Regional adaptation to mosquito vectors shapes *Plasmodium falciparum* populations

Duangkamon Loesbanluechai[1], Lauriane Sollelis[2,3], Megan Armstrong[1], Inish Menezes[1], Amelia Cox[1], Lizzie Bridget Tchongwe Divala[2], Anne Lawson[1], Sabyasachi Pradhan[1], Prince Chigozirim Ubiaru[1], Antrea Pallikara[1], Dorothy Armstrong[1], Gillian Parker[2], Chiyun Lee[4], Richard D Pearson[4], Barbara H Stokes[2,3], Lisa Ranford-Cartwright[1], Matthias Marti[2,3], Virginia Howick[1,*]

[1] School of Biodiversity, One Health and Veterinary Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, Scotland, U.K.

[2] School of Infection and Immunity, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, Scotland, U.K.

[3] Institute of Parasitology Zurich, VetSuisse and Medical Faculties, University of Zurich, Zurich, Switzerland

[4] Genomic Surveillance Unit, Wellcome Sanger Institute, Hinxton, U.K.

[*] Virginia.Howick@glasgow.ac.uk

Abstract

Plasmodium falciparum, the most virulent human malaria parasite, is transmitted by *Anopheles* mosquitoes whose community composition varies geographically. Population genomic analyses reveal highly differentiated *P. falciparum* loci with roles in mosquito infection, which may be driven by adaptation to regional vectors. To test this, we generated transgenic parasites in which reference alleles were replaced with geographically alternative variants at candidate transmission-stage genes. Transmissibility was compared across four mosquito species representing distinct geographic ranges (*An. gambiae*, *An. stephensi*, *An. minimus*, and *An. albimanus*). Two of five tested polymorphisms increased oocyst and sporozoite burdens in sympatric parasite–vector combinations. Both substitutions occurred in ookinete micronemal proteins, CTRP and WARP, within von Willebrand factor A domains, suggesting that regional allelic variation modulates *Plasmodium*–vector compatibility by altering midgut adhesion interactions.

Introduction

Malaria remains one of the world's most devastating infectious diseases, causing more than 500,000 deaths each year (1). The disease is caused by *Plasmodium* parasites transmitted through the bites of *Anopheles* mosquitoes. Among these parasites, *Plasmodium falciparum* is the most virulent and is widespread across the tropics. However, the community composition of *Anopheles* vectors varies geographically and may impose strong, regionalized selective pressures on the parasite (2–4).

Transmission through the mosquito represents the most severe population bottleneck in the *Plasmodium* life cycle, making it a critical target for transmission blocking interventions. After being ingested during an infectious blood meal, parasites must undergo rapid sexual differentiation, fertilization, and meiosis before forming invasive ookinetes that traverse the mosquito midgut epithelium (5). Only a small fraction of ookinetes successfully cross this barrier: parasite numbers often drop from thousands of gametocytes to fewer than ten oocysts (6). In incompatible vectors, parasites are eliminated by the mosquito immune system or fail to exit the midgut lumen (7, 8).

A key molecular determinant of this compatibility encodes the parasite antigen Pfs47, which mediates evasion of the mosquito immune response. Variants of Pfs47 act as the "key" that must match a mosquito "lock" for successful infection (9). Population genomic analyses reveal strong geographic structuring, with near fixation of distinct Pfs47 haplotypes across major continental regions, consistent with adaptation to regional vector communities (9–12). However, genome-wide analyses indicate that many other *P. falciparum* genes expressed during mosquito transmission show similarly strong geographic differentiation (13, 14). These patterns suggest that vector compatibility may be a polygenic trait, shaped by multiple interacting loci rather than a single lock-and-key mechanism.

Here, we tested this hypothesis by generating allelic replacement lines for five geographically differentiated transmission-stage genes and assaying their infectivity across four *Anopheles* species with distinct geographic ranges. We found that for two genes, sympatric parasite–vector combinations achieved higher transmission success than allopatric pairings. Both loci encode micronemal proteins, with substitutions localized to the same adhesive functional domain, suggesting direct interactions with mosquito midgut components. These findings support a model of vector compatibility driven by coevolution between *Plasmodium* and local mosquito populations through multiple molecular interfaces.

Results

Identification and allelic replacement of candidate loci

To identify loci involved in adaptation to regional mosquito vectors, we integrated population genomic and gene expression data to pinpoint variants that may influence parasite infectivity in the mosquito midgut. The MalariaGEN network analysed genomic variation in 7,000 *P. falciparum* samples and ranked genes by a global differentiation score based on the most divergent non-synonymous SNP per gene (13). We selected the top 5% of genes (278) from this list and cross-referenced them with Malaria Cell Atlas expression data (15, 16) to identify those associated with gametocytes or ookinetes and/or upregulated in *P. falciparum* relative to *P. berghei* during the ookinete stage (844 genes). These expression criteria enriched for genes likely involved in midgut traversal,

and we hypothesized that transcriptional differences between *P. falciparum* and *P. berghei* reflect vector-specific adaptations to their respective mosquito hosts (17). This intersection yielded 60 candidates (**Table S1**), from which eight were selected (**Figure 1A**), on the basis of high expression in mature gametocytes or ookinetes, and presence of high-frequency alternative alleles in geographic regions that coincided with mosquito species available in our insectary (**Table 1; Figure 1; Figure S1**). Six of the genes (CTRP, Pfs25, Pfs47, SOAP, WARP, P230) have established roles in transmission and are targets for potential transmission-blocking interventions (18–21). P230, P25, and P47 are surface proteins involved in gamete interaction, midgut traversal, and immune evasion (11, 22–25). SOAP, WARP, and CTRP localize to ookinete micronemes and mediate motility and midgut invasion (26–30), with CTRP knockout blocking parasite invasion of the mosquito midgut (31). The remaining two loci, HSP101 and Pf3D7_1004600, have not previously been linked to transmission but show higher expression in *P. falciparum* ookinetes compared to *P. berghei* (15). This species-specific expression suggests possible adaptation to distinct mosquito hosts. HSP101 encodes a PTEX translocon component active in asexual blood stages (32), while Pf3D7_1004600 encodes a protein of unknown function.

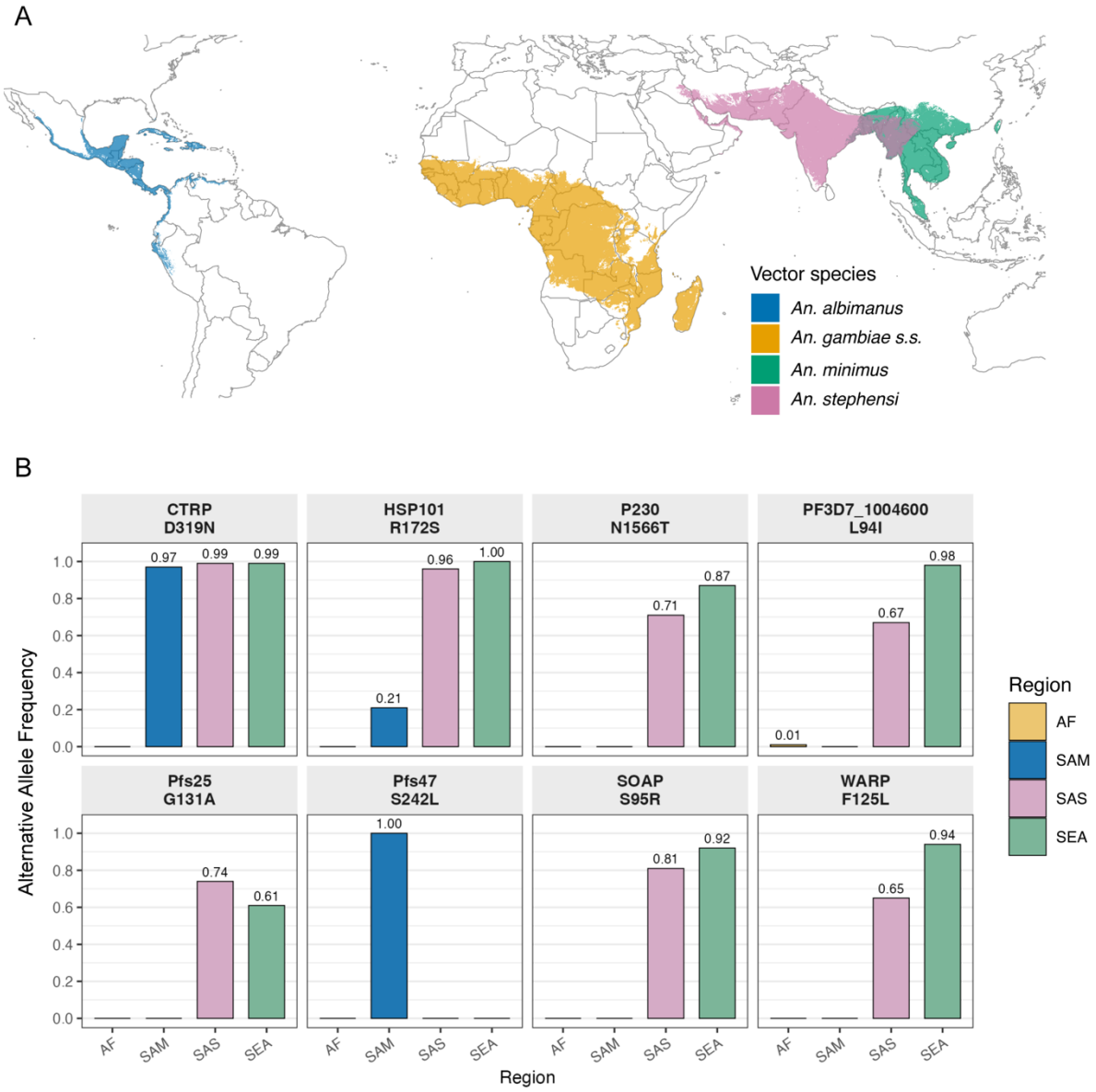

Figure 1. Distribution of *Anopheles* species and *P. falciparum* genotypes (A) The global distributions of *An. albimanus, An. gambiae s.s., An. minimus, and An. stephensi* from the Malaria Atlas Project (2, 33)**.** Shaded areas indicate the probability of presence is greater than 50%. (B) Alternative allele frequencies for the eight candidate loci across the four broad geographic areas associated with each of the vector species: Africa (AF), South America (SAM), South Asia (SAS), Southeast Asia (SEA). Data are from (34).

To assess the role of these non-synonymous SNPs in mosquito-specific adaptation, we used CRISPR-Cas9 to generate transgenic *P. falciparum* parasites by inserting the alternative allele into the reference parasite strain 3D7, which is of African origin, to generate a single amino acid change in each gene (**Figure S2**). For each gene, we also generated a synonymous (silent) control line that recoded the gRNA/PAM region to prevent Cas9 re-cutting while leaving the protein sequence unchanged, serving as a transgenesis-handling control for our phenotypic assays (**Table S2**). Successful genome editing was achieved for five of eight genes: CTRP, Pfs47, Pfs25, WARP, *and* HSP101. Attempts to edit SOAP, P230, and Pf3D7_1004600 using alternative guide RNAs were unsuccessful, as parasites did not recover post-transfection despite multiple efforts.

Table 1. List of candidate loci. Eight candidate genes were identified for allelic replacement. Shown are the gene ID, gene name, amino acid (AA) change, RNA expression pattern of the gene, global differentiation score with rank, and whether transfection was successful (✓) or not (✗). Gene expression data were analysed from the Malaria Cell Atlas (15, 16) (**Figure S1**). Global differentiation score, a gene-level *Fst*-derived metric, and rank out of the 5,561 *P. falciparum* genes are from the MalariaGEN Pf6 dataset (13).

