



OPEN ACCESS

EDITED BY

Mark C. Field,
University of Dundee, United Kingdom

REVIEWED BY

Mahamadou Ali Thera,
Université des Sciences, des Techniques et
des Technologies de Bamako, Mali
Marcus Lee,
University of Dundee, United Kingdom

*CORRESPONDENCE

Jesse Gitaka
✉ jgitaka@mku.ac.ke

RECEIVED 28 November 2023

ACCEPTED 25 March 2024

PUBLISHED 08 April 2024

CITATION

Maina M, Musundi S, Kuja J, Waweru H,
Kiboi D, Kanoi BN and Gitaka J (2024)
Genetic variation of the *Plasmodium*
falciparum circumsporozoite protein in
parasite isolates from Homabay
County in Kenya.
Front. Parasitol. 3:1346017.
doi: 10.3389/fpara.2024.1346017

COPYRIGHT

© 2024 Maina, Musundi, Kuja, Waweru, Kiboi,
Kanoi and Gitaka. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Genetic variation of the *Plasmodium falciparum* circumsporozoite protein in parasite isolates from Homabay County in Kenya

Michael Maina^{1,2}, Sebastian Musundi¹, Josiah Kuja¹,
Harrison Waweru^{1,2}, Daniel Kiboi², Bernard N. Kanoi¹
and Jesse Gitaka^{1*}

¹Centre for Malaria Elimination, Institute of Tropical Medicine, Mount Kenya University, Thika, Kenya,

²Department of Biochemistry, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

The *Plasmodium falciparum* Circumsporozoite Protein (PfCSP) has been used in developing the RTS,S, and R21 malaria vaccines. However, genetic polymorphisms within *PfCSP* compromise the effectiveness of the vaccine. Thus, it is essential to continuously assess the genetic diversity of *PfCSP*, especially when deploying it across different geographical regions. In this study, we assessed the genetic diversity of the *PfCSP* on isolates from Homabay County, a malaria-endemic region in western Kenya, and compared it against other isolates from Kenya. We extracted DNA from 27 microscopically confirmed *P. falciparum* positive samples and conducted Illumina sequencing to generate paired-end short reads. The sequences were then mapped to the Pf3D7 reference genome, and genetic variation was analyzed using bcftools. Additionally, we retrieved isolates from two other malaria-endemic regions in Kenya, Kisumu (n=58) and Kilifi (n=596), from MalariaGEN version 7 and compared their genetic diversity and natural selection. We also evaluated the predicted binding affinities for HLA class I and II supertype alleles for the identified haplotypes using NetMHCpan and NetMHCIIpan. Our results show that the N-terminal of PfCSP was relatively conserved with a notable mutation at A98G across all isolates. The number of NANP repeats varied across the three Kenyan sites within the central repeat region. Furthermore, the C-terminal region showed polymorphism within the Th2R and Th3R regions. Haplotype network analysis of the Kenyan isolates revealed 69 haplotypes, with the 3D7 reference being found in the most prevalent haplotype. When assessing the predicted binding affinities between superotypes in HLA class I and II with the identified haplotypes, we observed stronger predicted binding affinities to multiple haplotypes except for those containing the 3D7 reference. The results suggest the need to take into account the existing changes occurring in *PfCSP* while developing malaria vaccines.

KEYWORDS

circumsporozoite protein, vaccines, genetic diversity, *Plasmodium falciparum*, malaria

1 Introduction

Malaria poses a huge mortality burden in Sub-Saharan Africa and tropical countries. In 2021, an estimated 241 million infections were reported worldwide, causing about 627,000 deaths (World Health Organization, 2021). This burden was escalated by the COVID-19 pandemic disruptions of existing health systems (World Health Organization, 2021). Despite a declining trend of malaria incidences over the past decade (Rosenthal et al., 2019), pregnant women and children under five still bear the highest burden. *P. falciparum*, the deadliest among the *Plasmodium* spp (Murmu et al., 2021), has a complex life cycle alternating between human and mosquito vectors. This, coupled with a lack of effective vaccines, parasite resistance to the available antimalarials, and mosquito resistance to insecticides, hinder efforts to prevent, control and eliminate malaria (Dimala et al., 2018; Cowell and Winzeler, 2019; Maier et al., 2019).

The WHO rolled out the malaria vaccine initiative to support the fight against malaria (Birkett, 2010; Samuels et al., 2022). It is anticipated that a vaccine, in conjunction with other control measures, can synergize and complement each other to eliminate malaria (Dobanõ et al., 2019). One key target for vaccine development is *P. falciparum* circumsporozoite protein (PfCSP), the major abundant antigen expressed on the surface of sporozoites critical for the invasion of sporozoites in the human liver. Antibodies that target PfCSP have been shown to prevent sporozoite invasion of hepatocytes (Hollingdale et al., 1984; Cohen et al., 2010). The RTS,S (Mosquirix), and R21 are examples of vaccines based on *PfCSP* sequence. RTS,S and R21 are currently World Health Organization (WHO)-recommended malaria vaccines for widespread use among children living in malaria-endemic countries (Zavala, 2022).

Although RTS,S and R21 are based on *PfCSP*, slight differences exist structurally. RTS,S/AS01 is a sub-unit vaccine comprising of *PfCSP* region fused with hepatitis B surface antigen (HBsAg), alongside additional unfused HBsAg assembled into virus-like

particles (VLPs) which improve antigen presentation to unique T-cell epitopes (Sinnis and Fidock, 2022). Like RTS,S/AS01, R21/Matrix-M combines a VLP expressing PfCSP-HBsAg fusion protein with a saponin-based adjuvant. Unlike RTS,S, the R21 VLP excludes unfused HBsAg, enabling a 1:1 ratio between PfCSP and Hepatitis B antigen (Collins et al., 2017) (Figure 1). Notably, R21 showed at least 75% efficacy over 6 months among Burkina Faso children aged between 5–17 months in phase 2 trials after three rounds of immunization and similar levels of protection even after 1 year of the booster dose (Datoo et al., 2021, 2022). Some reports indicate diminishing effectiveness of vaccines, particularly against *P. falciparum* strains with diverse *PfCSP* alleles to the region used to make the RTS,S vaccine (Neafsey et al., 2015; Pringle et al., 2018; Laurens, 2019). Thus, understanding the role of genetic diversity of *PfCSP* has potential implications for vaccine rollout in countries with areas of highly seasonal malaria transmission (Mohamed et al., 2021; Merle et al., 2023).

