

ADDITIONAL METHODS

Western Blotting:

Male and female transgenic Vas1GFP mosquitoes were dissected in PBS to separate the ovaries or testes from the remaining gonad-less carcasses and placed in PBS with EDTA-free Protease Inhibitor Cocktail (Roche Applied Science). For female ovaries a time series containing lysates of non-bloodfed and 24 or 48 hours post-bloodfeeding was also collected. Samples were separated on 10–12% SDS-PAGE gels followed by electro-transfer to nitrocellulose membranes and incubated with antibodies against eGFP (1:3000 mouse monoclonal anti-GFP Living Colours JL-8) and as a loading control against alpha-tubulin (1:1000 mouse monoclonal anti- α -Tubulin Sigma Aldrich DMIA). Primary antibodies were detected using anti-mouse IgG HRP-conjugated secondary antibody and visualized by Western Lightning Chemiluminescence Reagent Plus kit (Perkin Elmer). Blots were stripped in 62.5 mM Tris-HCl, pH 6.7, 2% SDS, and 100 mM Mercaptoethanol at 60°C for 1 h.

Computational Modelling:

Targeting a somatic female fertility gene (invasive approach)

At the target locus there are 3 different alleles: wildtype (*wt*), HEG-containing knockout (*HEG*), and misrepaired mutant knockout (*M*). Therefore there are 6 genotypes: *HEG/HEG* (genotype1), *HEG/wt* (genotype2), *wt/wt* (genotype3), *M/HEG* (genotype4), *M/wt* (genotype5), *M/M* (genotype6).

Let the frequencies of these 6 genotypes in adult males be $m_1, m_2, m_3, m_4, m_5, m_6$, and in fertile adult females be $f_1, f_2, f_3, f_4, f_5, f_6$ (we exclude the sterile females because they do not contribute to the next generation). Then the frequencies of the 6 genotypes in the embryos of the next generation (after maternal deposition has had its effect) will be:

$$e1' = \frac{1}{4}f2 \left(2(1 + c(-1 + c)Dh_e + h_g) \right) m1 + m2 + m4 + c \left(h_g \left((2 + ch_g)m2 + m4 \right) + Dh_e(-2(-1 + c)(1 + ch_g)m2 + 2(m3 + ch_gm3) + m4 - cm4 + m5 + ch_gm5) \right)$$

$$e2' = \frac{1}{4}((2f3 + f5)(2m1 + m2 + ch_gm2 + m4) + f2(-1 + cD(1 + h_e))(2(-1 + c)m1 + 2(-1 + c)(1 + ch_g)m2 - 2m3 - m4 - m5 + c(m4 - h_g(2m3 + m5))))$$

$$e3' = \frac{1}{4}(-2f3 - f5 + (-1 + c)f2(-1 + cD(1 + h_e))^2)((-1 + c)m2 - 2m3 - m5)$$

$$e4' = \frac{1}{4}(-2c^3Df2h_gm2 + f5(2m1 + m2 + m4) + c^2f2(-2(-1 + h_g)h_gm2 + D(-2m1 + 2(-1 + h_g)m2 + 2h_gm3 - m4 + h_gm5)) + f2(m4 + m5 + 2m6) + c(f5h_gm2 + f2(2(1 + D - h_g)m1 + 2m2 + m4 + D(2m2 + 2m3 + m4 + m5) + h_g(-2m2 + m5 + 2m6))))$$

$$e5' = \frac{1}{4} \left(-2c^4D^2f2(1 + h_e)^2m2 + 2c^3Df2(1 + h_e) \left((2 + 2D(1 + h_e) - h_g)m2 + D(1 + h_e)(2m3 + m5) \right) + c^2f2 \left(-2(1 + D(1 + h_e)(3 + D + Dh_e - h_g) - h_g)m2 - D(1 + h_e)(2(3 + 2D(1 + h_e) - h_g)m3 - m4 + (2 + 2D(1 + h_e) - h_g)m5 - 2m6)) \right) + c \left(-(2f3(-1 + h_g) + f5h_g)m2 + f2(2(1 + D + Dh_e - h_g)m2 + 2m3 - m4 - h_g(2m3 + m5) + D(1 + h_e)(4m3 - m4 + m5 - 2m6) - 2m6) \right) + (f2 + 2f3)(m4 + m5 + 2m6) + f5(m2 + 2m3 + m4 + 2(m5 + m6)) \right)$$

$$e6' = \frac{1}{4} \left(c^4D^2f2(1 + h_e)^2m2 - c^3Df2(1 + h_e)(2(1 + D + Dh_e - h_g)m2 + D(1 + h_e)(2m3 + m5)) + c^2f2((1 + D + Dh_e - h_g)^2m2 + D(1 + h_e)(2(1 + D + Dh_e - h_g)m3 - m4 + (D + Dh_e - h_g)m5 - 2m6)) + f5(m4 + m5 + 2m6) + c(-f5(-1 + h_g)m2 + f2(1 + D + Dh_e - h_g)(m4 + m5 + 2m6)) \right)$$

where c is the rate of cleavage; cleaved alleles in the germline are repaired by homologous repair (HR) with probability h_g or by non-homologous repair (NHR) with probability $1-h_g$; cleaved alleles in the embryo are repaired by HR with probability h_e or by NHR with probability $1-h_e$; and D is the probability of maternal deposition (0 or 1 in the simulations presented in the text).

All males are fertile, but three of the female genotypes are sterile (HEG/HEG , HEG/M , and M/M).

