nature communications

Article

https://doi.org/10.1038/s41467-024-53631-5

# Improving the suppressive power of homing gene drive by co-targeting a distant-site female fertility gene

Received: 7 December 2023

Accepted: 16 October 2024

Published online: 26 October 2024

Check for updates

Nicky R. Faber[1,3] ✉, Xuejiao Xu[2,3], Jingheng Chen[2], Shibo Hou[2], Jie Du[2], Bart A. Pannebakker[1], Bas J. Zwaan[1], Joost van den Heuvel[1] & Jackson Champer[2] ✉

Gene drive technology has the potential to address major biological challenges. Well-studied homing suppression drives have been shown to be highly efficient in *Anopheles* mosquitoes, but for other organisms, lower rates of drive conversion prevent elimination of the target population. To tackle this issue, we propose a gene drive design that has two targets: a drive homing site where drive conversion takes place, and a distant site where cleavage induces population suppression. We model this design and find that the two-target system allows suppression to occur over a much wider range of drive conversion efficiency. Specifically, the cutting efficiency now determines the suppressive power of the drive, rather than the conversion efficiency as in standard suppression drives. We construct a two-target drive in *Drosophila melanogaster* and show that both components of the gene drive function successfully. However, cleavage in the embryo from maternal deposition as well as fitness costs in female drive heterozygotes both remain significant challenges for both two-target and standard suppression drives. Overall, our improved gene drive design has the potential to ease problems associated with homing suppression gene drives for many species where drive conversion is less efficient.

Gene drive technology could be a valuable tool to help address major challenges posed by populations of pest species. These populations include disease vectors, invasive species, and agricultural pests, for which often no effective method of control is available[1–5]. A gene drive is a genetic element that biases its inheritance in its own favour, allowing it to spread through a population over generations[6]. Synthetic gene drives can be designed for population modification, for example, to immunise a population of mosquitoes against malaria parasites[7], or for population suppression, for example, to eliminate a population of mosquitoes[8] or invasive pests[9,10]. Released in the target population, suppression gene drives are designed to spread at a rapid rate and carry a recessive fitness cost, thus causing a decline in population size or even complete elimination[8,11,12]. Suppression gene drives have benefits over conventional methods of control because they are species-specific (and thus more ecologically friendly), as well as potentially more efficient and more humane, though there are challenges regarding localisation and containment for some drive types[13].

There are many different types of gene drives, and what distinguishes them the most is how they handle the trade-off between efficiency of spread and confinement[2,5,6]. The most efficient and well-studied type of suppression drive is the CRISPR-Cas9-based homing drive. In drive heterozygotes, the gene drive copies itself to the

[1]Laboratory of Genetics, Department of Plant Sciences, Wageningen University & Research, Wageningen, The Netherlands. [2]Center for Bioinformatics, School of Life Sciences, Center for Life Sciences, Peking University, Beijing, China. [3]These authors contributed equally: Nicky R. Faber, Xuejiao Xu.
✉e-mail: nfaber@outlook.com; jchamper@pku.edu.cn

homologous chromosome in the germline through a process called "homing" or "drive conversion". The gene drive incurs a recessive fitness cost by being located inside a haplosufficient female fertility gene, so homozygous females are sterile. Ideally, female infertility is completely recessive, so that there are no fitness costs for drive heterozygotes. As the frequency of the gene drive increases, more sterile female offspring are created, thus imposing a genetic load on the population[14]. If the frequency of sterile females in the population is sufficiently high, the population can be completely eliminated[2,9].

Although homing gene drives are very promising in theory, in vivo tests in several organisms have revealed practical challenges, so complete population suppression is not yet attainable in most species (Fig. 1)[9,13,15–20]. Homing suppression gene drives face two major challenges, the first being the formation of functional resistance alleles[16,20–22]. A typical gene drive is active in the germline, and the primary way in which Cas9 cuts will be repaired is through homology-directed repair[23]. In this pathway, the cell uses the drive chromosome as a template, copying its sequence to the other chromosome. Sometimes, particularly if the cut is made outside of the time window in which this repair pathway is preferred, repair can occur via end-joining instead. This repair pathway is error-prone and often leads to small insertions and/or deletions. Because of these mutations, the gene drive guide RNA (gRNA) usually cannot recognise the new sequence, making it a resistance allele that can no longer be converted to a drive allele. Resistance alleles that preserve the function of the target gene ("r1" alleles) are problematic as there is a very strong selection for this allele which will rescue the population from the suppression drive[16]. Nonfunctional resistance alleles ("r2" alleles), on the other hand, will not be able to rescue the population because of their deleterious nature[20]. Although r2 alleles can reduce drive efficiency, the less common r1 alleles are the bigger problem because they will out-compete the gene drive in the population. There are several strategies to avoid the formation of r1 alleles in suppression drives, which include targeting an extremely conserved locus[8], multiplexing gRNAs[19,24,25], and using improved promoters[26–28].

Besides r1 alleles, a second challenge for suppression gene drives is low drive conversion, which is the rate at which wild-type alleles are converted into drive alleles[17,25]. In *Anopheles* mosquitoes, drive conversion has been shown to reach 95–99%[8,29]. In *Aedes* mosquitoes and *Drosophila melanogaster*, on the other hand, the conversion rate is usually significantly lower, around 50–70% for most constructs[20,25,30,31]. In addition, it appears that multiplexing gRNAs, which is a way to avoid r1 alleles, may further reduce conversion rates when using more than 2–4 gRNAs[19,25]. This low-efficiency results in partial population suppression instead of population elimination. The amount by which the population size is reduced compared to the expected population size in the absence of the gene drive is related to the genetic load of the gene drive when it reaches its equilibrium frequency in the population[14]. The genetic load needed to eliminate a target population increases for higher low-density growth rates (the reproductive advantage individuals experience in the absence of competition). Drive conversion efficiency seems to differ between species for the most commonly used promoters, though improvements might be possible by testing additional promoters[17,26,28,32–34], or perhaps by more strongly localising the gene drive mRNA to the nucleus[35].

To tackle this issue, we propose a gene drive design that has two targets: a homing site where drive conversion takes place, and a distant cutting site (where the drive is not present) for providing the fitness cost for population suppression (Fig. 1). For the homing site, we harness a modification drive that is located in an essential gene (thereby disrupting it) while also providing a rescue for this gene (Fig. 1)[36]. A previous study has already demonstrated the practical feasibility of using a homing gene drive while also targeting another gene for population modification[37], and another recent study has shown that

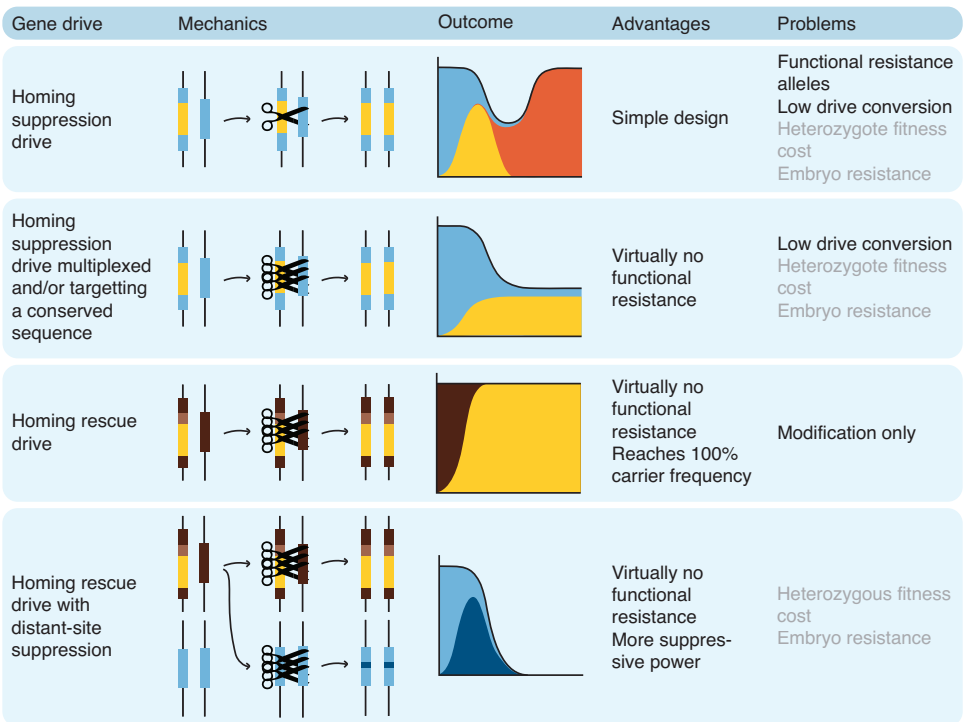


**Fig. 1 | Overview of several homing gene drive designs, their advantages, and major challenges they face.** In the outcomes graphs, light blue = female fertility gene, yellow = gene drive, orange/red = functional resistance allele, dark brown = essential gene, dark blue = non-functional resistance allele (only the female fertility target site is displayed for the distant-site suppression system). Additionally, in the schematics, light brown = essential gene rescue. Primary problems for drive designs are in black, and secondary problems are in grey.

https://doi.org/10.1038/s41467-024-53631-5

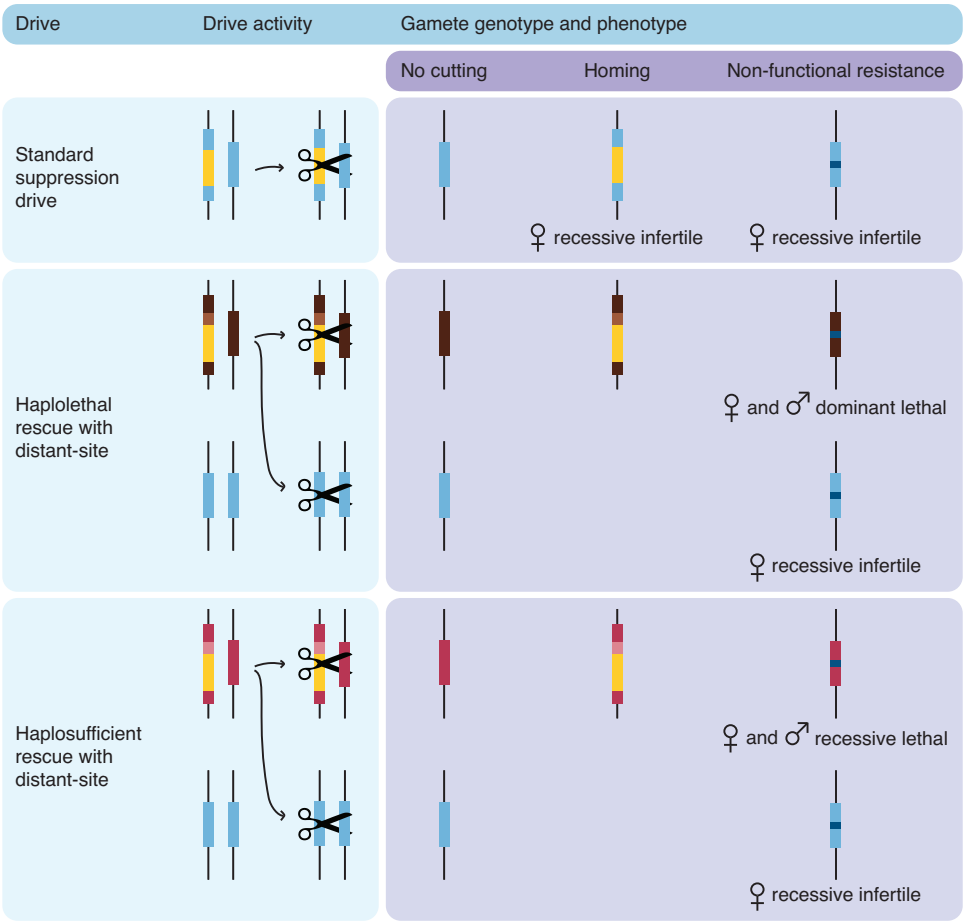


**Fig. 2 | Overview of our three modelled drives, their activity, and the resulting gamete genotypes and phenotypes.** We do not show functional resistance alleles, but their phenotypes would be the same as wild-type. Light blue = female fertility gene, yellow = gene drive, dark blue = non-functional resistance allele, dark brown = haplolethal essential gene, light brown = haplolethal essential gene rescue, dark pink = haplosufficient essential gene, light pink = haplosufficient essential gene rescue.

modifying a natural gene drive to also target a distant female fertility gene is feasible for population suppression[10]. Thus, we model the two-target design, comparing it to a standard homing suppression gene drive, and find that this two-target system allows suppression to occur over a much wider range of drive conversion efficiency. Most notably, the suppressive power now depends usually on cutting efficiency at the distant target instead of drive conversion, which is advantageous because the total cutting rate is easier to increase and has often been substantially higher than the drive conversion rate[17,18,20,34]. We construct this drive in *Drosophila melanogaster* and show that both components function successfully. We analyse the relevant genomic regions in which the drive components are inserted and find an unexpected sequence complexity. Nevertheless, the drive conversion rate is within the range generally observed in *D. melanogaster*, and the cut-rate at the distant-site target is very high. However, high fitness costs still thwart the drive's success in cage populations, a factor that could potentially be problematic in any suppression drive based on targeting essential genes. Nevertheless, our improved gene drive design enables the development of strong homing suppression gene drives for a wide array of species where drive conversion is less efficient.

