

Supplementary Materials

***Plasmodium* Evasion of Mosquito Immunity and Global Malaria Transmission: The Lock and Key Theory**

Alvaro Molina-Cruz , Gaspar E. Canepa, Nitin Kamath, Noelle V. Pavlovic, Jianbing Mu,
Urvashi N. Ramphul, Jose Luis Ramirez and Carolina Barillas-Mury

Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious
Diseases, National Institutes of Health, Rockville, MD 20852.

Correspondence should be addressed to: cbarillas@niaid.nih.gov (CB-M) and amolina-cruz@niaid.nih.gov (AM-C)

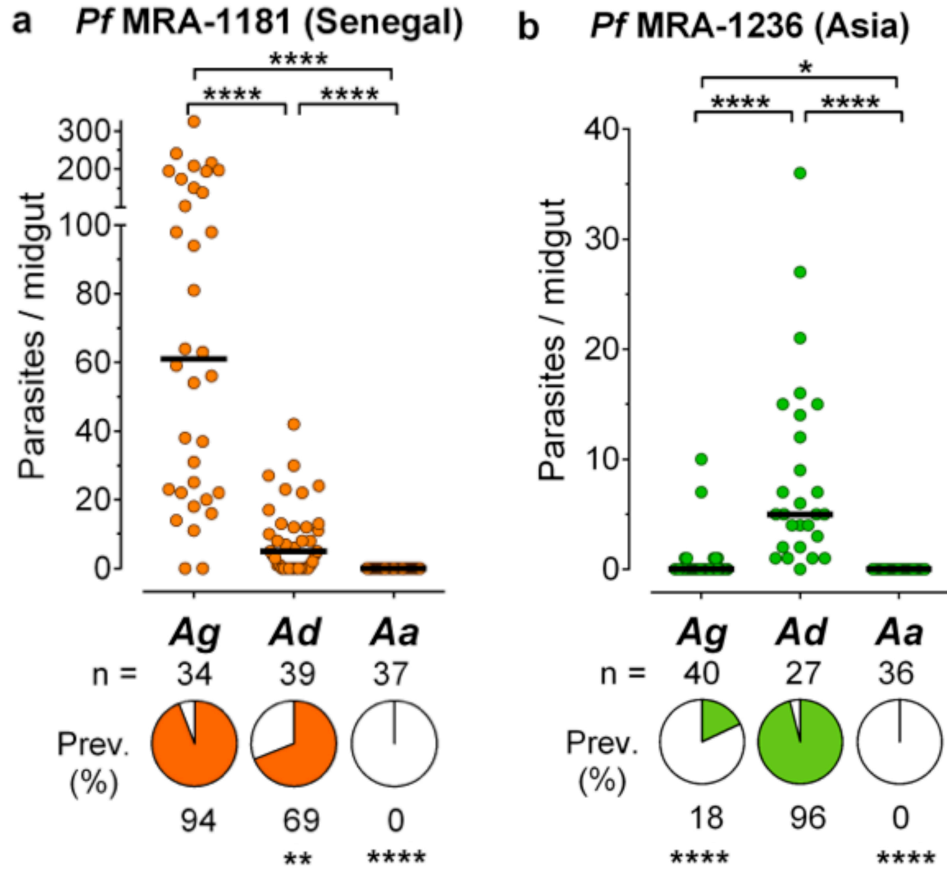


Fig.S1. Compatibility between *Plasmodium falciparum* isolates from different geographic origin and different anopheline mosquito species. **a,b**, Infection of *Anopheles gambiae* (Ag), *Anopheles dirus* (Ad) and *Anopheles albimanus* (Aa) mosquitoes with *P. falciparum* MRA-1181 (African) (**a**) or MRA-1236 (Asian) (**b**) lines. Each circle represents the number of oocysts present in the midgut of individual mosquitoes and the medians are indicated by the black line (n= number of midguts examined). Pie charts represent the prevalence of infection. All phenotypes were confirmed in two independent experiments. Medians were compared using the Mann-Whitney test; prevalence of infection were compared with the χ^2 test, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

