elimination.

Currently, there are two advanced vaccines against malaria; the RTS, S /AS01 and the R21/MM (Osier et al., 2016; Datoo et al., 2021), showing variable levels of efficacy against prevention of *P. falciparum* infection and severe disease. Both of these vaccines are based on the circumsporozoite protein (CSP) of *P. falciparum* (Datoo et al., 2021) as the principal vaccine constituent. The RTS,S / AS01, which provides a moderate protection (<50%) in children has been approved by the WHO for deployment in highly endemic areas (Lancet, 2021). Unfortunately, RTS,S and other candidate vaccine antigens have not been spared from parasite immune evasive mechanisms (Terheggen et al., 2014; Ouattara et al., 2018; Healer et al., 2004). The outcome of these is increased variability and variant expression of multiple surface antigens hampering the design of an efficacious vaccine against infections and disease. Various malaria parasite antigens have been targeted as candidates for vaccine design, some at the pre-clinical stage, while others are at a more advance stage (Osier et al., 2016).

Current vaccine design approaches target the pre-erythrocytic, erythrocytic and transmission-blocking stages of *P. falciparum* developmental cycle (Chauhan et al., 2010; Ouattara and Laurens, 2015; Mahmoudi and Keshavarz, 2018). Vaccines targeting the asexual blood stages have been largely explored, because this approach has a direct impact on reducing morbidity and mortality associated with disease in humans (Zhang et al., 2007), while those targeting the sporozoite and asexual liver stage aim to generate sterile immunity or prevent the erythrocytic cycle, responsible for most malaria morbidity and consequent deaths. Other approaches aim for transmission blocking vaccines (TBVs), based on the transmissible erythrocytic gametocyte stages or the ookinete stages of *P. falciparum* in the vector. The objective of TBVs is to elicit antibodies that inhibit infections or development of parasite in the *Anopheles* mosquito midgut, thereby preventing production of the transmissible sporozoite stages. This makes TBVs very attractive as additional tools for malaria elimination programs (Pirahmadi et al., 2019).

A potent anti-malaria vaccine should take into consideration the life cycle stage in which the candidate antigens is most expressed. The *P. falciparum* cell-traversal protein for ookinetes and sporozoite (*Pfceltos*) is involved in mechanisms of ookinete traversal of the mosquito midgut and sporozoites invasion of the human liver cells (Jimah et al., 2016). *Pfceltos* is said to be highly conserved among *Plasmodium* species, suggesting a critical functional role across plasmodial species. Therefore, focusing on immunological response in conserved protein regions is likely to confer protection against malaria infection with multiple species (Bergmann-Leitner et al., 2011). *Pfceltos*, is a 25-kDa protein expressed in both the mosquito ookinete stage and the sporozoites in the vertebrate host, including humans. It is associated with protection against symptomatic infection, humoral and cellular response in mouse model, inhibition of sporozoites infection and oocyst development from ookinete in mosquito (Bergmann-Leitner et al., 2011; Bergmann-Leitner et al., 2010).

Genetic diversity study on the full length of *Pfceltos* from global isolates have shown the presence of 39 non-synonymous single nucleotide polymorphisms (SNPs) and 69 haplotypes, usually in the C-terminus region, in comparison to the 3D7 reference strain. However, the epitopes that binds to B and T immune cells were identified to be in conserved regions, outside of the identified variants (Pirahmadi et al., 2018). In light of renewed focus on transmission blocking vaccines, and the promising results arising from *Pfceltos*, evaluating its diversity in *P. falciparum* isolates from different malaria endemic regions will inform on genetic variation and potential structural patterns that could compromise its efficacy in future vaccines, especially in regions where infections are due to heterologous parasite strains. To this end, we employed a population genetics approach to gain insight into variation pattern of *Pfceltos* gene from clinical isolates of *P. falciparum* from two regions of Nigeria.

2. Materials and methods

2.1. Ethics

The study was approved by the Institutional Review Board (IRB/16/347) of Nigerian Institute of Medical Research, Lagos; the Edo State Hospital Management Board and the Ministry of Health; and Cross River State Health Research Ethics Committee CRSMOH/RP/REC/2020/139. Study objectives, benefits and any potential risks were explained to all participants in English or Esan as the case maybe before enrolment into the study. Informed consent [written for adults and verbal assent for children (age $\geq 8 \leq 16$ years) with parental consent] was obtained from all participants.

2.2. Study sites and sample collection

Blood samples were drawn from patients presenting with clinical symptoms of malaria from two states in Nigeria – Edo State, located in southwestern Nigeria, and Cross River State in the South-South region, bordering Cameroon to the East. In Edo State, samples were collected from symptomatic participants from a community approach in Esan West Local Government Area. In Cross River State, two types of sample collection approaches were employed: a community survey and a hospital-based survey, from Biase Local Government Area. In both study areas, only rapid diagnostic test (RDT) positive participants with fever were enrolled. Epidemiologically, Cross River and Edo states are meso-endemic with a malaria prevalence rate of 35% in 6 months to ≤5 year old children (NMEP, 2014; Ebomwonyi et al., 2015). Sample collection in Cross River State was from January–March 2021 while in Edo State, samples were collected in February 2021.

In both states, febrile patients (both adults and children of <1 year of age) with some clinical symptoms of malaria (such as fever, temperature > 37 °C) were recruited, while pregnant women and other potential participants with other suspected underlying disease conditions were excluded. Blood samples were subsequently screened using PfHRP2-based RDT kits (SD Bioline, Abbot US) and positive samples used for preparing thick and thin blood films for microscopy. Furthermore, 3 drops of blood were spotted on Whatman® no. 3 filter paper (GE Healthcare Life Sciences) and allowed to air dry at room temperature. Each filter paper was kept in separate zip-locked sachet containing silica gel and stored at room temperature until use.

