A population modification gene drive targeting both *Saglin* and *Lipophorin* disables *Plasmodium* transmission in *Anopheles* mosquitoes

Emily Green, Etienne Jaouen, Dennis Klug, Roenick Proveti Olmo, Amandine Gautier, Stéphanie Blandin, Eric Marois*

Inserm U1257, CNRS UPR9022, Université de Strasbourg, Strasbourg, France

*Corresponding author: e.marois@unistra.fr

Abstract

Lipophorin is an essential, highly expressed lipid transporter protein that is secreted and circulates in insect hemolymph. We hijacked the *Anopheles gambiae Lipophorin* gene to make it co-express a single-chain version of antibody 2A10, which binds sporozoites of the malaria parasite *Plasmodium falciparum*. The resulting transgenic mosquitoes show a markedly decreased ability to transmit *Plasmodium berghei* expressing the *P. falciparum* circumsporozoite protein. To force the spread of this anti-malarial transgene in a mosquito population, we designed and tested several CRISPR/Cas9-based gene drives. One of these is installed in, and disrupts, the pro-parasitic gene *Saglin* and also cleaves wild type *Lipophorin,* causing the anti-malarial modified *Lipophorin* version to hitch-hike together with the *Saglin* drive. Although producing drive-resistant alleles, the *Saglin*-based gene drive reached high levels in caged mosquito populations and efficiently promoted the simultaneous spread of the antimalarial *Lipophorin::Sc2A10* allele. This combination is expected to affect parasite transmission by two different mechanisms. This work contributes to the design of novel strategies to spread antimalarial transgenes in mosquitoes, and illustrates some expected and unexpected outcomes encountered when establishing a population modification gene drive.

Introduction

Malaria-transmitting mosquitoes cause over 240 million malaria cases and 620,000 deaths annually (WHO, 2021). Vector control, mainly in the form of insecticide indoor residual spraying and insecticide-treated bed nets, has permitted a steady decrease in malaria cases during the last decades. However, this decline is stalling as mosquitoes become increasingly resistant to insecticides and as malaria parasites continue to evolve resistance to drugs. These ongoing genetic changes in vectors and parasites, under the selection pressure of two major pillars of malaria control, jeopardize the hard-won successes of the fight against malaria. To address the need for complementary malaria control approaches, intervention strategies based on intentional genetic changes in mosquitoes to decrease malaria transmission are under development. Recent advances in the field of gene drive (GD) provide a mean to push desirable transgenes to invade a target population of insects (Carballar-Lejarazú and James, 2017; Gantz and Akbari, 2018; Quinn and Nolan, 2020). GD interventions against malaria mosquitoes are

still in the laboratory phase of their development, and can be subdivided in two classes: (a) population suppression GDs that spread female sterility in mosquito populations with the aim to eliminate them (Hammond et al., 2016; Kyrou et al., 2018); (b) population modification GDs that leave mosquito populations in place, but genetically alter them to reduce their disease transmission capacity (Gantz et al., 2015; Pham et al., 2019; Carballar-Lejarazú et al., 2020). Both approaches raise ethical, safety and ecological concerns, and trigger controversy. One concern raised by elimination approaches is their possible consequences on the food web and ecosystems, even where dozens of sympatric mosquito species are potentially available to compensate for the loss of the target species. In some regions, elimination of one species could favor the installation of competing species that could also act as vectors, as can be feared for the *Aedes aegypti* / *Ae. albopictus* mosquito dyad vectoring dengue, Zika and chikungunya viruses. Modification approaches alleviate these concerns by leaving the target species in its ecological niche, but in turn, the indefinite persistence of transgenes utilized in modification drives could have secondary effects, so that both approaches should be planned with utmost caution. However, the option of rejecting gene drive technologies as a complementary tool to fight mosquito-borne pathogens also raises ethical issues, by ignoring a chance to curb the immense mortality and morbidity burden caused by mosquito-borne infections.

Anopheles gambiae (*sensu lato*), the major mosquito vector of malaria in Sub-Saharan Africa, is widely distributed, locally abundant during rainy seasons and its biomass might significantly contribute to the food chain, in a temporal fashion. Besides their possible negative impact on ecosystems, elimination approaches for this mosquito are likely to be complicated by re-colonization from residual pockets of surviving isolated populations, resulting in chasing dynamics of gene drive (Champer et al., 2021). For this species in its native ecosystem, population modification to decrease vector competence and preserve ecological functions could be a meaningful gene drive approach. Here, we present the engineering of the *A. gambiae Lipophorin* (*Lp*) essential gene, which naturally encodes a lipid transporter protein secreted in mosquito hemolymph. We modified *Lp* to co-express a single-chain antibody (ScFv) derived from 2A10, a well-established mouse monoclonal antibody that binds the *Plasmodium falciparum* circumsporozoite protein (CSP) (Zavala et al., 1983). Sporozoites are the parasite stage that mosquito females inject in vertebrate hosts while aquiring a blood meal. Once in the host, they quickly infect the liver and multiply to produce merozoites with red blood cell tropism. Through its binding to CSP (the major sporozoite surface antigen), 2A10 was shown to decrease sporozoite infectivity to cultured human hepatocytes (Hollingdale et al., 1984). Transgenic expression in *Anopheles stephensi* of a 2A10 ScFv, fused to the antimicrobial peptide Cecropin A and expressed under the control of the *Vitellogenin* promoter, decreased the number of salivary gland sporozoites (Isaacs et al., 2011; Isaacs et al., 2012). Expression in transgenic *A. gambiae* of a different CSP-binding ScFv without Cecropin decreased the sporozoites' ability to infect cultured hepatocytes and to cause malaria in mice (Triller et al., 2017). Here, we inserted a 2A10 ScFv coding sequence within the endogenous open reading frame of the *Lp* essential gene, exploiting a natural proteolytic cleavage site to separate the fusion protein moieties during their secretion. Using a mouse model of malaria, in which a transgenic rodent *Plasmodium berghei* strain expresses CSP from *P. falciparum*, we show a strong

decrease in the ability of the genetically engineered mosquitoes to transmit *Plasmodium*. In a second step, we tested different CRISPR/Cas9-based GD designs to force the spread of the *Lp::Sc2A10* transgene in laboratory mosquito populations. A first approach aimed to boost *Lp::Sc2A10* allelic frequency indirectly, using a suppression GD removing wild-type (WT) *Lp*. This failed, as both WT *Lp* and *Lp::Sc2A10* proved to be largely haploinsufficient in the presence of a *Lp* loss-of-function GD allele. In a second approach we inserted the GD cassette into, and disrupting, the *Saglin* gene, which encodes a *Plasmodium* agonist (Okulate et al., 2007 ; Ghosh et al., 2009; but see O'Brochta et al., 2019). In contrast to *Lp*, the loss of *Saglin* does not by itself impose a large fitness cost on laboratory mosquitoes (Klug et al., 2022). The *Saglin* GD also expressed *Lp* guide RNAs promoting homing of *Lp::Sc2A10* into WT *Lp*. These gRNAs targeting the essential *Lp* gene confer lethality to the *Saglin* GD and block its spread, unless the *Lp::Sc2A10* transgene is present. We show that this GD stimulates dual homing at both the *Lp* and *Saglin* loci, resulting in the simultaneous spread of the two transgenes in caged mosquito populations. These combined modifications are expected to decrease sporozoite transmission by two distinct mechanisms, i.e. depriving *Plasmodium* from the Saglin agonist plus attacking sporozoites with the 2A10 ScFv. GDs are susceptible to the emergence of GD-refractory mutant versions of their target sites (Hammond et al., 2017; Champer et al., 2018; Carballar-Lejarazú et al., 2022). Because our design employs multiplex gRNAs for both the *Saglin* and the *Lp* loci, we expected most failed homing events to result in loss-of-function *Lp* mutations that would be naturally counter-selected, and/or loss-of-function, GD-resistant *Saglin* alleles that should still contribute to decreasing vector competence. However, at least one functional allele that is refractory to homing was observed at the *Lp* locus.

Results

Development of an engineered Lp allele encoding an anti-sporozoite factor while preserving Lp function

The two most abundant proteins secreted in *Anopheles* hemolymph are the nutrient transporters Lipophorin (Lp) and Vitellogenin (Vg), as illustrated by Coomassie staining of electrophoresis gels revealing both proteins to be the major visible bands in hemolymph extracts (Fig. 1A and Rono et al., 2010). While Vg is induced in mosquito females only upon blood feeding and wanes within 3 days as eggs develop, Lp is expressed constitutively, with a boost of expression after blood feeding. We reasoned that Lp protein abundance, secretion in the hemolymph and persistence of expression outside gonotrophic cycles make *Lp* a particularly attractive host gene to co-express and secrete high levels of synthetic anti-*Plasmodium* factors. Here we tested a 2A10 scFv version without fusion to Cecropin as antimalarial effector, and hereinafter refer to it as Sc2A10. Using the CRISPR-Cas9 system, we knocked-in codon-optimized *Sc2A10* in frame after the endogenous first *Lp* coding exon, which encodes the Lp secretion signal peptide (Fig. 1B). As a result, transcription of the chimeric *Lp::Sc2A10* mRNA encodes the natural Lp signal peptide followed by Sc2A10 and the remainder of the Lp protein sequence. A natural proteolytic cleavage site located between the ApoLpII and ApoLpI subunits of the Lp polypeptide allows their separation during maturation in the Golgi

apparatus. We duplicated this motif to detach Sc2A10 from ApoLpII during protein maturation. The knock-in design also introduced silent mutations in the majority of possible Cas9 sgRNA target sites within *Lp* exon 1 and the beginning of intron 1, to facilitate selection of multiple sgRNAs targeting WT *Lp*, but not its engineered allele, when designing subsequent GDs. Finally, the 1644-bp first intron immediately following *Lp* exon 1 offered an ideal location to host a fluorescent selection marker to identify and track the modified gene. We placed a *GFP* marker gene under control of the synthetic *3xP3* promoter (Horn and Wimmer, 2000) and *D. melanogaster Tubulin56D* terminator in this intron, with care taken to preserve the intron splice junctions and in reverse orientation relative to *Lp* transcription so that the *Tub56D* terminator would not arrest *Lp* transcription. The *Lp* promoter is active in the fat body, while the *3xP3* promoter is active in the nervous system. These synthetic modifications in the *Lp* locus are schematized in Figure 1B and annotated in the provided sequence file (Suppl. file 1A).
Using automated flow cytometry-based (COPAS) selection of live neonate larvae based on differential GFP expression between homozygotes and heterozygotes (Marois et al., 2012), we established two stocks of *Lp::Sc2A10* mosquitoes, one heterozygous, the other homozygous. The homozygous stock was fertile and viable, devoid of obvious fitness cost. This indicates that the transgenic insertion into *Lp* did not disrupt its essential functions in development, physiology, flight or reproduction. A more careful evaluation of *Lp::Sc2A10* fitness costs was conducted by following the dynamics of the transgene in the heterozygous population over 17 generations. This revealed slow disappearance of the transgene over time (Figure 1C and Suppl. Table 1A) indicative of a subtle fitness cost that decreased transgene frequency by an average of 2.3% each generation. In fertility assays comparing wild-type and homozygous transgenic females, the latter consistently engendered fewer progeny than their WT sisters, pointing to a negative impact of the *Lp* modification on their reproductive capacity (Suppl. Table 1B).

The Lp::Sc2A10 *transgene directs Sc2A10 secretion into mosquito hemolymph*

To test whether the Sc2A10 single-chain antibody was expressed, and secreted in hemolymph via the Lp signal peptide, we extracted hemolymph from transgenic female mosquitoes and analyzed its protein content by mass spectrometry. Because the *Lp::Sc2A10* gene fusion produces a single mRNA, with the endogenous Lp signal peptide addressing the entire fusion protein to the secretory pathway, the Sc2A10 protein fragment must initially be stoechiometric to the Lp subunits ApoLpI and ApoLpII. Upon proteolytic cleavage dissociating Sc2A10, ApoLpII and ApoLpI and after their secretion into the hemolymph, each moiety possibly follows a different fate. In two mass spectrometry samples from the hemolymph of homozygous transgenic mosquitoes, for a total of 2629 peptides from the largest Lp subunit ApoLpI, we detected 672 peptides from smaller ApoLpII and 18 from Sc2A10 (Table 1). Corrected for size, this suggests that Sc2A10 molecules are 10 to 20-fold less abundant in the hemolymph compared to ApoLpI and ApoLpII. This observation suggests that while Sc2A10 is made and secreted into hemolymph, its disappearance by degradation, uptake in cells or stickiness to tissue, is faster than that of its sister Lp proteins. Of note, western blots using anti-ApoLpII antibodies on hemolymph samples from homozygous transgenic mosquitoes

showed no size upshift for ApoLpII, confirming that Sc2A10 is properly released from ApoLpII (Figure 1D). Overall higher concentrations of Lp in heterozygous compared to homozygous mosquitoes, and a 3 to 4-fold lower Sc2A10 : Lp ratio in heterozygous mosquitoes, indicate that the chimeric proteins from *Lp::Sc2A10* are less expressed than from WT *Lp*.