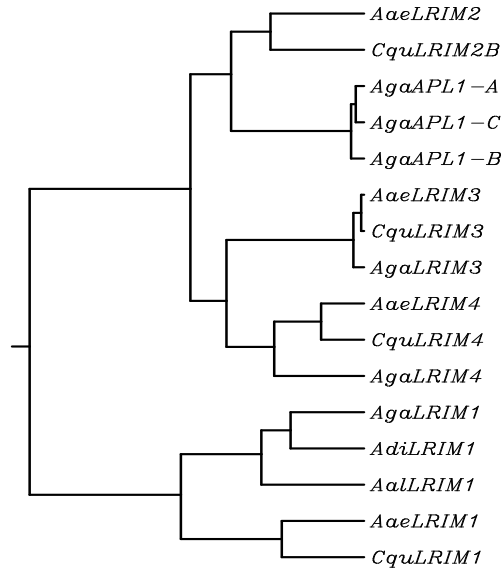


Fig S2. Phylogenetic tree (neighbor-joining) of LRIM family members (Long form) from *Anopheles gambiae* (Aga), *Aedes aegypti* (Aae), *Culex quinquefasciatus* (Cqu) and LRIM1 from *Anopheles dirus* (Adi) and *Anopheles albimanus* (Aal).

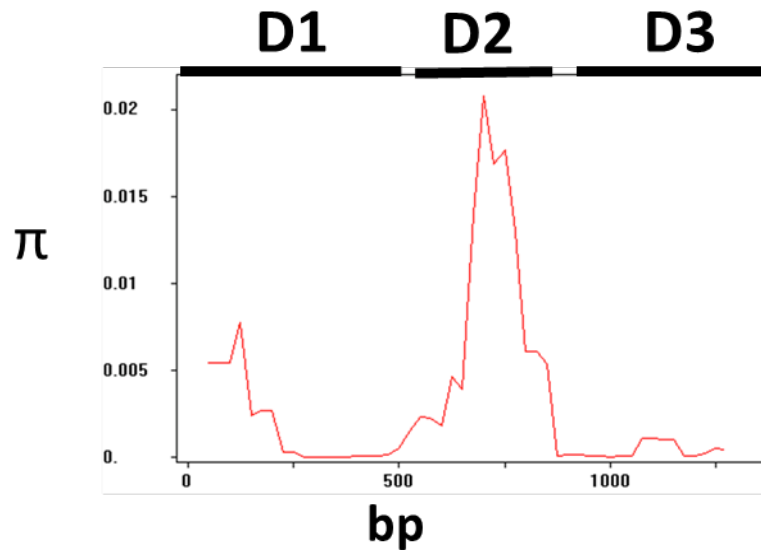


Fig S3. Nucleotide diversity along Pfs47 nucleotide sequence. The nucleotide diversity (π) in a 100bp sliding window along Pfs47 was estimated from 364 gene sequences from around the world. The predicted domain (D) 1, 2 and 3 position is indicated.

