Supplementary Information for

Cleave and Rescue, a novel selfish genetic element and general strategy for gene drive

Georg Oberhofer, Tobin Ivy, Bruce A. Hay

Corresponding author: Bruce A. Hay
Email: haybruce@caltech.edu

This PDF file includes:
Supplementary methods
Supplementary methods references
Fig. S1 to S10
Table S1 to S8

Other Supplementary Materials for this manuscript include the following:

Dataset S1. **DataS1.xlsx**; alignment of full rescue construct, *ClvR* constructs fasta files, primers
Dataset S2. **DataS2.xlsx**; genotype count table for drive experiments.
Dataset S3. **ClvR_2lAX.txt;** modeling Matlab source code, deterministic model of AX-ClvR, Autosomal element cleaving target gene on X
Dataset S4. **ClvR_2lAX_FCvIF.txt;** modeling Matlab source code, deterministic model of AX-ClvR, script calling the ClvR_2lAX.txt function to generate all AX-ClvR figures
Dataset S5. **ClvR_2lAA_multirun.txt;** modeling Matlab source code, deterministic model of the 2 locus, autosomal element cleaving target gene at separate autosomal locus, plots multiple single runs
Dataset S6. **ClvR_2lAA_multirun_support.txt;** modeling Matlab source code, script calling the ClvR_2lAA_multirun.txt function to generate all AA-ClvR figures
Dataset S7. **ClvR_2lAA_FCvIF.txt;** modeling Matlab source code, deterministic model of the 2 locus, autosomal element cleaving target gene at separate autosomal locus across a range of fitness costs and release ratios
Dataset S8. **ClvR_2lAA_FCvIF_support.txt;** modeling Matlab source code, function that supports ClvR_2lAA_FCvIF script by handling the calculations for a single run under a single set of parameters
Dataset S9. **ClvR_2lAX_wRwH.txt;** modeling Matlab source code, deterministic model of AX-ClvR, Autosomal element cleaving a target gene on X, with resistant alleles appearing at a frequency of 10^{-6}, in the absence or presence of homing

www.pnas.org/cgi/doi/10.1073/pnas.1816928116

Supplementary Methods

Cloning of *ClvR* constructs and fly germline transformation

We started with a plasmid having a dU6:3 promoter and a modified guide scaffold (1) separated by *BsmBI* cutsites. This is derived from our previous work (2), which was itself based on pCFD3-dU6:3gRNA, a gift from Simon Bullock (Addgene plasmid # 49410) (3). We digested with *BsmBI* and ligated annealed oligos (P1-68E FWD + P2-68E REV) as described on http://flycrispr.molbio.wisc.edu. The gRNA created target a region on the third chromosome (68E), which was chosen based on the location of an attP landing site in a widely used fly strain, zh-68E (4). Next, the plasmid was cut with *HindIII* and *SpeI* and the following 4 fragments were assembled in a Gibson reaction to yield plasmid tko-A (see *SI Appendix,* Fig. S4*A*):

Two homology arms, approximately 1kb in length up and downstream of the above gRNA target site, were amplified from genomic DNA with primers P3+P4 and P9+P10 (the second homology arm already had a *td-tomato* marker (5) driven by the ubiquitous *opie2* promoter (6) upstream of it); an attP site with primers P5+P6; and a 4.2 kb rescue fragment with primers P7+P8. The rescue fragment was based on the *tko* genomic region of *Drosophila virilis* (*Dv*), a distant *Drosophila* species (7). Additionally, 6 silent point mutations were introduced in the *Dv-tko* ORF to further reduce homology with *D. melanogaster tko*. The rescue transgene was gene synthesized by IDT as two gBlock fragments with an additional 2 point mutations introduced in the intron to work around a synthesis complexity issue (all introduced point mutations are indicated as blue letters in the alignment in Dataset S1).

Construct *tko*-A was injected into a fly strain expressing Cas9 in the germline under the control of regulatory sequences derived from the nanos gene (8) (Bloomington stock# 54591) (3). All injections were carried out by Rainbow Transgenic Flies.

Male injected G0 flies were outcrossed to *w*- and the progeny were scored for ubiquitous *td-tomato* expression. Male transformants were crossed to a TM3,*Sb*/TM6b,*Tb* balancer stock. Flies carrying the *td-tomato* marker and *Sb* were pooled and used as the injection strain for the 2nd construct described below. Note that we used only males in the outcrosses to get rid of Cas9 expressed from the injection strain X-chromosome.

For construct *tko*-B we first subcloned two constructs having two gRNAs each. We digested construct pU6:3-U6:1-tandem (2) (based on (3)) with *BsmBI* and ligated back in two gRNAs encoded in the primer overhangs: P11-tko-guidesA FWD + P12-tko-guidesA REV and P13-tko-guidesB FWD + P14-tko-guidesB REV.

A plasmid that had a *3xP3-GFP* marker gene, an attB site as well as parts of nos-Cas9-nos flanked by gypsy insulators was digested with *EcoRV* and *BglII*. In a three fragment Gibson reaction, full length nos-Cas9-nos, as well as the two gRNA cassettes from above were assembled to yield the final tko-B construct. Cas9 was amplified with primers P15-nosCas9 FWD + P16-nos-Cas9 REV, guide cassette A with P17-guidesA FWD + P18-guidesA REV, and guide cassette B with P19-guidesB FWD + P20-guidesB REV (*SI Appendix,* Fig. S4*B*).

Construct *tko*-B was injected along with a phiC31 helper plasmid (Rainbow Transgenic Flies) with the balanced *tko*-A flies as the injection strain. Injected G0 flies were outcrossed to *w*[1118]

and the progeny were screened for *3xP3-GFP* expression. Transgenic males were used to cross to the balancer stock TM3,*Sb*/TM6b,*Tb* as well as w^{1118}. Flies carrying the GFP marker over TM3,*Sb* were pooled to generate the balanced stock and flies homozygous for the $ClvR^{tko}$ construct were collected in the next generation.

Fly crosses and husbandry of $ClvR^{tko}$ flies

Heterozygous $ClvR^{tko}$ flies used in the crosses detailed in Fig. 3*B,C* were collected from the offspring of $ClvR^{tko}$ males crossed to w^{1118} females. In general all heterozygous $ClvR^{tko}$ flies used to set up crosses shown in experiments were collected from progeny of a cross between $ClvR^{tko}$ males and w^{1118} females (except crosses to the GDL strains, see below).

For the eclosion rate experiment (Text and *SI Appendix,* Table S1*A,B*) we took 4 females and males of the corresponding genotypes (10 replicates each) and let them lay eggs overnight in a vial. The next day we scored the number of eggs, waited for them to eclose, and then scored the number of viable adults beginning 10 days later.

For gene drive experiment 1 (Fig. 4*A*), we initiated the drive population by crossing 5 populations of 20 heterozygous $ClvR^{tko}$ /+ males to 30 w^{1118} virgins. This corresponds to a *ClvR* allele introduction frequency of 25%, and a *ClvR*-bearing (ClvR/+ or ClvR/ClvR) genotype frequency of 50%. For drive experiment 2 (Fig. 4*B*) 25 wildtype (+/+) males and 25 homozygous $ClvR^{tko}$ males were crossed with 30 w^{1118} females in separate bottles for two days. Mated females were separated from males and introduced in a bottle at a 1:1 ratio, again corresponding to a 25% *ClvR* allele introduction frequency, and a *ClvR*-bearing genotype frequency of 25%. For both sets of experiments adults were allowed to lay eggs in bottles for two days, after which they were removed. Progeny were allowed to develop, eclose and mate for another 12 days. All adult progeny were collected at this single time point. They were anesthetized using CO_2, and their genotypes with respect to *ClvR* (presence or absence) determined using the dominant *td-tomato* marker. Note that adult progeny continued to eclose after the time of collection. These were not counted or transferred into the subsequent generation population. Following counting, progeny were transferred to fresh bottles and allowed to lay eggs for two days, and the cycle repeated (counts in Dataset S2).

For the crosses to the GDL lines we mated heterozygous $ClvR^{tko}$ virgins with males of the corresponding GDL line. In the next generation we scored the progeny for the *td-tomato* marker, set aside 30 virgins and backcrossed them again to wildtype males of the corresponding GDL line. This cycle was repeated for 6 generations.

Sequencing

For sequence analysis of the *D. melanogaster tko* locus we first extracted genomic DNA from single male flies using the Qiagen DNeasy kit. A 450 bp DNA fragment spanning all 4 gRNA target sites was amplified with primers tko-seq1 FWD and tko-seq1 REV and sent for Sanger sequencing (Laragen). Sequence files were aligned in Benchling with MAFFT (9). Amino acid alignments in *SI Appendix,* Fig. S3*B* were generated with Clustal Omega (10).

