


OPEN ACCESS


EDITED BY

 Stefano D'Amelio,
Sapienza University of Rome, Italy

REVIEWED BY

 Mahamadou Ali Thera,
Université des Sciences, des Techniques
et des Technologies de Bamako, Mali
Joana Carneiro da Silva,
University of Maryland, United States

*CORRESPONDENCE

 Jonathan J. Juliano
 jjuliano@med.unc.edu

RECEIVED 18 November 2025

REVISED 15 March 2026

ACCEPTED 16 March 2026

PUBLISHED 07 April 2026

CITATION

 Ali IM, Vasquez Martinez B, Tchukenkam VPK, Sadler JM, Gorman CC, Nsango SE, Junior VV, Ateba JM, Wanda GY, Bailey JA, Dinglasan RR, Lin JT and Juliano JJ (2026) Haplotype diversity of *Plasmodium falciparum* circumsporozoite protein prior to malaria vaccine implementation in Kaelé Health District, Cameroon, 2022–2023. *Front. Parasitol.* 5:1749169. doi: 10.3389/fpara.2026.1749169

COPYRIGHT

 © 2026 Ali, Vasquez Martinez, Tchukenkam, Sadler, Gorman, Nsango, Junior, Ateba, Wanda, Bailey, Dinglasan, Lin and Juliano. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Haplotype diversity of *Plasmodium falciparum* circumsporozoite protein prior to malaria vaccine implementation in Kaelé Health District, Cameroon, 2022–2023

 Innocent M. Ali¹, Brenda Vasquez Martinez², Valery P. K. Tchukenkam¹, Jacob M. Sadler², Catherine C. Gorman², Sandrine E. Nsango³, Voundi Voundi Junior⁴, Joel Marcellin Ateba⁴, Grace Yimga Wanda¹, Jeffrey A. Bailey⁵, Rhoel R. Dinglasan⁶, Jessica T. Lin^{1,7} and Jonathan J. Juliano^{1,7,8*}
¹Department of Biochemistry, Faculty of Science, Université de Dschang, Dschang, Cameroon,

²Institute for Global Health and Infectious Diseases, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, ³Centre Pasteur du Cameroun, Yaoundé, Cameroon, ⁴National Malaria Control Program, Ministry of Public Health, Yaounde, Cameroon, ⁵Department of Pathology and Laboratory Medicine, Brown University, Providence, RI, United States, ⁶Department of Infectious Diseases and Immunology and the Emerging Pathogens Institute, College of Veterinary Medicine, University of Florida, Gainesville, FL, United States, ⁷Division of Infectious Diseases, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, ⁸Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Malaria remains a major public health concern in Cameroon, with *Plasmodium falciparum* responsible for most morbidity and mortality, particularly among children under five. In response to rising cases, Cameroon began implementing the RTS,S/AS01 malaria vaccine in early 2024. Given the vaccine's strain-specific efficacy, understanding haplotypic diversity is critical for evaluating long-term impact. We analyzed 100 *P. falciparum*-positive dried blood spots collected in Mapoussere, Kaele Health District (2022–2023). Using the 4CAST amplicon sequencing assay, we targeted four genes: *csp* (circumsporozoite protein), *ama1* (apical membrane antigen 1), *sera2* (serine repeat antigen 2), and *trap* (thrombospondin-related anonymous protein). Haplotypes were identified using SeekDeep, and diversity metrics including heterozygosity (H_e), nucleotide diversity (π), and selection statistics (Tajima's D , Fu and Li's D^* , F^*) were computed. We successfully genotyped *csp* in 35% of samples, identifying 22 haplotypes ($H_e = 0.908$; $\pi = 0.021$). The vaccine-matched haplotype was present in 20% of genotyped infections. The T cell TH2 and TH3 epitopes of *csp* showed signs consistent with balancing selection. Both *ama1* and *sera2* exhibited higher genotyping success and diversity, with *ama* showing significant Tajima's D

values. Within-participant haplotype counts were highest for *ama1* (mean = 2.8), followed by *sera2* (2.1), *csp* (1.3), and *trap* (1.2). This study provides a baseline of *P. falciparum* haplotype diversity in a vaccine-targeted region. The presence of vaccine-matched strains and high diversity in TH2/TH3 epitopes in *csp* may influence vaccine efficacy. Continued molecular surveillance is essential to monitor antigenic shifts and guide future strategies.

KEYWORDS

Cameroon, circumsporozoite (CS) protein, *Plasmodium falciparum*, malaria, RTS,S

Introduction

Malaria remains an important public health problem in Cameroon, with the entire country being endemic for malaria transmission, and ranking 11th on the list of countries with the most malaria cases (National Malaria Control Program, n.d). In 2023, there were over 7.3 million cases of malaria in Cameroon, with an estimated 11,300 deaths (World Health Organization, 2024). The majority of morbidity and mortality occurs in children under 5 years of age and is due to *Plasmodium falciparum*. Cameroon implements major interventions for controlling malaria, including treatment with effective artemisinin combination therapies (ACT), the use of long-lasting insecticide-treated bednets (LLIN), and indoor residual spraying (IRS). To date, validated and candidate mutations associated with artemisinin partial resistance have not been documented. Despite effective therapy being available, the number of cases increased by nearly 1.2 million per year between 2019 and 2023 (World Health Organization, 2024). In response, Cameroon was one of the first countries to implement RTS,S/AS01(MosquirixTM), beginning in January 2024 across 42 priority health districts, with the goal of countrywide coverage by the end of 2026 (Ndoula et al., 2024). Cameroon's Expanded Program on Immunization (EPI) reported that districts included in the initial implementation of the RTS,S/AS01 malaria vaccine experienced a 17% greater reduction in all-cause hospitalization and outpatient clinic visits for children under five, compared to districts without the vaccine (ReliefWeb, n.d). Additionally, around 60% of vaccinated districts saw a decline in overall under-five mortality, while 57% recorded fewer malaria-related deaths among children in this age group (ReliefWeb, n.d).