2.3. DNA extraction, PCR and processing for sequencing

Genomic DNA was extracted from 3 mm discs punched out of dried blood spots using the QIAamp DNA Blood Mini Kit (Qiagen®, Hilden, Germany). DNA was eluted in a 100 μl final volume per sample and stored at −20 °C until ready for use. All extracted gDNA was used to further confirm infection by a real-time PCR (RT-PCR) assay that targets the *P. falciparum* var. gene multi-copy terminal sequences (varATS) (Hofmann et al., 2015). Primers, probe sequence and thermal cycling conditions for RT-PCR validation of *P. falciparum* have been previously explained (Umunnakwe et al., 2019). A cycle threshold (CT) cut off of ≤38CT was used as threshold to score samples as positive for malaria. A total of 66 from Edo and 86 from Cross River were included for *Pfceltos* specific assay.

A fragment (475 bp) of *P. falciparum celtos* coding sequence was

amplified from each sample using newly designed primers (*Pfceltos* forward 5′- CAGAGGAAACAACGGACACA −3′ and *Pfceltos* reverse 5′-TTCGCACCTACAGCTGTTTC -3′) synthesised by IDTDNA (Corralville IA, USA). PCR amplification reactions were performed using the EconoTaq Plus Green 2× Master Mix (Lucigen, Middleton, WI). Each PCR reaction of 25 μl contained: 12.5 μl of the 2× master mix, 5 μl of nuclease free water, 5 μl of gDNA and 1.25 μl of each forward and reverse primer (which equates to 0.5 mM concentration of each primer). The cycling condition are 95 °C denaturation for 2 min, a second denaturation at 95 °C for 45 s, annealing at 58 °C for 45 s, extension of 72 °C for 1 min for 35 cycles, and a final extension at 72 °C for 5 min. All PCR products were run on 2% agarose gel stained with ethidium bromide, only amplified samples were further processed for sequencing.

Amplified PCR products were purified using the ExoSAP-IT™ enzyme (Thermo Scientific, MA, USA) prior to DNA sequencing as per manufacturer's instruction. Purified amplified PCR products were sequenced commercially by Genewiz-Azenta Life Sciences (South Plainfield, New Jersey- 684,847,215/687645926). For each fragment, sequencing was carried out from both the 5′ and 3′ directions (2× coverage).

2.4. Sequence analysis

Nucleotide base calling and chromatography visualization was performed with BioEdit Sequence Alignment Editor version 7.0.5.3 (Hall, 1999). Homologous DNA sequences for *Pfceltos* from the *P. falciparum* 3D7 reference strain (Pf 3D7_1216600), downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov), was used to detect single nucleotide polymorphisms (SNPs). To confirm the presence of any identified SNPs, re-visualization was carried out on the chromatogram. Furthermore, amino acid translation was carried out since the region sequenced was the coding region (exons), and the different amino acid haplotypes circulating in the two states were determined along with their frequencies.

2.5. Population genetic analysis

Various genetic test approaches were carried out to estimate the haplotype and nucleotide diversities, as well as evaluate linkage disequilibrium (LD) of *P. falciparum*. Translated amino acid sequences were used to determine the frequencies of different haplotypes for each population (Cross River and Edo). *Pfceltos* haplotype (Hd) (Nei, 1987) and nucleotide diversities (as estimated by θw and π) were determined independently for each sampling site. The average number of pairwise nucleotide difference per site was estimated by π (Tajima, 1989), and the θw value was based on the number of segregating sites in a population (Watterson, 1975). In order to determine the pattern of immune selection, the Tajima's D (TD) test was performed for the two different populations in Nigeria. The TD test provides the normalized difference between θw and π, where a negative value signifies abundance of low frequency polymorphism. This low frequency polymorphism is indicative of directional selection or population size expansion while a positive value indicates balancing selection or a reduction in the population size. All genetic and selection analyses were carried out with the DnaSP program version 5.0 (Librado and Rozas, 2009). In addition, the genealogical relatedness of the haplotypes from both states was constructed using the Templeton, Crandall and Sing (TCS) model (Clement et al., 2000) in the PopArt program (Leigh and Bryant, 2015). LD tests were performed using the HaploView program (Barrett et al., 2005) to determine possible association between the segregating SNPs in each population.

Genetic differentiation which is the difference in the average diversity *within* populations compared to that *among* populations denoted as the Wright Fixation index (Fst) and haplotype statistic index (Hst) was computed. Statistical significance was evaluated by non-parametric permutation-based test with 1000 random iterations of samples. These indices are proxy for determining the level of gene flow between two or more populations (Hudson et al., 1992). Fst value of <0.05 is denoted as low genetic differentiation or high gene flow between populations, 0.05–0.15 is interpreted as moderate genetic differentiation/gene flow, while Fst value >0.25 is said to have high genetic differentiation/low gene flow i.e. there is distinct difference between the populations.

Since this is a vaccine candidate antigen with high likelihood of multiple epitopes and being under immune selection, B-cell epitopes was predicted using the random forest algorithm. This algorithm is trained on annotating epitopes from antibody-antigen protein structures. FASTA format of the sequences from this study was uploaded on the BepiPred-2.0 sequential B-Cell epitope predictor server, using an epitope threshold of 0.5 (Jespersen et al., 2017). The predicted B-cell epitopes were further compared with regions having high TD value to ascertain if any of the predicted epitope is under immune selection.