Modeling

We use a deterministic, discrete-generation population frequency framework to model the spread of $ClvR^{tko}$ through a population. This model is an adjustment of one we have previously used

(11) that consists of a series of difference equations to calculate the expected frequency of each genotype based on the frequencies of all genotypes from the previous generation, augmented by fitness effects, cleavage events within the germline and as a result of maternal carryover, and is finally adjusted by a normalization factor. Within the model we assume that there is perfectly random mating, females produce all of their offspring as a result of a single mating; all releases are distributed uniformly throughout the population and release ratios are with respect to the new total population size; all presented fitness costs are for bearing two copies of the *ClvR* element and are relative to wildtype (thus all heterozygotes bear half that fitness cost); the chance of cleavage of the target gene by germline activity or by maternal carryover are the same, independent of whether the parent has one or two copies of *ClvR*; when an individual has two native copies of the target gene there is no homology directed repair of a cleaved target gene to repair it to the wild type version (so always repaired via NHEJ). The difference equations governing the genotype frequencies can be found within the provided Matlab source files (Dataset S3-8). The model in Dataset S3 was used to generate the simulated drive experiments in Fig. 5 (with the specified parameters described there) and the single parameter set runs of Fig. 6*D,E*. The model in Dataset S4 was used to generate the heatmaps in Fig. 2*A,B* and Fig. 6*F*. The model in Dataset S5,6 was used to generate the multiple single-parameter set runs of Fig. 2*D*. The model in Dataset S7,8 was used to generate the heatmap in Fig. 2*C*. Dataset S9 was used to generate the single-parameter set runs of Fig. 6B,E. The term "Release frequency (%)" for all heatmaps refers to the frequency of homozygous transgenic males compared to wild type males and females after a release has occurred (e.g. a 40% release means that 40% of the population is *ClvR/ClvR* male, 30% is +/+ male, and 30% is +/+ female).

Supplementary Methods References

1. Dang Y, et al. (2015) Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol* 16:280.

2. Oberhofer G, Ivy T, Hay BA (2018) Behavior of homing endonuclease gene drives targeting genes required for viability or female fertility with multiplexed guide RNAs. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1805278115.

3. Port F, Chen H-M, Lee T, Bullock SL (2014) Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proceedings of the National Academy of Sciences* 111(29):E2967–E2976.

4. Bischof J, Maeda RK, Hediger M, Karch F, Basler K (2007) An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proc Natl Acad Sci U S A* 104(9):3312–3317.

5. Shaner NC, et al. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol* 22(12):1567–1572.

6. Theilmann DA, Stewart S (1992) Molecular analysis of the trans-activating IE-2 gene of Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus. *Virology* 187(1):84–96.

7. Drosophila 12 Genomes Consortium, et al. (2007) Evolution of genes and genomes on the Drosophila phylogeny. *Nature* 450(7167):203–218.

8. Van Doren M, Williamson AL, Lehmann R (1998) Regulation of zygotic gene expression in Drosophila primordial germ cells. *Curr Biol* 8(4):243–246.

9. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30(4):772–780.

10. Sievers F, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.

11. Marshall JM, Hay BA (2014) Medusa: a novel gene drive system for confined suppression of insect populations. *PLoS One* 9(7):e102694.

12. Pelletier JN, Campbell-Valois FX, Michnick SW (1998) Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc Natl Acad Sci U S A* 95(21):12141–12146.

13. Grenier JK, et al. (2015) Global diversity lines - a five-continent reference panel of sequenced Drosophila melanogaster strains. *G3* 5(4):593–603.

14. Champer J, et al. (2017) Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations. *PLoS Genet* 13(7):e1006796.

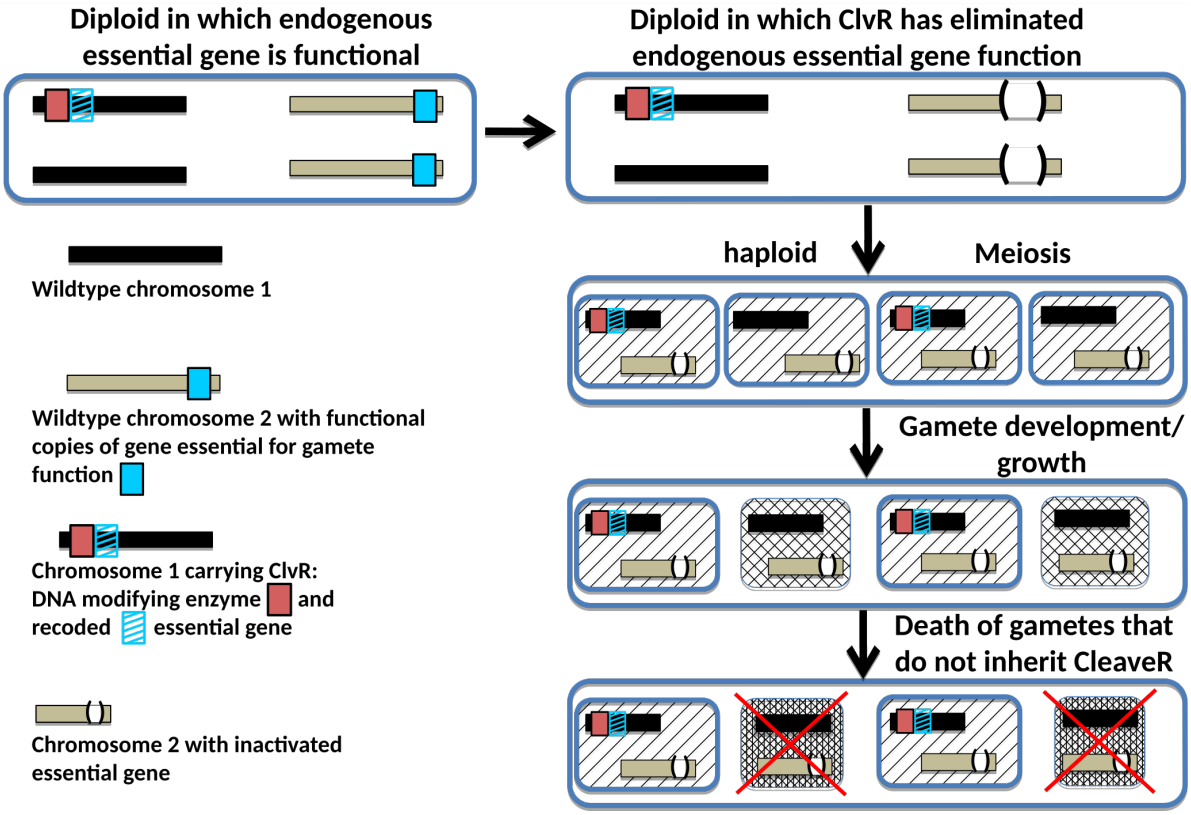

Fig. S1. *ClvR* can act as a gamete/spore killer. If expression of the essential gene being targeted is required in the products of meiosis, then those products that fail to inherit *ClvR* will die or otherwise be rendered non-functional (indicated by the cross-hatching that grows darker over time). The diploid survives regardless of whether the targeted gene is required generally or only in gametes, because it carries a recoded version of the essential gene (blue rectangle with diagonal lines). Haploid cells that inherit ClvR survive because they carry a recoded rescue that expresses during the haploid stage.