RTS,S/AS01 has been associated with a 55.1% (95% CI 50.5 to 59.3) reduction in clinical malaria in children aged 5–17 months during the 12 months of follow-up post-immunization, and 28.3% (95% CI 23.3 to 32.9) over 4 years. However, there is evidence that efficacy is strain-specific, with lower efficacy against heterologous strains, defined by the C-terminus of the *P. falciparum* circumsporozoite protein (*csp*), which contains two critical T cell epitopes, TH2 and TH3 (Agnandji et al., 2011; Neafsey et al., 2015; RTS,S Clinical Trials Partnership, 2015). Sequencing of *P. falciparum* circumsporozoite protein (*csp*), the gene encoding the antigen in the vaccine, shows differences in parasite populations by continent (Barry et al., 2009), but relatively few differences between populations within a continent (Aragam et al., 2013). This is potentially due to convergent evolution of the antigen due to interaction with the human immune system, leading the gene to be under slight balancing selection, or due to migration. Tracking of

csp variants in the population can be useful in understanding whether the selection pressure of the vaccine on the parasite population results in vaccine escape, whereby non-vaccine-matched genetic variants have a selective advantage and increase in frequency over time. To date, there is little information on *csp* diversity in Cameroon. One study reported 25 distinct haplotypes in the C-terminus of *csp* among 57 successfully sequenced samples from different areas of Cameroon collected in 2019 (Kojom Foko et al., 2024). A second study did not report haplotypes but found 51 single-nucleotide polymorphisms in *csp* from a subset of 117 parasitemic samples collected in 2024, primarily in the TH2 and TH3 epitopes (Efeti et al., 2025).

P. falciparum infections in Africa commonly contain more than one strain of malaria, and thus infections may contain strains representing more than one *csp* variant based on the TH2 and TH3 variation. The number of strains found within a single individual is loosely correlated to transmission intensity (Watson et al., 2021) and therefore varies by geographic region. In Cameroon, within-infection diversity has been evaluated by amplicon sequencing (Sadler et al., 2024), microsatellite analysis (Atuh et al., 2022; Biabi et al., 2023), and fragment length polymorphism analysis (Wanji et al., 2012; Biabi et al., 2023). Multi-strain infections were common across these studies.

As part of efforts to support the implementation and evaluation of RTS,S vaccine effectiveness in Cameroon, we assessed the diversity of the C-terminal TH2 and TH3 epitopes and flanking regions, as well as the within-individual haplotype diversity for 100 *P. falciparum* positive samples collected in the Kaelé health district in the Far North region of Cameroon during December 2022. This area is one of the original 42 districts where RTS,S/AS01 was implemented, providing an opportunity to assess baseline malaria diversity for the district. We leveraged a published multiplex amplicon sequencing assay, 4CAST, which targets variable regions of *P. falciparum csp*, apical membrane antigen 1 (*ama1*), serine repeat antigen 2 (*sera2*), and thrombospondin-related anonymous protein (*trap*) (LaVerriere et al., 2022).

Methods

Ethics statement

Ethical clearance for this study was obtained from the National Ethics Committee for Human Health Research under approval No

2021/12/1428/CE/CNERSH/SP on the 22nd of December, 2021. De-identified dried blood spots (DBS) were brought to the University of North Carolina at Chapel Hill for genotyping and are considered non-human subjects research (24-0777). Approved consent forms were used to obtain assent and/or parental authorization before initiating any data or blood collection.

Participant samples

Sample and data collection were conducted in participants' households, targeting asymptomatic children aged 3 months to 14 years as part of a malaria survey conducted in Kaelé health district in the Far North region of Cameroon during December 2022. Details of this study are provided in the [Supplementary Material](#). Trained research technicians administered consent forms in English, French, or Ffulde. Following written consent or parental authorization, fingerstick blood samples were obtained, and hemostasis was ensured using sterile bandages. Approximately 250 μ L of blood was collected via capillary tube; 5–6 μ L was immediately used for a malaria rapid diagnostic test (mRDT), and upon a positive result, two spots of roughly 80–100 μ L each were applied onto Whatman No. 3 filter paper. mRDT positive samples were confirmed with microscopy, and parasitemia was determined using the WHO standard protocol. The filter papers were dried overnight in a shaded, fly-free environment, carefully folded, stored in ziplock bags with silica gel desiccants, and transported to the laboratory. Samples were maintained at room temperature until DNA extraction was performed.

Targeted amplicon sequencing

Each DBS was punched into 96-well plates, and DNA was extracted from 3 X 6 mm punches using a Chelex-Tween extraction method ([dx.doi.org/10.17504/protocols.io.kxygx414kl8j/v1](https://doi.org/10.17504/protocols.io.kxygx414kl8j/v1) and [dx.doi.org/10.17504/protocols.io.dm6gpm5mjgzp/v1](https://doi.org/10.17504/protocols.io.dm6gpm5mjgzp/v1)). Multiplex amplicon sequencing of the *csp*, *ama1*, *sera2*, and *trap* genes was conducted using the 4CAST assay ([LaVerriere et al., 2022](#)). Details of the application of this method are provided at [protocols.io](https://doi.org/10.17504/protocols.io.eq2ly4qdm1x9/v1) ([dx.doi.org/10.17504/protocols.io.eq2ly4qdm1x9/v1](https://doi.org/10.17504/protocols.io.eq2ly4qdm1x9/v1)). Amplicons were pooled and sequenced on a 2X300bp MiSeq i100 25M run at the UNC High Throughput Sequencing Facility.

Data analysis

Demultiplexed amplicons were analyzed using SeekDeep ([Hathaway et al., 2018](#)) with a minimum within-sample haplotype cluster size of 500 reads and a minimum within-sample allele frequency cut-off of 5%, applying whichever was most restrictive per sample. Cutoffs were determined using sequence data from controls, including mocked DBS of *P. falciparum* 3D7 parasites (MRA-102) at different parasitemias (described elsewhere: [Gaither et al., 2025](#)), and 2 mocked mixtures of genomic DNA (MRA-102G for *P. falciparum* strain 3D7 and MRA-152G for strain 7G8) (Mix 1: 3D7 80% and 7G8 20%; Mix 2: 7G8 95% and 3D7 5%).