Gene ID	Gene name	AA change	Highest gene expression stage	Global differentiation score (rank)	Transfection (3D7)
PF3D7_1346800	Pfs47	S242L	Gametocyte	1 (1)	✓
PF3D7_0315200	CTRP	D319N	Ookinete	0.98 (5)	✓
PF3D7_1404300	SOAP	S95R	Ookinete	0.62 (120)	✗
PF3D7_0801300	WARP	F125L	Ookinete	0.66 (64)	✓
PF3D7_0209000	P230	N1566T	Gametocyte	0.61 (141)	✗
PF3D7_1031000	Pfs25	G131A	Ookinete	0.58 (205)	✓
PF3D7_1116800	HSP101	R172S	Ookinete	0.81 (7)	✓
PF3D7_1004600	.	L94I	Ookinete	0.67 (53)	✗

Two genes show higher infection with sympatric parasite-allele and vector-species combinations

Based on our hypothesis that these genes are involved in regional adaptation to vector communities, we expected that replacement of the reference allele with the alternative allele should increase infectivity in species that are sympatric for the alternative allele and decrease infectivity in allopatric ones. To test this, we performed Standard Membrane Feeding Assays (SMFAs) using five allelic-replacement lines and their silent controls across four *Anopheles* species with distinct geographic ranges (**Figure 1A**). Exflagellation rates prior to infectious feeding showed no differences between replacement and silent control lines (**Figure S3**), confirming comparable male gametocyte quality.

In total, 4,708 mosquitoes were infected across all experiments. Infection levels were assessed at the oocyst stage by measuring prevalence (proportion infected) and intensity (infection level among infected mosquitoes) (**Figures 2**). Oocysts were counted and measured 10 days post-infectious feed from dissected midguts. *An. gambiae* infected with the unmodified 3D7 clone was included as a positive control to confirm successful infection in each experimental replicate (**Figure S4**).

Significant mosquito-by-allele interactions were detected for two genes, CTRP and WARP, with sympatric combinations showing higher infection levels compared to allopatric ones. Both genes showed a mosquito-by-allele effect on oocyst prevalence ($p < 0.001$; **Figure 2B**). For CTRP D319N, oocyst prevalence was reduced in *An. gambiae* (allopatric; adj $p < 0.001$) but elevated in *An. stephensi* (sympatric; adj $p < 0.001$).

Similarly, the WARP F125L allele increased oocyst prevalence in *An. minimus* (sympatric; adj p = 0.018) and decreased it in *An. gambiae* (allopatric; adj p = 0.003) (**Figure 2A**).

WARP also showed a significant mosquito-by-allele effect on oocyst intensity ($p < 0.001$; **Figure 2C**), with the WARP F125L allele yielding higher infection loads in *An. minimus* (sympatric, adj p = 0.047) and *An. stephensi* (sympatric, adj p = 0.040) and lower loads in *An. gambiae* (allopatric, adj p = 0.003) (**Figure 2A**). No significant differences in intensity were observed for CTRP D319N. The remaining alleles (HSP101 R127S, Pfs25 G131A, and Pfs47 S242L) had no measurable effects on oocyst prevalence or intensity (**Figure 2A–C**). There was also no mosquito-by-allele effect on oocyst size at day 10 post-infectious feed in any of the genes tested (**Figure S5**).

For CTRP and WARP, an additional clone of both the allelic-replacement and silent control was tested to confirm the observed vector-specific transmission phenotype (**Figure S6**). We also quantified sporozoite prevalence and intensity in the mosquito salivary glands using qPCR and found a significant mosquito-by-allele interaction for WARP (p = 0.034), with the alternative allele (F125L) showing lower sporozoite prevalence compared to the reference in *An. gambiae* (allopatric, adj p = 0.018). Similarly, CTRP D319N was associated with reduced sporozoite prevalence in *An. gambiae* (adj p = 0.027) (**Figure S7**).

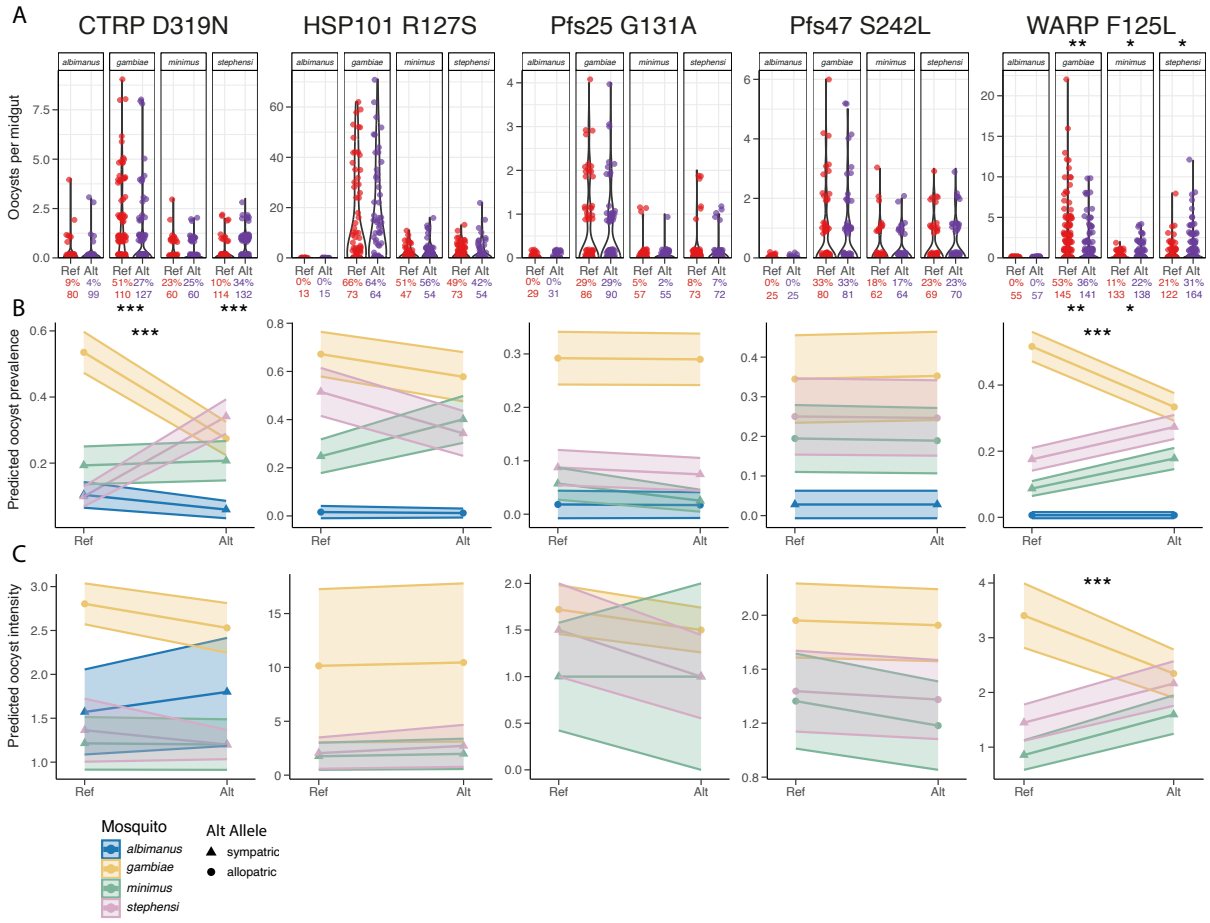

Figure 2. Infection of different *Anopheles* species by five *P. falciparum* allelic-replacement lines at oocyst stage. (A) The number of oocysts for each mosquito midgut counted 10 days post infectious feed across experimental replicates. For each candidate gene, the violin plot shows all oocyst data for each species (*An. albimanus*, *An. gambiae*, *An. minimus*, *An. stephensi*) and each parasite allele (reference (Ref) vs alternative (Alt)). The proportion of infected mosquitoes and total number of midguts assessed is displayed below each violin. Stars represent significant differences between the reference and alternative allele within an individual mosquito species based on post-hoc pairwise comparisons from oocyst prevalence models (stars below) or oocyst intensity models (stars above). (B) The estimated marginal means from the oocyst prevalence model (logistic regression) for each gene is plotted by species and allele. Stars represent a significant Mosquito x Allele interaction for that locus. (C) The estimated marginal means from the oocyst intensity model for each gene is plotted by species and allele. Stars represent a significant Mosquito x Allele interaction. Points in B and C are shaped by whether the alternative allele is sympatric (triangle) or allopatric (circle) for that mosquito species. Shaded regions around each reaction norm represent the standard error. *** $p < 0.001$, ** $p < 0.01$ * < 0.05.

CTRP D319N and WARP F125L both fall in von Willebrand factor A domains

Protein structure analysis of WARP and CTRP revealed that CTRP contains six von Willebrand factor A (vWFA) domains, whereas WARP has a single vWFA domain (**Figure 3A, 4D**)(28, 30, 31, 35). Both CTRP D319N and WARP F125L substitutions occur within these domains. The vWFA domains of CTRP are essential for gliding motility and onward transmission through the mosquito (36). By homology to integrin-like adhesion modules, these domains are predicted to mediate attachment to host surfaces through metal ion–dependent adhesion sites (MIDAS) (37, 38).

The D319N substitution in CTRP replaces a negatively charged aspartic acid with an uncharged asparagine (**Figure 3B, 4C**). Aspartic acid residues within vWFA domains may coordinate bound cations at or near the MIDAS motif, which could stabilize ligand binding and domain conformation (39, 40). Loss of this negative charge may weaken cation coordination and alter the structural dynamics of the vWFA domain, potentially modifying adhesion to mosquito midgut epithelial ligands or affecting downstream processes required for ookinete motility and invasion (26, 31, 36).

Similarly, the WARP F125L substitution replaces phenylalanine with leucine, removing an aromatic side chain (**Figure 3E, 3F**). Aromatic residues often contribute to hydrophobic packing within vWFA cores and π-interactions that stabilize protein–protein interfaces (41, 42). The loss of this aromatic interaction may subtly affect domain folding or reduce affinity for midgut-associated receptors such as basal lamina components and mosquito-derived lectins (30).

Given that vWFA domains mediate critical adhesion processes, these substitutions are likely to modulate protein–protein or protein–matrix binding in the midgut. We hypothesize that such alterations could underlie the observed allele-specific differences

in transmission efficiency by tuning parasite–vector compatibility at the molecular interface of midgut invasion.