## Results

In our main modelling results, we show the results of three gene drives: a standard female fertility homing suppression drive and two two-target drives that target a female fertility gene at a distant site for population suppression. These two-target drives are located in and provide rescue for a haplolethal or a haplosufficient gene. Drive conversion occurs normally in these, but they also cut and disrupt a distant-site female fertility target without rescue. Figure 2 shows the three drives' components, their activity, and also the alleles and phenotypes that can result in gametes from drive activity.

### Cut and conversion rate

Because the drive conversion rate is one of the most important parameters to determine the success of a suppression drive, we start by varying it, together with the total cut rate (referring to germline cutting)[12]. Any wild-type alleles that are cut but not converted to drive alleles are converted to nonfunctional resistance alleles (Fig. 3A). At the distant female fertility gene target, there is no drive allele present, so the total cut rate can only result in nonfunctional resistance alleles. Besides these, all other parameters are fixed at optimum values (no embryo resistance, no fitness costs, no functional resistance, and no effect of Cas9 saturation). Figure 3B shows population size over time after a gene drive introduction.

At 100% cutting and drive conversion, all three drives work equally well, rapidly eliminating the population. As the conversion rate decreases slightly, we still observe drive success, but stochasticity now plays an important role in achieving population elimination for the standard suppression drive. Further reducing drive conversion results first in the failure of the standard suppression drive, and then the haplosufficient rescue drive as well (though given more time it may still achieve complete suppression), whereas the haplolethal rescue drive remains successfully reliably and relatively efficient.

https://doi.org/10.1038/s41467-024-53631-5

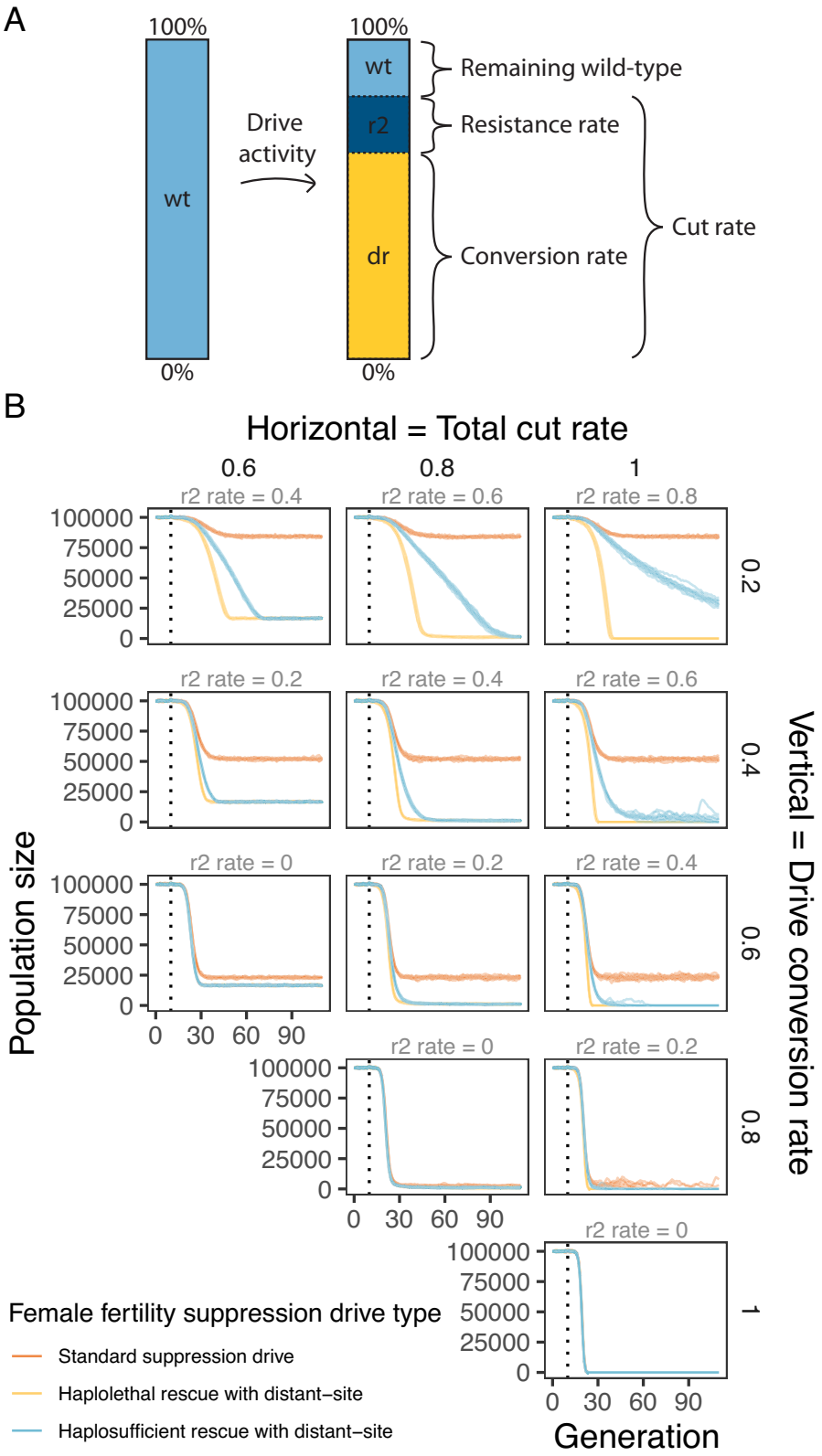


**Fig. 3 | Example population trajectories under varying cut and conversion rates. A** Schematic of drive activity rates in the germline. **B** Population suppression with varying total cut and drive conversion rates. The gene drive is introduced after generation 10 (the dotted line). The introduction frequency of gene drive heterozygotes is 0.01, and the total population size is 100,000. Note that both drive conversion rate and total cut rate are absolute rates, so the drive conversion rate can never be higher than the total cut rate. For each combination of parameters, we ran 10 model repetitions for each drive that are each shown as a translucent line.

https://doi.org/10.1038/s41467-024-53631-5

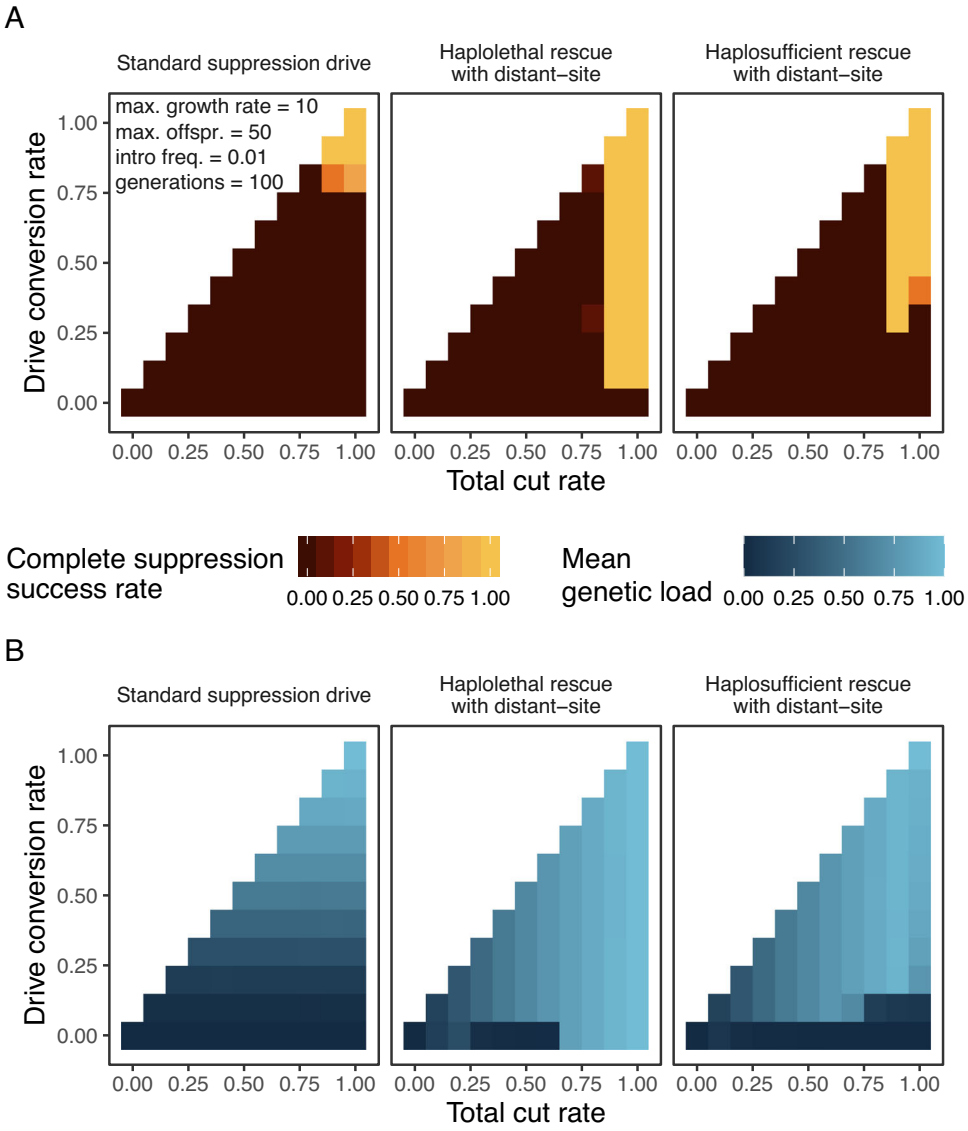


**Fig. 4 | Drive performance under varying cut and conversion rates. A** Complete suppression success rate and (**B**) mean genetic load with variable total cut and drive conversion rates. The complete suppression success rate is calculated as the number of repetitions in which population elimination occurs within 100 generations after drive introduction divided by the total number of repetitions. For each combination of parameters, we ran 10 model repetitions.

These major differences between drives can be explained by the standard suppression drive relying on drive conversion to both spread and suppress the population, whereas both two-target drives use two separate loci for this (Fig. 2). The haplolethal distant-site drive can make r2 alleles at the drive locus that are immediately removed due to the haplolethal nature of the gene (Supplementary Fig. S1A). Therefore, this gene drive spreads the most efficiently. The haplosufficient rescue drive and the standard suppression drive also form deleterious r2 alleles, but these can remain in the population due to slower removal, which impairs drive spread because drive conversion will not occur in drive/r2 heterozygotes (also an increase in r2 alleles cause a mild increase in stochasticity). In addition, in the standard suppression drive, not only do r2 alleles remain in the population and impair homing, but they also decrease the drive frequency (female drive/r2 heterozygotes are sterile) (Supplementary Fig. S1).

As the cutting rate decreases with constant drive conversion, we see that the standard suppression drive is not heavily impacted (its genetic load depends heavily on the conversion rate, with germline resistance alleles having little effect), whereas the distant-site suppression drives do not lose their effectiveness until the total cut rate falls significantly (Fig. 3 and Supplementary Fig. S2). Because the distant site drives the use of the total cut rate for disruption of the female fertility target, this total cut rate largely determines the suppressive power rather than drive conversion. Though the distant-site haplosufficient rescue drive is slower to spread due to the reduced ability to remove r2 alleles, they both eventually reach the same equilibrium population suppression (Supplementary Fig. S2).

