Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/3/4/e1601964/DC1

Supplementary Materials for

Evolutionary dynamics of CRISPR gene drives

Charleston Noble, Jason Olejarz, Kevin M. Esvelt, George M. Church, Martin A. Nowak

Published 5 April 2017, *Sci. Adv.* **3**, e1601964 (2017)
DOI: 10.1126/sciadv.1601964

This PDF file includes:

Supplementary Material: Evolutionary dynamics of CRISPR gene drives

Charleston Noble, Jason Olejarz,

Kevin M. Esvelt, George M. Church, and Martin A. Nowak

In this Supplementary Material, we mathematically study the evolutionary dynamics of a CRISPR gene drive construct with n guide RNAs. In section S1, we discuss relevant prior work on homing endonuclease gene drives. In section S2, we propose a simple model of population genetics of RNA-guided gene drives with multiple guide RNAs, and we analyze the selection pressure acting on an engineered drive construct. In section S3, we derive a condition for an engineered drive allele to invade a natural population. In section S4, we derive a condition for a population in which the drive has fixed to resist invasion by either wild-type or drive-resistant alleles. In section S5, we derive equations for interior equilibria permitted by our system. In section S6, we present numerical examples of the system's dynamics. In section S7, we extend the model from section S2 to include the effects of "neutral resistance".

section S1. Previous work on homing endonuclease gene drives

Deredec et al. (2008) (Ref. (21) in the main text) mathematically investigates the evolutionary dynamics of homing endonuclease gene drives. The authors begin with a two-allele model precluding resistance, consisting of a wild-type allele and a gene drive allele (pp. 2014–2016 of Deredec et al. (2008)). The model implicitly considers a single guide RNA because it was motivated by earlier single-target homing endonuclease genes. In their notation, p is the frequency of the wild-type allele, and q is the frequency of the drive allele. The authors assume Hardy-Weinberg proportions at all times, and they write a recurrence for q

$$q' = \frac{(1-s)q^2 + (1-sh)pq(1+e)}{1 - sq^2 - 2shpq}$$

Here, s is the fitness cost associated with a drive homozygote, sh is the fitness cost associated with a drive/wild-type heterozygote, and e is the probability that the HEG copies itself onto the homologous chromosome ("homes").

The authors identify that there are three possible fixed points:

$$q^* = 0$$
$$q^* = 1$$
$$q^* = \frac{e - (1+e)hs}{s(1-2h)}$$

The authors obtain the following invasion condition for the drive allele

$$s < \frac{e}{h(1+e)}$$

Intuitively, the fitness cost, sh, of a drive/wild-type heterozygote must be less than a monotonically increasing function of the homing rate, e, for the homing endonuclease gene to spread when rare. Low fitness costs of the drive and high homing rates facilitate the invasion of the drive. More specifically, the authors show that, if the drive/wild-type heterozygote has fitness close to the wild-type (i.e., h close to zero), then the drive invades and fixes (if s is small relative to e), coexists with the wild-type allele (if s is comparable in magnitude to e), or does not invade and is unstable (if s is large relative to e). The authors also show that, if the drive/wild-type heterozygote has fitness close to the drive homozygote (i.e., h close to one), then the drive invades and fixes (if s is small relative to e), is bistable with the wild-type allele (if s is comparable in magnitude to e), or does not invade and is unstable (if s is large relative to e). These are important insights into the evolutionary dynamics of homing endonuclease gene drives.

Deredec et al. then extend their model to consider also a single resistant allele (pp. 2018–2019 of Deredec et al. (2008)). In their notation, p is the frequency of the wild-type allele, q_H is the frequency of the drive allele, and q_M is the frequency of the misrepaired (resistant) allele. The authors assume Hardy-Weinberg proportions at all times, and they write recurrences for q_H and q_M

$$q'_H = \frac{q_H^2(1-s_H) + pq_H(1+e(1-\gamma))(1-h_H s_H) + q_M q_H(1-s_I)}{\overline{W}}$$
$$q'_M = \frac{q_M^2(1-s_M) + pq_M(1-h_M s_M) + pq_H(1-h_H s_H)e\gamma + q_M q_H(1-s_I)}{\overline{W}}$$

Here, $\overline{W}$ is the mean fitness of the population, and γ is the probability of misrepair.

The authors then consider a variety of special cases and make observations about each. A general theme is that low misrepair rates, high fitness of the drive, and low fitness of resistance alleles all act to improve drive spread. These are crucial points for understanding the evolutionary dynamics of homing endonuclease genes.

For a classic homing endonuclease gene drive, the latter two properties—high fitness of the drive and low fitness of resistance alleles—are naturally difficult to reconcile with each other, as we describe in the main text. Since cost-free resistance to a drive construct arises, alternative drive designs are necessary for effective genome editing. The recently developed CRISPR/Cas9 genome editing technology facilitates targeting arbitrary locations in a genome, greatly expanding the creative potential for manipulating wild populations. While CRISPR/Cas9 constructs offer enhanced opportunities for constructing functional gene drives, they also inevitably exhibit more complex dynamics that must be firmly understood.

section S2. Model for the evolutionary dynamics of a CRISPR gene drive with n gRNAs

Consider a wild population of diploid organisms. Our aim is to manipulate the population by modifying a particular locus which may be important, for example, for the organism's survival, reproduction, disease transmission, etc. Using CRISPR/Cas9 genome editing techniques, one can engineer a CRISPR gene drive with n guide RNAs to target this locus. See the main text and corresponding Fig. 1 for specific discussion of our proposed design.

To describe the evolutionary dynamics of such a construct, we consider a drive allele, D, a wild-type allele, 0, and n resistance alleles, i (with $1 \leq i \leq n$). (In the main text, we use the notation "W" for a wild-type allele instead of "0". The notation "0" is more natural for doing calculations.) There are $(n+2)(n+3)/2$ possible genotypes in the population: ij (with $0 \leq i \leq n$ and $0 \leq j \leq n$), iD (with $0 \leq i \leq n$), and DD. The drive mechanism works as follows:

Consider a type $0D$ individual; one allele is wild-type, and the other allele is the drive. There are n guide RNAs and therefore n targets for the drive to cut. At meiosis, the drive can cut any number of targets between 0 and n. If the drive cuts no targets, then the individual remains with genotype $0D$. If the drive cuts k targets (with $1 \leq k \leq n$), then one of several things can happen: One possibility is that homologous recombination copies the drive allele onto the damaged chromosome, so that the individual's genotype becomes DD. This is how the drive construct effects its spread through a population. Another possibility is that non-homologous end joining repairs the damaged chromosome without restoring the lost targets, so that the individual's genotype becomes iD (with $1 \leq i \leq n$). This is how resistance to the drive construct emerges. Yet another possibility is that non-homologous end joining perfectly repairs the damaged chromosome, so that the individual's genotype remains $0D$.

The drive allele can effect its spread as long as there is at least one remaining target. In an individual with genotype iD, either the drive cuts at no targets, with the individual's genotype remaining iD, or the drive cuts at some number, k, of the $n-i$ remaining targets (so that $1 \leq k \leq n-i$). After cutting, the individual can become homozygous in the drive allele (DD), the individual can lose additional targets by acquiring genotype jD (with $i+1 \leq j \leq n$), or the individual can remain with genotype iD.

