

The spatial spread and the persistence of gene drives are affected by demographic feedbacks, density dependence and Allee effects

Léna Kläy^{*,1}, Léo Girardin², Vincent Calvez², and Florence Débarre^{†1}

¹Institute of Ecology and Environmental Sciences Paris (IEES Paris), Sorbonne Université, CNRS, IRD, INRAE, Université Paris Est Creteil, Université de Paris, Paris Cedex 5, France.

²Institut Camille Jordan, UMR 5208 CNRS and Université Claude Bernard Lyon 1, France

Abstract

Homing gene drive alleles bias their own transmission by converting wild-type alleles into drive alleles. If introduced in a natural population, they might fix within a relatively small number of generations, even if they are deleterious. No engineered homing gene drive organisms have been released in the wild so far, and modelling is essential to develop a clear understanding of the potential outcomes of such releases. We use deterministic models to investigate how different demographic features affect the spatial spread of a gene drive. Building on previous work, we first consider the effect of the intrinsic population growth rate on drive spread. We confirm that including demographic dynamics can change outcomes compared to a model ignoring changes in population sizes, because changes in population density can oppose the spatial spread of a drive. Secondly, we study the consequences of including an Allee effect, and find that it makes a population more prone to eradication following drive spread. Finally, we investigate the effects of the fitness component on which density dependence operates (fecundity or survival), and find that it affects the speed of drive invasion in space, and can accentuate the consequences of an Allee effect. These results confirm the importance of checking the robustness of model outcomes to changes in the underlying assumptions, especially if models are to be used for gene drive risk assessment.

^{*}Corresponding author: lena.klay@univ-brest.fr

[†]Corresponding author: florence.debarre@normalesup.org

1 Introduction

A promising but controversial new strategy for the control of natural populations, artificial gene drive biases the transmission of particular alleles to the offspring, over expectations of regular Mendelian transmission [1–3]. Such alleles can be detrimental to the individuals carrying them, and yet spread in a population thanks to their transmission advantage. Artificial gene drive implementations, so far still restricted to laboratory settings, have achieved transmission rates of 99% in yeast (*Saccharomyces cerevisiae*, [4]), more than 90% in mosquitoes (*Anopheles gambiae*, [5]), and more than 85% in fruit flies (*Drosophila melanogaster*, [6]).

In “homing drives”, biased inheritance relies on gene conversion: in a heterozygous cell, the gene drive cassette present on one chromosome induces a double-strand break on a target site on the homologous chromosome, and repair by homologous recombination duplicates the cassette. The repetition of this process through generations favors the propagation of the drive allele in the population. Conversion can theoretically happen at different steps of the life-cycle, like in the germline of the parents, or in the zygote. Practical implementations in the lab have focused on conversion in the germline [7].

Biased transmission via gene conversion can lead to the spread of new, potentially deleterious traits in a population within a relatively small number of generations. Two main types of drive can be distinguished: *replacement drives*, aiming to change features of the target population without directly affecting its size, and *suppression drives*, aiming to reduce population size (an extreme being *eradication drives*). Because we are interested in exploring the effect of demographic dynamics on the spatial spread of gene drive alleles, our work here will focus on suppression drives. Experimental proofs of principle for this type of drive have been obtained with cage populations [8, 9], and the feasibility in large populations has been confirmed by theoretical studies [2, 10, 11].

Artificial gene drive like CRISPR-based homing drive holds promise for addressing a number of important real-world issues [12–14], among which the burden caused by vector-borne diseases like malaria. Artificial gene drive could be used to spread a new trait rendering progeny of vector mosquitoes unable to transmit disease [15], or simply leading to the reduction of vector mosquitoes population size over time [8, 16]. Applications of artificial gene drive are however not limited to human health. Gene drive could help conserve or even partially restore native ecosystems by disadvantaging invasive species or favouring endemic ones [17, 18]. It could also be used in agriculture to reverse insecticide resistance in pest animal species [19] or make weeds susceptible again to herbicides [20].

As of today, no artificial gene drive organisms have been released in the wild. Lab experiments, as well as mathematical and computational models, are crucial to evaluate the risks and benefits of gene drive, and to assess the safety of potential releases. Models are however simplifications of the living world, and it is crucial to understand the impact and importance of various modelling choices, and to test the robustness of results to changes in modelling assumptions.

The simplest theoretical models of gene drive represent well-mixed populations [21], and focus on allele frequencies changes over time [18, 22–25]. Here, we investigate the spatial spread of a gene drive allele, and how demographic features affect it. Previous work has shown that propagation of a drive in a well-mixed population did not necessarily imply that the drive would spread spatially. This is in particular the case when the drive is threshold-dependent, i.e. when, in a well-mixed population, it needs to be introduced in a high enough amount to increase in proportion [26, 27]. While changes in population density may be ignored when a drive barely affects reproduction or survival, it becomes important to consider them in the case of a suppression drive, because its increase in proportion directly affects population size. Previous work on a specific model [11, 28] found that demographic features can affect the speed of advance of a drive wave over a continuous space. Here, we will assess the robustness of this result to different modelling assumptions.

A population's growth rate is determined by birth and death rates [29]. Density regulation may affect the two differently, which has consequences for overall demographic dynamics [30]. Likewise, which fitness component is affected by the drive (i.e., whether the drive reduces fecundity or decreases survival) can also influence outcomes [31]. Finally, growth at low population density may be different from growth at high population densities, i.e. Allee effects may operate [32]. This can be caused by inbreeding depression, or difficulties to find a mate when the population density is low, for example [33]. Allee effects are frequently observed in the wild, including for animals considered as potential targets of control by artificial gene drive, like mosquito species affected by inbreeding depression [34–36]. The existence of Allee effects may also influence the outcome of the release of a drive affecting population size.

In this article, we consider a one-dimensional continuous environment, and we study the spatial spread (or failure thereof) of a drive allele invading an established wild-type population. We follow the densities of the different genotypes (drive homozygous, wild-type homozygous and heterozygous) over space and time using partial differential equations. We compare four types of demographic models, depending on the presence or absence of an Allee effect, and the fitness component (birth or death) on which density dependence operates. We characterise the spread of a drive by the existence and direction of its wave of advance, by the final total population density after the drive has spread (or failed to), and by the speed of the wave. We find that an Allee effect might help to eradicate or reduce the density of the targeted population, but that it might also lead to the failure of threshold-dependent drive invasions. We also find that the effect of demography on drive spread is limited in the case of density regulation on the birth rate, but is not when density regulation affects the death rate, where wave speed increases with intrinsic growth rate. This difference emerges because drive invasion over space primarily relies on the birth of new individuals. These results highlight the importance of ecological details on the outcome of the release of a drive.

2 Models and methods

2.1 Models

In this section, we build step-by-step the different models that we will compare. These models differ in their demographic components, which we first introduce.

2.1.1 Demographic terms

To assess how sensitive results might be to different demographic modelling choices, we will consider four models differing in their birth and death terms. We first illustrate these four demographic models in the case of a genetically and spatially homogeneous population, composed only of wild-type individuals. We will compare density dependence acting on the birth term (Models $\mathcal{B}\mathcal{N}$ and $\mathcal{B}\mathcal{A}$) or death term (Models $\mathcal{D}\mathcal{N}$ and $\mathcal{D}\mathcal{A}$), and the absence (Models $\mathcal{B}\mathcal{N}$ and $\mathcal{D}\mathcal{N}$) and the presence of an Allee effect (Models $\mathcal{B}\mathcal{A}$ and $\mathcal{D}\mathcal{A}$).

We denote by r the population's intrinsic growth rate, and by a the parameter controlling the Allee effect threshold (when there is an Allee effect, $-1 \leq a \leq 1$). In these models, the population's initial growth rate (i.e., when $n \rightarrow 0$) is r in the absence of Allee effect, and $-a r$ in the presence of Allee effect. When $-1 < a < 0$, the Allee effect is said to be weak (the initial growth rate remains positive), while when $0 < a < 1$, the Allee effect is said to be strong (the initial growth rate is negative; the population only grows if already at high enough density; see Appendix B for details).

Population density is scaled so that the carrying capacity in all models is 1, and time is scaled so that the death rate in the absence of density regulation is 1. Denoting by $n(t)$ population density at time t , the four models read, before introducing genetic diversity and spatial variation, as follows:

Model $\mathcal{B}\mathcal{N}$ (Density regulation on birth terms; no Allee effect)

$$\partial_t n(t) = \overbrace{(r(1-n(t))+1)n(t)}^{\text{births}} - \overbrace{n(t)}^{\text{deaths}} \quad (\forall t > 0), \quad (1a)$$

Model $\mathcal{B}\mathcal{A}$ (Density regulation on birth terms; Allee effect present)

$$\partial_t n(t) = \overbrace{(r(1-n(t))(n(t)-a)+1)n(t)}^{\text{births}} - \overbrace{n(t)}^{\text{deaths}} \quad (\forall t > 0), \quad (1b)$$

Model $\mathcal{D}\mathcal{N}$ (Density regulation on death terms; no Allee effect)

$$\partial_t n(t) = \overbrace{(r+1)n(t)}^{\text{births}} - \overbrace{(1+r n(t))n(t)}^{\text{deaths}} \quad (\forall t > 0), \quad (1c)$$

Model $\mathcal{D}\mathcal{A}$ (Density regulation on death terms; Allee effect present)

$$\partial_t n(t) = \overbrace{(r+1)n(t)}^{\text{births}} - \overbrace{(r(n(t)-1)(n(t)-a)+r+1)n(t)}^{\text{deaths}} \quad (\forall t > 0). \quad (1d)$$