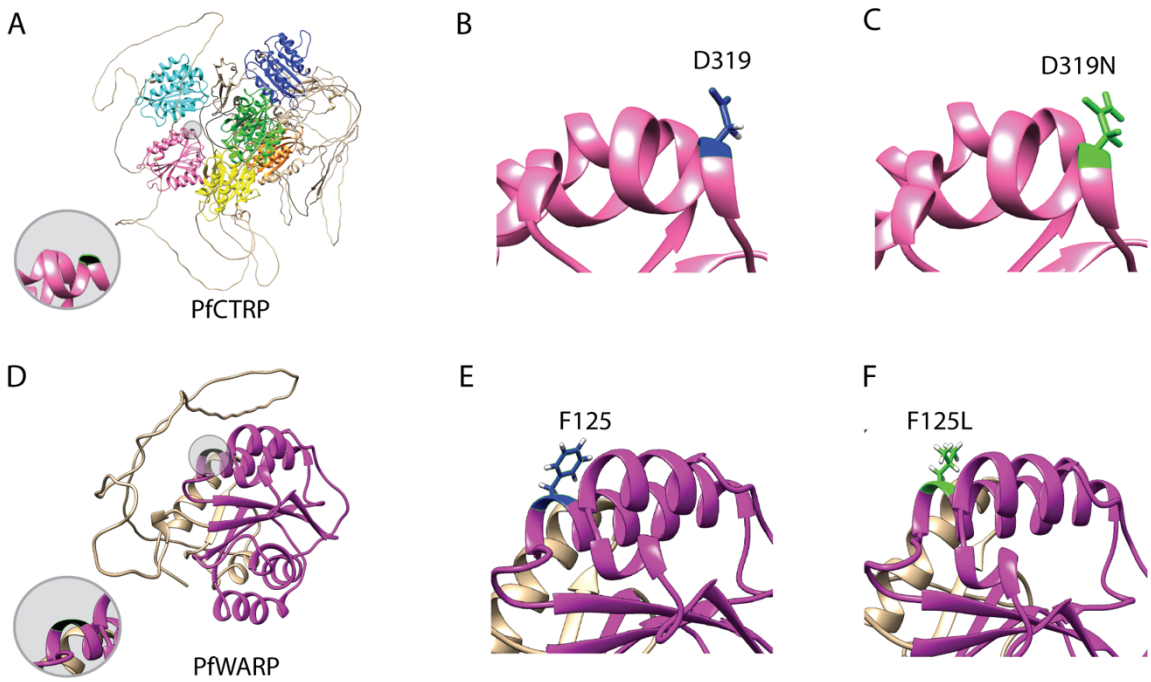

Figure 3. CTRP D319N and WARP F125L are in von Willebrand factor A (vWFA) domains. The protein structure of CTRP (A) and WARP (D) predicted by UCSF Chimera software (43). Distinct vWFA domains of CTRP shown in six colors (A), single vWFA domain for WARP highlighted in purple (D). The D319N substitution in CTRP (C) reduces negative charge on this residue (negative charge from aspartic acid (B) replaced with uncharged asparagine (C)). The F125L substitution in WARP removes an aromatic side chain (phenylalanine (E) replaced by leucine (F)).

Discussion

Understanding how spatial variation and local adaptation shape *P. falciparum* genetic diversity and transmission dynamics is fundamental to public health because treatment or intervention in one location may be ineffective in another. Here we have demonstrated how natural allelic variation in *P. falciparum* transmission-stage genes can produce vector-specific effects on parasite infectivity. By combining population genomic data with functional assays across multiple geographically diverse *Anopheles* species, we show that two variants, CTRP D319N and WARP F125L, modulate infection outcomes in a manner consistent with adaptation to sympatric vector hosts. These results provide direct experimental support for the long-standing hypothesis that *P. falciparum* is adapting to regional vector community composition. Additionally, our results challenge the lock-and-key paradigm that compatibility is determined by a single *P. falciparum*

locus (Pfs47), but instead indicate that this is a complex, polygenic trait where multiple molecular interactions determine vector specialisation.

CTRP and WARP encode micronemal proteins required for ookinete motility and midgut invasion and both contain vWFA domains that are likely involved in adhesion to the midgut epithelium and contained the polymorphisms of interest. The specific amino acid changes within the vWFA domains altered the charge of the residue for CTRP and removed an aromatic ring for WARP. These changes could potentially modify the strength or the specificity of binding to mosquito midgut ligands such as laminins, lectins or glycans (30, 44, 45). Future work to identify the binding partners of *P. falciparum* WARP and CTRP, and how the strength and specificity of these interactions change with different parasite alleles and sympatric vs allopatric vector midgut binding partners, could help elucidate the underlying molecular mechanisms driving the mosquito-by-allele phenotypes we observed.

We did not observe mosquito-by-allele interactions for Pfs47, HSP101, or Pfs25. We cannot rule out that these loci are also regionally adapted to different vectors, given the limited number of mosquito species tested and the fact that only one parasite genetic background was tested. Additionally, for Pfs47, previous work has shown that the full haplotype with four amino acid changes is needed to match the regional vector (*An. gambiae* vs *An. albimanus*) (46). However, the S242L mutation in an NF54 parasite background has previously been shown to result in immune detection and parasite melanisation by the *An. gambiae* L35 resistant strain (46). This difference from our results is likely explained by the use of different *An. gambiae* strains, as we do not regularly observe melanisation in our *An. gambiae* Kisumu colony. For all loci tested, future investigations could employ a full haplotype replacement strategy, as we only edited a single amino acid substitution — the variant with the highest *Fst* that contributed to the global differentiation score (13). Several additional non-synonymous SNPs showed similar fixation patterns within each gene (**Figure S8**), and together these may enhance compatibility between parasite and vector (34). Additionally, future work could test multiple genetic backgrounds and combine haplotypes from different genes to understand additive and epistatic effects on vector compatibility. *P. falciparum* lineages from these distinct geographic areas are genomically very different, which is likely due to very different selective pressures beyond vector species including transmission intensity and drug-pressure (3, 47, 48). Understanding how these selective pressures together shape the parasite genome and the spread of distinct lineages, including drug-resistant *P. falciparum*, is key to understanding the epidemiology of the disease.

The patterns of infection observed at the oocyst stage were not fully mirrored in the sporozoite data (**Figure S7**). Surprisingly, sporozoite prevalences were generally higher than oocyst prevalences, and sporozoites were detected in *An. albimanus* infected with the WARP-F125 parasite, but oocysts were not. These differences may reflect variation in

assay sensitivity or specificity, and recent work has seen the same discrepancy (higher sporozoite prevalence compared to oocysts) using the same qPCR primers, and suggests there is a higher false-positive rate in the qPCR assay (49). Nonetheless, the differences observed at the sporozoite stage between reference and alternative alleles in CTRP and WARP were consistent with an increase in transmissibility for sympatric combinations relative to allopatric ones.

Together, these findings demonstrate that naturally occurring amino acid substitutions in key transmission-stage proteins can alter *P. falciparum* infectivity in a vector-dependent manner. The reciprocal effects of CTRP D319N and WARP F125L across mosquito species provide experimental evidence that parasite–vector compatibility is an evolving, polygenic trait shaped by spatially variable ecological interactions. Such adaptation has important implications for the geographic structuring of malaria transmission and for the design of transmission-blocking interventions. Accounting for natural allelic variation and vector diversity will be essential to developing control strategies that remain effective across the heterogeneous parasite and mosquito populations that sustain global malaria transmission.

Materials and Methods

Generation of Transgenic Parasites

The pBgC plasmids (50, 51) were constructed by sequentially cloning the gRNA oligo and donor template. The plasmid was first digested with BsaI, then the sgRNA oligo was inserted by In-Fusion cloning (Takara Bio), following the manufacturer's protocol. The construct was transformed into XL-10 Gold competent cells, and colonies containing the sgRNA insert were selected. A donor template (300–600 bp) containing the desired mutation along with silent mutations was synthesized and cloned into the HindIII and EcoRI sites of the plasmid using In-Fusion. After transformation, four colonies were selected for sequencing to confirm insertion of the correct donor template.

P. falciparum 3D7 parasites were transfected with pBgC plasmids containing the desired non-synonymous mutation or silent control for each gene (**Table S2**). Synchronized ring-stage parasites (5–10% parasitemia) were transfected with 80–100 µg of plasmid. Transfected parasites were selected with 2.5 µg/ml Blasticidin S for 8 days, then returned to normal 1640 medium supplemented with 5% human serum, 25 mM HEPES, 50 uM hypoxanthine and 0.2% sodium bicarbonate. Cultures were incubated at 37°C in a gas mixture of 1% O_2, 3% CO_2, and 96% N_2. Once parasites recrudesced, genomic DNA was extracted, amplified using the primers in **Table S3** and sequenced to confirm genome editing. Limiting dilution cloning was done to select clones that had the desired mutation. Bulk cultures were diluted to 0.5% prior to serial dilution in 96-well plate. After two weeks, 50 µl from each well was stained with 1:5000 SYBR Green in PBS to detect

parasites by flow cytometery (MACSQuant). Clonal lines were expanded and sequenced to confirm successful editing.

Gametocyte Induction

Gametocytes were induced at 0.5–0.7% asexual parasitaemia and 6% hematocrit in 25 cm^2 flasks with 5 ml of culture following established protocols (52). Two flasks were set up for the day 17 gametocytes and three days later two flasks were set up for the day 14 gametocytes. Flasks were incubated at 37°C and media was changed daily. The media was increased to 1.5x when parasitaemia was 3-5%.

Mosquito rearing

An. albimanus PANAMA, *An. gambiae* KISUMU, *An. minimus* MINIMUS1, and *An. stephensi* SDA500 were reared for transmission experiments at 27 degrees C, 70% relative humidity and 12 hour by 12 hour light/dark cycle conditions. Larvae were fed on ground Tetramin and Tetra Pond pellets. Adults were maintained on a 5% D-glucose solution with 0.05% para-amino benzoic acid (PABA) via cotton pads placed on top of cages. *An. albimanus* were maintained on 10% glucose solution with 0.05% PABA.

Standard Membrane Feeding Assay and Exflagellation

The Standard Membrane Feeding Assay (SMFA) (53, 54) was performed by pooling gametocyte cultures from day 14 and day 17 flasks. The cultures were centrifuged at 500 × g for 7 minutes at 37 °C, then resuspended with an equal volume of human serum and diluted to 1% gametocytemia with 40% hematocrit. Blood meals were loaded in glass membrane feeders covered with goldbeater's skin, connected to a 37°C water bath. Mosquito pots that had 4-7 day old female mosquitoes were put under the glass feeder for 10 minutes to feed and then maintained at 27°C with 70% humidity.

Exflagellation was assessed from an aliquot of the prepared blood meal by diluting to 2.5% haematocrit in ookinete medium containing xanthurenic acid (10.4 g RPMI powder, 2 g of $NaHCO_3$, 50 mg of hypoxanthine in 1 L of dH_2O and 1 ml of 100 mM xanthurenic acid) (55). 10 ul of the dilution was added to a hemocytometer (Neubauer Improved C-Chip). After 10 minutes, exflagellation events were counted under 40X magnification for 10 minutes to estimate the number of events per µl at 1% gametocytemia.

Oocyst and Sporozoite Quantification

Ten days post feeding, mosquitoes were killed with 70% ethanol. Midguts were dissected in PBS and stained with 1% mercurochrome (25 µl per sample). Midguts were examined at 20X magnification on an Echo revolve microscope. The number oocysts per midgut was counted and oocyst size was measured.

Mosquitoes were frozen at day 17 post infectious feed at -20 °C for sporozoite measurement using qPCR. DNA extraction was performed on dissected mosquito heads and top of thorax using the DNeasy Blood & Tissue Kit (Qiagen). Quantitative PCR was performed by mixing 1X SYBR Green Master Mix (Applied Biosystems) and 600 nM of each primer targeting the *PfCOX* gene (forward: 5'-TTACATCAGGAATGTTATTGC-3', reverse: 5'-ATATTGGATCTCCTGCAAAT-3', total reaction 15 ul) (56). Quantitative PCR was run on a Viia 7 system with the following cycling conditions: denaturation: 98°C and annealing: 56°C. Melting curve analysis was used to confirm product specificity. Standard curves were generated using known numbers of ring-stage parasites; 100000, 10000, 1000, 100, 10 and 1.

