Supplementary Methods

Sophia H. Webster, Michael R. Vella, and Maxwell J. Scott on "Development and testing of a novel Killer-Rescue self-limiting gene drive system in *Drosophila melanogaster*", Proceedings Royal Society B, doi: 10.1098/rspb.2019.2994

Assembly of K and R Constructs

To generate the rescue (R) plasmid pUASGal80-PUbEGFP-attB, the PUb-EGFP marker was excised from pBXLII_PUbEGFP_fa (1) and inserted into pUASTattB (GenBank EF362409.1), essentially replacing the *white*+ marker with PUb-EGFP. To assemble pBC-UAS-hsp70-Gal80-p10, firstly the p10 polyA fragment was excised from pJFRC81 (Addgene #36342) by digestion with EcoRI and XbaI and inserted into pBSII KS(-) that had been cut with the same enzymes. Gal80 was excised from RIV gal80 (Drosophila Genomics Resource Center [DGRC] #1335) by digestion with NotI and XbaI and ligated with pBS-p10 that had been cut with the same enzymes. Lastly, the UAS-hsp70-syn21 fragment was amplified from pJFRC81 with primers 5'-TTTTTACGCGTCAAGCTTGCATGCCTGCAG-3' and 5'-GAAGATCTCTTGTTGTAGTCCATTTTGATTTTTTTTTTTAAGTTGGTACC-3'and Gal80-p10 was amplified from the pBS-Gal80-p10 plasmid with the primers 5'-GGTACCAACTTAAAAAAAAAAATCAAAATGGACTACAACAAGAGATCTTC-3' and 5'-TTCACTCGAGTAGATATCGAATTCGT-3'. The PCR products were mixed and a second amplification performed with the 5'-TTTTTACGCGTCAAGCTTGCATGCCTGCAG-3'and 5'-TTCACTCGAGTAGATATCGAATTCGT-3' primers. The final amplification product was ligated with pBC SK+ that had been digested with EcoRV. The resulting pBC-UAS-hsp70-Gal80-p10 plasmid sequence was confirmed by Sanger DNA sequencing. The UAS-hsp70-Gal80-p10 cassette was excised with the enzymes *MluI* and *EcoRV* and ligated into the pUASTattB-PUb-EGFP vector that had been cut with AscI and PspOMI (blunt) to make pUASGal80-PUbEGFP-attB. The UAS-hsp70 fragment present in the parental pUASTattB vector had been excised and thus was not present in the final construct.

To generate the pUAShsp70Gal4-PUbDsRed-attB killer construct with the hsp70 core promoter, first the PUb-DsRed marker was excised from pBXLII_attP220_PUbDsRed.T3_fa (1) and inserted into pUASTattB (GenBank EF362409.1) essentially replacing the *white*+ marker with PUb-DsRed. The Gal4 cds was excised from pCaSpeR-Gal4 (DGRC #1224) and inserted upstream of the p10 pA in pBS-p10 (see above). Lastly, the UAS-hsp70-syn21 fragment was amplified from pJFRC81 with primers 5'-TTTTTCTTAAGCAAGCTTGCATGCCTGCAG-3' and 5'-ATAGAAGACAGTAGCTTCATTTTGATTTTTTTTTTTAAGTTGGTACC-3'and Gal4-p10 was amplified from the pBS-Gal4-p10 plasmid with the primers 5'-GGTACCAACTTAAAAAAAAAAATCAAAATGAAGCTACTGTCTTCTAT-3' and 5'-TTTTTACTAGTGTTAACTCGAATCGCTA-3'. The PCR products were mixed and a second amplification performed with the 5'-TTTTTCTTAAGCAAGCTTGCATGCCTGCAG-3'and 5'-

TTTTTACTAGTGTTAACTCGAATCGCTA-3' primers. The final amplification product was digested with AflII and SpeI ligated with pSLfa1180fa (2) that had been digested with same enzymes. The resulting pUAS-hsp70-Gal4-p10 plasmid sequence was confirmed by Sanger DNA sequencing. The UAS-hsp70-Gal4-p10 cassette was excised with the enzymes AscI and PspOMI and ligated into the pUASTattB-PUb-DsRed vector that had been cut with the same enzymes. The UAS-hsp70 fragment present in the parental pUASTattB vector had been excised and thus was. not present in the final construct.