PfCSP comprises three regions: a conserved N-terminal region, a central repeat region, and a C-terminal region (Figure 1). The N-terminal region contains a conserved KLKQP motif required for hepatocyte entry (Sidjanski et al., 1997). The central repeat region is rich in NANP (Asn-Ala-Asn-Pro) tandem repeats and NVDP (Asn-Val-Asp-Pro) and is recognized as the site for antibody-mediated neutralization. The C-terminal region has two polymorphic Th2R and Th3R sub-regions recognized by CD8+ and CD4+ T-cells respectively (Putaporntip et al., 2009; He et al., 2022) (Figure 1). Natural selection related to host immunity is hypothesized to be responsible for the polymorphism of these sub-regions (Zheng et al., 2019; Huang et al., 2020). The interaction between PfCSP residues and T-cells is mediated by both HLA class I and II expressed on the surface of antigen-presenting cells (APC), which is essential for T-cell immune response. Polymorphisms observed in the Th2R and Th3R region may directly affect the binding affinity of different PfCSP amino acid residues and T-cells, which can enhance or reduce T-cell immune responses. Understanding vaccine candidates' genetic and immunological

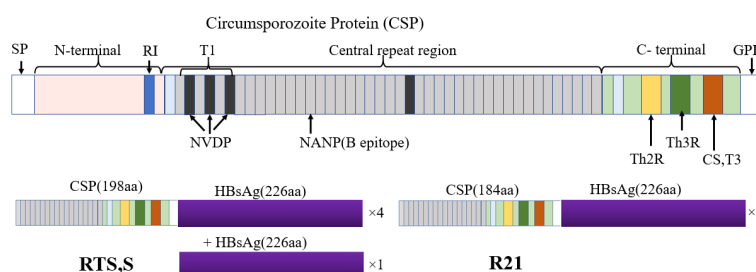


FIGURE 1

Structure of PfCSP and its associated subunit vaccines RTS,S and R21. PfCSP is divided into three regions; N-terminal region, central repeat region and C-terminal region. PfCSP also contains a signal peptide at the N-terminal end and a glycosylphosphatidylinositol (GPI) at the C-terminal. The N-terminal contains the R1 region, a conserved motif required for sporozoite invasion. The central repeat region contains NANP (Asn-Ala-Asn-Pro) tandem repeats and NVDP (Asn-Val-Asp-Pro). The central repeat region is primarily a B-cell epitope region. The C-terminal region contains three T-cell epitope regions including the polymorphic CD4+ T cell epitope (Th2R), CD8+ T cell (Th3R) and a conserved universal CD4 T cell epitope (CS.T3). The Th3R and CS.T3 region form part of Region II, a thrombospondin like domain. The entire structure of PfCSP is made up of 420 amino acids in the 3D reference. The RTS,S and R21 are subunit vaccines of PfCSP. RTS,S comprises 189 amino acids derived from the last 18 NANP repeats and the C-terminal region with the exception of the GPI anchor fused to four Hepatitis B surface Antigen (HBsAg) (226 amino acids). R21 comprises 184 amino acids derived from NANP repeats and C-terminal region but fused to one HBsAg.

interactions is critical in designing effective vaccines. In this study, we assessed the genetic diversity of *Pfcsf* utilizing paired-end short-read sequences of *P. falciparum* field isolates from Homabay County, a malaria-endemic region in western Kenya. We further compared our results with parasite populations from two malaria-endemic regions in Kenya using *Pfcsf* sequences retrieved from the global MalariaGEN database v7.0. Finally, we predicted the binding affinities of 3D7 reference and non-3D7 variants among the major HLA class I and II supertypes to assess their impact on T-cell recognition.

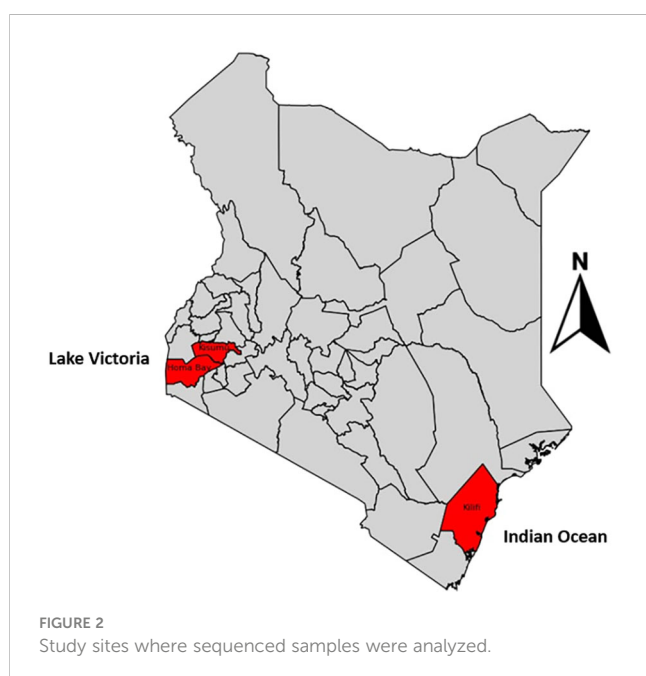
2 Methods

2.1 Study site

The study was conducted at Homabay Hospital in western Kenya, a malaria-endemic region associated with high malaria transmission intensities. We also included previously sequenced data from MalariaGEN version 7 from Kisumu and Kilifi. Homabay and Kisumu are lake malaria-endemic regions due to their closeness to Lake Victoria, while Kilifi is a coast-endemic region located next to the Indian Ocean (Figure 2).

2.2 Ethical approval

Approval for the study was obtained from the Mount Kenya University Ethics and Research Review Committee (MKU-ERC; approval P609/10/2014). All participants or guardians provided written consent, and the study adhered to relevant guidelines and regulations. The community was sensitized, and consent was sought before participation, with malaria cases being treated according to national guidelines.



2.3 Sample collection, genomic DNA extraction, and sequencing

The samples analyzed in this study were a subset of the larger surveillance studies, which have been extensively described elsewhere (Idris et al., 2016; Gitaka et al., 2017; Osborne et al., 2021). *P. falciparum* species were identified by microscopy following the WHO protocols and confirmed using established nested PCR assays as reported (Osborne et al., 2021). Monoclonal *P. falciparum* parasites obtained from the field were adapted for *in vitro* culture (Osborne et al., 2021). DNA was extracted from short-term parasite cultures (1 month) using the QIAamp DNA mini kit (Qiagen, Valencia, CA). Paired-end sequencing libraries were prepared using Nextera XT DNA library preparation Kit following the manufacturer's protocol (Illumina, USA). Whole genome sequencing was carried out on Illumina MiSeq technology generating reads of 300bp. These sequences are archived at the DDBJ BioProject, Accession number PRJDB12148 and EBI SRA Project accession PRJEB46180.