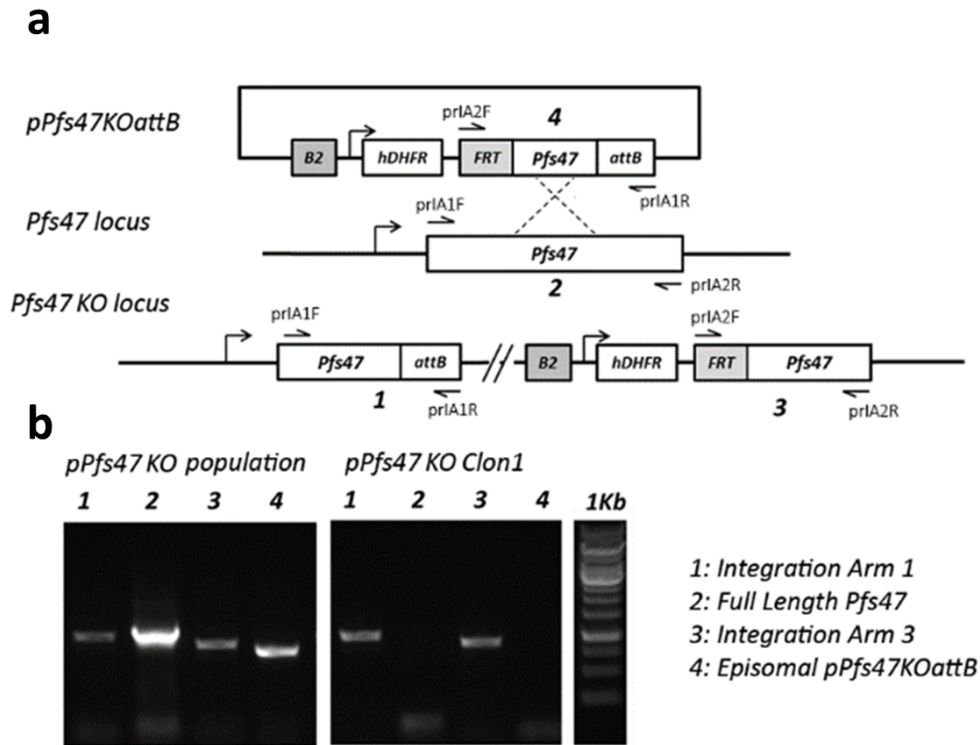


Fig.S4. Strategy for the generation of the *Pfs47KO* line and its genotype confirmation. **a**, Schematic representation of the predicted disruption of the *Pfs47* gene on chromosome 13 via a single-crossover strategy by transfection of the plasmid *pPfs47KOattB* in *P. falciparum* NF54. The drug selection cassette *hDHFR* and the recombination adaptor sites *B2*, *FRT* and *attB* are showed. **b**, Genotypic characterisation with integration-specific PCR products before and after cloning of parasites by minimal dilution, based on different combination of primers pair *prIA1F/prIA1R* and *prIA2F/prIA2R*. Lane 1 and 3 corresponds to the integration arms 1 and 2, respectively, lane 2 correspond to full length *Pfs47* and lane 4 correspond to the presence of episomal plasmid. Representation is not in scale and is for illustration purposes only.