3. Results

3.1. Characteristics of malaria patients from study sites

A total of 283 samples comprising of 187 and 96 were collected from Cross River and Edo states respectively. The mean age of the study participants from Cross River state is 8 years while that from Edo state is 11 years. Of these participants, 100 and 87 were males and females respectively (from Cross River) while 45 and 51 males and females were recruited from Edo state. Of the total number of recruited participants, 166 were detected as positive for *P. falciparum* infections (86 from Cross River and 80 from Edo). All *P. falciparum* positive samples from both states were subjected to conventional PCR targeting the *Pfceltos* gene; of this, 48 (55.8%) and 36 (45%) were amplified from Cross River and Edo state respectively. Twenty-five samples of the *Pfceltos* PCR-amplified samples (from each state) were subjected to Sanger sequencing (Table 1). All sequences generated in this study were submitted to the GenBank (accession numbers: ON866956-ON866996).

3.2. Evaluation of genetic diversity and haplotype frequency

Of the 25 *Pfceltos* amplified sequencing products from Cross River State, there were 11 segregating sites with 10 haplotypes. Though the number of segregating sites in Edo State was higher (34) than that of Cross River State, the number of haplotypes was lower (4) than that of Cross River. In addition, though the haplotype diversity between the two states were no different from each other, that from Edo state (0.757) was a bit lower than what was observed in Cross River state (0.890). The nucleotide diversity recorded for Edo ($\theta w = 0.02505$; $\pi = 0.03993$) was higher than what was observed in Cross River ($\theta w = 0.00930$; $\pi = 0.01033$). Consequently, the average number of nucleotide difference was higher in Edo (14.773) than in Cross River (3.825), but the recombination events in Cross River (2) was higher than that of Edo (Table 2).

Quality-trimmed nucleotide sequences were translated into amino acid (AA) and the sequences aligned with the protein reference strain in order to construct the haplotype. Mutations at AA positions 100, 102,

Table 1
Characteristics of recruited individuals from different study sites.

	Cross River State	Edo State
Number of participants	187	96
Proportion positive after molecular diagnosis	86 (45.9%)	80 (83.3%)
Proportion amplified for *Pfceltos* gene	48 (55.8%)	36 (45%)
Age (years)		
Mean	8	11
Range	1–12	3–15
Sex		
Male	100	45
Female	87	51

Table 2
Genetic diversity and pattern of selection of *Pfceltos* gene among malaria patients.

Genetic metric	Cross River State	Edo State
Number of isolates	25	25
SNPS/segregating sites	11	34
Number of haplotypes	10	4
Haplotype diversity	0.890	0.757
Nucleotide diversity θ	0.00930	0.02505
Π	0.01033	0.03993
Average number of nucleotide difference	3.823	14.773
Minimum number of recombination events (Rm)*	2	0
Test of neutrality		
Tajima's D	0.37982[a]	2.24344[b]
Fu and Li's D	0.65893[a]	1.72336[c]
Fu and Li's F	0.67066[a]	2.21302[c]

* = Sites segregating for three or four nucleotides were not considered, [a] = p > 0.10; [b] = p < 0.05; [c] = $p < 0.02$. All analyses were carried out with a sliding window of 100 in a 25-nucleotide stepwise length.

104, 114, 116, 117, 118 and 119 forming a reference haplotype of **KASLSFEN** (as per 3D7) were used to derive the sequence haplotypes. The most common haplotype in Cross River state was the KALLSVEQ with a 24% prevalence followed by the 3D7 KASLSFEN and KAFLSVES haplotypes, which occurred at similar frequency of 16%, while KAFLSVEK and KAHLSVEK occurred at 12%. Others were infrequent in occurrence. In Edo state however, of the 25 isolates sequenced, there were only four haplotypes with the most predominant being NAFLSFEK, which was found in 36% of the haplotypes, followed by KPSWPPKK and KASLSFEK, which occurred at 28% and 20% respectively. The least frequent haplotype was the KASLPVEK with a prevalence of 16% (Fig. 2; Table 3).

3.3. *Pfceltos selective signatures, genetic differentiation and gene flow at study sites*

In order to find out if *Pfceltos* gene from each state deviates from the neutral immune selection model, the Tajima's D (TD) test was performed. A moderately positive TD value was recorded for Cross River state (0.37982) while that for Edo state was highly positive (2.24344). Although, TD from Cross River did not statistically deviate from the neutral model ($p > 0.10$), that of Edo however statistically deviated from the neutral model ($p < 0.05$). The deviation from the neutral model was further confirmed by the Fu and Li D and F test (Table 2).

Genetic differentiation between Cross River and Edo state was computed with both Hst and Fst. Using the haplotype index (Hst), a statistically moderate genetic differentiation was observed (Hst = 0.08477, $p < 0.001$) between *Pfceltos* from both states. Similarly, the Fst value between both populations showed a strong genetic differentiation (Fst = 0.25599), thus implying low gene flow between parasite populations from both states (Table 4).

Table 3
Distribution of *Pfceltos* gene haplotypes from malaria patients in Nigeria.

