Signatures of selection and drivers for novel mutation on transmission-blocking vaccine candidate Pfs25 gene in western Kenya

Kevin O. Ochwedo1,2*, Shirley A. Onyango2,3, Collince J. Omondi1,2, Pauline W. Orondo2,3, Benny M. Ondeto1,2, Ming-Chieh Lee5, Harrysone E. Atieli2,6, Sidney O. Ogolla7, Andrew K. Githeko2,7, Antony C. A. Otieno1, Wolfgang R. Mukabana1,2, Guiyun Yan6, Daibin Zhong5*, James W. Kazura8

1 Department of Biology, Faculty of Science and Technology, University of Nairobi, Nairobi, Kenya, 2 Sub-Saharan Africa International Centre for Excellence in Malaria Research, Homa Bay, Kenya, 3 School of Zoological Sciences, Kenyatta University, Nairobi, Kenya, 4 Department of Biochemistry, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya, 5 Program in Public Health, College of Health Sciences, University of California, Irvine, Irvine, California, United States of America, 6 School of Public Health and Community Development, Maseno University, Kisumu, Kenya, 7 Centre for Global Health Research, Kenya Medical Research Institute, Kisumu, Kenya, 8 Centre for Global Health and Diseases, Case Western Reserve University, Cleveland, Ohio, United States of America

* kevinochwedo@gmail.com (KOO); dzhong@hs.uci.edu (DZ)

Abstract

Background

Leading transmission-blocking vaccine candidates such as Plasmodium falciparum surface protein 25 (Pfs25 gene) may undergo antigenic alterations which may render them ineffective or allele-specific. This study examines the level of genetic diversity, signature of selection and drivers of Pfs25 polymorphisms of parasites population in regions of western Kenya with varying malaria transmission intensities.

Methods

Dry blood spots (DBS) were collected in 2018 and 2019 from febrile outpatients with malaria at health facilities in malaria-endemic areas of Homa Bay, Kisumu (Chulaimbo) and the epidemic-prone highland area of Kisii. Parasites DNA were extracted from DBS using Chelex method. Species identification was performed using real-time PCR. The 460 base pairs (domains 1–4) of the Pfs25 were amplified and sequenced for a total of 180 P. falciparum-infected blood samples.

Results

Nine of ten polymorphic sites were identified for the first time. Overall, Pfs25 exhibited low nucleotide diversity (0.04×10⁻²) and low mutation frequencies (1.3% to 7.7%). Chulaimbo had the highest frequency (15.4%) of mutated sites followed by Kisii (6.7%) and Homa Bay (5.1%). Neutrality tests of Pfs25 variations showed significant negative values of Tajima’s D (-2.15, p<0.01) and Fu’s F (-10.91, p<0.001) statistics tests. Three loci pairs (123, 372), (364, 428) and (390, 394) were detected to be under linkage disequilibrium and none had...
history of recombination. These results suggested that purifying selection and inbreeding might be the drivers of the observed variation in Pfs25.

Conclusion
Given the low level of nucleotide diversity, it is unlikely that a Pfs25 antigen-based vaccine would be affected by antigenic variations. However, continued monitoring of Pfs25 immunogenic domain 3 for possible variants that might impact vaccine antibody binding is warranted.

Introduction
Over the past decade remarkable progress has been made in reducing the global malaria burden by coordinated public health interventions targeting both vectors and parasites [1]. However, progress in malaria control has been stalled partly due to the spread of insecticide and antimalarial drug resistance [1, 2]. One of the interventions needed to reduce the malaria burden is an effective P. falciparum vaccine [3–5]. Barriers to advances in vaccine development include allelic variation of target antigens due to mutations in their immunodominant protein domains [6–10]. Currently, there are attempts towards developing transmission-blocking vaccines (TBVs) that target P. falciparum sexual stage parasites [5, 11]. TBVs elicit host antibody responses that block the sporogonic cycle of the parasite in the mosquito vector, thereby thus reducing and, ultimately, stopping ongoing transmission in local communities.

Among the various TBVs under development a vaccine based on the Pfs25 protein sequence is perhaps the most advanced [12–15]. The cysteine-rich 25 kilo-Dalton molecule is expressed in the protease-rich mosquito vector midgut post-fertilization, where it facilitates ookinete epithelium penetration, aggregation, and maturation to oocysts [16, 17]. Protein expression of the Pfs25 protein is translationally repressed in the human host by the Pumilio/FBF family RNA-binding protein [18, 19]. The Pfs25 protein consists of four tandem epidermal growth factor (EGF) domains with 22 cysteine residues that is anchored on the parasite surface by a C-terminal glycosylphosphatidylinositol (GPI) [20]. Recombinant Pfs25 antigen in combination with other TBV antigens [12, 21] or alone have been shown to elicit anti-Pfs25 antibodies [13, 22–27] that block oocyst development in standard membrane feeding assays. Notably, translational repression of Pfs25 protein expression in the human host temper's enthusiasm for field deployment of a TBV based on this antigen alone due to the absence of natural boosting in malaria endemic communities [25, 27].

The Pfs25 exon is considered to be highly conserved despite the existence of various sites with mutations [28–30]. Documented mutations include locus 392 in P. falciparum isolates from China and India and locus 428 of isolates from Cambodia and India [29–31]. The aim of this study was to determine the distribution of polymorphic sites, level of genetic diversity and identify possible signatures of selection in the four domains of the Pfs25 protein from P. falciparum parasites circulating in malaria-endemic and epidemic-prone regions of western Kenya.