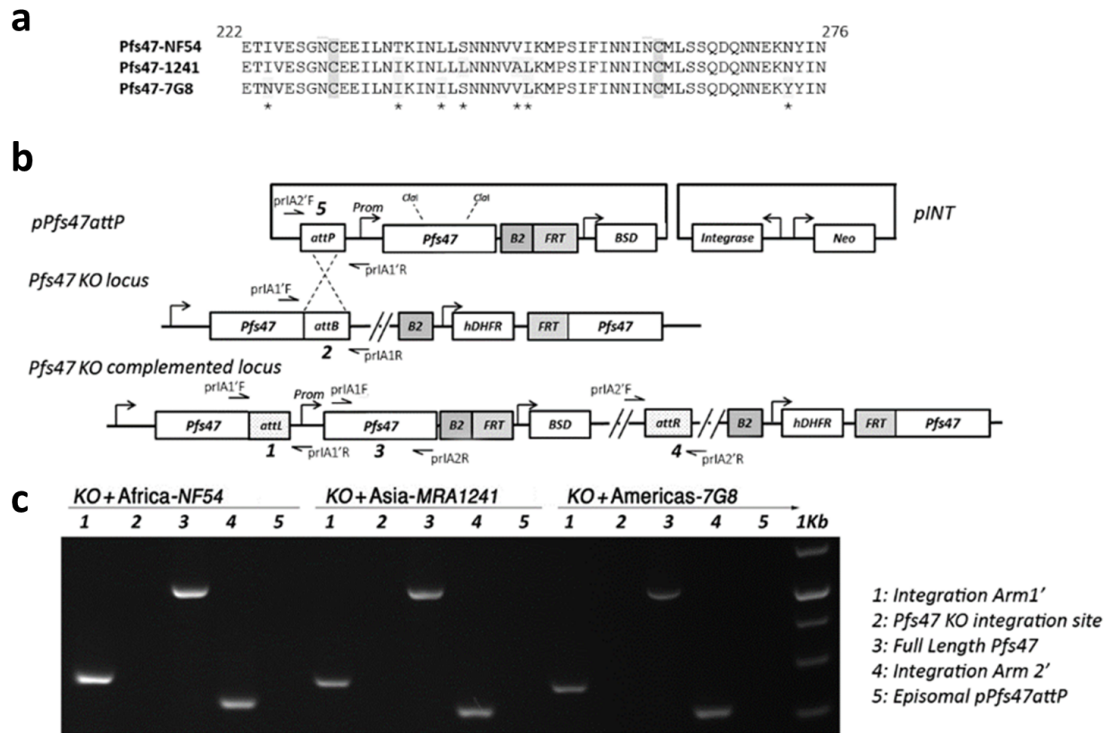


Fig. S5. Strategy for the complementation of the *Pfs47KO* line with representative haplotypes of *Pfs47* from different continents and its genotype confirmation. **a**, Amino acid alignment of *Pfs47* haplotypes (NF54, 1241, 7G8) used in this study to complement the *Pfs47KO* line. Only the region with polymorphisms within domain 2 (marked with an asterisk) is showed (aa222-aa276). **b**, Schematic representation of the integrase-mediated complementation of different *Pfs47* haplotypes on the *Pfs47KO* locus by the plasmid pPfs47attP. The drug selection cassettes hDHFR, BSD and Neo and the recombination adaptor sites B2, FRT and attB are showed. **c**, Genotype confirmation of the complementation of the *Pfs47KO*. Integration-specific PCR products after cloning of parasites by minimal dilution, based on different combination of the primers pairs prIA1'F/prIA1'R, prIA2'F/prIA2'R, prIA1F/prIA1R and prIA2F/prIA2R. Lane 1 and 4 corresponds to the integration arms 1 and 2, respectively, lane 2 correspond to the *Pfs47KO* integration site, lane 3 correspond to full length *Pfs47* and lane 5 corresponds to the presence of episomal plasmid. Representation is not in scale and is for illustration purposes only.

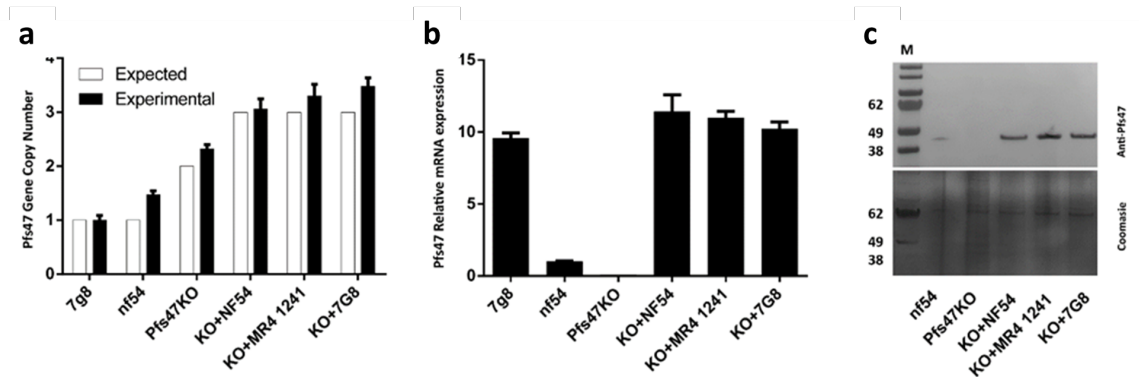


Fig. S6. Gene copy number and expression analysis of *Pfs47* in the *P. falciparum* lines generated in this study. **a**, *Pfs47* gene copy number of the different lines generated was assessed by qPCR. Primers were designed to amplify a 84bp sequence included in *Pfs47* 851 bp internal fragment in order to detect one copy in the wild type strain, two in the *Pfs47KO* and three in the complemented ones. **b**, Relative mRNA expression of *Pfs47* in wild-type (7G8 and NF54), *Pfs47* knockout (KO), and the KO line complemented with the NF54 or 7G8 or MR4 1241 allele of *Pfs47*. Relative mRNA expression of *Pfs47* was assessed by qPCR in stage IV–V gametocyte cultures. Detection of *Pfs47* mRNA in the complemented lines confirms gene expression upon complementation of the *Pfs47KO* line under the *P. falciparum* 7G8 promoter. **c**, Western blot analysis of expression of Pfs47 protein in equivalent amounts of gametocyte cultures of wild-type (NF54), *Pfs47KO*, and *Pfs47KO* line complemented with the NF54 or 7G8 or MR4- 1241 allele of *Pfs47*. Detection of *Pfs47* protein in the complemented lines confirms gene expression upon complementation of the *Pfs47KO* line.

