Designing gene drives to limit spillover to non-target populations

Gili Greenbaum[1], Marcus W. Feldman[1], Noah A. Rosenberg[1], and Jaehee Kim[1]

[1] *Department of Biology, Stanford University, Stanford, California, United States of America*

Abstract

The prospect of utilizing CRISPR-based gene drive technology for controlling populations, such as invasive and disease-vector species, has generated much excitement. However, the potential for spillovers of gene drive alleles from the target population to non-target populations — events that may be ecologically catastrophic — has raised concerns. Here, using two-population mathematical models, we investigate the possibility of limiting spillovers and impact on non-target populations by designing differential-targeting gene drives, in which the expected equilibrium gene drive allele frequencies are high in the target population but low in the non-target population. We find that achieving differential targeting is possible with certain configurations of gene drive parameters. Most of these configurations ensure differential targeting only under relatively low migration rates between target and non-target populations. Under high migration, differential targeting is possible only in a narrow region of the parameter space. When migration is increased, differential-targeting states can sharply transition to states of global fixation or global loss of the gene drive. Because fixation of the gene drive in the non-target population could severely disrupt ecosystems, we outline possible ways to avoid this outcome. Our results emphasize that, although gene drive technology is promising, understanding the potential consequences for populations other than the targets requires detailed analysis of gene drive spillovers, and that ways to limit the unintended effects of gene drives should be explored prior to the application of gene drives in natural settings.

1 Introduction

Gene drives are genetic constructs that can bias transmission of desired alleles to progeny, allowing them to rapidly increase in frequency even when the alleles are negatively-selected. They, therefore, have the potential to alter or even eradicate entire species. The population genetics of gene drives, under various genetic architectures, has been studied for several decades [1–8]. With recent innovation in technical engineering of gene drives using CRISPR/Cas9-based methods [9], gene drives have attracted considerable attention for their potential applications. In particular, engineered gene drives can potentially alter or eradicate disease vectors, agricultural pests, or invasive species [10–13].

However, the potential of this technology also raises significant concerns due to the possibility of gene drive spillovers to non-targeted populations [10, 14–17]. The effects of such spillovers could be

devastating, unintentionally driving species to extinction or permanently altering important traits, potentially leading to ecological cascades [15]. With invasive-species control, there is particular concern, because every invasive species is non-invasive in its native range. Moreover, because an invasion has occurred, it is likely that invaded regions are connected to native ones through migration. As a result, to prevent gene drive spillovers, every application of a gene drive in an invaded region must be designed to avoid them. For example, it has been suggested that CRISPR-based gene drives could be applied in New Zealand to eradicate invasive species, such as Australian possums, stoats, and rats [15, 17–19]. However, such plans must account for the possibility and potential consequences of spillovers of gene drives from New Zealand to the native ranges of these species. Therefore, understanding the dynamics of gene drives with CRISPR-based constructs under migration, in the context of spillovers to non-target populations, is crucial.

Meiotic drives are systems in which Mendel's law of equal segregation [20] is violated, with preferential transmission of particular alleles to subsequent generations [21]. Because many general principles of population-genetic theory are violated by non-Mendelian segregation [6], models of various types of meiotic drive have been studied extensively [1–3, 5, 22–26]. Of particular interest are cases in which alleles with enhanced transmission cause reduced fitness of individuals who carry them. Analyses of meiotic drives have focused on various types of genetic architectures and evolutionary models [11], such as modifier genes [5, 25] and sex-ratio distorters [3, 26].

CRISPR-based gene drives are particular cases of preferential transmission of alleles. In this case, the preferential transmission is generated at the zygote formation stage by conversion of heterozygotes carrying one copy of the gene drive allele to homozygotes with two copies of the gene drive allele [8, 9]. This conversion occurs when the CRISPR mechanism, which is incorporated in the gene drive allele, edits the other chromosome to replace the wild-type allele with the gene drive allele.

One of the features of meiotic drive models is the existence of polymorphic equilibria, namely the states where meiotic driver alleles persist in the population together with wild-type alleles, rather than sweeping to fixation or loss. These equilibria can be either stable or unstable, depending on the genetic architecture involved [1, 2, 4–6, 25, 27]. With CRISPR-based gene drive technology, it has been suggested that, in a single isolated population, unstable polymorphic equilibria could be utilized for generating a biosafety measure to address the dangers of accidental releases [8]. If a gene drive is initiated at a frequency below an unstable equilibrium, it is expected to be driven to loss, whereas it is expected to be driven towards fixation if initiated at frequencies above the equilibrium [1, 8, 28–30]. Similarly, in spatially continuous populations, it has been argued that a CRISPR-based gene drive can be engineered to be driven to fixation only once it is introduced over a large enough area [31].

However, few natural populations exist in isolation, and distinct populations are often connected via migration. Therefore, in order to understand potential consequences of gene drive spillovers, explicit incorporation of migration between popualtions into gene drive models is required. Some studies have examined migration models of meiotic drives with genetic architectures

other then CRISPR-based [7, 32–37]. Among these, some have found that polymorphic equilibria can exist for low migration rates, but not necessarily for high migration rates.

Polymorphic equilibrium states that represent stable conditions under which gene drive allele frequencies are high in one population but low in another, might potentially be exploited to mitigate spillovers. This approach would require identification of stable states in which gene drive frequencies are high in the target population and low in the non-target population, configuring the genetic architecture of a CRISPR-based gene drive to attain these states, and initiating the gene drive such that it would converge to these states. We term this approach *differential targeting*. In order to consider the prospect of mitigating gene drive spillovers through differential targeting, gene drive models that incorporate migration and CRISPR-based genetic architectures are required.

Here, we develop and investigate such models in the context of gene drive spillovers. We focus on identifying gene drive designs that allow for differential targeting, and we evaluate the feasibility of the approach for mitigating spillovers.

2 Modeling CRISPR-based gene drives with migration

The dynamics of gene drives introduced into a wild population depend in part on features that can, at least in principle, be configured by researchers, such as the gene drive phenotype and the conversion rate, and in part on features that cannot be controlled, such as the ecological circumstances and life-history traits of the species. In particular, the migration levels between populations are not a controlled feature of the gene drive. Here, in addition to migration, we consider several features of gene drives: (1) the selection coefficients of individuals carrying the gene drive allele, coefficients that are related to the designed gene drive phenotype; (2) the life-stage at which the gene drive phenotype is expressed and subjected to natural selection, specifically whether selection acts before or after the typical migratory life-stage; (3) the degree of dominance of the gene drive allele relative to the wild-type allele; (4) the efficiency of the gene drive conversion mechanism.

We model a CRISPR-based gene drive [9], following a previously described model [8]. We consider a population with a wild-type allele, a, and a gene drive allele, A, which is initially absent from the population. The gene drive is characterized by (1) the conversion rate c of heterozygotes carrying the gene-drive allele to homozygotes of the gene-drive allele (from Aa to AA), with $c = 1$ being full conversion and $c = 0$ being regular Mendelian inheritance; (2) the degree of dominance h of the gene-drive allele; (3) the selection coefficient s of a homozygote for the gene-drive allele relative to the wild type homozygote. In other words, the fitnesses are $1 - s$ for the AA genotype, $1 - hs$ for the Aa genotype, and 1 for the aa genotype. For $s > 0$, the gene-drive allele is negatively-selected, and we discuss only this case here, as a beneficial gene drive is expected to be driven to fixation in a deterministic model in all connected subpopulations.

2.1 One-deme model

We first introduce the recursion equation from [8] describing the change in gene drive allele frequency from one generation to the next, for a single isolated population. Assuming Hardy-Weinberg equilibrium, and denoting the gene drive allele frequency by q, the frequencies of the genotypes AA, Aa, and aa are q^2, $2q(1-q)$, and $(1-q)^2$, respectively. We denote the allele frequency in the subsequent generation as q'.

The AA homozygote fully contributes to q', and is subject to a selection coefficient of $1-s$; in total, the relative contribution of AA genotypes to q' is $q^2(1-s)$ (red arrow in Fig. 1A). Unconverted heterozygotes, Aa, contribute only half of the contribution of AA, since they pass on the A allele to offspring at probability 0.5. The unconverted heterozygotes represent $1-c$ of the total heterozygotes, and they are subject to a selection coefficient of $1-hs$; we denote $s_n = 0.5(1-c)(1-hs)$, and the total relative contribution of unconverted heterozygotes is $2q(1-q)s_n$ (green arrow in Fig. 1A). Converted heterozygotes represent a proportion c of the heterozygotes, but because they are converted to AA genotypes, they fully contribute to q' and are subject to a selection coefficient of $1-s$; we denote $s_c = c(1-s)$, and the relative contribution of converted heterozygotes is $2q(1-q)s_c$.

Combining the three cases and normalizing by the mean fitness in the population so that frequencies sum to 1, we obtain the following recursion relation for the gene drive allele frequency:

$$q' = \frac{q^2(1-s) + 2q(1-q)(s_n + s_c)}{\overline{w}}, \tag{1}$$

where the mean fitness is expressed as $\overline{w} = q^2(1-s) + 2q(1-q)(2s_n + s_c) + (1-q)^2$. The numerator in Eq. 1 describes the genotype frequencies weighted by their relative fitnesses and their contribution to the A allele frequency (grey box in Fig. 1A), and the denominator normalizes this relative contribution by the mean fitness in the population (orange arrow in Fig. 1A).

2.2 Two-deme models

To incorporate migration between the target population and a non-target population into this modeling framework, we consider possible extensions to the one-deme model. We assume two connected demes, each large and panmictic, with symmetric migration at a rate m between them (Fig. 1B). As in the one-deme model, we track the dynamics of the gene drive by obtaining the change in allele frequencies of the A allele in the two demes, q_1 and q_2, respectively.