Materials and methods
Study site and sampling
A cross-sectional study was conducted in 2018 and 2019 in health clinics in malaria-endemic areas of Homa Bay and Kisumu (Chulaimbo) Counties and the malaria epidemic-prone
highlands of Kisii County (Fig 1). A total of 302 patients presenting with fever and symptoms of uncomplicated malaria, e.g., myalgia, fatigue, non-localizing symptoms, were recruited to donate a finger prick blood sample. 138 patients were from Homa Bay County, 62 from Kisumu County (Chulaimbo) and 102 from Kisii County. Health facilities selected for sample collection in Homa Bay lie at 34.64190˚E-0.38000˚S and 1134–1330 metres above the sea level (asl), in Kisumu (Chulaimbo Sub County hospital) at 00.03572˚S-034.62196˚E and 1328–1458 m asl and in Kisii (Eramba health dispensary) at 34˚480E, 00˚350S; 1540–1740 m asl (Fig 1).

Four samples of 25 μL of blood were collected on coded Whatman™ Blood Stain Cards (GE Healthcare WB100014) containing details such as the patient’s study code, date of collection, age, and gender as previously described [32]. Each of the collected coded Whatman™ Blood-Stained Card was singly placed in individual plastic bags containing silica gel before being enclosed together in a labelled envelope and transported to the International Centre of Excellence for Malaria Research (ICEMR) laboratory located at Tom Mboya University College in Homa Bay town for storage at -20˚C and later molecular processing.

DNA extraction

Genomic DNA was extracted from the 302 DBS samples using the Chelex resin (Chelex -100) method following the published method [33] with modifications. Briefly, a 3mm ring of blotted filter paper containing a blood sample was punched into a sterile coded 1.5ml Eppendorf tube using a sterile craft store puncher. 950 μl of 1×PBS and 50μl of 10% saponin was added and mixed well before incubation at 4˚C overnight. The mixture was centrifuged at 12,000 rpm for 10 min at room temperature before discarding the supernatant content. Any remnant of saponin was washed by adding 1ml PBS and spinning at 12,000 rpm for 5 min. The PBS was discarded and the tube content spun for 15sec before removal of the liquid component using the P200 pipette. The Eppendorf tubes containing filter paper were air-dried at room temperature for 15min. 250μl (20%) of Chelex suspension was added and the mixture was incubated in a water bath maintained at 85˚C for 10min. During incubation, the mixture containing DNA
was vortexed for 5 min to suspend the extracted DNA. Upon completion, the mixture was spun at 12,000 rpm for 1 min and the DNA was transferred into a coded clean 0.5 ml Eppendorf tube and stored at -20°C. Cultured *P. falciparum* laboratory strain NF54 genomic DNA was also extracted and used as a positive control for PCR and sequencing.

**Detection of *Plasmodium falciparum* infections**

Real-time (RT-PCR) with species-specific 18s rRNA and probes were used to identify *P. falciparum* infections as previously described [34]. Briefly, RT-PCR was set with a final volume of 12 μl containing 2 μl of sample DNA, 6 μl of PerfeCTa™ qPCR ToughMix™, Low ROX™ Master mix (2X), 0.5 μl of the species-specific probe, 0.4 μl of the species-specific forward primers (10 μM), 0.4 μl of the species-specific reverse primers (10 μM) and 0.1 μl of double-distilled water. The thermal profile used was set at 50°C for 2 min, followed by 45 cycles of (95°C for 2 min, 95°C for 3 sec and 58°C for 30 sec). Three positive samples from laboratory strain and three negative samples from blank filter paper were used as a positive and negative control for RT-PCR respectively. The RT-PCR amplification was performed on QuantStudio 3 Real-Time PCR System (ThermoFisher, Carlsbad, CA, USA). A total of 180 samples were confirmed positive for *P. falciparum* by RT-PCR. The level of parasitemia ranged from 120 to 146,880 parasites/μL blood with a median of 4,920 parasites/μL.

**Pfs25 gene amplification and sequencing**

A total of 180 *P. falciparum* infected samples (78 from Homa Bay, 40 from Chulaimbo, and 62 from Kisii) confirmed positive by RT-PCR were used to amplify the target Pfs25 gene. Amplification of the 460bp exon was performed through nested PCR in T100™ Thermal Cycler (Bio-Rad) as previously described [35]. Briefly, for Nest I PCR, a final reaction volume of 13 μl was prepared by addition of 3 μl of parasite DNA to a mixture containing 5.75 μl of DreamTaq Green PCR Master Mix (2X), 0.5 μl of Nest I forward primer (10 μM), 0.5 μl of Nest I reverse primer (10 μM) and 3.25 μl of double-distilled water. For Nest II, a final reaction volume of 23 μl was prepared by addition of 3 μl of Nest I amplicon to a mixture containing 11.5 μl of DreamTaq Green PCR Master Mix (2X), 0.5 μl of Nest I forward primer (10 μM), 0.5 μl of Nest I reverse primer (10 μM) and 7.5 μl of double-distilled water. Nested I and II PCR conditions were set as follows: initial denaturation at 94°C for 3 minutes, 34 cycles of denaturation at 94°C for 30 sec, annealing at 50°C for 30 sec, primer extension at 68°C for 1 min and final extension at 72°C for 6 min. PCR amplification was performed on T100™ Thermal Cycler (Bio-Rad, Hercules, CA, USA). The quality of the 460bp amplicons from Nested II PCR were assessed by performing gel electrophoresis in 1.5% w/v agarose gel after amplification and before sequencing. Each of the 180 PCR products was purified by the addition of Exonuclease I and Shrimp Alkaline Phosphatase (ExoSAP-IT) and incubated at 37°C for 15 min. The ExoSAP-IT in purified PCR products was inactivated by heating at 80°C for 15 min. The cleaned PCR products with quantity above 25 ng/μl were selected for sequencing after testing 1 μl of each using NanoDrop™ Lite Spectrophotometer (Thermo Scientific™). Bi-directionally sequencing was done using BigDye® Terminator v3.1 Sequencing Standard kit on ABI PRISM® 3700 DNA Analyzer (Applied Biosystems, Foster City, CA, USA). Repeat sequencing was done on samples with poor reads to confirm mutated sites.