<i>P. falciparum</i> strain	Experiment 1			Experiment 2		
	<i>Ag</i>	<i>Ad</i>	<i>Aa</i>	<i>Ag</i>	<i>Ad</i>	<i>Aa</i>
NF54						
n (midguts)	38	40	36	32	32	30
Oocyst/mgt Median	21	2.5***	0****	91.5	6.5****	0****
(range)	(0-155)	(0-61)	(0-11)	(10-215)	(0-76)	(0-15)
Prevalence (%)	92	60**	30****	100	87*	40****
MRA-1241						
n (midguts)	40	30	30	45	28	30
Oocyst/mgt Median	2****	49.5	0*****	3****	37.5	0****
(range)	(0-54)	(0-166)	(0-2)	(0-69)	(0-125)	(0-0)
Prevalence (%)	68 ^{ns}	87	13****	69 ^{ns}	86	0****
Pf 7G8						
n (midguts)	40	39	38	37	29	32
Oocyst/mgt Median	1**	0****	2	1*	0****	2
(range)	(0-8)	(0-7)	(0-26)	(0-9)	(0-4)	(0-12)
Prevalence (%)	53 ^{ns}	28***	68	31*	3****	81

Table S1. Compatibility between *Plasmodium falciparum* isolates from different geographic origin and different anopheline mosquito species. Infection of *Anopheles gambiae* (*Ag*), *Anopheles dirus* (*Ad*) and *Anopheles albimanus* (*Aa*) mosquitoes with *P. falciparum* NF54 (putative African), MRA-1241 (Asian) and 7G8 (South America) lines. Table shows infection intensity and prevalence for two replicate experiments, results for Experiment 1 are also presented in Fig. 1 in the main text. Infection intensity medians were compared using the Mann-Whitney test; infection prevalences were compared with the χ^2 test, ns= no significant difference, * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.

<i>P. falciparum</i> strain	Experiment 1		Experiment 2	
NF54	<i>Anopheles dirus</i>			
	dsLacZ	<i>Ad</i> dsLRIM1	dsLacZ	<i>Ad</i> dsLRIM1
n (midguts)	19	19	29	30
Oocyst/mgt Median	4	28**	8	13.5***
(range)	(0-46)	(0-157)	(0-26)	(3-47)
Prevalence (%)	58	84 ^{ns}	96	100 ^{ns}
NF54	<i>Anopheles albimanus</i>			
	dsLacZ	<i>Aa</i> dsLRIM1	dsLacZ	<i>Aa</i> dsLRIM1
n (midguts)	38	35	32	32
Oocyst/mgt Median	0	184***	0	3****
(range)	(0-42)	(18-277)	(0-1)	(0-10)
Prevalence (%)	42	100****	3	83****
MRA-1241	<i>Anopheles albimanus</i>			
	dsLacZ	<i>Aa</i> dsLRIM1	dsLacZ	<i>Aa</i> dsLRIM1
n (midguts)	39	43	38	26
Oocyst/mgt Median	0	40****	0	31****
(range)	(0-14)	(3-115)	(0-11)	(0-289)
Prevalence (%)	43	100****	26	85****

Table S2. Effect of disrupting the mosquito complement-like system by silencing the leucine-rich repeat immune protein 1 (LRIM1) on susceptibility of *A. dirus* and *A. albimanus* to infection with incompatible *P. falciparum* strains. Female mosquitoes were injected with LacZ control or LRIM1 dsRNA 3 days before feeding on a gametocyte culture. Table shows infection intensity and prevalence for two replicate experiments, results for Experiment 1 are also presented in Fig. 1 in the main text. Infection intensity medians were compared using the Mann-Whitney test; infection prevalences were compared with the χ^2 test, ns = no significant difference, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