Haplotypes	Distribution Number (%)
Cross River (25 sequences)	
KASLSFEN	4 (12)
KAFLSVES	4 (16)
KAYLSVEN	1 (4)
KAFLSVEK	3 (12)
KALLSVEQ	6 (24)
KASLPVEK	2 (8)
NAFLSFEK	1 (4)
KAHLSVEK	3 (12)
KAILSVEK	1 (4)
Edo (25 sequences)	
KPSWPPKK	7 (28)
KASLPVEK	4 (16)
NAFLSFEK	9 (36)
KASLSFEK	5 (20)

NB: The consensus nucleotides of the isolates were translated into amino acids and compared with the amino acids of the reference strain. Mutations were located at amino acid positions 100, 102, 104, 114, 116, 117, 118 and 119 and the reference amino acids at those positions is **KASLSFEN.**

3.4. *Assessment of linkage disequilibrium between Pfceltos from study sites*

Linkage disequilibrium (LD) analysis was carried out in order to determine if SNPs present at different codons in the *Pfceltos* gene from both states were associated in a single parasite. The r2 value of each pairwise SNP association was computed separately for each population. The level of association of LD is denoted by the intensity of color (Fig. 1). SNP with a minor allele frequency of less than 0.001% were excluded as per the program requirement; hence for Cross River, SNPs at codons 304, 341, 350, 352 were excluded while SNPs at codons 310, 355 and 356 were also excluded. The observed association between pairwise SNPs in Cross River was moderate to weak with a few SNPs having a r2 value of 54, while majority recorded low r2 values. In contrast, strong multiple pairwise associations were observed between SNPs in Edo state as denoted by the dark square box and r2 values (Fig. 1).

3.5. *B-cell epitope prediction and immune selection*

Prediction of linear B-cell epitopes within the sequences revealed varying epitopes (Fig. 2). Some epitopes with amino acid sequence SSSLYNGSQFIEQ, occur in isolates from both states but not in regions with high TD. Other epitopes including SQSMNKIGDDL, LESQSMNKIGDDL, SVLQKNSPTFLESSFDI, GLPSVESLVAENVKPPKVDPA, and VGLPSVESLVAENVKPPKVDPA were in regions with high TD (both positive-Edo, negative-Cross River) in both states, a feature indicative of immune selection (Table 5; supplementary Fig. 1 and 2).

4. Discussion

Candidate antigens for malaria vaccine, especially those efficient against *P. falciparum,* have been in trial for a long time. A key challenge to having an effective vaccine however has been the variations in *P. falciparum* antigens targeted for vaccine development. These variations have hindered the effectiveness and efficacy of vaccine in trials. Continuous research into the genetic architecture and antigenic variation of candidate vaccine antigens among global *P. falciparum* populations would advance understanding of how this could affects vaccine efficacy, and facilitate approaches to design optimal, efficacious vaccine formulation (Escalante et al., 2002). *Pfceltos* is a promising malaria vaccine candidate, which has been shown to elicit both cellular and humoral immune response in an infected host (Bergmann-Leitner et al., 2011; Bergmann-Leitner et al., 2010; Hudson et al., 1992). However, not much is known about the diversity of this protein from different endemic regions and how it will affect vaccine efficacy. In this study, we evaluated the genetic diversity of *Pfceltos* gene, single nucleotide polymorphisms, pattern of selection and genetic differentiation between two states in Nigeria.

We detected high and moderate mutational events in Edo and Cross River states respectively, denoted by the number of segregating sites.

Table 4
Genetic differentiation and gene flow of *Pfceltos* gene between study sites.

Hst values	Cross River
Edo	0.08477***
Fst value	Cross River
Edo	0.25599

*** = $p < 0.001$.

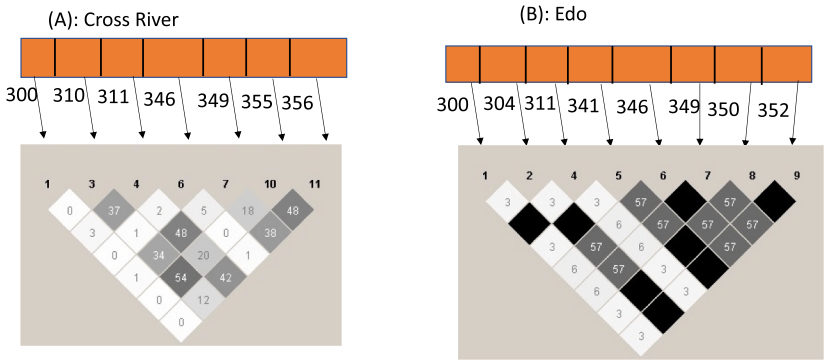


Fig. 1. Linkage disequilibrium between the SNPs of *Pfceltos* gene between malaria parasite isolates from (A) Cross River and (B) Edo. The strength of association between pair of SNPs is indicated by color intensity of the boxes. The darker boxes represent stronger association, and the lighter ones represent weaker ones. Values within the boxes are the r^2: higher value is indicative of stronger association, and the black boxes show perfect association ($r^2 = 1$).