2.4 Sequence analysis

In this study, the quality of the generated paired-end raw reads was assessed using FastQC. The Plasmodium reference genome (version 63) was retrieved from PlasmoDB and indexed using the Burrows-Wheeler Alignment (BWA) (Li and Durbin, 2009). Subsequently, the paired-end reads were mapped to the indexed Plasmodium genome, generating a SAM file, which was converted, sorted, and indexed into a BAM file. Samtools flagstat was then used to check for the percentage of reads that mapped to the genome. Comprehensive mapping statistics, including sequencing depth and coverage, were also assessed using Samtools (Li et al., 2009). Afterwards, Picard was used to pre-process the generated bam files for variant calling by sorting the order of coordinates, adding read groups, and marking duplicates. All the pre-processed bam files were merged using Samtools merge and were subsequently sorted and indexed. Variant calling analysis was performed using bcftools' mpileup and call functions. Single nucleotide polymorphisms (SNP) and Indels were then filtered to retrieve variants utilizing a haploid model for those with a base quality score >20 and minimum raw depth >=5. Bcftools view was used to filter biallelic SNPs with a minor allele frequency >0.01 and excluded samples which lacked greater than 10% of the genotypes. Samtools version 1.18 and bcftools consensus was employed to retrieve the CSP protein across all 27 samples (Li et al., 2009). The generated Fasta files were merged, nucleotides translated and aligned using MUSCLE.

We further retrieved the VCF files for chromosome 3 for Kisumu (n=58), a lake endemic region, and Kilifi (n=596), a coast endemic region in Kenya from the MalariaGEN database v7.0. The procedure on how the variant files were generated is previously described (Abdel Hamid et al., 2023). We utilized Gene Analysis Toolkit (GATK) Haplotype caller algorithm to call for SNPs and indels and SelectVariant algorithm to separate SNPs and indels. We also generated merged files containing SNPs and indels using GATK's MergeVCF algorithm. Subsequently, consensus FASTA

files were generated for Kisumu (n=58) and Kilifi (n=596) isolates for *Pfmsp* using Samtools and Bcftools consensus. This involved the replacement of the 3D7 reference nucleotide sequences at specific positions where SNPs and indels were identified. Following this, MEGA suite was employed to translate nucleotides to amino acid residues and aligned each isolate against the *P. falciparum* 3D7 reference using MUSCLE. Metadata for sequence IDs retrieved in MalariaGEN for Kisumu (n=58) and Kilifi (n=596) are provided in [Supplementary Table 1](#).

We also aimed to examine population genetic differentiation in the Kilifi samples (n=596) collected across different years (1994–2018). Population genetic differentiation was examined using the fixation index (Fst). Samples collected in individual year and same site were placed under the same population and Fst was calculated using Weir and Cockerham in vcftools. Fst values range from 0 to 1 with high Fst valued indicating a substantial degree of differentiation among the population. The Fws metric was also assessed; this evaluates the diversity relation within an individual sample compared to the population by assessing the probability of two random parasites having different alleles at specific positions. Fws ranges from 0 to 1 with values ≥ 0.95 representing high inbreeding rates resulting into low host diversity while values $Fws \leq 0.95$ signifying high inbreeding rates and therefore high host diversity. Thus, $Fws \geq 0.95$ represents monoclonal isolates while $Fws \leq 0.95$ represents polyclonal isolates. For Kisumu (n=58) and Kilifi(n=596) isolates, Fws was previously calculated as described ([Abdel Hamid et al., 2023](#)).

2.5 Genetic diversity of the C-terminal region

Genetic polymorphisms and neutrality tests were assessed the C-terminal region (909–1140 bp) using isolates from Homabay, Kisumu, and Kilifi using DnaSP version 6.12.03 ([Rozas et al., 2017](#)). Nucleotide diversity scores for the entire region spanning Th2R and Th3R were obtained and compared across the three sites using a sliding window of 10bp and a step size of 5bp. We also assessed the number of segregating sites (S), number of haplotypes (h), and haplotype diversity (Hd). We also assessed the neutrality of single nucleotide polymorphisms, Tajima D statistical test, Fu and Li's D and Fu and Li's F for the entire C-terminal region ([Tajima, 1989](#); [Fu and Li, 1993](#)). A web logo plot was also constructed for amino acid residues for Th2R (310–327 aa) and Th3R (352–363 aa) and compared against the 3D7 reference to analyze variant patterns using the WebLogo program. A haplotype network was created to analyze the relationship between the different isolates collected in the three sites in Kenya using the R package geneHapR ([Zhang et al., 2023](#)).

2.6 *In silico* prediction of the effect of mutations on vaccine efficacy

Since the RTS,S, and R21 vaccines serve as the structural basis for the CSP protein, we sought to examine the impact of individual

mutations on CD8+ and CD4+ T-cell responses implicated in CSP immune responses. The Th2R and Th3R regions are associated with CD8+ and CD4+ T-cell responses. As CD4+ T-cells recognize antigens presented through to MHC class II, we utilized NetMHCIIpan 4.0 to predict the binding strength of CSP peptides in the Th3R region utilizing PlasmoDB 3D7 reference sequence and a representative sequence from each of the predicted haplotype ([Reynisson et al., 2020](#)). Our selection comprised nine HLA supertypes class II HLA-DRB1 alleles (*0101, *0301, *0401, *0701, *0801, *0901, *1101, *1301, and *1501) that facilitate evaluation of over 95% of the human population HLA supertypes ([Reynisson et al., 2020](#); [Kotraiah et al., 2021](#)). We also assessed the impact of variations on the binding affinity of MHCII to CD8+ T-cells lined to Th2R utilizing NetMHCpan 4.1 ([Reynisson et al., 2021](#)). For this analysis, we utilized six class I supertype HLA alleles A*0101, A*0201, A*0301, A*2402, B*0702, and B*4403, which cover 90% of the population ([Tucker et al., 2021](#)).

3 Results

3.1 Genetic diversity in the N-terminal region

As expected, the N-terminal region remained relatively conserved across Kenyan isolates ([Figure 3](#)). When compared to the 3D7 reference genome, one variant A98G (6/27, 22.2%), was reported across Homabay isolates. Further comparison with other Kenyan isolates from MalariaGen version 7 revealed the presence of a similar variant at A98G in Kisumu (12/58, 20.7%) and Kilifi (133/596, 22.3%). In addition, low-frequency variants (<1%) were also reported in the Kilifi isolates at Q24A (1/596) and N49S (1/596) ([Figure 3](#)).

3.2 Genetic diversity in the central repeat region

Within the central repeat region, the NANP region remained relatively conserved across Kenyan isolates. Nonetheless, low-frequency variants were identified in Homabay at V126A (1/27), D127N (1/27), and A158V (1/27). In Kilifi variants were found at A186V (2/596), A242V (1/596), D288N (1/596), D294N (5/596). No minor variants were reported across NANP central repeat region in Kisumu isolates. One major variant, S301N, occurred at a high frequency (>80%) across Homabay (26/27, 96.29%), Kisumu (54/58, 93.10%), and Kilifi isolates (502/596, 84.22%) ([Figure 3](#)). The number of NANP repeats present in the Homabay samples ranged from 33 to 39, with 36 being the most frequent (15/27, 55.55%). In Kisumu, the range was 34–40, with the highest frequencies observed at 39 (29/58, 50%) and 38 (28/58, 48.28%). In Kilifi, the number of NANP repeats ranged from 29 to 41, with 33 being the most frequent (296/596, 49.66%).