Statistical analysis and graphing were completed with Graphpad Prism (v10.5.0) (Graphpad Software, Boston, MA). Heterozygosity (He) for each amplicon or haplotype, which is the probability that two haplotypes will be different, was determined as

$1 - (p_1^2 + p_2^2 + \dots + p_n^2)$, where p_1 , p_2 , etc. are the population frequencies of each haplotype and n is the total number of haplotypes in the population. Sequence logos were made using <https://weblogo.berkeley.edu/logo.cgi>. Maximum Likelihood phylogeny was generated using MEGA (v11).

Sequence analysis was carried out using DnaSP v.6 ([Rozas et al., 2017](#)). Nucleotide diversity (π) was used to quantify genetic variation within populations, where π represents the average number of nucleotide differences per site between all possible pairs of DNA sequences in a sample. To calculate π , aligned sequences were analyzed by computing the number of nucleotide differences between each pair, dividing by the sequence length to obtain per-site differences, and averaging across all unique sequence pairs. The formula is:

$$\pi = (1 / [n(n-1)/2]) \times \sum_{\{i < j\}} (\text{differences}_{\{ij\}} / \text{sequence length})$$

where n is the number of sequences. To detect deviations from neutral evolution, three standardized statistics were computed: Tajima's D , Fu and Li's D^* , and Fu and Li's F^* . These metrics compare different estimates of genetic diversity to infer selection or demographic history:

1. Tajima's D compares π to Watterson's θ (θ_t), which is based on the number of segregating sites. Under neutrality and constant population size, $\pi \approx \theta_t$ and $D \approx 0$. Deviations were interpreted as follows:

- $D < -2$: Suggests an excess of rare alleles, consistent with purifying selection or population expansion.
- $D > +2$: Indicates an excess of intermediate-frequency alleles, consistent with balancing selection or population bottlenecks.

2. Fu and Li's D^* compares the number of segregating sites (S) to the number of singletons (η_1), using the simplified formula:

$$D^* = (S - a \times \eta_1) / \sqrt{\text{Var}}$$

3. Fu and Li's F^* compares π to η_1 :

$$F^* = (\pi - b \times \eta_1) / \sqrt{\text{Var}}$$

where a and b are constants based on the sample size. For both D^* and F^* , values < -2 indicate an excess of singletons (suggesting recent expansion or purifying selection), while values $> +2$ suggest a deficit of singletons (consistent with balancing selection or population structure). These statistics were applied to each locus as a complete haplotype, as well as using a sliding window (window: 25bp, slide: 5bp).

Results

Participant characteristics

Participants were recruited directly from their homes, with a maximum of two children enrolled per household. This study included 100 samples that tested highly positive for *P. falciparum* by microscopy, with a geometric mean parasite density of 5744.12

parasites/ μL (95%CI: 3965.51-7522.73). Participants ranged in age from 3 months to 14 years. The most represented age group was 3 to 59 months, accounting for 60% of the samples. Children aged 6 to 14 years made up the remaining 40%. Males comprised 55% of the participants, and females 45%.

Sequencing quality and success

Results of sequencing controls are shown in [Supplementary Table 1](#). Overall, *ama1* and *sera2* performed the most robustly across samples, providing similar within-sample allele frequency estimates in the mixtures and most samples with recovered haplotypes. The *csp* amplicon worked well at higher parasitemias but tended to overestimate the frequency of the major variant in mixtures. The *trap* amplicon yielded poor results.

Among clinical samples, a total of 4,324,169 reads were extracted across the 4 amplicons, with 2,893,882 reads (66.9%) being used in the final haplotype determination. This was true on a gene-level basis for three genes, *csp*, *ama1*, and *sera2*. However, for *trap*, the majority of reads failed to produce haplotypes, primarily due to cluster sizes below our cutoff of 500 reads. Details of the sequencing yield for each target are shown in [Table 1](#). Individual-level genotyping data are shown in [Supplementary Table 2](#). Unique haplotypes for each gene and the number of samples with those haplotypes are shown in [Supplementary Tables 3–6](#).

Gene nucleotide diversity

For *csp*, we successfully genotyped 35/100 (35%) samples ([Table 1](#)). Twenty-two haplotypes were detected with a nucleotide heterozygosity of 0.908 in the population ([Supplementary Table 7](#)). The nucleotide diversity (π) was 0.021. The vaccine haplotype of *csp*, which is identical to the 3D7 strain, is represented by PfCSP.01 in our population. This haplotype was found in 7/35 (20%) of genotyped samples and represented 25,173/140,269 (17.9%) of the sequencing reads used to construct haplotypes.

Genetic diversity was higher, and more samples were successfully genotyped for *ama1* and *sera2* ([Table 1](#)). Thirty-five and 29 haplotypes were detected for *ama1* and *sera2*, respectively. Both had higher heterozygosity than *csp*, but π was higher for *ama1* and lower for *sera2*. The *trap* locus was the least diverse, with a π of 0.013, 10 haplotypes detected, and a heterozygosity of 0.876. However, it was also the locus with the fewest genotyped samples, with 11 samples genotyped.

Protein circumsporozoite protein (CS) diversity

Each of the haplotypes encoded a unique protein sequence ([Supplementary Figure 1](#)). Given that each nucleotide haplotype encodes a unique protein haplotype, the protein heterozygosity for CS is the same as for the nucleotide haplotype (0.908).

Within-sample haplotype diversity

The within-sample nucleotide haplotype diversity was assessed separately for each gene based on the number of haplotypes

TABLE 1 Sequencing success and characteristics of targeted genes.