P. falciparum NF54 transformant	Experiment 1			Experiment 2		
	Ag	Ad	Aa	Ag	Ad	Aa
Pfs47 KO						
n (midguts)	27	29	26	31	31	29
Oocyst/mgt Median	0	2 ^{ns}	0 ^{***}	0	0 ^{ns}	0 ^{ns}
(range)	(0-20)	(0-22)	0	(0-3)	(0-8)	0
Prevalence (%)	22	34 ^{ns}	0*	48	55 ^{ns}	0****
KO + Pfs47 Africa NF54						
n (midguts)	25	27	23	29	29	25
Oocyst/mgt Median	4	0****	0****	3	0****	0****
(range)	(0-18)	(0-2)	(0-3)	(0-14)	(0-4)	(0-2)
Prevalence (%)	72	34**	27**	73	38**	12****
KO + Pfs47 Asia MRA-1241						
n (midguts)	30	30	25	26	27	20
Oocyst/mgt Median	0***	6	0****	0***	4	0****
(range)	(0-4)	(0-22)	(0-1)	(0-6)	(0-10)	0
Prevalence (%)	45 ^{ns}	67	12****	47*	78	0***
KO + Pfs47 Americas 7G8						
n (midguts)	31	26	27	27	26	24
Oocyst/mgt Median	1 ^{ns}	0 ^{ns}	1	1 ^{ns}	0 ^{ns}	1
(range)	(0-7)	(0-8)	(0-8)	(0-5)	(0-3)	(0-4)
Prevalence (%)	55 ^{ns}	43 ^{ns}	63	56 ^{ns}	47 ^{ns}	67

Table S3. Effect of complementing NF54 *Pfs47*KO parasites with different *Pfs47* haplotypes on infectivity to different anopheline mosquito species. Infection of *Anopheles gambiae* (Ag), *Anopheles dirus* (Ad) and *Anopheles albimanus* (Aa) mosquitoes with the *P. falciparum* NF54 *Pfs47*KO line and complemented derivatives expressing *Pfs47* haplotypes from *P. falciparum* NF54 (African), MRA1241 (Asian) and 7G8 (South American) strains. Table shows infection intensity and prevalence for two replicate experiments, results for Experiment 1 are also presented in Fig. 3 in the main text. Infection intensity medians were compared using the Mann-Whitney test and infection prevalences using the Chi-square test, ns = no significant difference, * p<0.05, *** p<0.001, **** p<0.0001.

Table S4.

	Exxperiment 1		Experiment 2	
<i>Pfs47 KO</i>	<i>Anopheles gambiae</i>			
	<i>dsLacZ</i>	<i>Ag dsTEP1</i>	<i>LacZ</i>	<i>Ag dsTEP1</i>
n (midguts)	35	29	36	22
Oocyst/mgt Median	0	9****	0	0
(range)	(0-22)	(0-45)	0	(0-6)
Prevalence (%)	14	63****	0	41****
<i>Pfs47 KO</i>	<i>Anopheles dirus</i>			
	<i>dsLacZ</i>	<i>Ad dsLRIM1</i>	<i>dsLacZ</i>	<i>A.d. dsLRIM1</i>
n (midguts)	35	36	33	29
Oocyst/mgt Median	0	3***	0	1***
(range)	(0-5)	(0-22)	(0-4)	(0-8)
Prevalence (%)	29	62***	25	74***
<i>Pfs47 KO</i>	<i>Anopheles albimanus</i>			
	<i>dsLacZ</i>	<i>A.a. dsLRIM1</i>	<i>dsLacZ</i>	<i>A.a. dsLRIM1</i>
n (midguts)	25	25	22	31
Oocyst/mgt Median	0	6****	0	19****
(range)	0	(0-21)	0	(0-37)
Prevalence (%)	0	80****	0	91****
<i>KO + Pfs47 Africa NF54</i>	<i>Anopheles gambiae</i>			
	<i>dsLacZ</i>	<i>Ag dsTEP1</i>	<i>dsLacZ</i>	<i>Ag dsTEP1</i>
n (midguts)	28	34	32	36
Oocyst/mgt Median	3	4 ^{ns}	6.5	8 ^{ns}
(range)	(0-14)	(0-20)	(0-22)	(0-29)
Prevalence (%)	79	77 ^{ns}	82	81 ^{ns}
<i>KO + Pfs47 Asia MRA-1241</i>	<i>Anopheles dirus</i>			
	<i>dsLacZ</i>	<i>Ad dsTEP1</i>	<i>dsLacZ</i>	<i>Ad dsTEP1</i>
n (midguts)	30	28	28	33
Oocyst/mgt Median	2	2 ^{ns}	2	2.5 ^{ns}
(range)	(0-6)	(0-8)	(0-9)	(0-11)
Prevalence (%)	70	79 ^{ns}	72	70 ^{ns}
<i>KO + Pfs47 Americas 7G8</i>	<i>Anopheles albimanus</i>			
	<i>dsLacZ</i>	<i>Aa dsTEP1</i>	<i>dsLacZ</i>	<i>Aa dsTEP1</i>
n (midguts)	29	22	34	32
Oocyst/mgt Median	0	1.5 ^{ns}	1	2 ^{ns}
(range)	(0-6)	(0-9)	(0-6)	(0-14)
Prevalence (%)	69	63 ^{ns}	47	62 ^{ns}