**Ethics approval**

The study was approved by the Maseno University Ethics Review Committee (MUERC protocol No. 00456) and the University of California, Irvine Institutional Review Board (HS#2017-
Data analysis

The sequence assembly was done using Geneious version 11.1.5 software. Multiple sequence alignment was done using ClustalW, a matrix-based algorithm that is in build in Mega X software [36]. Segregating sites, mutated codons, and type of mutation were determined by inputting sequences from each region and reference sequence NF54 (Accession number X07802.1) in CodonCode Aligner version 9.0.1 (CodonCode Corporation, www.codoncode.com). Mixed haplotype infections were defined as that sample had sequencing chromatogram with double peaks (second peak higher that 50% of first peak). The PHASE function of DnaSP Version 6.12.03 was used to infer the haplotypes of mixed infections. The Pfs25 antigen structural delineation was done by inputting full-length sequence (X07802.1) on Protein Homology/analogY Recognition Engine (PHYRE2) version 2.0. The generated Pfs25 model by Phyre2 was visualized, modified and mutated sites marked in UCSF Chimera version 1.15 [37]. The genetic diversity indices; nucleotide diversity (\( \pi \)) [38], the mean number of pairwise difference (k) [39], number of segregating sites (S), number of haplotypes (h), haplotype diversity (\( H_d \)) were computed per region from the aligned sequences using DnaSP software [40] and results confirmed using Arlequin version 3.5.2 [41]. Haplotype data generated from DnaSP in nexus format with slight modification was used to draw a haplotype network using Population Analysis with Reticulate Trees (Popart) version 1.7 software [42]. The ratio of nonsynonymous to synonymous mutations (\( d_N/d_S \)) was computed to give a hint on selection pressure acting on Pfs25 antigen from this region [43–46]. Presence and type of natural selection or deviation from the standard neutral model were determined through computation of neutrality tests such Tajima’s \( D \) [47], Fu and Li’s \( D \), Fu and Li’s \( F \) and Fu’s \( F_s \) statistics [48]. The output of Tajima’s \( D \) values were exported to GraphPad version 8.3.0 and used to generate the Tajima’s \( D \) curve for both regions. The negative values from these tests correspond to a purifying natural selection whereas positive values are a representation of positive selection. Other factors or drivers sustaining the observed mutations such as recombination and inbreeding were confirmed by detecting the presence or absence of Recombination events (Rm) and Linkage disequilibrium (LD) were confirmed using DnaSP software.

Results

Distribution of polymorphisms in four domains of Pfs25 antigen

A total of 10 segregating or polymorphic sites (123, 124, 249, 330, 345, 364, 372, 390, 394 and 428) were identified across the 460bp of Pfs25 gene from P. falciparum isolates that were in circulation in western Kenya (S1 Fig). Homa Bay had 3/78 (3.8%) mutated sequences, Chulaimbo had 3/39 (7.7%), and Kisii had 3/60 (5%). The polymorphic sites were observed at a frequency of 7.9% (14/177) among sampled parasites (Table 1). The distribution of these polymorphic sites per study area was as follows: Homa Bay 5.1% (4/78), Chulaimbo 15.4% (6/39) and Kisii 6.7% (4/60). Most (80%) of the polymorphic sites (loci 123, 124, 249, 330, 345, 364, 372 and 428) were singletons or single nucleotide polymorphs (SNPs) whereas others (loci 390 and 394) were parsimony informative as they were observed in more than one sample. Three samples showed mixed infections of mutant/wild type (S1 Fig, S1 File). Transition and transversion equally contributed to the observed mutations across the 10 segregating sites (Table 1). The base substitution resulted in more (60%) synonymous amino acid or codon changes compared to nonsynonymous. Nonsynonymous changes in Pfs25 sequences from Homa Bay resulted in codon change L42M (Leucine to Methionine), whereas synonymous changes included H41H (Histidine to Histidine), C110C (Cysteine to Cysteine), and T142T (Threonine...
to Threonine). Each of these changes occurred at a 1.3% frequency, with two (H41H and T124T) occurring on the same sequence. On polymorphic sequences in Chulaimbo, synonymous T130T (Threonine to Threonine) and nonsynonymous mutation V132I (Valine to Isoleucine) occurred as a pair at a frequency of 7.7% each. Kisii parasites had codon changes I83I (Isoleucine to Isoleucine), C115W (Cysteine to Tryptophan), L122L (Leucine to Leucine), and V143G (Valine to Glycine), which were observed at a frequency of 1.61% each (Table 1). Polymorphic codons L122L and V143G occurred on a single sequence.

Out of the mutated 10 codons in Pfs25 antigen, 7 (330, 345, 364, 372, 390, 394 and 428) were present in D3 (Fig 2). Domain 2 had 1 (I83) mutated codon whereas D1 had two mutated codons (H41 and L42) which were close to D3 codons. In D3, mutated codon C110, L122, T130 and V132 were within structural proximity.

**Genetic diversity indices of Pfs25 gene**

Relatively low nucleotide diversity ($\pi$) was observed among the sequences analysed. The 177 Pfs25 sequences had $\pi$ of $0.04 \times 10^{-2}$, k of 0.16 and a nucleotide conservation index of 97.7% that slightly varied per study site (Table 2). Chulaimbo isolates had the highest observed $\pi$ ($0.07 \times 10^{-2}$) and Hd (0.15) followed by Kisii ($\pi = 0.03$ and Hd = 0.10) then Homa Bay ($\pi = 0.02$ and Hd = 0.08). There was no significant difference ($p = 0.400$) when the population pairwise FSTs were computed for Homa Bay and Kisii sequences (FST = 0.00). Similar results ($p = 0.160$) was observed for population differences between Chulaimbo and Kisii (FST = 0.037). No significant difference ($p = 0.110$) in observed allelic variation between Homa Bay and Chulaimbo sequences (FST = 0.039). The highest diversity index in Chulaimbo was due to a high frequency of T130T and V132I codon changes (Table 3). Homa Bay and Kisii had the highest number of Pfs25 haplotypes (four each) compared to Chulaimbo (Fig 3). Among analysed samples from Chulaimbo, blood samples of patients from the same household were found to harbour different haplotypes. Six blood samples, 3 from each different household had either Hap1 with no mutated sites or Hap5 with both mutated codon T130T and V132I. This was however not observed in sequences from Homa Bay and Kisii.