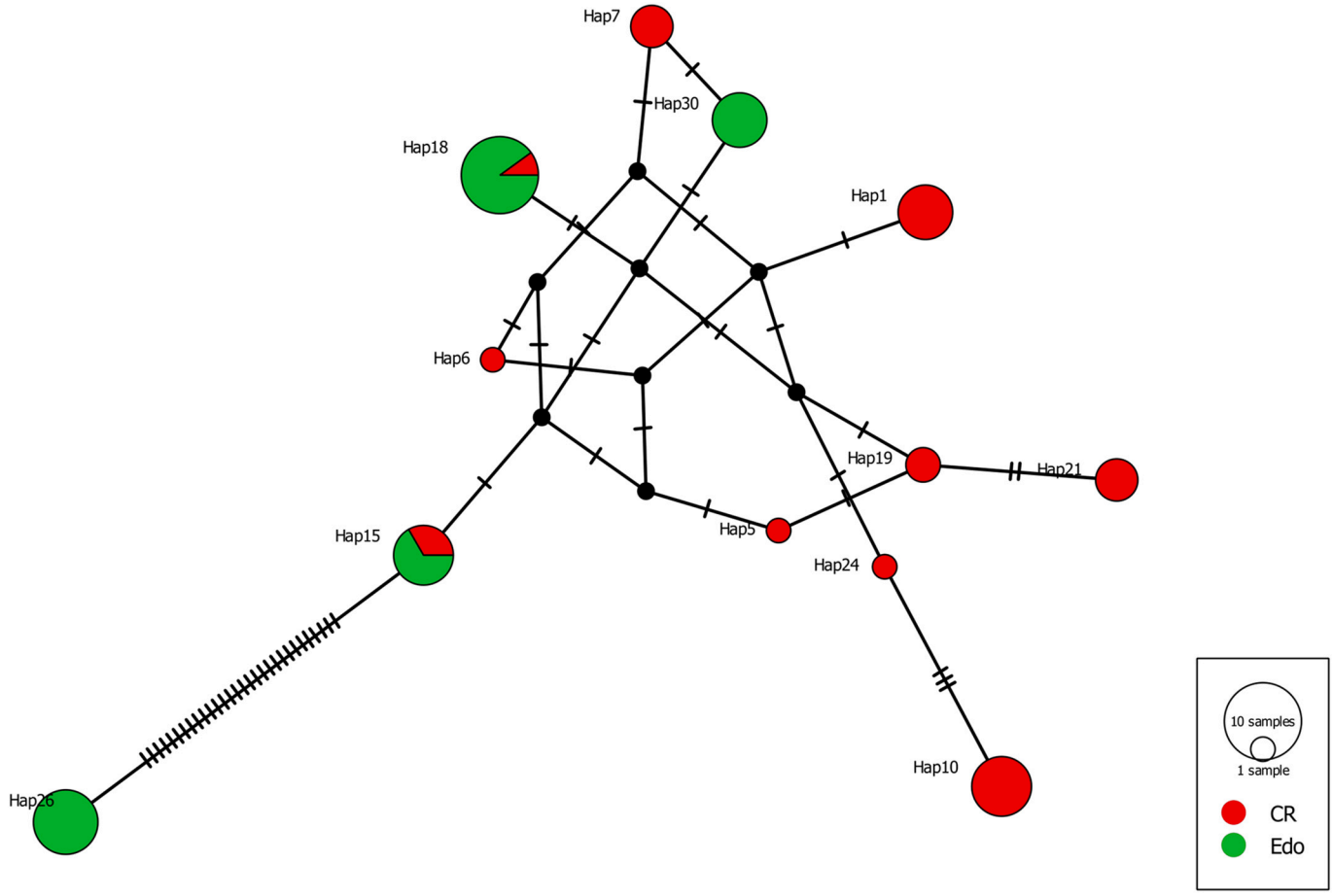


Fig. 2. Haplotype network analysis model.
Network summary of the diversity of *Pfceltos* haplotypes from 370 nucleotide sequences in Cross River and Edo ($N = 25$ each) was determined with the Templeton, Crandall, and Sing (TCS) model of the PopArt program. Size of circles are scaled to the frequency of the haplotypes. NB: Only two haplotypes are shared between both study locations as indicative of the color combination (green and pink) in Hap15 and Hap18.

Although the segregating sites in Edo is very high (34) than that in Cross River, a reverse observation was recorded with the number of haplotypes. The number of segregating sites in Edo is higher than that observed in The Gambia (25), Senegal (26) and Mali (20), while the estimate from Cross River is lower than that observed from all three West African countries, but comparable to Iran (9), French Guiana (22) and Uganda (14). Similar pattern was observed with haplotype diversities from both states in relation to these countries. However, the haplotype diversities from these study sites were comparable to that observed from the Gambia (0.897) and Mali (0.857) (Pirahmadi et al., 2018). Malaria parasite genetic diversity is one of the mechanisms of immune or drug evasion (Jespersen et al., 2017; Escalante et al., 2002; Ridzuan et al., 2016). Therefore, extensive genetic diversity in *Pfceltos,* a critical protein required for *P. falciparum* ookinetes to traverse mosquito midgut and sporozoites to invade the liver cells (Jimah et al., 2016), known to elicit substantial immune response (at the cellular and

Table 5
Predicted B-cell epitopes.

Predicted B-cell epitopes	Cross River	Edo
SSSLYNGSQFIEQ	24	25
SQSMNKIGDDL	13	10
LESQSMNKIGDDL	6	0
SVLQKNSPTFLESSFDI	13	12
QKNSPTFLESSFD	3	0
KVGLPSVESLVAENVKPPKVDPA	4	0
GLPSVESLVAENVKPPKVDPA	2	10
LPSVESLVAENVKPPKVDPA	8	5
VGLPSVESLVAENVKPPKVDPA	6	0
GWPPPKKPIAENVKPPKKNPT	0	25
LPSFENL	2	0
AENVKPPKVDPA	2	0

B cell epitopes were determined by the random forest algorithm on epitope annotation from antibody-antigen protein interaction; BepiPred program, version 2.0; threshold of 0.5.

humoral level), and is being considered for vaccine design, will have serious implications on the potency of a *Pfceltos*-subunit vaccine.