To make the pUASdscpGal4-PUbDsRed-attB plasmid, pUAShsp70Gal4-PUbDsRed-attB was modified using the site directed mutagenesis kit (New England Biolabs) to excise the *hsp70* core promoter and replace with a polylinker that contained unique EcoR1 and PacI sites. The primers used for site directed mutagenesis of pUAShsp70Gal4-PUbDsRed-attB were 5'-TCACCGGCGTCGGAAGAATTCGTACGCCTCGACGTCGCTAGCGCT-3' and 5'-TGCCCGGGCACGGTTAATTAAGCAGCGGGAAGAGAACTCTGAATAGATCTAAAAG GTAGG-3'. The site directed mutagenesis kit (New England Biolabs) was followed exactly as directed. The resulting plasmid was digested with EcoR1 and PacI and ligated with a synthesized DNA fragment that contained the DSCP core promoter and had been cut with the same enzymes. The sequence of the nucleotide fragment containing the DSCP core promoter and cloning sites was 5'-CACCGGCGGAATTCGAGCTCGCCCGGGGATCGAGCGCAGCGGTATAAAAGGGCG CGGGGTGGCTGAGAGCATCAGTTGTGAATGAATGTTCGAGCCGAGCAGACGTGC CGCTGCCTTCGTTAATATCCTTTGAATAAGCCAACTTTGAATCACAAGACGCATACC AAACTTAATTAA-3'

To make pUAShsp70Gal4-UAShid-PUbDsRed-attB, first the *hid* coding sequence was amplified from a full length cDNA clone (#11756, DGRC) with the PCR primers: For 5'-AAAAAGCGGCCGCGATCCCCGACACCAGACCAACT-3' and Rev 5'-CCCCCGGATCCGAAGAGAACTCTGAATAGGGAATTGGGA-3'. The amplification product was gel purified, digested with NotI and BamHI and ligated with the UASP transformation vector (DGRC, #1189) that had been cut with the same enzymes. UASP-hid was then excised by digestion with PspOMI and EcoRV and ligated with the pUAShsp70Gal4-PUbDsRed-attB vector that had been cut with PspOMI and BstZ17I.

The plasmid DNA sequences for the final constructs have been deposited in Genbank. The accession numbers are:
pUASGal80-PUbEGFP-attB: MN830806
pUAShsp70Gal4-PUbDsRed-attB: MN830807
pUASdscpGal4-PUbDsRed-attB: MN830808
pUAShsp70Gal4-UAShid-PUbDsRed-attB: MN830809

Fly Rearing and Strains

Plasmid DNA was prepared using Invitrogen Purelink HiPure Midiprep kit. The rescue construct pUAS-Gal80-PUbEGFP-attB was prepared in injection buffer at 750 ng/μL final concentration. For the pUAS-Gal4-PUbDsRed-attB killer constructs, injection mixes contained 500 ng/μL of Gal4-attB plasmid and 300 ng/μL of pBC-UAS-Gal80-p10

(lacks an attB site) as a source of transient expression of Gal80. The injection mixes were filtered through a 0.45μM Millipore Ultrafree-MC Centrifugal Filter (Cat #: UFC30HV00) to remove any additional debris before embryo injections.
The pUAS-Gal80-PUbEGFP-attB plasmid was injected into embryos from the nos-phiC31; attP40 strain (**Table S2**). Single G_0 adults that developed from injected embryos were crossed with the parental nos-phiC31; attP40 strain. Four independent lines were identified and kept separate in subsequent crosses. G_1 were interbred and homozygous G_2 offspring identified by fluorescence intensity. All transgenes confirmed to be on chromosome 2 through crosses with a CyO balancer strain. One line was selected as the rescue line for subsequent experiments.