Statistical Analysis

Statistical models for oocyst and sporozoite prevalence and intensity, as well as oocyst size, had the following core structure for each gene:

Response variable ~ Allele + Mosquito + Allele x Mosquito

Where Allele represents the reference or alternative allele for that locus and Mosquito represents the vector species in which the oocysts or sporozoites were measured (*An. albimanus*, *An. gambiae*, *An. minimus*, or *An. stephensi*). For oocyst prevalence, logistic regression (family = binomial(link = "logit")) was fitted with a biased-reduced generalized linear model using the brglm2 package in R with Replicate as a fixed effect in the model (57). This allowed for analysis of mosquito-allele combinations that had no oocysts detected (58). For oocyst intensity, non-zero samples were analysed using a generalized linear mixed model with a negative binomial distribution (family = nbinom2) using the glmmTMB package in R (59). Replicate and Clone were considered as random effects for CTRP and WARP where two clones were tested, and Replicate alone was considered a random effect for HSP101, Pfs25, and Pfs47. For sporozoite prevalence, glmmTMB was used for logistic regression again with Replicate and Clone as random effects. For sporozoite intensity, glmmTMB was used to fit a generalised linear mixed model with a Gaussian distribution with the same random factors as the sporozoite prevalence model. For oocyst size models, a linear mixed effects model was fitted using the lme4 package in R with Midgut from which the oocyst was measured as a random effect (60). Replicate was considered a fixed factor for CTRP, Pfs25, Pfs47 and WARP, but not for HSP101 as oocyst size was only measured for one of the three experimental replicates. For oocysts and sporozoite models, estimated marginal means and pairwise contrasts within species were calculated using the emmeans package in R (61).

Vector distribution mapping

Rasters downloaded from the Malaria Atlas Project (2, 33) are continuous probability of occurrence surfaces with values from 0 to 1. To produce a clear distribution footprint,

surfaces were converted to a presence polygon using a fixed threshold of 0.5. Shaded areas indicate the probability of presence is higher than 50%.

Protein structure visualisation

Protein structures were obtained from the AlphaFold database: AF-Q8IAL0-F1 (WARP) and AF-O97267-F1 (CTRP). Domain visualization, amino acid charge analysis, and mutation modelling were performed using UCSF Chimera version 1.18.

Acknowledgements

We thank Luca Nelli for creation of the map in Figure 1A. We thank Franzana Khaliq, Andy Waters, and Katarzyna Modrzynska for production of the *An. stephensi* mosquitoes. We thank Francesco Baldini, Arthur Talman, and Adam Dobson for helpful feedback on earlier versions of this manuscript. This work was funded by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant 220185/Z/20/Z) and a Lord Kelvin Adam Smith Leadership Fellowship awarded to VMH.

References

1. *World malaria report 2024: addressing inequity in the global malaria response* (Geneva: World Health Organization, 2024).

2. M. E. Sinka, *et al.*, A global map of dominant malaria vectors. *Parasit. Vectors* **5**, 69 (2012).

3. M. J. Mackinnon, K. Marsh, The selection landscape of malaria parasites. *Science* **328**, 866–871 (2010).

4. L. Sollelis, V. M. Howick, M. Marti, Revisiting the determinants of malaria transmission. *Trends Parasitol.* **40**, 302–312 (2024).

5. S. Bennink, M. J. Kiesow, G. Pradel, The development of malaria parasites in the mosquito midgut. *Cell. Microbiol.* **18**, 905–918 (2016).

6. R. C. Smith, J. Vega-Rodríguez, M. Jacobs-Lorena, The Plasmodium bottleneck: malaria parasite losses in the mosquito vector. *Mem. Inst. Oswaldo Cruz* **109**, 644–661 (2014).

7. A. Molina-Cruz, *et al.*, Some strains of Plasmodium falciparum, a human malaria parasite, evade the complement-like system of Anopheles gambiae mosquitoes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E1957-62 (2012).

8. L. A. Baton, L. C. Ranford-Cartwright, Ookinete destruction within the mosquito midgut lumen explains Anopheles albimanus refractoriness to Plasmodium falciparum (3D7A) oocyst infection. *Int. J. Parasitol.* **42**, 249–258 (2012).

9. A. Molina-Cruz, *et al.*, Plasmodium evasion of mosquito immunity and global malaria transmission: The lock-and-key theory. *Proc. Natl. Acad. Sci. U. S. A.* **112** (2015).

10. T. G. Anthony, S. D. Polley, A. P. Vogler, D. J. Conway, Evidence of non-neutral polymorphism in Plasmodium falciparum gamete surface protein genes Pfs47 and Pfs48/45. *Mol. Biochem. Parasitol.* **156**, 117–123 (2007).

11. A. Molina-Cruz, *et al.*, The human malaria parasite Pfs47 gene mediates evasion of the mosquito immune system. *Science* **340**, 984–987 (2013).

12. A. Molina-Cruz, *et al.*, Role of *Pfs47* in the dispersal of ancestral *Plasmodium falciparum* malaria through adaptation to different anopheline vectors. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2213626120 (2023).

13. MalariaGEN, *et al.*, An open dataset of Plasmodium falciparum genome variation in 7,000 worldwide samples. *Wellcome Open Res* **6**, 42 (2021).

14. C. V. Ukegbu, *et al.*, PIMMS43 is required for malaria parasite immune evasion and sporogonic development in the mosquito vector. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7363–7373 (2020).

15. E. Real, *et al.*, A single-cell atlas of Plasmodium falciparum transmission through the mosquito. *Nat. Commun.* **12**, 3196 (2021).

16. A. J. Reid, *et al.*, Single-cell RNA-seq reveals hidden transcriptional variation in malaria parasites. *Elife* **7**, e33105 (2018).

17. A. Cohuet, *et al.*, Anopheles and Plasmodium: from laboratory models to natural systems in the field. *EMBO Rep.* **7**, 1285–1289 (2006).

18. Y. Keleta, J. Ramelow, L. Cui, J. Li, Molecular interactions between parasite and mosquito during midgut invasion as targets to block malaria transmission. *NPJ Vaccines* **6**, 140 (2021).

19. A. Molina-Cruz, C. Barillas-Mury, Pfs47 as a malaria transmission-blocking vaccine target. *Am. J. Trop. Med. Hyg.* **107**, 27–31 (2022).

20. S. A. Healy, *et al.*, Pfs230 yields higher malaria transmission-blocking vaccine activity than Pfs25 in humans but not mice. *J. Clin. Invest.* **131** (2021).

21. R. M. Jones, *et al.*, A plant-produced Pfs25 VLP malaria vaccine candidate induces persistent transmission blocking antibodies against Plasmodium falciparum in immunized mice. *PLoS One* **8**, e79538 (2013).

22. A. N. Vermeulen, *et al.*, Sequential expression of antigens on sexual stages of Plasmodium falciparum accessible to transmission-blocking antibodies in the mosquito. *J. Exp. Med.* **162**, 1460–1476 (1985).

23. A. M. Tomas, *et al.*, P25 and P28 proteins of the malaria ookinete surface have multiple and partially redundant functions. *EMBO J.* **20**, 3975–3983 (2001).

24. S. Eksi, *et al.*, Malaria transmission-blocking antigen, Pfs230, mediates human red blood cell binding to exflagellating male parasites and oocyst production. *Mol. Microbiol.* **61**, 991–998 (2006).

25. B. C. L. van Schaijk, *et al.*, Pfs47, paralog of the male fertility factor Pfs48/45, is a female specific surface protein in Plasmodium falciparum. *Mol. Biochem. Parasitol.* **149**, 216–222 (2006).

26. J. T. Dessens, *et al.*, CTRP is essential for mosquito infection by malaria ookinetes. *EMBO J.* **18**, 6221–6227 (1999).

27. J. T. Dessens, *et al.*, SOAP, a novel malaria ookinete protein involved in mosquito midgut invasion and oocyst development. *Mol. Microbiol.* **49**, 319–329 (2003).

28. M. Yuda, K. Yano, T. Tsuboi, M. Torii, Y. Chinzei, von Willebrand Factor A domain-related protein, a novel microneme protein of the malaria ookinete highly conserved throughout Plasmodium parasites. *Mol. Biochem. Parasitol.* **116**, 65–72 (2001).

29. M. Yuda, H. Sakaida, Y. Chinzei, Targeted disruption of the plasmodium berghei CTRP gene reveals its essential role in malaria infection of the vector mosquito. *J. Exp. Med.* **190**, 1711–1716 (1999).

30. F. Li, *et al.*, Plasmodium ookinete-secreted proteins secreted through a common micronemal pathway are targets of blocking malaria transmission. *J. Biol. Chem.* **279**, 26635–26644 (2004).

31. T. J. Templeton, D. C. Kaslow, D. A. Fidock, Developmental arrest of the human malaria parasite Plasmodium falciparum within the mosquito midgut via CTRP gene disruption. *Mol. Microbiol.* **36**, 1–9 (2000).

32. J. R. Beck, V. Muralidharan, A. Oksman, D. E. Goldberg, PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes. *Nature* **511**, 592–595 (2014).

33. S. I. Hay, R. W. Snow, The malaria Atlas Project: developing global maps of malaria risk. *PLoS Med.* **3**, e473 (2006).

34. Malaria Genomic Epidemiology Network (MalariaGEN), *et al.*, Pf8: an open dataset of *Plasmodium falciparum* genome variation in 33,325 worldwide samples. *Wellcome Open Res.* **10**, 325 (2025).

35. F. Trottein, T. Triglia, A. F. Cowman, Molecular cloning of a gene from Plasmodium falciparum that codes for a protein sharing motifs found in adhesive molecules from mammals and plasmodia. *Mol. Biochem. Parasitol.* **74**, 129–141 (1995).

36. C. Ramakrishnan, *et al.*, Vital functions of the malarial ookinete protein, CTRP, reside in the A domains. *Int. J. Parasitol.* **41**, 1029–1039 (2011).

37. C. A. Whittaker, R. O. Hynes, Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. *Mol. Biol. Cell* **13**, 3369–3387 (2002).

38. T. Pihlajamaa, *et al.*, Structure of Plasmodium falciparum TRAP (thrombospondin-related anonymous protein) A domain highlights distinct features in apicomplexan von Willebrand factor A homologues. *Biochem. J.* **450**, 469–476 (2013).

39. J. O. Lee, P. Rieu, M. A. Arnaout, R. Liddington, Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18). *Cell* **80**, 631–638 (1995).

40. R. C. Liddington, M. H. Ginsberg, Integrin activation takes shape. *J. Cell Biol.* **158**, 833–839 (2002).

41. J. Emsley, M. Cruz, R. Handin, R. Liddington, Crystal structure of the von Willebrand Factor A1 domain and implications for the binding of platelet glycoprotein Ib. *J. Biol. Chem.* **273**, 10396–10401 (1998).

42. J. Takagi, B. M. Petre, T. Walz, T. A. Springer, Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. *Cell* **110**, 599–511 (2002).

43. E. F. Pettersen, *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

44. V. Mahairaki, T. Voyatzi, I. Sidén-Kiamos, C. Louis, The Anopheles gambiae gamma1 laminin directly binds the Plasmodium berghei circumsporozoite- and TRAP-related protein (CTRP). *Mol. Biochem. Parasitol.* **140**, 119–121 (2005).

45. R. R. Dinglasan, *et al.*, Plasmodium falciparum ookinetes require mosquito midgut chondroitin sulfate proteoglycans for cell invasion. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 15882–15887 (2007).

46. G. E. Canepa, A. Molina-Cruz, C. Barillas-Mury, Molecular analysis of Pfs47-mediated Plasmodium evasion of mosquito immunity. *PLoS One* **11**, e0168279 (2016).

47. M. Manske, *et al.*, Analysis of Plasmodium falciparum diversity in natural infections by deep sequencing. *Nature* **487**, 375–379 (2012).