We consider two types of models that lie at two extremes for the timing of migration relative to the life stage at which the gene drive phenotype affects fitness: (1) Migration occurs before selection, meaning that individuals migrate at a relatively early life stage, and the fitness consequences of the gene drive phenotypes are expressed in the deme to which the individuals have migrated (e.g. phenotypes expressed at late life-stages, such as during reproduction). We call this the m-before-s model. (2) Selection occurs before migration, at an early life stage, and individuals

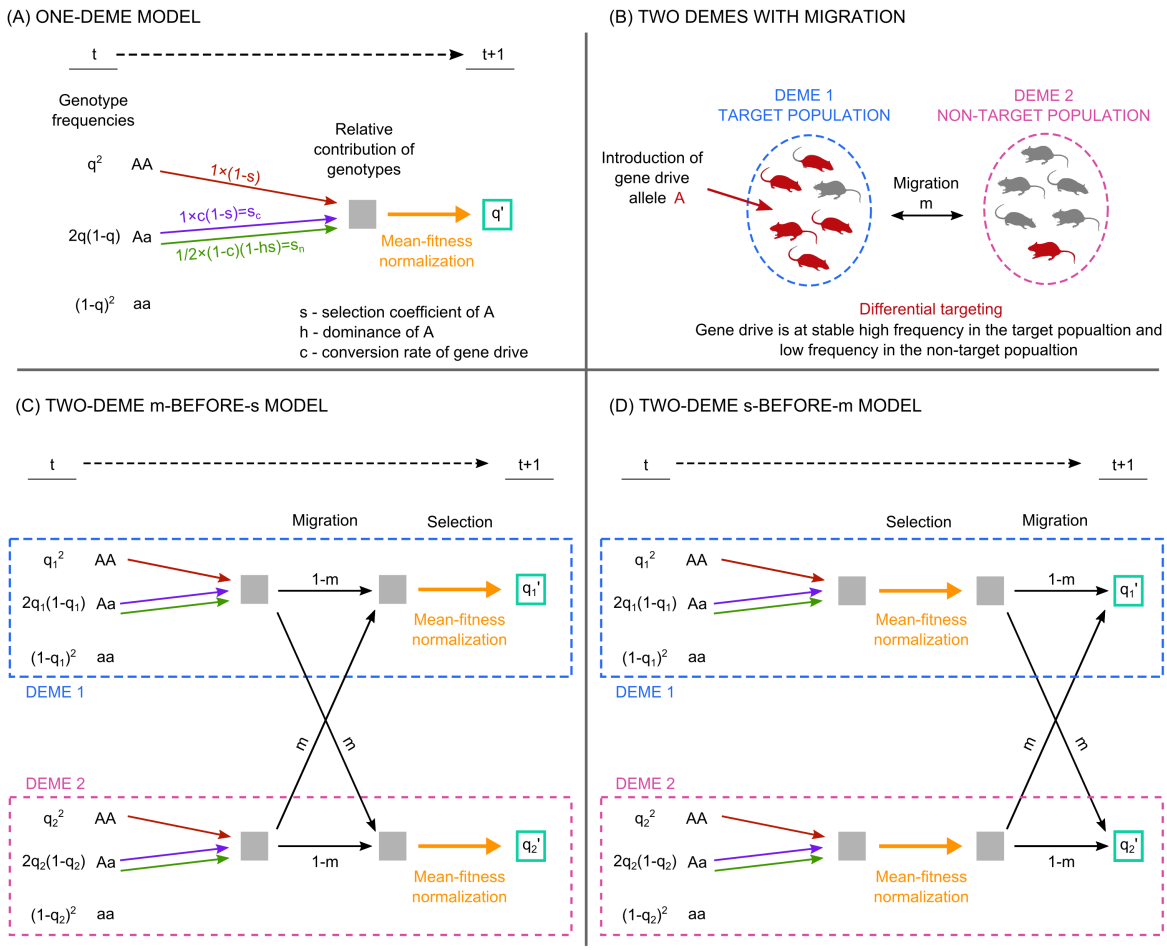

Figure 1: **Schematic depiction of gene drive models.** (A) Model of CRISPR-based gene drive in an isolated population (Eq. 1). Shown is the change of the allele frequency of the gene drive allele A over one generation, from t to $t+1$. Each genotype contributes to the frequency of A in generation $t+1$ depending on its frequency in generation t and the genotype selection coefficient: the AA genotypes contributes an A allele (red arrows), and the heterozygous genotype Aa contributes 1 (purple arrow) or 1/2 allelic copies (green arrow), depending on whether gene drive conversion occurs (at rate c). (B) Representation of a two-deme configuration with migration. The gene drive is introduced to the target population, and can spread to the non-target population through migration. Differential targeting, when possible, would produce a convergence to a stable state in which gene drive frequencies are high in deme 1 and low in deme 2. (C) Representation of the m-before-s two-deme model (Eq. 3). Black arrows denote migration. Migration occurs before selection (Eq. 2), and q_i' is calculated by normalizing the relative contributions to the frequency of A by the mean fitness of the post-migration population (orange arrow). (D) Representation of the s-before-m two-deme model (Eq. 4). Selection and mean-fitness normalization occur before migration, and q_i' is calculated by considering the effect of migration on the post-selection allele frequencies.

experience selection in their parental deme (e.g. phenotypes expressed at the juvenile stage). We call this the s-before-m model.

2.2.1 Migration before selection (m-before-s) model

In the the m-before-s model (Fig. 1C), we first consider the allele frequencies of A after migration but before selection, $\widetilde{q}_1$ and $\widetilde{q}_2$ in demes 1 and 2, respectively. These frequencies are obtained

by accounting for the relative contributions to the A allele frequencies of residents and migrants (Fig. 1C, black arrows):

$$\begin{aligned} \widetilde{q}_1 &= (1-m)q_1 + mq_2 \\ \widetilde{q}_2 &= (1-m)q_2 + mq_1. \end{aligned} \tag{2}$$

Next, we consider the frequencies of A after selection takes place on the post-migration gene pools (Fig. 1C, orange arrows). The changes in the allele frequencies between generations in each population are obtained as in Eq. 1, using the post-migration frequencies from Eq. 2:

$$\begin{aligned} q_1' &= \frac{\widetilde{q}_1^2(1-s) + 2\widetilde{q}_1(1-\widetilde{q}_1)(s_n+s_c)}{\overline{w}_1} \\ q_2' &= \frac{\widetilde{q}_2^2(1-s) + 2\widetilde{q}_2(1-\widetilde{q}_2)(s_n+s_c)}{\overline{w}_2}, \end{aligned} \tag{3}$$

where $\overline{w}_1$ and $\overline{w}_2$ are the post-migration mean fitnesses of demes 1 and 2, respectively, expressed as $\overline{w}_i = \widetilde{q}_i^2(1-s) + 2\widetilde{q}_i(1-\widetilde{q}_i)(2s_n+s_c) + (1-\widetilde{q}_i)^2$.

2.2.2 Selection before migration (s-before-m) model

In the s-before-m model, selection acts on the offspring in their parental deme, and we first determine the pre-migration frequencies of A by considering the relative contributions of the genotypes and normalizing by the mean fitness for each of the demes, as in Eq. 1 (Fig. 1D, orange arrows). These fitnesses, $\overline{w}_1$ and $\overline{w}_2$, are expressed as $\overline{w}_i = q_i^2(1-s) + 2q_i(1-q_i)(2s_n+s_c) + (1-q_i)^2$. Next, we consider the impact of migration on the pre-migration gene pools of the demes, considering the relative proportion of migrants and residents, as in Eq. 2 (Fig.1D, black arrows). Combining these two steps, we obtain:

$$\begin{aligned} q_1' &= (1-m)\frac{q_1^2(1-s) + 2q_1(1-q_1)(s_n+s_c)}{\overline{w}_1} + m\frac{q_2^2(1-s) + 2q_2(1-q_2)(s_n+s_c)}{\overline{w}_2} \\ q_2' &= (1-m)\frac{q_2^2(1-s) + 2q_2(1-q_2)(s_n+s_c)}{\overline{w}_2} + m\frac{q_1^2(1-s) + 2q_1(1-q_1)(s_n+s_c)}{\overline{w}_1}. \end{aligned} \tag{4}$$

The terms on the left in the sums represent the contributions of residents to the frequencies, and the terms on the right represent the contributions of migrants.

3 Differential-targeting equilibria (DTE)

In order to understand the evolutionary trajectories of the gene drive in the two-deme system, we study the equilibrium states of the models. This is done by solving Eq. 3, for the m-before-s model, or Eq. 4, for the s-before-m model, with the equilibrium conditions $q_1' = q_1$ and $q_2' = q_2$ (for the equilibrium states of the one-deme model, Eq. 1, see [8]). We denote these solutions, the equilibrium points in frequency space, as $(\hat{q}_1, \hat{q}_2)$, where $\hat{q}_1$ and $\hat{q}_2$ are the equilibrium frequencies

in demes 1 and 2, respectively. We consider only solutions for which both $\hat{q}_1$ and $\hat{q}_2$ lie in the interval $[0, 1]$, and we analyze the stability of these equilibria (see *Methods*).

A gene drive configuration is denoted as (s, c, h), where s, c, and h can have any value between 0 and 1, including 0 or 1. Under any gene drive configuration, there are two trivial equilibrium points, corresponding to global fixation, $(\hat{q}_1, \hat{q}_2) = (1, 1)$, and global loss, $(\hat{q}_1, \hat{q}_2) = (0, 0)$, of the gene drive allele. In the one-deme model, it has been demonstrated that for some gene drive configurations (s, c, h), there may exist a single non-trivial (polymorphic) equilibrium point, which could be either stable or unstable [8]. We term the set of gene drive configurations (s, c, h) for which a non-trivial stable equilibrium exists in the one-deme model by W. We denote configurations (s, c, h) for which a non-trivial unstable equilibrium exists by U. Thus, W and U are disjoint sets. Note that for configurations in W, the two trivial equilibria are unstable, whereas for configurations in U, the two trivial equilibria are stable.

We can leverage the results obtained for the one-deme model to explore equilibria in the two-deme models. For the two-deme models, the numbers of equilibrium solutions for Eq. 3 and Eq. 4 depend on the gene drive configurations (s, c, h), and also on the migration rate m (Fig. 2). For tractability, we first consider an ecologically uninteresting yet illustrative case, where $m = 0$. In this case of no migration, the system of two-deme becomes two independent one-deme system. As shown previously, because each deme has two trivial equilibria and one non-trivial equilibrium under W or U configuration, the number of equilibria of two disconnected demes is $3 \times 3 = 9$ for such configurations (for example, Fig. 2A). Of these equilibria, one equilibrium point, $(1, 0)$, represents a desired outcome, in which only the target deme (population 1) is affected by the gene drive, while it is absent from the non-target deme (population 2). However, this equilibrium is stable only for U configurations, and it is unstable for W configurations. For each stable equilibrium point, the region in allele frequency space for which gene drives initiated at those frequencies eventually converge to the stable equilibrium point is termed the *basin of attraction* of that equilibrium. If a gene drive with a configuration in U is initiated in the basin of attraction of the stable equilibrium point $(1, 0)$ (e.g., yellow region in Fig. 2A), then we expect that the desired outcome, the gene drive sweeping the target population and not the non-target population, will be reached.

Next, we consider $m > 0$, and we explore the possible number and type of equilibria attainable in Eqs. 3 and 4 by numerically exploring the entire parameter space of possible gene drive configurations and migration rates. We focus on W and U configurations, which are defined in respect to the one-deme model, and are not dependent on the migration rate m. For both the m-before-s and s-before-m models, for U configurations, and only for U configurations we numerically find that 9 equilibria still exist for some $m > 0$. For U configurations, 9 equilibria exist for low $m > 0$, whereas for higher migration rates there exist fewer equilibria (Figs. S1–S6). For example, 9 equilibria exist in Fig. 2A–F, but only 5 exist in Fig. 2G–J. For $m > 0$, when 9 equilibria exist, and only for these cases, there exists a single stable equilibrium point $(\hat{q}_1, \hat{q}_2)$ for which $\hat{q}_1 > \hat{q}_2$. We term such an equilibrium point a "differential-targeting" equilibrium (DTE), since by initiating

the gene drive in the attraction basin of the DTE, the system converges to frequencies that are higher in the target population than in the non-target population. In other words, we define an equilibrium point as a DTE if it is stable, and if the equilibrium gene drive frequency in the target deme is larger than the frequency in the non-target deme. For example, in Fig. 2A–F, the yellow points are DTEs, and the yellow regions around them are the corresponding basins of attraction. At the DTEs, the gene drive is not expected to be absent from the non-target population, but rather it is maintained at a frequency significantly lower than that in the target population.