Using these rules, we can formally express the rates at which each of the $n+2$ types of gametes are produced in terms of the frequencies of individuals in the population. We denote by $F_D(t)$ the rate (at time t) at which drive gametes (D) are produced by individuals in the population. We denote by $F_i(t)$ the rate (at time t) at which wild-type gametes ($i=0$) or gametes with varying levels of resistance ($1 \leq i \leq n$) are produced by individuals in the population. We have

$$\begin{aligned} F_D(t) &= f_{DD} x_{DD}(t) + \sum_{k=0}^{n} p_{kD,D} f_{kD} x_{kD}(t) \\ F_i(t) &= \sum_{k=0}^{i} p_{kD,i} f_{kD} x_{kD}(t) + \sum_{k=0}^{n} \frac{1+\delta_{ki}}{2} f_{ki} x_{ki}(t) \end{aligned} \tag{1}$$

δ_{ki} is the Kronecker delta. We use the following notation: $x_{ki}(t)$ denotes the frequency of

individuals (at time t) with only wild-type or resistance alleles, $x_{kD}(t)$ denotes the frequency of individuals (at time t) with one wild-type or resistance allele and one drive allele, and $x_{DD}(t)$ denotes the frequency of individuals (at time t) that are homozygous in the drive allele. (We define $x_{ki}(t)$ for $k \neq i$ and $x_{kD}(t)$ such that the ordering of the indices does not matter, i.e., $x_{ki}(t) = x_{ik}(t)$ is the frequency of individuals with one copy of the k allele and one copy of the i allele, and $x_{kD}(t) = x_{Dk}(t)$ is the frequency of individuals with one copy of the k allele and one copy of the drive allele.) f_{ki} denotes the fitness of individuals with only wild-type or resistance alleles, f_{kD} denotes the fitness of individuals with one wild-type or resistance allele and one drive allele, and f_{DD} denotes the fitness of individuals that are homozygous in the drive allele. $p_{kD,D}$ denotes the probability that an individual of genotype kD produces a D gamete. $p_{kD,i}$ denotes the probability that an individual of genotype kD produces an i gamete. From conservation of probability, we have the following identity

$$p_{kD,D} + \sum_{i=k}^{n} p_{kD,i} = 1$$

Notice that a type nD individual is fully resistant to being manipulated by the drive construct; such a fully resistant individual shows standard Mendelian segregation in its production of gametes. Thus, we have

$$p_{nD,n} = \frac{1}{2}$$

We understand Equations (1) as follows: Type DD individuals only produce type D gametes, hence the term $f_{DD}x_{DD}(t)$ in the equation for $F_D(t)$. Type kD individuals produce type D gametes with probability $p_{kD,D}$, hence the terms $p_{kD,D}f_{kD}x_{kD}(t)$ in the equation for $F_D(t)$. Type kD individuals produce type i gametes with probability $p_{kD,i}$, hence the terms $p_{kD,i}f_{kD}x_{kD}(t)$ in the equation for $F_i(t)$. Type ki individuals produce type i gametes with probability 1 if $k = i$ or with probability $1/2$ if $k \neq i$, hence the terms $[(1 + \delta_{ki})/2]f_{ki}x_{ki}(t)$ in the equation for $F_i(t)$.

The selection dynamics are modeled by the following system of equations

$$\begin{aligned} \dot{x}_{ij}(t) &= (2 - \delta_{ij})\, F_i(t)F_j(t) - \psi^2(t)x_{ij}(t) \\ \dot{x}_{iD}(t) &= 2F_i(t)F_D(t) - \psi^2(t)x_{iD}(t) \\ \dot{x}_{DD}(t) &= F_D^2(t) - \psi^2(t)x_{DD}(t) \end{aligned} \tag{2}$$

Here, an overdot denotes the time derivative, d/dt. In formulating the population dynamics, we assume random mating; i.e., two random gametes meet to form a new individual. Notice that the products $(2 - \delta_{ij})F_i(t)F_j(t)$, $2F_i(t)F_D(t)$, and $F_D^2(t)$ in Equations (2) represent the pairings of the different types of gametes to make new offspring. The quantity $\psi^2(t)$ represents a density-dependent death rate for the individuals in the population.

At any given time, t, we require that the total number of individuals sums to one

$$x_{DD}(t) + \sum_{i=0}^{n} x_{iD}(t) + \sum_{i=0}^{n}\sum_{j=0}^{i} x_{ij}(t) = 1 \tag{3}$$

To enforce this density constraint, we set

$$\psi(t) = F_D(t) + \sum_{i=0}^{n} F_i(t) \tag{4}$$

Throughout this SM, we choose to work in the framework of continuous time (Equations (2)), since we feel that this approach simplifies the mathematical analysis. In much of the remainder of this SM, we omit explicitly writing the time dependence on dynamical quantities for notational convenience.

section S3. Invasion of the drive construct

Consider a wild-type population in which all individuals have genotype 00. We perturb the wild-type population by introducing a small amount of the drive allele, D. What happens? Does the drive allele catalyze its own spread in the population, or is it eliminated?

For a perturbation to a wild-type population, we write the frequencies of the genotypes as

$$\begin{aligned}
x_{00} &= 1 \quad -\epsilon\delta_{00}^{(1)} - \epsilon^2\delta_{00}^{(2)} \quad -\mathcal{O}(\epsilon^3) \\
x_{0D} &= \quad +\epsilon\delta_{0D}^{(1)} + \epsilon^2\delta_{0D}^{(2)} \quad +\mathcal{O}(\epsilon^3) \\
x_{0i} &= \quad +\epsilon\delta_{0i}^{(1)} + \epsilon^2\delta_{0i}^{(2)} \quad +\mathcal{O}(\epsilon^3) \\
x_{ij} &= \quad\quad + \epsilon^2\delta_{ij}^{(2)} \quad +\mathcal{O}(\epsilon^3) \\
x_{iD} &= \quad\quad + \epsilon^2\delta_{iD}^{(2)} \quad +\mathcal{O}(\epsilon^3) \\
x_{DD} &= \quad\quad + \epsilon^2\delta_{DD}^{(2)} \quad +\mathcal{O}(\epsilon^3)
\end{aligned} \tag{5}$$

In Equations (5), it is implied that $1 \leq i \leq n$ and $1 \leq j \leq n$. The expansions (5) are understood as follows. The frequency of the wild-type allele is approximately one, since we only introduce a small amount of the drive allele. The frequency of the drive allele is of order $\epsilon \ll 1$. The small number of $0D$ individuals in the population also produce resistance alleles, and the frequency of these resistance alleles shortly after the perturbation is also small (i.e., of order $\epsilon \ll 1$). Notice that:

- New type 00 individuals are produced by pairing two wild-type gametes (each at a frequency $\mathcal{O}(1)$), so new type 00 individuals are generated at a rate $\mathcal{O}(1)$.
- New type $0D$ individuals are produced by pairing a wild-type gamete (at a frequency $\mathcal{O}(1)$) and a drive gamete (at a frequency $\mathcal{O}(\epsilon)$), so new type $0D$ individuals are generated at a rate $\mathcal{O}(\epsilon)$.
- New type $0i$ individuals (for $1 \leq i \leq n$) are produced by pairing a wild-type gamete (at a frequency $\mathcal{O}(1)$) and a resistant gamete (at a frequency $\mathcal{O}(\epsilon)$), so new type $0i$ individuals are generated at a rate $\mathcal{O}(\epsilon)$.
- New type ij individuals (for $1 \leq i \leq n$ and $1 \leq j \leq n$) are produced by pairing two resistant gametes (each at a frequency $\mathcal{O}(\epsilon)$), so new type ij individuals are generated at a rate $\mathcal{O}(\epsilon^2)$.
- New type iD individuals (for $1 \leq i \leq n$) are produced by pairing a resistant gamete (at a frequency $\mathcal{O}(\epsilon)$) and a drive gamete (at a frequency $\mathcal{O}(\epsilon)$), so new type iD individuals are generated at a rate $\mathcal{O}(\epsilon^2)$.

- New type DD individuals are produced by pairing two drive gametes (each at a frequency $\mathcal{O}(\epsilon)$), so new type DD individuals are generated at a rate $\mathcal{O}(\epsilon^2)$.