The number of haplotypes from Edo State is higher than that from Cross River State, but lower in comparison to Senegal and The Gambia. These high haplotypes recorded in Edo state could be due to expansion of the parasite population in this study area. High malaria endemicity can result in increased parasite recombination and ultimately increased diversity (Ridzuan et al., 2016). Only two haplotypes (KASLPVEK and NAFLSFEK) were shared between both states, but the prevalence of these haplotypes were higher in Edo (16% and 36%) than in Cross River (8% and 4%). The other haplotypes (two from Edo and seven from Cross River) were unique to each state suggesting the impact of host immune response on mutational changes in parasites. In addition, limited number of haplotypes and amino acid polymorphisms were observed in Edo state than in Cross River. This is a possible indication of the circulation of low frequency alleles that might pose challenge with a *Pfceltos* formulated vaccine efficacy.

Genetic exchange between parasite isolates from both states as determined by haplotype differentiation was low (Hst = 0.08477). This was contrary to the Fst value, which showed more genetically distinct (Fst = 0.25599) populations, making it somewhat difficult to predict the extent of gene flow between the two states. Nevertheless, this was lower than what was reported in Senegal, The Gambia, Mali and French Guiana but lower than Ugandan and Iranian isolates (Pirahmadi et al., 2018). This does not suggest any form of geographical heterogeneity as observed by previous authors (Pirahmadi et al., 2018). The distance between both study sites is <350 km, potentially allowing for the possibility of "parasite genetic communication" between both states due to human movements.

Interestingly, the TD value for *Pfceltos* from Edo state was highly positive and statistically significant, indicating that the gene is under balancing selection. This was further confirmed by the Fu and Li's D (2.21302) and F (1.72336) tests that were statistically significant ($p < 0.02$), an indication of deviation from the neutral selection model. Taken together with the high number of segregating sites and moderate number of haplotypes of *Pfceltos* in Edo, this suggest the effect of balancing selection occurring in this gene in Edo (Kar et al., 2016). Though, the TD and the Fu and Li's D/F value for Cross River did not suggest significant ($p > 0.10$) deviation from the neutral model, it can be presumed that the high number of haplotypes (10) and moderate number of segregating sites (11) might be a result of balancing selection that favor the occurrence of several low frequency of these haplotypes.

The observation of significant LD between the mutations in *Pfceltos* from Edo and a moderate association between mutations from Cross River will result in further genetic challenge to malaria control, should a *Pfceltos*-subunit vaccine be considered for prophylaxis in Nigeria. The detection of linear B-cell epitopes in regions with high TD values is indicative of high immunogenicity of the predicted epitopes, and the likelihood that these epitopes would be optimal for the design of an effective vaccine conjugate. These epitopes were some of those validated by Bergmann-Leitner and team earlier (Bergmann-Leitner et al., 2011; Bergmann-Leitner et al., 2010). Should these epitopes be considered in vaccine design, there is high probability that such a vaccine would produce allele-specific responses in different malaria endemic regions.

This "collaborative effort" by SNPs could have a substantial negative impact on the ongoing national malaria control intervention and eventually the global effort on malaria elimination by 2030. Due to the limited sample size and restriction in geographical locations, it is recommended that these studies be expanded to other parts of Nigeria, to generate more robust information on the genetic diversity and differentiation, molecular selection and structure and linkage disequilibrium of different *P. falciparum* vaccine candidate antigens.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2022.105369.

Data availability statement

All sequences analysed in this study have been deposited in the GenBank. The accession numbers are ON866956-ON866996.

Funding

This research received no external funding.

Contributions

MAO and BNT conceptualized this work; MAO, CI and MMM were instrumental with patient recruitment and sample collection; MAO and MNS carried out laboratory experiment and molecular genotyping; MAO carried out analyses and wrote the first draft of the manuscript; MAO, CI, OBM, AAN, MMM and BNT carried out critical review of manuscript. All authors read and approved the submission of this manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We acknowledge ongoing support and funding from the College of Health Sciences and Technology, Rochester Institute of Technology (BNT). MAO is supported through the American Association of Immunologists Intersect Fellowship Program for Computational Scientists and Immunologists. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Asare, K.K., Boampong, J.N., Duah, N.O., Afoakwah, R., Sehgal, R., Quashie, N.B., 2017. Synergism between Pfcrt and Pfmdr1 genes could account for the slow recovery of chloroquine sensitive Plasmodium falciparum strains in Ghana after chloroquine withdrawal. J Infect Public Health. 10, 110–119. https://doi.org/10.1016/j.jiph.2016.02.004.

Asua, V., Conrad, M.D., Aydemir, O., Duvalsaint, M., Legac, J., Duarte, E., et al., 2021. Changing prevalence of potential mediators of aminoquinoline, antifolate, and artemisinin resistance across Uganda. J Infect Dis. 223, 985–994.

Balikagala, B., Fukuda, N., Ikeda, M., Katuro, O.T., Tachibana, S.-I., Yamauchi, M., et al., 2021. Evidence of artemisinin-resistant malaria in Africa. N Engl J Med. 385, 1163–1171.

Barrett, J.C., Fry, B., Maller, J., Daly, M.J., 2005. Haploview : analysis and visualization of LD and haplotype maps. Bioinforma Appl Note. 21, 263–265.

Bergmann-Leitner, E.S., Mease, R.M., de la Vega, P., Savranskaya, T., Polhemus, M., Ockenhouse, C., et al., 2010. Immunization with pre-erythrocytic antigen CelTOS from plasmodium falciparum elicits cross-species protection against heterologous challenge with Plasmodium berghei. PLoS One. 5, e12294.

Bergmann-Leitner, E.S., Legler, P.M., Savranskaya, T., Ockenhouse, C.F., Angov, E., 2011. Cellular and humoral immune effector mechanisms required for sterile protection against sporozoite challenge induced with the novel malaria vaccine candidate CelTOS. Vaccine. 29, 5940–5949. https://doi.org/10.1016/j.vaccine.2011.06.053.