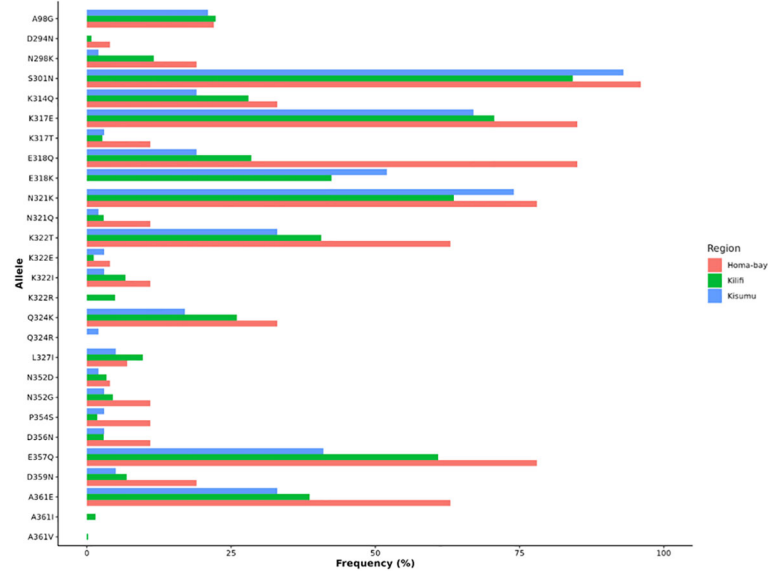


FIGURE 3
Frequency of mutations found in CSP across Homabay, Kisumu, and Kilifi isolates.

3.3 Genetic diversity in the C-terminal region across Kenyan samples

Amino acid changes largely occurred in the Th2R and Th3R region of the C-terminal region, as presented in Figure 4. A total of four logos are presented, with the first indicating the 3D7 reference strain, while the remaining show isolates from Homabay, Kisumu, and Kilifi. Compared to the 3D7 reference sequence, amino acid changes with high frequency variants (>5%) were observed across 13 positions in the Th2R (K314Q, K317E/T, E318Q/K, N321Q/K, K322T/E/I and Q324K/R, L327I) and Th3R regions (N352D/G, P354S, D356N, E357Q, D359N and A361E/I/V) (Figure 2). The

changes in K317E, E318K, N321K, and K322T were more common in the Th2R region, while Q324K, E357Q, and A361E in Th3R (Figure 3).

Nucleotide diversity of the C-terminal region was calculated using DnaSP for the Homabay, Kisumu and Kilifi isolates using a sliding window of 10 bp and a step size of 5 bp. The nucleotide diversity for Homabay isolates was (P_i : 0.02068), Kilifi (P_i : 0.02019), and Kisumu (P_i : 0.01776). The regions containing the T-cell epitopes in the Th2R and Th3R regions both showed high nucleotide diversity, while the region between Th2R and Th3R was relatively conserved (Figure 5). Other parameters associated with natural selection were also evaluated. Population genetic

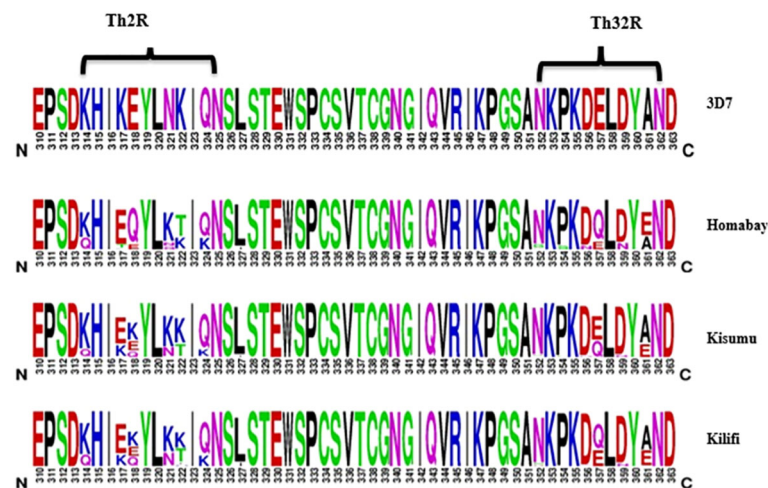


FIGURE 4
Non-synonymous mutations in the C-terminal region (311–363) of Kenya isolates from Homabay, Kisumu and Kilifi in comparison to the reference 3D7. The height of amino acid indicates their relative frequency at that position. Most amino acid changes were observed between the Th2R and Th3R region. The web logo plot was constructed using the WebLogo program.

differentiation analysis using F_{st} using Weir and Cockerham was 0.0047791 for the different population years in Kilifi suggesting a low degree of differentiation for the Kilifi isolates collected over the different years. Both Tajima D (0.306, $p > 0.1$), Fu and Li's D (1.22904, $p > 0.1$), and Fu and Li's F (1.10466) were slightly positive but not significant in Homabay samples and across isolates in Kisumu and Kilifi within the C-terminal region which could suggest weak balancing selection (Table 1). The haplotype diversity (H_d) for all isolates across the three regions was high (> 0.9), indicating high levels of genetic diversity (Table 1).

3.4 Haplotype network analysis

A total of 681 isolates from Homabay ($n=27$), Kisumu ($n=58$), and Kilifi ($n=596$) together with the 3D7, were used to construct the haplotype network. The sequences were clustered into 69 unique haplotypes (H001 to H069). Detailed information for each of the haplotypes is provided in Supplementary Table 2. The RTS,S vaccine haplotype, which contains the reference 3D7 strain and denoted as H001, represented 12% (86/681) of all the isolates. Notably haplotype H001 was absent in Homabay isolates (Figure 6). Among the haplotypes, H002, H003, H011, and H016 were present across all three sites, while haplotypes H031 and H069 were exclusively present in Homabay isolates. An additional 42 singletons were identified, suggesting that genetic polymorphisms were occurring within the Kenyan isolates.

3.5 Interaction between HLA class I and II and *Pf*csp haplotypes

For MHC class I, HLA-A alleles 01:01, 02:01, and 03:01 showed higher predicted binding affinities across the 69 haplotypes compared to the other supertype HLA-A alleles 03:01, 24:02, 07:02, and B*44:03 whose predicted binding affinities were below the threshold set for strong or weak binders. Of the three HLA-A

alleles which showed both strong and weak binding affinities, allele HLA-A 01:01 showed stronger binding affinities across multiple different haplotypes (Figure 7). For HLA class II, DRB1_1301 showed stronger binding affinities to multiple haplotypes as compared to the rest of the alleles (Figure 8). Nonetheless, HLA class II alleles DRB1_1501, DRB1_1101, DRB1_0801, and DRB1_0401 recognized multiple different HLA alleles with either strong or weak binding activity. The reference strain 3D7 showed weaker binding affinities outside the required thresholds for all HLA class I and II supertypes used in this analysis all haplotypes. Lastly, we did not observe any HLA class I or II allele supertype which showed strong binding affinity to any haplotype.