Also, notice that a nonzero amount of the drive allele and the resistance alleles are produced at order ϵ^2 by type ij, iD, and DD individuals, so there also exist terms of order ϵ^2 in the expansions for x_{0D} and x_{0i}. Hence, we arrive at the expansions (5).

Note that (5) and (3) impose a constraint on the $\mathcal{O}(\epsilon)$ terms in the genotype frequencies

$$\delta_{00}^{(1)} = \delta_{0D}^{(1)} + \sum_{i=1}^{n} \delta_{0i}^{(1)} \tag{6}$$

Also, note that (5) and (3) impose a constraint on the $\mathcal{O}(\epsilon^2)$ terms in the genotype frequencies

$$\delta_{00}^{(2)} = \delta_{0D}^{(2)} + \delta_{DD}^{(2)} + \sum_{i=1}^{n} \delta_{0i}^{(2)} + \sum_{i=1}^{n} \delta_{iD}^{(2)} + \sum_{i=1}^{n} \sum_{j=1}^{i} \delta_{ij}^{(2)}$$

Substituting (4), (1), (5), and (6) into the equation for $\dot{x}_{0D}$ in (2), we obtain

$$\dot{\delta}_{0D}^{(1)} = f_{00} \left(2 p_{0D,D} f_{0D} - f_{00}\right) \delta_{0D}^{(1)}$$

The drive allele invades a wild-type population if $\dot{\delta}_{0D}^{(1)} > 0$, i.e., if

$$2 p_{0D,D} f_{0D} > f_{00} \tag{7}$$

section S4. Stability of the drive construct

Consider a population in which the drive construct has fixed, so that all individuals have genotype DD. We perturb the DD population by introducing a small amount of the wild-type allele, 0. What happens? Is the DD population stable to perturbations, or does the wild-type allele or one of the resistance alleles invade the population?

For a perturbation to a population in which the drive construct has fixed, we write the frequencies of the genotypes as

$$\begin{aligned} x_{DD} = 1 \quad & -\epsilon \delta_{DD}^{(1)} - \epsilon^2 \delta_{DD}^{(2)} \quad && -\mathcal{O}(\epsilon^3) \\ x_{iD} = \quad & +\epsilon \delta_{iD}^{(1)} + \epsilon^2 \delta_{iD}^{(2)} \quad && +\mathcal{O}(\epsilon^3) \\ x_{ij} = \quad & + \epsilon^2 \delta_{ij}^{(2)} \quad && +\mathcal{O}(\epsilon^3) \end{aligned} \tag{8}$$

In Equations (8), it is implied that $0 \leq i \leq n$ and $0 \leq j \leq n$. The expansions (8) are understood as follows. The frequency of the drive allele is approximately one, since we only introduce a small amount of the wild-type allele. The frequency of the wild-type allele is of order $\epsilon \ll 1$. The small number of $0D$ individuals in the population also produce resistance alleles, and the frequency of these resistance alleles shortly after the perturbation is also small (i.e., of order $\epsilon \ll 1$). Notice that:

- New type DD individuals are produced by pairing two drive gametes (each at a frequency $\mathcal{O}(1)$), so new type DD individuals are generated at a rate $\mathcal{O}(1)$.

- New type iD individuals (for $0 \leq i \leq n$) are produced by pairing a non-drive gamete (at a frequency $\mathcal{O}(\epsilon)$) and a drive gamete (at a frequency $\mathcal{O}(1)$), so new type iD individuals are generated at a rate $\mathcal{O}(\epsilon)$.

- New type ij individuals (for $0 \leq i \leq n$ and $0 \leq j \leq n$) are produced by pairing two non-drive gametes (each at a frequency $\mathcal{O}(\epsilon)$), so new type ij individuals are generated at a rate $\mathcal{O}(\epsilon^2)$.

Also, notice that a nonzero amount of the non-drive alleles are produced at order ϵ^2 by type ij individuals, so there also exist terms of order ϵ^2 in the expansions for x_{iD}. Hence, we arrive at the expansions (8).

Note that (8) and (3) impose a constraint on the $\mathcal{O}(\epsilon)$ terms in the genotype frequencies

$$\delta_{DD}^{(1)} = \sum_{i=0}^{n} \delta_{iD}^{(1)} \tag{9}$$

Also, note that (8) and (3) impose a constraint on the $\mathcal{O}(\epsilon^2)$ terms in the genotype frequencies

$$\delta_{DD}^{(2)} = \sum_{i=0}^{n} \delta_{iD}^{(2)} + \sum_{i=0}^{n} \sum_{j=0}^{i} \delta_{ij}^{(2)} \tag{10}$$

Substituting (4), (1), (8), and (9) into the equations for $\dot{x}_{iD}$ in (2), we obtain

$$\dot{\delta}_{iD}^{(1)} = B_i \delta_{iD}^{(1)} + \sum_{k=0}^{i-1} A_{k,i} \delta_{kD}^{(1)} \tag{11}$$

Here, we use the shorthand notation

$$A_{k,i} = 2p_{kD,i} f_{kD} f_{DD}$$
$$B_i = A_{i,i} - f_{DD}^2$$

To solve (11), we take its Laplace transform. Using the notation $\Delta_{iD}^{(1)}(s) = \mathcal{L}\{\delta_{iD}^{(1)}(t)\}(s) = \int_0^\infty e^{-st} \delta_{iD}^{(1)}(t) dt$, we have

$$\Delta_{iD}^{(1)}(s) = \frac{1}{s - B_i} \delta_{iD}^{(1)}(0) + \frac{1}{s - B_i} \sum_{k=0}^{i-1} A_{k,i} \Delta_{kD}^{(1)}(s) \tag{12}$$

Here, we use $\delta_{iD}^{(1)}(0)$ to denote $\delta_{iD}^{(1)}(t)$ evaluated at time $t = 0$. Equation (12) specifies $\Delta_{iD}^{(1)}(s)$

in terms of each $\Delta_{kD}^{(1)}(s)$ for which $0 \leq k < i$. Simplifying, we have

$$\begin{aligned}\Delta_{iD}^{(1)}(s) = \frac{\delta_{iD}^{(1)}(0)}{s-B_i} + \sum_{k=0}^{i-1} \frac{\delta_{kD}^{(1)}(0)}{s-B_k} \Bigg[& \frac{A_{k,i}}{s-B_i} \\ & + \sum_{u=k+1}^{i-1} \frac{A_{k,u}A_{u,i}}{(s-B_u)(s-B_i)} \\ & + \sum_{u=k+1}^{i-2} \sum_{v=u+1}^{i-1} \frac{A_{k,u}A_{u,v}A_{v,i}}{(s-B_u)(s-B_v)(s-B_i)} \\ & + \sum_{u=k+1}^{i-3} \sum_{v=u+1}^{i-2} \sum_{w=v+1}^{i-1} \frac{A_{k,u}A_{u,v}A_{v,w}A_{w,i}}{(s-B_u)(s-B_v)(s-B_w)(s-B_i)} \\ & + \cdots \Bigg]\end{aligned} \tag{13}$$

We are interested in the time dependence of $\delta_{iD}^{(1)}(t)$. From Equation (13), notice that when the Laplace transform is inverted, the time dependence of each term in the resulting equation for $\delta_{iD}^{(1)}(t)$ has the form $t^\alpha \exp(B_j t)$, where $\alpha \geq 0$.