Notably, for $m > 0$, we observe that 9 equilibria exist only for U configurations and for low migration rates (Figs. S1–S6). For these U configurations, when migration exceeds a certain critical threshold, we observe that there are no longer 9 equilibrium solutions to Eq. 3 and Eq. 4, but only 5 or 3 solutions, none of which are both stable and allow differential targeting (Figs. S1–S6). We term this threshold of existence of a DTE as $m^*(s, c, h)$ (Fig. 2). In other words, $m^*(s, c, h)$ is defined as the supremum migration rate such that for all $m < m^*(s, c, h)$ there are 9 equilibria for Eq. 3 or Eq. 4 with the parameters s, c, h, and m. For (s, c, h) configurations that are not in U, $m^*(s, c, h)$ is undefined , because no DTEs exist, for any $m > 0$.

From the numerical solutions of equilibria across the parameter space, we observe that for scenarios where $m > m^*$, only two stable equilibria exist — global fixation and global loss. This means that differential-targeting of a gene drive is not possible when $m > m^*$, and spillover to deme 2 of a gene drive that efficiently affects deme 1 is unavoidable.

4 Migration and differential-targeting

We numerically obtained m^* for the entire possible parameter range U using the numerical equilibrium solution for Eqs. 3 and 4 (Fig. 3). We find that in both models, for most of the parameter range, m^* is very low (blue regions in Fig. 3), except for a narrow curved band across the parameter space (pale yellow bands in Fig. 3). m^* increases for higher conversion rates c, and it is maximal for $c = 1$ and $s \approx 0.7197$ under both the m-before-s ($m^* \approx 0.110$) and the s-before-m ($m^* \approx 0.109$) models; we denote the selection coefficient maximizing m^* as s^*. Note that h values are not relevant for a full-conversion gene drive ($c = 1$), since in such configurations the system has no heterozygous individuals.

4.1 Impact of differential-targeting on the non-target population

For a gene drive with configuration (s, c, h) in U applied in a two-deme system, and assuming that in practice m cannot be determined with certainty and is not necessarily fixed over time, it is important to understand the potential consequences for the non-target population of differential targeting, for a range of migration rates. In principle, initiation of the gene drive in the basin of attraction of the DTE leads to a high frequency of A in deme 1 and a low frequency in deme 2, and this state would be maintained until the population in deme 1 begins to collapse due to the

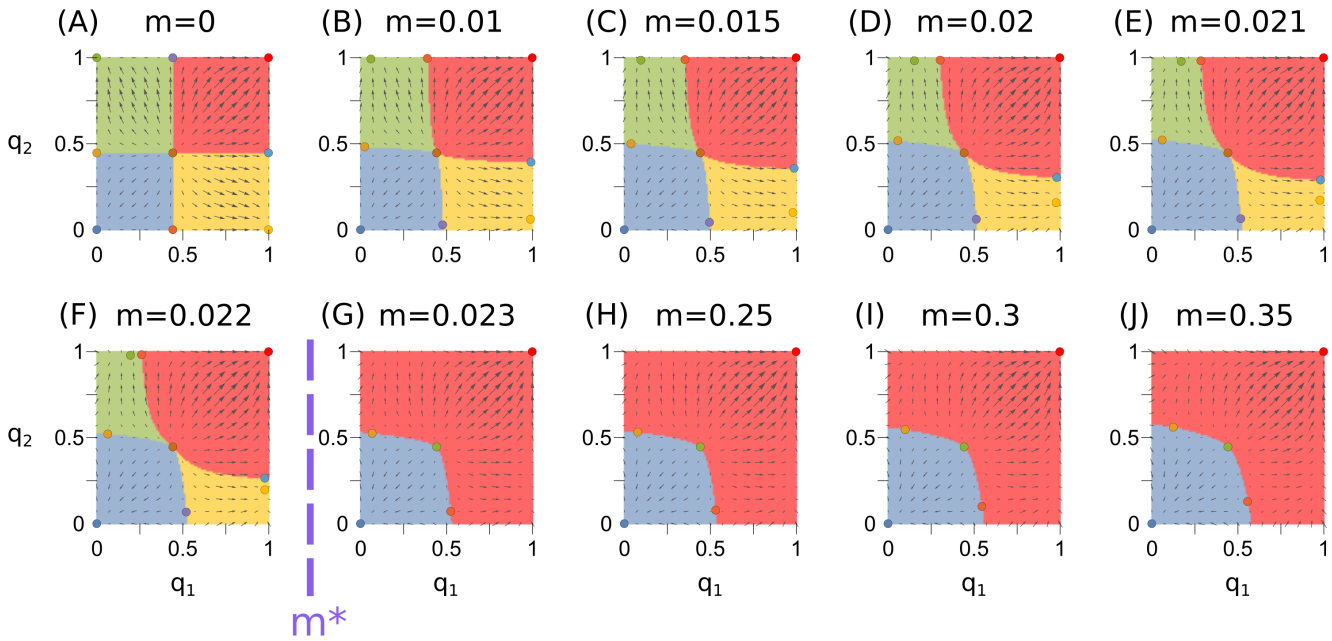

Figure 2: **Equilibria and basins of attractions for different migration rates.** Shown are results for the m-before-s model with a gene drive configuration of $h = 0, s = 0.6$, and $c = 0.8$ (a configuration in U). The circled points show the equilibria. The colored regions show the attraction basins, with the basin colors corresponding to the stable equilibria of each basin. The arrows show the vector field that describes the magnitude and direction of the change in allele frequencies at each point in frequency space. The differential-targeting equilibrium (DTE) is the stable yellow equilibrium point, which exists for migration rates lower than $m^* \approx 0.023$. .

population-level impact of the gene drive (assuming the goal is eradication of the population in deme 1). For these generations, the population in deme 2, the non-target population, experiences the burden of the gene drive at frequency $\hat{q}_2$ of the DTE. For a given configuration (s, c, h) of the gene drive, we define the maximum equilibrium frequency of the DTE in the non-target population for migration rates below m^* as $q_2^*(s, c, h)$. In other words, $q_2^*(s, c, h) = \sup_{m<m*} \hat{q}_2(s, c, h, m)$, where $\hat{q}_2(s, c, h, m)$ corresponds to the frequency of the DTE of a gene drive with configuration (s, c, h) under migration rate m.

We numerically computed q_2^* across the parameter range U (Fig. 4). In general, we find that q_2^* is positively correlated with m^* (Fig. 3 and Fig. 4), meaning that gene drive configurations that can sustain differential targeting for higher migration rates also potentially result in higher frequencies of the gene drives in the non-target population. The q_2^* values show qualitatively similar patterns for the two models we examined, although q_2^* values were slightly lower in the m-before-s model than in the s-before-m model. Considering all possible gene drive configurations in U, q_2^* was maximal for $c = 1$ and s^*, the same configuration that maximizes m^*, in both models (and any value of h). This maximal allele frequency of A in the non-target population was $q_2^* = 0.28$ for the m-before-s model and $q_2^* = 0.34$ for the s-before-m model. These configurations and q_2^* values delineate the conditions and strength for differential-targeting conditions impacting the non-target population most strongly in terms of frequency of the gene drive allele. Notably, if $c = 1$, q_2^* falls

sharply for $s > s^*$ and less so for $s < s^*$ (Fig. 4).

4.2 Exceeding of the critical migration rate m^*

The limit with which the possibility for differential targeting can withstand migration is described by m^*. However, we must also consider the consequences of application of the gene drives when migration exceeds m^*, because migration may unexpectedly change over time and may be difficult to measure. For a given gene drive configuration, when $m > m^*$, the two-deme system is expected to rapidly converge either to global loss or to global fixation of the gene drive allele. A gene drive configuration that converges to global loss if m^* is exceeded will result in a failure of the gene drive application. This case would, in general, be preferable to a gene drive that converges to global fixation when migration exceeds m^*, as this case would likely result in severe consequences for the non-target population.

In order to explore the outcomes for exceeding the threshold m^*, we assume that when m changes, the system converges rapidly to its equilibrium state, and we ignore the transient dynamics of changes between stable states. Under this assumption, we consider the continuous changes in the stable equilibrium point in frequency space due to continuous changes in migration rates. We can track the trajectory of the DTE $(\hat{q}_1, \hat{q}_2)$ with the increase of m from 0 to m^* for each gene drive configuration (s, c, h) (Fig. 5). To determine the consequence of a breach of m^*, we continue this trajectory above m^* with an instantaneous transition to either global fixation or global loss. To determine the state to which the trajectory will transition, we determined whether $(\hat{q}_1, \hat{q}_2)$, computed for the DTE with $m = m^* - \epsilon$, lies in the basin of attraction of global loss or of global fixation for basins computed for $m = m^* + \epsilon$ (Fig. 5). For example, in Figure 2, the yellow DTE for $m = 0.022$, just below m^*, is positioned in the red global-fixation basin of attraction for the scenario $m = 0.023$, just above m^*. We therefore conclude that, for the configuration in Figure 2, the gene drive would converge to global fixation if m^* were exceeded.

We computed trajectories of DTEs with increase in migration rates for $c = 1$ and different selection coefficients s (Fig. 5). For both models, we see that for selection coefficients lower than s^*, exceeding m^* results in global fixation, whereas for selection coefficients higher than s^*, exceeding the threshold m^* results in global loss (Fig. 5). In other words, for $c = 1$ and $s < s^*$, exceeding the m^* threshold results in global fixation of the gene drive allele, whereas for $s > s^*$, a breach results in global loss of the allele. This sharp transition, and the severity of the consequences of global fixation, suggest that a safer configuration of gene drive would be one with s somewhat higher than s^*, since maintaining a safe distance from this sharp transition would be prudent, although at a cost of producing a decrease in m^*. The qualitative behavior of the system with respect to breaching m^* is similar for $c < 1$, but the threshold selection coefficient s that distinguishes global fixation from global loss is different.