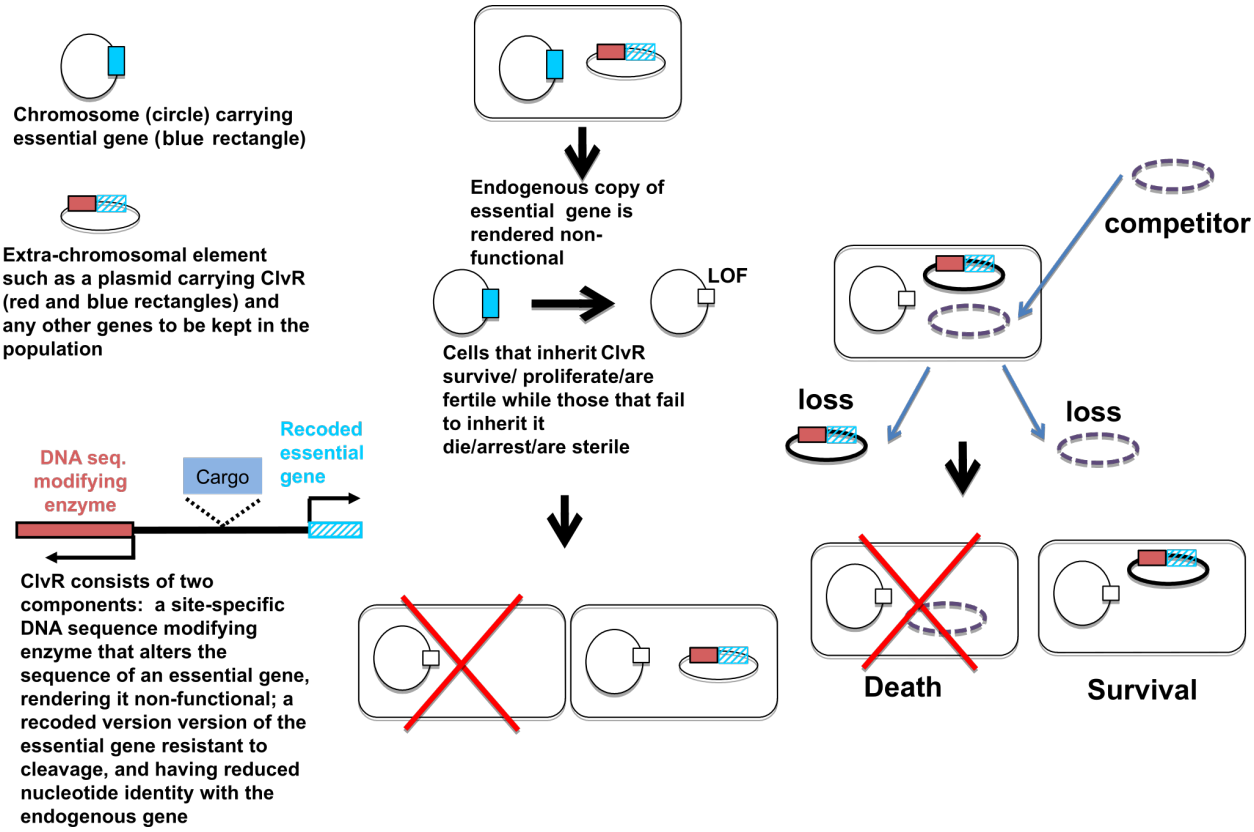

Fig. S2. *ClvR* promotes its own maintenance within a population. When *ClvR* is present on an extrachromosomal element that has some probability of not being transmitted to progeny, those who fail to inherit it die (Left). When competing genetic elements enter the same cellular environment as *ClvR*, individuals that inherit only the competitor die (Right). Similar points apply in the case of supernumerary or B chromosomes (not shown).

A

```
rescue          tacacttaacgcactcttgcaattccggacaagtcaagaggagacaactag--------a
Dm-tko          tatagtctagccactttctca----taatcaactaggagtgttacctctaaaaatgatca
                **.* *. *  ****.*. **    ... *** * .... *  **  ***.       *

rescue          cgttggcaatcggaaattggaagccttacagaaacactgcgtttataacttgttctcagc
Dm-tko          tgattgcaatctcaaacttggagcagt----taattttgttgtactggttttctttaagc
                .* * ******  ***.* *.***  *     **. .**.  *  *...** .*.* ***

rescue          tgtttctctctctcatcttgattacattgcagcgctgcagagcaattatttgtgtgccgc
Dm-tko          gacatatctaactc-tttcacttctttgccagccatccagtgcagctatgag---accgc
                 .. * ***  *** *.*.. ** . *  ****  * *** ***..***  *   .****

rescue          attgcgcggcatggcatcgttgaatcaaatgcatcgcactgggccgcatataaagaagcg
Dm-tko          cgtccgtggaatggcatcgctgcagcagatgcaccgcagcggaccgcacataaagacgcg
                  * **.** *********.** * **.*****.**** .**.*****.******* ***
                      guide1
rescue          tccgccacgtcagccCctggaCggtaaaccgtttgccaagggAgtGgtgctcaagacact
Dm-tko          tccgccgcgccagccgctggatggaaagcccttcgccaagggcgttgtcctgaagacgct
                ******.**.***** *****.** **.** **.******** ** ** ** *****.**
                                                guide2
rescue          gatcaagaagccaaagaaaccaaactcggcgaatcgtaaatgcgcgctggtgcgcttatc
Dm-tko          gataaagaagcccaagaagccgaactcggccaatcgaaagtgcgtgctggtgcgcctctc
                *** ******** *****.**.******** ***** **.****.**********.* **
                                                guide3
rescue          cacGggAaaggagatggtcgcctatatacccggcattggacataatctgcaggagcataa
Dm-tko          caccggcaaggagatggtggcctacatccccggcatcgggcacaacctgcaagagcacaa
                *** ** ********** *****.** ********.**.**.**.*****.*****.**

rescue          tattgtactgtgtcgcgtcggacgactgcaggatgtgcccggcgtcaagctgaaggcggt
Dm-tko          cattgtactgtgccgcgtgggggcgtctgcaggacgtgcccggcgttaagctgaaggctgt
                .**********.***** **.** ********.**********.********** **
                      guide4
rescue          gcgcggtgtctacgatctggcgcacgttatcaagaagggccaatgacaaccaactaccat
Dm-tko          gcGCGGAGTCTACGACTTGGCGcacgtcgTCAAGAAGAGCCAATAGTTAataaccacttc
                ****** ********..**********..********.******... * .***.**. .
                                                          STOP
```

B

```
Dvir-Tko-aa       --------------MNFLRQTFNVTKQLTAQALQSNYLCAALRGMASLNQMHRTGPHIKK
Dm-Tko-aa-C       MIMIAISNLEQLILLYWFSLSDISNSFTSLPAIQCSYE-TAVRGMASLQQMHRSGPHIKT
Dm-Tko-aa-B       --------------MNFLRQSFGITKQLASQAIQCSYE-TAVRGMASLQQMHRSGPHIKT
                                : ::  :   ..  :  *:*..*  :*:******:****:*****.

Dvir-Tko-aa       RPPRQPLDGKPFAKGVVLKTLIKKPKKPNSANRKCALVRLSTGKEMVAYIPGIGHNLQEH
Dm-Tko-aa-C       RPPRQPLDGKPFAKGVVLKTLIKKPKKPNSANRKCVLVRLSTGKEMVAYIPGIGHNLQEH
Dm-Tko-aa-B       RPPRQPLDGKPFAKGVVLKTLIKKPKKPNSANRKCVLVRLSTGKEMVAYIPGIGHNLQEH
                  ************************************.************************

Dvir-Tko-aa       NIVLCRVGRLQDVPGVKLKAVRGVYDLAHVIKKGQ*
Dm-Tko-aa-C       NIVLCRVGRLQDVPGVKLKAVRGVYDLAHVVKKSQ-
Dm-Tko-aa-B       NIVLCRVGRLQDVPGVKLKAVRGVYDLAHVVKKSQ-
                  *******************************:**.*
```

Fig. S3: Sequence alignments. (A) Shown are DNA alignments of the target region (*D.melanogaste*r *tko*; *Dm-tko*) and the *Dvir-tko* rescue (rescue). The target ORF is in red, the Rescue ORF in green, additionally modified bases in blue, gRNA target sites in pink, and the PAM in boldface. **(B)** Amino acid alignment of *D. virilis tko* (Dvir-Tko-aa) and the two annotated protein isoforms from *D. melanogaster* (Dm-Tko-aa-B and Dm-Tko-aa-C).