To demonstrate this, consider a set of real numbers $\{\beta_j\}$ and a set of positive integers $\{\nu_j\}$, and define $\mathcal{F}_k(s)$ for $k \geq 0$

$$\mathcal{F}_k(s) = \prod_{j=0}^{k} \frac{1}{(s-\beta_j)^{\nu_j}}$$

If the inverse Laplace transform of $\mathcal{F}_k(s)$, denoted by $\mathcal{L}^{-1}\{\mathcal{F}_k(s)\}(t)$, is equal to a sum of factors of the form $\mathcal{L}^{-1}\{1/(s-\beta_j)^\xi\}(t)$, where ξ is a positive integer, then each term in the solution for $\delta_{iD}^{(1)}(t)$ has the form $t^\alpha \exp(B_j t)$, where $\alpha \geq 0$.

To prove that $\mathcal{L}^{-1}\{\mathcal{F}_k(s)\}(t)$ is equal to a sum of factors of the form $\mathcal{L}^{-1}\{1/(s-\beta_j)^\xi\}(t)$, we use induction. Define

$$\mathcal{G}_{k+1}(t) = \mathcal{L}^{-1}\left\{\mathcal{F}_{k+1}(s)\right\}(t) = \mathcal{L}^{-1}\left\{\mathcal{F}_k(s)\frac{1}{(s-\beta_{k+1})^{\nu_{k+1}}}\right\}(t) \tag{14}$$

The inverse Laplace transform in (14) is calculated as follows

$$\mathcal{G}_{k+1}(t) = \int_0^t d\tau \ \left[\mathcal{L}^{-1}\left\{\mathcal{F}_k(s)\right\}(\tau)\right]\left[\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_{k+1})^{\nu_{k+1}}}\right\}(t-\tau)\right] \tag{15}$$

First, for the base case, consider Equation (14) for $k = 0$. We have

$$\mathcal{G}_1(t) = \mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_0)^{\nu_0}}\frac{1}{(s-\beta_1)^{\nu_1}}\right\}(t) \tag{16}$$

From (15), this becomes

$$\mathcal{G}_1(t) = \int_0^t d\tau \ \left[\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_0)^{\nu_0}}\right\}(\tau)\right]\left[\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_1)^{\nu_1}}\right\}(t-\tau)\right]$$

Substituting the expressions for $\mathcal{L}^{-1}\{1/(s-\beta_0)^{\nu_0}\}(\tau)$ and $\mathcal{L}^{-1}\{1/(s-\beta_1)^{\nu_1}\}(t-\tau)$, the equation for $\mathcal{G}_1(t)$ becomes

$$\mathcal{G}_1(t) = \int_0^t d\tau \ \left[\frac{\tau^{\nu_0-1}e^{\beta_0\tau}}{(\nu_0-1)!}\right]\left[\frac{(t-\tau)^{\nu_1-1}e^{\beta_1(t-\tau)}}{(\nu_1-1)!}\right]$$

Performing the integration over τ, we have

$$\begin{aligned}\mathcal{G}_1(t) = &\frac{(-1)^{\nu_0}}{(\nu_0-1)!(\nu_1-1)!(\beta_0-\beta_1)^{\nu_0}}\sum_{j=0}^{\nu_1-1}\binom{\nu_1-1}{j}\frac{(j+\nu_0-1)!}{(\beta_0-\beta_1)^j}\\ &\times\left[(\nu_1-j-1)!\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_1)^{\nu_1-j}}\right\}(t)\right.\\ &\left.-\sum_{k=0}^{j+\nu_0-1}(-1)^k\frac{(\nu_1-j+k-1)!}{k!}(\beta_0-\beta_1)^k\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_0)^{\nu_1-j+k}}\right\}(t)\right]\end{aligned}$$

Manipulating the indices and simplifying, we obtain

$$\begin{aligned}\mathcal{G}_1(t) = &\frac{(-1)^{\nu_0}}{(\nu_0-1)!(\nu_1-1)!(\beta_0-\beta_1)^{\nu_0+\nu_1}}\\ &\times\left[\sum_{j=1}^{\nu_1}\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_1)^j}\right\}(t)\binom{\nu_1-1}{\nu_1-j}(\nu_0+\nu_1-j-1)!(j-1)!(\beta_0-\beta_1)^j\right.\\ &-\sum_{j=1}^{\nu_1}\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_0)^j}\right\}(t)\sum_{k=0}^{j-1}(-1)^k\binom{\nu_1-1}{\nu_1-j+k}\frac{(\nu_0+\nu_1-j+k-1)!(j-1)!}{k!}(\beta_0-\beta_1)^j\\ &\left.-\sum_{j=\nu_1+1}^{\nu_0+\nu_1-1}\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_0)^j}\right\}(t)\sum_{k=0}^{\nu_1-1}(-1)^{\nu_1-j-k}\binom{\nu_1-1}{k}\frac{(\nu_0+k-1)!(j-1)!}{(j+k-\nu_1)!}(\beta_0-\beta_1)^j\right]\end{aligned} \tag{17}$$

We see that $\mathcal{G}_1(t)$ is equal to a sum of factors of the form $\mathcal{L}^{-1}\{1/(s-\beta_j)^{\xi}\}(t)$.

Next, consider Equation (14) for $k>0$. From (15), we have

$$\mathcal{G}_{k+2}(t) = \int_0^t d\tau \ \left[\mathcal{L}^{-1}\left\{\mathcal{F}_{k+1}(s)\right\}(\tau)\right]\left[\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_{k+2})^{\nu_{k+2}}}\right\}(t-\tau)\right]$$

This is equal to

$$\mathcal{G}_{k+2}(t) = \int_0^t d\tau \ \left[\mathcal{G}_{k+1}(\tau)\right]\left[\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_{k+2})^{\nu_{k+2}}}\right\}(t-\tau)\right] \tag{18}$$

For the inductive step, suppose that $\mathcal{G}_{k+1}(t)$ reduces to a sum of factors of the form $\mathcal{L}^{-1}\{1/(s-\beta_j)^{\xi}\}(t)$

$$\mathcal{G}_{k+1}(t) = \sum_j\sum_i\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_j)^{\xi_i}}\right\}(t) \tag{19}$$

Substituting (19) into (18), we have

$$\mathcal{G}_{k+2}(t) = \sum_j\sum_i\int_0^t d\tau \ \left[\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_j)^{\xi_i}}\right\}(\tau)\right]\left[\mathcal{L}^{-1}\left\{\frac{1}{(s-\beta_{k+2})^{\nu_{k+2}}}\right\}(t-\tau)\right]$$

This is equal to

$$\mathcal{G}_{k+2}(t) = \sum_j \sum_i \mathcal{L}^{-1}\left\{ \frac{1}{(s-\beta_j)^{\xi_i}} \frac{1}{(s-\beta_{k+2})^{\nu_{k+2}}} \right\}(t)$$

Then from Equations (16) and (17), we see that $\mathcal{G}_{k+2}(t)$ also necessarily reduces to a sum of factors of the form $\mathcal{L}^{-1}\{1/(s-\beta_j)^{\xi}\}(t)$, thus completing the proof.

Since $\delta_{iD}^{(1)}$ is equal to a sum of factors of the form $t^{\alpha}\exp(B_j t)$, where $\alpha \geq 0$, we see that if all $B_j < 0$, then all $\delta_{iD}^{(1)}$ approach zero in the long-time limit, and, from (9), we have that $\delta_{DD}^{(1)}$ approaches zero in the long-time limit. Therefore, if $B_j < 0$ for all values of $0 \leq j \leq n$, then the drive construct is evolutionarily stable.

If, instead, $B_j > 0$ for at least one value of j, then $\delta_{iD}^{(1)}$ has a term whose magnitude grows exponentially in time. The leading-order (in ϵ) terms in the expansions for x_{iD} in (8) are necessarily positive. Therefore, if the condition $B_j > 0$ is satisfied for at least one value of j, then $\delta_{iD}^{(1)}$ is positive and grows exponentially in time; i.e., the DD population is unstable to perturbations.