For each gene drive, we determined the complete suppression success rate and the genetic load (Fig. 4). The genetic load is the suppressive power of a drive, defined as the reduction in reproduction of the population compared to a wild-type population of the same size[14]. In Fig. 4A, we see that the standard suppression drive has the smallest area of population elimination success, followed by the haplosufficient two-target drive, and then the haplolethal two-target drive. The standard suppression drive requires a high drive conversion rate to eliminate the population, whereas both two-target drives rely mostly on the cut-rate alone. The same pattern is visible in the genetic load in Fig. 4B, where the genetic load of the standard suppression drive relies on the conversion rate only, whereas both two-target drives rely almost entirely on the cut-rate for their genetic load.

The haplosufficient rescue drive shows some more complex dynamics in some areas of parameter space in Fig. 4B. When the cut rate is 1, but the drive conversion is low, the genetic load is reduced because the drive itself is not able to reach high frequency due to r2 alleles are blocking its progress (Supplementary Fig. S3). The drive spreads best with the highest genetic load when the total cut rate is somewhat below 1. At the same time, high cut rates are still necessary at the distant sites to achieve population complete suppression (Supplementary Fig. S3).

Another notable dynamic for both two-target drives in Fig. 4B is the bottom row, where drive conversion is 0. Here, the two two-target drives become identical to a version of two toxin-antidote drives previously described called a TADE (Toxin Antidote Dominant Embryo) suppression drive and a TARE (Toxin Antidote Recessive Embryo) suppression drive[38]. These drives show density-dependent dynamics, so their ability to increase in frequency depends on their frequency as well as their total cut-rate, hence the lack of suppression success for the haplolethal rescue drive due to our small release size. In the case of the haplolethal two-target drive, due to the additional disruption of the distant-site female fertility gene, a cut-rate of at least 0.7 is necessary for the drive to increase in frequency. The TARE-like suppression drive is not able to reach a high genetic load in the first place[38]. Interestingly, at very low cut rates, we observe that the two-target drives are able to remain in the population long enough for the distant site to be disrupted up to a certain frequency, after which the gene drive can sometimes induce a small genetic load during the simulation.

### Embryo cut rate and somatic expression effects on fitness

Two other important determinants of drive success are the embryo cut rate and fitness costs in heterozygous females based on the disruption of wild-type alleles when they are needed for fertility[25,29]. Embryo cutting occurs when Cas9 and gRNA transcripts and proteins are maternally deposited into oocytes. Any wild-type alleles (at the drive or distant target site), especially paternal wild-type alleles after fertilisation, can be cleaved, which always results in resistance allele formation at this stage (since the window for homology-directed repair is over). This process impairs the drive's spread much more than germline resistance alleles because female drive progeny will be sterile, and male drive progeny will be unable to perform drive conversion. Similarly, undesired Cas9 expression in somatic cells, regardless of whether it results in drive conversion or resistance allele formation, will disrupt the wild-type alleles needed for female fertility, at least in some cells. Depending on where and when the fertility target gene is needed, even necessary germline expression could induce a similar fitness cost in female drive heterozygotes (or any female with at least one drive allele and at least one wild-type fertility target gene).

To assess the effects of the embryo cut rate and somatic expression fitness cost under moderate drive conversion, we model a total cut rate of 0.9 and drive conversion rate of 0.5. These parameter choices mimic values that have typically been achieved in species that seem recalcitrant to gene drives, such as *D. melanogaster* and *Aedes* mosquitoes[20,25,30,31]. Varying the embryo cut rate and somatic expression parameters over their full range (Fig. 5A), we see that the haplolethal two-target drive is sensitive to high embryo cut rates, whereas the haplosufficient two-target drive is not strongly affected by this. Both are affected negatively by female heterozygote fitness costs, which prevent success more easily for the weaker haplosufficient drive. Looking at genetic load (Fig. 5B), the standard suppression drive starts out weak, with increases in the embryo cut rate having roughly half the effect of reduction of female fitness. The two-target drives have a larger area of high genetic load compared to the area of complete suppression success, representing areas where the drive is still able to increase and eventually reach a higher genetic load, but not within the time frame of the normal simulations. Higher initial release frequencies could still allow these drives to succeed in a shorter time frame.

To further show drive dynamics besides the eventual genetic load, Supplementary Fig. S4 shows population size over time after the gene drive introduction. The standard suppression drive cannot perform well with these parameters, so it always reaches an equilibrium population size, which is higher with increasing embryo cutting and fitness costs. With no somatic fertility cost and no embryo cutting, both two-target drives suppress the population rapidly. With an increasing embryo cut-rate, all drives are slower for all three drives, but the haplolethal two-target drive is especially sensitive to this, losing all suppressive power (and the ability to increase in frequency at all) when the embryo cut rate is high. This is because high embryo cutting at the drive's site results in the immediate removal of drive alleles in the progeny of females (all progeny with nonfunctional resistance alleles at this site are nonviable).

With decreasing drive female fitness from somatic Cas9 cleavage, all drives are slower to suppress the population. Eventually, the drives lose the ability to increase in frequency in the first place, with the exact point being dependent on the drive (Supplementary Fig. S5). With low embryo cutting, the haplolethal drive remains successful for lower female fitness values, and this trend is reversed for higher embryo cutting.

### Functional resistance alleles

Functional r1 resistance alleles have been shown to be a major challenge for many gene drives, so we investigate how well these can be handled by the two-target drive design. We only model r1 alleles at the homing site of each drive, and not at the distant site. The rescue modification drives can suffer from r1 alleles that form due to incomplete homing, where only the rescue sequence is copied, but not the rest of the gene drive[19]. Thus, even with multiplexed gRNAs, there may be a lower limit for functional resistance for these sorts of drives. For a standard suppression drive, this is not an issue. However, while multiplexing can likely reduce functional resistance to low levels[25], this could reduce the drive conversion rate with more than 2-4 gRNAs, or with gRNAs that are further apart, both of which could cause drive failure due to lack of genetic load, thus leading to practical limits for multiplexing that may still permit functional resistance. At the distant site, multiplexing is more flexible because either homology-directed repair or end-joining can achieve the desired result. This could allow for more gRNAs that target the most highly conserved sites throughout the target gene, largely eliminating the chance to form functional resistance alleles, much like CRISPR toxin-antidote drives[39].

We compare the drives in a range of parameter spaces where all are always successful in the absence of r1 alleles (Supplementary Fig. S6). We see that at relatively low r1 rates, the standard suppression drive is the first to succumb to functional resistance allele formation in the majority of repetitions (Fig. 6). This is because r1 alleles provide immediate benefit, directly allowing females to be fertile, which results in a rapid increase in the frequency of the r1 allele at low population density (Supplementary Figs. S6, S7). Thus, if a resistance allele forms, it only needs to avoid stochastic loss for a small number of generations before it will prevent population elimination. Only the haplolethal two-target drive remains successful in most repetitions with a high r1 rate, whereas both other drives fail in most or all repetitions. The haplosufficient two-target drive remains successful in most replicates at intermediate r1 rates, whereas the standard suppression drive fails at relatively low r1 rates. In both two-target drives, r1 alleles have only a modest advantage over the drive allele, and this is indirect. Individuals will not experience the modest female heterozygote fitness cost if they have only r1 and wild-type alleles at the drive site, and the r1 allele will not itself disrupt the female fertility gene for progeny, potentially increasing its chance of being passed on to a fertile female. However, the drive will reach high frequency (Supplementary Fig. S7), meaning

https://doi.org/10.1038/s41467-024-53631-5

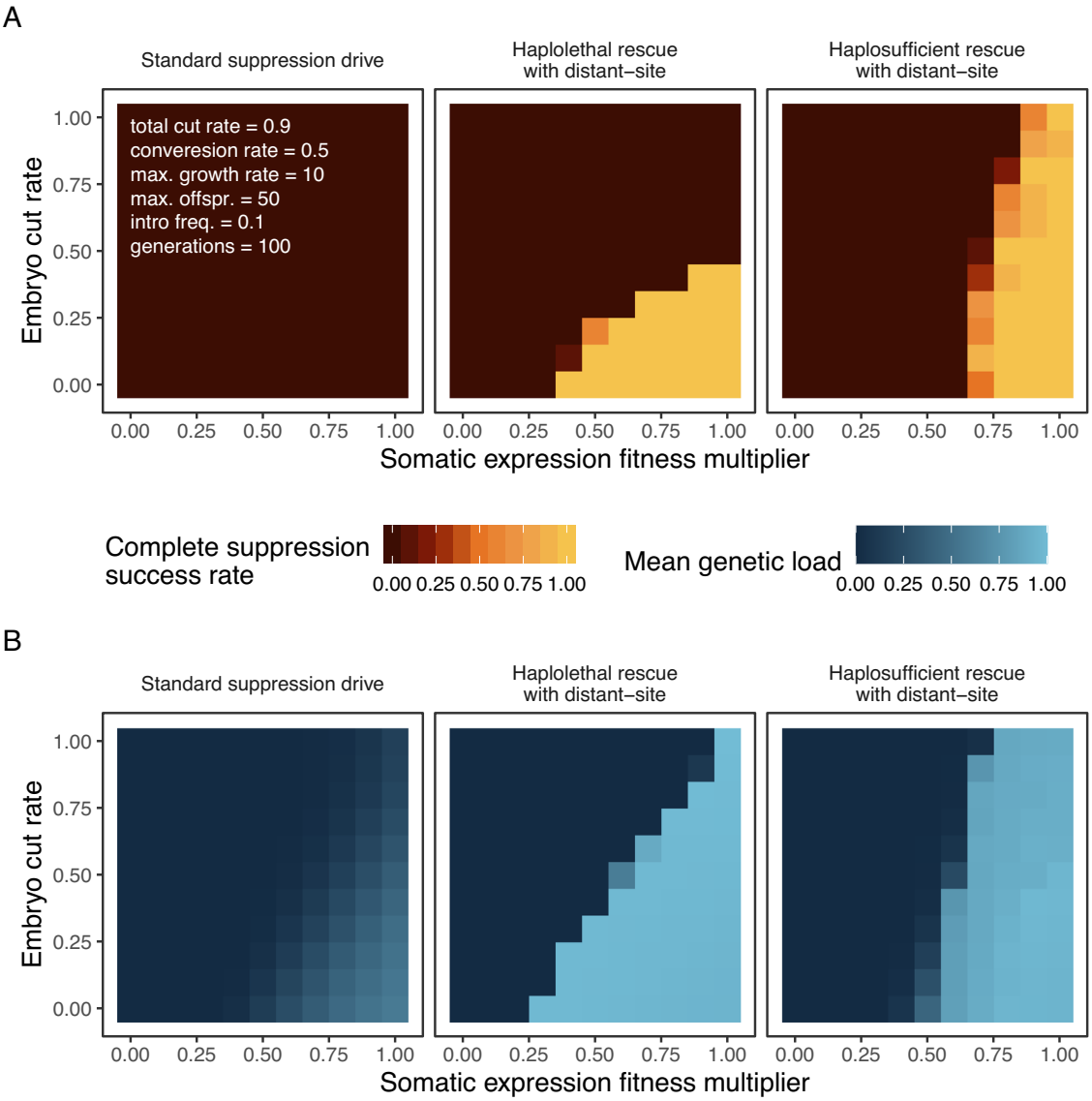


**Fig. 5 | Drive performance under varying somatic and embryonic drive activity.** **A** Complete suppression success rate and (**B**) mean genetic load under variable female fitness in drive heterozygotes from somatic expression fertility effects and variable embryo cut rates in the progeny of drive females. The complete suppression success rate is calculated as the fraction of simulations in which population elimination occurs within 100 generations after drive introduction. For (**A**), the introduction frequency of gene drive heterozygotes is 0.1, and the total population capacity is 100,000. The total cut rate is fixed to 0.9, and the drive conversion rate is 0.5. For each combination of parameters, we ran 10 model repetitions.

that most r1 alleles will be together with drive alleles, limiting their advantage and thus doing little to prevent suppression, which only requires one drive allele to disrupt the distant-site female fertility gene.

### gRNA saturation

Because the two-target drives need double the amount of gRNAs, the drive as a whole may suffer from additional gRNA saturation compared to a standard homing suppression drive. This means that the cutting efficiency of each individual gRNA is reduced due to a limited amount of Cas9 protein. Gene drives using multiplexing to avoid r1 alleles already face this challenge[19], which would be amplified in two-target drives. We do not model multiplexing explicitly, but instead assume that an equal number of gRNAs will be used for the homing site and the distant site, thus allowing us to reduce the cutting efficiency at each site proportionally. Here, we calculated the genetic load for each drive with various cut rates and gRNA saturation factors (Supplementary Fig. S8). The gRNA saturation factor is the relative Cas9 activity level with unlimited gRNAs (spread equally between all the gRNAs), with "1" being the activity in the presence of a single gRNA[19]. A saturation factor of 1 thus means that the total cut rate is immediately split between any number of gRNAs, while a rate of infinity (our default in previous simulations) means that the cut-rate at each gRNA target is the same as the cut rate of a 1-gRNA system.