Sample	ApoLpI	ApoLpII	Sc2A10
WT control	2075 (5.8)	486 (5.1)	0 (0)
hmz Sc2A10, sample 1	1627 (5.1)	405 (4.8)	11 (0.4)
hmz Sc2A10, sample 2	1002 (4.4)	267 (4.4)	7 (0.3)
het Sc2A10, sample 1	1557 (10.1)	351 (8.7)	2 (0.1)
het Sc2A10, sample 2	2290 (6.0)	664 (6.7)	2 (0.1)

Table 1: Mass spectrometry identification of peptides from Lipophorin and Sc2A10 in hemolymph from transgenic mosquitoes and wild-type sibling controls. The table shows the total number of peptide spectra detected in each hemolymph sample for the two Lp subunits and for Sc2A10. Hemolymph was collected from homozygous (hmz) and heterozygous (het) transgenic mosquitoes. Relative, normalized abundance of these peptides in the protein sample, calculated by dividing their spectral count by the respective protein's molecular weight (kDa) and by the total spectral counts from known abundant hemolymph proteins found in the sample (ApoLpIII, APLC1, LRIM1, Nimrod, Phenoloxidase, and TEP1), is indicated in parentheses. Resulting values were multiplied by a constant factor equal to the average number of spectra from these control proteins across samples.

The Lp::Sc2A10 *transgene decreases* Plasmodium *transmission*

To assess whether the amount of Sc2A10 present in hemolymph can affect *Plasmodium* transmission, we infected homozygous transgenic mosquitoes and their control wild-type siblings, allowed parasites to develop for 16-20 days and exposed naïve mice to be bitten by groups of 10 infected mosquitoes per mouse. The *Plasmodium berghei* strain we employed, *Pb-PfCSP*$^{\text{hsp70-GFP}}$, expresses GFP at a high level under control of the constitutive *hsp70* promoter and has its native CSP substituted with that from *P. falciparum* (Manzoni et al., 2014; Triller et al., 2017 and see Methods), which is the antigen recognized by 2A10. One day before exposure of naïve mice, the cold-anesthetized mosquitoes were examined under 488-nm light. Those lacking visible GFP parasites (oocysts in the abdomen, sporozoites visible through the cuticle in wing veins or salivary glands) were discarded to ensure that mice were exposed only to *Plasmodium*-infected mosquitoes. We controlled that each mouse was bitten by at least 6 infected mosquitoes, with exceptions indicated in Suppl. Table 2, by counting the blood-engorged females.

Consistent with previous reports (Isaacs et al., 2011; Isaacs et al., 2012), Sc2A10 significantly reduced transmission of *Pb-PfCSP*. Only 29.7% (11/37) of mice exposed to Lp::Sc2A10 homozygous mosquitoes became infected, compared to 97.1% (33/34)

exposed to control mosquitoes (Figure 2 and Suppl. Table 2). Additionally, 6 of the 11 mice that did become infected by transgenic mosquitoes experienced a 1-day delay in the development of infection (Suppl. Table 2), indicative of reduced liver infection (Reuling et al., 2020). To verify that the transmission-blocking phenotype of the *Lp::Sc2A10* transgene was specific to CSP from *P. falciparum*, we performed a control bite-back experiment comparing *Pb-PfCSP*$^{hsp70\text{-}GFP}$ to its *Pb-hsp70-GFP* parental parasite strain expressing endogenous *PbCSP*. While *Lp::Sc2A10* and wild-type mosquitoes transmitted the control *P. berghei* strain to mice with equal efficiency (n=5/5 and 4/4 mice infected, respectively) only the WT mosquitoes transmitted *Pb-PfCSP*$^{hsp70\text{-}GFP}$ (n=3/3 mice infected), whereas *Lp::Sc2A10* mosquitoes failed to transmit the chimeric strain (n=0/4 mice infected). This indicates that the transmission blocking capacity of the *Sc2A10* transgene is indeed PfCSP-specific (Table 2).

Mosquito genotype	WT		*Lp::Sc2A10*	
CSP in *P. berghei*	*PbCSP*	*PfCSP*	*PbCSP*	*PfCSP*
Number of infected mice	4/4	3/3	5/5	0/4

Table 2: *Lp::Sc2A10* transgenic mosquitoes are specifically affected in their transmission of *Plasmodium* expressing CSP from *P. falciparum*.

We also investigated if the transmission-reducing activity of the *Lp::Sc2A10* transgene is dependent on the dosage of Sc2A10, i.e. whether a single copy of the transgene would be as efficient as two. For this we exposed naïve mice to groups of 10 infected mosquitoes carrying zero, one or two *Lp::Sc2A10* copies. While 17 out of 18 mice bitten by control wild-type mosquitoes became infected, only 2 out of 5 mice bitten by homozygous transgenic mosquitoes and 4 out of 12 mice bitten by heterozygous mosquitoes, became infected (Suppl. Table 2). While this experiment is relatively small scale, it indicates that heterozygous mosquitoes also show a reduction in their *Plasmodium*-transmitting capacity.

Testing a suppression gene drive homing into wild-type Lp *to push* Lp::Sc2A10 *toward fixation*

To push the *Lp::Sc2A10* anti-*Plasmodium* factor towards fixation in a mosquito population, our first strategy was to employ a transiently acting Cas9-based GD destroying the wild-type *Lp* gene (Suppl. Fig. 1A). This aimed to exploit the natural tendency of suppression GDs to initially spread in the population only to be supplanted, over time, by GD-immune functional mutants of the target gene (Hammond et al., 2017). Here *Lp::Sc2A10* itself, lacking the gRNA target sites, would play the role of a GD-immune, functional *Lp* mutant allele. The advantage of this approach was that a *Cas9* and gRNA-carrying GD construct inserted in *Lp* need not persist indefinitely in the population, being supplanted by the technologically more benign *Lp::Sc2A10* transgene.

To minimize the spontaneous additional emergence of undesirable GD-immune and functional *Lp* alleles, we used a multiplex of 4 gRNAs targeting *Lp*. The rationale was that in cases of failed GD homing, multiple Cas9 cuts repaired by non-homologous end joining (NHEJ) would mainly generate loss-of-function alleles in the essential *Lp* gene (e.g. deletions between Cas9 cut sites, frameshifts at individual cut sites) doomed to be eliminated by natural selection. We constructed a $DsRed_{NLS}$-marked *Lp* suppression GD (Suppl. Fig. 1A and annotated sequence in Supplemental File 1B), in which an array of 4 gRNAs separated by tRNAs (Xie et al., 2015; Port and Bullock, 2016), targeting *Lp* but not *Lp::Sc2A10*, were expressed under control of a single U6 promoter. *Cas9* was placed under the control of *zpg* regulatory elements, currently considered best for use in suppression GDs due to tighter restriction of Cas9 activity to the germ line (Kyrou et al., 2018; Hammond et al., 2021). Knocking-in this GD into the *Lp* locus was achieved by injecting this plasmid into >800 mosquito eggs carrying a heterozygous copy of GFP-marked *Lp::Sc2A10*, in order to retain a protected functional copy of *Lp* during the disruption of wild-type *Lp* by Cas9. Genomic insertion in injected embryos relied on expression of the construct's own CRISPR/Cas9 components supplemented with an additional heterozygous copy of a YFP-marked *vasa-Cas9* transgene in the embryos.

By screening for the presence of red fluorescence in the G1 progeny of about 140 surviving injected G0 mosquitoes, we recovered several dozens of knock-in G1 larvae (most of which were $DsRed_{NLS}$ and GFP positive) and examined transgene inheritance by crossing them to wild-type mosquitoes. In this cross, $DsRed_{NLS}$ marking the *Lp-GD* construct was expected to strictly segregate away from GFP marking the *Lp::Sc2A10* allele. Puzzlingly however, the majority of knock-in larvae contained a compound *Lp* locus harboring the $DsRed_{NLS}$-marked GD cassette strictly linked to the GFP-marked *Lp::Sc2A10* transgene. Sequencing PCR products spanning parts of this insertion confirmed that a complex chromosomal rearrangement (Suppl. figure 1B) had repaired the Cas9-cleaved wild-type chromosome by incorporating the entire injected plasmid, and by recombining with the intact sister chromosome bearing the *Lp::Sc2A10* transgene. The resulting unintended transgenic modification at the *Lp* locus harbored all the components of a classical payload-bearing GD that we hereinafter refer to as *GFP-RFP*. Similarly to the *Lp::Sc2A10* allele, the *GFP-RFP Lp* allele was homozygous viable, functional Lp protein being expressed from its *Lp::Sc2A10* moiety. In spite of the presence of the *zpg-Cas9* and *gRNA*-encoding cassettes, the *GFP-RFP* locus showed little homing activity after crosses to WT, except for the appearance of rare GFP-only or RFP-only progeny larvae, raising concern about the efficiency of the *zpg-Cas9* and/or the gRNA-expressing cassettes. Providing an independent copy of *vasa-Cas9* on a different chromosome rescued this defect and resulted in high homing rates of various segments from the *GFP-RFP* cassette (mostly GFP alone, more rarely $DsRed_{NLS}$ alone or both) (Suppl. Fig. 2A), indicating that at least some of the gRNAs were active and that the defect lied in *zpg-Cas9* expression.

Fortunately, we also recovered a minority of knock-in larvae that had integrated the expected *Lp-GD* construct marked with $DsRed_{NLS}$ only (Suppl. Fig. 1A). However, no evidence of gene drive was observed in further generations, as DsRed inheritance did not exceed 50% when crossing heterozygotes to wild-type. On the contrary, DsRed frequency dropped rapidly in F2 and F3 generations, indicating that the non-functional

GD cassette inserted in *Lp* conferred a strong fitness cost. Most heterozygous mosquitoes carrying one copy of *Lp-GD* and either *GFP-RFP* or *Lp::Sc2A10* showed a striking developmental delay and died before adulthood (Suppl. Figure 2B). Heterozygous mosquitoes carrying *Lp-GD* / WT survived to adulthood but were small and short-lived. These observations indicate that WT *Lp* is largely haplo-insufficient in presence of a *Lp* loss-of-function allele, even more so if WT *Lp* is replaced by *Lp::Sc2A10*. This last point is consistent with the observed fitness cost and lower *Lp* expression from the *Lp::Sc2A10* allele, that already suggested that albeit functional, this allele is inferior to the wild-type.

From the haplo-insufficiency of *Lp* we concluded that it was impossible to create a loss-of function GD in this locus, precluding an indirect same-locus GD approach. In addition, the *zpg* promoter appeared to have low activity when integrated in the *Lp* locus.

Driving Lp::Sc2A10 distantly from Saglin with a zpg-Cas9 gene drive

To host a GD promoting the spread of *Lp::Sc2A10*, we turned our attention to the mosquito *Saglin* locus. Indeed, while not leading to complete transmission blockage, the loss of Saglin reduces parasite prevalence and loads, and has no fitness cost in homozygous mutant laboratory populations (Klug et al., 2022). Similar to an antimicrobial combination therapy, combining Sc2A10 expression with the knock-out of *Saglin* can be expected to reinforce the transmission-blocking properties of each strategy taken alone. Therefore, we sought to disrupt *Saglin* with a gene drive cassette that would home into wild-type *Saglin* and concomitantly promote *Lp::Sc2A10* homing into the *Lp* locus (Figure 3).

The *zpg* promoter still was an attractive choice to build a GD dually targeting *Saglin* and *Lp*, to avoid drawbacks from other characterized germline promoters (maternal Cas9 deposition resulting in failed homing events in the zygote accompanied by the formation of GD refractory alleles ; possible somatic mosaic loss-of-function of wild-type *Lp* leading to haplo-insufficiency in heterozygous mosquitoes carrying one *Lp::Sc2A10* copy). To exclude that the low *zpg-Cas9* activity we previously observed at the *Lp* locus was due to overlooked mutations in our construct, we subcloned the *zpg-Cas9* cassette from a validated GD construct (Kyrou et al., 2018) and incorporated it in a *Saglin* knock-in plasmid expressing Cas9, $DsRed_{NLS}$, 3 gRNAs directed against *Saglin* and 4 gRNAs against *Lp* (Figure 3, and annotated sequence of this plasmid in Suppl. File 1C). From injection of this plasmid into heterozygous *Lp::Sc2A10* / WT embryos, we crossed 36 G0 male survivors to WT, and recovered more than 60 transgenic G1 individuals. The sole source of Cas9 in this injection was the knock-in GD plasmid itself, showing that *zpg-Cas9* was efficiently active at least for initial insertion. 100% of the G1 transgenics arising from injected males were females, an indication that the construct had correctly homed onto the X chromosome, where *Saglin* is located. In addition, all G1 transgenics had also inherited a copy of *Lp::Sc2A10* (GFP positive), instead of the expected 50%. This suggests that the GD construct acted on the *Lp* locus, reflecting either efficient *Lp::Sc2A10* homing already in the germ line of G0 males, and/or toxicity of the construct's *Lp* gRNAs resulting in death of larvae that did not inherit a protected *Lp* copy.

In spite of the efficiency of *zpg-Cas9* for initial integration, this first Saglin GD (termed *SagGD*^{zpg}) showed disappointingly modest homing levels at both the *Saglin* and *Lp* loci in the G2 generation (Figure 4A). Using COPAS flow cytometry, we examined GFP and DsRed inheritance in 2327 larvae arising from the pooled G1 females crossed to WT males. It showed 57.7% GFP inheritance (*Lp::Sc2A10* locus) compared to 50% expected in the absence of homing. DsRed inheritance was more difficult to assess, as the low intensity of red fluorescence did not allow accurate gating to separate red fluorescent from negative larvae on COPAS diagrams. Approximate gating completed with visual examination of a sample of 200 larvae suggested 54.1% DsRed inheritance at the *Saglin* locus, an even more modest inheritance bias than for GFP at the *Lp* locus. Mild homing could be explained by position effects affecting the efficiency the *zpg* promoter, by a lack of recombination at the target loci, or by inefficiency of gRNAs. To distinguish among these possibilities, we generated females carrying single heterozygous copies of the three unlinked transgenes *SagGD*^{zpg}, *Lp::Sc2A10* and a non-fluorescent (puromycin resistance-marked) *vasa-Cas9*. The progeny of these triple-transgenic females crossed to wild-type not only showed markedly better homing rates (85% GFP inheritance) but also evidence of homing in the zygote (6% progeny appeared to be GFP homozygous) (Fig. 4B). Thus, *vasa-Cas*9 was able to significantly rescue the modest homing rate of *zpg-Cas9* from *SagGD*^{zpg}, resulting in an efficient split GD. This suggests that the cloned *zpg* promoter sequence is sensitive to positional effects, being poorly active when inserted in *Saglin* or *Lp* in contrast to other previously reported host loci such as *DSX* or *Nudel* (Kyrou et al., 2018; Hammond et al., 2021).