Genotype frequencies among fertile females are therefore:

$$f1' = 0$$

$$f2' = \frac{e2'}{e2' + e3' + e5'}$$

$$f3' = \frac{e3'}{e2' + e3' + e5'}$$

$$f4' = 0$$

$$f5' = \frac{e5'}{e2' + e3' + e5'}$$

$$f6' = 0$$

The frequencies of the *HEG*, *M* and *wt* alleles (at the embryo stage) are then:

$$e_1 + \frac{e_2}{2} + \frac{e_4}{2}$$

$$\frac{e_4}{2} + \frac{e_5}{2} + e_6$$

$$\frac{e_2}{2} + e_3 + \frac{e_5}{2}$$

respectively, and the frequency of sterile females is:

$$S = e_1 + e_4 + e_6$$

To determine the criteria for invasion we derived an expression for $p[3]/p[2]$, where $p[i]$ is the frequency of the HEG in the i^{th} generation, and took the limit where the introduction frequency was vanishingly small. When there is homing in the embryos ($h_e = h_g$), the line of neutrality for the HEG is given by

$$h_g = \frac{5 - 3c - \sqrt{25 - 22c + 5c^2}}{2(-2c + c^2)}$$

whereas when there is no homing in the embryo ($h_e = 0$), the line of neutrality is given by

$$h_g = -\frac{1}{2(-2 + c)}$$

The population mean fitness presented in the text was calculated as the proportion of females that are fertile, which is:

$$e_2 + e_3 + e_5$$

Targeting a germline female fertility gene (non-invasive approach)

If the target locus is a germline fertility gene then *HEG/wt* females can be sterilized by the action of the HEG in the germline. In this case the frequencies of the 6 genotypes among embryos will be:

$$e_1' = \frac{1}{4}f_2(2(m_1 + cDh_e m_1) + m_2 + m_4 + c(h_g m_2 + Dh_e((2 + c(-1 + h_g))m_2 + 2m_3 + m_4 + m_5)))$$

$$e_2' = \frac{1}{4}((2f_3 + f_5)(2m_1 + m_2 + ch_g m_2 + m_4) - f_2(-1 + cD(1 + h_e))(2m_1 + (2 + c(-1 + h_g))m_2 + 2m_3 + m_4 + m_5))$$

$$e_3' = -\frac{1}{4}(2f_3 + f_5 + f_2(-1 + cD(1 + h_e))^2)((-1 + c)m_2 - 2m_3 - m_5)$$

$$e_4' = \frac{1}{4}(c^2 Df_2(-1 + h_g)m_2 + f_5(2m_1 + m_2 + m_4) + c((f_2 - f_2 h_g + f_5 h_g)m_2 + Df_2(2m_1 + 2m_2 + 2m_3 + m_4 + m_5)) + f_2(m_4 + m_5 + 2m_6))$$

$$e_5' = \frac{1}{4}(2c^3 D^2 f_2(1 + h_e)^2 m_2 - c^2 Df_2(1 + h_e)((3 + 2D(1 + h_e) - h_g)m_2 + 2D(1 + h_e)(2m_3 + m_5)) + c(-(2f_3(-1 + h_g) + f_5 h_g)m_2 + f_2((1 + 2D(1 + h_e) - h_g)m_2 + D(1 + h_e)(4m_3 - m_4 + m_5 - 2m_6))) + (f_2 + 2f_3)(m_4 + m_5 + 2m_6) + f_5(m_2 + 2m_3 + m_4 + 2(m_5 + m_6))))$$

$$e_6' = \frac{1}{4}(-c^3 D^2 f_2(1 + h_e)^2 m_2 + c^2 Df_2(1 + h_e)((1 + D + Dh_e - h_g)m_2 + D(1 + h_e)(2m_3 + m_5)) + f_5(m_4 + m_5 + 2m_6) + c(-f_5(-1 + h_g)m_2 + Df_2(1 + h_e)(m_4 + m_5 + 2m_6))) ,$$

where all symbols have the same definition as above.

Again, the genotype frequencies of the adults are the same as for the embryos. All males are fertile, but three of the female genotypes are sterile (*HEG/HEG*, *HEG/M*, and *M/M*). Genotype frequencies among the subpopulation of females that are fertile are:

$$\begin{aligned}
 f1' &= 0 \\
 f2' &= \frac{e2' - ce2'}{e2' - ce2' + e3' + e5'} \\
 f3' &= \frac{e3'}{e2' - ce2' + e3' + e5'} \\
 f4' &= 0 \\
 f5' &= \frac{e5}{e2' - ce2' + e3' + e5'} \\
 f6' &= 0
 \end{aligned}$$

The $ce2$ term in the equation for $f2'$ arises because *HEG/wt* females can be sterilised as a result of HEG-induced cleavage and knockout mutation of the target gene in the germline.

The frequencies of the *HEG*, *M* and *wt* alleles (at the embryo stage) are then:

$$\begin{aligned}
 e1 + \frac{e2}{2} + \frac{e4}{2} \\
 \frac{e4}{2} + \frac{e5}{2} + e6 \\
 \frac{e2}{2} + e3 + \frac{e5}{2}
 \end{aligned}$$

respectively, and the frequency of sterile females is:

$$S = e1 + ce2 + e4 + e6$$

To calculate the total number of sterile females produced per released male in the t generations after release, we simulated a single introduction of homozygous *HEG/HEG* males and calculated

$$\frac{\sum_{i=1}^t S[i]}{init}$$

where $S[i]$ is the frequency of sterile females in the i^{th} generation after release and $init$ is the initial frequency of the released males, here taken to be 0.0001.