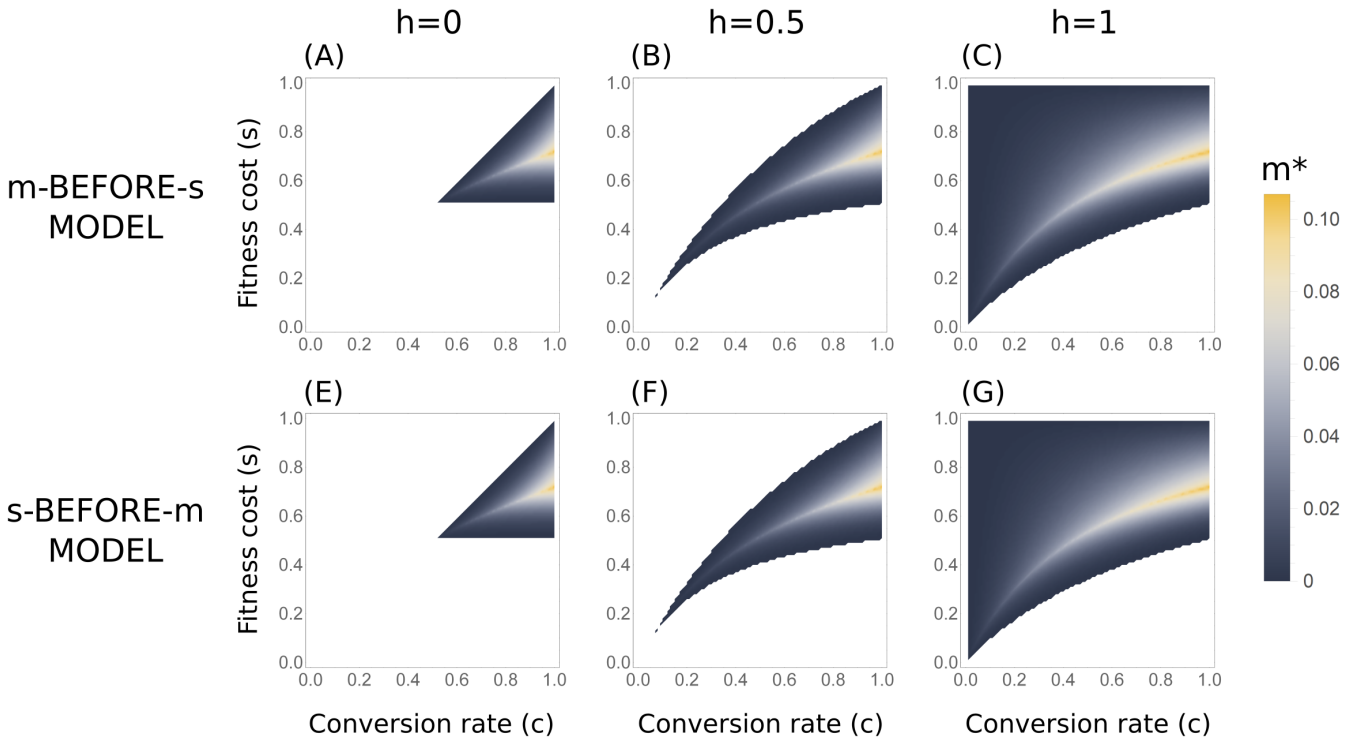

Figure 3: **Maximal migration rates m^* for which differential-targeting equilibria (DTEs) exist.** m-before-s model (Eq. 3) in top panels, s-before-m model (Eq. 4) in bottom panels. The colored regions denote the configurations for which DTEs exist (U configurations), and the regions in white denote configurations for which differential targeting of the demes is not possible. Over most of the parameter space, DTEs exist only for low migration rates (in blue), and only in a narrow band do DTEs exist with high migration (light yellow). In this band, m^* values are higher for higher conversion rates, and are they are maximal, over the entire parameter space, for $c = 1$, $s \approx 0.7197$, and any h value.

4.3 Perturbations from the DTEs

So far, we have evaluated gene drive spillovers by considering the equilibria to which the dynamics deterministically converge. However, we have not considered the possibility that stochastic events and processes might destabilize the system, allowing it to escape the DTE due to perturbations, and to reach a different stable state. In this section, we consider two types of perturbations: (1) perturbations due to genetic drift, and (2) perturbations due to an external event, such as a significant ecological disturbance, that can affect the allele frequencies of the demes.

The impact of genetic drift on the probability of escaping a DTE depends on the effective population size of demes, N_e. We estimated the *probability of escape* — the probability of leaving the basin of attraction of the DTE in 100 generations, starting at the DTE — by simulating the dynamics on a model similar to Eq. 3 and Eq. 4, but with allele frequencies also affected by genetic drift (Appendix B).

To examine the impact of genetic, we examined the scenario of $N_e = 100$ and $c = 1$, as an example (h is of no significance since there are no heterozygotes for $c = 1$). In general, for

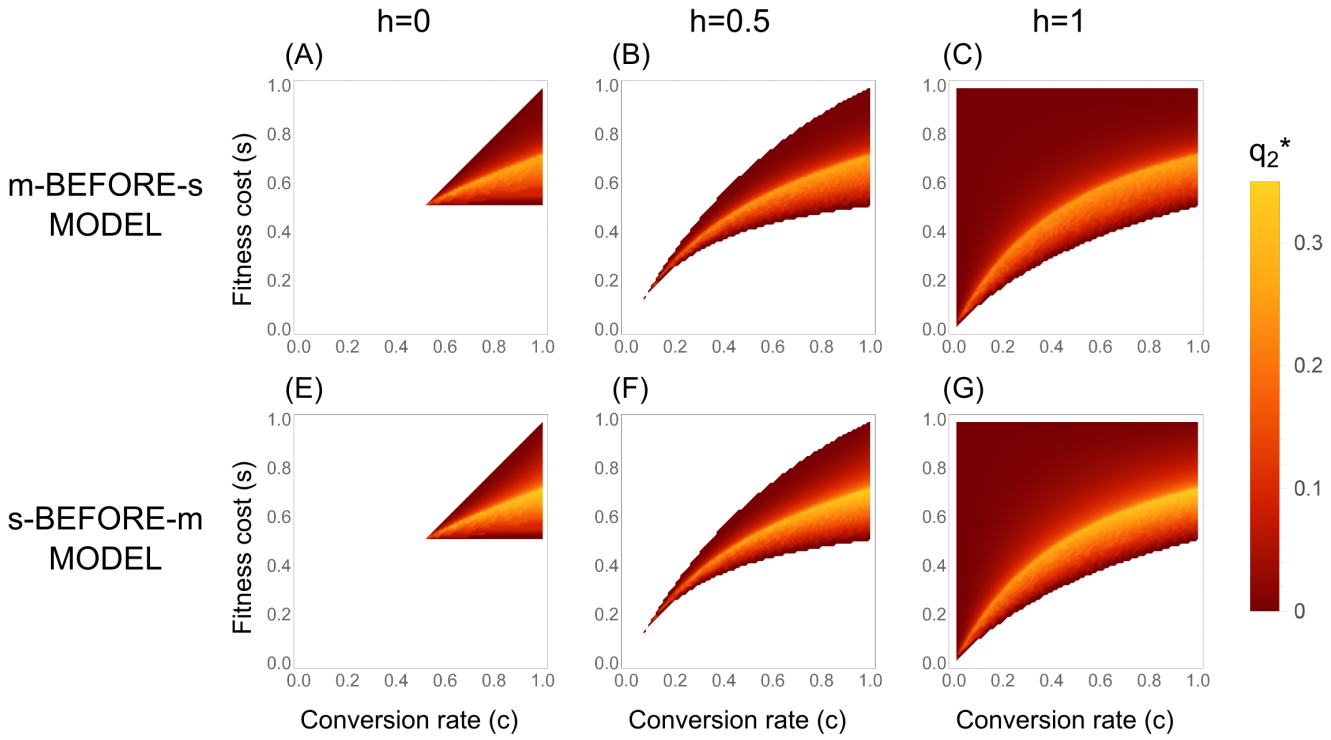

Figure 4: **Maximal frequencies in the non-target population q_2^* at differential-targeting equilibria (DTEs).** m-before-s model (Eq. 3) in top panels, s-before-m model (Eq. 4) in bottom panels. Comparing to Fig. 3, q_2^* values are correlated with m^* values.

this scenario, the probability of escape remains below 5% until m^* is approached (Fig. 6A and Fig. 6C). For larger populations with $N_e > 100$, the probabilities of escape are expected to be lower than those in Fig. 6 due to weaker genetic drift, whereas for populations with $N_e < 100$, the probabilities of escape are expected to be higher.

Genetic drift is a source of perturbation that is internal to the system, but we also consider external perturbations that alter allele frequencies in the demes. In this case, we evaluate the resilience to perturbations by examining the consequence of a single perturbation to the eventual DTE of the system. We define the safety radius of a gene drive configuration under a given migration rate to be the maximal magnitude of a perturbation from a DTE, in the allele frequency space, above which the system will not converge back to the DTE. In other words, the safety radius is defined as the shortest distance from the DTE to the boundary of its basin of attraction. For perturbations from the DTE larger than the stability radius, the system is expected to converge to a non-DTE. We show the stability radii of different gene drive configurations in Figure 6B. Similar to the the probability of escape from the DTE by genetic drift, the stability radii are particularly small when migration rates approach m^*.

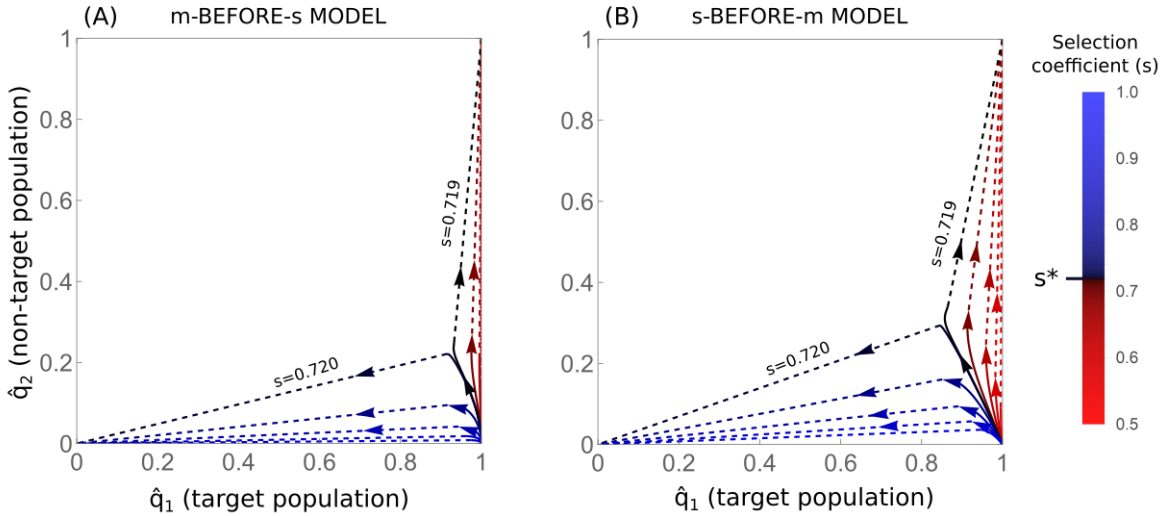

Figure 5: **Trajectories of DTEs with increasing migration rates.** Shown are trajectories for DTEs for gene drives with full conversion ($c = 1$) and different selection coefficients (s), marked in different colors. The arrows show the direction of increased migration; continuous lines show increase of migration below m^*, and the dashed lines show the abrupt transition of the equilibrium state when m^* is exceeded, to either global fixation or global loss. The color bars for s are centered such that the selection coefficient maximizing m^*, $s^* \approx 0.7197$, is shown in black, selection coefficients below s^* are in red, and those above s^* are in blue. In both models, s^* forms the threshold between convergence to global fixation or to global loss when the threshold m^* is exceeded.