Table S4. Effect of disrupting the mosquito complement-like system, by silencing the leucine-rich repeat immune protein 1 (LRIM1) in *Anopheles dirus* (*Ad*) and *Anopheles albimanus* or the thioester-containing protein 1 (TEP1) of *Anopheles gambiae* (*Ag*), on infection with the NF54 *Pfs47KO* line or complemented derivatives. Female mosquitoes were injected with dsLacZ control, dsTEP1 or dsLRIM1 three days before feeding on a gametocyte culture. Table shows infection intensity and prevalence for two replicate experiments, results for Experiment 1 are also presented in Fig. 1 in the text. Infection intensity medians were compared using the Mann-Whitney test and infection prevalence using the Chi-square test, *ns* = no significant difference, * $p < 0.05$, *** $p < 0.001$, **** $p < 0.0001$.

Population 1	Population 2	Fst
All populations		0.69052
Africa	Asia	0.79551
Africa	Americas	0.75416
Africa	PNG	0.46926
Asia	Americas	0.87647
Asia	PNG	0.47820
Americas	PNG	0.60479

Table S5. Index of fixation (Fst) for *Pfs47* haplotypes. The Fst was estimated across all populations and pair-wise between each continental population. The high Fst values between the different continental populations indicate strong geographic differentiation of *Pfs47* between those regions.

	SP (1-27aa)		Domain 1 (28-182aa)						Domain 2 (183-282aa)										Domain 3 (283-413aa)						GPI-anch. (413-439aa)		Number of haplotypes per region							
Protein Haplotype	13*	27	28	46	55	68	144	178	186	188	194	219	224	236	240	242	247	248	262	270	272	273	297	303	314	356	369	408	426	433	Africa	Asia	PNG	Americas
Hap 3D7	I	E	L	V	E	T	S	I	I	E	P	M	I	T	L	S	V	I	L	E	N	Y	I	I	I	T	P	V	I	F				
Hap 1	.	D	H	L	H	.	.	.	1	0	0	0
Hap 2	.	D	H	L	.	.	I	1	0	0	0	
Hap 3	.	D	H	L	.	.	I	1	0	0	0	
Hap 4	.	D	H	L	6	0	0	0	
Hap 5	.	D	H	L	27	0	0	0	
Hap 6	.	D	H	.	.	.	I	.	.	L	1	0	0	0	
Hap 7	H	.	.	.	I	.	.	L	1	0	0	0	
Hap 8	D	H	L	4	0	1	0	
Hap 9	H	L	S	.	.	.	1	0	0	0	
Hap 10	H	L	V	1	0	0	0	
Hap 11	N	.	.	.	H	L	2	0	0	0	
Hap 12	H	L	K	.	.	.	1	0	0	0	
Hap 13	H	L	102	0	0	0	
Hap 14	H	L	4	0	0	0	
Hap 15	T	H	L	H	.	.	1	0	0	0	
Hap 16	H	L	H	.	.	3	0	0	0	
Hap 17	H	L	H	.	.	6	0	0	0	
Hap 18	H	L	.	.	I	H	.	.	1	0	0	0	
Hap 19	H	L	.	.	I	H	.	.	4	0	0	0	
Hap 20	L	12	0	0	0	
Hap 21	L	L	.	Q	1	0	0	0	
Hap 22	H	L	.	.	Y	N	1	0	0	0	
Hap 23	.	D	H	L	.	.	Y	N	1	0	0	0	
Hap 24	H	I	L	.	.	Y	8	0	0	0	
Hap 25	H	L	.	.	I	1	0	0	0	
Hap 26	H	L	.	.	I	14	0	0	0	
Hap 27	D	H	L	.	.	I	5	0	0	0	
Hap 28	D	H	L	H	.	.	3	0	0	0	
Hap 29	D	H	L	1	0	0	0	
Hap 30	H	L	14	0	0	0	
Hap 31	H	.	.	I	.	.	.	L	0	1	0	0	
Hap 32	.	.	.	L	H	L	F	L	.	1	0	0	0
Hap 33	H	L	F	0	0	13	0	
Hap 34	H	.	.	I	.	L	A	L	1	0	1	28	
Hap 35	M	.	V	.	.	H	L	A	L	0	0	0	6	
Hap 36	V	.	.	.	H	L	0	0	0	2	
Hap 37	.	.	.	I	H	.	N	I	I	.	.	L	.	.	Y	0	1	0	0	
Hap 38	H	.	N	I	I	.	.	L	.	.	Y	L	V	0	1	0	0
Hap 39	H	.	N	I	I	.	.	L	.	.	Y	0	30	0	0	
Hap 40	.	.	.	I	.	K	.	.	.	K	H	.	N	I	I	.	.	L	0	5	0	0	
Hap 41	.	.	.	I	.	K	.	.	.	K	H	.	N	I	I	.	.	L	L	.	.	.	I	.	0	1	0	0	
Hap 42	.	.	.	I	.	K	H	.	N	I	I	.	.	L	.	.	Y	1	34	8	0	