Gene	Total extracted reads	# Read used for final haplotypes (% of extracted)	Range of reads per successful sample	# of successful genotyped samples	# Haplotypes	He	Tajima's D	Fu and Li D	Fu and Li F	π
<i>csp</i>	192,801	140,269 (72.8%)	516–27,370	35	22	0.908	0.22 (p>0.10)	0.24	0.23	0.021
<i>ama1</i>	2,836,417	1,837,489 (64.8%)	554–49,957	83	35	0.952	1.94 (0.1>p>0.05)	1.74	2.51	0.041
<i>sera2</i>	1,252,534	895,025 (71.5%)	513–49,419	75	29	0.919	-0.449 (p>0.1)	-0.85	-0.32	0.010
<i>trap</i>	42,417	21,099 (49.7%)	540–7,626	11	10	0.876	0.041 (p>0.1)	-0.15	-0.10	0.013

detected. Among the four genes analyzed, *ama1* exhibited the widest range and highest number of haplotypes across samples (Figure 1). The proportion of polyclonal infections was 28.6% (10/35) for *csp*, 68.7% (57/83) for *ama1*, 65.3% (49/75) for *sera2*, and 18.2% (2/11) for *trap*. Mean within-sample haplotype counts (range) were: *csp* – 1.3 (1-2), *ama1* – 2.8 (1-9), *sera2* – 2.1 (1-6), and *trap* – 1.2 (1-2). Median within-sample haplotype counts (95% CI of the median) were: *csp* – 1 (1–1), *ama1* – 2 (2–3), *sera2* – 2 (2–2), and *trap* – 1 (1–2). Per participant haplotype counts for the antigens are shown in Supplementary Figure 2.

Selection on *csp*

None of the 4 genes showed significant Tajima's *D*, Fu and Li *D*^{*}, or Fu and Li *F*^{*} values for the complete gene (Table 1). The TH2 epitope (amino acids 311–327) and the TH3 epitope (amino acids 352–363) of *csp* showed elevated Tajima's *D* values, but not reaching a value of 2, on sliding window analysis (Figure 2A). The translated protein shows high levels of diversity in these regions (Figure 2D; Supplementary Figure 3). Sliding window Tajima's *D* analysis shows significantly elevated Tajima's *D* at multiple windows in *ama1*, but other genes lack significant windows in either direction (Supplementary Figure 4).

Diversity of *csp* in Cameroon

In order to assess the relationship of *csp* antigen diversity in Kaele to previously published data in Cameroon, we downloaded a subset of sequences from 22 isolates from Pette, also in Northern Cameroon, from a previous report (Supplementary Table 8) (Kojom Foko et al., 2024). Data from Efeti et al. was not available for download (Efeti et al., 2025). Using these data, we generated a median-joining phylogenetic tree of sequences trimmed to the same length (Figure 3). The isolates appear to be from a single population,

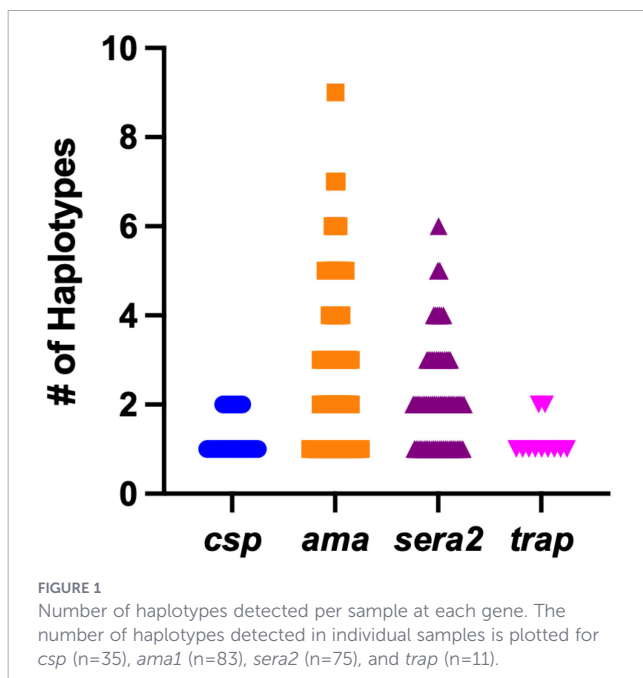
with isolates from both studies intermixed. Identical samples with no branch length differences were found between the studies.

Discussion

P. falciparum remains a major public health challenge in Cameroon. Despite the availability of effective interventions, the number of malaria cases has been increasing over the last 5 years (World Health Organization, 2024). Implementation of the malaria vaccine RTS,S/AS01 has shown promising reductions in the districts where it has been implemented so far, with plans to expand to the entire country (ReliefWeb, n.d). Here we describe the diversity of *csp*, the antigen in RTS,S/AS01, as well as three other malaria antigens (*ama1*, *sera2*, and *trap*), in a northern district of Cameroon prior to vaccine implementation. Given that there is evidence of strain-specific vaccine efficacy, understanding the longitudinal dynamics of the *csp* antigen in the population may be helpful in understanding long-term vaccine efficacy (Neafsey et al., 2015). We found that the vaccine-matched *csp* haplotype was found in 20% of genotyped infections. The overall diversity of the antigen was similar to previous reports of *csp* diversity in Cameroon (Kojom Foko et al., 2024). Haplotypes detected in this study were intermixed with haplotypes from Pette in Northern Cameroon, with identical haplotypes detected (Figure 3). The TH2 and TH3 epitopes have high levels of diversity and are consistent with balancing selection, similar to other studies (Figure 2) (Aragam et al., 2013).