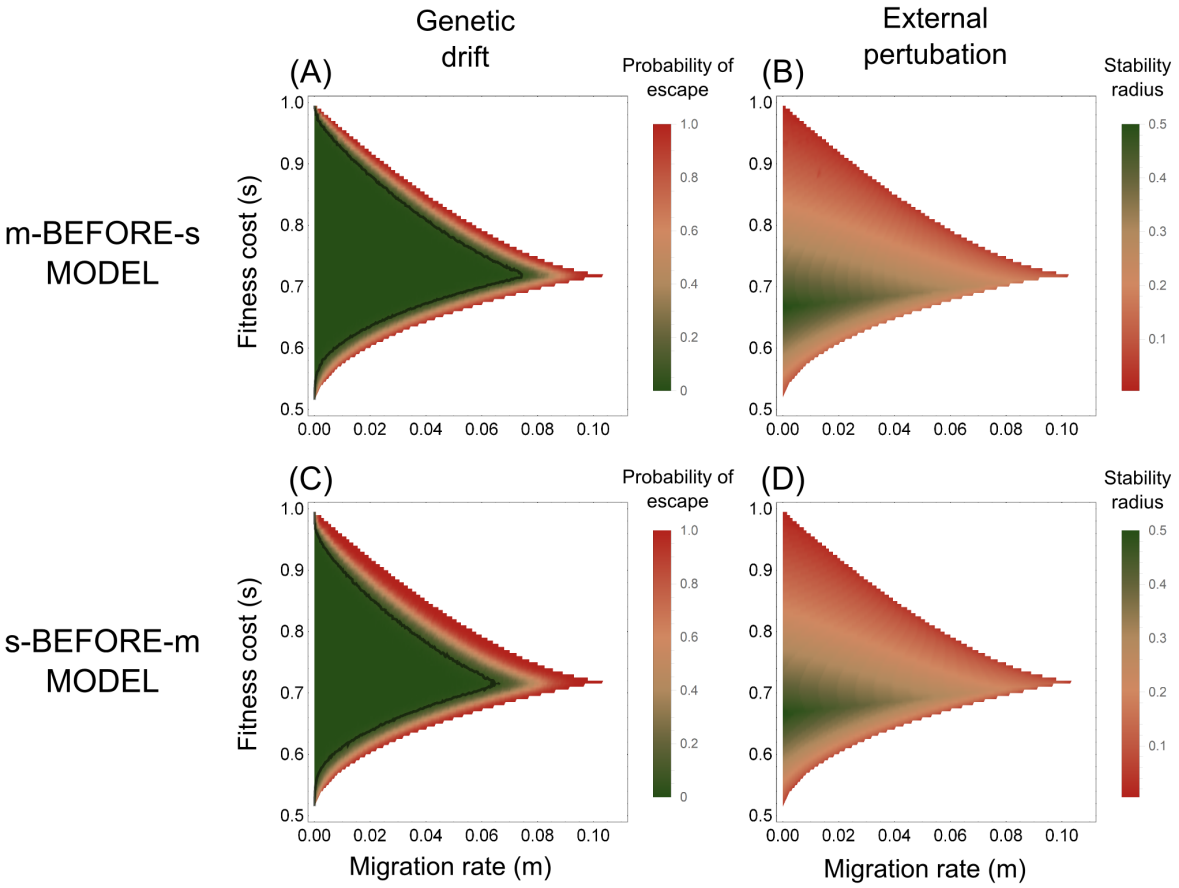

Figure 6: **Perturbation from DTEs.** Results are shown for gene drives with full conversion, $c = 1$, and consequently any h value. (A) The probability of escape from the DTE for different migration rates m and selection coefficients s due to genetic drift, for the m-before-s model. The probability is defined as the probability of departing from the attraction basin of the DTE over 100 generations with genetic drift, in a population with $N_e = 100$. The black line denotes 5% probability of escape. (B) The safety radii of the DTE for the m-before-s model. (C) The probability of escape for the s-before-m model. (D) The safety radii of the DTE for the s-before-m model. In both (A) and (B), the white regions denote scenarios for which $m > m^*$.

5 Discussion

We have explored the possibility of effectively applying a gene drive in a target population while limiting the exposure of a non-target population to spillovers. By investigating equilibria of the expected evolutionary dynamics, we show that for some gene drive configurations, those for which a polymorphic unstable equilibrium exists in the one-deme model (U configurations), it is possible to initiate the gene drive such that differential targeting — higher equilibrium gene drive allele frequencies in the target population than in the non-target population — is possible. However, we also show that for these configurations, upon increasing the migration rate m, a sharp transition occurs to a state in which the differential-targeting equilibrium no longer exists, and the two populations face similar fates, either global loss or global fixation of the gene drive allele. We numerically identify an "optimal" gene drive configuration, with $c = 1$ and $s^* \approx 0.7197$, as the configuration that allows for differential targeting under the highest migration rate, $m^* \approx 0.110$.

For most U configurations, other than in from a narrow region of the parameter space, differential targeting can be achieved only for low migration rates (Fig. 3). Consequently, for relatively high migration rates, it is unlikely that gene drives could be configured with enough accuracy to provide sufficient confidence that differential targeting could be achieved. In general, gene drive spillovers to non-target population should therefore be considered likely, and prevention of spillovers should not rely on differential targeting. Only when the gene drive configuration has been validated with high accuracy and confidence, or when low migration rates between target and non-target populations can be assured, will it be possible to view differential targeting as a practical measure for limiting spillovers.

We found that the configuration that maximizes the critical migration threshold m^* ($s^* \approx 0.7197$, $c = 1$, and any h value) is also the one that maximizes the impact on the non-target population (Fig. 4). This same configuration also differentiates between scenarios for which increasing migration above m^* results in global loss or global fixation of the gene drive allele (Fig. 5). Such breaches may occur as a result of unexpected increases in migration, or from errors in approximation of actual migration rates. As a precaution, it would be wise to design gene drives for which exceeding of m^* results in global loss rather than the alternative, so as to avoid full exposure of the non-target population to the gene drive.

In this regard, considering practical difficulties of accurately specifying gene drive parameters such as s, c, and h, it would perhaps be better to avoid gene drive specifications that are too close to this sharp transition of s^*, and to aim for s values that are somewhat higher than s^*. This strategy would also result in lower impact on the non-target population in terms of gene drive allele frequencies (q_2^*), compared to values that are slightly lower than s^* (Fig. 4), albeit at the cost of decreasing m^* (Fig. 3). Although some minor differences exist between the outcomes for the two models we examined, in general, the order of migration and selection have little affect on the properties studied here.

It is important to note that the migration rates m, which are defined as migration probabilities

of individuals, do not necessarily refer to the mean migration probabilities of individuals across the entire target or non-target population, but rather could refer to smaller subpopulations in the region from which migration occurs. For example, considering a target island with an invasive rodent species and a non-target island with native rodents, and that the rodents migrate mostly via major ports on these islands, the subpopulations of the port and its immediate surroundings would probably be the appropriate unit for estimating m and N_e, and not the populations of the entire islands. This is because if a gene drive is established in a subpopulation within the non-target island, it would be very difficult to stop further expansion of the gene drive in that island. Within-population migration rates are typically much higher than migration rates between populations, and the established gene drive will likely spread further. Therefore, for migration between two highly-populated islands, for example, if we measure m across the entire island population, and not among appropriate subpopulations, then we might be significantly underestimating m for the purpose of our models.

Many challenges must be overcome before CRISPR-based gene drives can be efficiently and safely applied, such as the potential evolution of resistance to the gene drive modification [38–44]. For the type of differential targeting described here, accurate configuration of gene drive parameters, such as s, c, and h, is required. However, it is unclear whether such accuracy and arbitrary choice of parameters can be feasibly obtained. For example, to date, conversion rates (c) above 90% have been demonstrated only in some invertebrates [9, 45–48] and yeast [49], with much lower conversion efficiencies occurring in mammals [50]. Another difficulty for differential targeting is that implementation might require a large concentrated effort to initiate the gene drive in the basin of attraction of the desired equilibrium (the DTE). This effort could require high gene drive allele frequencies, and hence it might involve engineering and releasing many engineered individuals, especially for high migration rates close to m^*. A third problem is that we do not yet understand the consequences of more complex population structure, involving more than two populations or populations that are internally structured. Given these difficulties, to understand and mitigate the dangers of spillovers, it would be important to study more elaborate population structures in models of gene drives, as well as to explore additional biosafety mechanisms, such as daisy-chain gene drives [51] or other molecular safe-guarding techniques [37, 52–56].

CRISPR-based gene-drive technology is potentially a powerful tool for species control, but it is also extremely dangerous. The conclusions of this paper should not be read as a call for applications of gene drives in wild settings, or as a blueprint for CRISPR-based gene drive design. Instead, we believe the consequences of gene drive application and the potential for spillovers should be further discussed in the scientific community and in society at large before actual gene drive applications are realized, and a goal of this paper is to facilitate and inform these discussions.

6 Methods

6.1 Determining equilibria and basins of attraction

For a given model (m-before-s or s-before-m), gene drive configuration (s, c, h), and migration rate m, the equilibrium points were computed by solving numerically the systems of equations Eq. 3 or 4, with the equilibrium conditions $q_1' = q_1$ and $q_2' = q_2$, using Mathematica software [57]. For each equilibrium point, $(\hat{q}_1, \hat{q}_2)$, the stability was determined using stability analysis, described in the following section. To generate Fig. 2 and Figs. S1–S6, this was done for all triple (s, c, h) combinations, with s ranging from 0 to 1 in increments of 0.01, c ranging from 0 to 1 in increments of 0.01, and $h = 0, 0.5$ or 1. To generate Fig. 3 and Fig. 4, these configurations were each computed for migration rates m between 0 and 0.15 at increments of 0.0001.

For a stable equilibrium, there exists a region in allele frequency space around the equilibrium, called the *basin of attraction* [58], such that any trajectory starting within it will eventually converge to the stable point. To determine basins of attraction for a configuration (s, c, h) and migration rate m, we start from each point (q_1, q_2) in the frequency space, and track its trajectory for 1000 generations by repeatedly using Eq. 3 or 4. We denote the end point of the trajectory with initial condition (q_1, q_2) by $(q_1, q_2)_{1000}$. The attractor associated with (q_1, q_2) was defined as the equilibrium point $(\hat{q}_1, \hat{q}_2)$ of Eq. 3 or 4 that was at a distance of less than 0.001 in frequency space from $(q_1, q_2)_{1000}$. If there were no such equilibria or more than one such equilibrium, then the attractor of (q_1, q_2) was left undefined. The basins of attraction were computed for a grid of points in frequency space at a resolution of 0.01×0.01. For all computations used to generate Figs. 2, 5, and 6, attractors were defined for all points computed.

6.2 Stability analysis

We determine the stability of an equilibrium by examining how the trajectories of the allele frequencies in the two populations behaved in its vicinity. For a given gene drive configuration, (s, c, h), migration rates, m, and model (m-before-s or s-before-m), the equilibria of the system were found by solving the system of discrete-time recursion equations that describe the change of allele frequencies over a single generation:

$$\begin{aligned} q_1' &= f_1(q_1, q_2) \\ q_2' &= f_2(q_1, q_2), \end{aligned} \tag{5}$$

where f_i (for $i = 1, 2$) is defined in Eq. 3 or 4, depending on the model.