4 Discussion

Vaccine development is a crucial aspect of malaria control and elimination. Various antigens have been identified as potential vaccine candidates for malaria, including RTS,S and R21, both based on *Pf*csp. A key challenge in malaria vaccine development is the high degree of antigenic diversity due to immune pressure, resulting in different antigenic variants from the reference strain circulating within populations and therefore reducing the effectiveness of current malaria vaccines. This limitation is particularly relevant in endemic malaria regions. It could explain the partial protection reported in RTS,S within the first 12 months and the subsequent waning of the immune response over the next 4 years in children aged between 5 and 17 months (RTS, 2015).

The N-terminal of PfCSP remained relatively conserved across Kenyan isolates, which coincides with other isolates across other regions (Lê et al., 2018; Mohamed et al., 2019; Huang et al., 2020). The N-terminal of CSP binds to heparan sulfate proteoglycans, and cleavage to a conserved region I (KLKQP) by a parasite protease results in to change from being adhesive to invasive to hepatocytes. Specifically, antibodies to the region I found between amino acid ⁸²(DNEKLRKPKHKKLLKQPADG)¹⁰⁰ of the N-terminal inhibit the

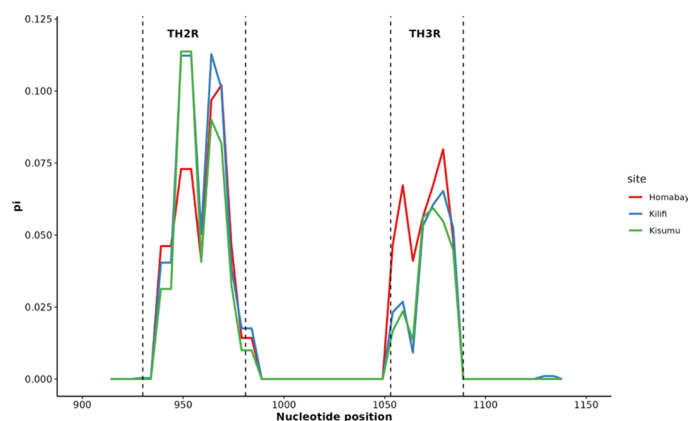


FIGURE 5

Nucleotide diversity in Th2R and Th3R C-terminal region across Homabay, Kilifi and Kisumu isolates. Nucleotide diversity was calculated using DnaSP with a sliding window of 10 bp with a step size of 5 bp.

TABLE 1 Nucleotide diversity and tests for neutrality in the C-terminal region of *Pfcs*p in Kenyan isolates.

	n	H	S	Hd \pm SD	Tajima D	Fu and Li's D	Fu and Li's F
Homabay	27	13	17	0.915 \pm 0.00085	0.30688*	1.22904*	1.10466*
Kilifi	596	104	21	0.950 \pm 0.00002	0.76834*	0.82698*	0.97593*
Kisumu	58	29	17	0.962 \pm 0.00010	0.87513*	0.87513*	0.97593*

(n, Number of isolates; H, Number of haplotypes; S, Number of segregating polymorphic sites; Hd \pm SD, Haplotype diversity; *p>0.10).

cleavage of CSP and impair invasion both *in vivo* and *in vitro* (Espinosa et al., 2015). In the current results, the non-synonymous mutation A98G was observed in high frequency, as previously reported across most global isolates within region I (Huang et al., 2020). The presence of this non-synonymous mutation suggests the area is under selection pressure and may result in structural changes that have an impact on the sporozoite invasion process in the hepatocyte. Despite this, the conserved motif (KLKQP) involved in heparin sulfate binding remained well conserved across Kenyan isolates. The conserved nature of the N-terminal region, coupled with its key functions, collectively makes it an attractive component of the PfCSP vaccine (Bongfen et al., 2009).

The central repeat region in CSP contains species-specific 4-amino acids NANP/NVDP repeats, which make up the part of the immunodominant B-cell epitopes recognized by neutralizing antibodies. Furthermore, the region has been shown to contribute to sporozoite growth in the mosquito. The different number of tetrapeptide NANP repeats was observed across Kenyan samples, contributing to this region's polymorphisms. The results are consistent with the number of NANP repeats reported across other Kenyan isolates (Lê et al., 2018) and African isolates (Huang et al., 2020). The differing number of repeats has been shown to affect the PfCSP structure. More specifically, Escalante et al. reported that the stability of type- β PfCSP is increased with the number of repeats (Escalante et al., 2002). The current RTS, S vaccines have 19 NANP tetrapeptide repeats, and to date, there is no

direct evidence linking the number of NANP repeats to vaccine efficacy. Nonetheless, targeting regions with repeat regions has been shown to result in less durable antibody responses. Repeat regions such as NANP have been suggested to be decoy regions used by the parasite to prevent the immune system from mounting a strong neutralizing effect.

A high level of polymorphism was observed within the Th2R and Th3R regions of the C-terminal of PfCSP. The Th2R region was more polymorphic compared to the Th3R regions, as previously reported (Amegashie et al., 2020; Dieng et al., 2023). Considering ThR2 and ThR3 regions are involved in the CD4+ and CD8+ T cell responses, the presence of polymorphisms may help the parasite escape the immune pressure from the host, which has a direct impact on the overall efficacy of the vaccine. These results are similar to previous reports, which showed between one and six changes in amino acid occurring within the Th2R and Th3R regions of Zambia, Ghana, and DRC populations (Pringle et al., 2018).