The standard suppression drive is not impacted by this model of gRNA saturation because this drive only targets a single site (it can be thought of as the "baseline" in this scenario, even if it would have multiple gRNAs in practice) (Supplementary Fig. S8). The two target drives are both impacted by gRNA saturation, which reduces the genetic load by a moderate amount if the gRNA saturation factor is low. This reduction is low if the total cut rate is high (Supplementary Fig. S9). For the haplosufficient drive with a high cut rate, gRNA saturation actually has a positive effect on the drive because its optimal cut rate is somewhat below 1 (Fig. 4 and Supplementary Fig. S10). The current best estimate of the gRNA saturation factor is 1.5[19], though this could potentially vary substantially between systems. Overall, gRNA saturation is potentially problematic, though the two-target

https://doi.org/10.1038/s41467-024-53631-5

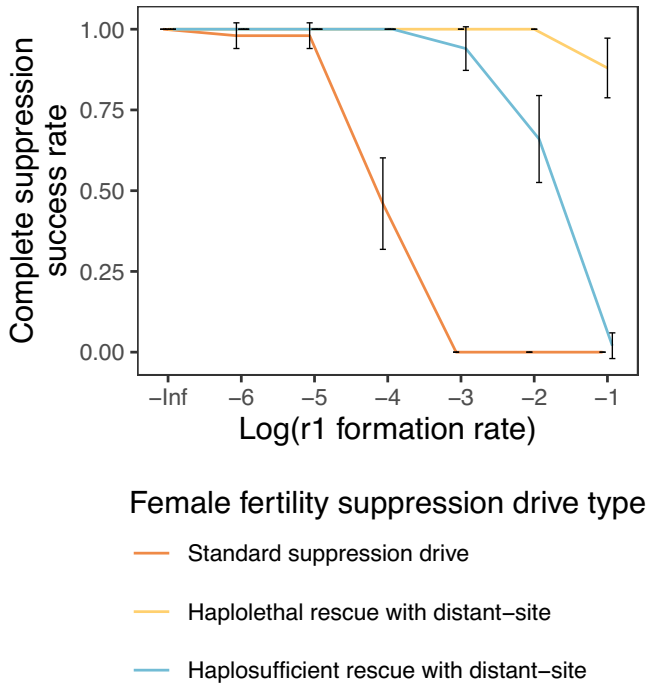


**Fig. 6 | Impact of functional resistance alleles on drive performance.** Successful population elimination under various functional resistance allele formation rates. The r1 formation rate is the relative rate of resistance alleles that are functional, rather than nonfunctional r2 alleles. The model is run for 100 generations after an introduction of drive heterozygotes at 0.01 frequency into a population of 100,000. To compare the drives where they were all successful in the absence of functional resistance alleles, we used a total cut rate of 1.0, a drive conversion rate of 0.9, a somatic expression female fertility fitness effect of 0.9, and an embryo cut rate of 0.1 in the progeny of female drive carriers. For each combination of parameters, we ran 50 model repetitions. Error bars show the 95% confidence intervals around the mean.

drives would still usually be expected to have higher power than a standard suppression drive, especially when the total cut rate is high.

### Test of two-part gene drive in *Drosophila melanogaster*

We construct the two-target gene drive design in *D. melanogaster* by reusing two previously built gene drives, a standard female fertility suppression drive and a haplolethal modification drive, and combining them[25,36]. The homing construct is integrated in the haplolethal gene *RpL35A*, whereas the distant target site is the haplosufficient female fertility gene *yellow-g* (Fig. 7A). This drive was constructed based on another successful drive reported previously[36], by adding extra gRNAs targeting *yellow-g* from a previous suppression drive[25] to convert the original modification drive into a suppression drive.

Two independent EGFP-marked drive lines (named line 1A and line 1D) were generated. These drive lines were respectively crossed to a DsRed-marked *nanos*-Cas9 line to produce heterozygous individuals carrying one copy of the drive and one copy of the Cas9 allele, after which these heterozygotes were crossed to $w^{1118}$ flies for assessing drive efficiency. The drive inheritance rates of Line 1A were 78% for drive males and 71% for drive females, while the inheritance rates of Line 1D were 76% for drive males and 67% for drive females. In comparison, the inheritance for the drive in the absence of Cas9 was all approximately 50%. Our fitted model indicates significant effects of the cross (control or drive) ($\chi^2(1) = 112.60$, $p < 0.001$) and the sex of the driving parent ($\chi^2(1) = 13.89$, $p < 0.001$), but not between lines or any interactions (Fig. 7B and Source Data Set S1). The drive inheritance rates were much lower than the original haplolethal homing drive, where drive inheritance rates for both male and female heterozygotes were 91%. This difference is likely caused by gRNA saturation, meaning that the two gRNAs of the homing drive had lower cut rates because they shared the same amount of Cas9 with the additional four gRNAs targeting *yellow-g*. Besides the difference between males and females, the two gRNA lines performed slightly differently in drive conversion rate as well.

### Cage experiments

To characterise the population suppression activity of this drive, cage experiments were set up by mixing drive carriers (heterozygous for the drive allele and homozygous for Cas9) and a small fraction of Cas9 homozygous flies. These cages were followed for several discrete generations, measuring drive carrier frequency by phenotyping all adults. Though there was some stochasticity in population size due to the single-bottle nature of the cages, the dynamic of the drive was clear. Drive carrier proportions of most cages quickly decreased within five generations (Fig. 8A). The population size for the two lower frequency releases was unaffected, and for one of the high-release frequency cages, the population size recovered after an initial reduction (Fig. 8B, Source Data Set S2). However, the other high frequency release cage showed complete population suppression in the second generation even though drive carrier frequency was declining. In the suppressed population, the five remaining flies all happened to be females, whereas this was not the case in the population that recovered. There, the five flies in that generation consisted of four females (one of which lacked the drive) and one male (also without the drive), allowing the population to grow immensely in the next generation. This result implies that the distant-site suppression drive was functional, though overall efficiency was low due to drive fitness costs outweighing biased drive inheritance. It also confirms that near-fixation of a gene drive will not be enough to suppress a population of *D. melanogaster*, given their ability to produce many offspring from just a few remaining individuals under the conditions of these cages.

### Resistance alleles and fertility

While functional resistance is a possible explanation for the rapid decline of a suppression drive from high frequency, we considered this unlikely due to the use of multiplexed gRNAs and previous studies in larger cages with similar drives finding no evidence of functional resistance[25,36]. To explore the possible reason for poor cage performance, a fertility test was conducted to assess the fitness of our drive carriers (which was found to be reduced in both drives this two-target drive is based on, but not to the extent of eliminating them from cages[25,36]). We observed a considerable reduction in egg viability in females compared to males, particularly in those carrying the drive. For females with the drive, 24% of eggs were viable in line 1A, and 19% were viable in line 1D. In contrast, for males carrying the drive, 96% of eggs were viable in line 1A, and 84% were viable in line 1D. In individuals not carrying the drive, the average egg viability is similar between males and females (Fig. 9 and Source Data Set S3). Interestingly, outliers in both drive-carrying and non-drive-carrying females indicate that a similar mechanism might be behind the reduction in egg viability. Our fitted model indicates significant effects of the sex of the drive parent ($\chi^2(1) = 16.62$, $p < 0.001$), and line (1A or 1D) ($\chi^2(1) = 8.12$, $p = 0.004$), as well as of interactions between cross (drive carrying or not) and sex of drive parent ($\chi^2(1) = 12.42$, $p < 0.001$), and between cross and line ($\chi^2(1) = 5.12$, $p = 0.024$). We collected 16 and 18 drive daughters from drive females (which were offspring of the above cross) of line 1A and line 1D, respectively, and crossed them to $w^{1118}$ males. 87.50% from line 1A and 77.78% from line 1D were sterile (Supplementary Table S1), which was likely caused by resistance allele formation in the embryo at the *yellow-g* target site due to maternal deposition of Cas9 and gRNAs. Indeed, sterile females were randomly collected for sequencing, which indicated 100% cutting in the *yellow-g*

https://doi.org/10.1038/s41467-024-53631-5

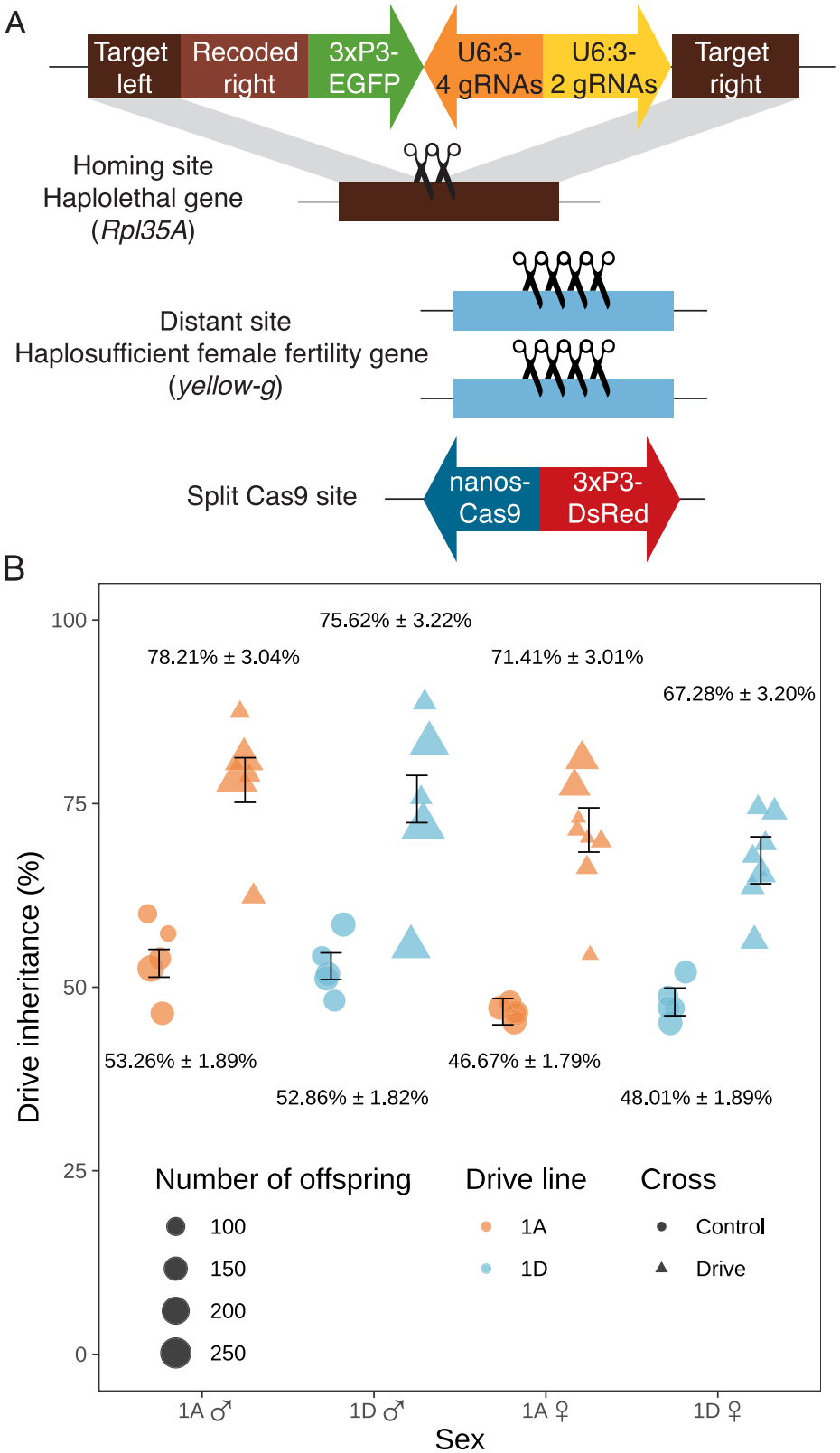


**Fig. 7 | Drive construct and inheritance rates. A** The drive allele is inserted into the haplolethal gene *RpL35A* and contains a recoded version as a rescue to avoid gRNA cleavage. It is expected that only offspring inheriting two functional copies of *RpL35A* (which can be either a drive or wild-type allele) can survive. The two scissors at the *RpL35A* site show the gRNA sites for cleaving and homing, while the four scissors at the *yellow-g* site show the gRNA sites for disrupting *yellow-g*, a haplosufficient female fertility gene located at a distant site. **B** The drive inheritance rate indicates the percentage of offspring with EGFP fluorescence from the cross between heterozygotes (containing one copy of drive and one copy of Cas9) and $w^{1118}$ flies (without drive and Cas9). The size of each dot represents the total number of offspring in that batch, and whiskers indicate the standard error of the mean. For each group from left to right, $n = 691, 760, 751, 968, 780, 777, 702$, and 800 flies. Source data are provided as a Source data file (Source Data Set S1).