In the following, we denote the birth rate by $B(n(t))$ and the death rate by $D(n(t))$ such that the four equations can all be written as:

$$\partial_t n(t) = \overbrace{B(n(t)) n(t)}^{\text{births}} - \overbrace{D(n(t)) n(t)}^{\text{deaths}} \quad (\forall t > 0). \quad (2)$$

2.1.2 Drive and wild-type

The demographic models being defined, we now add genetic diversity to the models, following the same approach as in [11, 28]. The variable $n(t)$ becomes the total density, and we denote by n_i the density of individuals with genotype i . There are two possible alleles at the locus that we consider: the wild-type allele (W) and the drive allele (D), so that there are three different genotypes: wild-type homozygotes ($i = WW$), drive homozygotes ($i = DD$) and heterozygotes ($i = DW$). The fitness effect of a genotype is represented by a coefficient f_i acting on the birth term. It represents the selective disadvantage conferred by the drive to the individual carrying it. Wild-type homozygotes have fitness $f_{WW} = 1$, drive homozygotes have fitness $f_{DD} = 1 - s$, where s is the fitness cost of the drive, and drive heterozygotes have fitness $f_{DW} = 1 - sh$, where h is the dominance parameter. We assume that mating occurs at random: the probability that a genotype l mates with a genotype k is equal to $\frac{n_l n_k}{n^2}$. Finally, we denote by $\pi_{l,k}^i$ the probability for a couple of parents with genotypes l and k to have offspring of type i . This probability depends on the moment at which gene conversion takes place and on the probability c that gene conversion takes place and is successful ($0 \leq c \leq 1$). Here we assume that gene conversion takes place in the germline, because this is the timing currently successfully implemented in the lab [7, 37], unlike gene conversion in the zygote. We assume that failed gene conversion leaves the wild-type allele intact. In other words, we do not consider here the generation of resistance alleles. With these assumptions, the dynamics are now given by the following equations:

$$\partial_t n_i(t) = B(n(t)) f_i n(t) \underbrace{\sum_{l,k} \pi_{l,k}^i \frac{n_l(t)}{n(t)} \frac{n_k(t)}{n(t)}}_{\text{Mating term}} - D(n(t)) n_i(t) \quad (\forall t > 0) (\forall i). \quad (3)$$

The formulas for $\pi_{l,k}^i$ are included in the full equations presented in Appendix C.1.

2.1.3 Space

Our equations so far did not include space; we now add this component. We assume that the movement of individuals is described by a diffusion term with equal diffusion coefficients for all genotypes. Space is scaled such that these coefficients are normalised to 1. We obtain the following equations:

$$\partial_t n_i(t, x) = B(n(t)) f_i n(t) \sum_{l,k} \pi_{l,k}^i \frac{n_l(t)}{n(t)} \frac{n_k(t)}{n(t)} - D(n(t)) n_i(t, x) + \partial_{xx}^2 n_i(t, x) \quad (\forall t > 0) (\forall x \in \mathbb{R}) (\forall i). \quad (4)$$

All parameters of the models are summarised in Table 1.

Parameters	Range values	Description
r	$(0, +\infty)$	Intrinsic growth rate
c	$[0, 1]$	Conversion rate
s	$(0, 1)$	Fitness cost of drive homozygotes
h	$[0, 1]$	Drive dominance
a	$[-1, 1]$	Allee effect threshold

Table 1: Model parameters.

We have presented equations with genotype densities n_i (condensed model in equation (4); full equations for each model are given in Appendix C). The model can be rewritten to follow allele densities instead (see Appendix C.2), or total population size and allele frequencies (see Appendix C.3); different steps of the analysis may require different formulations of the model.

2.2 Traveling waves

The introduction of drive individuals in a wild-type population will give rise to a wave of change in genotype densities through space, called a traveling wave (except in the *gene drive clearance* case, see below). Traveling waves propagate with a constant speed, while maintaining their shape in space. We consider an initial condition in which the left half of the domain is full of drive ($n_{DD} = 1$), and the right half is full of wild-type ($n_{WW} = 1$), as illustrated in Figure S1. In this article, we are not exploring the effect of inoculum size and distribution, which is a question in itself, and arises in particular in the case of threshold-dependent drives [38]. We therefore choose an initial condition maximizing the possibility of drive spread. The model is then solved numerically. We classify the outcomes into five categories, present in the four models, depending on: the existence or not of a traveling wave; whether the population persists or is eradicated; and in the former case, the genotype(s) present at the end. These outcomes are illustrated in Figure 1.

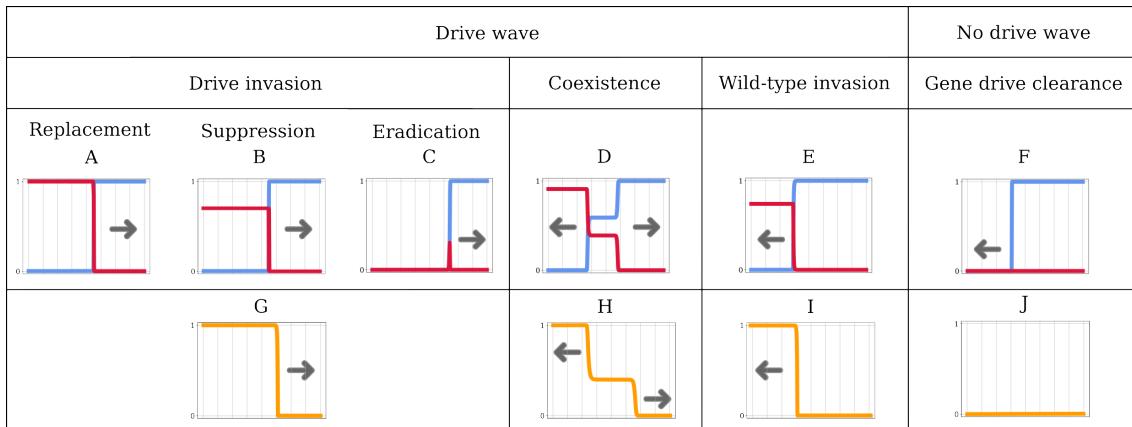


Figure 1: Types of spatial dynamics. Panels A–F correspond to allele densities, with the drive allele in red and the wild-type allele in blue. Arrows represent the direction of advance of the wave. Panels G–J show the equivalent with the drive allele frequency. The horizontal axis represents space.

It can happen that the model leads to the decay of the drive allele uniformly in space. This case arises in particular when a well-mixed population composed only of drive individuals is not sustainable. The introduced drive subpopulation just dies out, freeing space; in this case, there is no drive traveling wave. We describe this as *gene drive clearance* (Figure 1F, J). The wild-type population then recolonizes the emptied space, at a speed described in the standard Fisher-KPP traveling wave problem (see [39–41]).

When the drive traveling wave does exist, we distinguish between two cases, depending on the sign of the speed v . When $v > 0$, the wave moves to the right: it is a *drive invasion*. When $v < 0$, the wave moves to the left: it is a *wild-type invasion*. In some specific cases, drive and wild-type invasions can happen simultaneously: the waves decompose into two sub-traveling wave solutions over half of the domain. They move in opposite directions and lead to the coexistence of both alleles in-between (Figure 1D, H).

When the drive invades and replaces the wild-type (Figure 1G), we distinguish three cases depending on the state of the population in the wake of the front(s): i) in the case of *replacement* drives, the population persists in the wake of the front(s) at the same density as the original wild-type population (Figure 1A); ii) in the case of *suppression* drives, the population persists in the wake of the front(s), albeit at a lower density than the original wild-type population (Figure 1B); iii) in the case of *eradication* drives, the population is eradicated in the wake of the drive invasion front(s), leaving empty space behind (Figure 1C). In the latter case, two scenarios are possible: persistence of drive homozygotes only (as in Figure 1G); but also possibly persistence of all genotypes (as in Figure 1H).

The code for these simulations is available on GitHub (https://github.com/LenaKlay/gd_project_1/deterministic). We ran our simulations in Python 3.6, with the Spyder environment. Heatmaps in Figures 3 were computed thanks to the INRAE Migale bioinformatics facility.

3 Results

3.1 Demography and dominance can affect the final allelic proportions

Here, we focus on the importance of demography in the model, i.e., on the role played by the intrinsic growth rate r on the final allelic proportions. Analytical results can be obtained for $r = 0$ and for $r \rightarrow \infty$; intermediate cases are investigated numerically.

When $r = 0$, deaths and births compensate each other in a fully wild-type population. In this limit case, models \mathcal{BN} , \mathcal{BA} , \mathcal{DN} and \mathcal{DA} are the same (given in equation (C.5)). Both the final densities of all genotypes and the speed of the wave are therefore the same for all models, which we will characterise below, recalling results from our previous work [28].

Leaving aside the density-dependence constraint, the bigger r is, the faster the wild-type population grows. When $r \rightarrow \infty$, final allelic proportions are the same in models \mathcal{BN} , \mathcal{BA} , \mathcal{DN} and \mathcal{DA} (see Section C.4). This is however not necessarily the case for the total population density and for the wave speed.

Following previous work [28], let us introduce:

$$s_1 := \frac{c}{1 - h(1 - c)}, \quad s_{2,g} := \frac{c}{2ch + h(1 - c)} = \frac{c}{h(1 + c)}. \quad (5)$$

These are threshold values of the fitness cost s determining qualitatively different outcomes. When the drive allele is recessive ($h < 1/2$), $s_1 < s_{2,g}$; when the drive allele is dominant ($h > 1/2$), $s_1 > s_{2,g}$.

When the fitness cost s is low enough ($s < \min(s_1, s_{2,g})$), there is a wave of advance of the drive for both $r = 0$ and $r \rightarrow \infty$ (drive invasion, as in Fig. 1G).