**Signature of selection and other mutation drivers on Pfs25 gene**

The computed ratio of nonsynonymous and synonymous substitutions ($d_\text{N}/d_\text{S}$) for all Pfs25 sequences from western Kenya was 5.37 (Table 1). The $d_\text{N}/d_\text{S}$ ratio was not only greater than 1

### Table 1. Polymorphic sites in Pfs25 sequences from Homa Bay, Kisumu (Chulaimbo) and Kisii region in western Kenya.

| Segregating sites | Domain | Allelic frequency | Substituted bases | Type of Substitution | Codon change | Type of mutation |
|-------------------|--------|------------------|------------------|---------------------|--------------|-----------------|
|                   |        | Homa Bay n (%)   | Chulaimbo n (%)  | Kisii n (%)         |              |                 |
| 123               | D1     | 1 (1.3)          | -                | -                   | T/C         | Transition      | H41H | Syn |
| 124               | D1     | 1 (1.3)          | -                | -                   | T/A         | Transversion    | L42M | Nsyn |
| 249               | D2     | -                | -                | 1 (1.7)             | A/T         | Transversion    | I83I | Syn |
| 330               | D3     | 1 (1.3)          | -                | -                   | T/C         | Transition      | C110C | Syn |
| 345               | D3     | -                | -                | 1 (1.7)             | T/G         | Transversion    | C115W | Nsyn |
| 364               | D3     | 1 (1.7)          | -                | -                   | T/C         | Transition      | L122L | Syn |
| 372               | D3     | 1 (1.3)          | -                | -                   | A/G         | Transition      | T124T | Syn |
| 390               | D3     | -                | 3 (7.7)          | -                   | T/A         | Transversion    | T130T | Syn |
| 394               | D3     | -                | 3 (7.7)          | -                   | G/A         | Transition      | V132I | Nsyn |
| 428               | D3     | -                | -                | 1 (1.7)             | T/G         | Transversion    | V143G | Nsyn |

n represents the proportion of sequences per study site with polymorphic sites.

https://doi.org/10.1371/journal.pone.0266394.t001
across all sequences but also in each of the three study sites. Neutrality test results (Tajima’s $D$) for all the Pfs25 sequences were negative, thus signifying the presence of natural selection. The Tajima’s $D$ value was -2.12 ($p < 0.05$), Fu and Li’s $D^*$ test statistic, $FLD^*$: -4.78 ($p < 0.02$), Fu and

![Fig 2. Predicted structure of Pfs25 antigen and distribution of polymorphic codons on its domain 1, 2 and 3.](https://doi.org/10.1371/journal.pone.0266394.g002)

Table 2. Summary of genetic diversity indices of Pfs25 gene in western Kenya.

| Study site   | N     | C (%) | S  | $\pi$ ($\times 10^{-2}$) | H   | $Hd$ | $Ks$ or $d_s$ ($\times 10^{-2}$) | $Ka$ or $d_N$ ($\times 10^{-2}$) | $d_N/d_S$ | Tajima’s $D$ | Fu’s $F_S$ |
|--------------|-------|-------|----|--------------------------|-----|------|---------------------------------|---------------------------------|-----------|--------------|-------------|
| Homa Bay     | 78    | 99.10 | 4  | 0.02                     | 4   | 0.08 | 0.09                            | 0.01                            | 12.29     | -1.81*       | -4.42*      |
| Chulaimbo    | 39    | 99.50 | 2  | 0.07                     | 2   | 0.15 | 0.16                            | 0.04                            | 3.93      | -0.73        | 0.80        |
| Kisii        | 60    | 99.10 | 4  | 0.03                     | 4   | 0.10 | 0.07                            | 0.02                            | 3.89      | -1.84*       | -3.90*      |
| All sites    | 177   | 97.70 | 10 | 0.04                     | 8   | 0.10 | 0.10                            | 0.02                            | 5.37      | -2.15*       | -10.91*     |

N: Sample size; C: Conservation index; S: Segregating sites; $\pi$: nucleotide diversity; H: haplotype; $Hd$: Haplotype diversity; $Ks$ or $d_s$: The number of synonymous (or silent) substitutions per synonymous (or silent) site by Jukes and Cantor; $Ka$ or $d_N$: The number of nonsynonymous substitutions per nonsynonymous site by Jukes and Cantor; $d_N/d_S$: Ratio of nonsynonymous and synonymous substitutions

*: Significance ($p < 0.05$).

https://doi.org/10.1371/journal.pone.0266394.t002
Li’s $F^*$ test statistic, FLF: -4.58 ($p < 0.02$) and Fu’s $F_S$ statistic: -10.91 ($p < 0.001$). A similar trend was observed on Tajima’s $D$: -1.81 ($p < 0.05$), and Fu’s $F_S$: -4.42 ($p = 0.011$) values of Pfs25 sequences from Homa Bay. Other test results FLD: -3.72 and FLF: -3.50 were also significant ($p < 0.05$). The selection pressure exacted its effect on loci existing between 98-148bp and 323-373bp in Pfs25 sequences from Homa Bay (Fig 4). These loci harbour the observed
segregating sites in Homa Bay sequences. Among sequences from Chulaimbo parasites, Taji-
ma’s $D$ and Fu’s $F_S$ were -0.73 and 0.8, respectively. All these test results were not significant
($p > 0.05$), with FLD* and FLF* values of 0.77 and 0.37 respectively. Deviation from the stan-
dard neutral model among the Chulaimbo samples was observed on loci existing between 348-
423bp where segregating sites were observed (Fig 4). Sequences from Kisii had significant
($p < 0.05$) Tajima’s $D$ (-1.84) and Fu’s $F_S$ (-3.9) values. The FLD* and FLF* results were -3.54
($p < 0.02$) and -3.52 ($p < 0.02$) respectively. Purging selection on loci blocks ranging from 223-
469bp was observed for these sequences (Fig 4).