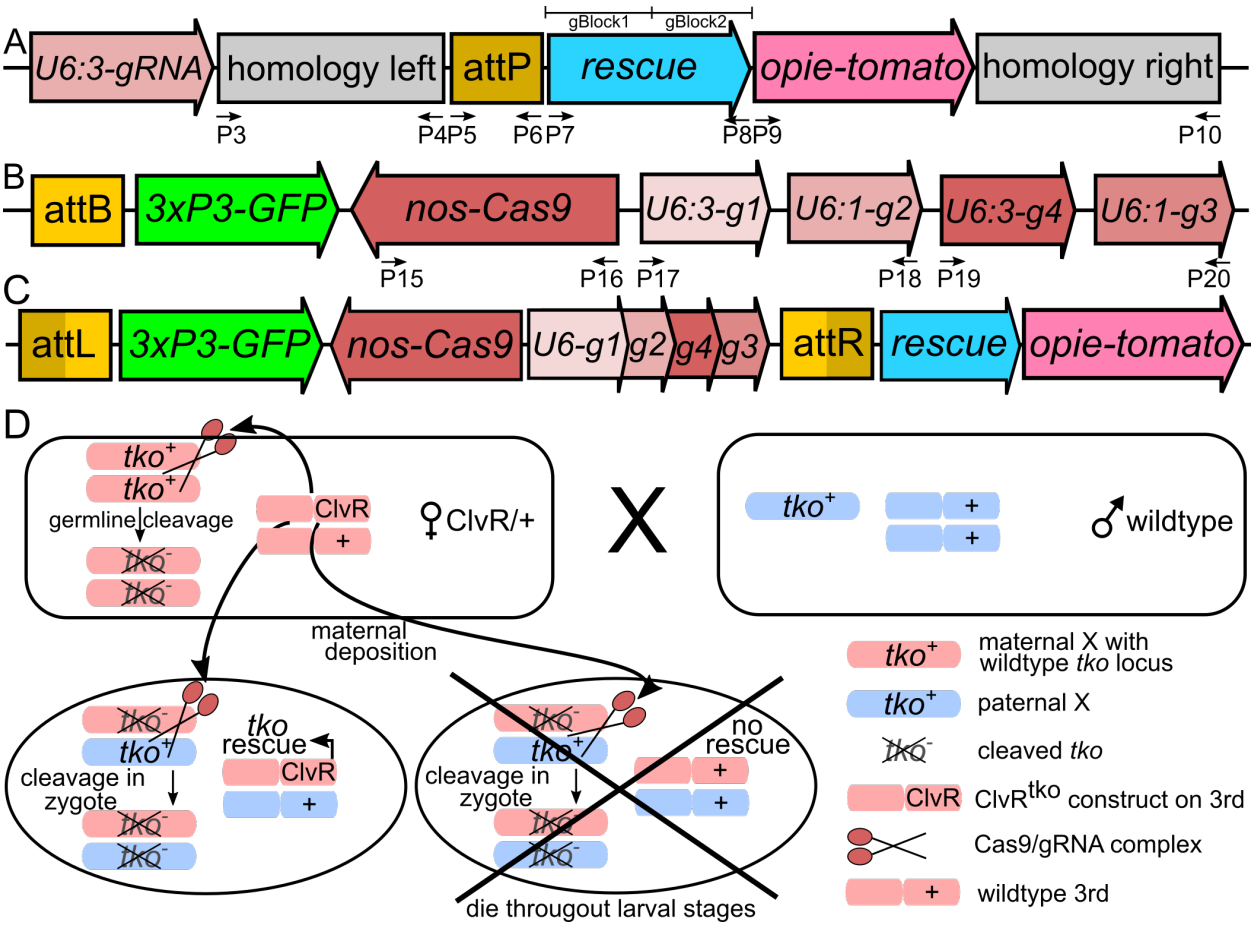

Fig. S4: *ClvR* construct design and principle. (A) Construct A with a U6:3-gRNA, an attP site, the *tko* rescue copy based on *Drosophila virilis tko* and a ubiquitous *opie2-td-tomato* marker. Only elements between the homology arms were inserted into a neutral site (68E) on the 3rd chromosome via Cas9 mediated HR. Cloning primers for Gibson assembly are indicated as arrows. **(B)** Construct B with an attB site, a *3xP3-GFP* marker, Cas9 driven by nanos regulatory elements, and a set of four U6 driven gRNAs. Construct B was integrated into the attP landing site of construct A via phiC31 integrase. **(C)** Final construct after B was integrated into A. **(D)** Principle by which *ClvR* acts. Females heterozygous for the *ClvR* construct create cleaved and LOF *tko* alleles in the germline. Additionally, active Cas9/gRNA complex is deposited maternally to all embryos (ovals in lower row), where subsequently paternal alleles are cleaved and rendered LOF (this conversion is indicated by the vertical arrow). Offspring without the Rescue copy from the *ClvR* element die (large X).

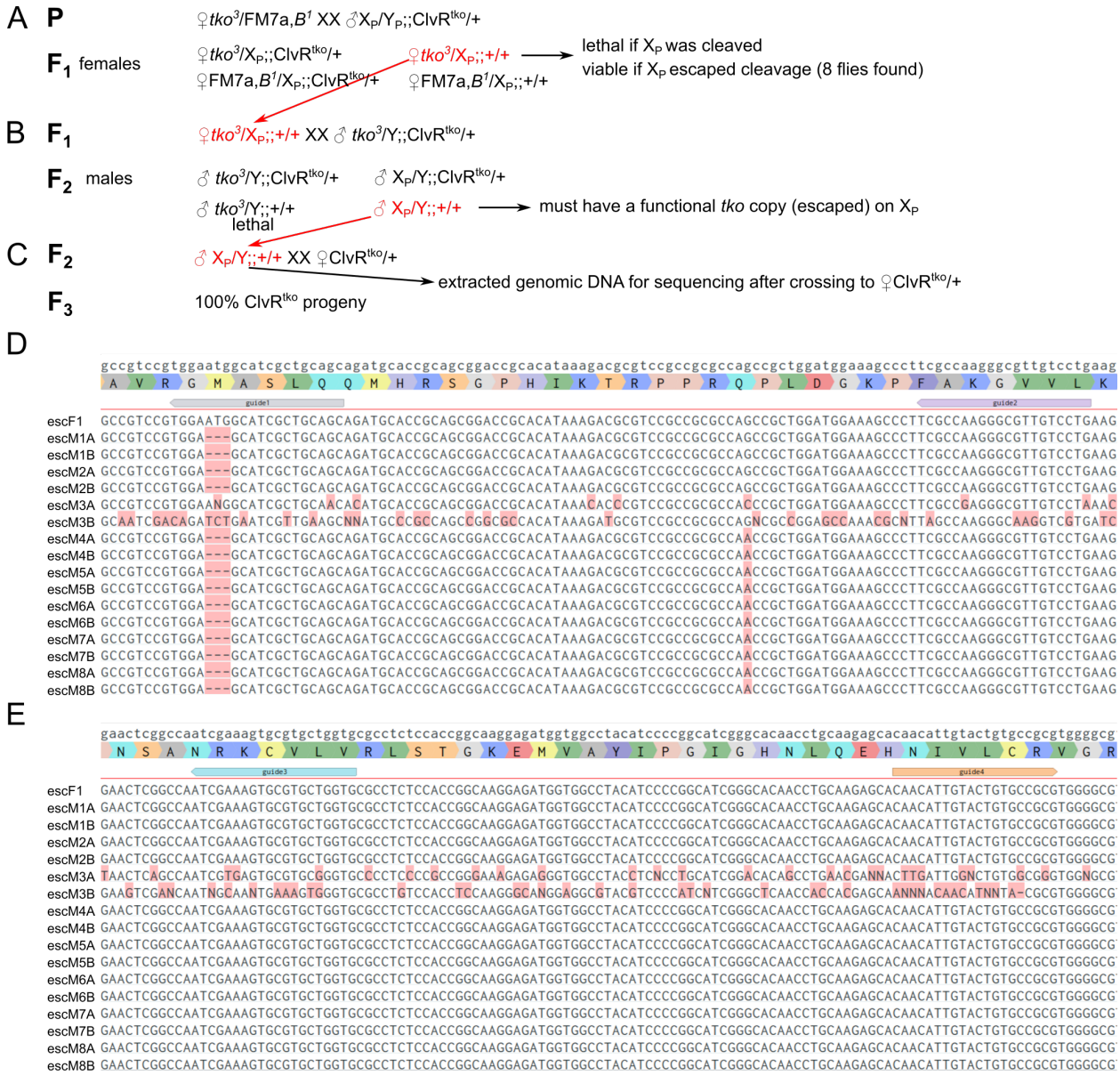

Fig. S5: (A-C) Mating scheme to isolate X chromosomes in which the *D. melanogaster tko* locus was not rendered non-functional (escapers) in the germline of male parents heterozygous for *ClvR*^{tko}. **(D-E)** Sequencing alignments to target sites 1,2 (D) and 3,4 (E). Escaper "escF1" from bottle 2 of female *ClvR*^{tko}/+ XX *w*^{1118} (see Table S2). Escapers M1-M3 from bottle 1, M4-M8 from bottle 2 of male *ClvR*^{tko} XX *tko*^{3}/FM7,*B*^{1} (see Table S3). Male escapers from bottle 2 have a common SNP (G to A between gRNA1 and gRNA2) not present in escapers from bottle 1. Thus, it is possible that the 8 isolates from males represent multiple isolates of two or more germline events. Note that the large number of sequence polymorphisms in escM3A and escM3B reflects ambiguous sequencing signal at a variety of positions. The basis for this remains unclear. We speculate that this reflects nuclear mosaicism, which could occur if the F1 *ClvR*^{tko}-bearing males provided some level of paternal carryover that altered the *tko* locus from the Xp chromosome in some nuclei of the F2 males used for sequencing and crosses to the *ClvR*^{tko}-bearing female.