The weak positive Tajima's D values reported in the C-terminal region of *Pfcs*p across the three different sites in Kenya which could suggest balancing selection primarily due to host immune pressure on the C-terminal. Nonetheless, the positive Tajima's D values across the three sites were non-significant and therefore inconclusive. Evidence of balancing selection has previously been reported across isolates from Navirongo in Ghana and populations from Malawi (Bailey et al., 2012; Amegashie et al., 2020). The weak positive Tajima's D values could be further explained by the fact

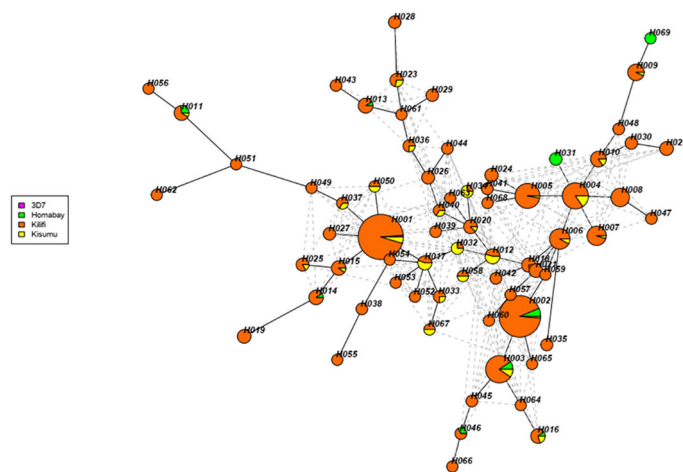


FIGURE 6

Haplotype network analysis of Kenyan *Pfcs*p C-terminal region. The haplotype network was constructed using the R package gene "geneHapR". The networks show a total of 69 haplotypes found in 691 isolates and the reference 3D7. The branch length is proportion to the divergence while the size of each node indicates the frequencies of each node. The color of each node represents the different site.

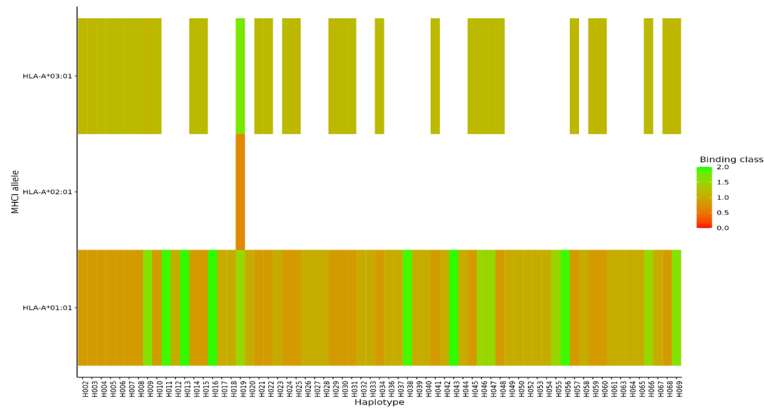


FIGURE 7

Binding affinities of peptide sequences of each haplotype after proteasomal cleavage of the C-terminal region in MHC I. Peptides classified as strong binders where the % Rank < 0.5% and weak binders where % Rank > 0.5% but less than 2% for MHC I.

that Homabay, Kisumu, and Kilifi are regarded as malaria endemic regions within Kenya, and therefore, raising the possibility that the variants could be maintained due to their selective advantage, such as conferring better immune response against malaria or being resistant to malaria. The overall values of haplotype diversity for the C-terminal were also high, suggesting a higher level of genetic diversity across Kenyan isolates which is a common feature of most African isolates (Huang et al., 2020). Further analysis of the nucleotide diversity was similar to other global isolates, showing a relatively similar plot composed of high genetic variation within the Th2R and Th3R regions.

Haplotype analysis of the C-terminal region revealed a branched pattern in terms of haplotype diversity. In addition, a total of 42 singletons were identified, suggesting Kenyan isolates show high levels of genetic diversity similar to other African isolates. The 3D7 reference genome used in the development of the RTS,S, and R21 vaccines shared the highest frequency with most isolates from Kisumu and Kilifi but not from Homabay. The genetic

heterogeneity might result in different levels of vaccine efficacy, especially in areas where the vaccine haplotype was absent (Pringle et al., 2018). Therefore, consistent monitoring of the existing haplotypes while administering vaccines that target the C-terminal region of CSP will be valuable, especially in high transmission areas. In addition, developing specific individual vaccines that target the most prevalent haplotype in malaria-endemic regions is another solution.

HLA class I and II typically exhibit different binding affinities among diverse populations. Polymorphisms in specific regions recognized by HLA alleles can either provide protection against a variety of pathogens or increase susceptibility to diseases. The strength of binding between antigenic peptides and the groove of the HLA, as well as with T-cell receptor peptides, determines TCR activation, signaling, and cell responsiveness (Baumgaertner et al., 2022). Given that regions Th2R and Th3R are crucial for evoking CD4+ and CD8+ responses, examining their interaction with HLA alleles offers insights into the impact of existing variants on T-cell



FIGURE 8

Binding affinities of peptide sequences of each haplotypes after proteasomal cleavage of the C-terminal region in MHCII. Peptides classified as strong binders if the % Rank < 2% and weak binders where % Rank > 2% but less than 10% for MHCII.

responses. Our predicted results indicate that multiple haplotypes exhibit better binding affinity compared to the reference strain 3D7 for both HLA class I and II supertypes. These findings align with observations from the evaluation of Ghanaian isolates. In that context, the 3D7 reference clearly bound to the HLA class II allele, suggesting that the reference might be binding to additional epitopes present in PfCSP, which were not assessed in this study (Dieng et al., 2023).

The samples used in this study were collected from Homabay in Kenya, one of the malaria-endemic regions in Kenya. The different malaria endemic regions have unique parasite transmission diversities that may continuously impact CSP diversity. To understand the complete CSP diversity in Kenya, it is important to assess mutations across the country continuously. Also, the study only utilized the major HLA class I and II supertypes present globally, and might not provide a representative picture of Kenyan HLA class I and II supertypes since which are currently not resolved especially among malaria endemic regions. Other factors such as that may influence HLA variation and T-cell responses, such as exposure differences, T-cell receptor variation, and vaccination status, were not examined. Lastly, the sequenced isolates used in analyzing the current data were cultured *in vitro* and therefore, there is also possibility of some mutations being introduced during the culturing process.

5 Conclusion

This study provides essential information on the genetic diversity and polymorphism of *PfCSP* from Kenya. The N-terminal region and central repeat region within the Kenyan isolates remained relatively conserved, while the Th2R and Th3R in the C-terminal region showed high levels of polymorphism. Lastly, the 3D7 reference did not bind weakly or strongly to supertype alleles, although multiple different haplotypes were able to be recognized by the same HLA alleles.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

The studies involving humans were approved by Mount Kenya University Ethics and Research Review Committee (MKU-ERC). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. The animal study was approved by Mount Kenya University Ethics and Research Review Committee (MKU-

ERC). The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

MM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SM: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – review & editing. JK: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – review & editing. HW: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – review & editing. DK: Data curation, Investigation, Supervision, Validation, Visualization, Writing – review & editing. BK: Conceptualization, Data curation, Formal analysis, Investigation, Supervision, Validation, Writing – review & editing. JG: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by BK an EDCTP Fellow under EDCTP2 programme supported by the European Union grant number TMA2020CDF-3203. JG received support from the African Academy of Sciences. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpara.2024.1346017/full#supplementary-material>