Chauhan, V.S., Yazdani, S.S., Gaur, D., 2010. Malaria vaccine development based on merozoite surface proteins of Plasmodium falciparum. Hum Vaccin. 6, 757–762 http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20861668.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: A computer program to estimate gene genealogies. Mol Ecol. 9, 1657–1659.

Datoo, M.S., Natama, M.H., Somé, A., Traoré, O., Rouamba, T., Bellamy, D., et al., 2021. Efficacy of a low-dose candidate malaria vaccine, R21 in adjuvant Matrix-M, with seasonal administration to children in Burkina Faso: a randomised controlled trial. Lancet. 397, 1809–1818.

Ebomwonyi, A., Omoregie, A.O., Noutcha, M.A.E., Okiwelu, S.N., 2015. Trend in Malaria Incidence Rates (2006–2013) in Edo State, Nigeria. Int J Trop Dis Heal 7, 40–48. https://doi.org/10.9734/IJTDH/2015/15093.

Escalante, A.A., Grebert, H.M., Isea, R., Goldman, I.F., Basco, L., Magris, M., et al., 2002. A study of genetic diversity in the gene encoding the circumsporozoite protein (CSP) of Plasmodium falciparum from different transmission areas - XVI. Asembo Bay Cohort Project. Mol Biochem Parasitol. 125, 83–90.

Fadel, A.N., Ibrahim, S.S., Tchouakui, M., Terence, E., Wondji, M.J., Tchoupo, M., et al., 2019. A combination of metabolic resistance and high frequency of the 1014F kdr mutation is driving pyrethroid resistance in Anopheles coluzzii population from Guinea savanna of Cameroon. Parasites and Vectors. 12, 1–13. https://doi.org/10.1186/s13071-019-3523-7.

Fukuda, N., Tachibana, S.I., Ikeda, M., Sakurai-Yatsushiro, M., Balikagala, B., Katuro, O. T., et al., 2021. Ex vivo susceptibility of Plasmodium falciparum to antimalarial drugs in Northern Uganda. Parasitol Int. 81, 102277 https://doi.org/10.1016/j.parint.2020.102277.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/96/NT. In: Nucleic Acids Symposium Series, 41, pp. 95–98.

Healer, J., Murphy, V., Hodder, A.N., Masciantonio, R., Gemmill, A.W., Anders, R.F., et al., 2004. Allelic polymorphisms in apical membrane antigen-1 are responsible for evasion of antibody-mediated inhibition in Plasmodium falciparum. Mol Microbiol. 52, 159–168.

Hofmann, N., Mwingira, F., Shekalaghe, S., Robinson, L.J., Mueller, I., Felger, I., 2015. Ultra-sensitive detection of plasmodium falciparum by amplification of multi-copy subtelomeric targets. PLoS Med. 12, e1001788.

Hudson, R.R., Slatkin, M., Maddison, W.P., 1992. Estimation of levels of gene flow from DNA sequence data. Genetics. 132, 583–589.

Jespersen, M.C., Peters, B., Nielsen, M., Marcatili, P., 2017. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 45, W24–W29.

Jimah, J.R., Salinas, N.D., Sala-Rabanal, M., Jones, N.G., David Sibley, L., Nichols, C.G., et al., 2016. Malaria parasite CelTOS targets the inner leaflet of cell membranes for pore- dependent disruption. Elife. 5, e20621.

Kar, N., Chauhan, K., Nanda, N., Kumar, A., Carlton, J.M., Das, A., 2016. Comparative assessment on the prevalence of mutation in the Plasmodium falciparum drug-resistant genes in two different ecotypes of Odisha State. India. Infect Genet Evol. 41, 78–83.

Lancet, The, 2021. Malaria vaccine approval: a step change for global health. Lancet. 398, 1381. https://doi.org/10.1016/S0140-6736(21)02235-2.

Leigh, J.W., Bryant, D., 2015. POPART: Full-feature software for haplotype network construction. Methods Ecol Evol. 6, 1110–1116.

Librado, P., Rozas, J., 2009. DnaSP v6 : a software for comprehensive analysis of DNA polymorphism data. Bioinforma Appl Note. 25, 1451–1452.

Mahmoudi, S., Keshavarz, H., 2018. Malaria vaccine development: The need for novel approaches: A review article. Iran J Parasitol. 13, 1–10.

Ndiaye, D., Daily, J.P., Sarr, O., Ndir, O., Gaye, O., Mboup, S., et al., 2005. Mutations in Plasmodium falciparum dihydrofolate reductase and dihydropteroate synthase genes in Senegal. Trop Med Int Heal. 10, 1176–1179.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Ngassa-Mbenda, H.G., Das, A., 2016. Analysis of genetic diversity in the chloroquine-resistant gene Pfcrt in field Plasmodium falciparum isolates from five regions of the southern Cameroon. Infect Genet Evol. 44, 450–458.

NMEP, 2014. A Description of the Epidemiology of Malaria to Guide the Planning of Control in Nigeria http://www.inform-malaria.org/wp-content/uploads/2014/04/Nigeria-Epi-Report-V2–060214.pdf.

Olusegun-Joseph, T., Oboh, M., Awoniyi, A., Adebowale, A., Agbaso, M., Fagbohun, I., 2020. Efficacy of piperonyl butoxide (PBO) synergist on pyrethroid and dichlorodiphenyl trichloroethane (DDT) resistant mosquitoes in Lekki, Lagos State, Nigeria. Anim Res Int. 17, 3821–3828.