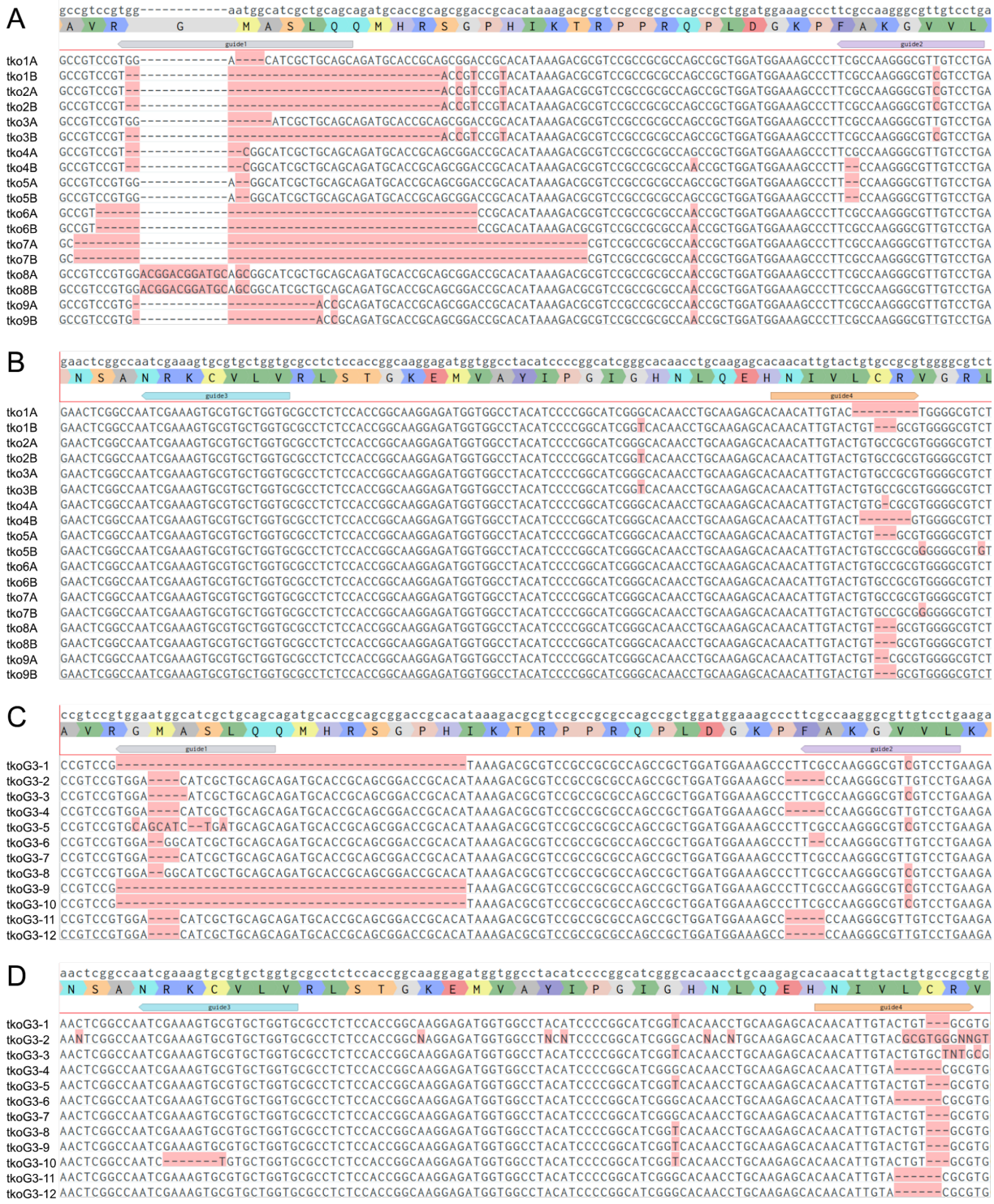

Fig. S6: Molecular analysis of cleavage events. Shown are the alignments of the *tko* locus of male progeny coming from *ClvR*^{tko}/+ mothers (two flies each, selected from 9 crosses, tko1A, tko1B, … tko9B) **(A,B)** or from a homozygous stock inbred for 3 generations (12 flies selected from bottles, tkoG3-1 to tkoG3-12) **(C,D)**. Alignments were split for ease of visibility. gRNA1 and gRNA2 target sites are shown in A and C, and gRNA3 and gRNA4 target sites in B and D. Top row shows the template with annotated gRNA target sites and amino acid sequence. Mismatches in the alignments are highlighted in red.

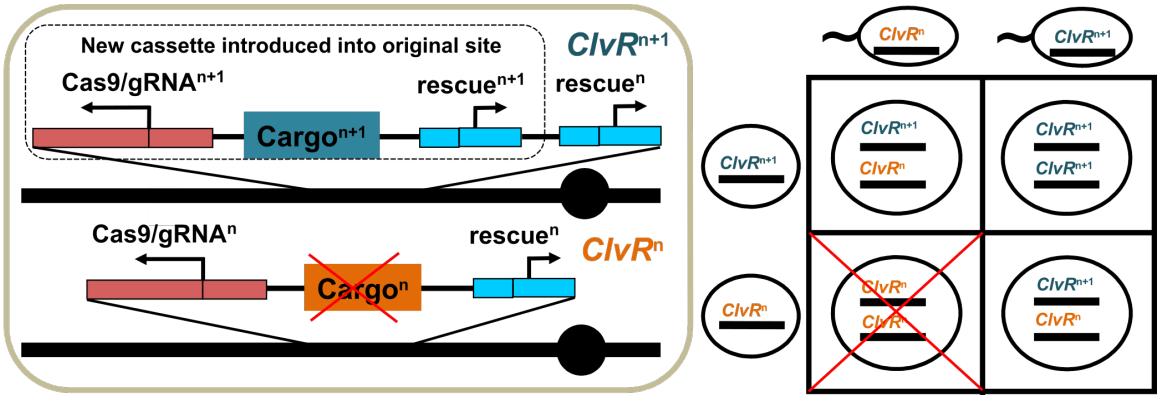

Fig. S7. Removal of a first generation ClvR, coupled with replacement by a second generation *ClvR* element. Multiple rounds of population replacement can be carried out by locating *ClvR*$^{n+1}$ at the same site as *ClvR*^{n}, with *ClvR*$^{n+1}$ targeting essential gene^{n+1}, while also carrying the original rescuing copy of essential gene^{n}. Because progeny carrying *ClvR*^{n} are sensitive to loss of essential gene^{n+1}, only those carrying *ClvR*$^{n+1}$survive, regardless of their status with respect to *ClvR*^{n}. The function of *ClvR*$^{n+1}$ can be made completely orthogonal to that of *ClvR*^{n} through the use of Cas9/gRNA variants from other species that cannot load the gRNAs generated by *ClvR*^{n}.