Table S6. Sequence polymorphisms and regional frequencies of *Pfs47* protein haplotypes. Amino acid changes are indicated for each *Pfs47* haplotype in reference to *Pfs47* 3D7. The number of isolates with a given *Pfs47* haplotype is indicated for different geographic regions. The predicted *Pfs47* domain structure is indicated (SP, signal peptide; GPI-anch, GPI anchorage). Order of *Pfs47* haplotypes and color shading corresponds to the phylogenetic tree in Fig. 2 of the main text. *Indicates amino acid position.

Table S7. Sequence of oligonucleotide primers used in this study

Primer name	Sequence
prFRTFfs47F	ATCGATAACTCCATGGGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGAATCTCATTTATATTCTGCTTA
prattBPfs47R	CCTAGGAGTTCATGGCCGCGATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGTTATAACATATACATGCCTTCCTTAT
prB2F	AGAATACTCGCGGCCGCGAGTTTCATTAAGGAATAACTAATTCCTAATGAAACTCCGCGGGGAGGACTAGTCCG
prB2R	CGGACTAGTCTCCCCGCGGGAGTTTCATTAGGGAATTAGTTATTCCTTAATGAAACTCGCGGCCGCGAGTATTCT
prB2FRTF	GCAGCCCGGGGATCCGAGTTTCATTAAGGAATAACTAATTCCTAATGAAACTCGCGGCCGCGAAGTTCCTATTCTCTA
prB2FRTR	AAATACTAGTGGATCGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGCGGCCGCGAGTTTCATTAGGGAATTA
prattPF	ATGTTTTTATAAACCAAGCTTGGGGTTTGTAACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCGAGGG
prattPR	AGCTGGAGTCCACCCTCGAGGTTTGTCTGGTCAACCACCGCGTCTCAGTGGTGTACGGTACAAACCCCAAGCTTGG
prClaF	AAAAATTAGATATCGATAAAAAAGTATGCAATAAATTCATC
prClaR	GTCTTTTGCGAATCGATTTCCTTTTCATATAATACATGTTTA
prIA1F	GTATGGGAAGAATGATCAGC
prIA1R	CCGCGATGATCCTGACG
prIA2F	CCTATACTTTCTAGAGAATAGGAACTTC
prIA2R	TCATATGCTAACATACATGTAAAAAA
prIA1'F	TCATATGCTAACATACATGTA
prIA1'R	ATGTTATACACATAATGTGTTAGGA
prIA2'F	GATGGGTGTGATTTTACGAA
prIA2'R	CCATATGCGGTGTGAAATACCG
0203F	GATGCTGCTGGAAAACTAC
0203R	CCTACATCCCATACGGTAAA
PF13_0248F	GCAGGCATTAAATGTCCATA
PF13_0248R	CTTTTGCGAATCGATTTCCT