When the fitness cost s is high enough ($s > \max(s_1, s_{2,g})$), and the intrinsic growth rate is high ($r \rightarrow \infty$), the drive wave retreats (wild-type invasion, as in Fig. 1I). When the intrinsic growth rate is low ($r = 0$), $s > \max(s_1, s_{2,g})$ results in drive clearance (as in Fig. 1J): the drive is just too costly even for a full-drive population.

What happens for intermediate fitness cost ($\min(s_1, s_{2,g}) < s < \max(s_1, s_{2,g})$) and high growth rate depends on the dominance parameter h . If $h < 1/2$, drive and wild-type alleles coexist eventually (coexistence, as in Fig. 1H). If $h > 1/2$, there is a bistability, the drive is threshold-dependent: the final outcome is either drive invasion or wild-type invasion, and depends on the initial conditions.

These results are summarised in Table 2.

(a) When $h < 1/2$			
	$0 < s < s_1$	$s_1 < s < s_{2,g}$	$s_{2,g} < s < 1$
$r \rightarrow \infty$	Drive invasion	Coexistence	Wild-type invasion
$r = 0$	Drive invasion**		Gene drive clearance

(b) When $h > 1/2$			
	$0 < s < s_{2,g}$	$s_{2,g} < s < s_1$	$s_1 < s < 1$
$r \rightarrow \infty$	Drive invasion	Bistability	Wild-type invasion
$r = 0$	Drive invasion**		Gene drive clearance

Table 2: Types of model outcomes for Models \mathcal{BN} , \mathcal{BA} , \mathcal{DN} and \mathcal{DA} , depending on the fitness cost s , intrinsic growth rate r and dominance parameter h . The outcomes are in terms of allele proportions, as in Fig. 1G–J.

These results illustrate the importance of taking demography into account. Threshold-dependent drives (i.e. drives leading to bistabilities), are considered more socially responsible than threshold independent drives, as they are potentially localised and reversible [25, 26, 42]. The intrinsic growth rate r is a key component to reach this threshold dependence, as r has to be sufficiently large for the bistability to happen. Indeed, a small r would result in the systematic decay of gene drive alleles (Table 2) and no possibility of drive invasion at all.

As in models without demography nor spatial structure [22, 23], the dominance parameter h determines whether threshold-dependence can be attained or not: a bistable outcome only exists when

**The term “drive invasion” here is slightly ambiguous, as it does not specify the genetic composition in the wake of the eradication wave. This exponentially small population might contain the three possible genotypes or only the drive-homozygous one, depending on parameter choices.

$h > \frac{1}{2}$, i.e. when the fitness of heterozygous individuals is closer to the fitness of drive homozygous individuals than to that of wild-type homozygous individuals. This result was already given in a simpler panmictic model [22]: indeed, the birth and death terms in our models \mathcal{BN} and \mathcal{BA} tend to this panmictic model for large values of r .

3.2 The strength of the Allee effect and the choice of the fitness component targeted by the density-dependence affect the final allelic density

In the previous section, we have only described outcomes in terms of allele frequencies. In this section, we compare the final population density n^* in the four models, and in particular conditions for which the population goes extinct ($n^* = 0$). We detail the final densities in all three types of invasions: drive invasion, wild-type invasion, and coexistence. In case of gene drive clearance (decay of the drive allele uniformly in space), the final density is equivalent to the one obtained after a wild-type invasion: population size goes back to carrying capacity 1.

In all three types of invasion, there are up to three possible regimes: population eradication ($n^* = 0$); population persistence ($n^* = n^+ > 0$); and bistability (the final total population size is either 0 or n^+ depending on the initial condition relative to a specific density n^*). Note that the bistability involved here is different from the bistability on allele frequencies as seen in the previous section; the bistability that we consider in this section is about population densities.

We can write the final population densities in a generic manner for the three types of invasion. We define the mean fitness \mathcal{F} :

$$\mathcal{F}(p_{\text{D}}) = (1-s)(p_{\text{D}})^2 + 2(1-sh)p_{\text{D}}(1-p_{\text{D}}) + (1-p_{\text{D}})^2, \quad (6)$$

and p^* the final proportion of the drive allele in the population. It verifies:

$$((2h-1)s p_{\text{D}}^* + (1-sh)(1+c)-1) p_{\text{D}}^* (1-p_{\text{D}}^*) = 0, \quad (7)$$

according to the allelic frequency systems detailed in Section C.3. For wild-type invasion, $p_{\text{D}}^* = 0$ and $\mathcal{F}(p_{\text{D}}^*) = 1$; for drive invasion, $p_{\text{D}}^* = 1$ and $\mathcal{F}(p_{\text{D}}^*) = 1-s$; for coexistence,

$$p_{\text{D}}^* = \frac{1-(1-sh)(1+c)}{s(2h-1)} \in (0, 1). \quad (8)$$

and $\mathcal{F}(p^*)$ is still given by equation (6). Note that $\mathcal{F}(p_{\text{D}}^*) \in [1-s, 1]$ as \mathcal{F} is a decreasing function of p_{D}^* .

The final densities n^* are then computed by solving the allelic frequency systems detailed in Section C.3 with the relevant value of p_{D}^* . These results, holding for all values of the intrinsic growth rate r , are summarised in Table 3 and illustrated in Figure 2 with $c = 0.85$ and $h = 0.9$.

Model	Regime	n^+ and n^τ (if it exists)
$\mathcal{B}\mathcal{N}$	Eradication if $r < \frac{1-\mathcal{F}(p_D^*)}{\mathcal{F}(p_D^*)}$ Persistence if $r > \frac{1-\mathcal{F}(p_D^*)}{\mathcal{F}(p_D^*)}$	$n^{+(\mathcal{B}\mathcal{N})} = 1 - \frac{1-\mathcal{F}(p_D^*)}{r\mathcal{F}(p_D^*)}$
$\mathcal{B}\mathcal{A}$	Eradication if $r < \frac{1-\mathcal{F}(p_D^*)}{\left(\frac{1-a}{2}\right)^2 \mathcal{F}(p_D^*)}$ Bistability if $r > \frac{1-\mathcal{F}(p_D^*)}{\left(\frac{1-a}{2}\right)^2 \mathcal{F}(p_D^*)}$ and $-ar < \frac{1-\mathcal{F}(p_D^*)}{\mathcal{F}(p_D^*)}$ Persistence if $r > \frac{1-\mathcal{F}(p_D^*)}{\left(\frac{1-a}{2}\right)^2 \mathcal{F}(p_D^*)}$ and $-ar > \frac{1-\mathcal{F}(p_D^*)}{\mathcal{F}(p_D^*)}$	$n^{\tau(\mathcal{B}\mathcal{A})} = \frac{1+a-\sqrt{(1+a)^2-4(a+\frac{1-\mathcal{F}(p_D^*)}{r\mathcal{F}(p_D^*)})}}{2}$ $n^{+(\mathcal{B}\mathcal{A})} = \frac{1+a+\sqrt{(1+a)^2-4(a+\frac{1-\mathcal{F}(p_D^*)}{r\mathcal{F}(p_D^*)})}}{2}$
$\mathcal{D}\mathcal{N}$	Eradication if $r < \frac{1-\mathcal{F}(p_D^*)}{\mathcal{F}(p_D^*)}$ Persistence if $r > \frac{1-\mathcal{F}(p_D^*)}{\mathcal{F}(p_D^*)}$	$n^{+(\mathcal{D}\mathcal{N})} = 1 - \frac{(r+1)(1-\mathcal{F}(p_D^*))}{r}$
$\mathcal{D}\mathcal{A}$	Eradication if $r[\left(\frac{1-a}{2}\right)^2 - (1-\mathcal{F}(p_D^*))] < 1-\mathcal{F}(p_D^*)$ Bistability if $r[\left(\frac{1-a}{2}\right)^2 - (1-\mathcal{F}(p_D^*))] > 1-\mathcal{F}(p_D^*)$ and $r[-a - (1-\mathcal{F}(p_D^*))] < 1-\mathcal{F}(p_D^*)$ Persistence if $r[\left(\frac{1-a}{2}\right)^2 - (1-\mathcal{F}(p_D^*))] > 1-\mathcal{F}(p_D^*)$ and $r[-a - (1-\mathcal{F}(p_D^*))] > 1-\mathcal{F}(p_D^*)$	$n^{\tau(\mathcal{D}\mathcal{A})} = \frac{1+a-\sqrt{(1+a)^2-4(a+\frac{(r+1)(1-\mathcal{F}(p_D^*))}{r})}}{2}$ $n^{+(\mathcal{D}\mathcal{A})} = \frac{1+a+\sqrt{(1+a)^2-4(a+\frac{(r+1)(1-\mathcal{F}(p_D^*))}{r})}}{2}$

Table 3: Regimes and final densities in Models $\mathcal{B}\mathcal{N}$, $\mathcal{B}\mathcal{A}$, $\mathcal{D}\mathcal{N}$ and $\mathcal{D}\mathcal{A}$, where p_D^* is the final proportion of the drive allele in the population and $\mathcal{F}(p_D^*)$ the mean fitness. We consider three regimes regarding the value of the final population density, n^* : population eradication ($n^* = 0$), population persistence ($n^* = n^+ > 0$) and bistability (the final total population size is either 0 or n^+ depending on the initial condition relative to the threshold density n^τ).

3.2.1 When the Allee effect gets stronger, the population is more prone to eradication

In the models without Allee effects, there is no bistability for the final population size once the type of invasion is known (see Table 3). However, and as seen before, the type of invasion itself might change depending on the initial conditions (bistability on allele frequency, see Table 2b), and consequently can still affect the final population size. The situation is different in models with Allee effect, where the final population size might depend on the initial conditions. In case of a weak Allee effect ($-1 < a < 0$), the three possible regimes are eradication, persistence and bistability. In case of a strong Allee effect ($0 < a < 1$), the only two possible regimes are eradication and bistability.