To analyze the stability of the system in the vicinity of an equilibrium point $(\hat{q}_1, \hat{q}_2)$, we linearized the nonlinear functions f_1 and f_2 by taking the first-order approximation of the Taylor expansion about the equilibrium. From this linearization, in order to understand the consequences of small perturbations from the equilibrium, we examined the matrix of the partial derivatives of f_1 and f_2, evaluated at the equilibrium point (i.e., the Jacobian matrix of Eq. 5):

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial q_1} & \frac{\partial f_1}{\partial q_2} \\ \frac{\partial f_2}{\partial q_1} & \frac{\partial f_2}{\partial q_2} \end{bmatrix}. \tag{6}$$

The stability type of the equilibrium point $(\hat{q}_1, \hat{q}_2)$ was determined by examining the two eigenvalues λ_1 and λ_2 of $\mathbf{J}$ evaluated at $(\hat{q}_1, \hat{q}_2)$. If both eigenvalues have modulus less than 1 ($|\lambda_1| < 1$ and $|\lambda_2| < 1$), then the equilibrium point $(\hat{q}_1, \hat{q}_2)$ is stable. If at least one eigenvalue has modulus greater than 1, then the equilibrium point is unstable [59].

6.3 Perturbation analysis

To determine the probability of escape from DTEs, we formulated the models as a stochastic process in which the expected frequency of the gene drive allele A after one generation is the same as in Eq. 3 and Eq. 4, but the actual frequency is binomially distributed with $2N_e$ trials, following the Wright-Fisher model for diploid populations, for each population. Genetic drift was implemented after both selection and migration, in both models. Full details appear in Appendix B.

For each gene drive configuration and migration rate, the equilibria and basins of attraction were determined. Next, for each gene drive configuration and migration rate, 1000 simulations were initiated, starting at the DTE $(\hat{q}_1, \hat{q}_2)$, and were run for 100 generations by repeatedly applying Eq. S1 or S2. In each simulation, it was determined whether or not the allele frequencies remained in the basin of attraction of the DTE for the 100 generations. The probability of escape was computed as the fraction of the simulations for which the allele frequencies did not remain in the attraction basin of the DTE. Trajectories that escaped the basin of attraction but then returned to the basin of attraction, within the 100 generations, were considered as escaped trajectories, so as to provide a conservative estimate of probability of escape.

In order to compute the stability radius for a given model, gene drive configuration, and migration rate, first the equilibria and basins of attraction were computed. The distance, in frequency space, between the DTE and all points in its basin of attraction were then computed. The stability radius was defined as the supremum over distances, in frequency space, between the DTE and any point in its basin of attraction. In other words, for a DTE $(\hat{q}_1, \hat{q}_2)$ with its basin of attraction $B_{(\hat{q}_1,\hat{q}_2)}$, the stability radius was defined as $\sup_{(q_1,q_2)\in B_{(\hat{q}_1,\hat{q}_2)}} \sqrt{(\hat{q}_1 - q_1)^2 + (\hat{q}_2 - q_2)^2}$. In practice, the stability radius was determined by computing the maximal distance between the DTE and any point within its basin of attraction, for a grid of points in the two-dimensional frequency space at a resolution of 0.01×0.01.

References

1. Hartl, D. L. Analysis of a general population genetic model of meiotic drive. *Evolution* **24,** 538–545 (1970).
2. Thomson, G. J. & Feldman, M. W. Population genetics of modifiers of meiotic drive. II linkage modification in the segregation distortion system. *Theoretical Population Biology* **5,** 155–162 (1974).
3. Thomson, G. J. & Feldman, M. W. Population genetics of modifiers of meiotic drive: IV. On the evolution of sex-ratio distortion. *Theoretical Population Biology* **8,** 202–211 (1975).
4. Hartl, D. L. Modifier theory and meiotic drive. *Theoretical Population Biology* **7,** 168–174 (1975).
5. Thomson, G. J. & Feldman, M. W. Population genetics of modifiers of meiotic drive III. Equilibrium analysis of a general model for the genetic control of segregation distortion. *Theoretical Population Biology* **10,** 10–25 (1976).
6. Charlesworth, B. & Hartl, D. L. Population Dynamics of the Segregation Distorter Polymorphism of *Drosophila melanogaster*. *Genetics* **89,** 171–92 (1978).
7. Marshall, J. M. & Hay, B. A. Confinement of gene drive systems to local populations: A comparative analysis. *Journal of Theoretical Biology* **294,** 153–171 (2012).
8. Unckless, R. L., Messer, P. W., Connallon, T. & Clark, A. G. Modeling the manipulation of natural populations by the mutagenic Chain reaction. *Genetics* **201,** 425–431 (2015).
9. Gantz, V. M. & Bier, E. The mutagenic chain reaction: A method for converting heterozygous to homozygous mutations. *Science* **348,** 442–444 (2015).
10. Oye, K. A. *et al.* Regulating gene drives. *Science* **345,** 626–628 (2014).
11. Champer, J., Buchman, A. & Akbari, O. S. Cheating evolution: Engineering gene drives to manipulate the fate of wild populations. *Nature Reviews Genetics* **17,** 146–159 (2016).
12. Piaggio, A. J. *et al.* Is It Time for Synthetic Biodiversity Conservation? *Trends in Ecology and Evolution* **32,** 97–107 (2017).
13. Prowse, T. A. *et al.* Dodging silver bullets: Good CRISPR gene-drive design is critical for eradicating exotic vertebrates. *Proceedings of the Royal Society B: Biological Sciences* **284,** 20182048 (2017).
14. Webber, B. L., Raghu, S. & Edwards, O. R. Opinion: Is CRISPR-based gene drive a biocontrol silver bullet or global conservation threat?: Fig. 1. *Proceedings of the National Academy of Sciences* **112,** 10565–10567 (2015).
15. Esvelt, K. M. & Gemmell, N. J. Conservation demands safe gene drive. *PLoS Biology* **15,** 1–8 (2017).

16. Noble, C., Adlam, B., Church, G. M., Esvelt, K. M. & Nowak, M. A. Current CRISPR gene drive systems are likely to be highly invasive in wild populations. *eLife* **7,** e33423 (2018).

17. Rode, N. O., Estoup, A., Bourguet, D., Courtier-Orgogozo, V. & Débarre, F. Population management using gene drive: molecular design, models of spread dynamics and assessment of ecological risks. *Conservation Genetics* (2019).

18. Owens, B. The big cull: Can New Zealand pull off an audacious plan to get rid of invasive predators by 2050? *Nature* **541,** 148–150 (2017).

19. Dearden, P. K. *et al.* The potential for the use of gene drives for pest control in New Zealand: a perspective. *Journal of the Royal Society of New Zealand* **48,** 225–244 (2018).

20. Mendel, G. Versuche ber pflanzen-hybriden. *Verhandlungen des Naturforschenden Vereines in Brunn* (1865).

21. Sandler, L & Novitski, E. Meiotic drive as an evolutionary force. *The American Naturalist* **91,** 105–110 (1957).

22. Lewontin, R. C. & Dunn, L. C. The evolutionary dynamics of a polymorphism in the house mouse. *Genetics* **45,** 705 (1960).

23. Lewontin, R. C. Interdeme selection controlling a polymorphism in the house mouse. *The American Naturalist* **96,** 65–78 (1962).

24. Karlin, S. & McGregor, J. *Equilibria for genetic systems with weak interaction* in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Biology and Health* (Berkeley, CA, 1972), 79–87.

25. Prout, T., Bundgaard, J. & Bryant, S. Population genetics of modifiers of meiotic drive I. The solution of a special case and some general implications. *Theoretical Population Biology* **4,** 446–465 (1973).

26. Curtsinger, J. W. & Feldman, M. W. Experimental and theoretical analysis of the 'sex-ratio' polymorphism in *Drosophila pseudoobscura*. *Genetics* **94,** 445–466 (1980).

27. Hiraizumi, Y., Sandler, L & Crow, J. F. Meiotic drive in natural populations of *Drosophila melanogaster*. III. Populational implications of the segregation-distorter locus. *Evolution* **14,** 433–444 (1960).

28. Davis, S., Bax, N. & Grewe, P. Engineered underdominance allows efficient and economical introgression of traits into pest populations. *Journal of Theoretical Biology* **212,** 83–98 (2001).

29. Magori, K. & Gould, F. Genetically engineered underdominance for manipulation of pest populations: A deterministic model. *Genetics* **172,** 2613–2620 (2006).

30. Altrock, P. M., Traulsen, A., Reeves, R. G. & Reed, F. A. Using underdominance to bi-stably transform local populations. *Journal of Theoretical Biology* **267,** 62–75 (2010).

31. Tanaka, H., Stone, H. A. & Nelson, D. R. Spatial gene drives and pushed genetic waves. *Proceedings of the National Academy of Science* **114,** 8752–8457 (2017).

32. Karlin, S. *et al.* Application of method of small parameters to multi-niche population genetic models. *Theoretical Population Biology* **3,** 186–209 (1972).

33. Karlin, S. & McGregor, J. Polymorphisms for genetic and ecological systems with weak coupling. *Theoretical Population Biology* **3,** 210–238 (1972).

34. Huang, Y., Lloyd, A. L., Legros, M. & Gould, F. Gene-drive into insect populations with age and spatial structure: a theoretical assessment. *Evolutionary Applications* **4,** 415–428 (2011).

35. Marshall, J. M. & Hay, B. A. Inverse Medea as a novel gene drive system for local population replacement: A theoretical analysis. *Journal of Heredity* **102,** 336–341 (2011).

36. Akbari, O. S. *et al.* A synthetic gene drive system for local, reversible modification and suppression of insect populations. *Current Biology* **23,** 671–677 (2013).

37. Dhole, S., Vella, M. R., Lloyd, A. L. & Gould, F. Invasion and migration of spatially self-limiting gene drives: A comparative analysis. *Evolutionary Applications* **11,** 794–808 (2018).

38. Unckless, R. L., Clark, A. G. & Messer, P. W. Evolution of resistance against CRISPR/Cas9 gene drive. *Genetics* **205,** 827–841 (2017).

39. Church, G. M., Noble, C., Nowak, M. A., Esvelt, K. M. & Olejarz, J. Evolutionary dynamics of CRISPR gene drives. *Science Advances* **3,** e1601964 (2017).

40. Champer, J. *et al.* Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations. *PLoS Genetics* **13,** 1–18 (2017).

41. Gemmell, N. J. & Tompkins, D. M. Gene Drives and Rodent Control: Response to Piaggio et al. *Trends in Ecology and Evolution* **32,** 314–315 (2017).

42. Zentner, G. E., Wade, M. J., Drury, D. W., Siniard, D. J. & Dapper, A. L. CRISPR/Cas9 gene drives in genetically variable and nonrandomly mating wild populations. *Science Advances* **3,** e1601910 (2017).

43. Marshall, J. M., Buchman, A., Sánchez, C. H. M. & Akbari, O. S. Overcoming evolved resistance to population-suppressing homing-based gene drives. *Scientific Reports* **7,** 3776. ISSN: 20452322 (2017).