To further assess the presence or history of drivers shaping allelic patterns within the four
domains of the Pfs25 antigen, we assessed a pair of loci that were under LD as well as having a
history of recombination events (Rm). Three pairs of loci (123, 372), (364, 428) and (390, 394)
within polymorphic Pfs25 sequences were observed to be under LD (Table 3). The three loci
pairs had highly significant ($P < 0.001$) LD values. The presence of significant LD values is an
indicator of inbreeding as a driver for the observed sequence changes in each study site. None
of mutated Pfs25 sequences from western Kenya had a history of recombination.

**Discussion**

This study revealed novel polymorphic sites across the four immuno-dominant domains of
the Pfs25 antigen. Most of these sites were dimorphic sites and involved Pfs25 D3. They
resulted in 9 haplotypes that were unique (“private”) not only to Homa Bay, Kisumu or Kisii
highlands but also to western Kenya. The Pfs25 gene exhibited low nucleotide diversity and a
high conservation index. Natural selection was identified to be the leading cause of these muta-
tions with purifying selection purging on specific loci blocks. Additionally, inbreeding was
established as an alternative driver for linked polymorphic loci in analysed sequences from
each study site. There was high level of interbreeding in *P. falciparum* population from Homa Bay, Chulaimbo and Kisii zones.

Generally, low nucleotide diversity was observed in Pfs25 sequences from parasites population in western Kenya compared to Asia ($0.11 \times 10^{-2}$), Thailand ($0.09 \times 10^{-2}$), Cambodia ($0.12 \times 10^{-2}$), India ($0.31 \times 10^{-2}$), Brazil ($0.18 \times 10^{-2}$) and Vietnam ($0.26 \times 10^{-2}$) [49]. Chulaimbo sequences had high nucleotide variation compared to Homa Bay and Kisii. The observed variation between the two malaria-endemic sites and a close similarity between Homa Bay and Kisii implies that ongoing indoor residual spraying targeting malaria vectors within Homa Bay [50] may be directly or indirectly reducing the numbers of parasite clones or genetic diversity. Despite having high number of haplotypes across western Kenya, their observed number was within the set limit of the total number of segregating plus one (S+1), thus implying the population is under selective sweep or growth [48, 51]. Apart from the dimorphic site 428 (codon V143) which had previously been reported in Pfs25 sequences from Cambodia [31, 49], India [30, 31, 49] and now in Kisii highland, the remaining nine sites were novel. The amino acid change observed in this codon (V143G), however, did not correspond to documented V143A mutation following differences in base substitution at locus 428 (T/G as opposed to T/C) of the gene in one of the sequences from the Kisii area. None of the sequences from western Kenya had G131A mutations that had been previously observed in the parasite population from India and Cambodia [31]. However, three sequences from Chulaimbo study sites had mutations at a nearby codon V132I, and was observed at a frequency of 7.7%. Compared to its analogue in *Plasmodium vivax*, two mutated codons T130 and V132 identified in Pfs25 had been described in Pv25 in Asia [52, 53].

The majority of the polymorphic sites we discovered were in the Pfs25 D3 region. This domain had previously been described to be highly immunogenic compared to others, and is also known for harboring binding sites of 1D2 and 4B7 human monoclonal antibodies used in standard membrane feeding assay studies of *P. falciparum* mosquito infectivity [14, 25, 54]. Among mutated codons in D3, codon change C110C and C115W targeted two of the 21 cysteine residues that are known to be highly conserved and function in maintaining the folding pattern of the Pfs25 polypeptide [14, 31, 55, 56]. The presence of such silent and nonsynonymous mutation points to the possibility of a weak or strong selection pressure acting on nearby codons within the Pfs25 antigen [57]. If pronounced, such mutation may interfere with the folding pattern of the Pfs25 antigen, and thereby impact interaction between host antibodies elicited by Pfs25 TBV and modify interactions with other B-cell epitopes as observed with other proteins [58] and cysteine residues in the D2 region of Pfs47 [58].

As opposed to previous reports describing the absence of selective pressure acting on the Pfs25 gene [28], our studies suggest that natural selection may play a role in shaping allelic diversity of the Pfs25 in western Kenya. Specifically, purifying selection was suggested based on computation of the ratio of nonsynonymous and synonymous substitutions [43–46]. This was evident across all mutated codon since they displayed a significant Tajima’s $D$ value. The loci pairs under LD not only reaffirmed the history of selection but also confirmed inbreeding as another factor in the spread and sustenance of such mutations [59]. We speculate that selection pressure may be arising from unknown factor in protease-rich midgut of female Anophe- line vectors.

**Conclusion**

The Pfs25 gene from the malaria-endemic and epidemic-prone region of western Kenya revealed varying levels of genetic diversity and novel haplotypes. Purifying selection and inbreeding were predicted as the cause and drivers for the observed variations. The low level of
nucleotide diversity is an indicator that TBV based on Pfs25 sequences is less likely to be affected by antigenic variations. However, this study recommends continued monitoring of the Pfs25 gene from different malaria-prone regions including areas where clinical trials have been conducted. This will not only aid in unravelling new polymorphic sites that could have an effect on antibody binding especially on immunogenic D3 but also monitor the durability of the Pfs25 gene as a potential TBV.

**Supporting information**

**S1 Fig.** Sequence chromatograms for mutated loci within Pfs25 gene from Plasmodium falciparum in western Kenya. The sequence chromatograms are aligned to the reference sequence (PF3D7_1031000). The vertical red arrows indicate the position of the SNPs. The asterisk (*) represents samples with mixed haplotype infections. The sequences are also deposited in GenBank with accession numbers MT212735-MT212808 and MT225462-MT225527. (PDF)

**S1 File.** Raw sequencing chromatograms. (RAR)

**S1 Data.**

(FAS)

**S2 Data.**

(FAS)

**S3 Data.**

(FAS)

**Acknowledgments**

Sincere thanks to Carey O. Osemba (Laboratory technician at Eramba health facility), Homa Bay and Chulaimbo health facilities staff, study participants and the entire ICEMR personnel who participated in this research study. We are also grateful to the editor and two anonymous reviewers for their constructive comments that have helped to improve this manuscript.