The condition for eradication is the same in all four models if we set $a = -1$. However as a grows, i.e. when the Allee effect gets stronger, the ranges of s (fitness disadvantage for drive) and r (intrinsic growth rate) leading to population eradication get larger in the models with Allee effect $\mathcal{B}\mathcal{A}$ and $\mathcal{D}\mathcal{A}$ (Appendix D and G).

Similarly when a grows, the ranges of s and r leading to population persistence get smaller in Models $\mathcal{B}\mathcal{A}$ and $\mathcal{D}\mathcal{A}$ (Appendix E and G). We observe that the larger a is, the more the “persistence” and “bistable” regimes are restricted to high values of r and small values of s (Appendix G). In the case of a strong Allee effect ($0 < a < 1$), the “persistence” regime even disappears (Appendix E).

If the drive persists at the end, i.e. if there is a drive invasion or a coexistence state with a non-zero final population density, then, given how our models are formulated, the stronger the Allee effect, the smaller the final population size in Models $\mathcal{B}\mathcal{A}$ and $\mathcal{D}\mathcal{A}$ (Appendix F2).

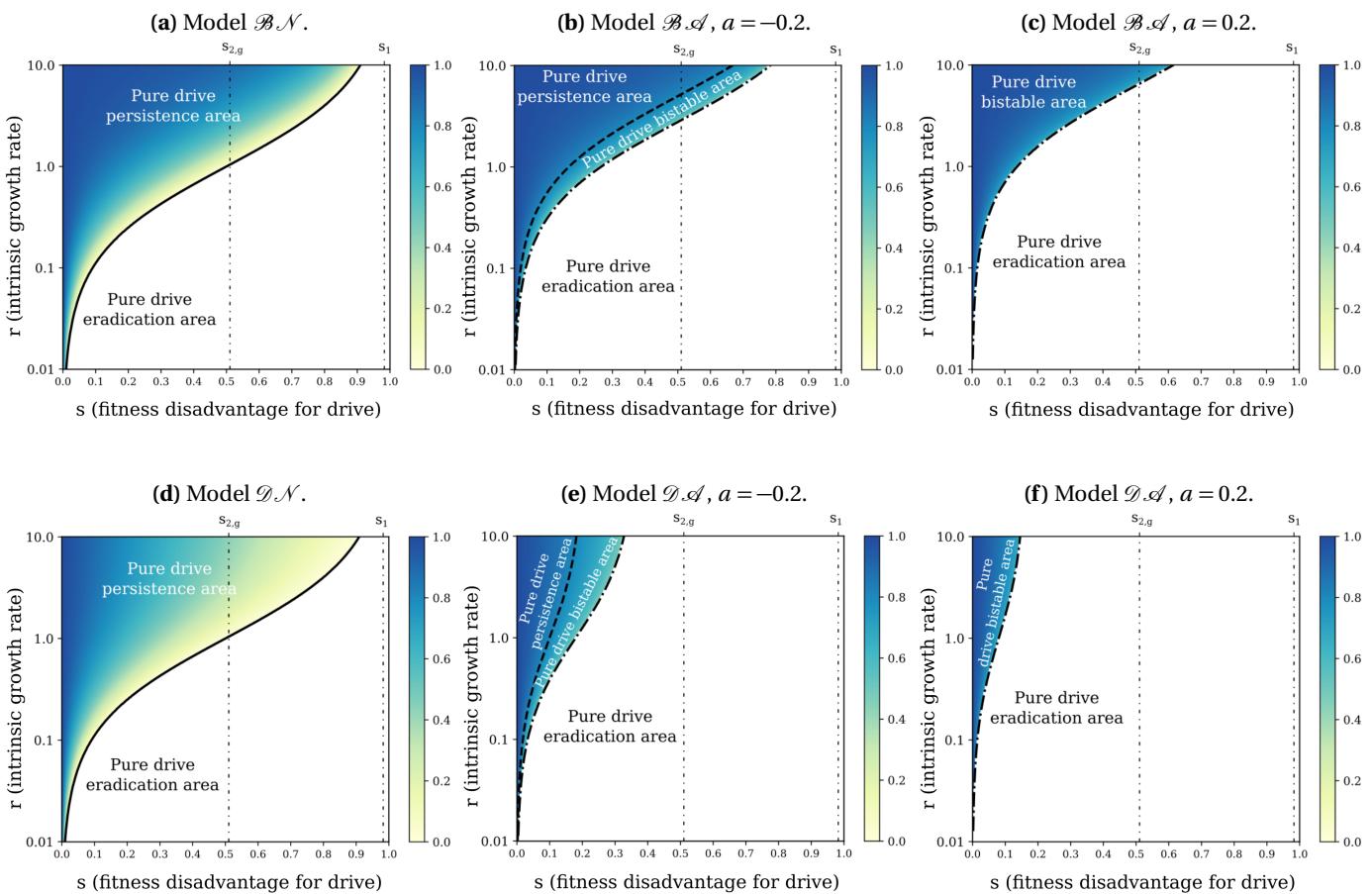


Figure 2: Value of the final population density $n^* = \max(0, n^+)$ in case of a drive invasion, shown in shades of color, with parameters $c = 0.85$ and $h = 0.9$. Since $h > \frac{1}{2}$, drive invasion always occurs for $s < s_{2,g}$, is not systematic for $s_{2,g} < s < s_1$, and never occurs for $s_1 < s$ (see Table 2b). The “pure drive” areas correspond to the final population density expected in case of a drive invasion: persistence ($n^* = n^+$), eradication ($n^* = 0$) or bistability (either $n^* = n^+$ or $n^* = 0$ depending on the initial condition). These final densities are given in Table 3 with $p_D^* = 1$, because the proportion of drive alleles in the final population is one in case of a drive invasion without coexistence.

3.2.2 A density-dependence constraint on survival instead of fecundity might lead to a higher chance of eradication and reduces the final allelic density in case of drive persistence

The conditions for eradication or persistence are the same in Models $\mathcal{B}\mathcal{N}$ and $\mathcal{D}\mathcal{N}$, i.e., they do not depend on whether the density dependence acts on births or deaths (no “bistable” regime for these models, Table 3 and Figure 2). However if we consider the models with Allee effect for a given a value, there is a higher chance of eradication when density dependence acts on deaths (Model $\mathcal{D}\mathcal{A}$) than when it acts on births (Model $\mathcal{B}\mathcal{A}$) (Appendix D). Interestingly, when $r \rightarrow \infty$, eradication is still possible in Model $\mathcal{D}\mathcal{A}$ (for $\mathcal{F}(p_D^*) < 1 - \left(\frac{1-a}{2}\right)^2$) while it is not in Model $\mathcal{B}\mathcal{A}$ (Figure 2).

If the drive persists in models without Allee effect, i.e. if there is a drive invasion or a coexistence state with a non-zero final population density, the final population size is lower when the density-dependence constraint acts on the death rate (Model $\mathcal{D}\mathcal{N}$) than on the birth rate (Model $\mathcal{B}\mathcal{N}$, see Appendix F.1). The same conclusion holds for models with Allee effect: for a given a value, if the drive persists, the final population is less dense in Model $\mathcal{D}\mathcal{A}$ than in Model $\mathcal{B}\mathcal{A}$ (Appendix F.2).

3.3 A density-dependence constraint on the deaths instead of the births results in a faster invasion

We now focus on the speed of a drive invasion, i.e. the speed of the traveling wave emerging from a drive invasion (see Section 2.2).

A speed v of the wave can be calculated when the models are simplified (linearised) assuming low drive density. This speed corresponds to the speed of drive invasion when the movement of individuals is caused by the few drive individuals at the expansion edge, where the drive density is low. In this case, the wave is called a “pulled wave”. This happens when such small populations have high growth rates, because the movement is then mainly driven by new births. When movement is brought about by individuals in the bulk of the wave (i.e., in the case of a “pushed wave”), the calculated speed corresponds to a lower estimate of the speed: the real speed is higher, but cannot be calculated in general. In a previous article [28], we showed that the calculated speed v corresponds to the speed of a drive invasion when the dominance parameter h is lower than $\frac{1}{2}$, or for a drive fitness cost s small enough when $h > \frac{1}{2}$ (for a precise condition, see [28]). This result was rigorously proven for large and small values of the intrinsic growth rate r , and numerically observed for all r . We calculate and compare this speed value in our four models (mathematical details are given in Appendix C.5).

In models $\mathcal{B}\mathcal{N}$ and $\mathcal{B}\mathcal{A}$ with density dependence acting on the birth term, this speed is given by:

$$v_{\mathcal{B}\mathcal{N}, \mathcal{B}\mathcal{A}} = 2\sqrt{(1-sh)(1+c)-1}. \quad (9)$$

In models $\mathcal{D}\mathcal{N}$ and $\mathcal{D}\mathcal{A}$ with density dependence acting on the death term, the speed becomes:

$$v_{\mathcal{D}\mathcal{N}, \mathcal{D}\mathcal{A}} = 2\sqrt{(1+r)[(1-sh)(1+c)-1]}. \quad (10)$$

The speeds only exist for $(1-sh)(1+c) > 1$ (or equivalently $s < s_2$, with $s_{2,g}$ given in equation (5)), which is the necessary condition to have a strictly positive drive allele production at the front of the wave. To understand why, first note that the density of drive alleles is very low at the front of the wave. Therefore, we can make the approximation that at least one parent in each couple formed at the front of the wave is a wild-type homozygote WW. Consequently, the offspring carrying a drive allele are necessarily heterozygotes: in the front of the wave, the production of drive alleles only relies on the heterozygotes. These heterozygotes have a fitness of $(1-sh)$ and produce drive alleles at rate $(1+c)$: therefore, for a drive invasion to be possible, the production rate $(1-sh)(1+c)$ of drive alleles should be above the rate 1 at which they disappear. The higher the production rate is, the faster the wave moves.