https://doi.org/10.1038/s41467-024-53631-5

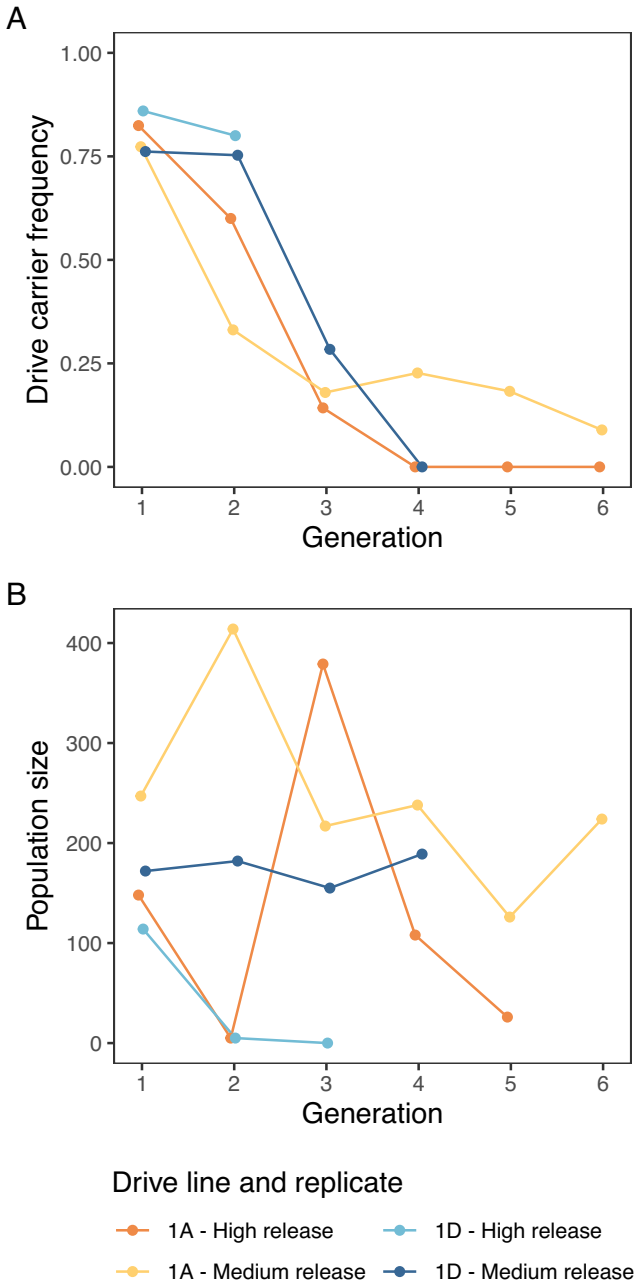


**Fig. 8 | Small cage experiments.** Heterozygous flies carrying one copy of the drive allele and two copies of Cas9 were used to generate progeny for "generation zero" of four small cages (the two medium release cages also included a non-drive female that was also homozygous for Cas9). **A** Drive carrier frequency and (**B**) population size in each generation. Source data are provided as a Source data file (Source Data Set S2).

target site. Another 16 non-drive progeny from a driven female and Cas9 male cross were randomly collected for sequencing, showing 87.5% (14 out of 16) cutting in *yellow-g*. These results indicate that the poor cage performance in this study is likely due to the fitness cost of female drive heterozygotes together with high levels of embryo cutting and subsequent resistance allele formation at both the female fertility site and the haplolethal drive site, though it is currently unclear to what extent each of these factors contributes.

In addition to the egg to adult viability, we measured the drive inheritance of these offspring. The drive inheritance rate of line 1A was 67.82% ± 2.83% in drive males and 74.93% ± 3.73% in drive females, while it was 75.99% ± 2.67% in drive males and 79.47% ± 4.62% in drive females in line 1D (Supplementary Fig. S11). The drive inheritance rates of these flies were somewhat higher than our previous test, where drive heterozygotes had only a single copy of Cas9, except for drive males of line 1A. The reason for this anomaly is unclear, particularly considering that the lack of Cas9 appears to be a limiting factor in the performance of the drive compared to the original modification drive it is based on. In addition, although drive inheritance rates in the previous test were higher in drive males than in drive females, this difference was not seen in this later drive conversion test.

### Whole-genome sequencing

To further investigate the difference in performance between lines 1A and 1D, as well as the generally lower-than-expected homing efficiency, we sequenced the genome of both drive lines using Nanopore sequencing, because initial screening indicated the presence of repeated regions. For line 1D, a single contig covered the entire gene drive, showing that although the line contained the full construct as expected, the insertion was duplicated, including the cloning vector in between the two copies (Supplementary Fig. S12C). This means that line 1D contained an extra copy of the distant-site gRNAs cassette. For line 1A, there were five separate contigs covering the gene drive. We found that the distant-site drive construct was incorporated thrice, with two vector backbones in between, and with inversions between two gRNA cassettes (Supplementary Fig. S12B). This was likely caused by recombination between gRNA cassettes during transformation. Line 1A contained four gRNA cassettes, but both distant-site ones were only partially present (two out of four distant-site gRNAs were eliminated due to the recombination).

## Discussion

In this study, we analysed the possibility of converting a modification to a suppression drive by the addition of gRNAs targeting female fertility or other essential genes without rescue. We found via modelling that this can substantially increase the genetic load (suppressive power) of the drive compared to a similar standard suppression drive unless the drive conversion rate is already very high. Instead of the genetic load being mainly determined by the drive conversion rate, it is determined by the total germline cut rate in these two target drives. We then tested such a drive in the model organism *D. melanogaster*, which successfully biased its inheritance while cutting the distant female fertility gene target at high rates. However, fitness costs in driving heterozygous females were quite high for this drive, which prevented success in most cage replicates despite high release frequency. Nevertheless, two-target drive systems may be quite desirable tools for the suppression of pests where high drive conversion is difficult to achieve but not high cut rates. Such high drive conversion has only been consistently achieved in *Anopheles* mosquitoes, while several other species have had substantially lower or inconsistent drive conversion, such as *Drosophila*, *Aedes*, and mice[20,24,25,30,34,37]. In these, relatively high cut rates have often been achieved when drive conversion is low, which would be necessary for the two-target drive design to work in these species.

In the two-target suppression drive, population suppression is largely decoupled from drive frequency increase. This is in contrast to the standard homing suppression rive, where the drive allele itself is mainly responsible for disrupting the suppression target. Thus, the two-target drive can keep increasing in frequency and eventually cut most female fertility target alleles. The standard suppression drive reaches an equilibrium that allows many wild-type alleles to remain, mostly accounting for its genetic load. Though the two-target suppression drive allows for higher suppressive power, it still needs to be able to increase in frequency in the first place. Thus, drive conversion must be sufficiently high to overcome fitness costs due to somatic

https://doi.org/10.1038/s41467-024-53631-5

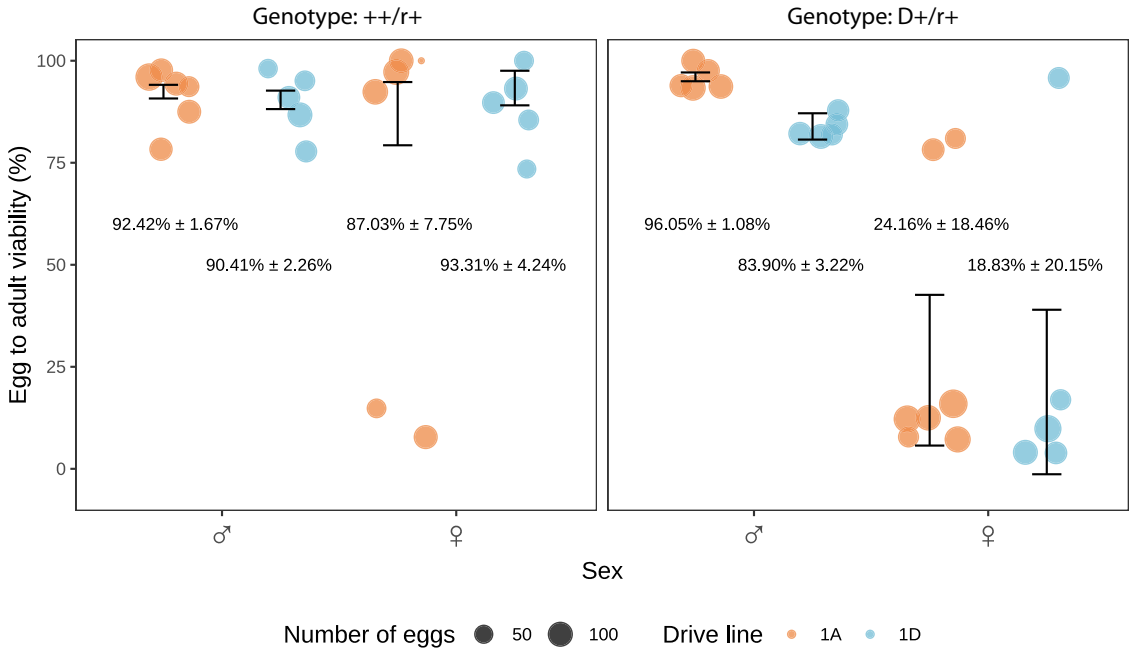


**Fig. 9 | Egg viability test.** The individuals tested in this experiment were all homozygous for Cas9, but different in the drive site and distant site, for which genotypes are listed at the top. "+", "D" and "r" respectively represent wild-type allele, drive allele and nonfunctional (*yellow-g* target site) resistant allele. Note that while cut rates at the *yellow-g* site were high, it is possible that a small fraction of flies marked as having an "r" at the target site were, in fact, homozygous for wild-type at this site. The size of each dot represents the total number of eggs laid in that batch, and whiskers indicate the standard error of the mean. For each group from left to right, $n = 534$, 361, 463, 332, 453, 404, 681, and 434 eggs. Source data are provided as a Source data file (Source Data Set S3).

expression and embryo cutting (the latter of which is mostly important at the distant site, though it could have a substantial negative effect at the drive site if the drive targets a haplolethal gene there). These performance parameters are still highly dependent on the Cas9 promoter and, to a lesser extent, the target gene. For the target gene, additional possibilities are available for two-target drives because they can keep a high genetic load even when targeting X-linked female fertility genes.

While our modelling results are promising for two-target drives that lack high drive conversion rates, it should be noted that drive conversion is still essential for determining the speed at which the drive's frequency will increase, which determines the time to population elimination. Though this system can increase the genetic load, such drives will still take longer to eliminate a population than a drive with high drive conversion efficiency (regardless of whether such a drive is a standard suppression drive or a two-target system). In more complex environments such as two-dimensional continuous space, this slowed rate of increase could potentially result in failed suppression due to the "chasing" effect, where wild-type individuals migrate back into an area previously cleared by the gene drive, resulting in cycles of suppression and re-establishment[12]. However, this same study included a powerful but slow drive that avoided this effect, which may be analogous to a two-target system. Additional modelling is thus needed to explore the outcomes of two-target drives in complex scenarios[40].

It should be noted that, like standard homing drives, measures to limit or localise this drive may be necessary. The designs modelled in this study could be threshold-dependent under certain narrow parameter combinations (for example, certain somatic expression fertility effects and embryo cut rates). Outside of those ranges, such as standard homing drives, they could be designed to be self-limiting or threshold-dependent by incorporating a daisy-drive[41] or tethered design[42]. In addition, the fact that two-target drives target multiple loci could have interesting potential for targeting a specific population only based on locally fixed alleles[43]. In fact, the locally fixed allele strategy would be easier to implement with a two-target drive because the gRNAs targeting the female fertility gene would not need to be close together to achieve high drive conversion efficiency.