44. Champer, J. *et al.* Reducing resistance allele formation in CRISPR gene drive. *Proceedings of the National Academy of Sciences* **115,** 5522–5527 (2018).

45. Gantz, V. M. *et al.* Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proceedings of the National Academy of Sciences* **112,** E6736–E6743 (2015).

46. Hammond, A. *et al.* A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nature Biotechnology* **34,** 78–83 (2016).

47. Buchman, A., Marshall, J. M., Ostrovski, D., Yang, T. & Akbari, O. S. Synthetically engineered Medea gene drive system in the worldwide crop pest *Drosophila suzukii*. *Proceedings of the National Academy of Sciences* **18,** 4725–4730 (2018).

48. Kyrou, K. *et al.* A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nature Biotechnology,* 1062–1066 (2018).

49. Shapiro, R. S. *et al.* A CRISPR-Cas9-based gene drive platform for genetic interaction analysis in *Candida albicans*. *Nature Microbiology* **3,** 73–82 (2017).

50. Grunwald, H. A. *et al.* CRISPR Cas9 in the female mouse germline. *Nature* **566,** 105–109 (2019).

51. Noble, C. *et al.* Daisy-chain gene drives for the alteration of local populations. *Proceedings of the National Academy of Science* **116,** 8275–8282 (2019).

52. Akbari, O. S. *et al.* Safeguarding gene drive experiments in the laboratory. *Science* **349,** 927–929 (2015).

53. DiCarlo, J. E., Chavez, A., Dietz, S. L., Esvelt, K. M. & Church, G. M. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nature Biotechnology* **33,** 1250–1255 (2015).

54. Vella, M. R., Gunning, C. E., Lloyd, A. L. & Gould, F. Evaluating strategies for reversing CRISPR-Cas9 gene drives. *Scientific Reports* **7,** 11038 (2017).

55. Marshall, J. M. & Akbari, O. S. Can CRISPR-based gene drive be confined in the wild? A question for molecular and population biology. *ACS Chemical Biology* **13,** 424–430 (2018).

56. Champer, J. *et al.* Molecular safeguarding of CRISPR gene drive experiments. *eLife* **8,** 1–10 (2019).

57. Wolfram Research, I. & Wolfram Research. Mathematica, Version 11.3 (2018).

58. Elaydi, S. *Discrete Chaos: With Applications in Science and Engineering* 2nd editio (Chapman and Hall/CRC, 2007).

59. Galor, O. in *Discrete Dynamical Systems* 93–105 (Springer, 2007).

Appendix A - Characterization of equilibria

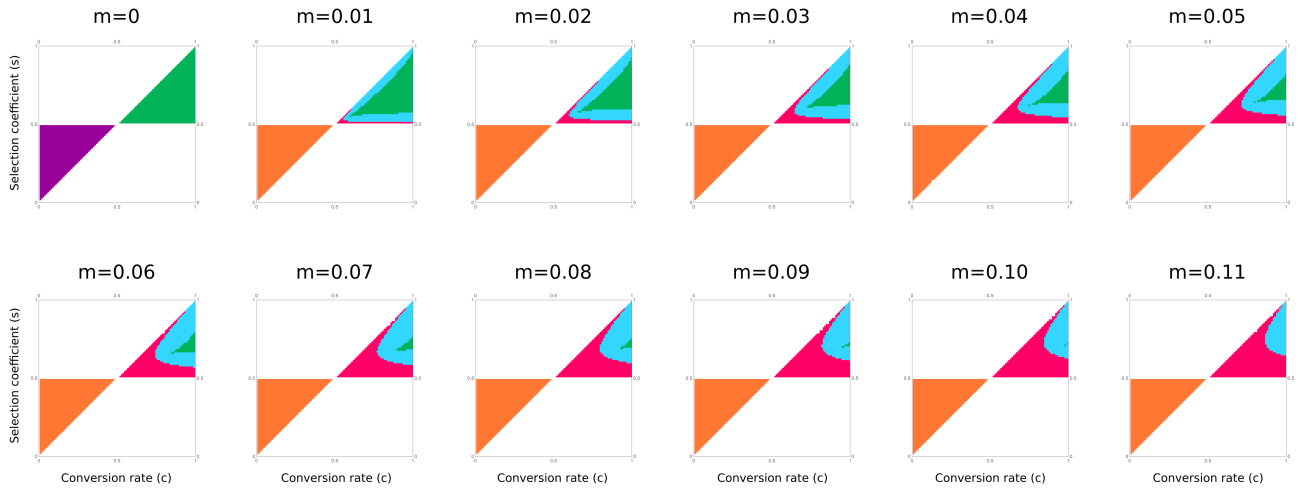

Figure S1: **Partitioning of the parameter space by the characterization of the equilibria for the m-before-s model (Eq. 3), for recessive gene drives ($h = 0$), under different migration rates.** Green — 9 equilibria, four of which are stable, and one of which is a DTE (stable and $\hat{q}_1 > \hat{q}_2$); Purple — 9 equilibria, one of which is stable and symmetric ($\hat{q}_1 = \hat{q}_2$); Blue — 5 equilibria, 2 stable trivial (global fixation and global loss) and 3 unstable; Red — 3 equilibria, 2 stable trivial and one unstable; Orange – 3 equilibria, 2 unstable trivial and one stable symmetric ; White — 2 trivial equilibria, one stable and one unstable. In the bottom white region, the stable equilibrium is global fixation of the gene drive allele, and in the top white region, the stable equilibrium is global loss. The only gene drive configurations for which differential targeting is possible appear in green.

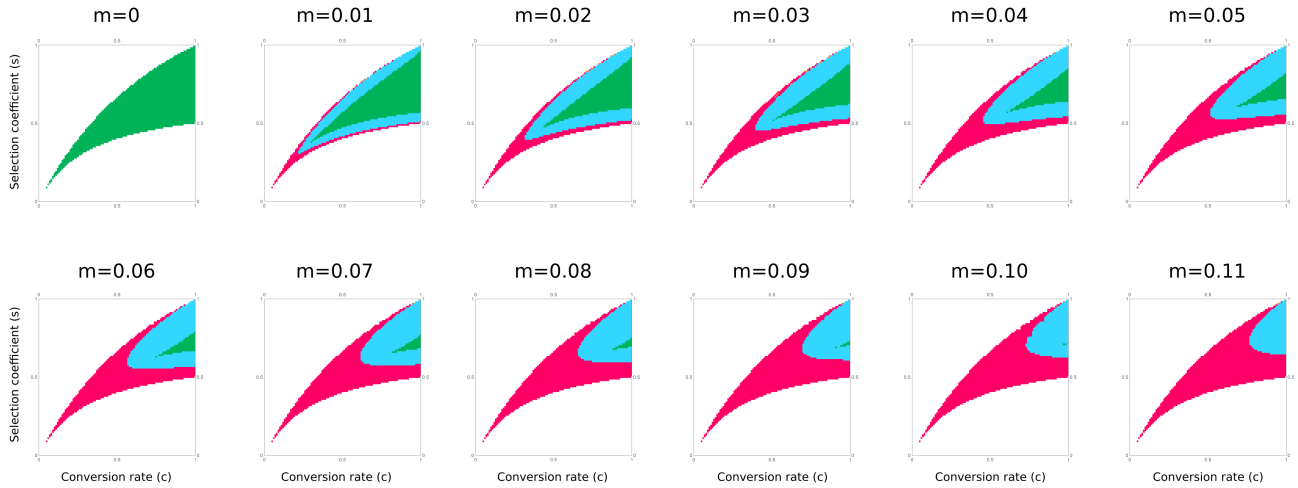

Figure S2: **Partitioning of the parameter space by the characterization of the equilibria for the m-before-s model (Eq. 3), for additive gene drives ($h = 0.5$), under different migration rates.** Green — 9 equilibria, four of which are stable, and one of which is a DTE (stable and $\hat{q}_1 > \hat{q}_2$); Purple — 9 equilibria, one of which is stable and symmetric ($\hat{q}_1 = \hat{q}_2$); Blue — 5 equilibria, 2 stable trivial (global fixation and global loss) and 3 unstable; Red — 3 equilibria, 2 stable trivial and one unstable; White — 2 trivial equilibria, one stable and one unstable. In the bottom white region, the stable equilibrium is global fixation of the gene drive allele, and in the top white region, the stable equilibrium is global loss. The only gene drive configurations for which differential targeting is possible appear in green.

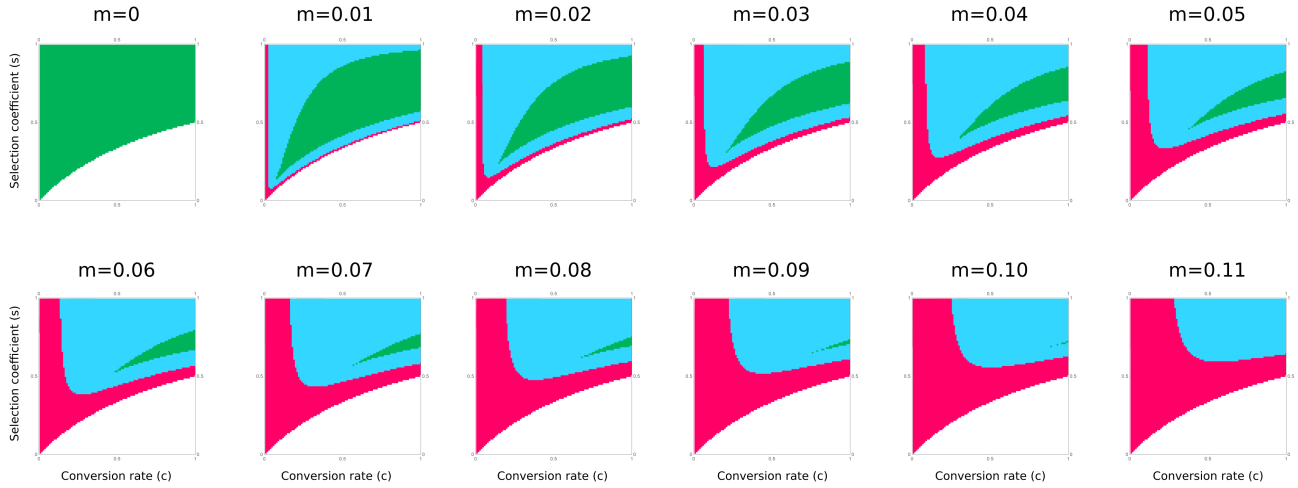

Figure S3: **Partitioning of the parameter space by the characterization of the equilibria for the m-before-s model (Eq. 3), for dominant gene drives ($h = 1$), under different migration rates.** Green — 9 equilibria, four of which are stable, and one of which is a DTE (stable and $\hat{q}_1 > \hat{q}_2$); Purple — 9 equilibria, one of which is stable and symmetric ($\hat{q}_1 = \hat{q}_2$); Blue — 5 equilibria, 2 stable trivial (global fixation and global loss) and 3 unstable; Red — 3 equilibria, 2 stable trivial and one unstable; White — 2 trivial equilibria, one stable and one unstable. In the bottom white region, the stable equilibrium is global fixation of the gene drive allele. The only gene drive configurations for which differential targeting is possible appear in green.