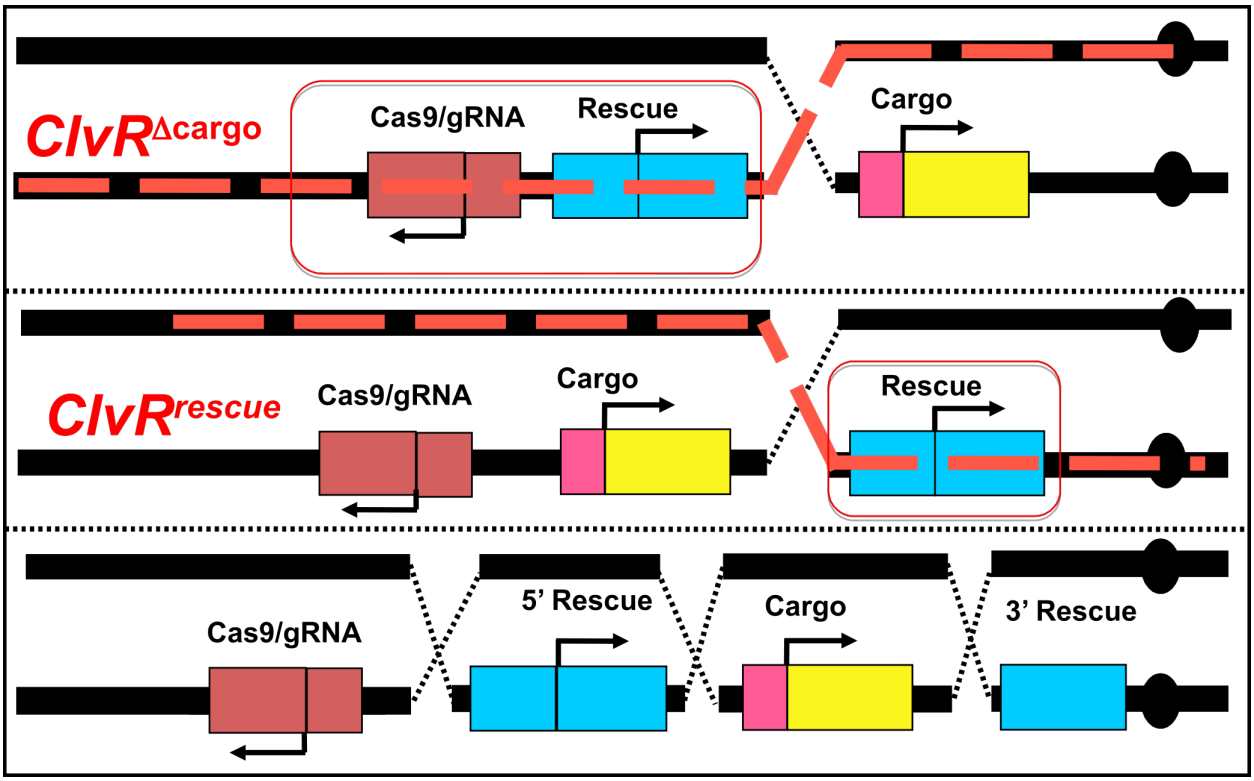

Fig. S8. Separation of a functional Rescue from the Cargo can be prevented by locating the Cargo in an intron of the Rescue. Cargo and recoded rescue will often have minimal homology with surrounding sequences on homologous chromosomes, and thus are unlikely to recombine away from each other through traditional homologous recombination during meiosis. However, a break between the two genes followed by reciprocal end joining with the same region on the homologous chromosome could potentially separate them, though the frequency of this kind of event is unclear. Locating the *ClvR* cargo in an intron of the Rescue transgene (bottom panel) prevents breakage and end joining-mediated separation of a functional Rescue (the key component driven into the population by *ClvR*) from the Cargo. Separation could otherwise generate empty *ClvR* elements ($ClvR^{\Delta cargo}$, top panel), or Rescue only elements ($ClvR^{rescue}$, middle panel), the spread of which provide no beneficial function. Crossed lines indicate sites of chromosome breakage and end joining with a similar position on a homologous chromosome. Recombinant products of interest are indicated by the thick orange dotted lines.

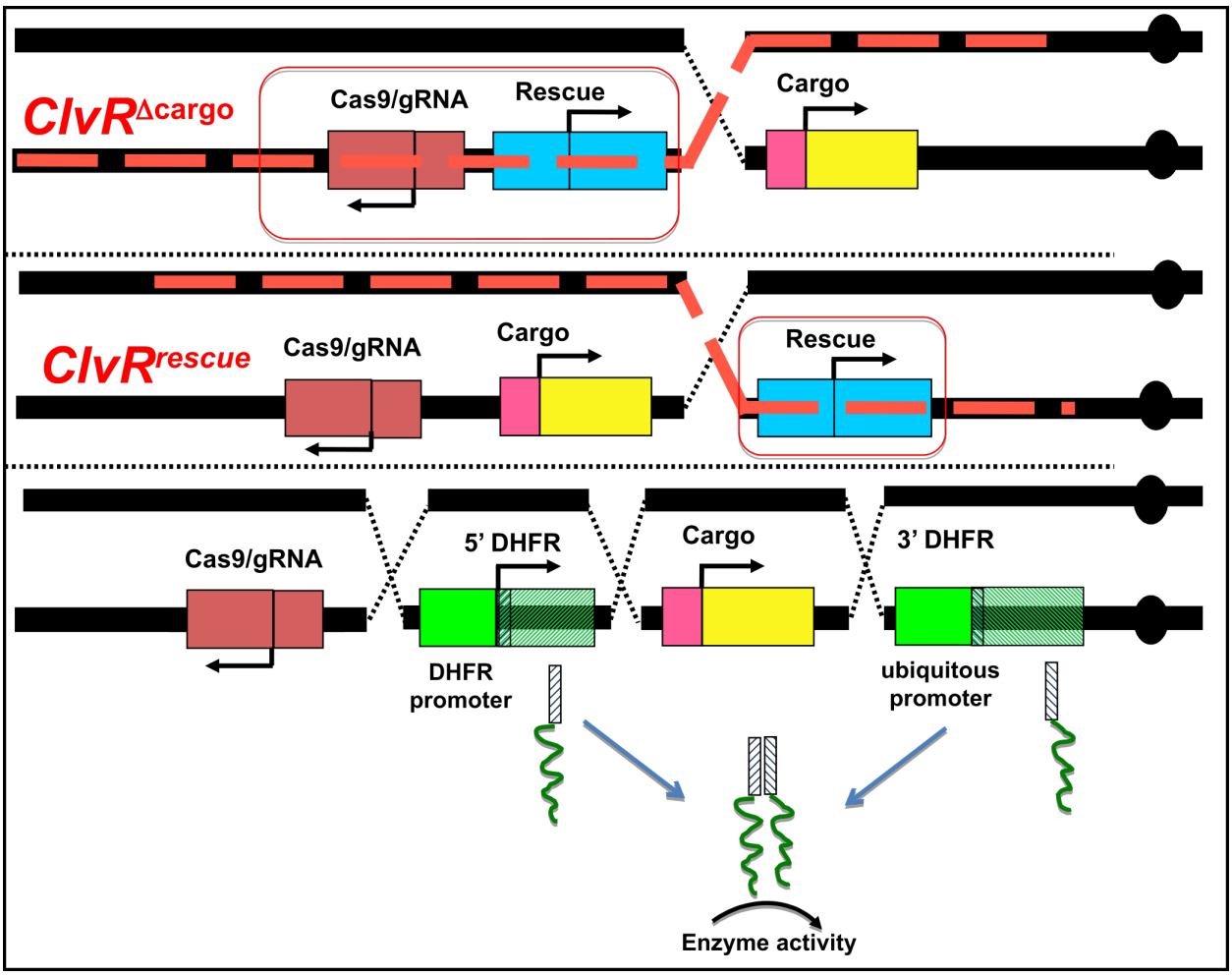

Fig. S9. Separation of a functional Rescue from the Cargo can be prevented by locating the Cargo between two transgenes whose co-expression is required to produce a functional Rescue essential enzyme, such as dihydrofolate reducatse. The 5' half of DHFR is driven by its own promoter. The 3' half is driven by a strong ubiquitous promoter. The two domains are brought together to form an active enzyme through heterodimerization, mediated by specific domains at the N-terminus of each protein (12) (boxes with diagonal lines).

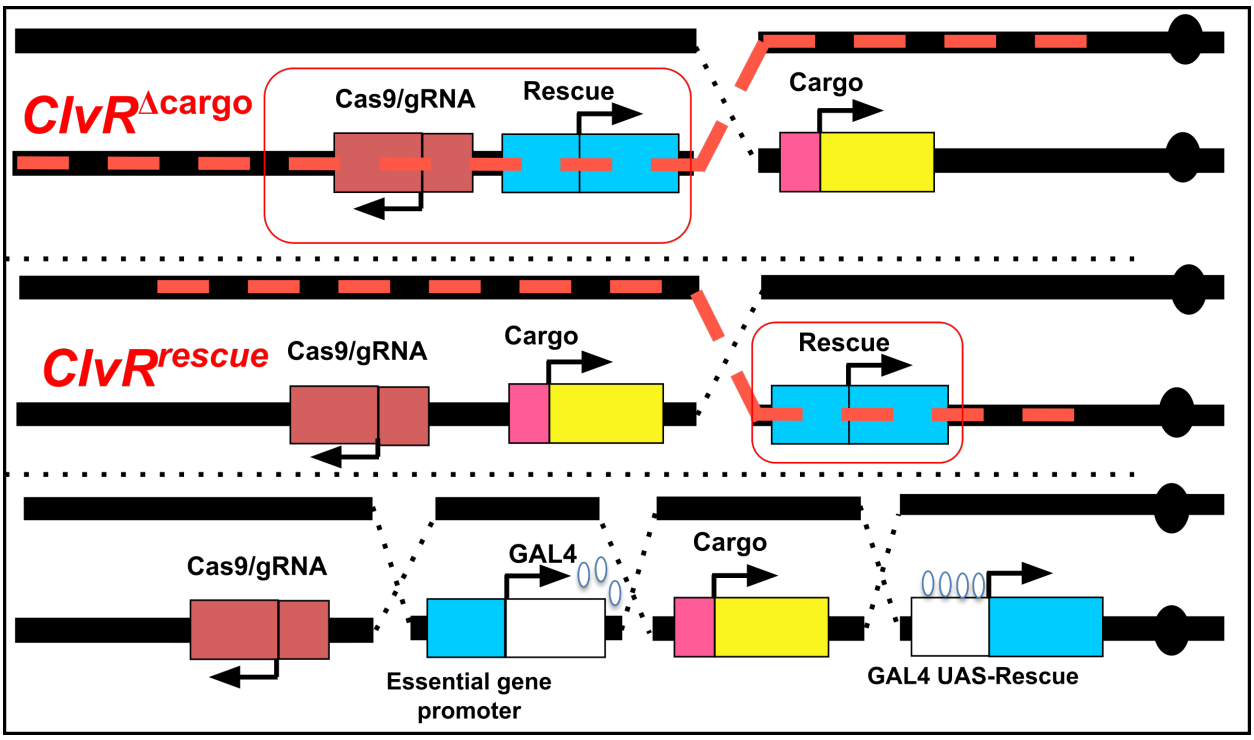

Fig. S10. Separation of a functional Rescue from the Cargo can be prevented by locating the Cargo between two transgenes whose co-expression is required to produce a functional Rescue protein. Here this is achieved using a two-component transcription-based system. The essential gene promoter drives the expression of a heterologous transcriptional activator such as GAL4. The Rescue transgene contains GAL4 UAS binding sites sufficient to drive GAL4-dependent expression, upstream of an otherwise promoterless, recoded Rescue transgene.