**Author Contributions**

**Conceptualization:** Kevin O. Ochwedo, Andrew K. Githeko, Antony C. A. Otieno, Wolfgang R. Mukabana, Guiyun Yan, Daibin Zhong, James W. Kazura.

**Data curation:** Kevin O. Ochwedo, Daibin Zhong.

**Formal analysis:** Kevin O. Ochwedo.

**Funding acquisition:** Guiyun Yan, James W. Kazura.

**Investigation:** Kevin O. Ochwedo.

**Methodology:** Kevin O. Ochwedo, Shirley A. Onyango, Collince J. Omondi, Pauline W. Orondo, Benyl M. Ondeto, Ming-Chieh Lee, Wolfgang R. Mukabana, Guiyun Yan, Daibin Zhong, James W. Kazura.

**Project administration:** Kevin O. Ochwedo, Harrysone E. Atieli, Sidney O. Ogolla.

**Resources:** Kevin O. Ochwedo.

**Software:** Kevin O. Ochwedo, Daibin Zhong.
Supervision: Kevin O. Ochwedo, Wolfgang R. Mukabana, Guiyun Yan.

Validation: Kevin O. Ochwedo, Daibin Zhong.

Visualization: Kevin O. Ochwedo, Ming-Chieh Lee.

Writing – original draft: Kevin O. Ochwedo.

Writing – review & editing: Kevin O. Ochwedo, Shirley A. Onyango, Sidney O. Ogolla, Andrew K. Githeko, Antony C. A. Otieno, Wolfgang R. Mukabana, Guiyun Yan, Daibin Zhong, James W. Kazura.

References

1. WHO. World malaria report 2019. 2019 [cited 17 Mar 2022]. Available: https://www.who.int/publications/i/item/9789241565721

2. Alout H, Roche B, Dabirié RK, Cohuet A. Consequences of insecticide resistance on malaria transmission. Knoll LJ, editor. PLOS Pathogens. 2017; 13: e1006499. https://doi.org/10.1371/journal.ppat.1006499 PMID: 28880906

3. Draper SJ, Sack BK, Nielsen CM, Rayner JC, Higgins MK, et al. Cell Host & Microbe Review Malaria Vaccines: Recent Advances and New Horizons. 2018. https://doi.org/10.1016/j.chom.2018.06.008

4. Kaslow DC, Biernoix S. RTS,S: Toward a first landmark on the Malaria Vaccine Technology Roadmap. Vaccine. 2015; 33: 7425–7432. https://doi.org/10.1016/j.vaccine.2015.09.061 PMID: 26431982

5. Ouattara A, Laurens MB. Vaccines Against Malaria. Clinical Infectious Diseases. 2015; 60: 930–936. https://doi.org/10.1093/cid/ciu954 PMID: 25452593

6. Bergmann-Leitner ES, Duncan EH, Mease RM, Angov E. Impact of pre-existing MSP142-allele specific immunity on potency of an erythrocytic Plasmodium falciparum vaccine. Malaria Journal. 2012; 11: 315. https://doi.org/10.1186/1475-2875-11-315 PMID: 22958482

7. Genton B, Betuela I, Felger I, Al-Yaman F, Anders RF, Saul A, et al. A Recombinant Blood-Stage Malaria Vaccine Reduces Plasmodium falciparum Density and Exerts Selective Pressure on Parasite Populations in a Phase 1–2b Trial in Papua New Guinea. The Journal of Infectious Diseases. 2002; 185: 820–827. https://doi.org/10.1086/339342 PMID: 11920300

8. Miura K, Herrera R, Diouf A, Zhou H, Mu J, Hu Z, et al. Overcoming allelic specificity by immunization with five allelic forms of Plasmodium falciparum apical membrane antigen 1. Infection and immunity. 2013; 81: 1491–1501. https://doi.org/10.1128/IAI.01414-12 PMID: 23429537

9. Ogutu BR, Apollo OJ, McKinney D, Okoth W, Siangla J, Dubovskij F, et al. Blood Stage Malaria Vaccine Eliciting High Antigen-Specific Antibody Concentrations Confers No Protection to Young Children in Western Kenya. Beeson JG, editor. PLoS ONE. 2009; 4: e4708. https://doi.org/10.1371/journal.pone.0004708 PMID: 19262754

10. Takala SL, Coulibaly D, Thera MA, Dicko A, Smith DL, Guindo AB, et al. Dynamics of Polymorphism in a Malaria Vaccine Antigen at a Vaccine-Testing Site in Mali. Olliaro PL, editor. PLoS Medicine. 2007; 4: e93. https://doi.org/10.1371/journal.pmed.0040093 PMID: 17355170

11. Kengne-Ouafou JA, Sutherland CJ, Binka FN, Awandare GA, Urban BC, Dinko B. Immune responses to the sexual stages of Plasmodium falciparum parasites. Frontiers in Immunology. 2019; 10. https://doi.org/10.3389/fimmu.2019.00136 PMID: 30804940

12. Brod F, Miura K, Taylor I, Li Y, Marini A, Salman AM, et al. Combination of RTS,S and Pfs25-IMX313 Induces a Functional Antibody Response Against Malaria Infection and Transmission in Mice. Frontiers in immunology. 2018; 9: 2780. https://doi.org/10.3389/fimmu.2018.02780 PMID: 30564231

13. Chichester JA, Green BJ, Jones RM, Shoji Y, Miura K, Long CA, et al. Safety and immunogenicity of a plant-produced Pfs25 virus-like particle as a transmission blocking vaccine against malaria: A Phase 1 dose-escalation study in healthy adults. Vaccine. 2018; 36: 5865–5871. https://doi.org/10.1016/j.vaccine.2018.08.033 PMID: 30126674