pUAS-Gal4-PUbDsRed-attB and pUAS-Gal4-PUbDsRed-UAS-hid-attB plasmids (with pBC-UAS-Gal80-p10 plasmid that lacked attB) were injected into embryos from the nos-phiC31; attP2 strain (Table S2). Single G_0 adults that developed from injected embryos were then mated to the homozygous UAS-Gal80 rescue line so that all G_1 offspring would have one copy of UAS-Gal80 to repress lethal effects of Gal4 overexpression. G_1 adults were then crossed again with homozygous UAS-Gal80 rescue and G_2 offspring that had one copy of the Gal4 killer and were homozygous for Gal80 rescue selected by fluorescence intensity. The G_2 adults were interbred and, if possible, G_3 offspring that were homozygous for the Gal4 killer selected by fluorescence intensity.

For the gene drive experiments, the attP40 (BDSC Stock # 25709) was used as the wild-type strain because the attP40 background is the same as that of the UAS-Gal80 rescue. The initial cage was comprised of 50 genetically engineered virgin females, 50 genetically engineered males, 25 attP40 virgin females, and 25 attP40 males. Five biological replicates of each Killer-Rescue gene drive experiment were conducted using 8oz round bottom bottles (Cat # 32-129F) from Genesee Scientific filled with 75 mL of fly food. The flies were placed in the bottles for three days to acclimatize to the conditions. The crosses were then transferred to a second bottle, which was used to count genotypes for the next generation of offspring. Three days later the flies in the second bottle were removed and discarded. Fifteen days after the second bottle was set, all the adult flies were genotyped by marker screening under the fluorescent scope and sorted into their respective genotypic groupings based on fluorescence intensity (**Fig. S4**). The experiment was carried out for six (UAS-Gal80, hsp70-Gal4, hsp70-Gal4-hid) or nine (DSCP-Gal4) generations. All fly counts are shown in Table S1.

RNA isolation and qRT-PCR

RNA Preparation: Pre-filled tubes containing zirconium beads were used for sample collection (Benchmark Scientific Triple-Pure High Impact Zirconium Beads 1.5 mm Cat # D1032-15). 500 μL of Trizol ® was pipetted into each pre-filled tube and then the samples were directly collected into the Trizol ®. After the tissues were disrupted using the homogenizer (OPS Diagnostic HT mini homogenizer 230V Cat # BM-D1030) for 3-4 minutes at max speed (4000 rpm), they were processed immediately using a phenol-chloroform extraction protocol and Qiagen RNEasy Mini Kit (Cat# 74104).

Thermo Fisher Scientific dsDNAse (Cat # EN0771) was added to the purified RNA to remove any contaminating DNA. cDNA was synthesized using the Invitrogen (Cat#18080-400 Invitrogen) SuperScript III First-Strand Synthesis SuperMix kit.

To measure relative transcript levels, qRT-PCR was performed with the cDNA template diluted 1:4 with nuclease-free water then pipetted into quadruplicate wells of a 384 well optical plate (Cat#4309849 Applied Biosystems). Thermo Maxima SYBR Green/Rox qPCR Master Mix 2X (Cat#K0221) was added to 10μM primers to create a master mix, which was then dispensed into wells using a multichannel pipette. The primer sequences for Gal4, Gal80, Rpl32, and 18S rRNA are listed in supplementary (Table S3). The qPCR run was performed on a BioRad CFX384 C1000 Touch Thermocycler.

The relative normalized expression, ΔΔCq, was calculated to quantify the expression of Gal4 and Gal80 present in the samples collected. The calculations for the qPCR were performed as follows: (i) the mean Cq for each biological replicate was calculated by averaging the four technical replicates, (ii) the standard deviation and standard error of the mean were calculated, (iii) the ΔCq was calculated for each reference gene (Rpl32 and 18S rRNA) by taking the difference between the mean Cq for each sample biological replicate and the corresponding mean Cq of the reference gene, and (iv) the ΔΔCq for each sample was calculated as $2^{-\Delta\Delta Cq}$.