Tracking one mosquito population of *SagGD*^{zpg} combined with *Lp::Sc2A10* (initial allele frequencies: 33% and 25%, respectively) over 7 generations showed maintenance of both transgenes over time, at elevated frequencies (G7: GFP allelic frequency=59.2%, phenotypic expression of DsRed in >90% of larvae, n=4282 larvae), indicating that the modest GD strength of *SagGD*^{zpg} was still able to increase the frequency of the *SagGD*^{zpg} allele over time and counteract the *Lp::Sc2A10* fitness cost.

Driving Lp::Sc2A10 distantly from Saglin with a vasa-Cas9 gene drive

In order to achieve a stronger GD that may push *Lp::Sc2A10* to fixation, we re-built the *Saglin* GD construct with the *vasa* promoter to express *Cas9* (*SagGD*^{vasa}, Fig. 3 and Suppl. File 1D), in spite of its known maternal deposition potentially resulting in undesired zygotic and somatic mutation at gRNA target loci.

About 35 surviving G0 injected male mosquitoes (which were homozygous for *Lp::Sc2A10* to prevent any premature damage on *Lp*), outcrossed to WT females, yielded about 20 DsRed positive G1 females and no positive male, again an indication of proper insertion of the GD cassette onto the X chromosome, where *Saglin* is located. Interestingly, these G1 females (heterozygous at both *SagGD*^{vasa} on the X chromosome and *Lp::Sc2A10* on the third chromosome) developed normally to adulthood and were fertile when crossed to wild-type males. High homing rates at both loci were recorded in G2 by examining the progeny of the individual G1 females crossed to WT, a large

majority of progeny inheriting both a *SagGD*^{vasa} copy (DsRed$_{NLS}$) and a *Lp::Sc2A10* copy (GFP) (Table 3).

Female #	Total larvae	Negatives	GFP only	DsRed only	GFP+ DsRed	GFP inheritance (*Lp* locus)	DsRed inheritance (*Saglin* locus)
1	89	0	0	0	89	100 %	100 %
2	52	0	4	0	48	100 %	84.6 %
3	59	0	0	0	59	100 %	100 %
4	59	0	6	0	53	100 %	79.6 %
5	50	0	5	0	45	100 %	80 %
6	57	0	3	0	54	100 %	89.4 %

Table 3: The *SagGD*^{vasa} GD shows high homing rates in the G2 generation at both the *Lp* and *Saglin* loci. Six individual *SagGD*^{vasa} G1 females mated to WT oviposited in individual tubes and their larval progeny was scored visually for GFP and DsRed fluorescence.

We also sought to examine the homing rates at the two loci at a larger scale, using COPAS flow cytometry to analyse the progeny from [*SagGDvasa*/Y ; *Lp::Sc2A10*/+] males crossed *en masse* to wild-type females, and from [*SagGDvasa*/+ ; *Lp::Sc2A10*/+] females crossed *en masse* to wild-type males. The male cross yielded only 405 larvae, of which 391 were GFP positive (GFP inheritance: 96.5%, corresponding to a homing rate of 93%). *DsRed* on the X chromosome was, as expected, passed on to 50% of the progeny (daughters). The female cross yielded 5197 larvae showing 95.1% *GFP* inheritance and 80.7% *DsRed* inheritance, corresponding to homing rates of 90.2% at the *Lp* locus and 81.4% at the *Saglin* locus. These G0 crosses were the funders of Population 7 and 8 further monitored below (Fig. 5).

Thus, in contrast to *zpg*, the *vasa* promoter was highly active in germ cells when inserted in the *Saglin* gene. The selected gRNAs, or homing at the *Lp* locus, appeared to be more effective than at the *Saglin* locus. Efficient homing lifted doubt on the activity of *Lp* and *Saglin* gRNAs in the GD constructs, or on the propensity of the *Saglin* locus to undergo homologous recombination.

To verify that *SagGD*^{vasa} abolished *Saglin* expression, we dissected salivary glands from control and homozygous DsRed+ mosquito females, and subjected them to western blotting using anti-Saglin antibodies. As expected, Saglin was undetectable in *SagGD*^{vasa} mosquitoes (Fig 3, inset).

Temporal dynamics of the SagGD^{vasa} *and* Lp::Sc2A10 *transgenes*

We initially established two independent populations of [*SagGD*^{vasa}; *Lp::Sc2A10*] mosquitoes, with a starting frequency of the autosomal *Lp::Sc2A10* transgene of 25%, whereas X-linked *SagGD*^{vasa} initial frequency was 33.3% (wild-type females crossed to *SagGD*^{vasa}/Y; *Lp::Sc2A10*/+ G2 males). We maintained them for >21 generations and, after 10 and 16 generations, derived six additional populations from the basal ones by outcrossing to wild type. To examine the long-term dynamics of both transgenes, we exploited GFP fluorescence from *Lp::Sc2A10* and DsRed fluorescence from *SagGD*^{vasa}

to COPAS-analyze successive generations of all populations, with large sample sizes. COPAS analyses allowed not only to distinguish negative from transgenic larvae but also, at least partially, heterozygotes from homozygotes of each transgene (see snapshots of COPAS diagrams in Suppl. File 2). Visual examination of the successive COPAS diagrams provides an accurate sense of the dynamics of both transgenes generation after generation, and an appreciation of sample sizes. Figure 5 offers a graphical representation of the same temporal dynamics, after extracting approximate percentages of larvae of each genotype from the COPAS diagrams. In all populations, mosquitoes lacking *Lp::Sc2A10* disappeared rapidly, and all young populations rapidly consisted of a majority of homozygous *Lp::Sc2A10* mosquitoes (Figure 5 and Suppl. File 2). Despite the absence of physical linkage between the two transgenes, DsRed-only individuals were strikingly absent from all populations at early generations, indicative of the toxicity of the *SagGD*^{vasa} transgene in the absence of a protected *Lp* allele. Importantly, toxicity of *SagGD*^{vasa} was due to its destructive action on the *Lp* locus rather than to a fitness cost of its disrupted *Saglin* host locus, since individuals carrying both DsRed and GFP thrived in all populations. Several populations (notably populations 1, 3, 4, 5, 6) rapidly seemed fixed for *Lp::Sc2A10*. A trend for reversal, with *Lp::Sc2A10* progressively giving way to GFP negative individuals (presumably functional *Lp* mutants refractory to homing), was observed from G16 on for population 2, although GFP positive individuals still represented a large majority of the population (92% at G21). For the *Saglin* locus, COPAS diagram interpretation must take into account the position of this gene on the X chromosome, with male mosquitoes bearing a single copy and appearing in the same larval cloud as heterozygous females (dosage compensation not being apparent at this locus). Drive will be slower at this locus as it can occur only in females. While the majority of mosquitoes in all populations rapidly carried at least one copy of *SagGD*^{vasa}, we observed that a substantial fraction of DsRed negative individuals (15 to 30% depending on the population), likely carrying GD refractory *Saglin* mutants, persisted in all populations except population 4 (Figure 5 and Suppl. File 2).

Detection of homing-refractory Saglin *mutants*

As a non-essential gene, *Saglin* is likely to accumulate mutations in evolving mosquito populations due to failed homing events of the GD construct associated with NHEJ repair at the three gRNA target sites. Indeed, DsRed negative mosquitoes (lacking the GD construct in *Saglin*) persisted at a relatively stable frequency of approximately 30 % in most experimental populations (Fig. 5). To characterize *Saglin* mutants, we COPAS-extracted 750 and 150 DsRed-negative larvae from a surplus of larvae from generation 4 of population 1 and generation 3 of population 2. DsRed-negative individuals represented 18 and 31 % of these populations, respectively, with the *Saglin* gene on their X chromosome exposed to between one and four rounds of Cas9 activity. We PCR amplified a *Saglin* fragment spanning the 3 gRNA target sites, and subjected the amplicon to high-throughput sequencing. While the majority of sequenced DsRed negative alleles still corresponded to wild-type *Saglin* having thus far escaped GD activity, a wide variety of *Saglin* mutations, different between the two sampled populations, was readily identified (Fig. 6a,b). Interestingly, a majority of these *Saglin*

haplotypes were mutated only at the target site of gRNA2, suggesting that this gRNA was the most active of the three and that Cas9, confronted to a pool of three *Saglin* and four *Lp* gRNAs, will not necessarily cleave all target sites in a given germline cell. However, combinations of mutations at gRNA1, gRNA2 and/or gRNA3 target sites were also identified. Several haplotypes were consistent with iterative action of Cas9 in lineages of haplotypes, a given mutation occurring either singly, or associated to one or two additional mutations at the other gRNA target sites. Contrarily to our expectations, deletions spanning two gRNA target sites were rare in our samples (one instance of such a haplotype was detected, carrying a deletion between gRNA1 and gRNA2). Overall, Cas9 was not highly active at the Saglin locus, some target sites being missed at each generation and combinations of mutations accumulating iteratively as germ cells were exposed to Cas9 in successive generations. Therefore, a multiplex gRNA GD experiencing failed homing at a given generation resulting in mutation of one target site may still retain a potential for successful homing in later generations, as long as all gRNA targets are not yet mutated. Inexorably however, *Saglin* mutants fully refractory to the GD will form; we recovered at least one haplotype with mutations in all three target sites after 4 generations (Figure 6a).

Appearance of a GD-refractory Lp mutant

NHEJ mutants in *Lp* were expected to emerge at much lower frequency than *Saglin* mutants in evolving populations, given that *Lp* is an essential gene under strong selective pressure. While tracking the frequency of transgenes in [*SagGD*^{vasa}; *Lp::Sc2A10*] mosquitoes, we noted the appearance of a few GFP negative larvae in populations 2 and 3. These amounted to 1.5% of total larvae in generation 16 of population 2. We COPAS-isolated a sample of these larvae and grew them to adulthood, expecting that they might die during development due to unrescued deleterious mutations in *Lp*. However, these GFP negative mosquitoes developed to healthy and fertile adults. Sequencing the *Lp* gene in 12 individuals revealed that all were homozygous for the same array of mutations within three of the four *Lp* gRNA target sequences (Figure 6). One mutation removed the PAM of gRNA2 and deleted a single valine from the Lp signal peptide (MWVLGGRRLLWSFL<u>V</u>SLVLIQSVSA, missing valine underlined). The next mutation is a C to T substitution 3 bases before the PAM of gRNA3, and is silent. The last is a 5-base deletion within the target of gRNA4 in the *Lp* first intron. Viability of the homozygous mutant indicates that this combination of mutations, while rendering *Lp* refractory to Cas9-mediated gene conversion, preserved *Lp* function. This showed that (i) the first of the four gRNAs expressed from the tRNA-gRNA array was inactive; and (ii) combinations of mutations preserving the function of target essential genes can still emerge despite the cumulative activity of three gRNAs.

*Vector competence of [*SagGD^{vasa}; Lp::Sc2A10*] mosquitoes*

We expected that the combination of Sc2A10 expression with *Saglin* loss-of-function would enhance the *Plasmodium* transmission-blocking phenotype. To assess this, we

mixed *[SagGD^{vasa}; Lp::Sc2A10]* mosquito larvae from the 20th generation of population 2 with an equal number of non-fluorescent wild-types. The resulting cage of adults was blood-fed on a mouse infected with *Pb-PfCSP*$^{hsp70\text{-}GFP}$. Seventeen days later, females carrying visible GFP parasites were selected and split into six cups of ten transgenic, and four cups of ten wild-type. All four mice exposed on day 18 to the infected control females developed parasitemia, whereas all six mice exposed to infected transgenics remained negative (Suppl. File 2). Six additional mice were exposed to 8-12 bites from *[*SagGD^{vasa}; Lp::Sc2A10*]* oocyst-carrying females of the 4th generation of population 6, which derived from *[*SagGD^{vasa}; Lp::Sc2A10*]* males of the 16th generation crossed to WT . Only two of these six mice became infected (Suppl. Table 2). While the small scale of these experiments cannot strictly confirm whether the *Saglin* knockout GD enhanced protection against transmission (especially as we discarded less-infected mosquitoes, which are expected to result from *Saglin* loss-of-function), the absence of infection in 10/12 mice showed that *[SagGD^{vasa}; Lp::Sc2A10]* mosquitoes maintained a decreased ability to transmit *Plasmodium* after >20 generations of existence.