References

- Abdel Hamid, M. M., Abdelraheem, M. H., Acheampong, D. O., Ahouidi, A., Ali, M., Almagro-Garcia, J., et al. (2023). Pf7: an open dataset of *Plasmodium falciparum* genome variation in 20,000 worldwide samples. *Wellcome Open Res.* 8, 22. doi: 10.12688/wellcomeopenres.18681.1
- Amegashie, E. A., Amenga-Etego, L., Adobor, C., Ogoti, P., Mbogo, K., Amambua-Ngwa, A., et al. (2020). Population genetic analysis of the *Plasmodium falciparum* circumsporozoite protein in two distinct ecological regions in Ghana. *Malar. J.* 19, 473. doi: 10.1186/s12936-020-03510-3
- Bailey, J. A., Mvalo, T., Aragam, N., Weiser, M., Congdon, S., Kamwendo, D., et al. (2012). Use of massively parallel pyrosequencing to evaluate the diversity of and selection on *Plasmodium falciparum* csp T-cell epitopes in Lilongwe, Malawi. *J. Infect. Dis.* 206, 580–587. doi: 10.1093/infdis/jis329
- Baumgaertner, P., Schmidt, J., Costa-Nunes, C. M., Bordry, N., Guillaume, P., Luescher, I., et al. (2022). CD8 T cell function and cross-reactivity explored by stepwise increased peptide-HLA versus TCR affinity. *Front. Immunol.* 13. doi: 10.3389/fimmu.2022.973986
- Birkett, A. J. (2010). PATH malaria vaccine initiative (MVI): Perspectives on the status of malaria vaccine development. *Hum. Vaccin.* 6, 139–145. doi: 10.4161/hv.6.1.10462
- Bongfen, S. E., Ntsama, P. M., Offner, S., Smith, T., Felger, I., Tanner, M., et al. (2009). The N-terminal domain of *Plasmodium falciparum* circumsporozoite protein represents a target of protective immunity. *Vaccine* 27, 328–335. doi: 10.1016/j.vaccine.2008.09.097
- Cohen, J., Nussenzweig, V., Nussenzweig, R., Vekemans, J., and Leach, A. (2010). From the circumsporozoite protein to the RTS,S/AS candidate vaccine. *Hum. Vaccin.* 6, 90–96. doi: 10.4161/hv.6.1.9677
- Collins, K. A., Snaith, R., Cottingham, M. G., Gilbert, S. C., and Hill, A. V. S. (2017). Enhancing protective immunity to malaria with a highly immunogenic virus-like particle vaccine. *Sci. Rep.* 7, 46621. doi: 10.1038/srep46621
- Cowell, A. N., and Winzeler, E. A. (2019). The genomic architecture of antimalarial drug resistance. *Brief. Funct. Genomics* 18, 314–328. doi: 10.1093/bfpg/elz008
- Dattoo, M. S., Natama, H. M., Somé, A., Bellamy, D., Traoré, O., Rouamba, T., et al. (2022). Efficacy and immunogenicity of R21/Matrix-M vaccine against clinical malaria after 2 years' follow-up in children in Burkina Faso: a phase 1/2b randomised controlled trial. *Lancet Infect. Dis.* 22, 1728–1736. doi: 10.1016/S1473-3099(22)00442-X
- Dattoo, 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. doi: 10.1016/S0140-6736(21)00943-0
- Dieng, C. C., Ford, C. T., Lerch, A., Doniou, D., Vegesna, K., Janies, D., et al. (2023). Genetic variations of *Plasmodium falciparum* circumsporozoite protein and the impact on interactions with human immunoproteins and malaria vaccine efficacy. *Infect. Genet. Evol.* 110, 105418. doi: 10.1016/j.meegid.2023.105418
- Dimala, C. A., Kika, B. T., Kadia, B. M., and Blencowe, H. (2018). Current challenges and proposed solutions to the effective implementation of the RTS,S/AS01 Malaria Vaccine Program in sub-Saharan Africa: A systematic review. *PLoS One* 13, 1–11. doi: 10.1371/journal.pone.0209744
- Dobanõ, C., Ubbilos, I., Jairoce, C., Gyan, B., Vidal, M., Jiménez, A., et al. (2019). RTS,S/AS01E immunization increases antibody responses to vaccine-unrelated *Plasmodium falciparum* antigens associated with protection against clinical malaria in African children: A case-control study. *BMC Med.* 17, 157. doi: 10.1186/s12916-019-1378-6
- 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. doi: 10.1016/S0166-6851(02)00216-5
- Espinosa, D. A., Gutierrez, G. M., Rojas-Lopez, M., Noe, A. R., Shi, L., Tse, S. W., et al. (2015). Proteolytic cleavage of the *Plasmodium falciparum* circumsporozoite protein is a target of protective antibodies. *J. Infect. Dis.* 212, 1111–1119. doi: 10.1093/infdis/jiv154
- Fu, Y. X., and Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics* 133, 693–709. doi: 10.1093/genetics/133.3.693
- Gitaka, J. N., Takeda, M., Kimura, M., Idris, Z. M., Chan, C. W., Kongere, J., et al. (2017). Selections, frameshift mutations, and copy number variation detected on the surf 4.1 gene in the western Kenyan *Plasmodium falciparum* population. *Malar. J.* 16, 98. doi: 10.1186/s12936-017-1743-x
- He, Z., Zhang, Q., Wang, D., Hu, Y., Zhou, R., Qian, D., et al. (2022). Genetic polymorphism of circumsporozoite protein of *Plasmodium falciparum* among Chinese migrant workers returning from Africa to Henan Province. *Malar. J.* 21, 248. doi: 10.1186/s12936-022-04275-7
- Hollingdale, M. R., Nardin, E. H., Tharavanij, S., Schwartz, A. L., and Nussenzweig, R. S. (1984). Inhibition of entry of *Plasmodium falciparum* and *P. vivax* sporozoites into cultured cells; an *in vitro* assay of protective antibodies. *J. Immunol.* 132, 909–913. doi: 10.4049/jimmunol.132.2.909
- Huang, H. Y., Liang, X. Y., Lin, L. Y., Chen, J. T., Ehapo, C. S., Eyi, U. M., et al. (2020). Genetic polymorphism of *Plasmodium falciparum* circumsporozoite protein on Bioko Island, Equatorial Guinea and global comparative analysis. *Malar. J.* 19, 1–13. doi: 10.1186/s12936-020-03315-4
- Idris, Z. M., Chan, C. W., Kongere, J., Gitaka, J., Logedi, J., Omar, A., et al. (2016). High and heterogeneous prevalence of asymptomatic and sub-microscopic malaria infections on islands in lake victoria, Kenya. *Sci. Rep.* 6, 36958. doi: 10.1038/srep36958
- Kotraiah, V., Phares, T. W., Terry, F. E., Hindocha, P., Silk, S. E., Nielsen, C. M., et al. (2021). Identification and immune assessment of T cell epitopes in five *Plasmodium falciparum* blood stage antigens to facilitate vaccine candidate selection and optimization. *Front. Immunol.* 12. doi: 10.3389/fimmu.2021.690348
- Laurens, M. B. (2019). RTS,S/AS01 vaccine (Mosquirix™): an overview. *Hum. Vaccines Immunother.* 16, 480–489. doi: 10.1080/21645515.2019.1669415
- Lê, H. G., Kang, J. M., Moe, M., Jun, H., Thái, T. L., Lee, J., et al. (2018). Genetic polymorphism and natural selection of circumsporozoite surface protein in *Plasmodium falciparum* field isolates from Myanmar. *Malar. J.* 17, 361. doi: 10.1186/s12936-018-2513-0
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Maier, A. G., Matuschewski, K., Zhang, M., and Rug, M. (2019). *Plasmodium falciparum*. *Trends Parasitol.* 35, 481–482. doi: 10.1016/j.pt.2018.11.010
- Merle, C. S., Badiane, N. A., Affoukou, C. D., Affo, S. Y., Djogbenou, S. L., Hounto, A., et al. (2023). Implementation strategies for the introduction of the RTS,S/AS01 (RTS,S) malaria vaccine in countries with areas of highly seasonal transmission: workshop meeting report. *Malar. J.* 22, 1–14. doi: 10.1186/s12936-023-04657-5
- Mohamed, N. S., Abdelbagi, H., Elsadig, A. R., Ahmed, A. E., Mohammed, Y. O., Elssir, L. T., et al. (2021). Assessment of genetic diversity of *Plasmodium falciparum* circumsporozoite protein in Sudan: the RTS,S leading malaria vaccine candidate. *Malar. J.* 20, 1–12. doi: 10.1186/s12936-021-03971-0
- Mohamed, N. S., Ali Albsheer, M. M., Abdelbagi, H., Siddiq, E. E., Mohamed, M. A., Ahmed, A. E., et al. (2019). Genetic polymorphism of the N-terminal region in circumsporozoite surface protein of *Plasmodium falciparum* field isolates from Sudan. *Malar. J.* 18, 1–10. doi: 10.1186/s12936-019-2970-0
- Murmu, L. K., Sahu, A. A., and Barik, T. K. (2021). Diagnosing the drug resistance signature in *Plasmodium falciparum*: a review from contemporary methods to novel approaches. *J. Parasitol. Dis.* 45, 869–876. doi: 10.1007/s12639-020-01333-2
- Neafsey, D. E., Juraska, M., Bedford, T., Benkeser, D., Valim, C., Griggs, A., et al. (2015). Genetic diversity and protective efficacy of the RTS,S/AS01 malaria vaccine. *N. Engl. J. Med.* 373, 2025–2037. doi: 10.1056/NEJMoa1505819
- Osborne, A., Manko, E., Takeda, M., Kaneko, A., Kagaya, W., Chan, C., et al. (2021). Characterizing the genomic variation and population dynamics of *Plasmodium falciparum* malaria parasites in and around Lake Victoria, Kenya. *Sci. Rep.* 11, 1–11. doi: 10.1038/s41598-021-99192-1
- Pringle, J. C., Carpi, G., Almagro-Garcia, J., Zhu, S. J., Kobayashi, T., Mulenga, M., et al. (2018). RTS,S/AS01 malaria vaccine mismatch observed among *Plasmodium falciparum* isolates from southern and central Africa and globally. *Sci. Rep.* 8, 1–8. doi: 10.1038/s41598-018-24585-8
- Putaporntip, C., Jongwtiwees, S., and Hughes, A. L. (2009). Natural selection maintains a stable polymorphism at the circumsporozoite protein locus of *Plasmodium falciparum* in a low endemic area. *Infect. Genet. Evol.* 9, 567–573. doi: 10.1016/j.meegid.2009.02.006