In our experimental demonstration, the constructs functioned in terms of drive conversion while cutting the distant site at high rates. However, drive efficiency at the homing site was considerably lower than the original modification rescue drive[36], for which duplications in our drive lines could be the cause. Although the drive seems stable in our lines based on the pooled genomic data of 25 individuals, it may not have been during drive conversion due to the duplication of homology arms, resulting in the loss of some regions. More importantly, during homology-directed repair, when the gene drive induces a double-stranded break, DNA from the homologous chromosome might form secondary structures due to the repeats in the gene drive. This could potentially obstruct the repair machinery and cause replication stress[44], reducing the relative efficiency of homology-directed repair. In another recent homing rescue drive, a low homing efficiency was observed, which increased when a repeated region was replaced with a similar element that contained several mutations[45]. Considering these, it seems advisable to avoid repeated regions as much as possible in gene drive constructs, which could be done readily by simply combining the gRNAs into a single cassette or using different gRNA promoters.

Besides the structural problems, the lower efficiency could also be due to Cas9 saturation from the additional gRNAs targeting the distant site, since we used the same Cas9 as the original homing rescue drive[36]. The fact that line 1A contained more homing gRNAs (both proportionally and absolutely) could partially explain the higher inheritance rate we observed. Since cut rates were very high at the distant site, gRNA saturation likely had a smaller impact there. The cut-rate at the homing site would be expected to decrease more than at the distant site, especially because the cut rate for the original modification drive was not high in the first place[36], at least compared to other

https://doi.org/10.1038/s41467-024-53631-5

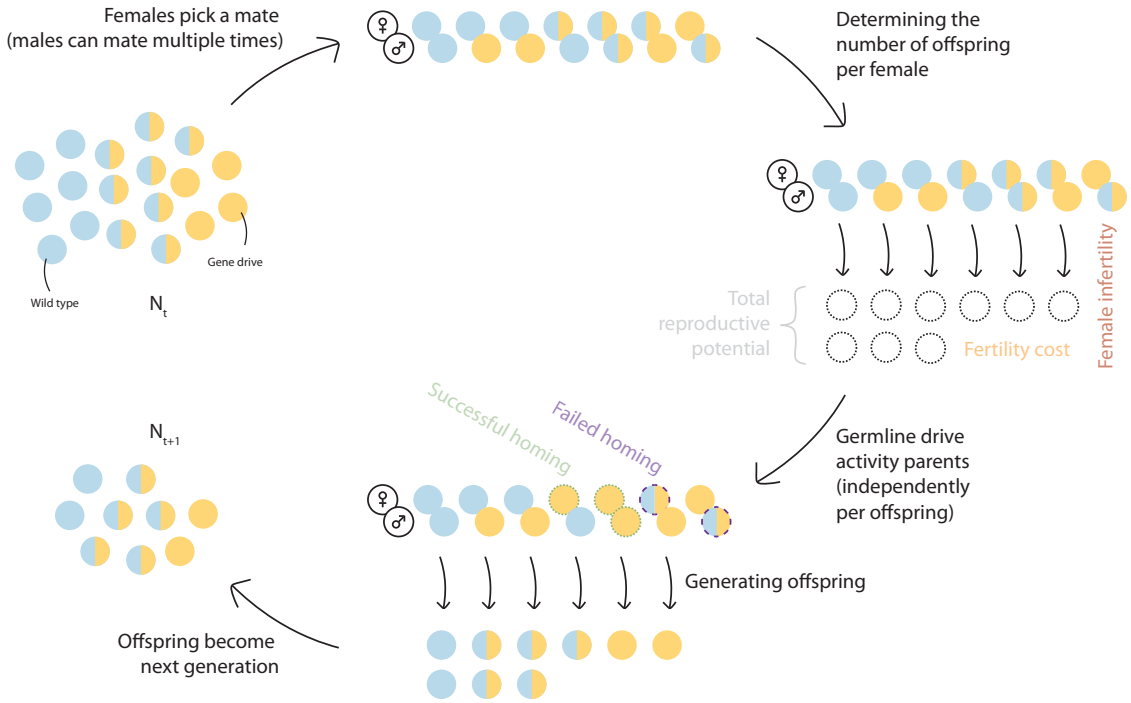


**Fig. 10 | Model overview, showing a simplified sequence of steps the model goes through every generation.** Each circle represents an individual, and its colour shows the genotype. We show a simple homing gene drive with only wild type (blue) and drive (yellow) alleles. The two-target drives will show different dynamics dependent on their design (Fig. 2). After each female picks a male for mating, we calculate the reproductive potential of these females. Then, based on drive genotype, females can be either completely fertile, or completely sterile (drive homozygotes), or partially fertile (drive heterozygotes). Offspring is generated, and for each individual offspring, Mendelian inheritance is modified by drive activity in the parental germline. This means that drive conversion could be successful for one offspring of a single drive parent, but not another. Additional drive activity from maternally deposited Cas9 is also modelled, but not shown in this figure. Finally, these offspring become the next generation of adults.

published drives in *D. melanogaster*. These issues could be more readily avoided using entirely new designs with balanced numbers of high-activity gRNAs for both target sites. In general, however, the fact that two-target drives use a higher number of gRNAs than standard homing drives could be a potential weakness due to gRNA saturation reducing the cut-rate at each site. We expect this reduction to be generally limited, though, because cutting rates are high for many organisms, thus reducing the impact of gRNA saturation.

Arguably more problematic than the homing efficiency, the high fitness cost in female heterozygotes prevented success in population cages, and embryo cutting at both the drive site and the distant target site further contributed to drive failure. This fitness cost was expected to an extent because our two-target drive is based on a suppression drive showing reduced fitness as well[25]. In the haplolethal modification drive, drive carrier males showed no significant reduction in offspring egg-to-pupa survival, but females did show a significant effect[36]. This fitness reduction would represent an additional fitness cost in our two target drives and is likely from nonfunctional resistance alleles rendering offspring nonviable. If formed in the germline, these alleles would not have significant negative effects, but if formed in the embryo, they could further contribute to reduced drive success. In the earlier standard suppression drive targeting *yellow-g*, at first, offspring from neither male or female drive parents showed a significant reduction in the egg-to-pupa survival rate, but later crosses with drier food conditions did show a significant reduction in egg-to-pupa survival rate, and inference from cage experiments showed even higher fitness costs. This fitness cost likely indicates either leaky somatic expression or disruption of *yellow-g* too early in the germline, where it still might need to function. Together, these factors could all contribute to the more severe fitness cost for female drive carriers with our two-target drive, though we currently do not know what factor contributes to what extent. In the original suppression drive, offspring of drive carrier females were found to be sterile at high rates. This infertility indicates significant embryo activity, which also had a similar negative effect on the two-target drive. In the two-target drive, gRNA saturation could have slightly ameliorated these effects due to the presence of more gRNAs and cut rates being slightly reduced, but the different genomic sites may change this, and reduction in drive conversion efficiency for similar reasons would likely outweigh any benefits.

All of these fitness costs will likely be similar for two-target drives targeting different genes using *nanos*-Cas9, with the exception of *yellow-g* potentially being required in the early germline. Therefore, finding promoters that are exclusively expressed in the (late) germline is essential for the future development of two-target drives (and any suppression drive, for that matter), and promoters with little to no embryo activity are also preferred. In the standard suppression drive, however, the higher drive conversion rate (and lack of embryo cutting at the haplolethal target) still allowed it to reach a moderate equilibrium instead of declining to zero. If the two-target drive in our study had similar or even somewhat lower drive conversion, but lower fertility cost or embryo resistance allele formation, our model indicates that it would have been able to increase in frequency in the population and eventually eliminate it with high genetic load. Even in its current form, our drive could potentially be successful with repeated releases, analogous to a more powerful form of sterile insect technique, which is far less powerful than a gene drive but still potentially useful in several situations due to its self-limiting nature[46].

Overall, we have shown that for scenarios of low to moderate drive conversion and high total germline cut rates, two-target drives may offer substantially increased suppressive power. With further efforts to find better Cas9 promoters and target gene combinations with lower fitness costs, this unique drive design may unlock the potential for strong population suppression in many scenarios.

## Methods

### Ethical statement

This research complies with all relevant ethical regulations and was approved by the Peking University biosafety office.

### Modelling

We use a stochastic, individual-based model with non-overlapping generations in a randomly mating population of fixed carrying capacity. We use the population genetics modelling software SLiM version 4.0.1[47] in combination with R version 4.1.0[48]. Our code can be found on GitHub at https://github.com/NickyFaber/Two-target_drive.

**Drive types.** We have modelled 10 different types of homing suppression drives. For ease of visualisation, we have chosen to show the three most representative drives in the main results. A list of the rest of the drives and their modelling results are in the Supplementary Materials.

1. Drive with female fertility target
2. Haplolethal rescue drive with distant-site female fertility target
3. Haplosufficient rescue drive with distant-site female fertility target

For a detailed explanation of the mechanics and phenotypes of these three drives, see the results. To model each drive, we include two loci in SLiM: the homing site and an optional distant site. At the homing site, there are four potential alleles: wild type, drive, r1 (functional resistance), and r2 (nonfunctional resistance). At the distant site, we only model wild-type and r2 alleles (full freedom to choose any set of gRNAs is assumed to reduce functional r1 resistance to negligible levels, see results). The structure and steps of the model are described below.

**Model structure.** We have used previous work by Champer et al. (2020) as a starting point for our modelling[19]. That study modelled complex drive activity in the germline, including gRNA multiplexing (the number of gRNAs used to target a site), timing of drive activity (early or late germline), and gRNA saturation (the effect that a limited supply of Cas9 reduces the cutting efficiency at each targeted site with an increasing number of targets). Since our objective is to compare various gene drive designs focusing on genetic load instead of on resistance alleles, we have removed some of those complexities for this study. We model the population and introduction of the gene drives as follows:

- Generation 1: Population Initialisation
- Generation 1-10: Population equilibration without gene drive
- Generation 11: Introduction of heterozygous gene drive individuals
- Generation 111 or 161: End of the model. As default, we run the model for 111 generations (100 generations with gene drive). When we calculate the genetic load, we run it for an additional 50 generations to increase the number of generations in which we can determine the equilibrium genetic load.

In each generation, the model executes the following steps in order (Fig. 10):

**(1) Reproduction.** In our model, the number of offspring that is generated by each female is a product of the fertility status of the female, her ability to find a mate, the carrying capacity, and the fitness of the female.

- Fertility status check. A female could be infertile due to the drive mechanism, both at the homing site and the distant site. These loci are both checked for the presence of nonfunctional alleles, that is, either a drive allele or an r2 allele. If the female has two of these in any combination at least one locus, she does not generate any offspring.
- Selecting a mate. A female randomly selects a male from the population. If there are no males available, she does not generate any offspring.
- Generating offspring. In this study, we model density-dependent fertility (rather than density-dependent mortality or a combination of the two). This is computationally efficient and still allows easy density regulation. See Dhole et al. (2020) for a discussion on how genetic load could translate to population size reductions for different species-specific scenarios[14]. Different models would be more suitable for particular species, but genetic load itself would be dependent on the drive mechanism, regardless of the exact form of density dependence. For each female, the number of offspring ($O$) she generates in that generation ($i$) is based on a binomial distribution:

$$O_i = B(O_{max}, p_i), \tag{1}$$

where $i$ is the current generation, $O_{max}$ is the maximum amount of offspring per female, and $p$ is the average fraction of this maximum amount of offspring that will be generated. This fraction is normally defined as $\frac{2}{O_{max}}$ (each female must generate two offspring to maintain the population at carrying capacity). However, our model includes two further influences. First, $p$ can be increased when the population is below carrying capacity because offspring will have more resources to survive, or vice versa, which we call the carrying capacity factor ($F_{CC}$). Second, $p$ can be decreased in female drive carriers due to somatic expression of the drive reducing fertility (because some wild-type female fertility alleles are disrupted in somatic cells where they are needed for fertility), which we call the somatic expression fertility factor ($F_{SEF}$). Thus, $p$ is defined as:

$$p_i = F_{CC_i} * F_{SEF_i} * \frac{2}{O_{max}}. \tag{2}$$

The carrying capacity factor is defined so that at very low population densities, it is close to the maximum growth rate ($r$), at carrying capacity ($K$), it is close to 1, and above carrying capacity, it is between 0 and 1. This leads to a logistic growth curve:

$$F_{CC_i} = \frac{r}{(r-1) * \frac{N_i}{K} + 1}, \tag{3}$$

where $N$ is the number of adults in the population.