DISCUSSION

We hijacked an essential gene, *Lipophorin* (*Lp*), to force *Anopheles gambiae* mosquitoes to co-express an anti-malarial factor. The choice of the *Lp* gene was motivated by its exceptionally strong expression level (Lp being the most abundant constitutively expressed protein in mosquito hemolymph) and by its endogenous signal peptide, which we exploited to obtain concomitant secretion of Sc2A10 with Lp. The existence of an endogenous proteolytic cleavage site within the Lp protein, that we duplicated to detach Sc2A10 from Lp, was also advantageous, as well as the natural intron immediately following the signal peptide-encoding exon 1 that could accept a visual selection marker for easy tracking of the genetic modification. This approach is related to the integral gene drive strategy (Hoermann et al., 2021) in that anti-malarial factors are expressed together with an endogenous gene, but here the modified endogenous gene does not comprise any component of a gene drive —in the case of integral gene drives, gRNAs are inserted in a synthetic intron. The genetic modification does not disrupt expression or function of the *Lp* essential gene. On the contrary, Lp protein expression is intended to remain as unaffected as possible, but obligatorily proceeds through translation of the Sc2A10 antimalarial factor inserted between the Lp signal peptide and the rest of the Lp protein. This design minimizes the risk of spontaneous transgene loss by mutation, since frameshift or nonsense mutations in the *Sc2A10* coding sequence would prevent *Lp* expression, resulting in their elimination by natural selection. Furthermore, the absence of repeated motifs in the synthetic sequence decreases the risk of in-frame deletions. The possibility of a loss of the transgene's anti-*Plasmodium* activity should therefore be limited to missense mutations or in-frame deletions arising in the *Sc2A10* coding sequence, or to *Plasmodium* evolution, notably under selection pressure imposed by the transgene itself (Marshall et al., 2019). However, the 2A10 antibody targets a highly repeated epitope on the CSP protein (the NANP motif), that in addition to being repetitive is also conserved across *P. falciparum* strains. A combination of mutations leading to the complete loss of this epitope therefore

seems unlikely, and would probably cause a dramatic loss of *Plasmodium* fitness as CSP is required in many processes in the sporozoite's journey between two hosts (Balaban et al., 2021; Coppi et al., 2011). Hijacking Lp and its signal peptide is an approach that could be widely applicable to other possible factors acting in the hemolymph to attack various parasite stages, such as ookinetes, which represent a particularly vulnerable stage of parasite development in the mosquito. To make mosquitoes that express and combine additional secreted factors against parasites, *Vitellogenin* (AGAP004203) could also be engineered in the same manner, taking advantage of its own signal peptide, extremely high expression level but also transient induction following a blood meal. Blood meal inducibility would be particularly suitable for antimalarial factors specifically targeting ookinetes, as they emerge on the hemolymph-bathed side of the midgut exactly within the short time window when *Vitellogenin* is expressed. Mosquitoes would not be exposed during their development to a factor expressed from *Vg*, limiting potential fitness costs. As proposed by Hoermann et al. (2021) *Carboxipeptidase* (*Cp*) represents an additional attractive host locus for antimalarial factors, sharing many of the advantages of *Vg*. A single GD expressing multiplexed gRNAs could stimulate the simultaneous spread of *Lp, Vg and Cp* modified, anti-parasitic alleles.

In spite of the high Lp protein secretion in mosquito hemolymph, we only detected less than 1/10 of this molecular amount for Sc2A10 in hemolymph by mass spectrometry. This could be explained by faster removal from the hemolymph and /or degradation of Sc2A10 compared to Lp. Different, more stable anti-Plasmodium peptides that could accumulate to higher levels may show even better efficacy as transmission-blocking agents compared to Sc2A10. This also highlights the benefit of choosing a strong promoter such as that of *Lp* to ensure maximal production of antimalarial effectors in the hemolymph. A single copy of the *Lp::Sc2A10* transgene still showed a reduction in transmission. Of note, the rodent model of malaria used here (*P. berghei* expressing CSP from *P. falciparum*) is characterized by very high infection rates, typically producing dozens of oocysts per midgut, each of which can release thousands of sporozoites in mosquito hemolymph. The ability of secreted Sc2A10 to decrease sporozoite transmission in this high infection context is highly encouraging, given the much lower levels of *P. falciparum* (typically 1 to a few oocysts) in field infections. An even greater anti-*Plasmodium* activity of Sc2A10 may be reached by fusing it to antimicrobial peptides such as CecropinA, as achieved by Isaacs et al. in conventional transgenic *Anopheles stephensi* mosquitoes (Isaacs et al., 2011; Isaacs et al., 2012) and subsequently in the first published GD for *A. stephensi* (Gantz et al., 2015). In these studies, the Sc2A10::CecropinA fusion was expressed intermittently under the control of the blood-meal inducible *Vitellogenin* promoter, and combined with a second scFv targeting *Plasmodium falciparum* ookinete Chitinase 1 expressed under the control of the blood meal inducible *Cp* promoter. While the insertion of *Sc2A10::CecropinA* within *Lp* coding sequences could further improve the efficiency of our approach, there is a risk that CecropinA expressed at all developmental stages, and an increased complexity of the Lp fusion protein, worsen the fitness cost already observed in *Lp::Sc2A10* mosquitoes.

To force the spread of the *Lp::Sc2A10* transgene in mosquito populations, we tested several gene drive (GD) designs. A same-locus suppression drive (installed in the *Lp* gene) was unsuccessful, as *Lp* proved largely haplo-insufficient. Distant-locus GDs disrupting *Saglin*, a non-essential but pro-parasitic gene, were efficient at promoting invasion of both the *Saglin* and *Lipophorin* transgenic alleles. With this design, *Cas9* and gRNA-coding genes may persist indefinitely within the invaded mosquito population. Although for this reason we initially preferred the same-locus design, in which insertion of CRISPR/Cas9 components abolished *Lp* function thereby causing their progressive disappearance by natural selection, the added benefit of a GD installed in *Saglin* is that it should further reduce the transgenic mosquitoes' vectorial capacity, as a combination of two distinct mechanisms to decrease sporozoite transmission will be more robust to durably combat *Plasmodium*. The genetic design established here can be viewed as a two-component GD where one component (*Lp::Sc2A10*) has no drive capacity by itself but (i) rescues the lethality and (ii) hitchhikes on the drive capacity of the other ($SagGD^{vasa}$) ensuring its own hyper-Mendelian spread. The two transgenic loci are, therefore, reciprocally dependent on each other: *Lp::Sc2A10* depends on *SagGD* for its long-term persistence and spread in a population, *SagGD* depends on *Lp::Sc2A10* for its survival. Importantly, we observed that same-locus or distant-locus GDs carrying guide RNAs that target wild-type *Lp* are inviable, and rapidly go extinct in the absence of a recoded *Lp* allele like *Lp::Sc2A10* (or spontaneous *Lp* mutants immune to GD that may arise in GD populations). Since $SagGD^{vasa}$ will destroy mosquitoes not carrying *Lp::Sc2A10,* this system could be expected to initiate population replacement with a large proportion of laboratory-derived genetic background, rather than population modification by introgression of the transgene in a field genetic background ; but this is tempered by (i) viability of dually heterozygous hybrids and (ii) the moderate fitness cost of *Lp::Sc2A10* that can allow it to diffuse in the field in a Mendelian way independently of $SagGD^{vasa}$, leading to influx of field genetic background in the resulting final population. Because *Saglin* knockout mutants bear no fitness cost at least in laboratory conditions, GD-refractory *Saglin* alleles generated by failed homing events are certain to accumulate over time (all the more so as their lack of toxicity confers them a high selective advantage over the *Saglin* GD that is toxic towards WT *Lp*). If the demise of the *Saglin* GD occurs sufficiently late, *Lp-Sc2A10* could still reach high levels or even fixation in a mosquito population. If, however, GD refractory mutations in *Saglin* accumulate rapidly and no longer support *Lp::Sc2A10* maintenance, the elimination of both $SagGD^{vasa}$ and *Lp::Sc2A10* will be inexorable. Thus wild-types alleles are likely to re-establish at both loci once the frequency of active *Saglin* GDs drops, especially under conditions of strong mosquito immigration from regions adjacent to the initial GD spread range. Still, even as genetic resistance to the GD develops, the resulting high frequency of Saglin loss-of-function mutant alleles arising from failed homing events, with their transient selective advantage as long as $SagGD^{vasa}$ and wild *Lp* alleles co-exist, should decrease vector competence. Once a GD installed in the *Saglin* locus is defeated by the accumulation of GD-refractory mutations, new generations of GD installed at different loci could be launched to re-ignite the spread of the *Lp-Sc2A10* allele. For instance, GDs built for the same mosquito species in the frame of other projects could also incorporate gRNAs re-igniting the spread of *Lp::Sc2A10*, promoting antimalarial approaches similar to combination therapy (Ellis et al., 2022). In this respect, transgenic systems permitting the expression of multiplexes of many gRNAs would be useful. The constructs

presented here incorporate multiplex gRNAs combining individual U6-gRNA cassettes with a U6-gRNA::tRNA quadruple gRNA expression cassette, with the drawback of displaying many repeated motifs. We are currently working to improve this design by increasing the number and variety of tRNA spacers, so that the 76-bp constant region of each gRNA is the only remaining repeated element. Other pathways to improvement of the system presented here include the use of promoters better restricted to the germline to control Cas9 expression and limit GD refractory mutation emergence. Interestingly, the *zpg* promoter that in theory possesses this advantage showed low activity once inserted in the *Lp* and *Saglin* loci, in contrast to the *DSX* or *Nudel* loci (Hammond et al., 2016; Kyrou et al., 2018). This suggests that the *zpg* promoter is more sensitive to the local chromatin context than the *vasa2* promoter, and that its use in GD constructs may be restricted to a subset of genomic loci. The *nanos* promoter (Gantz et al., 2015; Terradas et al., 2021; but see Hammond et al., 2021) may represent an interesting alternative.

A simpler design to promote the spread of antimalarial factors such as Sc2A10 in mosquito populations would be a single-locus GD (for example installed in *Saglin* or other pro-parasitic genes if the idea of consolidating the anti-*Plasmodium* activity via two mechanisms is retained). Few loci, however, can provide the benefits offered here by *Lp* that include a suitable expression pattern, high expression, an endogenous secretion signal, endogenous proteolytic cleavage, a natural intron ideally positioned to host a fluorescent marker, and the status of essential gene exploited by the described translational fusion to limit spontaneous transgene loss by mutation. Besides, adding the *Cas9* and gRNA-coding cassettes to an intron of an essential gene chosen as GD host would add >5 kb of foreign DNA sequence, thereby significantly increasing the complexity of the modification and the risk of affecting functionality of the essential gene. For example, inadvertent intron splice sites in synthetic sequences can preclude proper mRNA splicing (Hoermann et al., 2022); we also encountered this problem when inserting a different anti-*Plasmodium* ScFv in *Lp* (Green, 2019).

In conclusion, the *Sag-GD*^{vasa} transgene presented here constitutes a viable GD only in the presence of an engineered *Lp* allele, which by itself has no drive capacity but will hitchhike with *Sag-GD*^{vasa} to spread at hyper-Mendelian rates. In caged mosquito populations, both genetic loci initially spread together and decrease mosquito capacity to transmit *Plasmodium* expressing CSP from *P. falciparum*. Mutations at both loci, especially in *Saglin* that is under low selection pressure, arise due to failed homing events, have a selective advantage over the toxic *SagGD*^{vasa} allele, and over time will replace the GD, so that the system carries the seeds of its own removal. The anti-*Plasmodium Lp-Sc2A10* allele will slowly disappear due to its fitness cost, unless its spread is re-ignited by updated distant-locus GDs. Though such a reversibility could be seen as a safety advantage for modification GDs (and may be enhanced by intentionally releasing laboratory-selected resistance alleles), future work should focus on establishing GD constructs in *Saglin* that limit the emergence of GD refractory alleles to prolong their potential field lives.

Materials and Methods

Plasmid construction

We used Golden Gate Cloning (Engler and Marillonnet, 2014) to assemble all parts of the *Lp* (AGAP001826) knock-in plasmid shown in Fig. 1B in destination vector pENTRR4-ATCC-LacZ-GCTT (Addgene# 173668). The full annotated plasmid sequence is provided in Suppl. File 1B. The recoded beginning of *Lp*, the Sc2A10 coding sequence and intronic 3xP3-GFP-Tub56D marker cassette were ordered as synthetic DNA gBlocks (IDT DNA, Belgium) flanked by appropriate *Bsa*I restriction sites. The Lp endogenous Furin cleavage site (RFRR) was replicated along with 3 N-terminal and 4 C-terminal flanking amino acids between the Sc2A10 and ApoLpII amino acid sequences. This, and retaining 7 endogenous nucleotides upstream of the intron 5' splice junction after the *Sc2A10* coding sequence, resulted in the addition of 7 non-natural amino acids (GIRESAA) to the N-terminus of the ApoLpII moiety. 1520 bp of 5' and 1027 bp of 3' flanking sequence were cloned on either side of the modified *Lp* sequence to be knocked-in. Adjacent to the 3' *Lp* homology arm, we cloned three guide RNA (gRNA) expression modules (U6 promoter — gRNA coding sequence — terminator) that specify Cas9 cleavage within the first exon and intron of the genomic (but not recoded) *Lp* sequence. The three gRNA-expressing modules were prepared in pKSB-sgRNA1—3 (Addgene #173671—173673) as described (Dong et al., 2018) with gRNAs targeting the motifs: GAATCAGCACTAGCGACACCAGG, GCGTGAAATATCGTCAGGGATGG and GGAACGATGTGGGTCCTCGGTGG (PAMs underlined) in *Lp*. One gRNA target site (with PAM) was included on the 3' flanking homology arm's distal extremity to favor linearization of the donor plasmid after egg microinjection. This donor plasmid was called *pENTR R4 3xgRNA Lp::Sc2A10*. Plasmids encoding gene drive components for insertion into the *Lp* and *Saglin* sequences were assembled by Golden Gate Cloning in the same manner, including *promoter-Cas9-SV40* and 3xP3-$DsRed_{NLS}$-SV40 modules, the sequences of which are provided in the complete plasmid sequence (Suppl. File 1A-D). The *zpg-Cas9* cassette in $SagGD^{zpg}$ was subcloned from p17410 (Kyrou et al., 2018). To obtain a puromycin resistant, non-fluorescent, Cas9-expressing mosquito line we constructed *att*B docking plasmid pDSAP-*vasa-Cas9-SV40* by Golden Gate Cloning of a *vasa* promoter module and of a Cas9-SV40 module into plasmid pDSAP (Volohonsky et al., 2015).