14. Scally SW, McLeod B, Bosch A, Miura K, Liang Q, Carroll S, et al. Molecular definition of multiple sites of antibody inhibition of malaria transmission-blocking vaccine antigen Pfs25. Nature communications. 2017; 8: 1568. https://doi.org/10.1038/s41467-017-01924-3 PMID: 29146922

15. University of Oxford. A Phase Ia Study to Assess Safety and Immunogenicity of the Plasmodium Falciparum Malaria Vaccine Candidate Pfs25-IMX313 in Matrix-M1 Adjuvant in Healthy Adults Living in the UK. clinicaltrials.gov; 2020 Sep. Report No.: NCT04130282. Available: https://clinicaltrials.gov/ct2/show/NCT04130282
16. Baton LA, Ranford-Cartwright LC. Do malaria ookinete surface proteins P25 and P28 mediate parasite entry into mosquito midgut epithelial cells? | Malaria Journal | Full Text. 2005 [cited 17 Mar 2022]. Available: https://malariajournal.biomedcentral.com/articles/10.1186/1475-2875-4-15 PMID: 15733320
17. Tomas AM, Margos G, Dimoupolous G, Lin LHMV, Koning-Ward TFD, Sinha R, et al. P25 and P28 proteins of the malaria ookinete surface have multiple and partially redundant functions. EMBO Journal. 2001; 20: 3975–3983. https://doi.org/10.1093/embob/20.15.3975 PMID: 11483501
18. Miao J, Fan Q, Parker D, Li X, Li J, Cui L. Puf mediates translation repression of transmission-blocking vaccine candidates in malaria parasites. PLoS pathogens. 2013; 9: e1003268. https://doi.org/10.1371/journal.ppat.1003268 PMID: 23637595
19. Stowers AW, Keister DB, Muratova O, Kaslow DC. A Region of Plasmodium falciparum Antigen Pfs25 That Is the Target of Highly Potent Transmission-Blocking Antibodies. 2000; 68: 5530–5538. https://doi.org/10.1128/IAI.68.10.5530-5538.2000 PMID: 10992450
20. Kaslow DC, Quakyl IA, Syin C, Raum MG, Keister DB, Coligan JE, et al. A vaccine candidate from the sexual stage of human malaria that contains EGF-like domains. Nature. 1988; 333: 74–76. https://doi.org/10.1038/333074a0 PMID: 3283563
21. Huang WC, Deng B, Mabrouk MT, Seffouh A, Ortega J, Long C, et al. Particle-based, Pfs230 and Pfs25 immunization is effective, but not improved by duplexing at fixed total antigen dose. Malaria Journal. 2020; 19: 309. https://doi.org/10.1186/s12936-020-03368-5 PMID: 32859199
22. Kumar R, Angov E, Kumar N. Potent malaria transmission-blocking antibody responses elicited by Plasmodium falciparum Pfs25 expressed in Escherichia coli after successful protein refolding. Infection and Immunity. 2014; 82: 1453–1459. https://doi.org/10.1128/IAI.01438-13 PMID: 24421036
23. Lee S-M, Wu C-K, Plieskatt J, McAdams DH, Miura K, Ockenhouse C, et al. Assessment of Pfs25 expressed from multiple soluble expression platforms for use as transmission-blocking vaccine candidates. Malaria Journal. 2016; 15: 405. https://doi.org/10.1186/s12936-016-1464-6 PMID: 27515826
24. Li Y, Leneghan DB, Miura K, Nikolaeva D, Brian IJ, Dicks MDJ, et al. Enhancing immunogenicity and transmission-blocking activity of malaria vaccines by fusing Pfs25 to IMX313 multimerization technology. Scientific Reports. 2016; 6: 1–13. https://doi.org/10.1038/s41598-016-0001-8 PMID: 28442746
25. McGuire KA, Miura K, Wiethoff CM, Williamson KC. New adenovirus-based vaccine vectors targeting Pfs25 elicit antibodies that inhibit Plasmodium falciparum transmission. Malaria Journal. 2017; 16: 254. https://doi.org/10.1186/s12936-017-1896-7 PMID: 28619071
26. McLeod B, Miura K, Scally SW, Bosch A, Nguyen N, Shin H, et al. Potent antibody lineage against malaria transmission elicited by human vaccination with Pfs25. Nature Communications. 2019; 10. https://doi.org/10.1038/s41598-019-11980-6 PMID: 31551421
27. Talaat KR, Ellis RD, Hurd J, Hentrich A, Gabriel E, Hynes NA, et al. Safety and immunogenicity of Pfs25-EPA/ alhydro gel1, a transmission blocking vaccine against Plasmodium falciparum: An open label study in malaria naïve adults. PLoS ONE. 2016; 11. https://doi.org/10.1371/journal.pone.0163144 PMID: 27749907
28. Escalante AA, Lal AA, Ayala FJ. Genetic polymorphism and natural selection in the malaria parasite Plasmodium falciparum. Genetics. 1998; 149: 198–202. https://doi.org/10.1093/genetics/149.1.189 PMID: 9584096
29. Kaur H, Sehgal R, Goyal K, Makkar N, Yadav R, Bharti PK, et al. Genetic diversity of Plasmodium falciparum merozoite surface protein-1 (block 2), glutamate-rich protein and sexual stage antigen Pfs25 from Chandigarh, North India. Tropical medicine & international health: TM & IH. 2017; 22: 1590–1598. https://doi.org/10.1111/tmi.12990 PMID: 29029367
30. Patel P, Bhati PK, Bansal D, Raman RK, Mohapatra PK, Sehgal R, et al. Genetic diversity and antibody responses against Plasmodium falciparum vaccine candidate genes from Chhattisgarh, Central India: Implication for vaccine development. Tetteh KKA, editor. PLOS ONE. 2017; 12: e0182674. https://doi.org/10.1371/journal.pone.0182674 PMID: 29029677
31. Juliano JJ, Parobek CM, Brazeau NF, Ngasala B, Randrianarivelojosia M, Lon C, et al. Pooled Ampli-con Deep Sequencing of Candidate Plasmodium falciparum Transmission-Blocking Vaccine Antigens. Am J Trop Med Hyg. 2016; 94: 143–146. https://doi.org/10.4269/ajtmh.15-0571 PMID: 26503281
32. Coombs RW, Fiscus S. Processing of Dried Blood Spots Standard Operating Procedure Title: Processing of Dried Blood Spots Standard Operating Procedure Approved By (Network): Network Name, Title Signature Date Revision History Version Effective Date (dd/mm/yyyy) Comments. 2009.
33. Bereczky S, Mårtensson A, Gil JP, Färnert A. Short report: Rapid DNA extraction from archive blood spots on filter paper for genotyping of Plasmodium falciparum. American Journal of Tropical Medicine and Hygiene. 2005; 72: 249–251. https://doi.org/10.4269/AJTMH.2005.72.249 PMID: 15772315
34. Veron V, Simon S, Carme B. Multiplex real-time PCR detection of P. falciparum, P. vivax and P. malariae in human blood samples. Experimental parasitology. 2009; 121: 346–351. https://doi.org/10.1016/j.exppara.2008.12.012 PMID: 19124021