Mathematical Model

We used a mathematical model to calculate expected population genetics of the cage experiments. The model is discrete-time with non-overlapping generations and assumes random mating and equal sex ratios. The model includes fitness parameters for each genotype, which we assume are assessed prior to mating, The population is normalized every generation, and we track frequencies of each genotype, denoted by subscript *i*. A genotype's fitness, w_i, gives the proportion of individuals of that genotype that enter the mating pool, i.e., survive to adulthood after any viability and embryonic fitness costs are assessed, relative to wild-type fitness of 1. At generation t, the relative proportion of individuals of each genotype entering the next generation, $B_i(t+1)$, is a function of fitness, current adult frequencies A_i, and the probability, $P(i|m,n)$, that a mating between female of genotype m and male of genotype n produces offspring of genotype i:

$$B_i(t+1) = w_i \sum_m A_m(t) \sum_n P(i|m,n)\, A_n(t)\,.$$

Then, the genotype frequencies of adults in the next generation's mating pool is given by:

$$A_i(t+1) = \frac{B_i(t+1)}{\sum_j B_j(t+1)}.$$

Mathematical Model with Stochasticity

A stochastic model was used to generate synthetic experimental data. Like the deterministic model, the stochastic model is discrete-time with non-overlapping generations and assumes random mating. Unlike the deterministic model, the stochastic model separately tracks females and males and incorporates mating stochasticity. We let $F_i(t)$ and $M_i(t)$ represent the adult females and males, respectively, of genotype i at time t. By assuming random mating, the number of females of genotype *m* that mate with males of each other genotype n, $F'_{m,n}$ for $j = 1,2 \ldots 9$, is distributed multinomially with probabilities equal to the proportion of males of each genotype in the previous generation:

$$F'_{m,n=1:9}(t) \sim \text{Multinomial}\left(n = F_m(t-1), p = \frac{M_{1:9}(t-1)}{\sum_j M_j(t-1)} \right).$$

We assume the number of offspring each female produces can be modeled by a Poisson distribution, for which we let $\lambda = 6$ based on empirical approximation. The total number of offspring produced by females of genotype *m* mating with males of genotype *n* is

$$E_{m,n}(t) \sim \text{Poisson}\left(F'_{m,n} * \lambda\right),$$

and the offspring genotypes will be based on the probabilities $P(i \mid m, n)$, that a mating between female of genotype m and male of genotype n produces offspring of genotype i.

$$E'_{m,n,i=1:9} \sim \text{Multinomial}\left(n = E_{m,n}(t), p = P(i = 1{:}\,9 | m, n)\right).$$

The proportion of individuals of genotype i that survive to adulthood and contribute toward the next generation is given by that genotype's fitness, w_i, which is defined relative to wildtype fitness of 1. We assume the number of survivors, $B_i(t)$, is binomially distributed:

$$B_i(t) \sim \text{Binomial}\left(n = \sum_{m,n} E'_{m,n,i}\,, p = w_i \right).$$

Finally, we assume sex is determined at emergence to adulthood, with equal probability of males and females:

$$F_i(t) \sim \text{Binomial}(n = B_i(t), p = 0.5),$$

$$M_i(t) = B_i(t) - F_i(t).$$

With the stochastic model, confidence intervals for the parameter estimates were obtained through parametric bootstrapping. First, the stochastic model (with the

estimated parameter values) was used to generate synthetic data. Then, parameter estimation was performed on the synthetic data. This process was repeated $n = 2000$ times for each experiment, and the confidence intervals give the 0.025 and 0.975 quantiles for each parameter.

Supplementary Methods References

1. Schetelig MF, Handler AM. Strategy for enhanced transgenic strain development for embryonic conditional lethality in *Anastrepha suspensa*. Proc Natl Acad Sci U S A. 2012;109(24):9348-53.
2. Horn C, Wimmer EA. A versatile vector set for animal transgenesis. Dev Genes Evol. 2000;210(12):630-7.