The speeds $v_{\mathcal{B}\mathcal{N}, \mathcal{B}\mathcal{A}}$ and $v_{\mathcal{D}\mathcal{N}, \mathcal{D}\mathcal{A}}$ are very similar but differ by one coefficient: $v_{\mathcal{D}\mathcal{N}, \mathcal{D}\mathcal{A}}$ is $\sqrt{r+1}$ times larger than $v_{\mathcal{B}\mathcal{N}, \mathcal{B}\mathcal{A}}$. This difference relies on the density-dependence constraint, affecting either the births or the deaths. At the front of the wave, the population density, composed nearly only of wild-type individuals, reaches the maximum carrying capacity. Consequently, the density-dependence constraint hinders any increase in the population density and this happens in two different ways: in models $\mathcal{B}\mathcal{N}$ and $\mathcal{B}\mathcal{A}$, it limits the births so that they do not exceed the deaths, whereas in models $\mathcal{D}\mathcal{N}$ and $\mathcal{D}\mathcal{A}$, it increases the death rate to compensate the births. As a result, the turnover rate is greater in models $\mathcal{D}\mathcal{N}$ and $\mathcal{D}\mathcal{A}$, which induces a faster invasion because the wave movement is mainly driven by new births. Details of the speed calculations are given in Appendix C.5. To illustrate this result, we plot the speed of the wave for the four models in Figure 3 and observe that the speed of the drive invasion always increases with r in models $\mathcal{D}\mathcal{N}$ and $\mathcal{D}\mathcal{A}$, in contrast with models $\mathcal{B}\mathcal{N}$ and $\mathcal{B}\mathcal{A}$, for which the speed of the wave does not depend on r . Note however that while the speed $v_{\mathcal{B}\mathcal{N}, \mathcal{B}\mathcal{A}}$ is independent of r for a drive fitness cost s small enough, it is not the case for the final density n_D^* , which depends on r (Table 3 and Figure 2, models $\mathcal{B}\mathcal{N}$ and $\mathcal{B}\mathcal{A}$). As a result, for a small enough s in models $\mathcal{B}\mathcal{N}$ and $\mathcal{B}\mathcal{A}$, the wave travels at a constant speed no matter the density of population left behind.

3.4 The Allee effect might cause the failure of threshold dependent drives

Finally, in Figure 3, we observe that for $s > s_{2,g}$, the Allee effect might prevent drive invasion. Yellow-orange areas in heatmaps (a) (resp. (d)) are becoming blue in heatmaps (b) and (c) (resp. (e) and (f)) due to the Allee effect in the model. Noticeably, the drive invasions for $s > s_{2,g}$ are threshold-dependent drive invasions, often considered as more socially responsible than threshold independent drives invasions [25, 26]. Once again this influence is accentuated when the Allee effect gets stronger (for larger values of a). Allee effects may therefore hamper the spread of threshold-dependent drives.

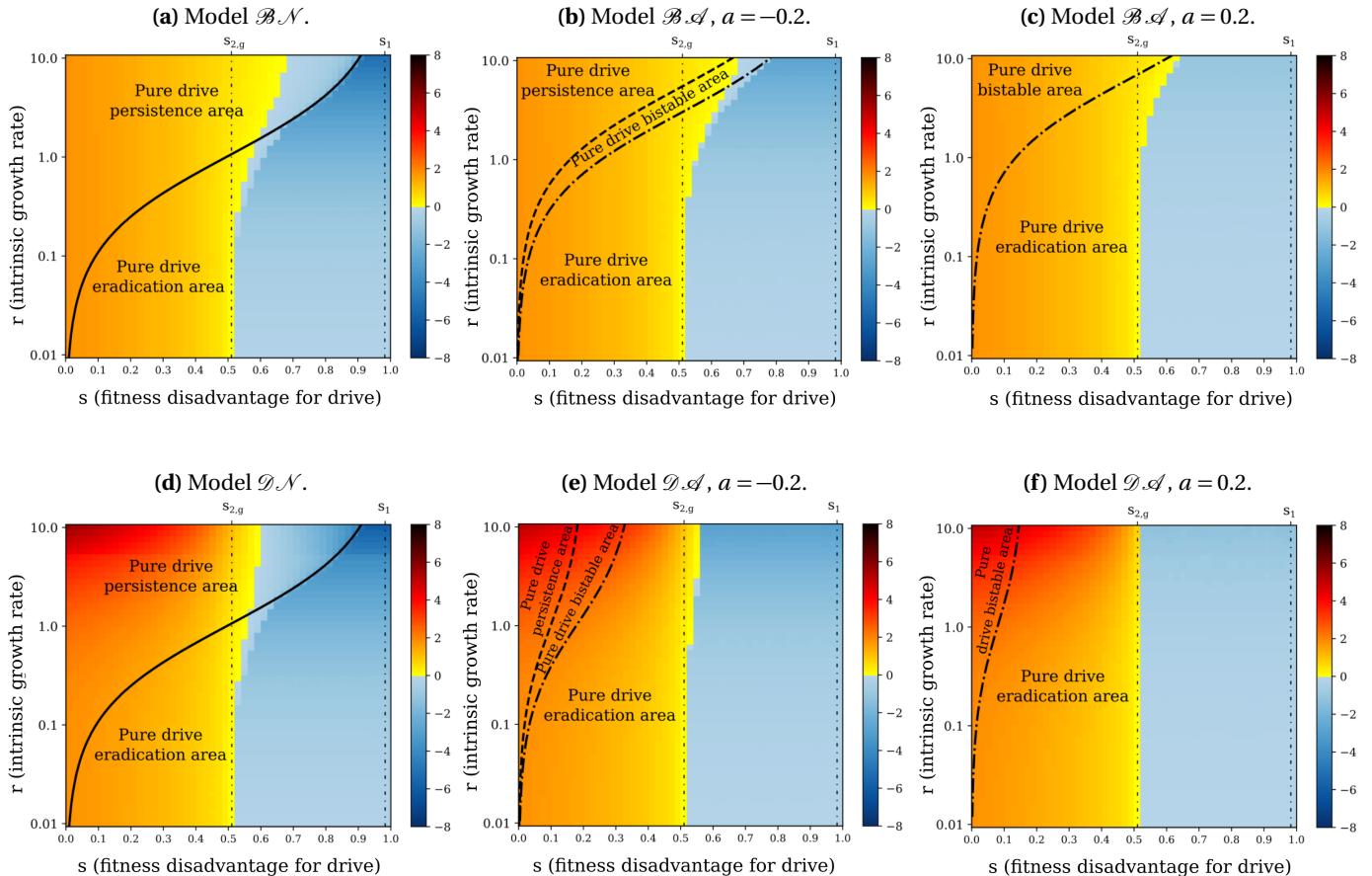


Figure 3: Speed of the wave of drive alleles shown as shades of color, with parameters $c = 0.85$ and $h = 0.9$. When the drive invades the population, the speed is positive (in yellow-orange). On the contrary, when the wild-type invades the population, the speed is negative (in blue). The “pure drive” areas correspond to the final population density expected in case of a drive invasion: persistence ($n^* = n^+$), eradication ($n^* = 0$) or bistability (either $n^* = n^+$ or $n^* = 0$ depending on the initial condition). Since $h > \frac{1}{2}$, coexistence of drive and wild-type alleles in the final population is impossible (see Table 2b). Each model is initiated with the initial conditions described in Figure S1. For $s_{2,g} < s < s_1$, the speed value may vary if we consider different initial conditions (bistability on the final allele frequencies, see Table 2b). On the left of the vertical line $s = s_{2,g}$, the speed value is independent of r when the density-dependence constraint acts on fecundity (Models $\mathcal{B}\mathcal{N}$ and $\mathcal{B}\mathcal{A}$) while it grows with r when the density-dependence constraint acts on survival (Models $\mathcal{D}\mathcal{N}$ and $\mathcal{D}\mathcal{A}$).

4 Discussion

Understanding the conditions for the spatial spread of an artificial gene drive and its consequences on a targeted population is essential before considering any field release. Laboratories experiments provide information on gene drive dynamics in a small confined and controlled environment, and mathematical models can help gain further insights at small and larger scales.

Theoretical models are meant to provide insights on real-world dynamics, so it is important to assess how the result of a model depends on modelling choices. In this article, we investigated the influence of considering i) demography, and more precisely different values of the intrinsic growth rate of the target population, ii) the presence/absence of an Allee effect and iii) which fitness component (birth or death) is affected by density dependence. We considered the effects of these features on the type of outcome, on final population density, and on the speed of the wave.

We first described the different qualitative outcomes, extending results from our previous studies [11, 28] on the importance of taking into account demography in the models. We confirm that the intrinsic growth rate r qualitatively affects results at intermediate values of the fitness cost s . A high intrinsic growth rate leads to a threshold-dependent drive invasion, while a low intrinsic growth rate results in the decay of drive alleles uniformly in space. Models not considering population densities but focusing on frequencies [e.g. 18, 22–25] have dynamics similar to our models provided $r \rightarrow \infty$.

As intuitively expected, an Allee effect makes the population more susceptible to eradication, widening the range of s (fitness disadvantage for drive) and r (intrinsic growth rate) values leading to population eradication after a drive invasion. This phenomenon is accentuated when the Allee effect gets stronger (for larger values of α). In addition, given the way our models are formulated, in models with Allee effect, the larger α , the lower the final population density in case of drive persistence, meaning that an Allee effect might represent a non-negligible helping force to eradicate or suppress natural populations. However, we also showed that an Allee effect might reduce the range of s and r values leading to a threshold-dependent drive invasion, often considered as more socially responsible than threshold independent drives invasions [25, 26].