- Reynisson, B., Alvarez, B., Paul, S., Peters, B., and Nielsen, M. (2021). NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* 48, W449–W454. doi: 10.1093/NAR/GKAA379
- Reynisson, B., Barra, C., Kaabinejadian, S., Hildebrand, W. H., Peters, B., Peters, B., et al. (2020). Improved prediction of MHC II antigen presentation through integration and motif deconvolution of mass spectrometry MHC eluted ligand data. *J. Proteome Res.* 19, 2304–2315. doi: 10.1021/acs.jproteome.9b00874
- Rosenthal, P. J., John, C. C., and Regina Rabinovich, N. (2019). Perspective piece malaria: How are we doing and how can we do better? *Am. J. Trop. Med. Hyg.* 100, 239–241. doi: 10.4269/ajtmh.18-0997
- Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248
- RTS, S. C. T. P. (2015). Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: Final results of a phase 3, individually randomised, controlled trial. *Lancet* 386, 31–45. doi: 10.1016/S0140-6736(15)60721-8
- Samuels, A. M., Ansong, D., Kariuki, S. K., Adjei, S., Bollaerts, A., Ockenhouse, C., et al. (2022). Efficacy of RTS,S/AS01E malaria vaccine administered according to different full, fractional, and delayed third or early fourth dose regimens in children aged 5–17 months in Ghana and Kenya: an open-label, phase 2b, randomised controlled trial. *Lancet Infect. Dis.* 22, 1329–1342. doi: 10.1016/S1473-3099(22)00273-0
- Sidjanski, S. P., Vanderberg, J. P., and Sinnis, P. (1997). Anopheles stephensi salivary glands bear receptors for region I of the circumsporozoite protein of Plasmodium falciparum. *Mol. Biochem. Parasitol.* 90, 33–41. doi: 10.1016/S0166-6851(97)00124-2
- Sinnis, P., and Fidock, D. A. (2022). The RTS,S vaccine—a chance to regain the upper hand against malaria? *Cell* 185, 750–754. doi: 10.1016/j.cell.2022.01.028
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi: 10.1093/genetics/123.3.585
- Tucker, K. D., Schanen, B. C., Phares, T. W., Sassano, E., Terry, F. E., Hindocha, P., et al. (2021). Identification, selection and immune assessment of liver stage CD8 T cell epitopes from plasmodium falciparum. *Front. Immunol.* 12. doi: 10.3389/fimmu.2021.684116
- World Health Organization. (2021). World malaria report 2021. Available at: <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021>.
- Zavala, F. (2022). RTS,S: the first malaria vaccine. *J. Clin. Invest.* 132, e156588. doi: 10.1172/JCI156588
- Zhang, R., Jia, G., and Diao, X. (2023). geneHapR: an R package for gene haplotypic statistics and visualization. *BMC Bioinf.* 24, 199. doi: 10.1186/s12859-023-05318-9
- Zheng, J., Pan, H., Gu, Y., Zuo, X., Ran, N., Yuan, Y., et al. (2019). Prospects for malaria vaccines: pre-erythrocytic stages, blood stages, and transmission-blocking stages. *BioMed. Res. Int.* 2019, 9751471. doi: 10.1155/2019/9751471