48. M. K. Rono, *et al.*, Adaptation of Plasmodium falciparum to its transmission environment. *Nature Ecology & Evolution* **2**, 377–387 (2018).

49. E. Guissou, *et al.*, Intervention reducing malaria parasite load in vector mosquitoes: No impact on Plasmodium falciparum extrinsic incubation period and the survival of Anopheles gambiae. *PLoS Pathog.* **19**, e1011084 (2023).

50. M. Filarsky, *et al.*, GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science* **359**, 1259–1263 (2018).

51. E. Hitz, *et al.*, The catalytic subunit of Plasmodium falciparum casein kinase 2 is essential for gametocytogenesis. *Commun. Biol.* **4**, 336 (2021).

52. R. Carter, L. Ranford-Cartwright, P. Alano, The culture and preparation of gametocytes of Plasmodium falciparum for immunochemical, molecular, and mosquito infectivity studies. *Methods Mol. Biol.* **21**, 67–88 (1993).

53. T. S. Churcher, *et al.*, Measuring the blockade of malaria transmission--an analysis of the Standard Membrane Feeding Assay. *Int. J. Parasitol.* **42**, 1037–1044 (2012).

54. K. Miura, *et al.*, Qualification of Standard Membrane-Feeding Assay with Plasmodium falciparum Malaria and Potential Improvements for Future Assays. *PLoS One* **8**, e57909 (2013).

55. M. J. Delves, *et al.*, Routine in vitro culture of P. falciparum gametocytes to evaluate novel transmission-blocking interventions. *Nat. Protoc.* **11**, 1668–1680 (2016).

56. A. Boissière, *et al.*, Application of a qPCR assay in the investigation of susceptibility to malaria infection of the M and S molecular forms of An. gambiae s.s. in Cameroon. *PLoS One* **8**, e54820 (2013).

57. I. Kosmidis, D. Firth, Jeffreys-prior penalty, finiteness and shrinkage in binomial-response generalized linear models. *Biometrika* **108**, 71–82 (2021).

58. P. C. Ubiaru, L. C. Ranford-Cartwright, When a mean can be meaningless: evaluating mosquito infections with Plasmodium parasites. *Parasitology* 1–9 (2025).

59. M. E. Brooks, *et al.*, glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal* **9**, 378–400 (2017).

60. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using **lme4**. *J. Stat. Softw.* **67**, 1–48 (2015).

61. S. R. Searle, F. M. Speed, G. A. Milliken, Population marginal means in the linear model: An alternative to least squares means. *Am. Stat.* **34**, 216–221 (1980).

62. C. Lee, *et al.*, Pf-HaploAtlas: an interactive web app for spatiotemporal analysis of Plasmodium falciparum genes. *Bioinformatics* **40**, btae673 (2024).

Supplementary Information

Supplementary Figures

Figure S1. Expression of candidate genes in Malaria Cell Atlas data

Figure S2. Genome and Protein alignment for candidate loci

Figure S3. Exflagellation from five allelic-replacement parasite lines and silent controls

Figure S4. Level of 3D7 infection in *An. gambiae* for each experimental replicate

Figure S5. Oocyst size across parasite lines and vector species

Figure S6. Oocyst by clone for WARP and CTRP

Figure S7. Sporozoite data for WARP and CTRP

Figure S8. CTRP and WARP haplotypes

Supplementary Tables

Table S1. List of 60 candidates with ranks and gene expression data

Table S2. Donor template sequence for plasmid construct of mutation and silene control

Table S3. Guide oligonucleotide and primer sequences.