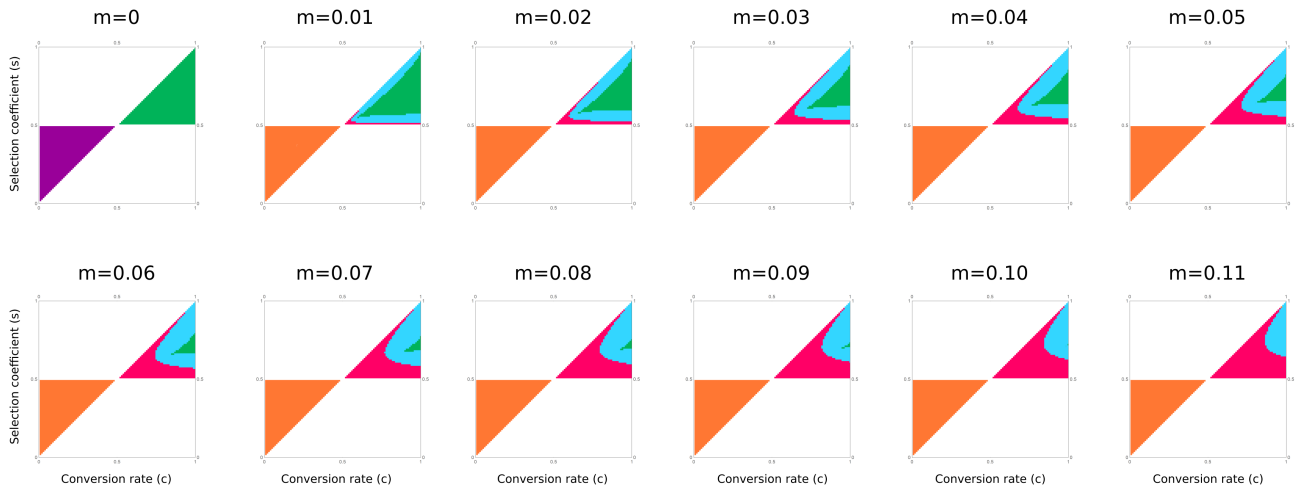

Figure S4: **Partitioning of the parameter space by the characterization of the equilibria for the s-before-m model (Eq. 3), for recessive gene drives ($h = 0$), under different migration rates.** Green — 9 equilibria, four of which are stable, and one of which is a DTE (stable and $\hat{q}_1 > \hat{q}_2$); Purple — 9 equilibria, one of which is stable and symmetric ($\hat{q}_1 = \hat{q}_2$); Blue — 5 equilibria, 2 stable trivial (global fixation and global loss) and 3 unstable; Red — 3 equilibria, 2 stable trivial and one unstable; Orange – 3 equilibria, 2 unstable trivial and one stable symmetric ; White — 2 trivial equilibria, one stable and one unstable. In the bottom white region, the stable equilibrium is global fixation of the gene drive allele, and in the top white region, the stable equilibrium is global loss. The only gene drive configurations for which differential targeting is possible appear in green.

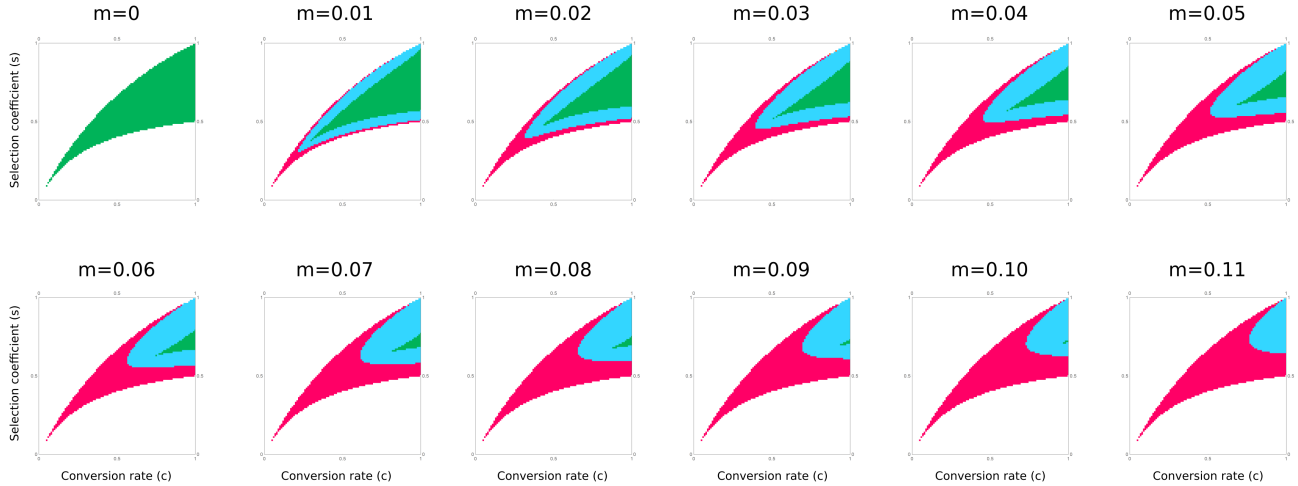

Figure S5: **Partitioning of the parameter space by the characterization of the equilibria for the s-before-m model (Eq. 3), for additive gene drives ($h = 0.5$), under different migration rates.** Green — 9 equilibria, four of which are stable, and one of which is a DTE (stable and $\hat{q}_1 > \hat{q}_2$); Purple — 9 equilibria, one of which is stable and symmetric ($\hat{q}_1 = \hat{q}_2$); Blue — 5 equilibria, 2 stable trivial (global fixation and global loss) and 3 unstable; Red — 3 equilibria, 2 stable trivial and one unstable; White — 2 trivial equilibria, one stable and one unstable. In the bottom white region, the stable equilibrium is global fixation of the gene drive allele, and in the top white region, the stable equilibrium is global loss. The only gene drive configurations for which differential targeting is possible appear in green.

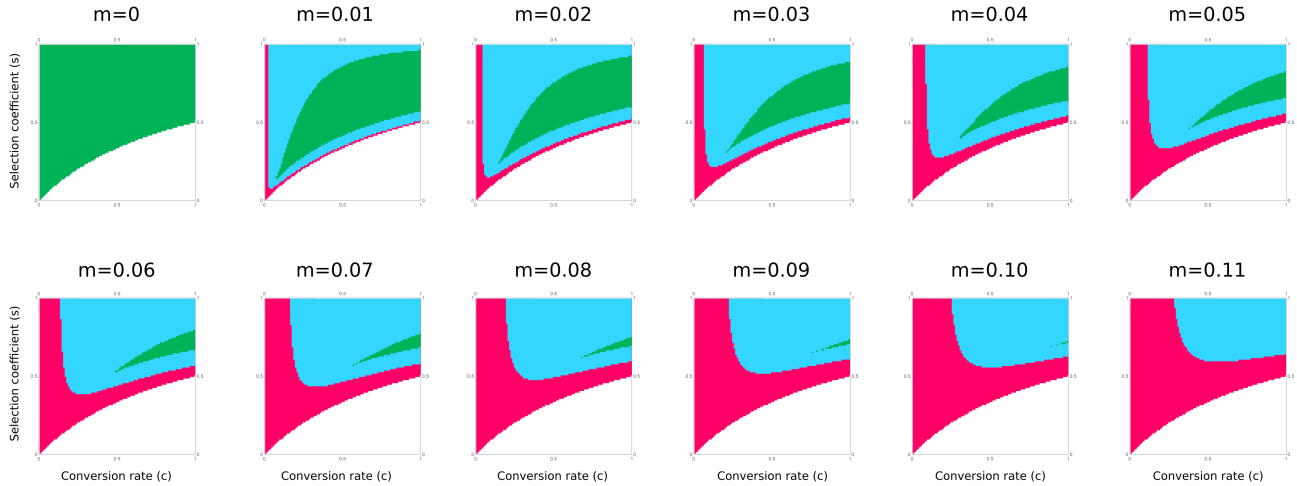

Figure S6: **Partitioning of the parameter space by the characterization of the equilibria for the s-before-m model (Eq. 3), for dominant gene drives ($h = 1$), under different migration rates.** Green — 9 equilibria, four of which are stable, and one of which is a DTE (stable and $\hat{q}_1 > \hat{q}_2$); Purple — 9 equilibria, one of which is stable and symmetric ($\hat{q}_1 = \hat{q}_2$); Blue — 5 equilibria, 2 stable trivial (global fixation and global loss) and 3 unstable; Red — 3 equilibria, 2 stable trivial and one unstable; White — 2 trivial equilibria, one stable and one unstable. In the bottom white region, the stable equilibrium is global fixation of the gene drive allele. The only gene drive configurations for which differential targeting is possible appear in green.

Appendix B - Models with genetic drift

Equations 3 and 4 are deterministic. However, we can modify the models to incorporate genetic drift by treating the allele frequencies in the demes, q_1 and q_2, as binomially distributed random variables, and applying genetic drift after selection and migrations have been accounted for. The distribution following one generation in the m-before-s model is given by:

$$\begin{aligned} q_1' &\sim Binomial\left(\frac{\widetilde{q}_1^2(1-s)+\widetilde{q}_1(1-\widetilde{q}_1)(s_n+s_c)}{\overline{w}_1}, 2N_{e,1}\right) \times \frac{1}{2N_{e,1}} \\ q_2' &\sim Binomial\left(\frac{\widetilde{q}_2^2(1-s)+\widetilde{q}_2(1-\widetilde{q}_2)(s_n+s_c)}{\overline{w}_2}, 2N_{e,2}\right) \times \frac{1}{2N_{e,2}} \end{aligned} \tag{S1}$$

where $N_{e,1}$ and $N_{e,2}$ are the effective population sizes in demes 1 and 2, respectively, and $\widetilde{q}_i$ and $\overline{w}_i$ are defined in the main text.

Similarly, for the s-before-m model, the equations describing change in frequency of A over one generation are:

$$\begin{aligned} q_1' &\sim Binomial\left((1-m)\frac{q_1^2(1-s)+q_1(1-q_1)(s_n+s_c)}{\overline{w}_1} + m\frac{q_2^2(1-s)+q_2(1-q_2)(s_n+s_c)}{\overline{w}_2}, 2N_{e,1}\right) \times \frac{1}{2N_{e,1}} \\ q_2' &\sim Binomial\left((1-m)\frac{q_2^2(1-s)+q_2(1-q_2)(s_n+s_c)}{\overline{w}_2} + m\frac{q_1^2(1-s)+q_1(1-q_1)(s_n+s_c)}{\overline{w}_1}, 2N_{e,2}\right) \times \frac{1}{2N_{e,2}} \end{aligned} \tag{S2}$$

With these models, we can simulate the allele frequency trajectories in the two demes, accounting for the effective population sizes. For Fig. 6A and Fig. 6C, we generated the trajectories for $N_{e,1} = N_{e,2} = 100$ starting at the DTE frequencies. This means that we consider the possibility of escape once the DTE is reached, and not the probability of escape during the transient convergence to the stable state, if the gene drive is initiated at frequencies other than the DTE frequencies.