Table S1*A*: Summary of Table S1*B*. Shown are the average genotype frequencies (*ClvR*, *td-tomato* and *w*-) and eclosion rates in % with standard deviations from 10 replicates.

	Cross	td-tomato+	w-	Eclosion rate
A	♀w^{1118} XX ♂w^{1118}	0	100	95.9 ±2.0
B	♀w^{1118} XX ♂$ClvR^{tko}$/+	49.9 ±3.0	50.1 ±3.0	95.2 ±2
C	♀$ClvR^{tko}$/+ XX ♂w^{1118}	100	0	46.4 ±1.3
D	♀$ClvR^{tko}$ XX ♂$ClvR^{tko}$	100	0	95.1±1.7

Table S1*B*: Survival assay. Flies of the indicated cross were allowed to lay eggs in a vial for 18 hours. Afterwards, eggs were counted and allowed to develop to adulthood. Eclosed adults were scored for genotype, with *ClvR*-bearing flies identified by the presence of *td-tomato* (tom+).

♂$ClvR^{tko}$/+ XX ♀w^{1118}					
	eggs	tom+	tom-	eclosion rate	ratio
	78	39	38	0.987	0.506
	62	28	29	0.919	0.491
	38	17	20	0.974	0.459
	55	26	26	0.945	0.5
	83	41	38	0.952	0.519
	65	30	32	0.954	0.484
	22	11	10	0.955	0.524
	47	22	23	0.957	0.489
	24	10	13	0.958	0.435
	69	34	30	0.928	0.531
sum	543	258	259		
eclosion rate(SD):	0.952			SD=	0.02
ratio:		0.499	0.501	SD=	0.03
♀w^{1118} XX ♂w^{1118}					
	eggs	tom+	tom-	eclosion rate	

	82	0	79	0.963	
	69	0	67	0.971	
	38	0	35	0.921	
	16	0	16	1	
	68	0	65	0.956	
	61	0	58	0.951	
	53	0	51	0.962	
	54	0	51	0.944	
	93	0	90	0.968	
	78	0	75	0.962	
sum	612	0	587		
hatch rate(SD):	0.959			SD=	0.02
ratio:		0	1		
♀ClvR^{tko}/+ XX ♂w^{1118}					
	eggs	tom+	tom-		
	38	17	0	0.447	
	126	59	0	0.468	
	46	22	0	0.478	
	70	33	0	0.471	
	52	25	0	0.481	
	50	23	0	0.46	
	53	24	0	0.453	
	49	23	0	0.469	
	61	27	0	0.443	
	107	49	0	0.458	
sum	545	253	0		
hatch rate(SD):	0.464			SD=	0.013
ratio:		1	0		
♀ClvR^{tko}/ClvR^{tko} XX ♂ClvR^{tko}/ClvR^{tko}					

	eggs	tom+	tom-		
	56	53	0	0.946	
	64	62	0	0.969	
	50	47	0	0.94	
	73	69	0	0.945	
	42	39	0	0.929	
	45	43	0	0.956	
	58	56	0	0.966	
	51	47	0	0.922	
	59	56	0	0.949	
	87	82		0.943	
sum	388	369	0		
hatch rate(SD):	0.951			SD=	0.017
ratio:		1	0		

Table S2: Crosses to determine rate of *D. melanogster tko* gene inactivation due to female germline cleavage and maternal carry over-dependent cleavage. Shown are the offspring genotype frequencies for a cross between w^{1118}; *ClvR*^{tko}/+ females and w^{1118} males. Flies were scored as *ClvR*-bearing based on the presence of the *td-tomato* marker. Of 3736 flies scored, one did not have the *td-tomato* marker, resulting in a cleavage rate of 0.9997. All crosses were single fly crosses if not otherwise noted (pool = a few flies; bottle = many flies (~50)).

cross	tomato+	tomato-	ratio	note
1	61	**0**	1	
2	50	**0**	1	
3	63	**0**	1	
4	62	**0**	1	
5	49	**0**	1	
6	48	**0**	1	
7	50	**0**	1	
8	127	**0**	1	pool
9	55	**0**	1	
10	33	**0**	1	
11	52	**0**	1	
12	203	**0**	1	pool
13	99	**0**	1	pool
14	45	**0**	1	
15	42	**0**	1	
16	72	**0**	1	
17	53	**0**	1	
18	23	**0**	1	
19	49	**0**	1	
20	49	**0**	1	
21	38	**0**	1	
22	32	**0**	1	
23	39	**0**	1	
24	12	**0**	1	
25	46	**0**	1	
26	7	**0**	1	
bottle1	868	**0**	1	bottle
bottle2	736	**1**	0.9986	bottle

bottle3	672	**0**	1	bottle
SUM	3735	**1**	0.99973	

Table S3: Crosses to determine male germline cleavage rate. Shown are the offspring genotype frequencies for crosses between *ClvR*^{tko}/+ males and *tko*^{3}/FM7a,*B*^{1} females. Flies having the *ClvR* element were scored by the presence of the *td-tomato* marker. The *tko*^{3} mutant allele is on a *w*+ X chromosome; The X_P paternal X chromosome is *w*- (*w*^{1118}); The *ClvR*^{tko} element on the third chromosome is marked by the presence of *td-tomato*; The FM7a,*B*^{1} Balancer X chromosome is identifiable by virtue of the *Bar* dominant eye marker (*B*^{1}); + refers to a wildtype third chromosome; Y refers to the Y chromosome. The male germline cleavage rate was calculated as the ratio of 8 (*tko*^{3}/X_p;;+) / 907 (*tko*^{3}/X_p;;*ClvR*^{tko}) = 0.9911. The 5 escapers from bottle 2 share a common polymorphism (*SI Appendix,* Fig. S5), and thus may represent multiple isolates of the same adult male germline cleavage and repair event.

cross	♀*tko*^{3}/X_P;; *ClvR*^{tko}	♀FM7a,*B*^{1}/ X_P;;*ClvR*^{tko}	♀*tko*^{3}/X_P;;+	♀Fm7a,*B*^{1}/ X_P;;+	♂*tko*^{3}/Y_P;; *ClvR*^{tko}	♂FM7a,*B*^{1}/ Y_P;;*ClvR*^{tko}	♂*tko*^{3}/Y_P;;+	♂Fm7a,*B*^{1}/ Y_P;;+
1	15	6	**0**	12	2	5	0	8
2	5	6	**0**	7	4	3	0	1
3	8	7	**0**	8	6	0	0	0
4	7	5	**0**	2	9	1	0	1
5	16	13	**0**	15	14	4	0	1
6	10	11	**0**	14	16	5	0	2
7	16	14	**0**	13	23	5	0	3
8	15	13	**0**	16	15	6	0	3
9	24	23	**0**	8	16	1	0	3
10	19	9	**0**	9	9	4	0	3
11	12	13	**0**	10	22	2	0	4
12	11	15	**0**	8	19	5	0	4
13	14	8	**0**	12	20	4	0	1
14	7	7	**0**	2	5	4	0	4
15	18	7	**0**	15	23	2	0	4
16	14	23	**0**	15	19	2	0	1
17	32	21	**0**	18	12	2	0	1
18	13	7	**0**	16	19	4	0	2
19	8	4	**0**	4	2	3	0	2
20	11	18	**0**	13	23	1	0	2
21	8	6	**0**	6	5	3	0	6
22	27	19	**0**	13	16	1	0	2
23	17	6	**0**	15	11	1	0	4
24	14	17	**0**	19	17	6	0	1