The resulting condition is that the DD population is stable to perturbations with a wild-type allele if

$$2\max\left(p_{kD,k} f_{kD}\right) < f_{DD} \tag{20}$$

4.1 Completely recessive fitness cost for a resistance mutation

Here, we consider a special case in which the fitness cost associated with having resistance to the drive is completely recessive. If the fitness of each heterozygote with a resistance allele, f_{kD}, exactly equals f_{DD} for all k, then is the DD population stable to perturbations? We expect that $p_{kD,k} < 1/2$ for all $0 \leq k < n$. Therefore, if $f_{kD} = f_{DD}$ for all k, then the inequality (20) is satisfied for all $k < n$ and becomes an equality for $k = n$.

All resistance alleles with at least one target ($0 \leq k < n$) are removed from the population by selective forces. We must focus on the fully resistant allele, n. To probe the stability of the DD population, we substitute (4), (1), (8), (9), and (10) into (2), and we keep terms that are $\mathcal{O}(\epsilon^2)$. We have

$$-\dot{\delta}_{DD}^{(2)} = f_{DD}\left(f_{DD} - 2f_{nn}\right)\delta_{nn}^{(2)} + \frac{1}{4} f_{DD}^2 \left[\delta_{DD}^{(1)}\right]^2 \tag{21}$$

We also have

$$\dot{\delta}_{nn}^{(2)} = -f_{DD}^2 \delta_{nn}^{(2)} + \frac{1}{4} f_{DD}^2 \left[\delta_{DD}^{(1)}\right]^2 \tag{22}$$

We can integrate (22). We get

$$\delta_{nn}^{(2)} = \frac{1}{4}\left[\delta_{DD}^{(1)}\right]^2 \left[1 - \exp\left(-f_{DD}^2 t\right)\right] \tag{23}$$

We are interested in the regime $1 \ll t \ll \epsilon^{-1}$. We must consider the sign of $\dot{\delta}_{DD}^{(2)}$ at large times $t \gg 1$ but before the terms in (8) become similar in magnitude. Our condition for stability of the DD population is therefore

$$\lim_{\substack{\epsilon t \to 0 \\ t \to \infty}} \dot{\delta}_{DD}^{(2)} < 0$$

Shortly after the perturbation, the exponential in the solution for $\delta_{nn}^{(2)}$ will approach zero. Substituting (23) into (21) and simplifying, we see that the DD population is stable to perturbations if

$$f_{nn} < f_{DD} \tag{24}$$

section S5. Interior equilibria

A drive construct increases in frequency when rare if Equation (7) is satisfied. A drive construct that has already fixed is stable to perturbations if Equation (20) is satisfied (or if Equation (24) is satisfied for the case of a completely recessive fitness cost for resistance). But if a small amount of the drive construct is introduced into a wild-type population, then does the drive spread completely to fixation?

To answer this question, it is helpful to know if the model for the drive dynamics, Equations (2), admits an interior equilibrium. Notice that, if all time derivatives are zero, then Equations (2) simplify to

$$\begin{aligned} x_{ij} &= \frac{(2-\delta_{ij})\,F_i F_j}{\psi^2} \\ x_{iD} &= \frac{2F_i F_D}{\psi^2} \\ x_{DD} &= \frac{F_D^2}{\psi^2} \end{aligned}$$

Next, we define x_i to equal the frequency of allele i in the population. Thus, x_0 is the frequency of the wild-type allele, and x_i for $1 \leq i \leq n$ is the frequency of a resistance allele with i damaged targets. Also, x_D is the frequency of the drive allele. These allele frequencies can be calculated from the frequencies of individuals of the various genotypes

$$\begin{aligned} x_i &= \frac{1}{2}x_{iD} + \sum_{j=0}^{n} \frac{1+\delta_{ij}}{2} x_{ij} \\ x_D &= x_{DD} + \frac{1}{2}\sum_{i=0}^{n} x_{iD} \end{aligned}$$

Similarly to Equation (3), the sum of all allele frequencies equals one at all times

$$x_D + \sum_{i=0}^{n} x_i = 1$$

We directly compute the following results

$$\begin{aligned} x_i^2 &= \left(\frac{1}{2}x_{iD} + \sum_{j=0}^{n}\frac{1+\delta_{ij}}{2}x_{ij}\right)^2 = \frac{F_i^2}{\psi^4}\left(F_D + \sum_{j=0}^{n}F_j\right)^2 = \frac{F_i^2}{\psi^2} = x_{ii} \\ x_D^2 &= \left(x_{DD} + \frac{1}{2}\sum_{j=0}^{n}x_{jD}\right)^2 = \frac{F_D^2}{\psi^4}\left(F_D + \sum_{j=0}^{n}F_j\right)^2 = \frac{F_D^2}{\psi^2} = x_{DD} \\ 2x_i x_D &= \frac{2F_i F_D}{\psi^2} = x_{iD} \\ (2-\delta_{ij})\,x_i x_j &= \frac{(2-\delta_{ij})\,F_i F_j}{\psi^2} = x_{ij} \end{aligned} \tag{25}$$

In summary, we obtain

$$\begin{aligned} x_{ij} &= (2-\delta_{ij})\,x_i x_j \\ x_{iD} &= 2x_i x_D \\ x_{DD} &= x_D^2 \end{aligned} \tag{26}$$

From (25), we have that

$$\begin{aligned} \psi x_i &= F_i \\ \psi x_D &= F_D \end{aligned} \tag{27}$$

By substituting Equation (4) for ψ and Equations (1) for F_i and F_D into (27), and substituting (26), we obtain

$$\left(f_{DD}x_D^2 + 2\sum_{k=0}^{n}f_{kD}x_k x_D + \sum_{j=0}^{n}\sum_{k=0}^{j}(2-\delta_{jk})\,f_{jk}x_j x_k\right)x_i = 2\sum_{k=0}^{i}p_{kD,i}f_{kD}x_k x_D + \sum_{k=0}^{n}f_{ki}x_k x_i \tag{28}$$

We also obtain

$$\left(f_{DD}x_D^2 + 2\sum_{k=0}^{n}f_{kD}x_k x_D + \sum_{j=0}^{n}\sum_{k=0}^{j}(2-\delta_{jk})\,f_{jk}x_j x_k\right)x_D = f_{DD}x_D^2 + 2\sum_{k=0}^{n}p_{kD,D}f_{kD}x_k x_D \tag{29}$$

Equations (28) and (29) must be simultaneously satisfied for $0 \leq x_D \leq 1$ and $0 \leq x_i \leq 1$ for each i at each interior fixed point. If Equations (28) and (29) cannot be simultaneously solved for a given set of parameter values, then no interior fixed point exists.

section S6. Numerical examples

Numerical simulations of Equations (2) are helpful for understanding the evolutionary dynamics of a drive construct. For simplicity, we consider a single guide ($n = 1$), and we choose the following parameter values

$$\begin{aligned} f_{00} = f_{10} &= 1 \\ f_{0D} = f_{1D} = f_{DD} &= 1-c \\ f_{11} &= 1-s \\ p_{0D,0} &= 0 \end{aligned} \tag{30}$$

We make the following assumptions: The fitness cost of the drive, c, is dominant. The fitness cost of the resistant allele, s, is recessive. Also, the drive construct in a $0D$ heterozygote always cuts at the target, and either the drive allele is copied by homologous recombination or resistance emerges. Thus, we have $p_{0D,0} = 0$.

In Fig. S1 (a and b), numerical simulations demonstrate evolutionary invasion of the drive construct. For these simulations, the initial condition is $x_{AA} = 1 - 10^{-4}$ and $x_{DD} = 10^{-4}$. The relevant condition for determining evolutionary invasion is Equation (7).