Supplementary Figures

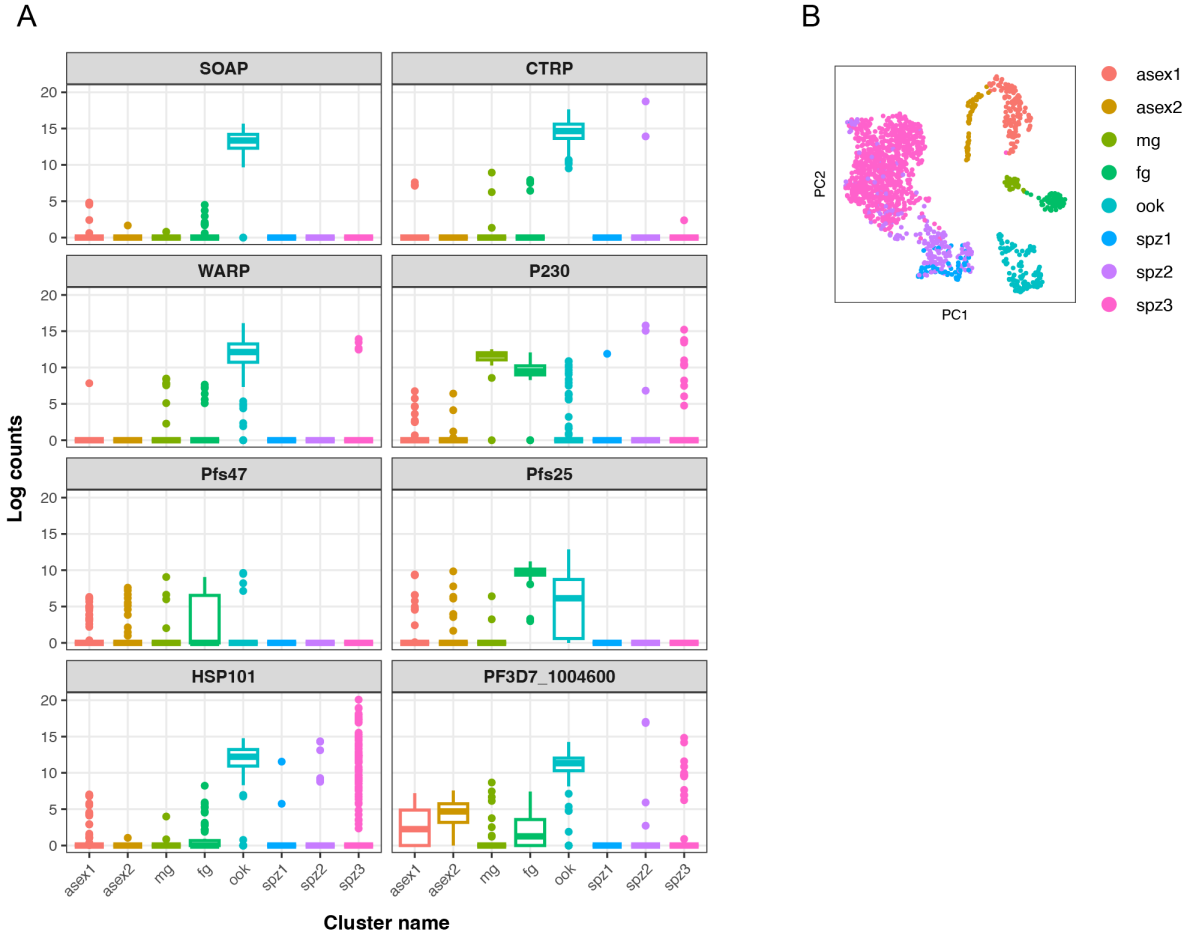

Figure S1. Expression of candidate genes in Malaria Cell Atlas data

Expression of the eight candidate genes across Malaria Cell Altas Smart-Seq2 *Plasmodium falciparum* data from (15, 16). Expression is shown for each gene in each stage (A) based on annotated clusters (B). Candidate genes showed high expression in the mature gametocytes or ookinete stages. Annotated clusters are ordered by developmental progression. Asex = asexual, mg = male gametocyte, fg = female gametocyte, ook=ookinete, and spz = sporozoite. Targeted candidate genes were selected from the intersection of genes in the top 5% of global differentiation scores from MalariaGEN data (13) and genes in the Malaria Cell Atlas data that either fell in gene clusters 7-12 or were more highly expressed in *P. falciparum* ookinetes compared to *P. berghei* (Table S1). Gene clusters 7-12 (g7-g12) are associated with high expression in *P. falciparum* gametocytes or ookinetes (15).

```
WARP_F125L_chr8:102454
Genomic
PF3D7_0801300        GAATCTATGGTCTTTAATTCGAATGTAAGTAAGGATAAGGCTCATATGTCAGTTGTCCTT     420
PfHB3_080006400      GAATCTATGGTCTTTAATTCGAATGTAAGTAAGGATAAGGCTCATATGTCAGTTGTCCTT     420
Pf7G8_080006600      GAATCTATGGTCTTTAATTCGAATGTAAGTAAGGATAAGGCTCATATGTCAGTTGTCCTT     420
PfDd2_080006700      GAATCTATGGTCCTTAATTCGAATGTAAGTAAGGATAAGGCTCATATGTCAGTTGTCCTT     420
                     ************ ************************************************
Protein
PF3D7_0801300.1-p1           ESMVFNSNVSKDKAHMSVVLFAGKQRVVVPFTDESSQDKNMLLEKINSIDEGIATSSDTL    180
Pf7G8_080006600-t41_1-p1     ESMVFNSNVSKDKAHMSVVLFAGKQRVLVPFTDESSQDKNMLLEKINSIDEGIATSSDTL    180
PfHB3_080006400-t41_1-p1     ESMVFNSNVSKDKAHMSVVLFAGKQRVLVPFTDESSQDKNMLLEKINSIDEGIATSSDTL    180
PfDd2_080006700-t41_1-p1     ESMVLNSNVSKDKAHMSVVLFAGKQRVFVPFTDESSQDKNMLLEKINSIDEGIATSSDTL    180
                             ****:**********************.*******************************
_____________
CTRP_D319N_chr3:617426
Genomic
PfHB3_030019000      GAGGATTATTACGATTTAACTTTAATTTTAGATGAGTCACGAAGTATAACTCTCAACAAA     960
PF3D7_0315200        GAGGATTATTACGATTTAACTTTAATTTTAGATGAGTCACGAAGTATAACTCTCGACAAA     960
Pf7G8_030021000      GAGGATTATTACGATTTAACTTTAATTTTAGATGAGTCACGAAGTATAACTCTCAACAAA     960
PfDd2_030020200      GAGGATTATTACGATTTAACTTTAATTTTAGATGAGTCACGAAGTATAACTCTCAACAAA     960
                     ***************************************************** *****
Protein
PfHB3_030019000-t41_1-p1     EDYYDLTLILDESRSITLNKWKKDVVPFAEKVLNNLNIDKDKIHVGIMRFAKSMKTDIGY    360
PF3D7_0315200.1-p1           EDYYDLTLILDESRSITLDKWKKDVVPFAEKVLNNLNIDKDKIHVGIMRFAKSMKTDIGY    360
Pf7G8_030021000-t41_1-p1     EDYYDLTLILDESRSITLNKWKKDVVPFAEKVLNNLNIDKDKIHVGIMRFAKSMKTDIGY    360
PfDd2_030020200-t41_1-p1     EDYYDLTLILDESRSITLNKWKKDVVPFAEKVLNNLNIDKDKIHVGIMRFAKSMKTDIGY    360
                             ******************:*****************************************
_____________
HSP101_R172S_chr11:638298
Genomic (intron included)
Pf7G8_110020400      GATGAAGCCATTAAAGCAATATTAAAAAGCAAATTTGAAAAAACGAAAAAGGATAAAGAT     1221
PfHB3_110019900      GATGAAGCCATTAAAGCAATATTAAAAAGAAAATTTGAAAAAACGAAAAAGGATAAAGAT     1253
PF3D7_1116800        GATGAAGCCATTAAAGCAATATTAAAAAGAAAATTTGAAAAAACGAAAAAGGATAAAGAT     1243
PfDd2_110019800      GATGAAGCCATTAAAGCAATATTAAAAAGCAAATTTGAAAAAACGAAAAAGGATAAAGAT     1236
                     **************************** ******************************
Protein
PF3D7_1116800.1-p1           TLALAEAAANKYKSPKVDVEHLLSGLSNDELVNEIFNEVYLTDEAIKAILKKKFEKTKKD    180
PfHB3_110019900-t41_1-p1     TLALAEAAANKYKSPKVDVEHLLSGLSNDELVNEIFNEVYLTDEAIKAILKKKFEKTKKD    180
Pf7G8_110020400-t41_1-p1     TLALAEAAANKYKSPKVDVEHLLSGLSNDELVNEIFNEVYLTDEAIKAILKSKFEKTKKD    180
PfDd2_110019800-t41_1-p1     TLALAEAAANKYKSPKVDVEHLLSGLSNDELVNEIFNEVYLTDEAIKAILKSKFEKTKKD    180
                             ************************************************** ********
_____________
P47_S242L_chr13:1879470
Genomic
PfDd2_130052600      TTATCAAATAATAATGTGGTTTTAAAAATGCCTTCCATATTTATAAATAATATTAATTGT     780
Pf7G8_130051200      TTATTAAATAATAATGTGGCTTTAAAAATGCCTTCCATATTTATAAATAATATTAATTGT     780
PF3D7_1346800        TTATCAAATAATAATGTGGTTATAAAAATGCCTTCCATATTTATAAATAATATTAATTGT     780
PfHB3_130053100      TTATTAAATAATAATGTGGCTTTAAAAATGCCTTCCATATTTATAAATAATATTAATTGT     780
                     **** ************** * ***************************************
Protein
PfDd2_130052600-t41_1-p1     LSNNNVVLKMPSIFINNINCMLSSQDQNNEKYYINLKADKTKHIDGCDFTKPKGKGIYKN    300
PF3D7_1346800.1-p1           LSNNNVVIKMPSIFINNINCMLSSQDQNNEKNYINLKADKTKHIDGCDFTKPKGKGIYKN    300
Pf7G8_130051200-t41_1-p1     LLNNNVALKMPSIFINNINCMLSSQDQNNEKNYINLKADKTKHIDGCDFTKPKGKGIYKN    300
PfHB3_130053100-t41_1-p1     LLNNNVALKMPSIFINNINCMLSSQDQNNEKNYINLKADKTKHIDGCDFTKPKGKGIYKN    300
                             * ****.:*********************** ****************************
_____________
Pfs25_G131A_chr10:1253677
Genomic
PF3D7_1031000        ATATTAGATACAAGCAATCCTGTTAAAACTGGAGTTTGCTCATGTAATATAGGCAAAGTT     420
Pf7G8_100035300      ATATTAGATACAAGCAATCCTGTTAAAACTGGAGTTTGCTCATGTAATATAGGCAAAGTT     420
PfHB3_100035300      ATATTAGATACAAGCAATCCTGTTAAAACTGGAGTTTGCTCATGTAATATAGGCAAAGTT     420
PfDd2_100036200      ATATTAGATACAAGCAATCCTGTTAAAACTGCAGTTTGCTCATGTAATATAGGCAAAGTT     420
                     ****************************** ****************************
Protein
PF3D7_1031000.1-p1           ILDTSNPVKTGVCSCNIGKVPNVQDQNKCSKDGETKCSLKCLKENETCKAVDGIYKCDCK    180
Pf7G8_100035300-t41_1-p1     ILDTSNPVKTGVCSCNIGKVPNVQDQNKCSKDGETKCSLKCLKENETCKAVDGIYKCDCK    180
PfHB3_100035300-t41_1-p1     ILDTSNPVKTGVCSCNIGKVPNVQDQNKCSKDGETKCSLKCLKENETCKAVDGIYKCDCK    180
PfDd2_100036200-t41_1-p1     ILDTSNPVKTAVCSCNIGKVPNVQDQNKCSKDGETKCSLKCLKENETCKAVDGIYKCDCK    180
                             **********.*************************************************
_________________

P230_N1566T_chr10:375134-375135
Genomic
PfDd2_020011800      TTAAATATAAAATTTGATACATATGTTACTGGCACTGATCAAAATAAATATCTTATGACA     4740
Pf7G8_020013700      TTAAATATAAAATTTGATAATTATGTTACTGGCACTGATCAAAATAAATATCTTATGACA     4737
PF3D7_0209000        TTAAATATAAAATTTGATAATTATGTTACTGGCACTGATCAAAATAAATATCTTATGACA     4737
PfHB3_020013700      TTAAATATAAAATTTGATAATTATGTTACTGGCACTGATCAAAATAAATATCTTATGACA     4725
                     ******************  ***************************************
Protein
PfDd2_020011800-t41_1-p1     LNIKFDTYVTGTDQNKYLMTYMDLHLSHKRNYLKELFHDLGKKKPADTDANPESIIESLS    1620
Pf7G8_020013700-t41_1-p1     LNIKFDNYVTGTDQNKYLMTYMDLHLSHKRNYLKELFHDLGKKKPADTDANPESIIESLS    1619
PF3D7_0209000.1-p1           LNIKFDNYVTGTDQNKYLMTYMDLHLSHKRNYLKELFHDLGKKKPADTDANPESIIESLS    1619
PfHB3_020013700-t41_1-p1     LNIKFDNYVTGTDQNKYLMTYMDLHLSHKRNYLKELFHDLGKKKPADTDANPESIIESLS    1615
                             ******.*****************************************************
```

```
_______________
SOAP_S95R_chr2:150282
Genomic
PF3D7_1404300        TCAATGAAATATAGACATGCCAGTAAGGGATCTAAAGGATATAGTAAAAAAATGACAAAG     300
Pf7G8_140009700      TCAATGAAATATAGACATGCCAGTAAGGGATCTAAAAGATATAGTAAAAAAATGACAAAG     300
PfHB3_140010200      TCAATGAAATATAGACATGCCAGTAAGGGATCTAAAAGATATAGTAAAAAAATGACAAAG     300
PfDd2_140008900      TCAATGAAATATAGACATGCCAGTAAGGGATCTAAAAGATATCGTAAAAAAATGACAAAG     300
                     ********************************** ***** *****************
Protein
PF3D7_1404300.1-p1            KCLPDNFCECECSCKNKTGFSMKYRHASKGSKGYSKKMTKTSSRSRQCNDLPLEKSEQET    120
Pf7G8_140009700-t41_1-p1      KCLPDNFCECECSCKNKTGFSMKYRHASKGSKRYSKKMTKTSSRSRQCNDLPLEKSEQET    120
PfHB3_140010200-t41_1-p1      KCLPDNFCECECSCKNKTGFSMKYRHASKGSKRYSKKMTKTSSRSRQCNDLPLEKSEQET    120
PfDd2_140008900-t41_1-p1      KCLPDNFCECECSCKNKTGFSMKYRHASKGSKRYRKKMTKTSSRSRQCNDLPLEKSEQET    120
                              ******************************** * *************************
_______________
Unknown protein_L94I_chr10:216319
Genomic
PF3D7_1004600        AATATTCATGGAACATTATTTATTATAAGTTTAATTTGTTATTCTCTTGAAAAAATTATT     616
Pf7G8_100009100      AATATTCATGGAACATTATTTATTATAAGTTTAATTTGTTATTCTCTTGAAAAAATTATT     612
PfDd2_100010000      AATATTCATGGAACATTATTTATTATAAGTTTAATTTGTTATTCTATTGAAAAAATTATT     498
PfHB3_100009100      AATATTCATGGAACATTATTTATTATAAGTTTAATTTGTTATTCTCTTGAAAAAATTATT     660
                     ******************************************* **************
Protein
PF3D7_1004600.1-p1            LFICSGLKEDENHGNITTNIHGTLFIISLICYSLEKIIPFSFCSDFLFIFLLSLSCTLPF    120
PfHB3_100009100-t41_1-p1      LFICSGLKEDENHGNITTNIHGTLFIISLICYSLEKIIPFSFCSDFLFIFLLSLSCTLPF    120
Pf7G8_100009100-t41_1-p1      LFICSGLKEDENHGNITTNIHGTLFIISLICYSLEKIIPFSFCSDFLFIFLLSLSCTLPF    120
PfDd2_100010000-t41_1-p1      LFICSGLKEDENHGNITTNIHGTLFIISLICYSIEKIIPFSFCSDFLFIFLLSLSCTLPF    120
                              *********************************:*************************
```

Figure S2. Genome and protein alignments for candidate genes. Non-synonymous SNPs identified in each gene selected for this study. Alternative alleles are highlighted in green within the protein sequence alignment, while blue indicates nucleotide variations responsible for the observed SNPs.

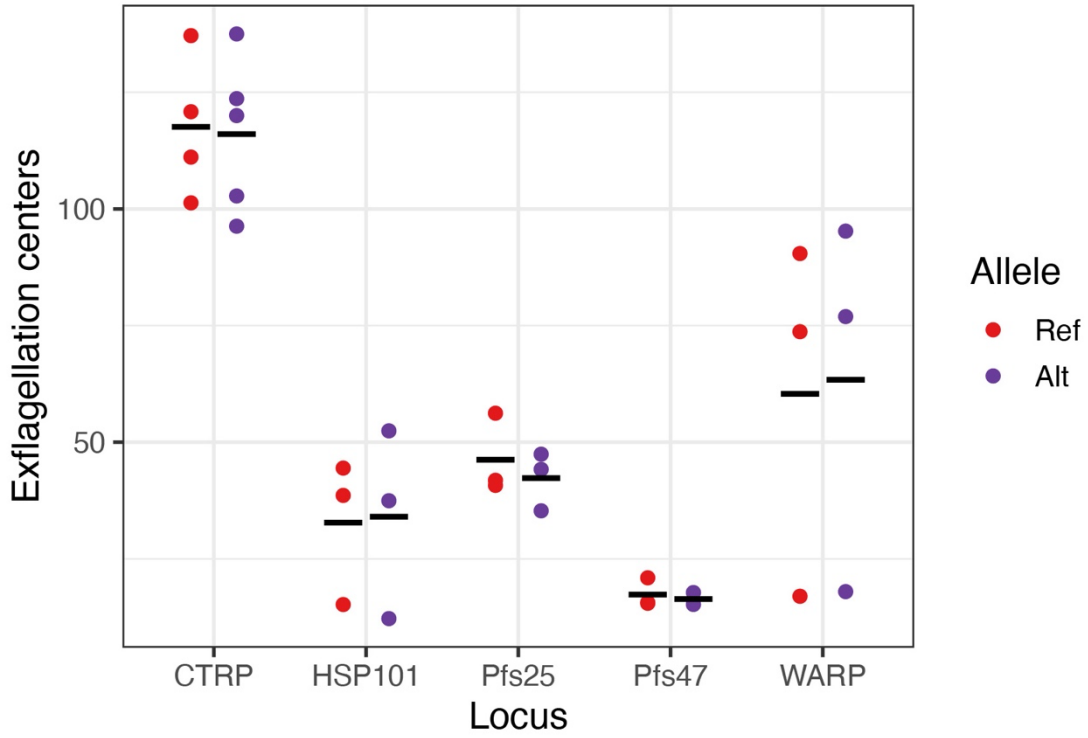

Figure S3 Exflagellation from five allelic-replacement parasite lines and silent controls

Exflagellation was measured for each infectious blood meal. Each point represents a single replicate and the black bar represents the mean. No difference was observed between reference and alternative alleles in any of the five candidate loci (Linear model: Exflagellation centers ~ Allele, $p > 0.05$ for all loci).