25	11	8	**0**	3	8	3	0	4
26	11	10	**0**	8	11	1	0	0
27	14	14	**0**	13	15	1	0	4
28	18	18	**0**	14	18	1	0	3
29	19	18	**0**	10	27	0	0	2
30	16	17	**0**	11	23	6	0	3
31	16	17	**0**	13	12	0	0	1
32	18	13	**0**	16	17	0	0	2
33	15	13	**0**	13	22	3	0	2
34	15	17	**0**	11	15	4	0	4
35	11	11	**0**	11	13	1	0	3
bottle1	219	165	**3**	200	216	21	0	11
bottle2	183	169	**5**	154	156	33	0	19
sum	907	768	**8**	747	880	148	0	121
total flies counted	3579							

Table S4: Analysis of escapers. Shown are the alterations in the gRNA target sites of escaper flies. Flies are numbered based on the cross they were coming from (escF1 from bottle 2 of female *ClvR*^{tko}/+ mothers; escM1A-escM8B from male *ClvR*^{tko}/+ fathers. See *SI Appendix,* Fig. S5 for mating scheme to isolate the escaper X-chromosome). '+' indicates an unaltered target site, numbers indicate the size of the deletion. The last two columns show the number of progeny from an outcross of the escaper males to *ClvR*^{tko}/+ females, and the fraction carrying the *ClvR* marker *td-tomato* (tom+) or lacking it (tom-). The two males escM3A and esc M3B gave a mixed sequencing signal, which could not be aligned unambiguously (ND, not determined). All escapers were still sensitive to *ClvR* drive, as shown by the results of the outcross to *ClvR*^{tko}/+ females, which resulted in a progeny population in which all individuals carried *ClvR*^{tko} (tom+), indicating that the *D. melanogaster tko* locus had been disrupted in all non-*ClvR*^{tko}-bearing individuals.

escaper	g1	g2	g3	g4	tom+	tom-
escF1	+	+	+	+	62	0
escM1A	3	+	+	+	31	0
escM1B	3	+	+	+	65	0
escM2A	3	+	+	+	66	0
escM2B	3	+	+	+	62	0
escM3A	ND	ND	ND	ND	45	0
escM3B	ND	ND	ND	ND	37	0
escM4A	3	+	+	+	34	0
escM4B	3	+	+	+	48	0
escM5A	3	+	+	+	79	0
escM5B	3	+	+	+	87	0
escM6A	3	+	+	+	50	0
escM6B	3	+	+	+	68	0
escM7A	3	+	+	+	62	0
escM7B	3	+	+	+	57	0
escM8A	3	+	+	+	73	0
escM8B	3	+	+	+	85	0

Table S5: *ClvR*^{tko} genotype frequencies during introgression into 5 different GDL genetic backgrounds. *ClvR*^{tko}/+ females were mated each generation with GDL males. Labels of GDL lines from (13, 14) are given in the column headers. Progeny were counted and their genotypes were scored with respect to the presence of the *ClvR td-tomato* marker. After each generation 30 virgins were collected and backcrossed to wildtype males of the corresponding GDL stock. Shown are the numbers of scored flies with the *ClvR* marker *td-tomato*. Flies without the marker are indicated in brackets. Maternal germline and carryover-dependent mutation of the *D. melanogaster tko* locus was efficient since progeny lacking *ClvR*^{tko} were not observed, 0/7882

Generation	B12	I02	N23	T01	ZW140
1	103(0)	73(0)	84(0)	90(0)	85(0)
2	184(0)	206(0)	217(0)	212(0)	194(0)
3	272(0)	221(0)	259(0)	211(0)	236(0)
4	304(0)	447(0)	316(0)	350(0)	253(0)
5	342(0)	228(0)	297(0)	206(0)	249(0)
6	540(0)	406(0)	453(0)	429(0)	415(0)
SUM	1745	1581	1626	1498	1432
Total flies scored	7882				

Table S6: Sequence polymorphisms in the tko gRNA target sites used in this study, in *Drosophila* strains from the 1000 fly genomes project. Shown are pre-existing polymorphisms (SNP) in these strains, with the location and type of the SNP in the corresponding gRNA target site. The last column gives the number of gRNA target sites used in this work that are not altered in each strain. The gRNA2 target site was polymorphic in about half of the 1000 fly genomes, and was also present at some frequency in the lab strain used in our experiments, w^{1118}. With this data available it should be possible to select more conserved target sites, e.g. acagccttcagcttaacgccGGG (conserved in all), and gtgctggtgcgcctctccacCGG (SNP in one strain), though it remains to be determined if gRNAs corresponding to these sequences are highly active (see our results in the main text with gRNA3).

strain	gRNA1	gRNA2	gRNA3	gRNA4	Functional gRNAs
US103	+	G-->A (bp10)	+	C-->A (bp13)	2
GU6	+	+	+	T-->C (bp10)	3
KR39	+	A-->G (bp7)	+	C-->A (bp13)	2
RAL149	+	+	+	C-->A (bp13)	3
RAL808	+	G-->A (bp10)	+	C-->A (bp13)	2
SP188	+	A-->G (bp7)	+	C-->A (bp13)	2
ZI420	+	G-->A (bp10)	+	C-->A (bp13)	2
ZI508	+	+	+	C-->A (bp13)	3
CO10N	C-->T (bp8)	+	+	+	3
ZI251N	C-->T (bp8)	G-->A (bp10)	+	+	2

Table S7: Molecular analysis of *ClvR* induced mutations at the target locus. (A) Shown are the type of cleavage events observed at the different gRNA target sites (g1-g4) in male progeny of *ClvR*^{tko}/+ mothers (from Fig. 3*B*). Unaltered target sites are indicated as '+', polymorphisms predicted to render the target site resistant to cleavage are indicated by 'SNP', and gRNA target site mutations likely to result in LOF as 'indel'. **(B)** As with (A), but with males coming from a homozygous *ClvR*^{tko} stock inbred for 3 generations. Note how mutations accumulate over multiple generations.

A	fly	g1	g2	g3	g4		B	fly	g1	g2	g3	g4
	1.1	indel	+	+	indel			1	indel	SNP	+	indel
	1.2	indel	SNP	+	indel			2	indel	indel	+	indel
	2.1	indel	SNP	+	+			3	indel	SNP	+	indel
	2.2	indel	SNP	+	+			4	indel	indel	+	indel
	3.1	indel	+	+	+			5	indel	SNP	+	indel
	3.2	indel	SNP	+	+			6	indel	indel	+	indel
	4.1	indel	+	+	indel			7	indel	+	+	indel
	4.2	indel	indel	+	indel			8	indel	SNP	+	indel
	5.1	indel	indel	+	indel			9	indel	SNP	+	indel
	5.2	indel	indel	+	+			10	indel	SNP	indel	indel
	6.1	indel	+	+	+			11	indel	indel	+	indel
	6.2	indel	+	+	+			12	indel	indel	+	indel
	7.1	indel	+	+	+							
	7.2	indel	+	+	+							
	8.1	indel	+	+	indel							
	8.2	indel	+	+	indel							
	9.1	indel	+	+	indel							
	9.2	indel	+	+	indel							

Table S8: Allele frequency of *ClvR*^{tko} in drive experiment 1. Shown are male outcrosses taken from the drive experiment to *w*^{1118} virgins. We took approximately 100 males and outcrossed them individually to *w*^{1118} virgins. A male was considered to be homozygous if all progeny had the *ClvR td-tomato* marker and heterozygous if not. Replicates A-E are from the individual drive 1 replicates, M is the ratio inferred from the model. Note that not all of the 100 set up outcrosses produced offspring (sum of scored crosses ranged from 90-106). Allele frequency was calculated as (1*heterozygotes + 2*homozygotes)/(2*sum). Data shown here was used to plot Fig. 5*D*.

replicate	Generation	allele_frequency	sum	homozygotes	heterozygotes
A	7	73.93617021	94	45	49
B	7	76.34408602	93	49	44
C	7	83.85416667	96	65	31
D	7	80.97826087	92	57	35
E	7	80.43478261	92	56	36
A	10	78.94736842	95	55	40
B	10	84.23913043	92	63	29
C	10	88.94736842	95	74	21
D	10	84.89583333	96	67	29
E	10	87.5	96	72	24
A	16	87.22222222	90	67	23
B	16	90.20618557	97	78	19
C	16	90.65934066	91	74	17
D	16	86.79245283	106	78	28
E	16	91.5	100	83	17
M	0	25			
M	1	25			
M	2	36.35326813			
M	3	46.17784137			
M	4	56.56169314			
M	5	66.18114755			
M	6	73.74922128			
M	7	79.03286509			
M	8	82.64573218			
M	9	85.21007274			
M	10	87.11550319			

M	11	88.58611344			
M	12	89.75541776			
M	13	90.70740524			
M	14	91.49750742			
M	15	92.16378133			
M	16	92.73322195			
M	17	93.22550916			