Consistent with previous studies in Cameroon, we found no strong evidence of evolutionary pressure acting on the C-terminal of the *csp* gene, as Tajima's *D*, Fu & Li's *F*^{*}, and Fu & Li's *D* values were not statistically significant (Kojom Foko et al., 2024). At the level of the complete gene, our results aligned with earlier reports from other countries in Africa as well (Bailey et al., 2012; Aragam et al., 2013). However, similar to those studies, we observed peaks in Fu & Li's *D*^{*}, Fu & Li's *F*^{*}, and Tajima's *D* around the TH2 and TH3 regions, indicating an excess of intermediate-frequency alleles (Figure 2) (Weedall et al., 2007; Bailey et al., 2012; Aragam et al., 2013). These patterns are similar to previously reported work in other regions of Africa (Bailey et al., 2012; Aragam et al., 2013). This suggests that genetic diversity at these loci is maintained within the population, potentially benefiting the parasite and reflecting weak selective pressure; likely driven by host immunity, as supported by structural modeling (Bailey et al., 2012; Aragam et al., 2013). Notably, we detected a high proportion (20%) of vaccine-matched infections, significantly exceeding previous reports from Cameroon (2%) and other African sites (Aragam et al., 2013; Neafsey et al., 2015; Kojom Foko et al., 2024). This discrepancy may be partly due to differences in sequencing methods between the previous study (sanger sequencing) in Cameroon and the present one. Importantly, our sequenced fragment includes the full TH2 and TH3 epitopes, which are known to be under allele-specific selection.

The *csp* C-terminal sequences from our study in Northern Cameroon show strong phylogenetic continuity with those previously reported from the same region (Pette), with isolates



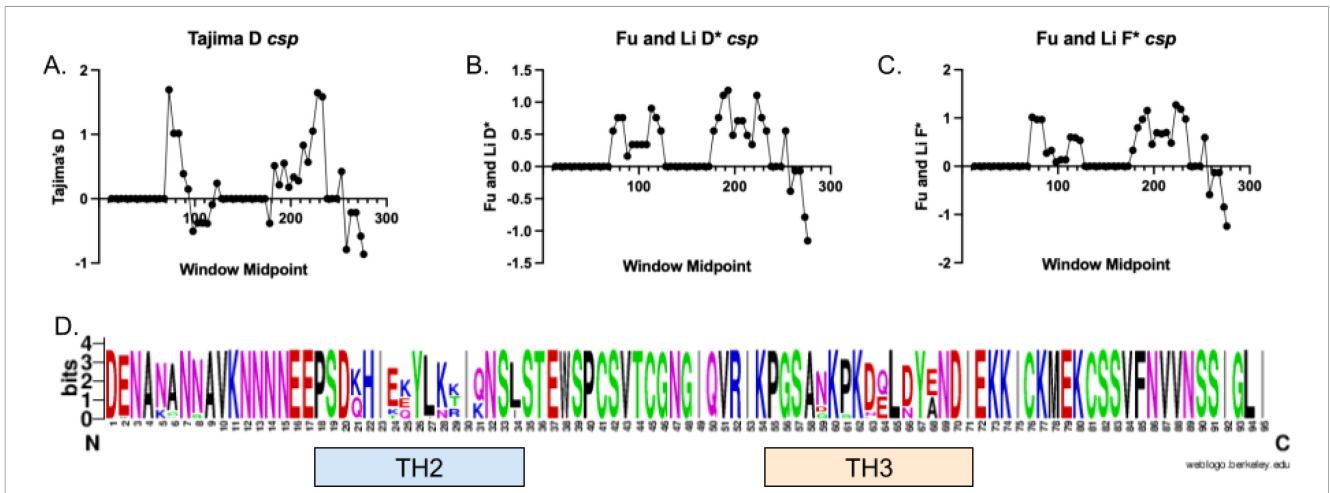


FIGURE 2 TH2 and TH3 diversity and selection in *csp*. (A–C) show scans for selection across the amplicon using a 25bp widow with a slide of 5bp. No values for Tajima's D (A), Fu and Li D* (B), or Fu and Li F* (C) reached statistical significance. (D) shows a web logo of amino acids for the region amplified with the TH2 (amino acids 311–327) region and TH3 region (amino acids 352–363) marked.