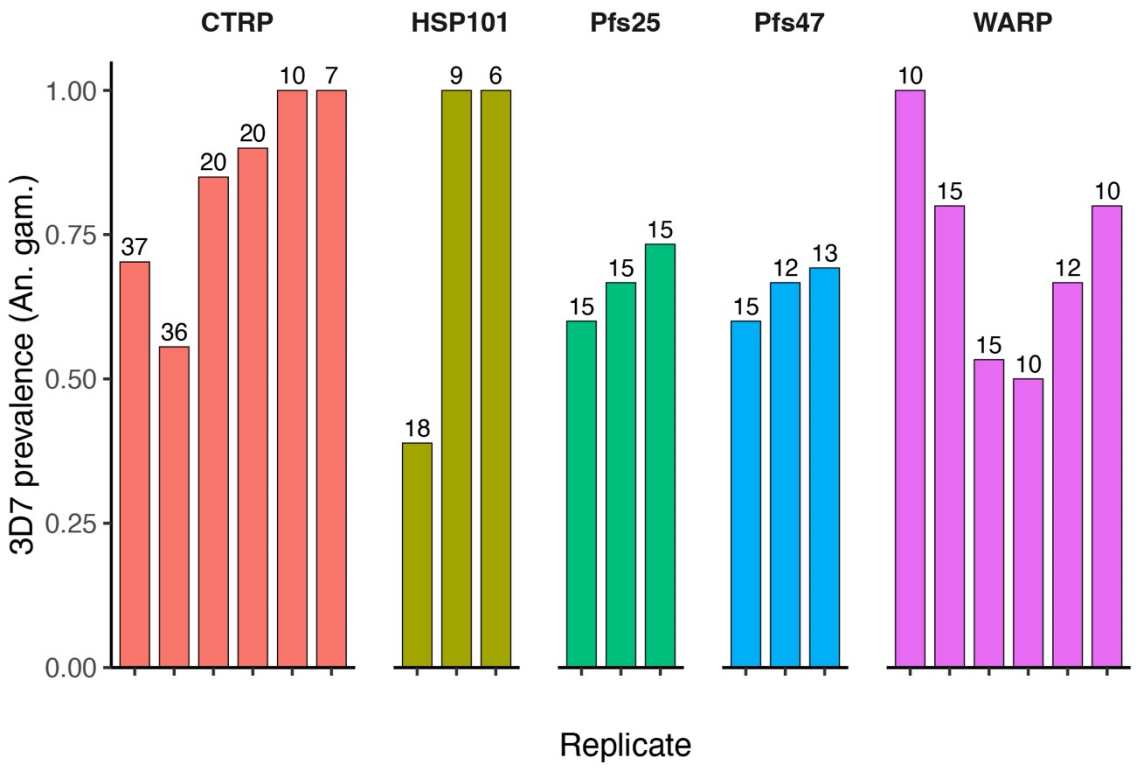

Figure S4. **Oocyst prevalence for wild-type 3D7 in *An. gambiae* in each experimental replicate.** An. gambiae infected with unedited wild-type 3D7 was included as a control in all experimental replicates to confirm that infection was achieved. The number above each bar represents the number of midguts examined.

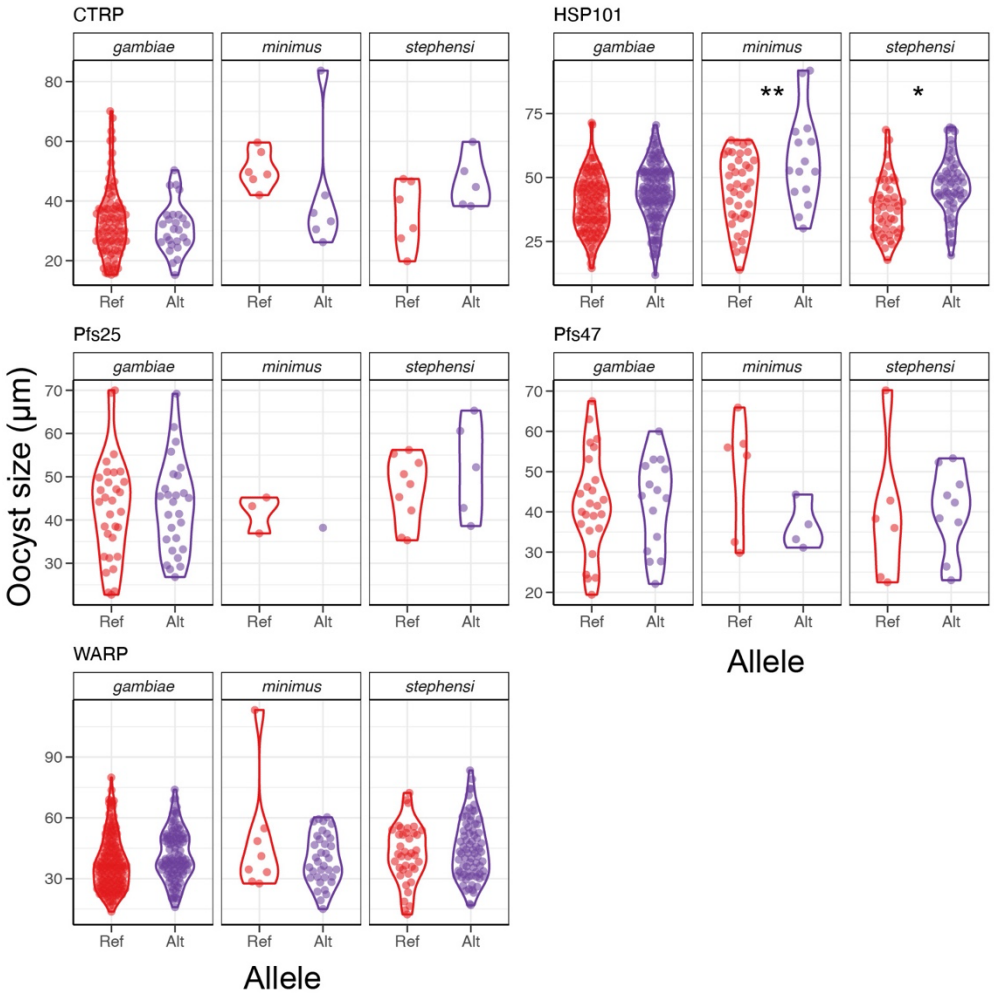

Figure S5. Oocyst diameter across parasite lines and vector species. No significant mosquito-by-allele interaction was found for any locus. However, for HSP101, the alternative allele had larger oocysts for *An. minimus* (adj $p = 0.003$) and *An. stephensi* (adj $p = 0.040$). Only a single replicate was measured for HSP101 compared to two or three replicates for all other loci.

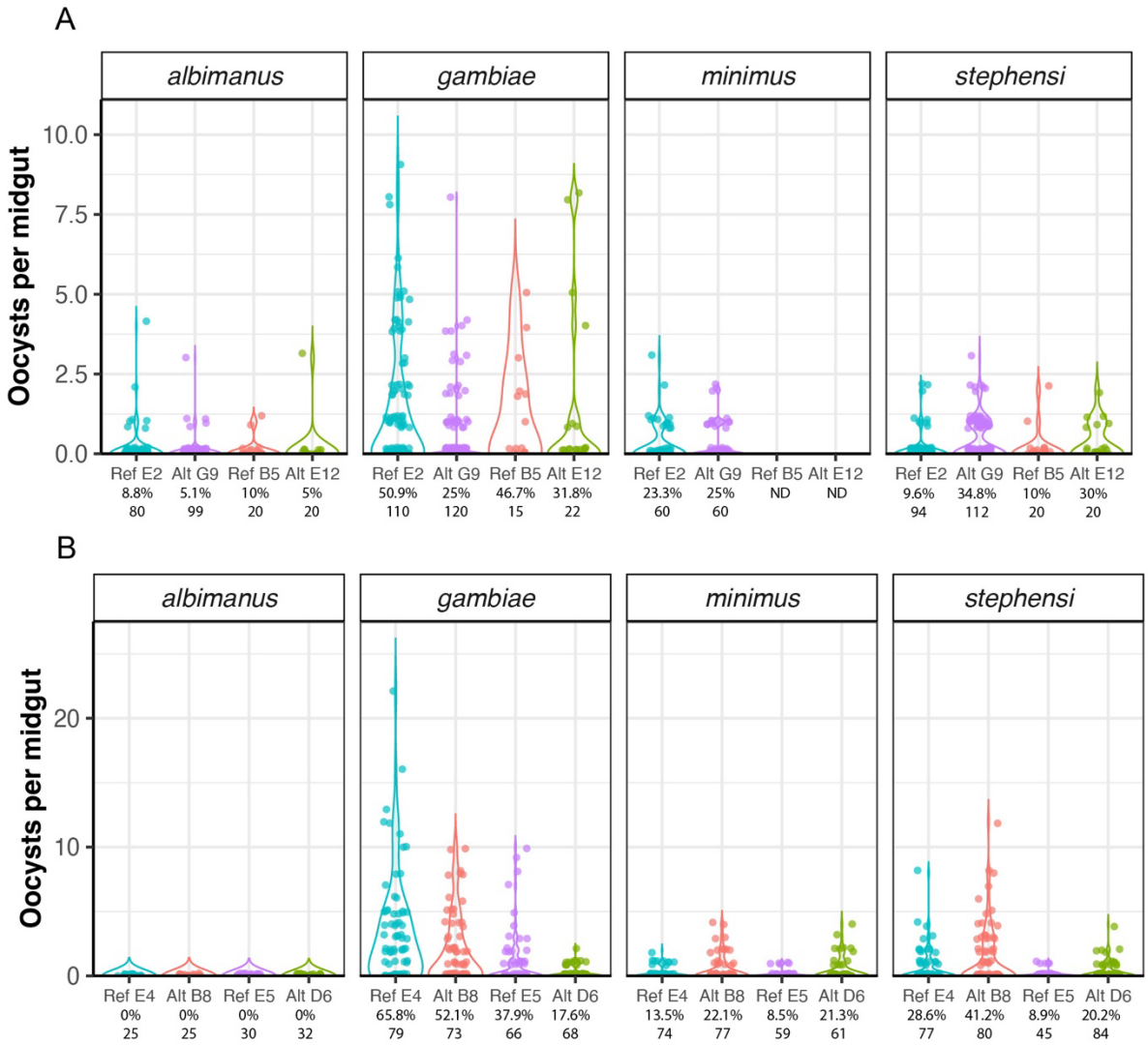

Figure S6. Oocysts per midgut for CTRP and WARP broken down by clone. A second clone was tested for each allele for CTRP (A) and WARP (B) to confirm that vector-specific transmission phenotype was repeatable across clones. For CTRP, Ref E2 and Alt G9 were the primary clones tested and Ref B5 and Alt E12 were the secondary clones tested. For WARP, Ref E4 and Alt B8 were the primary clones tested, whereas Ref E5 and Alt D6 were the secondary clones tested. Each dot represents oocyst counts from a single mosquito midgut.