- In Fig. S1 (a), we set $p_{0D,D} = 0.75$ and $s = 0.4$. From Equation (7), the critical value of c for invasion is $1/3$. If $c = 0.34$ (green curve), then the drive construct does not invade. If $c = 0.33$ (blue curve), then the drive construct invades.

- In Fig. S1 (b), we set $p_{0D,D} = 0.65$ and $s = 0.3$. From Equation (7), the critical value of c for invasion is approximately 0.23. If $c = 0.235$ (green curve), then the drive construct does not invade. If $c = 0.225$ (blue curve), then the drive construct invades.

In Fig. S1 (c and d), numerical simulations demonstrate evolutionary stability of the drive construct. For these simulations, the initial condition is $x_{DD} = 1 - 10^{-2}$ and $x_{AA} = 10^{-2}$. From (30), notice that the condition (20) becomes an equality. Therefore, the relevant condition for determining evolutionary stability is Equation (24).

- In Fig. S1 (c), we set $p_{0D,D} = 0.75$ and $c = 0.32$. From Equation (24), the critical value of s for stability is 0.32. If $s = 0.315$ (green curve), then the drive construct is unstable. If $s = 0.325$ (blue curve), then the drive construct is stable.

- In Fig. S1 (d), we set $p_{0D,D} = 0.65$ and $c = 0.2$. From Equation (24), the critical value of s for stability is 0.2. If $s = 0.195$ (green curve), then the drive construct is unstable. If $s = 0.205$ (blue curve), then the drive construct is stable.

In Fig. S1 (e and f), numerical simulations demonstrate the behavior of the drive construct at intermediate frequencies. For these simulations, the initial condition is $x_{AA} = 1 - 10^{-4}$ and $x_{DD} = 10^{-4}$. If Equations (28) and (29) cannot simultaneously be solved numerically, then there is no interior equilibrium.

- In Fig. S1 (e), we set $p_{0D,D} = 0.75$ and $c = 0.32$. From numerical analysis of Equations (28) and (29), values of s that are slightly below approximately 0.815 permit an interior equilibrium, while values of s that are slightly above approximately 0.815 do not. If $s = 0.81$ (green curve), then the drive construct reaches an equilibrium frequency that is strictly between 0 and 1. If $s = 0.82$ (blue curve), then the drive construct spreads to fixation.

- In Fig. S1 (f), we set $p_{0D,D} = 0.65$ and $c = 0.2$. From numerical analysis of Equations (28) and (29), values of s that are slightly below approximately 0.285 permit an interior equilibrium, while values of s that are slightly above approximately 0.285 do not. If $s = 0.28$ (green curve), then the drive construct reaches an equilibrium frequency that is strictly between 0 and 1. If $s = 0.29$ (blue curve), then the drive construct spreads to fixation.

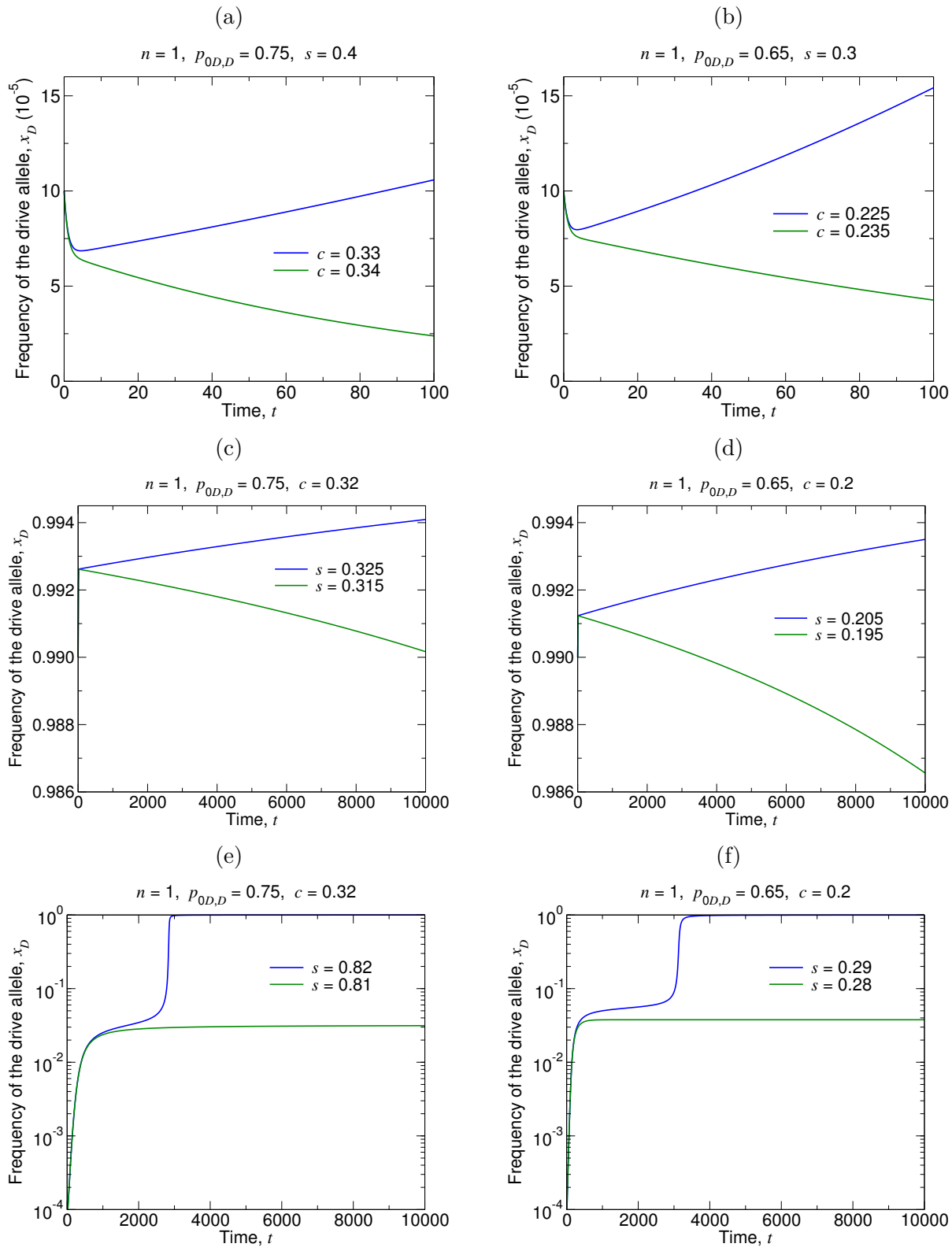

fig. S1. Numerical simulations of the evolutionary dynamics.

section S7. Neutral resistance

In this section, we present an extension of the model which accounts for the phenomenon of "neutral resistance". This can occur if non-homologous end joining results in repair at a cut site which disrupts the recognition sequence of a guide RNA while nonetheless leaving the function of the target gene intact. This can occur, for example, via an in-frame insertion or deletion or a synonymous mutation. The resulting allele is similar (with respect to the drive mechanism) to the resistant alleles discussed in previous sections: the repaired target is immune to cutting by its corresponding guide RNA. However, the mutation conferring this resistance is not deleterious.

We represent this scenario by an extension of our original model (section S2). We consider a drive allele, D, n "costly" resistant alleles, R_i (with $1 \leq i \leq n$), n "neutral" resistant alleles, S_i (with $1 \leq i \leq n$), and the wild-type allele, S_0. The drive mechanism works as follows (see Fig. 2 in the main text for an illustration):

Consider a type S_0D individual; one allele is wild-type, and the other allele is the drive. There are n guide RNAs and therefore n targets for the drive to cut. At meiosis, the drive can cut any number of targets between 0 and n. If the drive cuts no targets, then the individual remains with genotype S_0D. If the drive cuts k targets (with $1 \leq k \leq n$), then one of several things can happen: One possibility is that homologous recombination copies the drive allele onto the damaged chromosome, so that the individual's genotype becomes DD. Another possibility is that non-homologous end joining repairs the damaged chromosome without restoring the lost targets, and the resulting resistant allele is either costly, in which case the individual's genotype becomes R_iD (with $1 \leq i \leq n$), or cost-free, in which case the individual's genotype becomes S_iD (with $1 \leq i \leq n$). Yet another possibility is that non-homologous end joining perfectly repairs the damaged chromosome, so that the individual's genotype remains S_0D.