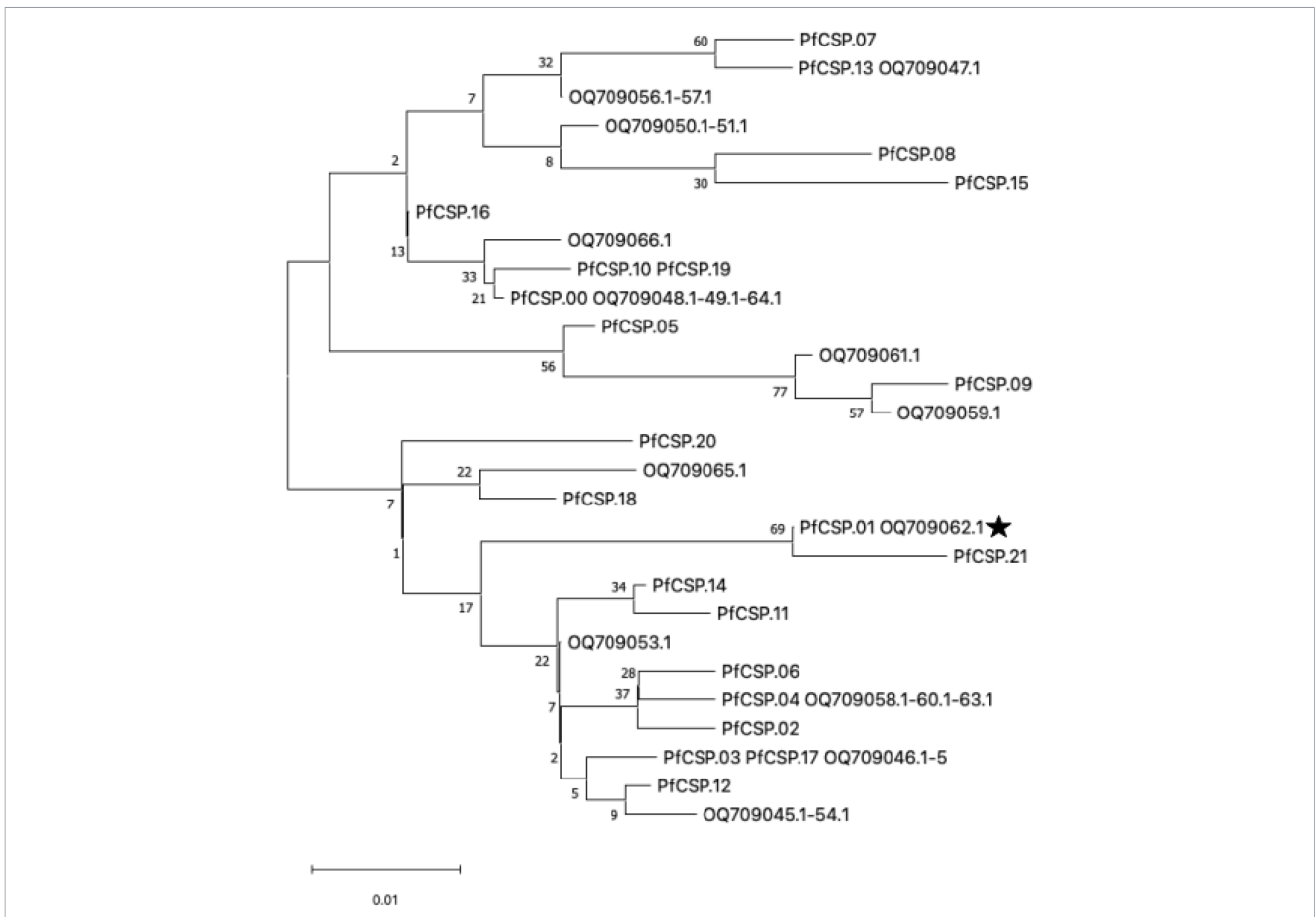


FIGURE 3 Maximum likelihood phylogenetic tree of *csp* nucleotide sequences. Twenty-two public sequences from Pette collected in 2019 (start with OQ) were compared to the unique 22 haplotypes detected in 2022–23 (start with PfCSP). The public data represents all sequences from the study site and thus has repeated haplotypes. Trimming from the full 4CAST sequencing length to the 194bp length of public data resulted in certain haplotypes becoming identical (e.g., PfCSP.03 and PfCSP.17). Branches with multiple samples with identical sequences have all sequence names listed (multiple OQ samples are listed by the -XX.X denotation in the phylogenetic branch). The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model. The tree with the highest log likelihood (-619.88) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Node values are determined by 1000 bootstrap values. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. There were a total of 194bp in the final dataset. The RTS,S vaccine haplotype is marked with a star.

from both datasets intermixed and some sequences being identical. This pattern indicates the presence of a single circulating *Plasmodium falciparum* population. Such genetic homogeneity contrasts with the higher diversity observed in southern Cameroonian sites such as Douala, where novel mutations like E357L and D359Y, which may be detrimental based on protein structure, were detected in previous studies (Kojom Foko et al., 2024). The absence of these mutations in our dataset, along with the low prevalence (<5%) of K322I, suggests limited selective pressure acting on these loci within the northern eco-epidemiological zone (Supplementary Figure 1). Importantly, the Cameroonian *csp* sequences have previously been shown to cluster phylogenetically with isolates from Sierra Leone, Ghana, Zambia, and Equatorial Guinea, highlighting regional genetic connectivity within sub-Saharan Africa (Kojom Foko et al., 2024). Given the close relatedness of our isolates to those from Pette, it is likely that this broader phylogeographic relationship remains consistent.

Other genes showed weak signals of selection, with *ama1* being the only locus showing statistically significant selection in sliding window analysis. The positive Tajima's *D* values for *ama1* are consistent with balancing selection, as reported in other studies (Osier et al., 2010; Arnott et al., 2014; Naung et al., 2022). This gene also showed the highest nucleotide diversity and second highest number of haplotypes (Table 1). This is similar to other reports of *ama1* diversity in Cameroon, which reported a similar *He* (0.976) and a lower nucleotide diversity (0.016) (Hawadak et al., 2024). Tajima's *D* in that study was 2.058. Whole genome sequencing data from Mt. Cameroon area showed similar levels consistent with balancing selection in *ama1* and *trap* (Tajima's *D* >1) as shown here (Apinjoh et al., 2023).

Within-host haplotype diversity estimates from the 4 genes varied, with *ama1* providing the highest mean estimate, the highest median estimate, and the greatest range of haplotypes detected. Polyclonal infections were commonly detected by *ama1* and *sera2*, but not *csp* and *trap*. Within-host diversity estimates were similar to those estimated by amplicon deep sequencing and molecular inversion probes in Dschang, Cameroon, which estimated COI (based on the loci with the highest number of haplotypes during genotyping) to be 2.7 (Range 1-5) and 1.2 (range 1-3), respectively (Sadler et al., 2024). In that study, however, *ama1* was not the most informative locus for within-sample diversity, providing the largest number of haplotypes in 43 of the 100 samples genotyped. Rather, an amplicon at a different hypervariable region (HeOME A) provided the highest within-sample haplotype diversity estimate for more samples (Sadler et al., 2024). The majority of other work in Cameroon regarding polyclonality has relied on the detection of size polymorphisms in merozoite surface protein genes or microsatellites (Atuh et al., 2022; Biabi et al., 2023; Bouopda-Tuedom et al., 2025). While within-host diversity is generally found to be lower with those methods than amplicon deep sequencing, polyclonal infections were commonly found.