Finally, we considered the impact of whether the density-dependence constraint targets births or deaths: close to the maximal carrying capacity, in case of rarefaction of the resources, the net growth of the population is limited by either a low number of offspring per generation or a high death rate.

In this study, we show that when density dependence acts on deaths, it acts in concert with the Allee effect by enlarging the eradication conditions and reducing the final density, compared to when density dependence acts on births. How density dependence acts also strongly impacts the speed of propagation: a drive invasion would be $\sqrt{r+1}$ times faster for a density-dependence constraint on the death rate instead of the birth rate. This prediction holds for a fitness cost reducing the birth rate (individuals carrying drive alleles have fewer offspring than wild-type ones). However, the conclusions might change for a fitness cost increasing the death rate instead, as shown in a different model of CRISPR-based homing drives [31].

Our models are deterministic. This deterministic framework can describe population dynamics at large scales, but cannot capture “chasing” events, whereby wild-type individuals recolonize empty space in the wake of the wave of an eradication drive, and which can arise at low population densities [43–47]. Stochastic fluctuations are likely to be important in particular in the case of suppression and eradication drives, and are left for future investigation.

Among the deterministic models in the literature, the models we develop are generalist: they could be applied to different species, and any gene drive construct reducing the fitness of the individual carrying it. These models do not aim to bring precise and quantitative predictions, for which more specific models need to be developed, but rather get some insights into the possible outcomes, and dissect the roles played by different model elements. However, this generalist approach naturally comes with simplifications.

In our models, we assume that gene conversion either successfully takes places, or does not take place. We did not include resistance alleles which can emerge when conversion fails and repairs by non-homologous end-joining occur, or resistance due to standing genetic variation at the target locus. The emergence of resistance alleles can alter the propagation of the drive, but can be mitigated by specific constructs [18, 48–50].

Some other simplifications are directly related to the biological characteristics of the species. The polyandrous mating system of mice populations can limit the spread of gene drives [51, 52] or mate search capabilities [53]. In mosquito populations, the plural life stages (egg, larva, pupa and adults) might influence the modelling conclusions and need to be taken into account by including corre-

sponding age structure in models [54–56]. In bee populations, the haploid phases of the life cycle result in less powerful drives: the conditions for fixation are narrower and the spread is slower [57, 58]. Finally, it is not rare that males and females have different fitnesses in transgenic mosquitoes [8, 16, 47, 59]: more specific models than ours would need to include sex differences.

Finally and more broadly, species do not live in isolation, and interactions of the targeted species within its ecosystem would need to be considered. Competing species or predators can facilitate drive-based suppression [57], and environmental conditions such as seasonality (dry or wet season) can highly impact the eradication of mosquito populations, for example [44, 46, 47]. It is of public utility to also consider the impact of gene drive on the whole ecosystem and anticipate the potential risks: the probability of transmission of the gene drive cassette to another species [60], or the cascade of population dynamics and evolutionary processes potentially initiated by the eradication of a species [61].

Overall, we have shown the importance of considering precise population dynamics on the outcome of the release of a drive. This approach through theoretical models gives first interesting insights that now need to be enhanced with ecological knowledge on specific systems.

Appendix

A Initial conditions

All our simulations in one-dimensional space are initiated with the initial conditions described in the main text, and illustrated in Figure S1:

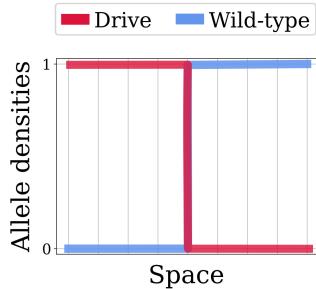


Figure S1: Initial conditions used in the simulations. The left half of the domain is full of drive ($n_{DD} = n_D = 1$), and the right half is full of wild-type ($n_{WW} = n_W = 1$).

B Allee effect

We consider the equation describing the dynamics of the population density n :

$$\partial_t n - \partial_{xx}^2 n = n(1-n)\sigma(n) = f(n) \quad \text{with } n \in [0, 1]. \quad (\text{B.1})$$

The Allee effect characterises a positive correlation between population density and the per capita population growth rate.

Without Allee effect, the population growth rate $f(n)$ is always positive and the per capita population growth rate $\frac{f(n)}{n}$ is maximum as the population density n tends to zero. This happens for example when $\sigma(n) = 1$ in Equation (B.1) (Figure S2a). Mathematically, we write:

$$\max_{n \in [0, 1]} \left(\frac{f(n)}{n} \right) \leq f'(0). \quad (\text{B.2})$$

With a weak Allee effect, the population growth rate $f(n)$ is still positive, but the maximum of the per capita population growth rate $\frac{f(n)}{n}$ is reached at a strictly positive population density. This happens for example when $\sigma(n) = (n - a)$ with $-1 < a < 0$ in Equation (B.1) (Figure S2b). Mathematically:

$$\max_{n \in [0, 1]} \left(\frac{f(n)}{n} \right) > f'(0) > 0. \quad (\text{B.3})$$

Finally, with a strong Allee effect, the population growth rate $f(n)$ is negative for small population density, and positive after. This happens for example when $\sigma(n) = (n - a)$ with $0 < a < 1$ in Equation (B.1) (Figure S2c). Mathematically:

$$\exists a > 0 \text{ such that } \forall n \in]0, a[\quad \frac{f(n)}{n} < 0. \quad (\text{B.4})$$

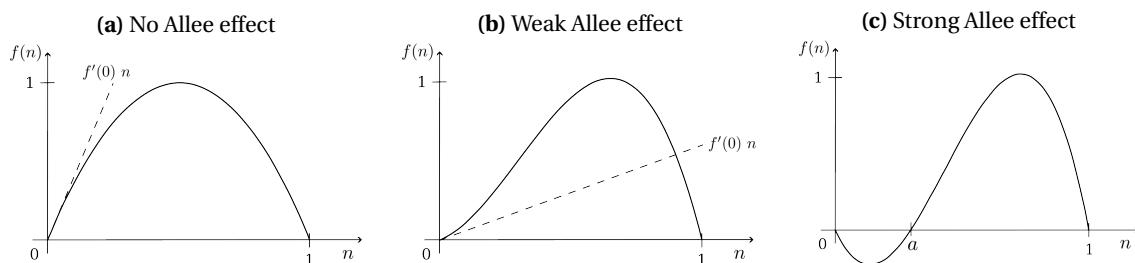


Figure S2: Illustration of the three cases concerning the Allee effect.

C Models

C.1 Genotype densities

For the sake of clarity, we omit variables in the notation ($n_i = n_i(t, x)$) in the following. Each model contains three equations for the three genotype densities: homozygote drive n_{DD} , heterozygote n_{DW} and homozygote wild-type n_{WW} .

Model \mathcal{BN} :

$$\left\{ \begin{array}{l} \partial_t n_{\text{DD}} - \partial_{xx}^2 n_{\text{DD}} = (1-s)(r(1-n)+1) \frac{\frac{1}{4}(1+c)^2 n_{\text{DW}}^2 + (1+c) n_{\text{DW}} n_{\text{DD}} + n_{\text{DD}}^2}{n} - n_{\text{DD}}, \\ \partial_t n_{\text{DW}} - \partial_{xx}^2 n_{\text{DW}} = (1-sh)(r(1-n)+1) \frac{(1+c) n_{\text{WW}} n_{\text{DW}} + 2 n_{\text{WW}} n_{\text{DD}} + \frac{1}{2}(1-c^2) n_{\text{DW}}^2 + (1-c) n_{\text{DW}} n_{\text{DD}}}{n} - n_{\text{DW}}, \\ \partial_t n_{\text{WW}} - \partial_{xx}^2 n_{\text{WW}} = (r(1-n)+1) \frac{n_{\text{WW}}^2 + (1-c) n_{\text{WW}} n_{\text{DW}} + \frac{1}{4}(1-c)^2 n_{\text{DW}}^2}{n} - n_{\text{WW}}. \end{array} \right. \quad (\text{C.1})$$

Model \mathcal{BA} :

$$\left\{ \begin{array}{l} \partial_t n_{\text{DD}} - \partial_{xx}^2 n_{\text{DD}} = (1-s)(\max(r(1-n)(n-a)+1, 0)) \frac{\frac{1}{4}(1+c)^2 n_{\text{DW}}^2 + (1+c) n_{\text{DW}} n_{\text{DD}} + n_{\text{DD}}^2}{n} - n_{\text{DD}}, \\ \partial_t n_{\text{DW}} - \partial_{xx}^2 n_{\text{DW}} = (1-sh)(\max(r(1-n)(n-a)+1, 0)) \frac{(1+c) n_{\text{WW}} n_{\text{DW}} + 2 n_{\text{WW}} n_{\text{DD}} + \frac{1}{2}(1-c^2) n_{\text{DW}}^2 + (1-c) n_{\text{DW}} n_{\text{DD}}}{n} - n_{\text{DW}}, \\ \partial_t n_{\text{WW}} - \partial_{xx}^2 n_{\text{WW}} = (\max(r(1-n)(n-a)+1, 0)) \frac{n_{\text{WW}}^2 + (1-c) n_{\text{WW}} n_{\text{DW}} + \frac{1}{4}(1-c)^2 n_{\text{DW}}^2}{n} - n_{\text{WW}}. \end{array} \right. \quad (\text{C.2})$$