Fertility scaling is done for females with at least one drive allele. Somatic expression of the drive can impact female fertility by prematurely disrupting wild-type fertility gene alleles. One of our modelled drives targets a female fertility gene at the homing site without rescue and the other two at the distant site. We only model somatic expression fitness costs for the drive sites without rescue. The total fertility cost is calculated per female as being additive as follows:

$$F_{SEF} = m^{H_{wt}} * m^{\frac{1}{D_{wt}}}, \tag{4}$$

where $m$ is the somatic expression fertility cost multiplier, from 0 (complete sterility) to 1 (no fertility cost), $H_{wt}$ is the number of wild-type alleles at the homing female fertility site (which can only be 0 or 1, since the other allele must be the drive), and $D_{wt}$ is the number of wild-type alleles at the distant female fertility target. At the distant locus, individuals with two wild-type alleles will be less impacted by the fertility cost, since it is less likely that both copies of the fertility gene are disrupted due to somatic expression.

**(2) Gene drive activity.** After all the offspring are generated, each offspring's genotype is modified based on the drive activity in the parents' germline. Drive activity in the parental germlines is modelled for each offspring independently. Cutting, homing, and the creation of resistance alleles are stochastic.

- Germline gene drive activity. For both parents of each offspring, the presence of the drive in their genome is checked. If at least one copy of the drive is present, the gene drive is active in the germline. First, the cut-rate is a parameter in our model, but it can be impacted by gRNA saturation as follows:

$$P_{c_s} = 1 - (1 - P_c)^{\frac{s}{s+l-1}}, \quad (5)$$

where $P_{c_s}$ is the cut-rate adjusted for gRNA saturation, $l$ is the number of loci, either 1 or 2 depending on whether the drive targets a distant site in addition to the homing site (we assume equal gRNA multiplexing at both sites), $P_c$ is the global cut-rate, and $s$ is the Cas9 saturation factor, which can range from 1 (which means 1 gRNA is enough to saturate all Cas9 proteins, so the cut-rate for each gRNA rapidly declines as more gRNAs are added) to infinite (which means no amount of gRNAs is enough to saturate the Cas9 proteins, so the cut-rate at each gRNA remains the same regardless of the number of gRNAs). If a randomly generated number between 0 and 1 is higher than this cut-rate, cutting occurs. We do this separately for the distant site alleles as well, if present. At the distant site, cutting always results in an r2 allele. We made this choice based on the following assumptions: (1) with a high cut rate, wildtype loci on both chromosomes are likely to be cut simultaneously, reducing the chances of successful homology-directed repair, (2) alleles repaired back to wildtype can be re-cut multiple times, with a lower likelihood of homology-directed repair occurring each time, and (3) if the cut rate is high, most gene drive individuals will inherit one r2 allele from their drive parent, making it more likely that homology-directed repair will produce a second r2 allele. Our experiments in *D. melanogaster* showed that most offspring had r2 alleles at the *yellow-g* locus, indicating that homology-directed repair back to a wild-type allele is rare in our system. At the drive locus, however, homing can occur if there is successful cutting based on the homing success rate ($P_h$). The homing success rate is defined at the beginning of the model based on the conversion rate ($P_{conv}$), which is a parameter:

$$P_h = \frac{P_{conv}}{P_c}, \quad (6)$$

where $P_{conv}$ can only be as high as $P_c$. If homing is successful, again based on $P_h$, the allele is converted to a drive allele. If there was cutting, but no homing, the locus is converted into an r1 allele with a probability equal to the r1 formation rate. Otherwise, it will become an r2 allele.

- Embryo gene drive activity. In the early embryo, we model maternal deposition of drive Cas9 and gRNAs. Potential cuts that occur here always result in a resistance allele (which can be r1 or r2 as above). The cut-rate ($P_{e_s}$) is calculated the same as in Formula (5), except the exponent is additionally multiplied by a maternal deposition factor ($d$) that accounts for either a drive heterozygous or homozygous mother:

$$P_{e_s} = 1 - (1 - P_e)^{\frac{s*d}{s+l-1}}, \quad (7)$$

where $P_e$ is the embryo cut rate, which is a parameter in our model. The maternal deposition factor is based on experimental data showing that the cut-rate is higher in embryos with a drive/wild-type heterozygous mother than expected due to drive conversion, so $d$ is 1.83 and 2 for drive/wild-type heterozygotes and drive homozygotes, respectively. It would remain 1 for fertile drive/resistance allele heterozygotes[19].

- Offspring viability. After drive activity in the parental germline, inheritance, and embryo activity, we check if the offspring's resulting genotype is still viable. All drives that target either a haplolethal gene (where any nonfunctional resistance allele makes individuals nonviable) or a haplosufficient gene (where only nonfunctional resistance allele homozygotes are nonviable) could result in nonviable offspring. These offspring are removed from the population.

**(3) Mortality.** We model discrete, non-overlapping generations, so we remove the entire parental generation after offspring have been generated.

**(4) Calculating genetic load.** Because our model is stochastic, populations can be suppressed if the genetic load is close to but lower than the deterministic requirement, which is below 1 and depends on the maximum growth rate of the population. Therefore, in order to calculate genetic load with precision for drives with high genetic loads, we run a module in the model in which we artificially raise the number of offspring a female produces by dividing it by a certain bonus factor (that ranges between 0 and 1)[19]. These are later corrected for when we calculate the genetic load. Thus, the bonus factor has no biological meaning and is purely intended to precisely measure an outcome under stochastic variation. With this approach, the population is not eliminated unless the genetic load is practically 1, allowing for precise measurements of high genetic load for several generations when the drive is at its equilibrium frequency. This bonus factor ($F_B$) is calculated as follows:

$$F_{B_i} = \frac{N_{FF_i}}{N_{F_i}}, \quad (8)$$

where $N_F$ is the number of females, and $N_{FF}$ is the number of fertile females.

Then, in the next generation, the number of offspring calculated in Formula (1) is increased as follows:

$$O_{F_{B_i}} = \frac{O_i}{F_{B_i}}, \quad (9)$$

rounded to a whole number. In addition, in genetic load simulations, gene drive carriers are introduced at 0.5 frequency, and the model is run for 150 generations after introducing the gene drive. The mean genetic load is calculated as the mean genetic load over the last 10 generations. The low-density growth rate and the maximum number of offspring are set at a factor 10 higher than default (so 100 and 500 instead of 10 and 50, respectively) in order to calculate the genetic load precisely despite stochasticity.

**(5) Tracking outcomes of interest**. For each generation, we calculate population size, genetic load, and genotype frequencies at both the homing site and the distant site. We calculate the genetic load ($G$) based on the observed and expected population size in the next generation ($N_{i+1}$ and $N_{exp_{i+1}}$, respectively):

$$G_i = 1 - \frac{N_{i+1} * F_{B_i}}{N_{exp_{i+1}}}, \quad (10)$$

where $F_B$ is the above mentioned bonus factor we apply.

$N_{exp}$ is based on the number of females ($N_F$) and the carrying capacity factor $F_{CC}$ defined in Equation (3):

$$N_{exp_i} = 2 * N_{F_i} * F_{CC_i}, \quad (11)$$

where $F_{CC}$ is the same carrying capacity factor defined in Equation (3), and again multiplying by 2 because each female must generate two offspring to maintain the population at carrying capacity.

## Experimental work

**Plasmid construction.** The starting plasmids TTTgRNAt and TTTgRNAtRNAi were used for building gRNA helper plasmids for knock-in[36]. The gRNA cassette used in the donor plasmid was obtained from HSDygU4[25]. A two-step assembly process was done to generate donor plasmids (SI Appendix, Methods). Q5 High-Fidelity DNA Polymerase used for PCR and enzymes for digestion were purchased from New England Biolabs. PCR and restriction digestion products were purified with Zymo Research Gel DNA Recovery Kit, and plasmids were assembled by using HiFi DNA Assembly Cloning Kit and subsequently transformed into DH5$\alpha$ competent cells from TIANGEN. ZymoPure Midiprep Kit (Zymo Research) was used to prepare donor constructs for embryo injection. Oligo synthesis and Sanger sequencing was done by BGI Genomics. All the primers, plasmids, and construction procedures used in this study can be found in the Supplementary Information. Plasmid maps are available on GitHub (https://github.com/NickyFaber/Two-target_drive) in GenBank format created using ApE software[49].

**Generation of transgenic lines.** All of the transgenic lines were generated via CRISPR/Cas9 homology-directed repair. *D. melanogaster* injections were completed by UniHuaii, with a mixture consisting of the donor plasmid, Cas9 helper plasmid, and gRNA helper plasmid targeting the insertion site. To be specific, the donor plasmid HSDrgU2U4v2 (518 ng/ul) was injected into AHDr352v2 flies[36] along with TTTrgU2t (150 ng/ul), which provided gRNAs for transformation, and Cas9-expressing helper plasmid TTChsp70c9 (450 ng/ul) to generate the drive line. The other donor plasmid SNc9NR (506 ng/ul) was injected into $w^{1118}$ flies along with helper plasmids BHDabg1 (100 ng/ul) and TTChsp70c9 (459 ng/ul) to construct a *nanos*-Cas9 line. Surviving G0 flies were crossed to $w^{1118}$ flies, and G1 adults were screened for transgenic inserts based on the presence of green or red fluorescence in the eyes. Flies were reared in an incubator at 25 °C following a 14/10 h day/night cycle.

For phenotyping, flies were first anaesthetised with $CO_2$ and then screened for fluorescence using the NIGHTSEA system (SFA-GR). The homozygosity of flies was scored by the fluorescence intensity and confirmed by sequencing.

**Drive conversion.** Drive (gRNA-expressing line) males were crossed to Cas9 females to generate heterozygous offspring, which was subsequently out-crossed to $w^{1118}$. The drive inheritance and sex of offspring were recorded. To confirm whether the distant target site in *yellow-g* was disrupted, individuals containing both drive and Cas9 alleles were randomly collected for genomic DNA extraction and genotyping. A fragment covering *yellow-g* target sites was amplified with primers 52_YGLeft_S_F3 and 54_YGRight_S_R6 (see GitHub DNA files).

**Small cage study.** Drive heterozygous males with red fluorescence from the Cas9 allele were crossed to homozygous Cas9 females for several generations to produce a line that was heterozygous for the drive and homozygous for Cas9 (D/+; Cas9/Cas9). Two experimental groups were set up with different initial release frequencies. In the higher release frequency group, four drive females were crossed to four drive males, while in the medium release frequency group, four drive females were crossed with four drive males, and one Cas9 virgin female was crossed to Cas9 males. Thus, the high drive frequency release (generation number 0) was 0.5 (1.0 carrier frequency), and the medium drive frequency release was 0.4 (0.8 carrier frequency). These adults were allowed to mate in vials for one day before moving females into a separate bottle for oviposition. Females were allowed to lay eggs (which represented "generation 0") for three days and were then removed from bottles. When most pupae eclosed to adults, they were moved to a new bottle for a one-day oviposition before being removed and phenotyped. Hereafter, only one-day oviposition was conducted in each generation. The adults of each generation were scored for eye fluorescence phenotype and sex.

**Fecundity and fertility test.** To minimise batch effects caused by food quality or population density flies with different genotypes used for this test were generated from the same parental cross and reared in the same bottle. First, males that were heterozygous for the drive and homozygous for Cas9 were crossed to Cas9 homozygous females, generating offspring with different genotypes. Next, these offspring were individually crossed to Cas9 homozygous males or females and allowed to lay eggs for three days. Adults in the same vial were moved to a new vial each day, and the number of eggs was counted in each vial. Offspring were allowed to hatch, and the egg-to-adult survival rate and adult phenotypes of these offspring were scored. Female drive offspring were randomly collected and crossed to Cas9 males, after which the sterile females were genotyped for the *yellow-g* distant site.

**Statistics & reproducibility.** All samples are listed in the figure captions and Source Data Sets. No data were excluded from the analyses. No statistical method was used to predetermine the sample size. The experiments were not randomised because the experimenter needed to set up and track specific crosses. The investigators were not blinded to allocation during experiments and outcome assessment because the main experimenter needed to set up and track specific crosses.