Mosquito egg microinjection and recovery of transgenics

Embryo microinjection to obtain the *Lp::Sc2A10* transgene by CRISPR-Cas9 knockin was performed as for classical transgenesis (Volohonsky et al., 2015) with 400 ng/µl of donor plasmid DNA and addition of 2 µM of the drug Scr7 (APExBio) to the injected plasmid mix. Transgenic *Anopheles coluzzii* mosquitoes of the Ngousso genetic background expressing the eSpCas9 variant with reduced off-target activity (Slaymaker et al., 2016) in the germ line under control of the *vasa2* promoter (Papathanos et al., 2009) were used as a source of eggs for microinjection (vasa-eSpCas9, 3xP3-DsRedNLS strain). We injected the knock-in plasmid into 400 eggs, recovered 140 surviving G0 adults and out-crossed them *en masse* to wild-type mosquitoes. In the G1 progeny of the injected G0 males, about 100 GFP positive, putative knock-in individuals were recovered. We established independent families from ten GFP positive G1 females

outcrossed to wild-type, and counter-selected the Cas9 transgene (marked with 3xP3-DsRed) in these families. All 10 founder females were confirmed by PCR to carry the intended transgene insertion in the target Lp locus. We pooled two families to yield the *Lp::Sc2A10* line. To avoid confusion between two DsRed markers, injection of the *Lp-GD* suppression GD construct (DsRedNLS-marked) was performed in eggs expressing wild type Cas9 (vasa-Cas9, 3xP3-YFP G3 strain). Injection of the *SagGD*^{zpg} and *SagGD*^{vasa} constructs was performed in embryos lacking any Cas9 transgene, but respectively carrying a heterozygous or homozygous copy of *Lp::Sc2A10*. Numbers of injected eggs and recovered transgenics are provided in the main text. A puromycin resistant, non fluorescent Cas9-expressing mosquito strain was obtained by inserting plasmid pDSAP-*vasa-Cas9-SV40* into the attP docking site of *A. gambiae* line X1 as described (Volohonsky et al., 2015). This line was balanced over a YFP fluorescent marker inserted in the same locus by crossing to the YFP vasa-Cas9 line, with selection of the F1 progeny on puromycin as described (Volohonsky et al., 2015), and made homozygous by counter-selecting YFP fluorescence in the F2 progeny of the self-crossed puromycin resistant F1.

Plasmodium berghei Pb-PfCSP$^{hsp70\text{-}GFP}$ parasite strain

We generated a novel *P. berghei* strain combining the strong GFP fluorescence of *Δp230p-GFP* (Manzoni et al., 2014) with the substitution of endogenous *P. berghei* CSP for CSP from *P. falciparum* from strain *Pb-PfCSP* (Triller et al., 2017), in which GFP fluorescence was too faint for selecting live infected mosquitoes examined under 488-nm light. For this, 150 mosquito females were allowed to blood feed on a mouse co-infected with both parental parasite strains at a ratio of 1 strongly fluorescent Δ*p230p*-GFP for 40 *Pb-PfCSP*. Sexual reproduction of *P. berghei* in the mosquito generates hybrid haploid sporozoites, some of which inherit both Δ*p230p*-GFP and *Pb-PfCSP*. 17 days after their infective blood meal, live female mosquitoes were screened under 488-nm light to select those displaying visible GFP sporozoites trapped at the base of their wing veins. 20 positive females were offered a blood meal on a naïve mouse. When parasitemia reached 0.1%, 3000 strongly GFP positive blood stage parasites were sorted by flow cytometry and injected intravenously into two naïve mice, only one of which developed parasitemia 11 days after passage. Its blood was then passaged into a new mouse. When parasitemia reached 0.4%, 10 new mice were injected with a blood dilution corresponding to 1 parasite each, although this number was probably underestimated as all 10 mice developed parasitemia. Of these, four mice tested PCR positive for *PfCSP* (PCR primers GGCCTTATTCCAGGAATACCAGTGCT / GGATCAGGATTACCATCCGCTGGTTG) and negative for PbCSP (PCR primers GAAGAAGTGTACCATTTTAGTTGTAGCGTC / TGGGTCATTTGGGTTTGGTGGTG). We selected one clone for passage into naive mice, confirmed its PCR negativity for PbCSP and positivity for PfCSP, this clone was called *Pb-PfCSP*$^{hsp70\text{-}GFP}$.

Mosquito infections and bite-back experiments

Neonate larvae from the heterozygous stock of *Lp::Sc2A10* mosquitoes were sorted by COPAS flow cytometry (Marois et al., 2012) to assemble a mix of equal numbers of homozygous GFP+ larvae (or heterozygous, in some experiments) and negative control siblings. Maintaining the transgene in a heterozygous population ensured the absence of genetic divergence that otherwise may change vector competence between control and transgenic mosquitoes over time. Following COPAS sorting, transgenic and control larvae were co-cultured, and adults kept in the same cage, to equalize potential environmental influences (including unequal microbial communities) that could differentially affect the vector competence of the two genotypes if grown separately. Mosquitoes were infected together by blood-feeding from the same mouse with a parasitemia between 2 and 4%. 16-19 days after infection, mosquitoes were anesthetized and *Lp::Sc2A10* females separated from negative control females based on fluorescence of the transgenesis marker in their eyes. In addition, only *Plasmodium*-carrying mosquitoes were retained, by screening for mosquitoes with GFP-positive midgut oocysts and wing vein or salivary gland sporozoites visible through their cuticle. On day 17-20 after infection, naïve mice were individually exposed to groups of 10 *Lp::Sc2A10* or negative control infected mosquitoes. The number of engorged mosquitoes was recorded after infectious feeding, only mice bitten by at least 6 mosquitoes were included in subsequent analyses with the exception of two mice bitten by 5 and 4 highly infected transgenic and control mosquitoes, respectively, in one experiment in which the mouse bitten by the 4 control mosquitoes became infected. Mouse parasitemia was monitored by flow cytometry using an Accuri™ flow cytometer between day 4 and 12 after bite-back. Mice reaching >1.5% parasitemia were sacrificed, and mice remaining negative 12 days post infectious feeding were considered uninfected.

Mass spectrometry analysis of hemolymph

Hemolymph from 25-30 cold-anesthetized female mosquitoes was collected directly into 1x Laemmli buffer 48 hours post-blood feeding by clipping the proboscis and gently pressing their abdomen. Samples were precipitated and digested with trypsin, and 1/5 of the digestion product was analyzed on a Q Exactive Plus Mass Spectrometer coupled to an Easy-nanoLC1000 (Thermo). The acquired data was searched against the *Anopheles* UniProt database plus the Sc2A10 protein sequence using Mascot and the total number of spectra corresponding to ApoLpI, ApoLpII, Sc2A10 and a selection of additional hemolymph proteins that served for normalization were counted.

Protein gels and Western blotting

Mosquito salivary glands or carcass were dissected in 20 µl Laemmli buffer, crushed with a pestle and denatured at 65 °C for 5 minutes. Samples were centrifuged for 3 min at 15,800 g, loaded on a Mini-PROTEAN TGX Stain-Free Precast Gel (BioRad) along with 6 µL of PageRuler Plus Prestained Protein Ladder (ThermoFisher) as standard and electrophoresed at 170 V using the Mini Trans-Blot cell system (BioRad). Gels were

blotted on PVDF membranes (Trans-Blot Turbo Mini 0.2 µm PVDF Transfer Pack; BioRad) using the mid-range program of a Pierce Fast blotter (Thermo Fisher Scientific). Membranes were blocked for one hour in PBS + 0.1% Tween (PBST) supplemented with 5% fat-free milk powder, incubated overnight at 4°C with primary antibody, washed three times in PBST and incubated one hour at room temperature in secondary antibody conjugated to horseradish peroxidase (HRP). Antibodies were diluted in PBST, 3% milk. Membranes were then washed three times for 10 minutes with PBS. Antibody binding was revealed using the Super signal WestPico Plus kit (ThermoFisher). After a 1-2 minute incubation, images were acquired using the Chemidoc software (Biorad). Before incubation with further primary antibodies to visualize additional proteins, membranes were stripped for 20 to 30 minutes in Restore PLUS Western Blot Stripping Buffer (ThermoFisher) and washed three times for 10 minutes in PBST followed by a new blocking incubation. Saglin polyclonal antibody and recombinant Saglin protein sample were a kind gift from Eric Calvo (NIAID, Rockville, MD). Human antibodies used to control protein loading of salivary gland samples were obtained from 2 ml venous blood taken from a human volunteer arm-feeding *Anopheles* cages over several years. Blood was incubated at room temperature for 1 hr to allow coagulation and centrifuged for 10 min at ~16.000 g. Serum supernatant was aliquoted and stored at -20°C until use. For Coomassie staining of electrophoresed hemolymph samples, hemolymph was collected 42h post blood feeding by clipping the proboscis of cold-anesthetized mosquitoes and gently pressing their abdomen, collecting drops of hemolymph into a pipet tip containing protein sample buffer (1 µl of sample buffer per mosquito). Samples were denatured for 2 minutes at 90°C and hemolymph from 10 mosquitoes was loaded on each lane of a 4 - 20% 1.5mm polyacrylamide gel (Invitrogen). After electrophoresis, gels were rinsed in water and stained in SimplyBlue safeStain solution (ThermoFisher).

COPAS analyses

A COPAS Select instrument (Union Biometrica) was used as described (Marois et al., 2012; Bernardini et al., 2014) to quantify fluorescent transgenes in populations of neonate mosquito larvae or to establish mixed cultures of defined proportions. The use of unfed neonate larvae in clean water was crucial for precise gating and sorting. For sorting, flow rate was kept under 20 larvae/second in Pure mode allowing superdrops. PMTs were set at 500 (GFP), 600-900 (RFP). Delay, Width, signal threshold and minimum TOF parameters were set on 8, 6, 100 and 150, respectively.

Amplicon sequencing

A 494-bp region of *Saglin* encompassing the 3 gRNA target sites was amplified from genomic DNA extracted from 150 and 750 COPAS-sorted, DsRed negative neonate larvae with the Blood and Tissue kit (Qiagen) using PCR primers <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u>GCAGAAGCAGCTCGACGC and <u>GACTGGAGTTCAGACGTGTGCTCTTCCGATCT</u>GCAGCTGCCGGAAGTGCT (Illumina adapters underlined) and Phire DNA polymerase (Thermofisher Scientific), an annealing temperature of 68°C and 32 amplification cycles. Resulting amplicons were gel-purified

and sent for sequencing using the Genewiz Amplicon EZ service. Returned paired-end reads (250 nt length each) were assembled to form single contigs using PANDAseq version 2.11 (Masella et al., 2012); unmerged reads were manually combined pairwise and added to the assembled contigs. Data was analysed using a Docker Desktop Personal version 20.10.16 (https://www.docker.com) -based CRISPResso2 version 2.2.8 (Clement et al., 2019) followed by manual curation.

Acknowledgements

We thank the laboratory of Chris Janse for providing the parental *P. berghei-PfCSP* strain from which the *Pb-PfCSP*$^{hsp70\text{-}GFP}$ strain was derived. We thank Andrea Crisanti, Andrew Hammond and Kyrous Kyrou for the gift of plasmid p17410 to subclone the *zpg-Cas9* expression cassette. We thank Lauriane Kuhn, Johana Chicher and Philippe Hammann of the Strasbourg IBMC Proteomics Platform for the mass spectrometry sample preparation and analysis. We thank Muriel Philipps and Claudine Ebel from the Illkirch IGBMC flow cytometry platform for sorting *P. berghei*-infected red blood cells. We thank Mallory Kastner for help during the project.

Agreements

Work on mice was evaluated by the CREMEAS Ethics committee and authorized by Ministère de l'Enseignement Supérieur et de la Recherche (MESRI) under reference APAFIS #20562-2019050313288887v3. Work with genetically modified mosquitoes was evaluated by Haut Conseil des Biotechnologies and authorized by MESRI (déclaration d'utilisation d'OGM en milieu confiné #3243 and agreement #3912).

Figure legends

Fig. 1. Characterization and modification of the endogenous *Lipophorin* gene. A: Coomassie staining of electrophoresed hemolymph protein. Mosquitoes were injected with the indicated double-stranded RNA to silence *Vg* or *Lp* as described (Rono et al., 2010) and offered a blood meal after 2 days. Hemolymph was collected 42h post blood feeding. B: scheme (not to scale) depicting the insertion of synthetic sequences into the natural Lp gene. The Signal Peptide (SP) was recoded to remove gRNA target sites. Sc2A10 immediately follows the SP and ends with a RFRR furin cleavage site (red star) for separation from ApoLpII in the secretory pathway, duplicated from the ApoLpI/II natural cleavage site found downstream. C: spontaneous dynamics of *Lp::Sc2A10* transgene frequency over 17 generations. Between 1700 and 9100 neonate larvae in generations 1, 2, 3, 4, 7, 10, 11, 12, and 17, and 522 larvae in generation 16, were analysed by COPAS flow cytometry. The proportions of homozygotes (brightest GFP), heterozygotes (intermediate GFP) and GFP negative (WT) were estimated for each

analyzed generation by gating each larval population in the COPAS software. Values are provided in Suppl. Table 1. D: western blot with anti-ApoLpII antibody on hemolymph samples from 10 female mosquitoes. First and last lanes: protein size ladder (molecular weights indicated), second lane: hemolymph from WT mosquitoes. Lanes c, d: hemolymph from homozygous *Lp::Sc2A10* mosquitoes. Failed separation between scFv and ApoLpII would result in a 13 kDa molecular weight upshift compared to the control. Lanes a, b: hemolymph from mosquitoes expressing a distinct *Lp::ScFv* fusion not further discussed in this work.