The drive allele can effect its spread as long as there is at least one remaining target. In an individual with genotype R_iD or S_iD, either the drive cuts at no targets, with the individual's genotype remaining R_iD or S_iD, or the drive cuts at some number, k, of the $n-i$ remaining targets (so that $1 \leq k \leq n-i$). After cutting, the individual can become homozygous in the drive allele (DD), the individual can lose additional targets by acquiring genotype R_jD or S_jD (with $i+1 \leq j \leq n$), or the individual can remain with genotype R_iD or S_iD. We assume that costly resistant alleles R_i cannot convert to cost-free resistant alleles S_j, but cost-free resistant alleles S_i can convert to costly resistant alleles R_j.

Using these rules, we can formally express the rates at which each of the $2n+2$ types of gametes are produced in terms of the frequencies of individuals in the population. We denote by $F_D(t)$ the rate (at time t) at which drive gametes (D) are produced by individuals in the population. We denote by $F_{S_i}(t)$ the rate (at time t) at which wild-type gametes ($i=0$) or gametes with varying levels of cost-free resistance ($1 \leq i \leq n$) are produced by individuals in the population. And we denote by $F_{R_i}(t)$ the rate (at time t) at which gametes with varying levels of costly resistance ($1 \leq i \leq n$) are produced by individuals in the population. We

have

$$F_D(t) = f_{DD}x_{DD}(t) + \sum_{k=1}^{n} p_{R_kD,D} f_{R_kD} x_{R_kD}(t) + \sum_{k=0}^{n} p_{S_kD,D} f_{S_kD} x_{S_kD}(t)$$

$$F_{S_i}(t) = \sum_{k=0}^{n} \frac{1+\delta_{ki}}{2} f_{S_kS_i} x_{S_kS_i}(t) + \frac{1}{2}\sum_{k=1}^{n} f_{R_kS_i} x_{R_kS_i}(t) + \sum_{k=0}^{i} p_{S_kD,S_i} f_{S_kD} x_{S_kD}(t)$$

$$\begin{aligned} F_{R_i}(t) = &\sum_{k=1}^{n} \frac{1+\delta_{ki}}{2} f_{R_kR_i} x_{R_kR_i}(t) + \frac{1}{2}\sum_{k=0}^{n} f_{R_iS_k} x_{R_iS_k}(t) \\ &+ \sum_{k=1}^{i} p_{R_kD,R_i} f_{R_kD} x_{R_kD}(t) + \sum_{k=0}^{i-1} p_{S_kD,R_i} f_{S_kD} x_{S_kD}(t) \end{aligned}$$

δ_{ki} is the Kronecker delta. $x_{IJ}(t)$ denotes the frequency of individuals (at time t) with genotype IJ, where $I, J = D, S_0, S_1, \ldots, S_n, R_1, \ldots, R_n$. Similarly, f_{IJ} is the fitness of IJ individuals, and $p_{IJ,K}$ denotes the probability of an individual with genotype IJ producing a K gamete. From conservation of probability, we have the following identities

$$p_{R_kD,D} + \sum_{i=k}^{n} p_{R_kD,R_i} = 1$$

$$p_{S_kD,D} + \sum_{i=k}^{n} p_{S_kD,S_i} + \sum_{i=k+1}^{n} p_{S_kD,R_i} = 1$$

Notice that type R_nD and type S_nD individuals are fully resistant to being manipulated by the drive construct; such a fully resistant individual shows standard Mendelian segregation in its production of gametes. Thus, we have

$$p_{R_nD,R_n} = p_{S_nD,S_n} = \frac{1}{2}$$

The selection dynamics are modeled by the following system of equations

$$\begin{aligned} \dot{x}_{DD}(t) &= F_D^2(t) - \psi^2(t) x_{DD}(t) \\ \dot{x}_{R_iD}(t) &= 2F_{R_i}(t)F_D(t) - \psi^2(t) x_{R_iD}(t) \\ \dot{x}_{S_iD}(t) &= 2F_{S_i}(t)F_D(t) - \psi^2(t) x_{S_iD}(t) \\ \dot{x}_{R_iS_j}(t) &= 2F_{R_i}(t)F_{S_j}(t) - \psi^2(t) x_{R_iS_j}(t) \\ \dot{x}_{R_iR_j}(t) &= (2-\delta_{ij})F_{R_i}(t)F_{R_j}(t) - \psi^2(t) x_{R_iR_j}(t) \\ \dot{x}_{S_iS_j}(t) &= (2-\delta_{ij})F_{S_i}(t)F_{S_j}(t) - \psi^2(t) x_{S_iS_j}(t) \end{aligned}$$

The quantity $\psi^2(t)$ represents a density-dependent death rate for the individuals in the population.

At any given time, t, we require that the total number of individuals sums to one

$$x_{DD}(t) + \sum_{i=1}^{n} x_{R_iD}(t) + \sum_{i=0}^{n} x_{S_iD}(t) + \sum_{i=1}^{n}\sum_{j=0}^{n} x_{R_iS_j}(t) + \sum_{i=1}^{n}\sum_{j=1}^{i} x_{R_iR_j}(t) + \sum_{i=0}^{n}\sum_{j=0}^{i} x_{S_iS_j}(t) = 1$$

To enforce this density constraint, we set

$$\psi(t) = F_D(t) + \sum_{i=1}^{n} F_{R_i}(t) + \sum_{i=0}^{n} F_{S_i}(t)$$

7.1 Invasion of the drive construct

The steps for determining if the drive construct invades when there is neutral resistance are the same as in section S3. The drive allele invades a wild-type population if

$$2p_{S_0D,D}f_{S_0D} > f_{S_0S_0}$$

7.2 Stability of the drive construct

The steps for determining if the drive construct is stable when there is neutral resistance are the same as in section S4. The DD population is stable to perturbations with a wild-type allele if

$$2 \max_{A \in S \cup R} (p_{AD,A} f_{AD}) < f_{DD}$$

7.3 Explicit cellular models of CRISPR gene drive

We now specify values of the inheritance probabilities, $p_{AB,C}$, and fitness values, f_{AB}, which explicitly describe possible scenarios by which a CRISPR gene drive acts within individuals. First, we specify a parameter set that corresponds with the behavior of CRISPR gene drives as described in prior literature. Then, we specify a parameter set that corresponds with our newly proposed CRISPR gene drive construct. These specified parameter sets for the previous and newly proposed drive constructs are used for the simulations of the previous and newly proposed drive constructs, respectively, in the main text.

7.3.1 Previous drives

For CRISPR gene drives as described in prior literature, $n = 1$. Reasonable choices for the fitness values and inheritance probabilities are as follows:

The wild-type has the maximum fitness of $f_{S_0S_0} = 1$, and the cost-free resistant allele, S_1, is identical to the wild-type allele, S_0, with respect to fitness. Disruption of the target gene produces a recessive fitness cost, s, and the gene drive construct produces a dominant fitness cost, c. However, since the previously demonstrated drive constructs copied themselves by inserting at (and thus disrupting) the target sequence, the drive allele contains a disrupted copy of the target gene. Thus, DD and RD individuals incur both the cost of the drive construct, c, and the cost of resistance, s. These two costs can be assumed to be independent so that the corresponding fitness effects are multiplicative, i.e., $(1-c)(1-s)$. Therefore, we have the following fitness values: $f_{DD} = f_{RD} = (1-c)(1-s)$, $f_{SD} = 1-c$, $f_{RR} = 1-s$, and $f_{RS} = f_{SS} = 1$.