**Phenotype data analysis.** For the statistical analysis of our experimental data, we fit generalised linear mixed-effects models containing our fixed experimental factors (genotype of the cross, drive individuals' sex, and transformed line 1A or 1D), as well as a random effect to correct for batch effects (each vial was considered an independent batch). Briefly, we fit two models with a binomially distributed response variable using a logit link and maximum likelihood estimation: one for gene drive inheritance and one for egg-to-adult survival. First, we fit a full model with all fixed effects and interactions and compared what random effect structure fits the data best. In the best model for gene drive inheritance, the random batch effect depended on the genotype (control crosses showed normal Mendelian inheritance, whereas the drive crosses showed increased variance in inheritance rates, which is likely an effect of pre-meiotic germline drive activity). In the best model for egg-to-adult survival, the batch effect depended on the sex of the drive parents (which is likely an effect of maternal deposition of Cas9 and gRNAs, though this data shows an interesting binary-like distribution). Then, we sequentially removed the least significant fixed effect from each model (starting with interactions) until only significant effects were left. To calculate the mean and standard error for each group, we fit a model with the same random batch effect, but with each group as a fixed effect. The analysis was performed using R (4.4.0) with support from the glmmTMB (1.1.9)[50] and DHARMa (0.4.6)[51] packages. The code and model summaries are available on GitHub (https://github.com/NickyFaber/Two-target_drive).

**Whole genome sequencing.** From lines 1A and 1D each, 25 flies were pooled for DNA isolation using QIAGENs DNeasy Blood & Tissue Kit. A Nanopore library was prepared and sequenced on a MinION Mk1C (Oxford Nanopore). We used built-in adaptive sampling software to enrich the 8 Mb region around the gene drive on chromosome 3R[52]. An overall sequencing depth of 40-50X coverage was reached in non-enriched regions of the genome, while around 70X coverage was obtained in the enriched region. Base-calling was done using Oxford Nanopore's basecaller Dorado[53], de novo genomes were assembled

using Flye software[54], and reads were mapped to validate and manually curate the assembly around the gene drive sequences using IGV[55]. Annotation of the consensus sequence was done manually in ApE[49]. Annotated sequences for both lines are available in GenBank format on GitHub (https://github.com/NickyFaber/Two-target_drive), and raw sequencing data is deposited on the European Nucleotide Archive (PRJEB80216).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

All data supporting the findings of this study are available within the paper and its Source Data files. Our code is on GitHub (https://github.com/NickyFaber/Two-target_drive), and sequencing data is on the European Nucleotide Archive under accession code PRJEB80216. Source data are provided in this paper.

## Code availability

All code supporting the findings of this study is available on GitHub (https://github.com/NickyFaber/Two-target_drive)[56].

## References


1. Esvelt, K. M., Smidler, A. L., Catteruccia, F. & Church, G. M. Concerning RNA-guided gene drives for the alteration of wild populations. *ELife*. **3**, e03401 (2014).
2. Champer, J., Buchman, A. & Akbari, O. S. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nat. Rev. Genet*. **17**, 146–159 (2016).
3. Godfray, H. C. J., North, A. & Burt, A. How driving endonuclease genes can be used to combat pests and disease vectors. *BMC Biol*. **15**, 81 (2017).
4. Bier, E. Gene drives gaining speed. *Nat. Rev. Genet*. **23**, 5–22 (2022).
5. Wang, G. H. et al. Symbionts and gene drive: two strategies to combat vector-borne disease. *Trends Genet*. **38**, 708–723 (2022).
6. Alphey, L. S., Crisanti, A., Randazzo, F. F. & Akbari, O. S. Opinion: Standardizing the definition of gene drive. *Proc. Natl. Acad. Sci*. **117**, 30864–30867 (2020).
7. Gantz, V. M. et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. *Proc. Natl. Acad. Sci*. **112**, E6736–E6743 (2015).
8. Kyrou, K. et al. A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nat. Biotechnol*. **36**, 1062–1066 (2018).
9. Prowse, T. A. A. et al. Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. *Proc. R. Soc. B Biol. Sci*. **284**, 20170799 (2017).
10. Gierus, L. et al. Leveraging a natural murine meiotic drive to suppress invasive populations. *Proc. Natl. Acad. Sci*. **119**, e2213308119 (2022).
11. Wilkins, K. E., Prowse, T. A. A., Cassey, P., Thomas, P. Q. & Ross, J. V. Pest demography critically determines the viability of synthetic gene drives for population control. *Math. Biosci*. **305**, 160–169 (2018).
12. Champer, S. E. et al. Modeling CRISPR gene drives for suppression of invasive rodents using a supervised machine learning framework. *PLOS Comput. Biol*. **17**, e1009660 (2021).
13. Esvelt, K. M. & Gemmell, N. J. Conservation demands safe gene drive. *PLOS Biol*. **15**, e2003850 (2017).
14. Dhole, S., Lloyd, A. L. & Gould, F. Gene drive dynamics in natural populations: The importance of density dependence, space, and sex. *Annu. Rev. Ecol. Evol. Syst*. **51**, 505–531 (2020).
15. Price, T. A. R. et al. Resistance to natural and synthetic gene drive systems. *J. Evol. Biol*. **33**, 1345–1360 (2020).
16. Hammond, A. M. et al. The creation and selection of mutations resistant to a gene drive over multiple generations in the malaria mosquito. *PLOS Genet*. **13**, e1007039 (2017).
17. Pfitzner, C. et al. Progress toward zygotic and germline gene drives in mice. *CRISPR J*. **3**, 388–397 (2020).
18. Grunwald, H. A. et al. Super-Mendelian inheritance mediated by CRISPR-Cas9 in the female mouse germline. *Nature* **566**, 105–109 (2019).
19. Champer, S. E. et al. Computational and experimental performance of CRISPR homing gene drive strategies with multiplexed gRNAs. *Sci. Adv*. **6**, eaaz0525 (2020).
20. Champer, J. et al. Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations. *PLOS Genet*. **13**, e1006796 (2017).
21. Carrami, E. M. et al. Consequences of resistance evolution in a Cas9-based sex conversion-suppression gene drive for insect pest management. *Proc. Natl. Acad. Sci*. **115**, 6189–6194 (2018).
22. Pham, T. B. et al. Experimental population modification of the malaria vector mosquito, Anopheles stephensi. *PLOS Genet*. **15**, e1008440 (2019).
23. Rode, N. O., Estoup, A., Bourguet, D., Courtier-Orgogozo, V. & Débarre, F. Population management using gene drive: molecular design, models of spread dynamics and assessment of ecological risks. *Conserv. Genet*. **20**, 671–690 (2019).
24. Oberhofer, G., Ivy, T. & Hay, B. A. Behavior of homing endonuclease gene drives targeting genes required for viability or female fertility with multiplexed guide RNAs. *Proc. Natl. Acad. Sci*. **115**, E9343–E9352 (2018).
25. Yang, E. et al. A homing suppression gene drive with multiplexed gRNAs maintains high drive conversion efficiency and avoids functional resistance alleles. *G3*. **12**, jkac081 (2022).
26. Hammond, A. et al. Regulating the expression of gene drives is key to increasing their invasive potential and the mitigation of resistance. *PLOS Genet*. **17**, e1009321 (2021).
27. Carballar-Lejarazú, R. et al. Next-generation gene drive for population modification of the malaria vector mosquito, Anopheles gambiae. *Proc. Natl. Acad. Sci*. **117**, 22805–22814 (2020).
28. Du, J. et al. Germline Cas9 promoters with improved performance for homing gene drive. *Nat. Commun*. **15**, 4560 (2024).
29. Hammond, A. et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. *Nat. Biotechnol*. **34**, 78–83 (2016).
30. Li, M. et al. Development of a confinable gene drive system in the human disease vector Aedes aegypti. *ELife*. **9**, e51701 (2020).
31. Reid, W. et al. Assessing single-locus CRISPR/Cas9-based gene drive variants in the mosquito Aedes aegypti via single-generation crosses and modeling. *G3* **12**, jkac280 (2022).
32. Chan, Y. S., Huen, D. S., Glauert, R., Whiteway, E. & Russell, S. Optimising homing endonuclease gene drive performance in a semi-refractory species: The drosophila melanogaster experience. *PLoS ONE* **8**, e54130 (2013).
33. Carballar-Lejarazú, R., Tushar, T., Pham, T. & James, A. A. Cas9-mediated maternal effect and derived resistance alleles in a gene-drive strain of the African malaria vector mosquito, Anopheles gambiae. *Genetics*. **221**, iyac055 (2022).
34. Anderson, M. A. E. et al. Closing the gap to effective gene drive in Aedes aegypti by exploiting germline regulatory elements. *Nat. Commun*. **14**, 338 (2023).
35. Yadav, A. K. et al. CRISPR/Cas9-based split homing gene drive targeting doublesex for population suppression of the global fruit pest Drosophila suzukii. *Proc. Natl. Acad. Sci*. **120**, e2301525120 (2023).
36. Champer, J. et al. A CRISPR homing gene drive targeting a haplolethal gene removes resistance alleles and successfully spreads

through a cage population. *Proc. Natl. Acad. Sci.* **117**, 24377–24383 (2020).
37. Kandul, N. P. et al. Assessment of a split homing based gene drive for efficient knockout of multiple genes. *G3* **10**, 827–837 (2020).
38. Champer, J., Champer, S. E., Kim, I. K., Clark, A. G. & Messer, P. W. Design and analysis of CRISPR based underdominance toxin antidote gene drives. *Evol. Appl.* **14**, 1052–1069 (2021).
39. Champer, J., Kim, I. K., Champer, S. E., Clark, A. G. & Messer, P. W. Performance analysis of novel toxin-antidote CRISPR gene drive systems. *BMC Biol.* **18**, 27 (2020).
40. Zhu, Y. & Champer, J. Simulations reveal high efficiency and confinement of a population suppression CRISPR toxin-antidote gene drive. *ACS Synth. Biol.* **12**, 809–819 (2023).
41. Noble, C. et al. Daisy-chain gene drives for the alteration of local populations. *Proc. Natl. Acad. Sci.* **116**, 8275–8282 (2019).
42. Metzloff, M. et al. Experimental demonstration of tethered gene drive systems for confined population modification or suppression. *BMC Biol.* **20**, 119 (2022).
43. Sudweeks, J. et al. Locally fixed alleles: A method to localize gene drive to island populations. *Sci. Rep.* **9**, 15821 (2019).
44. Saada, A. A., Lambert, S. A. & Carr, A. M. Preserving replication fork integrity and competence via the homologous recombination pathway. *DNA Repair*. **71**, 135–147 (2018).
45. Hou, S et al. A homing rescue gene drive with multiplexed gRNAs reaches high frequency in cage populations but generates functional resistance. *J. Genet. Genom*. (2024).
46. Spinner, S.A. et al. New self-sexing Aedes aegypti strain eliminates barriers to scalable and sustainable vector control for governments and communities in dengue-prone environments. *Fronti. Bioeng. Biotechnol.* **10**, (2022).
47. Haller, B. C. & Messer, P. W. SLiM 4: Multispecies eco-evolutionary modeling. *Am. Nat.* **201**, E127–E139 (2022).
48. R Core Team. R: A language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (2024).
49. Davis, M. W. & Jorgensen, E. M. ApE, a plasmid editor: a freely available DNA manipulation and visualization program. *Front. Bioinform.* **2**, 818619 (2022).
50. Brooks, M. E. et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R. J.* **9**, 378–400 (2017).
51. Hartig, F, Hartig, M.F. Package 'DHARMa'. R package. (2017).
52. Payne, A. et al. Readfish enables targeted nanopore sequencing of gigabase-sized genomes. *Nat. Biotechnoli.* **39**, 442–450 (2021).
53. Oxford Nanopore Technologies. Dorado. https://github.com/nanoporetech/dorado (2023).
54. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
55. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
56. Faber, N.R. Code accompanying this paper. Repository NickyFaber/Two-target_drive: Code with manuscript. https://doi.org/10.5281/zenodo.13759565 (2024).

## Acknowledgements

This study was supported by grants from the Graduate School for Production Ecology & Resource Conservation (PE&RC) in the Netherlands (NRF), and laboratory startup funds from Peking University (J.Champer), the Centre for Life Sciences (J.Champer), the Li Ge Zhao Ning Life Science Youth Research Fund (J.Champer), and grants from the National Science Foundation of China (32302455 (XX) and 32270672 (J.Champer)).

## Author contributions

N.R.F. and J.Champer conceived the study. N.R.F. conducted the modelling. Experimental work was led by X.X., with assistance from J.Chen, S.H., J.D., and N.R.F. Analysis was completed by N.R.F. and X.X. J.Champer, B.A.P., J.H., and B.J.Z. provided supervision. N.R.F., X.X., and J.Champer wrote the manuscript, and all authors reviewed it.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-53631-5.

**Correspondence** and requests for materials should be addressed to Nicky R. Faber or Jackson Champer.

**Peer review information** *Nature Communications* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.