Model \mathcal{DN} :

$$\left\{ \begin{array}{l} \partial_t n_{\text{DD}} - \partial_{xx}^2 n_{\text{DD}} = (1-s)(r+1) \frac{\frac{1}{4}(1+c)^2 n_{\text{DW}}^2 + (1+c) n_{\text{DW}} n_{\text{DD}} + n_{\text{DD}}^2}{n} - (r n + 1) n_{\text{DD}}, \\ \partial_t n_{\text{DW}} - \partial_{xx}^2 n_{\text{DW}} = (1-sh)(r+1) \frac{(1+c) n_{\text{WW}} n_{\text{DW}} + 2 n_{\text{WW}} n_{\text{DD}} + \frac{1}{2}(1-c^2) n_{\text{DW}}^2 + (1-c) n_{\text{DW}} n_{\text{DD}}}{n} - (r n + 1) n_{\text{DW}}, \\ \partial_t n_{\text{WW}} - \partial_{xx}^2 n_{\text{WW}} = (r+1) \frac{n_{\text{WW}}^2 + (1-c) n_{\text{WW}} n_{\text{DW}} + \frac{1}{4}(1-c)^2 n_{\text{DW}}^2}{n} - (r n + 1) n_{\text{WW}}. \end{array} \right. \quad (\text{C.3})$$

Model \mathcal{DA} :

$$\left\{ \begin{array}{l} \partial_t n_{\text{DD}} - \partial_{xx}^2 n_{\text{DD}} = (1-s)(r+1) \frac{\frac{1}{4}(1+c)^2 n_{\text{DW}}^2 + (1+c) n_{\text{DW}} n_{\text{DD}} + n_{\text{DD}}^2}{n} - (r(n-1)(n-a) + r+1) n_{\text{DD}}, \\ \partial_t n_{\text{DW}} - \partial_{xx}^2 n_{\text{DW}} = (1-sh)(r+1) \frac{(1+c) n_{\text{WW}} n_{\text{DW}} + 2 n_{\text{WW}} n_{\text{DD}} + \frac{1}{2}(1-c^2) n_{\text{DW}}^2 + (1-c) n_{\text{DW}} n_{\text{DD}}}{n} - (r(n-1)(n-a) + r+1) n_{\text{DW}}, \\ \partial_t n_{\text{WW}} - \partial_{xx}^2 n_{\text{WW}} = (r+1) \frac{n_{\text{WW}}^2 + (1-c) n_{\text{WW}} n_{\text{DW}} + \frac{1}{4}(1-c)^2 n_{\text{DW}}^2}{n} - (r(n-1)(n-a) + r+1) n_{\text{WW}}. \end{array} \right. \quad (\text{C.4})$$

Note that all four models reduce to a single model for $r = 0$. This model is:

$$\left\{ \begin{array}{l} \partial_t n_{\text{DD}} - \partial_{xx}^2 n_{\text{DD}} = (1-s) \frac{\frac{1}{4}(1+c)^2 n_{\text{DW}}^2 + (1+c) n_{\text{DW}} n_{\text{DD}} + n_{\text{DD}}^2}{n} - n_{\text{DD}}, \\ \partial_t n_{\text{DW}} - \partial_{xx}^2 n_{\text{DW}} = (1-sh) \frac{(1+c) n_{\text{WW}} n_{\text{DW}} + 2 n_{\text{WW}} n_{\text{DD}} + \frac{1}{2}(1-c^2) n_{\text{DW}}^2 + (1-c) n_{\text{DW}} n_{\text{DD}}}{n} - n_{\text{DW}}, \\ \partial_t n_{\text{WW}} - \partial_{xx}^2 n_{\text{WW}} = \frac{n_{\text{WW}}^2 + (1-c) n_{\text{WW}} n_{\text{DW}} + \frac{1}{4}(1-c)^2 n_{\text{DW}}^2}{n} - n_{\text{WW}}. \end{array} \right. \quad (\text{C.5})$$

C.2 Allelic densities

For our analysis, it is convenient to introduce the allelic (half-) densities ($n_{\text{D}}, n_{\text{W}}$). For a conversion occurring in the germline, we have $n_{\text{D}} = n_{\text{DD}} + \frac{1+c}{2} n_{\text{DW}}$ and $n_{\text{W}} = n_{\text{WW}} + (1 - \frac{1+c}{2}) n_{\text{DW}}$ (see section 3.2 in [28] for more details). We deduce the following systems.

Model $\mathcal{B}\mathcal{N}$:

$$\begin{cases} \partial_t n_D - \partial_{xx}^2 n_D = n_D \left[\frac{r(1-n)+1}{n} \left[(1-s)n_D + (1-sh)(1+c)n_W \right] - 1 \right], \\ \partial_t n_W - \partial_{xx}^2 n_W = n_W \left[\frac{r(1-n)+1}{n} \left[n_W + (1-sh)(1-c)n_D \right] - 1 \right]. \end{cases} \quad (\text{C.6})$$

Model $\mathcal{B}\mathcal{A}$:

$$\begin{cases} \partial_t n_D - \partial_{xx}^2 n_D = n_D \left[\frac{\max(r(1-n)(n-a)+1, 0)}{n} \left[(1-s)n_D + (1-sh)(1+c)n_W \right] - 1 \right], \\ \partial_t n_W - \partial_{xx}^2 n_W = n_W \left[\frac{\max(r(1-n)(n-a)+1, 0)}{n} \left[n_W + (1-sh)(1-c)n_D \right] - 1 \right]. \end{cases} \quad (\text{C.7})$$

Model $\mathcal{D}\mathcal{N}$:

$$\begin{cases} \partial_t n_D - \partial_{xx}^2 n_D = n_D \left[\frac{r+1}{n} \left[(1-s)n_D + (1-sh)(1+c)n_W \right] - (rn+1) \right], \\ \partial_t n_W - \partial_{xx}^2 n_W = n_W \left[\frac{r+1}{n} \left[n_W + (1-sh)(1-c)n_D \right] - (rn+1) \right]. \end{cases} \quad (\text{C.8})$$

Model $\mathcal{D}\mathcal{A}$:

$$\begin{cases} \partial_t n_D - \partial_{xx}^2 n_D = n_D \left[\frac{r+1}{n} \left[(1-s)n_D + (1-sh)(1+c)n_W \right] - (r(n-1)(n-a)+r+1) \right], \\ \partial_t n_W - \partial_{xx}^2 n_W = n_W \left[\frac{r+1}{n} \left[n_W + (1-sh)(1-c)n_D \right] - (r(n-1)(n-a)+r+1) \right]. \end{cases} \quad (\text{C.9})$$

C.3 Allelic frequencies

It may sometimes be more appropriate to write the models in terms of the proportion of drive allele $p_D = \frac{n_D}{n_D + n_W}$ and total population size $n = n_D + n_W$. The models become the following.

Model $\mathcal{B}\mathcal{N}$:

$$\begin{cases} \partial_t n - \partial_{xx}^2 n = (r(1-n)+1) \left((1-s)p_D^2 + 2(1-sh)p_D(1-p_D) + (1-p_D)^2 \right) n - n, \\ \partial_t p_D - \partial_{xx}^2 p_D = 2 \partial_x \log(n) \partial_x p_D + (r(1-n)+1) \left((2h-1)s p_D + (1-sh)(1+c) - 1 \right) p_D (1-p_D). \end{cases} \quad (\text{C.10})$$

Model $\mathcal{B}\mathcal{A}$:

$$\begin{cases} \partial_t n - \partial_{xx}^2 n = (\max(r(1-n)(n-a)+1, 0)) \left((1-s)p_D^2 + 2(1-sh)p_D(1-p_D) + (1-p_D)^2 \right) n - n, \\ \partial_t p_D - \partial_{xx}^2 p_D = 2 \partial_x \log(n) \partial_x p_D + (\max(r(1-n)(n-a)+1, 0)) \left((2h-1)s p_D + (1-sh)(1+c) - 1 \right) p_D (1-p_D). \end{cases} \quad (\text{C.11})$$

Model $\mathcal{D}\mathcal{N}$:

$$\begin{cases} \partial_t n - \partial_{xx}^2 n = (r+1) \left((1-s)p_D^2 + 2(1-sh)p_D(1-p_D) + (1-p_D)^2 \right) n - (rn+1) n, \\ \partial_t p_D - \partial_{xx}^2 p_D = 2 \partial_x \log(n) \partial_x p_D + (r+1) \left((2h-1)s p_D + (1-sh)(1+c) - 1 \right) p_D (1-p_D). \end{cases} \quad (\text{C.12})$$

Model $\mathcal{D}\mathcal{A}$:

$$\begin{cases} \partial_t n - \partial_{xx}^2 n = (r+1) \left((1-s)p_D^2 + 2(1-sh)p_D(1-p_D) + (1-p_D)^2 \right) n - (r(n-1)(n-a)+r+1) n, \\ \partial_t p_D - \partial_{xx}^2 p_D = 2 \partial_x \log(n) \partial_x p_D + (r+1) \left((2h-1)s p_D + (1-sh)(1+c) - 1 \right) p_D (1-p_D). \end{cases} \quad (\text{C.13})$$

Equations on p_D differ from the standard equation often used in population genetics, as they contain an advection term $2 \partial_x(\log n) \partial_x p_D$. This term appears when calculating $\partial_{xx}^2 p_D = \partial_{xx}^2 \frac{n_{DD}}{n}$ and represents a demographic flux from denser to less dense areas, due to variations in population density. It is opposed to the spread of the costly drive allele (see Figure 2 [11]).