Fig. 2: infection status of mice bitten by Sc2A10-expressing vs. WT mosquitoes infected with *Pb-PfCSP*$^{hsp70\text{-}GFP}$. Cumulative results from 13 independent experiments performed over a period of two years. Pie charts show the total number of infected (orange) and uninfected (blue) mice after exposure to the bite of *Lp::Sc2A10* mosquito females or their negative siblings. Aggregated data from
independent experiments were analyzed using Fisher's exact test with a two-tailed p-value.

Fig. 3: Scheme of *Saglin*-based gene drives that also promote *Lp* locus modification. A drive cassette comprising *Cas9* (under either the *zpg* or *vasa* promoter) and an array of 7 gRNA-coding modules is inserted disruptively in the endogenous *Saglin* open reading frame on chromosome X, along with a *3XP3-DsRedNLS* fluorescence marker to track the genetic modification. Four gRNAs separated by a repeated glycine tRNA are expressed under control of one U6 promoter, and target wild-type *Lipophorin* on chromosome III (symbolized by green arrow) and promote *Lp::Sc2A10* homing. Three gRNAs, each expressed under control its own U6 promoter, target wild-type Saglin (purple arrow) and promote homing of the gene drive cassette. Inset: western-blot using Saglin antibodies shows the absence of Saglin protein in the salivary glands of dissected SagGD^{vasa} homozygous females.

Fig. 4: *Sag-GD*^{zpg} triggers modest homing at the *Lp* locus, which can be rescued by *vasa-Cas9*. A: COPAS analysis of the G2 progeny from G1 *SagGD*^{zpg} ; *Lp::Sc2A*10 females crossed to WT. Weak DsRed fluorescence does not allow accurate separation of DsRed+ / DsRed- larvae. Note that GFP inheritance is only slightly higher than the 50% expected in the absence of homing. B: COPAS analysis of the progeny from triple transgenic *SagGD*^{zpg} /+; *Lp::Sc2A*10/+; *vasa-Cas9*/+ females crossed to WT. Note the higher GFP homing rate and the existence of GFP homozygotes revealing homing in early embryos rather than in the germ line, due to maternal Cas9 deposition. DsRed positive and negative larvae were indistinguishable with the COPAS settings used for this experiment.

Fig. 5: Temporal dynamics of the *Sag-GD*^{vasa} and *Lp::Sc2A10* transgenes in 8 mosquito populations. Transgenic mosquitoes carrying both transgenes were crossed to wild-type (populations 1, 2, 5, 6, 7, 8) or mixed with wild-types (populations 3, 4) and the frequency of each transgene in 1000-4000 neonate larvae of successive generations

was tracked by flow cytometry (COPAS) using the GFP fluorescent marker of *Lp::Sc2A10* and the DsRed fluorescent marker of *Sag-GD*^{vasa}. Calculated transgene frequency, taking homozygous and heterozygous larvae into account (continous lines), and percentage of fluorescent mosquitoes (dotted lines) are shown on the graphs. Frequency of each genotype in the population was calculated after gating corresponding clouds of larvae on COPAS diagrams (see Suppl. File 2) and recording the associated percentage measured by COPAS software, or by opening COPAS files and gating in WinMDI software. Gating is approximative due to partial overlap between different clouds of larvae. Supplemental File 2 shows strips of all COPAS diagrams that served to extract this data, which provide a more accurate sense of transgene dynamics over generations.

Figure 6: (a, b) Characterization of *Saglin* failed homing mutations. A PCR product spanning the three *Saglin* gRNA target sites was amplified from 150 DsRed-negative larvae from generation 3 of population 2 (a) or from 750 DsRed-negative larvae from generation 4 of population 1 (b) and subjected to high-throughput amplicon sequencing. The different *Saglin* mutant haplotypes discovered in each sample are aligned to the WT sequence (top) with target site and protospacer-adjacent motifs (PAMs) indicated. Arrows point to Cas9 cleavage sites. Deleted nucleotides are highlighted in green, inserted nucleotides in blue. // indicates portions of WT sequence not represented on the figure. The distance between two gRNA target sites (number of nucleotides) is indicated above the WT sequence. **(c) gRNA target sites and failed homing-induced mutations in the *Lp* gene.** The WT sequence of the 5' region of the *Lp* gene shows exon and intron sequences (blue and black regions, respectively), the position of gRNA target sites with PAMs, and the ATG initiator codon (bold-underlined). Nucleotides shown in red were deleted in twelve sequenced homozygous mutant mosquitoes, except for the C in the third gRNA target site that was replaced with a T in the mutants. The additional, 11-bp deletion in the intron, distant from gRNA target sites, was unexpected.

Suppl. File 2: Temporal dynamics of the *Sag-GD*^{vasa} and *Lp::Sc2A10* transgenes in 8 mosquito populations. Each horizontal strip of panels corresponds to an independent caged mosquito population. Each panel is a COPAS analysis snapshot of neonate larvae at the indicated generation. Successive panels provide a global view of each transgene's evolution on samples of 1000-4000 neonate larvae at each generation, with fluorescence intensity correlating with transgene copy number (GFP marks the *Lp::Sc2A10* transgene on the Y axis, DsRed marks *Sag-GD*^{vasa} on the X axis). Up to 9 partially overlapping distinct larval populations corresponding to the 9 possible genotypes (0, 1, or 2 copies of each fluorescence, sketched in the top panel) can be resolved on COPAS diagrams. Dots indicated by a white arrow correspond to debris (dead larvae, egg shells) and are difficult to distinguish from non-fluorescent larvae. Yellow arrow in population 2, G14 indicates the appearance of GFP negative larvae that represent a *Lp* homing refractory mutant (see text). Tracking Populations 1 and 2 started shortly after initial transgenesis, by crossing G2 [*Sag-GD*^{vasa} / Y; *Lp::Sc2A10* / +] males to WT females (considered the G0 cross for temporal tracking of genotypes). Population 3 was assembled by mixing 20 transgenics from generation 12 of population 2 with 190 WT. Population 4 was assembled by mixing 60 transgenics from generation 12 of population 2 with 189 WT. In these new G0 mixes reproductive

success of transgenics proved higher than that of the WT (due to differences in the quality of the parental mosquitoes), as G1 pupae consisted of 189 GFP transgenics for 358 negatives = 34.5% instead of expected 10% (population 3) or 346 GFP transgenics for 304 negatives = 53.2% instead of expected 24.1% (population 4). In all populations, note the initial rapid convergence of genotypes towards *Lp::Sc2A10* (GFP) homozygosity, the initial rapid disappearance of *Sag-GD*^{vasa} (DsRed) positives lacking GFP, and the persistence of a fraction of DsRed negative larvae.

Suppl. Fig. 1: Insertion of a *Lipophorin* suppression gene drive and its expected and unexpected outcomes. A: scheme of the suppression drive construct. Upon chromosome cleavage by Cas9 (shown by scissors), homologous recombination (dotted lines) via 5' and 3' regions of homology is expected to insert the intervening elements (gRNA-encoding array, zpg-Cas9 expression cassette, DsRedNLS marker cassette) into the beginning of the Lp gene, disrupting its function. Note that this recombination event knocking out the essential *Lp* gene was designed to occur in the presence of an intact *Lp::Sc2A10* sister chromosome, not shown in the scheme but contributing to the unexpected outcome shown in B. B: most knock-in events unexpectedly arose from compound recombinations between the broken WT chromosome and both the repair template plasmid and the intact Lp::Sc2A10 sister chromosome, resulting in insertion of the gene drive elements genetically linked to the *Lp::Sc2A10* transgene. Shown is a hypothetical scheme interpreting the compound insertion, inferred from the genetic linkage observed between GFP and DsRedNLS and from sequencing a PCR product obtained with primers AG105 and EM890 (indicated). Δ indicates a 500-bp deletion in the cloned Lp 3'region, which arose following micro-injection of the plasmid because one of the gRNA target sites was retained by error in this part of the construct (indicated by scissors). Presence of the gRNA-tRNA array (black arrows) was confirmed by drive activity in presence of *vasa-Cas9*.

Suppl. Fig. 2: Characterization of the two types of Lp-GD integration events. A: The GFP-RFP line showed high homing rates when complemented with an independent vasa-Cas9 transgene. COPAS diagrams show the fluorescence of progeny neonate larvae from heterozygous GFP-RFP mosquitoes (third chromosome) that also carried a non-fluorescent, puromycin resistance-marked vasa-Cas9 transgene (second chromosome), crossed to non-fluorescent partners (WT* indicates non-fluorescent partners actually carrying one copy of the puromycin resistant vasa-Cas9 transgene, not influencing the crossing outcome). N indicates the total number of larvae analyzed in the diagrams; percentages of larvae of each fluorescence are indicated. Non-fluorescent progeny would amount to 50% in the absence of homing. Note that the percentage of GFP-RFP is close to 50%, mainly reflecting Mendelian inheritance of the parental transgene. Thus most homing events involved only a GFP-containing segment from the parental GFP-RFP transgene. Rarer instances of RFP-only segment homing (arrows) were accompanied by a decrease of DsRed fluorescence intensity, indicating that the *3xP3* promoter is less active than in the GFP-RFP context. **B**: Haploinsufficiency at the *Lipophorin* locus. Photograph shows larval and pupal progeny from heterozygous *Lp-GD* crossed to heterozygous *Lp::Sc2A10*. Right panel focuses on a subset of larvae observed under fluorescent light. Larvae that inherited one copy of the *Lp-GD* loss-of-

function allele (red eyes) and one copy of *Lp::Sc2A10* (green eyes/brain) are developmentally delayed and will die before adulthood. Individuals that inherited a WT *Lp* copy and either modified allele will complete development, with strong loss of fitness for WT/*Lp-GD* individuals (red-eyed pupa).

Suppl. Table 2: mouse parasitemia after infection by mosquito bites.
Mice are grouped by categories according to the genotype of the *Plasmodium*-carrying mosquitoes biting them or by the strain of parasite infecting them. Groups of 10 infected transgenic mosquito females biting a mouse were matched by groups of 10 wild-type infected mosquitoes, siblings of the transgenics and grown in the same culture, biting a mouse. Colored cells indicate the stage at which infected mice were sacrificed to prevent the appearance of malaria symptoms. n.d= not determined. 'D' in the third column indicates a mouse bitten by transgenic mosquitoes that showed a clearly delayed parasitemia compared to the controls. In two cases (indicated in third column), mouse was bitten by fewer than 6 mosquitoes.

Bibliography

Balaban, A.E., Kanatani, S., Mitra, J., Gregory, J., Vartak, N., Sinnis-Bourozikas, A., Frischknecht, F., Ha, T., Sinnis, P., 2021. The repeat region of the circumsporozoite protein is an elastic linear spring with a functional role in Plasmodium sporozoite motility. https://doi.org/10.1101/2021.05.12.443759

Bernardini, F., Galizi, R., Menichelli, M., Papathanos, P.-A., Dritsou, V., Marois, E., Crisanti, A., Windbichler, N., 2014. Site-specific genetic engineering of the Anopheles gambiae Y chromosome. Proc. Natl. Acad. Sci. U. S. A. 111, 7600–7605. https://doi.org/10.1073/pnas.1404996111

Carballar-Lejarazú, R., James, A.A., 2017. Population modification of Anopheline species to control malaria transmission. Pathog. Glob. Health 111, 424–435. https://doi.org/10.1080/20477724.2018.1427192

Carballar-Lejarazú, R., Ogaugwu, C., Tushar, T., Kelsey, A., Pham, T.B., Murphy, J., Schmidt, H., Lee, Y., Lanzaro, G.C., James, A.A., 2020. Next-generation gene drive for population modification of the malaria vector mosquito, Anopheles gambiae. Proc. Natl. Acad. Sci. U. S. A. 117, 22805–22814. https://doi.org/10.1073/pnas.2010214117

Carballar-Lejarazú, R., Tushar, T., Pham, T.B., James, A.A., 2022. Cas9-mediated maternal effect and derived resistance alleles in a gene-drive strain of the African malaria vector mosquito, Anopheles gambiae. Genetics iyac055. https://doi.org/10.1093/genetics/iyac055

Champer, J., Kim, I.K., Champer, S.E., Clark, A.G., Messer, P.W., 2021. Suppression gene drive in continuous space can result in unstable persistence of both drive and wild-type alleles. Mol. Ecol. 30, 1086–1101. https://doi.org/10.1111/mec.15788

Champer, J., Liu, J., Oh, S.Y., Reeves, R., Luthra, A., Oakes, N., Clark, A.G., Messer, P.W., 2018. Reducing resistance allele formation in CRISPR gene drive. Proc. Natl. Acad. Sci. U. S. A. 115, 5522–5527. https://doi.org/10.1073/pnas.1720354115

Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., Pinello, L., 2019. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol. 37, 224–226. https://doi.org/10.1038/s41587-019-0032-3

Coppi, A., Natarajan, R., Pradel, G., Bennett, B.L., James, E.R., Roggero, M.A., Corradin, G., Persson, C., Tewari, R., Sinnis, P., 2011. The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host. J. Exp. Med. 208, 341–356. https://doi.org/10.1084/jem.20101488

Dong, Y., Simões, M.L., Marois, E., Dimopoulos, G., 2018. CRISPR/Cas9 -mediated gene knockout of Anopheles gambiae FREP1 suppresses malaria parasite infection. PLoS Pathog. 14, e1006898. https://doi.org/10.1371/journal.ppat.1006898

Ellis, D.A., Avraam, G., Hoermann, A., Wyer, C.A.S., Ong, Y.X., Christophides, G.K., Windbichler, N., 2022. Testing non-autonomous antimalarial gene drive effectors using self-eliminating drivers in the African mosquito vector Anopheles gambiae. PLoS Genet. 18, e1010244. https://doi.org/10.1371/journal.pgen.1010244

Engler, C., Marillonnet, S., 2014. Golden Gate cloning. Methods Mol. Biol. Clifton NJ 1116, 119–131. https://doi.org/10.1007/978-1-62703-764-8_9

Gantz, V.M., Akbari, O., 2018. Gene editing technologies and applications for insects. Curr. Opin. Insect Sci. 28. https://doi.org/10.1016/j.cois.2018.05.006

Gantz, V.M., Jasinskiene, N., Tatarenkova, O., Fazekas, A., Macias, V.M., Bier, E., James, A.A., 2015. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. Proc. Natl. Acad. Sci. U. S. A. 112, E6736-6743. https://doi.org/10.1073/pnas.1521077112

Ghosh, A.K., Devenport, M., Jethwaney, D., Kalume, D.E., Pandey, A., Anderson, V.E., Sultan, A.A., Kumar, N., Jacobs-Lorena, M., 2009. Malaria parasite invasion of the mosquito salivary gland requires interaction between the Plasmodium TRAP and the Anopheles saglin proteins. PLoS Pathog. 5, e1000265. https://doi.org/10.1371/journal.ppat.1000265

Green, E., 2019. Development of genetic control strategies for insect pests using CRISPR/Cas9 (phdthesis). Université de Strasbourg.