Unfortunately, we were only able to generate *csp* genotypes from 35 of 100 samples sequenced, limiting the generalizability of the data and potentially the within-population frequency estimates. The high number of distinct haplotypes may not allow Tajima's *D* to clearly differentiate between an excess of rare alleles from an

increase in intermediate-frequency alleles. Diverse *ama1* and *sera2* genes were genotyped from more samples and likely represent the population diversity more accurately. The diversity of *trap* is likely under-estimated due to the low number of samples that were successfully sequenced. We used highly restrictive filtering for determining haplotypes with a high within-sample minimum frequency (>5%) and a high threshold for the number of reads needed (500 reads) to determine a haplotype. This likely impacted both the ability to detect minority variants in mixed infections, as well as the overall success of genotyping samples. Still, this report provides new, important information about parasite antigenic diversity in Cameroon, which can inform vaccine implementation evaluation programs.

In summary, this study provides a critical baseline assessment of *P. falciparum* haplotype diversity in Kaelé health district, Cameroon, prior to the implementation of the RTS,S/AS01 malaria vaccine. Our findings reveal substantial diversity in the *csp* gene, particularly within the TH2 and TH3 epitopes, consistent with balancing selection and previous reports from other regions of Africa. Interestingly, 20% of genotyped infections contain a strain that matched the vaccine strain. The overall diversity and phylogenetic relationships suggest a largely homogeneous parasite population structure in Far North Cameroon. The within-host diversity estimates underscore the prevalence of polyclonal infections, which may influence vaccine efficacy and parasite evolution. Limitations in genotyping success, particularly for *csp* and *trap*, highlight the need for optimized sequencing approaches to capture minority variants and improve resolution. Despite these constraints, our data contribute valuable insights into the genetic landscape of *P. falciparum* in a vaccine-targeted region and underscore the importance of continued molecular surveillance to monitor antigenic shifts and inform long-term vaccine strategies.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI SRA, BioProject PRJNA1445970.

Ethics statement

The studies involving humans were approved by National Ethics Committee for Human Health Research. 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.

Author contributions

IA: Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing, Formal analysis,

Resources. BV: Investigation, Writing – review & editing, Data curation, Methodology. VT: Data curation, Investigation, Writing – original draft, Writing – review & editing. JS: Data curation, Methodology, Writing – review & editing. CG: Data curation, Investigation, Writing – review & editing. SN: Funding acquisition, Resources, Supervision, Writing – review & editing. VJ: Investigation, Resources, Writing – review & editing. AM: Investigation, Resources, Writing – review & editing. GW: Investigation, Resources, Writing – review & editing. JB: Software, Writing – review & editing. RD: Funding acquisition, Resources, Supervision, Writing – review & editing. JL: Project administration, Supervision, Writing – review & editing. JJ: Conceptualization, Formal analysis, Funding acquisition, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was funded by the National Institute for Allergy and Infectious Diseases, National Institutes of Health (R01AI165537 to JJ, RD, and SN; 1U19AI181594 to RD and JJ; K24AI134990 to JJ) and EDCTP (Grant #CSA2018HS-2520) which supported field sample collection via NMCP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

The following reagent was obtained through BEI Resources, NIAID, NIH: *Plasmodium falciparum*, Strain 3D7, MRA-102, contributed by Daniel J. Carucci. The following reagent was obtained through BEI Resources, NIAID, NIH: Genomic DNA from *Plasmodium falciparum*, Strain 3D7, MRA-102G, contributed by Daniel J. Carucci. The following reagent was obtained through BEI Resources, NIAID, NIH: Genomic DNA from *Plasmodium falciparum*, Strain 7G8, MRA-152G, contributed by David Walliker.

References

- Agnandji, S. T., Lell, B., Soulanoudjingar, S. S., Fernandes, J. F., Abossolo, B. P., RTS,S Clinical Trials Partnership, et al. (2011). First results of phase 3 trial of RTS,S/AS01 malaria vaccine in African children. *N. Engl. J. Med.* 365, 1863–1875. doi: 10.1056/NEJMoa1102287
- Apinogh, T. O., Ajonina, M. U., Abera, D., Chi, H. F., Tata, R. B., Mugri, R. N., et al. (2023). Genomic analysis of *Plasmodium falciparum* isolates across different altitudinal zones along the slope of Mount Cameroon. *Front. Malar.* 1. doi: 10.3389/fmala.2023.1075755
- Aragam, N. R., Thayer, K. M., Nge, N., Hoffman, I., Martinson, F., Kamwendo, D., et al. (2013). Diversity of T cell epitopes in *Plasmodium falciparum* circumsporozoite protein likely due to protein-protein interactions. *PLoS One* 8, e62427. doi: 10.1371/journal.pone.0062427
- Arnott, A., Wapling, J., Mueller, I., Ramsland, P. A., Siba, P. M., Reeder, J. C., et al. (2014). Distinct patterns of diversity, population structure and evolution in the AMA1 genes of sympatric *Plasmodium falciparum* and *Plasmodium vivax* populations of Papua New Guinea from an area of similarly high transmission. *Malar. J.* 13, 233. doi: 10.1186/1475-2875-13-233
- Atuh, N. I., Anong, D. N., Jerome, F.-C., Oriero, E., Mohammed, N. I., D'Alessandro, U., et al. (2022). High genetic complexity but low relatedness in infections from Western Savannah Highlands and coastal equatorial Lowlands of Cameroon. *Pathog. Glob. Health* 116, 428–437. doi: 10.1080/20477724.2021.1953686
- 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
- Barry, A. E., Schultz, L., Buckee, C. O., and Reeder, J. C. (2009). Contrasting population structures of the genes encoding ten leading vaccine-candidate antigens of the human malaria parasite, *Plasmodium falciparum*. *PLoS One* 4, e8497. doi: 10.1371/journal.pone.0008497
- Biabi, M. F. A. B., Fogang, B., Essangui, E., Maloba, F., Donkeu, C., Keumoe, R., et al. (2023). High prevalence of polyclonal infections and association with poor IgG antibody responses in a hyper-endemic area in Cameroon. *Trop. Med. Infect. Dis.* 8 (8), 390. doi: 10.3390/tropicalmed8080390

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer MAT declared a past collaboration with the author IA to the handling editor.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. Generative AI was used in the development of this manuscript. The authors take full responsibility for the content.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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.2026.1749169/full#supplementary-material>