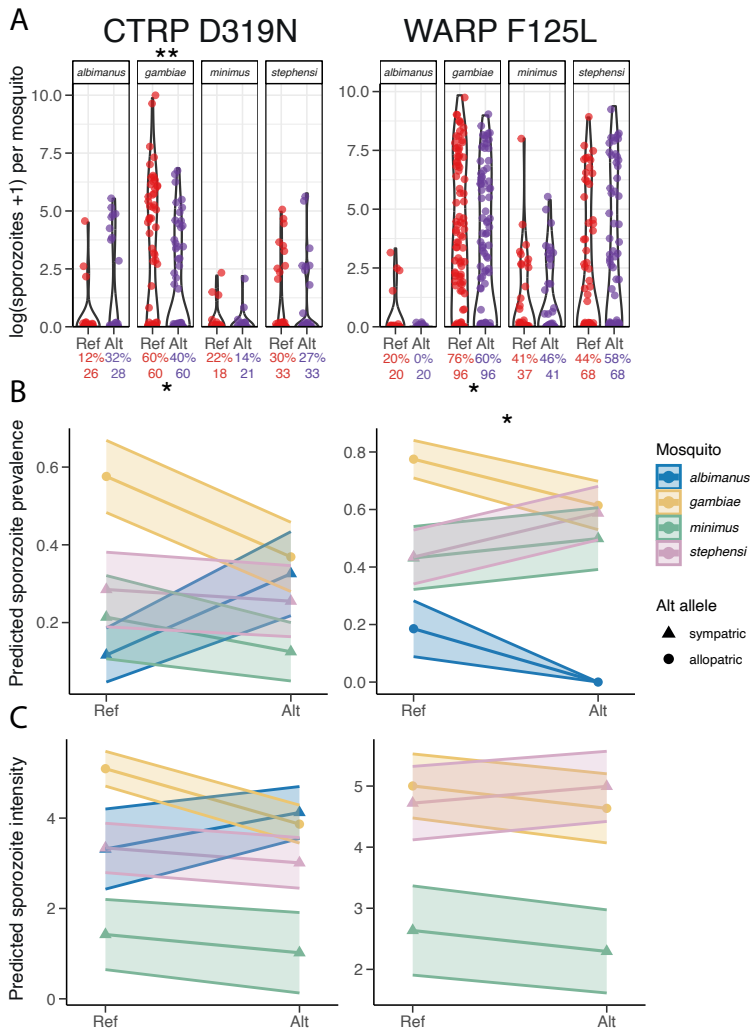

Figure S7. Infection of different *Anopheles* species by CTRP and WARP allelic-replacement lines at sporozoite stage. (A) The log(sporozoites +1) quantified by qPCR for each mosquito head/thorax at 17 days post infectious feed. For each gene, the violin plot shows all sporozoite data for each species (*An. albimanus*, *An. gambiae*, *An. minimus*, *An. stephensi*) and each parasite allele (reference (Ref) vs alternative (Alt)). The proportion of infected mosquitoes and total number of midguts assessed is displayed below each violin. Stars represent significant differences between the reference and alternative allele within an individual mosquito species based on post-hoc pairwise comparisons from sporozoite prevalence models (stars below) or sporozoite intensity models (stars above). The WARP alternative allele showed lower sporozoite prevalence compared to the reference in *An. gambiae* (allopatric, adj p = 0.018). Similarly, CTRP D319N was associated with reduced sporozoite prevalence and intensity in *An. gambiae* (prevalence adj p = 0.027, intensity adj p = 0.002). (B) The estimated marginal means from the sporozoite prevalence model for each gene is plotted by species and allele. We detected a significant mosquito-by-allele interaction for WARP (p = 0.034) (C) The estimated marginal means from the sporozoite intensity model for each gene is plotted by species and allele. Stars represent a significant Mosquito x Allele interaction. Points in B and C are shaped by whether the alternative allele is sympatric (triangle) or allopatric

(circle) for that mosquito species. Shaded regions around each reaction norm represent the standard error. *** $p < 0.001$, ** $p < 0.01$, * < 0.05.

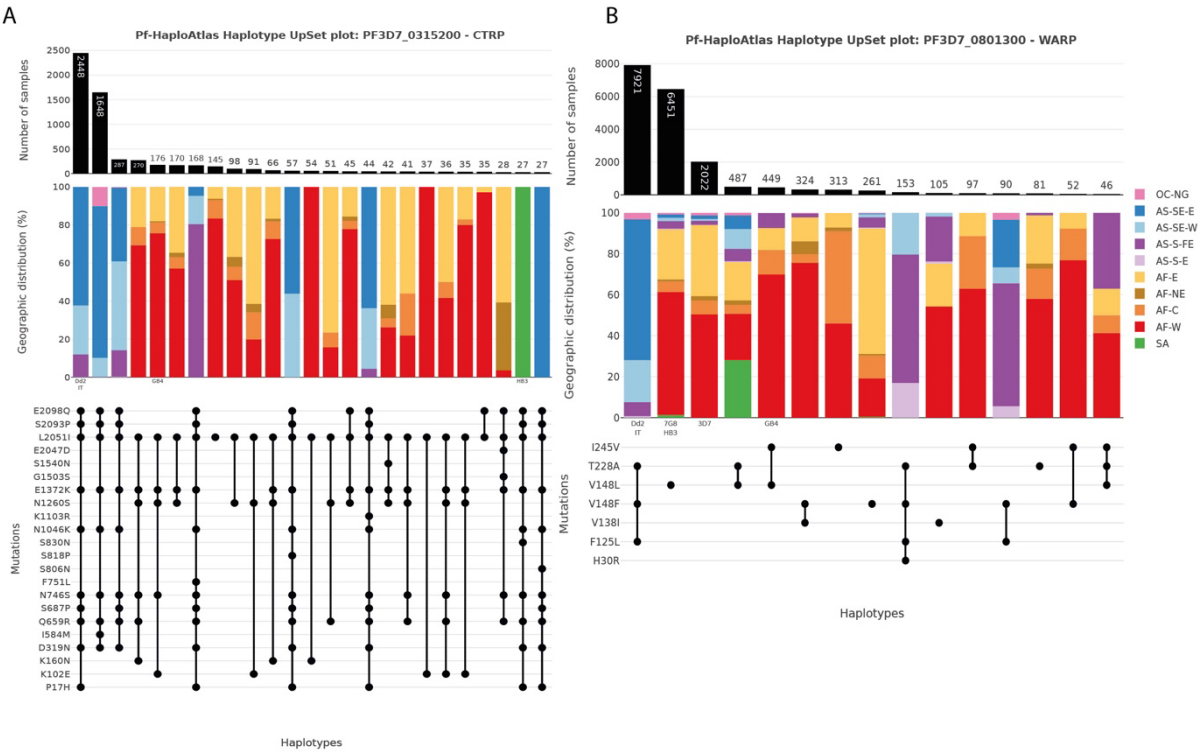

Figure S8 CTRP and WARP haplotypes. Images were created using the MalariaGEN Pf-HaploAtlas (62). See https://pf-haploatlas.streamlit.app/?gene_id=PF3D7_0315200 and https://pf-haploatlas.streamlit.app/?gene_id=PF3D7_0801300 for more information. Legend terms: OC-NG = Oceania,New Guinea; AS-SE-E = Southeast Asia (East); AS-SE-W (Southeast Asia (West)); AS-S-FE = South Asia (Far East); AS-S-E = South Asia (East); AF-E = Africa (East); AF-NE = Africa (North-East); AF-C = Africa (Central); AF-W = Africa (West); SA = South America

Table S1. 60 Candidate genes

See sup file

Table S2. Donor template sequences for mutant and silent control plasmids

See sup file

Table S3. Guide oligonucleotide and primer sequences. The guide oligos were cloned into a plasmid construct to facilitate targeted gene editing, and the primer sets were employed to confirm the successful insertion of the SNP.

Primer name	Sequence	Purpose
Pf1004600_L94I_g1F	TAAGTATATAATATTAAATGGTAACGTACATGATAGTTTTAGAGCTAGAA	gRNA oligo primer
Pf1004600_L94I_g1R	TTCTAGCTCTAAAACTATCATGTACGTTACCATTTAATATTATATACTTA	gRNA oligo primer
Pf1004600_L94I_g2F	TAAGTATATAATATTTATGTTCATAAAACACAAGCGTTTTAGAGCTAGAA	gRNA oligo primer
Pf1004600_L94I_g2R	TTCTAGCTCTAAAACGCTTGTGTTTTATGAACATAAATATTATATACTTA	gRNA oligo primer
PfP230_N1566T_g2F	TAAGTATATAATATTATGTTCTGCATCTACATCCCGTTTTAGAGCTAGAA	gRNA oligo primer
PfP230_N1566T_g2R	TTCTAGCTCTAAAACGGGATGTAGATGCAGAACATAATATTATATACTTA	gRNA oligo primer
PfSOAP_S95R_g1F	TAAGTATATAATATTGAAATATAGACATGCCAGTAGTTTTAGAGCTAGAA	gRNA oligo primer
PfSOAP_S95R_g1R	TTCTAGCTCTAAAACTACTGGCATGTCTATATTTCAATATTATATACTTA	gRNA oligo primer
PfHSP101_R172S_g1F	TAAGTATATAATATTTATTGCTTTAATGGCTTCATGTTTTAGAGCTAGAA	gRNA oligo primer
PfHSP101_R172S_g1R	TTCTAGCTCTAAAACATGAAGCCATTAAAGCAATAAATATTATATACTTA	gRNA oligo primer
PfP47_S242L_g1F	TAAGTATATAATATTGAAGAAACTATTGTAGAATCGTTTTAGAGCTAGAA	gRNA oligo primer
PfP47_S242L_g1R	TTCTAGCTCTAAAACGATTCTACAATAGTTTCTTCAATATTATATACTTA	gRNA oligo primer
PfWARP_F125L_g1F	TAAGTATATAATATTAGAGTCAAGTTTTGTTCAGAGTTTTAGAGCTAGAA	gRNA oligo primer
PfWARP_F125L_g1R	TTCTAGCTCTAAAACTCTGAACAAAACTTGACTCTAATATTATATACTTA	gRNA oligo primer
PfCTRP_D319N_g1F	TAAGTATATAATATTGAAGTATAACTCTCGACAAAGTTTTAGAGCTAGAA	gRNA oligo primer
PfCTRP_D319N_g1R	TTCTAGCTCTAAAACTTTGTCGAGAGTTATACTTCAATATTATATACTTA	gRNA oligo primer
Pfs25_G131A_g1F	TAAGTATATAATATTTAAGCATTTTAATGAGCATTGTTTTAGAGCTAGAA	gRNA oligo primer
Pfs25_G131A_g1R	TTCTAGCTCTAAAACAATGCTCATTAAAATGCTTAAATATTATATACTTA	gRNA oligo primer
For1_out_HSP101	CAAACAAGAGCGGGAGCGgtatg	primer for testing the insertion of SNP in genome
Rev1_out_HSP101	CGGTCATATCTACTGAAGGAGG	primer for testing the insertion of SNP in genome
For1_out_P47	TGATTCCAGCGCATATCATTTAG	primer for testing the insertion of SNP in genome
Rev1_out_P47	GCGATATGTAATTCCATTACTGC	primer for testing the insertion of SNP in genome
For1_out_WARP	CCCGTTTTATCTTTATTATTCG	primer for testing the insertion of SNP in genome
Rev1_out_WARP	GTTCAGTCTTGTTCATACATATTG	primer for testing the insertion of SNP in genome
For1_out_CTRP	GAGATTACGTCACTTATGGTGATG	primer for testing the insertion of SNP in genome
Rev1_out_CTRP	GATCCCGTATTACATATCTTATCCG	primer for testing the insertion of SNP in genome
For2_out_WARP	GTTGTTGTATAATAAGAGAGAG	primer for testing the insertion of SNP in genome
Rev2_out_WARP	CATCTTATGATTTATTCTTATCAC	primer for testing the insertion of SNP in genome
For2_out_HSP101	gCCATTGGATAATAAAAGCAAAATTG	primer for testing the insertion of SNP in genome
Rev2_out_HSP101	CTTCTTTCGAAAGCTGAACAAC	primer for testing the insertion of SNP in genome