Hammond, A., Galizi, R., Kyrou, K., Simoni, A., Siniscalchi, C., Katsanos, D., Gribble, M., Baker, D., Marois, E., Russell, S., Burt, A., Windbichler, N., Crisanti, A., Nolan, T., 2016. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. Nat. Biotechnol. 34, 78–83. https://doi.org/10.1038/nbt.3439

Hammond, A., Karlsson, X., Morianou, I., Kyrou, K., Beaghton, A., Gribble, M., Kranjc, N., Galizi, R., Burt, A., Crisanti, A., Nolan, T., 2021. Regulating the expression of gene drives is key to increasing their invasive potential and the mitigation of resistance. PLoS Genet. 17, e1009321. https://doi.org/10.1371/journal.pgen.1009321

Hammond, A.M., Kyrou, K., Bruttini, M., North, A., Galizi, R., Karlsson, X., Kranjc, N., Carpi, F.M., D'Aurizio, R., Crisanti, A., Nolan, T., 2017. The creation and selection of mutations resistant to a gene drive over multiple generations in the malaria mosquito. PLoS Genet. 13, e1007039. https://doi.org/10.1371/journal.pgen.1007039

Hoermann, A., Habtewold, T., Selvaraj, P., Corsano, G.D., Capriotti, P., Inghilterra, M.G., Kebede, T.M., Christophides, G.K., Windbichler, N., 2022. Gene Drive Mosquitoes Can Aid Malaria Elimination by Retarding Plasmodium Sporogonic Development. https://doi.org/10.1101/2022.02.15.480588

Hoermann, A., Tapanelli, S., Capriotti, P., Del Corsano, G., Masters, E.K., Habtewold, T., Christophides, G.K., Windbichler, N., 2021. Converting endogenous genes of the malaria mosquito into simple non-autonomous gene drives for population replacement. eLife 10, e58791. https://doi.org/10.7554/eLife.58791

Hollingdale, M.R., Nardin, E.H., Tharavanij, S., Schwartz, A.L., Nussenzweig, R.S., 1984. Inhibition of entry of Plasmodium falciparum and P. vivax sporozoites into cultured cells; an in vitro assay of protective antibodies. J. Immunol. 132, 909–913.

Horn, C., Wimmer, E.A., 2000. A versatile vector set for animal transgenesis. Dev. Genes Evol. 210, 630–637. https://doi.org/10.1007/s004270000110

Isaacs, A.T., Jasinskiene, N., Tretiakov, M., Thiery, I., Zettor, A., Bourgouin, C., James, A.A., 2012. Transgenic Anopheles stephensi coexpressing single-chain antibodies resist Plasmodium falciparum development. Proc. Natl. Acad. Sci. U. S. A. 109, E1922-1930. https://doi.org/10.1073/pnas.1207738109

Isaacs, A.T., Li, F., Jasinskiene, N., Chen, X., Nirmala, X., Marinotti, O., Vinetz, J.M., James, A.A., 2011. Engineered resistance to Plasmodium falciparum development in transgenic Anopheles stephensi. PLoS Pathog. 7, e1002017. https://doi.org/10.1371/journal.ppat.1002017

Klug, D., Gautier, A., Calvo, E., Marois, E., Blandin, S.A., 2022. The salivary protein Saglin facilitates efficient midgut colonization of Anopheles mosquitoes by malaria parasites. https://doi.org/10.1101/2022.04.25.489337

Kyrou, K., Hammond, A.M., Galizi, R., Kranjc, N., Burt, A., Beaghton, A.K., Nolan, T., Crisanti, A., 2018. A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. Nat. Biotechnol. 36, 1062–1066. https://doi.org/10.1038/nbt.4245

Manzoni, G., Briquet, S., Risco-Castillo, V., Gaultier, C., Topçu, S., Ivănescu, M.L., Franetich, J.-F., Hoareau-Coudert, B., Mazier, D., Silvie, O., 2014. A rapid and robust selection procedure for generating drug-selectable marker-free recombinant malaria parasites. Sci. Rep. 4, 4760. https://doi.org/10.1038/srep04760

Marois, E., Scali, C., Soichot, J., Kappler, C., Levashina, E.A., Catteruccia, F., 2012. High-throughput sorting of mosquito larvae for laboratory studies and for future vector control interventions. Malar. J. 11, 302. https://doi.org/10.1186/1475-2875-11-302

Marshall, J.M., Raban, R.R., Kandul, N.P., Edula, J.R., León, T.M., Akbari, O.S., 2019. Winning the Tug-of-War Between Effector Gene Design and Pathogen Evolution in Vector Population Replacement Strategies. Front. Genet. 10, 1072. https://doi.org/10.3389/fgene.2019.01072

Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G., Neufeld, J.D., 2012. PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics 13, 31. https://doi.org/10.1186/1471-2105-13-31

O'Brochta, D.A., Alford, R., Harrell, R., Aluvihare, C., Eappen, A.G., Li, T., Chakravarty, S., Sim, B.K.L., Hoffman, S.L., Billingsley, P.F., 2019. Is Saglin a mosquito

salivary gland receptor for Plasmodium falciparum? Malar. J. 18, 2. https://doi.org/10.1186/s12936-018-2634-5

Okulate, M.A., Kalume, D.E., Reddy, R., Kristiansen, T., Bhattacharyya, M., Chaerkady, R., Pandey, A., Kumar, N., 2007. Identification and molecular characterization of a novel protein Saglin as a target of monoclonal antibodies affecting salivary gland infectivity of Plasmodium sporozoites. Insect Mol. Biol. 16, 711–722. https://doi.org/10.1111/j.1365-2583.2007.00765.x

Papathanos, P.A., Windbichler, N., Menichelli, M., Burt, A., Crisanti, A., 2009. The vasa regulatory region mediates germline expression and maternal transmission of proteins in the malaria mosquito Anopheles gambiae: a versatile tool for genetic control strategies. BMC Mol. Biol. 10, 65. https://doi.org/10.1186/1471-2199-10-65

Pham, T.B., Phong, C.H., Bennett, J.B., Hwang, K., Jasinskiene, N., Parker, K., Stillinger, D., Marshall, J.M., Carballar-Lejarazú, R., James, A.A., 2019. Experimental population modification of the malaria vector mosquito, Anopheles stephensi. PLoS Genet. 15, e1008440. https://doi.org/10.1371/journal.pgen.1008440

Port, F., Bullock, S.L., 2016. Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. Nat. Methods 13, 852–854. https://doi.org/10.1038/nmeth.3972

Quinn, C.M., Nolan, T., 2020. Nuclease-based gene drives, an innovative tool for insect vector control: advantages and challenges of the technology. Curr. Opin. Insect Sci. 39, 77–83. https://doi.org/10.1016/j.cois.2020.03.007

Reuling, I.J., Mendes, A.M., de Jong, G.M., Fabra-García, A., Nunes-Cabaço, H., van Gemert, G.-J., Graumans, W., Coffeng, L.E., de Vlas, S.J., Yang, A.S.P., Lee, C., Wu, Y., Birkett, A.J., Ockenhouse, C.F., Koelewijn, R., van Hellemond, J.J., van Genderen, P.J.J., Sauerwein, R.W., Prudêncio, M., 2020. An open-label phase 1/2a trial of a genetically modified rodent malaria parasite for immunization against Plasmodium falciparum malaria. Sci. Transl. Med. 12, eaay2578. https://doi.org/10.1126/scitranslmed.aay2578

Rono, M.K., Whitten, M.M.A., Oulad-Abdelghani, M., Levashina, E.A., Marois, E., 2010. The major yolk protein vitellogenin interferes with the anti-plasmodium response in the malaria mosquito Anopheles gambiae. PLoS Biol. 8, e1000434. https://doi.org/10.1371/journal.pbio.1000434

Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., Zhang, F., 2016. Rationally engineered Cas9 nucleases with improved specificity. Science 351, 84–88. https://doi.org/10.1126/science.aad5227

Terradas, G., Hermann, A., James, A.A., McGinnis, W., Bier, E., 2021. High-resolution in situ analysis of Cas9 germline transcript distributions in gene-drive Anopheles mosquitoes. G3 Bethesda Md jkab369. https://doi.org/10.1093/g3journal/jkab369

Triller, G., Scally, S.W., Costa, G., Pissarev, M., Kreschel, C., Bosch, A., Marois, E., Sack, B.K., Murugan, R., Salman, A.M., Janse, C.J., Khan, S.M., Kappe, S.H.I., Adegnika, A.A., Mordmüller, B., Levashina, E.A., Julien, J.-P., Wardemann, H., 2017. Natural Parasite Exposure Induces Protective Human Anti-Malarial Antibodies. Immunity 47, 1197-1209.e10. https://doi.org/10.1016/j.immuni.2017.11.007

Volohonsky, G., Terenzi, O., Soichot, J., Naujoks, D.A., Nolan, T., Windbichler, N., Kapps, D., Smidler, A.L., Vittu, A., Costa, G., Steinert, S., Levashina, E.A.,

Blandin, S.A., Marois, E., 2015. Tools for Anopheles gambiae Transgenesis. G3 Bethesda Md 5, 1151–1163. https://doi.org/10.1534/g3.115.016808

WHO, 2021. World Malaria report 2021.

Xie, K., Minkenberg, B., Yang, Y., 2015. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc. Natl. Acad. Sci. U. S. A. 112, 3570–3575. https://doi.org/10.1073/pnas.1420294112

Zavala, F., Cochrane, A.H., Nardin, E.H., Nussenzweig, R.S., Nussenzweig, V., 1983. Circumsporozoite proteins of malaria parasites contain a single immunodominant region with two or more identical epitopes. J. Exp. Med. 157, 1947–1957. https://doi.org/10.1084/jem.157.6.1947