- Bouopda-Tuedom, A. G., Abate, L., Feufack-Donfack, B. L., Ngou, C. M., Bayibeki, A. N., Moukoko, C. E. E., et al. (2025). Genetic polymorphism of *msp2* in *Plasmodium falciparum* isolates among asymptomatic malaria infections from two ecological settings in Cameroon. *Malar. J.* 24, 1–14. doi: 10.1186/s12936-025-05414-6
- Efeti, M. T., Fankem, S. N., Diallo, M. T., Lahngong, M. S., Acha, N. L., Shey, R. A., et al. (2025). Targeted sequencing identifies SNPs associated with antimalarial drug resistance and the CSP vaccine antigen in *Plasmodium falciparum* from southwest Cameroon. *Int. J. Mol. Sci.* 26, 10764. doi: 10.3390/ijms262110764
- Gaither, C., Morgan, C., Kirby, R., Karema, C., Gashema, P., Zuromski, J., et al. (2025). Prevalence of asymptomatic non-falciparum and falciparum malaria in the 2014–15 Rwanda Demographic Health Survey. *PLoS One* 20, e0330480. doi: 10.1371/journal.pone.0330480
- Hathaway, N. J., Parobek, C. M., Juliano, J. J., and Bailey, J. A. (2018). SeekDeep: single-base resolution de novo clustering for amplicon deep sequencing. *Nucleic Acids Res.* 46, e21. doi: 10.1093/nar/gkx1201
- Hawadak, J., Kojom Foko, L. P., Dongang Nana, R. R., Yadav, K., Pande, V., Das, A., et al. (2024). Genetic diversity and natural selection of apical membrane antigen-1 (*ama-1*) in Cameroonian *Plasmodium falciparum* isolates. *Gene* 894, 147956. doi: 10.1016/j.gene.2023.147956
- Kojom Foko, L. P., Hawadak, J., Eboumbou Moukoko, C. E., Das, A., and Singh, V. (2024). Genetic analysis of the circumsporozoite gene in *Plasmodium falciparum* isolates from Cameroon: Implications for efficacy and deployment of RTS,S/AS01 vaccine. *Gene* 927, 148744. doi: 10.1016/j.gene.2024.148744
- LaVerriere, E., Schwabl, P., Carrasquilla, M., Taylor, A. R., Johnson, Z. M., Shieh, M., et al. (2022). Design and implementation of multiplexed amplicon sequencing panels to serve genomic epidemiology of infectious disease: A malaria case study. *Mol. Ecol. Resour.* 22, 2285–2303. doi: 10.1111/1755-0998.13622
- National Malaria Control Program (Cameroon: National Malaria Control Program). Available online at: <https://pnlp.cm> (Accessed November 8, 2025).
- Naung, M. T., Martin, E., Munro, J., Mehra, S., Guy, A. J., Laman, M., et al. (2022). Global diversity and balancing selection of 23 leading *Plasmodium falciparum* candidate vaccine antigens. *PLoS Comput. Biol.* 18, e1009801. doi: 10.1371/journal.pcbi.1009801
- Ndoula, S. T., Mboussou, F., Njoh, A. A., Nembot, R., Baonga, S. F., Njinkeu, A., et al. (2024). Malaria vaccine introduction in Cameroon: early results 30 days into rollout. *Vaccines (Basel)* 12 (4), 346. doi: 10.3390/vaccines12040346
- 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
- Osier, F. H. A., Weedall, G. D., Verra, F., Murungi, L., Tetteh, K. K. A., Bull, P., et al. (2010). Allelic diversity and naturally acquired allele-specific antibody responses to *Plasmodium falciparum* apical membrane antigen 1 in Kenya. *Infect. Immun.* 78, 4625–4633. doi: 10.1128/iai.00576-10
- ReliefWeb One-year anniversary malaria vaccine rollout underscores remarkable progress offers new data potential impact. Available online at: <https://reliefweb.int/report/cameroon/one-year-anniversary-malaria-vaccine-rollout-underscores-remarkable-progress-offers-new-data-potential-impact> (Accessed November 8, 2025).
- Rozas, J., Ferrer-Mata, A., Sánchez-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 Clinical Trials Partnership (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: 10.1056/nejmoa1208394
- Sadler, J. M., Simkin, A., Tchuengkam, V. P. K., Gerdes Gyuricza, I., Fola, A. A., Wamae, K., et al. (2024). Application of a new highly multiplexed amplicon sequencing tool to evaluate antimalarial resistance and relatedness in individual and pooled samples from Dschang, Cameroon. *Front. Parasitol.* 3. doi: 10.3389/fpara.2024.1509261
- Wanji, S., Kengne-Ouafu, A. J., Eyong, E. E. J., Kimbi, H. K., Tendongfor, N., Ndamukong-Nyanga, J. L., et al. (2012). Genetic diversity of *Plasmodium falciparum* merozoite surface protein-1 block 2 in sites of contrasting altitudes and malaria endemicities in the Mount Cameroon region. *Am. J. Trop. Med. Hyg.* 86, 764–774. doi: 10.4269/ajtmh.2012.11-0433
- Watson, O. J., Okell, L. C., Hellewell, J., Slater, H. C., Unwin, H. J. T., Omedo, I., et al. (2021). Evaluating the performance of malaria genetics for inferring changes in transmission intensity using transmission modeling. *Mol. Biol. Evol.* 38, 274–289. doi: 10.1093/molbev/msaa225
- Weedall, G. D., Preston, B. M. J., Thomas, A. W., Sutherland, C. J., and Conway, D. J. (2007). Differential evidence of natural selection on two leading sporozoite stage malaria vaccine candidate antigens. *Int. J. Parasitol.* 37, 77–85. doi: 10.1016/j.ijpara.2006.09.001
- World Health Organization (2024). *World malaria report 2024* (Geneva, Switzerland).