Opondo, K.O., Jawara, M., Cham, S., Jatta, E., Jarju, L., Camara, M., et al., 2019. Status of insecticide resistance in *Anopheles gambiae* (s.l.) of the Gambia. Parasites and Vectors. 12, 1–8. https://doi.org/10.1186/s13071-019-3538-0.

Osier, F.H., Mackinnon, M.J., Crosnier, C., Fegan, G., Wanaguru, M., Ogada, E., et al., 2016. New antigens for multi-component blood-stage vaccines against Plasmodium falciparum malaria, 6. https://doi.org/10.1126/scitranslmed.3008705.New.

Ouattara, A., Laurens, M.B., 2015. Vaccines against malaria. Clin Infect Dis. 60, 930–936.

Ouattara, A., Tran, T.M., Doumbo, S., Matthew, A., Agrawal, S., Niangaly, A., et al., 2018. Extent and dynamics of polymorphism in the malaria vaccine candidate plasmodium falciparum reticulocyte-binding protein Homologue-5 in Kalifabougou, Mali. Am J Trop Med Hyg. 99, 43–50.

Pirahmadi, S., Zakeri, S., Mehrizi, A.A., Djadid, N.D., 2018. Analysis of genetic diversity and population structure of gene encoding cell-traversal protein for ookinetes and sporozoites (CelTOS) vaccine candidate antigen in global Plasmodium falciparum populations. Infect Genet Evol. 59, 113–125. https://doi.org/10.1016/j.meegid.2018.01.023.

Pirahmadi, S., Zakeri, S., Mehrizi, A., Djadid, N., Raz, A.A., Sani, J., et al., 2019. Cell-traversal protein for ookinetes and sporozoites (CelTOS) formulated with potent TLR adjuvants induces high-affinity antibodies that inhibit Plasmodium falciparum infection in Anopheles stephensi. Malar J. 18, 146. https://doi.org/10.1186/s12936-019-2773-3.

The malERA Refresh Consultative Panel on Tools for Malaria E. malERA: An updated research agenda for insecticide and drug resistance in malaria elimination and eradication. PLoS Med. 14, 2017 e1002450.

Pwalia, R., Joannides, J., Iddrisu, A., Addae, C., Acquah-Baidoo, D., Obuobi, D., et al., 2019. High insecticide resistance intensity of *Anopheles gambiae* (s.l.) and low efficacy of pyrethroid LLINs in Accra, Ghana. Parasites and Vectors. 12, 1–9. https://doi.org/10.1186/s13071-019-3556-y.

Rasmussen, S.A., Ceja, F.G., Conrad, M.D., Tumwebaze, P.K., Byaruhanga, O., Katairo, T., et al., 2017. Changing antimalarial drug sensitivities in Uganda. Antimicrob Agents Chemother. 61, e01516.

Ridzuan, M., Abd, M., Sastu, U.R., Norahmad, N.A., Abdul-karim, A., Muhammad, A., et al., 2016. Genetic diversity of plasmodium falciparum populations in malaria declining areas of. PLoS One. 11, e0152415.

Straimer, J., Gandhi, P., Renner, K.C., Schmitt, E.K., 2022. High prevalence of plasmodium falciparum K13 mutations in rwanda is associated with slow parasite clearance after treatment with artemether-lumefantrine. J Infect Dis. 225, 1411–1414.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123, 585–595.

Terheggen, U., Drew, D.R., Hodder, A.N., Cross, N.J., Mugyenyi, C.K., Barry, A.E., et al., 2014. Limited antigenic diversity of Plasmodium falciparumapical membrane antigen 1 supports the development of effective multi-allele vaccines. BMC Med. 12, 183. https://doi.org/10.1186/s12916-014-0183-5.

Tumwebaze, P.K., Katairo, T., Okitwi, M., Byaruhanga, O., Orena, S., Asua, V., et al., 2021. Drug susceptibility of Plasmodium falciparum in eastern Uganda: a longitudinal phenotypic and genotypic study. The Lancet Microbe. 2, e441–e449. https://doi.org/10.1016/s2666-5247(21)00085-9.

Umunnakwe, F.A., Idowu, E.T., Ajibaye, O., Etoketim, B., Akindele, S., Shokunbi, A.O., et al., 2019. High cases of submicroscopic Plasmodium falciparum infections in a suburban population of Lagos. Nigeria. Malar J. 18, 433. https://doi.org/10.1186/s12936-019-3073-7.

Uwimana, A., Legrand, E., Stokes, B.H., Ndikumana, J.L.M., Warsame, M., Umulisa, N., et al., 2020. Emergence and clonal expansion of in vitro artemisinin-resistant Plasmodium falciparum kelch13 R561H mutant parasites in Rwanda. Nat Med. 26, 1602–1608.

Van Loon, W., Oliveira, R., Bergmann, C., Habarugira, F., Ndoli, J., Sendegeya, A., et al., 2022. In vitro confirmation of artemisinin resistance in plasmodium falciparum from patient isolates, Southern Rwanda, 2019. Emerg Infect Dis. 28, 852–855.

Watterson, G.A., 1975. On the number of segregating sites in genetical models without recombination. Theor Popul Biol. 276, 256–276.

Weedall, G.D., Mugenzi, L.M.J., Menze, B.D., Tchouakui, M., Ibrahim, S.S., Amvongo-Adjia, N., et al., 2019. A cytochrome P450 allele confers pyrethroid resistance on a major African malaria vector, reducing insecticide-treated bednet efficacy. Sci Transl Med. 11.

WHO, 2021. World Malaria Report.

Zhang, Q., Xue, X., Qu, L., Pan, W., 2007. Construction and evaluation of a multistage combination vaccine against malaria. Vaccine. 25, 2112–2119.