We then compute the drive-heterozygote gamete production probabilities as follows:

- R_1D individuals produce R_1 gametes and D gametes equiprobably because the single target site is resistant to cutting, so we have

$$p_{R_1D,R_1} = p_{R_1D,D} = \frac{1}{2}$$

- S_1D individuals produce S_1 gametes and D gametes equiprobably because the single target site is resistant to cutting, so we have

$$p_{S_1D,S_1} = p_{S_1D,D} = \frac{1}{2}$$

- S_0D individuals produce S_0 gametes precisely when no cutting occurs. Since cutting occurs with probability q, we have

$$p_{S_0D,S_0} = \frac{1-q}{2}$$

- S_0D individuals produce D gametes by inheriting the existing D allele, or by cutting at the single target site with probability q and undergoing HR repair with probability P. We have

$$p_{S_0D,D} = \frac{1}{2} + \frac{qP}{2}$$

- S_0D individuals produce S_1 gametes by cutting at the single target site with probability q, undergoing NHEJ repair with probability $1-P$, and repairing the cut perfectly with probability γ. We have

$$p_{S_0D,S_1} = \frac{q(1-P)\gamma}{2}$$

- S_0D individuals produce R_1 gametes by cutting at the single target site with probability q, undergoing NHEJ repair with probability $1-P$, and repairing the cut imperfectly with probability $1-\gamma$. We have

$$p_{S_0D,R_1} = \frac{q(1-P)(1-\gamma)}{2}$$

7.3.2 Newly proposed drives

For our newly proposed CRISPR gene drive construct, any $n \geq 1$ is valid. Reasonable choices for the fitness values and inheritance probabilities are as follows:

The wild-type has the maximum fitness of $f_{S_0S_0} = 1$, and cost-free resistant alleles, S_i, are identical to the wild-type allele, S_0, with respect to fitness. The cost, c, conferred by the drive is dominant, while the cost, s, conferred by costly resistant alleles—which are disrupted copies of the target gene—is recessive. Furthermore, we assume that the drive allele contains a functional copy of the target gene, so DD and RD individuals do not incur the recessive fitness cost for target disruption. Thus, we have $f_{DD} = f_{RD} = f_{SD} = 1-c$, $f_{RR} = 1-s$, and $f_{RS} = f_{SS} = 1$.

We then assign values to the drive-heterozygote gamete production probabilities according to the biological description outlined in the main text and illustrated in Fig. 2B. We first define a probability density, $P_K(k \mid n, i, q)$, which describes the probability that k target sites undergo cutting, given that there are n total targets, of which i are currently resistant to cutting, and where each of the $n-i$ susceptible targets are cut independently with probability q. This distribution is binomial, specifically:

$$P_K(k \mid n, i, q) = \binom{n-i}{k} q^k (1-q)^{n-i-k}$$

This distribution is defined for $0 \leq k \leq n - i$.

In the case that two or more cuts occur, we assume that all target sites between the two outermost cuts are lost due to loss of the intervening DNA sequence. To account for this effect, we further define a probability density, $P_L(l \mid k, n, i)$, which describes the probability that l targets are lost given k cuts, n total target sites, and i currently resistant sites. This distribution can be straightforwardly computed

$$P_L(l \mid k, n, i) = (n - i - l + 1)\binom{l-2}{k-2} \Big/ \binom{n-i}{k}$$

This distribution is defined for $2 \leq k \leq l \leq n - i$.

We then compute the drive-heterozygote gamete production probabilities as follows:

- R_iD individuals produce D gametes by inheriting the existing D allele, or by cutting at one or more sites on the R_i chromosome (each with probability q) and undergoing HR repair (with probability P). We have

$$p_{R_iD,D} = \frac{1}{2} + \frac{P}{2}(1 - (1-q)^{n-i})$$

- R_iD individuals produce R_i gametes precisely when no cutting occurs. Each of the $n - i$ sites is susceptible to cutting, and cutting occurs independently at each with probability q, so we have

$$p_{R_iD,R_i} = \frac{1}{2}(1-q)^{n-i}$$

- R_iD individuals produce R_{i+1} gametes (with $i < n$) by cutting at exactly one target site (where each is cut independently with probability q) and undergoing NHEJ repair (with probability $1 - P$). Since we assume that costly resistant alleles cannot convert back to cost-free alleles, we do not consider the efficacy of repair by NHEJ. In this case, we have

$$p_{R_iD,R_{i+1}} = \frac{1-P}{2}(n-i)q(1-q)^{n-i-1}$$

- R_iD individuals produce R_k gametes (with $i + 2 \leq k \leq n$) by losing $k - i$ target sites and undergoing NHEJ repair (with probability $1 - P$). Since we assume that costly resistant alleles cannot convert back to cost-free alleles, we do not consider the efficacy of repair by NHEJ. In this case, we have

$$p_{R_iD,R_k} = \frac{1-P}{2} \sum_{j=2}^{k-i} P_L(k-i \mid j, n, i) P_K(j \mid n, i, q)$$

  The sum is over the number of simultaneous cuts, j, which could possibly give rise to a loss of $k - i$ targets.

- S_iD individuals produce D gametes by inheriting the existing D allele, or by cutting at one or more sites on the S_i chromosome (each with probability q) and undergoing HR repair (with probability P). We have

$$p_{S_iD,D} = \frac{1}{2} + \frac{P}{2}(1-(1-q)^{n-i})$$

- S_iD individuals produce S_i gametes precisely when no cutting occurs. Each of the $n-i$ sites is susceptible to cutting, and cutting occurs independently at each with probability q, so we have

$$p_{S_iD,S_i} = \frac{1}{2}(1-q)^{n-i}$$

- S_iD individuals produce S_{i+1} gametes (with $i < n$) by cutting at exactly one target site (where each is cut independently with probability q), undergoing NHEJ repair (with probability $1-P$), and repairing the cut perfectly (with probability γ). We have

$$p_{S_iD,S_{i+1}} = \frac{1-P}{2}(n-i)q(1-q)^{n-i-1}\gamma$$

- S_iD individuals do not produce S_k gametes when $k \geq i+2$. This is because cutting at two or more target sites would lead to a large deletion in the intervening DNA sequence, resulting in loss of target gene function. Thus

$$p_{S_iD,S_{i+2}} = \cdots = p_{S_iD,S_n} = 0$$

- S_iD individuals produce R_{i+1} gametes (with $i < n$) by cutting at exactly one target site (where each is cut independently with probability q), undergoing NHEJ repair (with probability $1-P$), and repairing the cut imperfectly (with probability $1-\gamma$). We have

$$p_{S_iD,R_{i+1}} = \frac{1-P}{2}(n-i)q(1-q)^{n-i-1}(1-\gamma)$$

- S_iD individuals produce R_k gametes (with $i+2 \leq k \leq n$) by losing $k-i$ target sites and undergoing NHEJ repair (with probability $1-P$). This is because cutting at two or more target sites would lead to a large deletion in the intervening DNA sequence, resulting in loss of target gene function. Thus we have

$$p_{S_iD,R_k} = \frac{1-P}{2}\sum_{j=2}^{k-i} P_L(k-i \mid j,n,i)P_K(j \mid n,i,q)$$

  The sum is over the number of simultaneous cuts, j, which could possibly give rise to a loss of $k-i$ targets.

For the numerical simulations of both the previous and newly proposed drive constructs shown in the main text, we set $q = P = 0.95$ and $\gamma = 1/3$.