35. Mlambo G, Vasquez Y, LeBlanc R, Sullivan D, Kumar N. A filter paper method for the detection of Plasmodium falciparum gametocytes by reverse transcription polymerase chain reaction. The American journal of tropical medicine and hygiene. 2008; 78: 114–116. PMID: 18187793

36. Kumar Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution. 2018; 35: 1547–1549. https://doi.org/10.1093/molbev/msy096 PMID: 29722887

37. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—A visualization system for exploratory research and analysis. Journal of Computational Chemistry. 2004; 25: 1605–1612. https://doi.org/10.1002/jcc.20084 PMID: 15264254

38. Nei M. MOLECULAR PHYLOGENY AND GENETIC DIVERSITY ANALYSIS. 1997. Available: www.bio.psu.edu/imeg/neilab.htm.

39. Tajima F. The effect of change in population size on DNA polymorphism. Genetics. 1989; 123. https://doi.org/10.1093/genetics/123.3.585 PMID: 2513255

40. Fu YX, Li WH. Statistical tests of neutrality of mutations. Genetics. 1993; 133: 693–709. https://doi.org/10.1093/genetics/133.3.693 PMID: 8454210

41. Sookpongthai P, Kaewthamasorn M, Wangkanont K. Global Diversity of the Gene Encoding the Pfs25 Protein—A Plasmodium Falciparum Transmission-Blocking Vaccine Candidate. 2021. https://doi.org/10.1186/s13071-021-00578-2 PMID: 34749796

42. President’s Malaria Initiative. Kenya–The U.S. President’s Malaria Initiative VectorLink Project. 2021 [cited 17 Mar 2022]. Available: https://pmivectorlink.org/where-we-work/kenya/

43. Depaulis F, Veuille M. Neutrality tests based on the distribution of haplotypes under an infinite-site model. Molecular Biology and Evolution. 1998; 15: 1788–1790. https://doi.org/10.1093/oxfordjournals.molbev.a025905 PMID: 9917213

44. Peng H, Zheng L, Zhu X, Wang G, Pan Y, Li Y, et al. Genetic diversity of transmission-blocking vaccine candidates Pvs25 and Pvs28 in Plasmodium vivax isolates from Yunnan Province, China. Parasites & Vectors 2011 4:1. 2011; 4: 1–8. https://doi.org/10.1186/1756-3305-4-224 PMID: 22117620

45. Lé HG, Kang JM, Jun H, Lee J, Moe M, Thái TL, et al. Genetic diversity and natural selection of transmission-blocking vaccine candidate antigens Pvs25 and Pvs28 in Plasmodium vivax Myanmar isolates. Acta Tropica. 2019; 198: 105104. https://doi.org/10.1016/j.actatropica.2019.105104 PMID: 31396059

46. Patel PN, Tolia N. Structural vaccinology of malaria transmission-blocking vaccines. Expert Review of Vaccines. Taylor and Francis Ltd.; 2021. pp. 199–214. https://doi.org/10.1080/14760584.2021.1873135 PMID: 33430656

47. Duffy PE, Kaslow DC. A novel malaria protein, Pfs28, and Pfs25 are genetically linked and synergistic as falciparum malaria transmission-blocking vaccines. Infection and Immunity. 1997; 65: 1109–1113. https://doi.org/10.1128/IAI.65.3.1109-1113.1997 PMID: 9038325
Saxena AK, Singh K, Su HP, Klein MM, Stowers AW, Saul AJ, et al. The essential mosquito-stage P25 and P28 proteins from Plasmodium form tile-like triangular prisms. Nature Structural and Molecular Biology. 2006; 13: 90–91. https://doi.org/10.1038/nsmb1024 PMID: 16327807

Lawrie DS, Messer PW, Hershberg R, Petrov DA. Strong Purifying Selection at Synonymous Sites in D. melanogaster. Plotkin JB, editor. PLoS Genetics. 2013; 9: e1003527. https://doi.org/10.1371/journal.pgen.1003527 PMID: 23737754

Canepa GE, Molina-Cruz A, Yenkoidok-Douti L, Calvo E, Williams AE, Burkhardt M, et al. Antibody targeting of a specific region of Pfs47 blocks Plasmodium falciparum malaria transmission. npj Vaccines. 2018; 3: 26. https://doi.org/10.1038/s41541-018-0065-5 PMID: 30002917

Larrañaga N, Mejía RE, Hormaza JI, Montoya A, Soto A, Fontecha GA. Genetic structure of Plasmodium falciparum populations across the Honduras-Nicaragua border. Malaria Journal. 2013; 12. https://doi.org/10.1186/1475-2875-12-354 PMID: 24093629