Figure 1

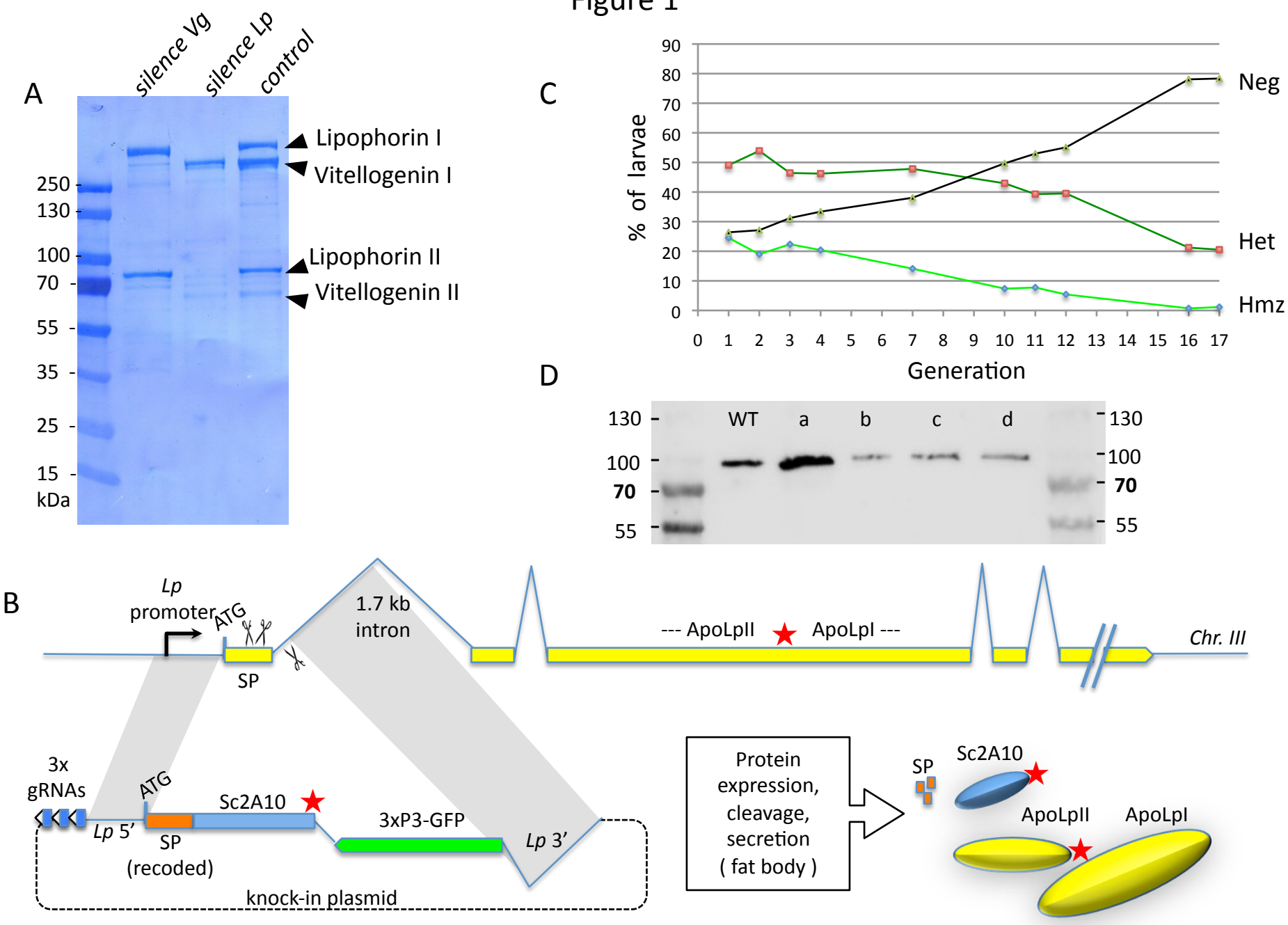

A
silence Vg
silence Lp
control
Lipophorin I
Vitellogenin I
Lipophorin II
Vitellogenin II
250
130
100
70
55
35
25
15
kDa
C
% of larvae
Generation
Neg
Het
Hmz
D
WT
a
b
c
d
130
100
70
55
B
Lp promoter
ATG
SP
1.7 kb intron
--- ApoLpII
ApoLpI ---
Chr. III
3x gRNAs
ATG
Lp 5'
SP (recoded)
Sc2A10
3xP3-GFP
Lp 3'
knock-in plasmid
Protein expression, cleavage, secretion (fat body)
SP
Sc2A10
ApoLpII
ApoLpI

Figure 2

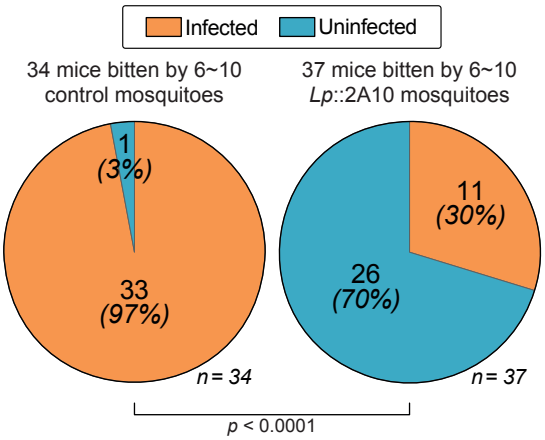

Infected
Uninfected
34 mice bitten by 6~10 control mosquitoes
37 mice bitten by 6~10 *Lp*::2A10 mosquitoes
1 (3%)
33 (97%)
n = 34
11 (30%)
26 (70%)
n = 37
p < 0.0001

Figure 3

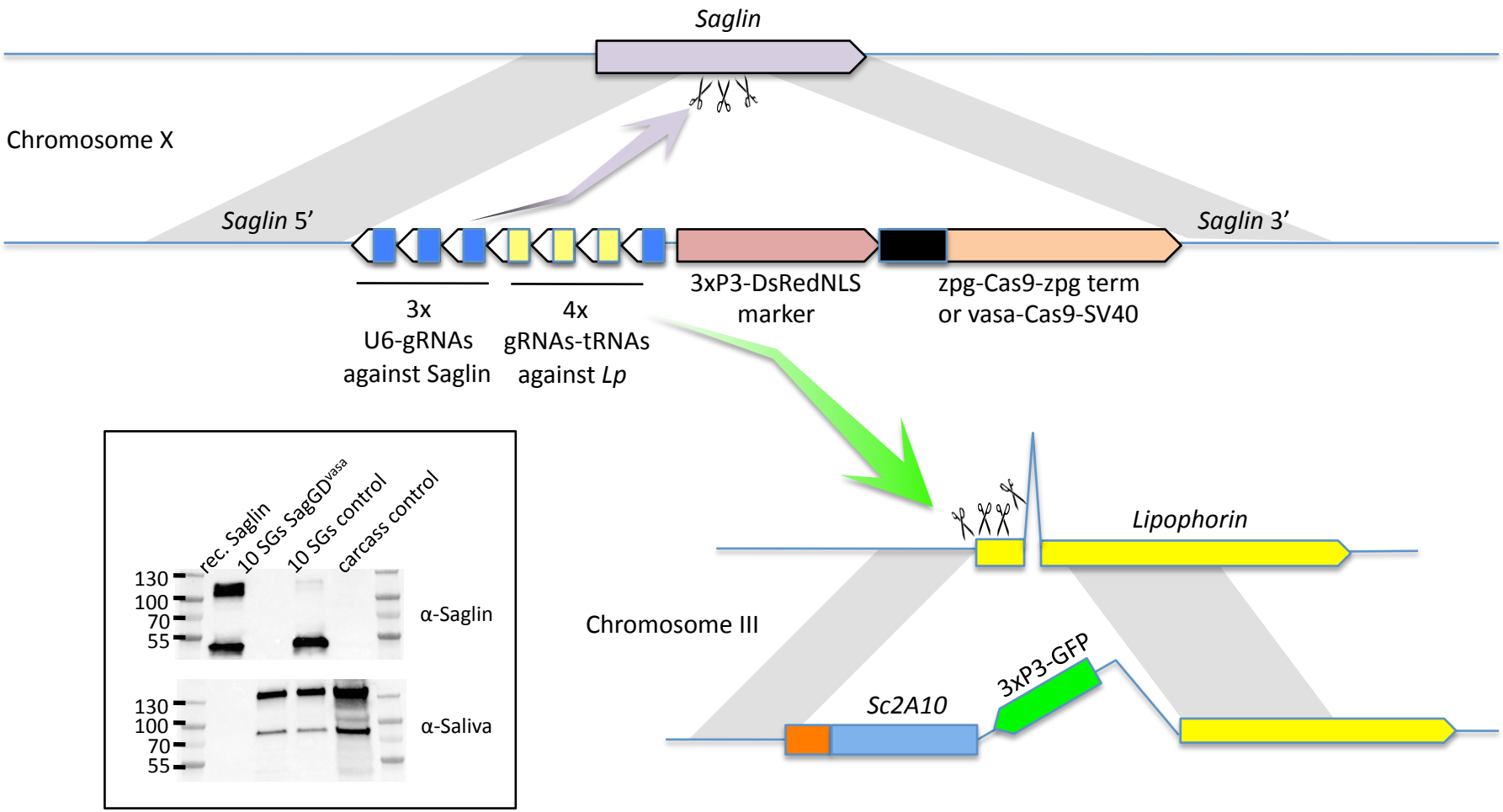
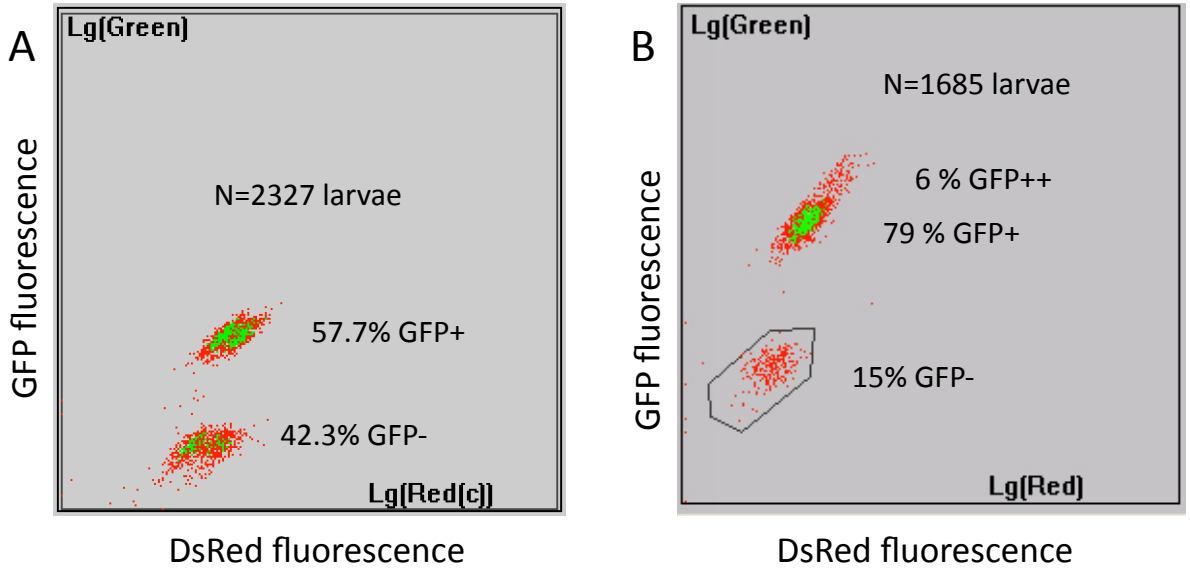

Saglin
Chromosome X
Saglin 5'
Saglin 3'
3x
U6-gRNAs
against Saglin
4x
gRNAs-tRNAs
against Lp
3xP3-DsRedNLS
marker
zpg-Cas9-zpg term
or vasa-Cas9-SV40
rec. Saglin
10 SGs SagGD^Vasa
10 SGs control
carcass control
130
100
70
55
α-Saglin
α-Saliva
Lipophorin
Chromosome III
Sc2A10
3xP3-GFP

Figure 4


A
Lg(Green)
N=2327 larvae
57.7% GFP+
42.3% GFP-
Lg(Red(c))
GFP fluorescence
DsRed fluorescence
B
Lg(Green)
N=1685 larvae
6 % GFP++
79 % GFP+
15% GFP-
Lg(Red)
GFP fluorescence
DsRed fluorescence

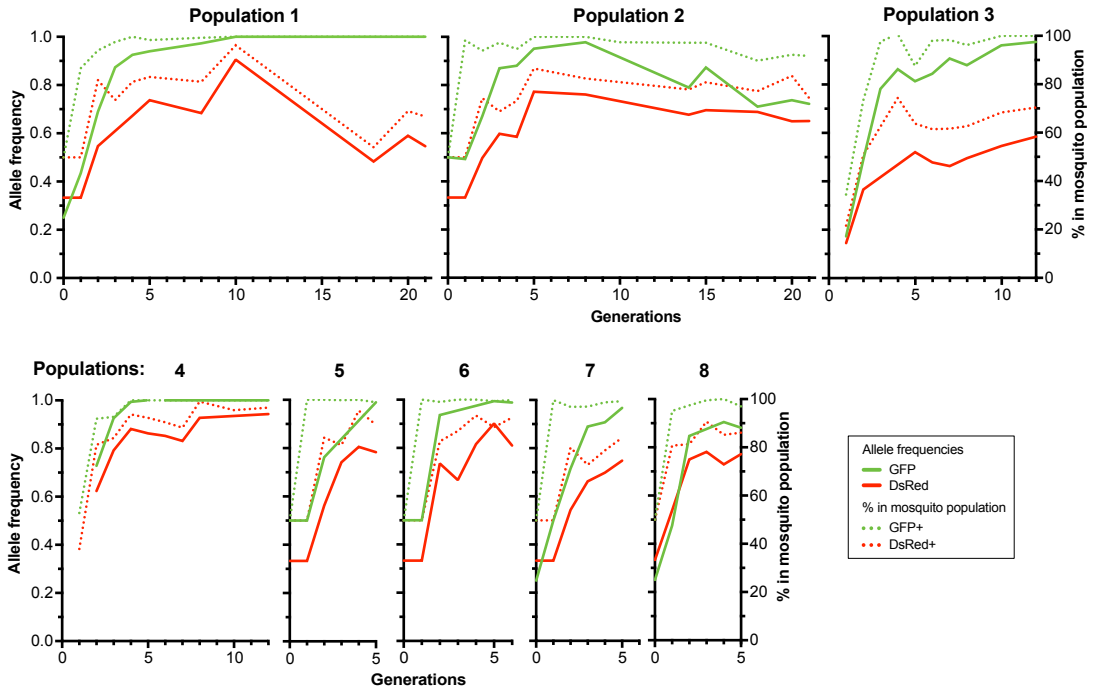

Population 1
Population 2
Population 3
Allele frequency
% in mosquito population
Generations
Populations:
4
5
6
7
8
Allele frequencies
GFP
DsRed
% in mosquito population
GFP+
DsRed+

Figure 6

a
gRNA1
gRNA2
gRNA3
PAM
43 nt
310 nt
124 nt deletion
b
gRNA1
gRNA2
gRNA3
PAM
43 nt
310 nt
c
Chr. 2R
11,129,164
gRNA1
gRNA2
gRNA3
gRNA4
PAM
AGAP001826 exon 1
AGAP001826 intron 1
M W V L G G R R L L W S F L V S L V L I Q S V S A
open reading frame
11,128,928
AGAP001826 intron 1