C.4 Final allelic proportions for r small and large

In models \mathcal{BN} and \mathcal{BA} for large values of r , using the Strugarek-Vauchelet rescaling [62] in (C.10) and (C.11), the systems reduce to one limit equation on p_D :

$$\partial_t p_D - \partial_{xx}^2 p_D = \frac{\left((2h-1) s p_D + (1-sh)(1+c)-1 \right) p_D (1-p_D)}{(1-s) p_D^2 + 2(1-sh) p_D (1-p_D) + (1-p_D)^2 - 1}. \quad (\text{C.14})$$

Equation (C.14) has already been studied in [18]: a heatmap illustrates the final proportions in the case $c = 0.85$ (Figure 4 in [18]). In models \mathcal{DN} and \mathcal{DA} , the equation on p_D in (C.10) and (C.11) is:

$$\partial_t p_D - \partial_{xx}^2 p_D = 2 \partial_x \log(n) \partial_x p_D + \underbrace{(r+1) \left((2h-1) s p_D + (1-sh)(1+c)-1 \right) p_D (1-p_D)}_{\text{reaction term}}. \quad (\text{C.15})$$

The reaction term in equation (C.15) becomes larger as r increases: this indicates that the traveling wave has an infinite speed when r tends to infinity, meaning that the equilibrium is reached instantaneously. Therefore, the term $2 \partial_x \log(n) \partial_x p_D$ is instantaneously zero and the final proportions are the same as for models \mathcal{BN} and \mathcal{BA} .

As a consequence, all models \mathcal{BN} , \mathcal{BA} , \mathcal{DN} and \mathcal{DA} share the same final proportions for large values of r . This conclusion also holds for $r = 0$, as the models are equal (see Appendix C.1). These proportions have already been determined in a previous article [28], for the \mathcal{BN} case. We recall these results below in section F and generalise them to our four models.

C.5 Speed of the problem simplified at low drive density

In Section 3.3, we focus on drive invasion and therefore consider low drive density and high wild-type density at the front of the wave. The speed v of the wave can be calculated when the models are simplified (linearised) at low drive density: it is deduced from the reproduction of the few drive individuals at the front wave,

$$v = 2 \sqrt{\lim_{n_D \rightarrow 0} \left(\frac{F(n_D)}{n_D} \right)}, \quad (\text{C.16})$$

where F represents the net production of drive alleles. The formulas for F in the different models are the following.

Model \mathcal{BN} :

$$F^{(\mathcal{BN})}(n_D) = n_D \left[(r(1-n)+1) \left[(1-s) \frac{n_D}{n} + (1-sh)(1+c) \frac{n_w}{n} \right] - 1 \right] \quad (\text{C.17})$$

Model \mathcal{BA} :

$$F^{(\mathcal{BA})}(n_D) = n_D \left[\max(r(1-n)(n-a)+1, 0) \left[(1-s) \frac{n_D}{n} + (1-sh)(1+c) \frac{n_w}{n} \right] - 1 \right]. \quad (\text{C.18})$$

Model \mathcal{DN} :

$$F^{(\mathcal{DN})}(n_D) = n_D \left[(r+1) \left[(1-s) \frac{n_D}{n} + (1-sh)(1+c) \frac{n_w}{n} \right] - (rn+1) \right]. \quad (\text{C.19})$$

Model \mathcal{DA} :

$$F^{(\mathcal{DA})}(n_D) = n_D \left[(r+1) \left[(1-s) \frac{n_D}{n} + (1-sh)(1+c) \frac{n_w}{n} \right] - (r(n-1)(n-a)+r+1) \right]. \quad (\text{C.20})$$

Considering high wild-type density at the front of the wave ($n_w \approx n \approx 1$), the speed in models \mathcal{BN} and \mathcal{BA} is given by:

$$v_{\mathcal{BN}, \mathcal{BA}} = 2 \sqrt{(1-sh)(1+c)-1}. \quad (\text{C.21})$$

In models \mathcal{DN} and \mathcal{DA} , we have:

$$v_{\mathcal{DN}, \mathcal{DA}} = 2 \sqrt{(1+r)[(1-sh)(1+c)-1]}. \quad (\text{C.22})$$

To understand why $v_{\mathcal{D},\mathcal{N},\mathcal{D},\mathcal{A}}$ is greater by a coefficient of $\sqrt{1+r}$ than $v_{\mathcal{B},\mathcal{N},\mathcal{B},\mathcal{A}}$, we have to understand the population dynamics at the front of the wave. There, the density is close to the maximum carrying capacity 1, with low drive density ($n_{\text{D}} \approx 0$) and high wild-type density ($n_{\text{W}} \approx n \approx 1$). On one hand, in models \mathcal{B},\mathcal{N} and \mathcal{B},\mathcal{A} , the density-dependence constraint is placed on the birth term, reducing the production rate of drive alleles to $(1-s)h(1+c)$ while individuals disappear at rate 1 (C.17, C.18). On the other hand, in models \mathcal{D},\mathcal{N} and \mathcal{D},\mathcal{A} , the density-dependence constraint is placed on the death term increasing to $(r+1)$ the rate at which the drive alleles disappears, while they are produced at rate $(r+1)(1-sh)(1+c)$ (C.19, C.20). Consequently, the net production remains constant $(1-sh)(1+c)$, but the turnover rate is $r+1$ times greater. As the wave movement largely relies on the reproduction, this reflects in the speed formula: the propagation is $\sqrt{r+1}$ times faster.

D Comparison of the conditions leading to eradication

To compare the conditions leading to eradication, we refer to the results summarized in Table 3.

In Models \mathcal{B},\mathcal{N} and \mathcal{D},\mathcal{N} , eradication occurs when:

$$r\mathcal{F}(p_{\text{D}}^*) < 1 - \mathcal{F}(p_{\text{D}}^*), \quad (\text{D.1})$$

in Model \mathcal{B},\mathcal{A} , when:

$$\left(\frac{1-a}{2}\right)^2 r\mathcal{F}(p_{\text{D}}^*) < 1 - \mathcal{F}(p_{\text{D}}^*). \quad (\text{D.2})$$

and in Model \mathcal{D},\mathcal{A} , when:

$$r\mathcal{F}(p_{\text{D}}^*) - r\left[1 - \left(\frac{1-a}{2}\right)^2\right] < 1 - \mathcal{F}(p_{\text{D}}^*), \quad (\text{D.3})$$

With $\left(\frac{1-a}{2}\right)^2 \in [0, 1]$ and $\mathcal{F}(p_{\text{D}}^*) \in [1-s, 1]$, we obtain the following inequalities:

$$r\mathcal{F}(p_{\text{D}}^*) - r\left[1 - \left(\frac{1-a}{2}\right)^2\right] \leq \left(\frac{1-a}{2}\right)^2 r\mathcal{F}(p_{\text{D}}^*) \leq r\mathcal{F}(p_{\text{D}}^*) \quad (\text{D.4})$$

Inequalities (D.2) and (D.3) both imply (D.1). In other words, eradication is easier to achieve in Models \mathcal{B},\mathcal{A} and \mathcal{D},\mathcal{A} (with Allee effect) than in Models \mathcal{B},\mathcal{N} and \mathcal{D},\mathcal{N} (without Allee effect) in the sense that the range of parameters leading to eradication is larger.

Inequality (D.3) also implies (D.2), i.e., in models with Allee effect, for a given a value, the density-dependence constraint on the deaths (Model \mathcal{D},\mathcal{A}) makes the eradication easier than if it is on the births (Model \mathcal{B},\mathcal{A}).

Note that when $a = -1$, the conditions leading to eradication are equivalent in all four models. When a becomes larger, the range of parameters leading to eradication in models \mathcal{B},\mathcal{A} and \mathcal{D},\mathcal{A} also becomes wider.

E Comparison of the conditions leading to persistence

To compare the conditions leading to persistence, we refer to the results summarized in Table 3.

In Models \mathcal{B},\mathcal{N} and \mathcal{D},\mathcal{N} , persistence occurs when:

$$r\mathcal{F}(p_{\text{D}}^*) > 1 - \mathcal{F}(p_{\text{D}}^*), \quad (\text{E.1})$$

in Model \mathcal{B},\mathcal{A} , when:

$$\begin{cases} \left(\frac{1-a}{2}\right)^2 r\mathcal{F}(p_{\text{D}}^*) > 1 - \mathcal{F}(p_{\text{D}}^*), \\ -a r\mathcal{F}(p_{\text{D}}^*) > 1 - \mathcal{F}(p_{\text{D}}^*), \end{cases} \quad (\text{E.2})$$

and in Model \mathcal{D},\mathcal{A} , when:

$$\begin{cases} r\mathcal{F}(p_{\text{D}}^*) - r\left[1 - \left(\frac{1-a}{2}\right)^2\right] > 1 - \mathcal{F}(p_{\text{D}}^*), \\ r\mathcal{F}(p_{\text{D}}^*) - r(a+1) > 1 - \mathcal{F}(p_{\text{D}}^*). \end{cases} \quad (\text{E.3})$$

First note that persistence is only possible in case of a weak Allee effect ($a < 0$) in Models $\mathcal{B.A}$ and $\mathcal{D.A}$ (second lines in Systems (E.2) and (E.3)). With $(\frac{1-a}{2})^2 \in [0, 1]$ and $\mathcal{F}(p_{\text{D}}^*) \in [1-s, 1]$, we obtain the following inequalities:

$$r\mathcal{F}(p_{\text{D}}^*) \geq \left(\frac{1-a}{2}\right)^2 r\mathcal{F}(p_{\text{D}}^*) \geq r\mathcal{F}(p_{\text{D}}^*) - r\left[1 - \left(\frac{1-a}{2}\right)^2\right] \quad (\text{E.4})$$

$$r\mathcal{F}(p_{\text{D}}^*) \geq -ar\mathcal{F}(p_{\text{D}}^*) \geq r\mathcal{F}(p_{\text{D}}^*) - r(1+a) \quad (\text{E.5})$$

Inequality (E.1) imply both (E.2) and (E.3). In other words, persistence is easier to achieve in Models $\mathcal{B.N}$ and $\mathcal{D.N}$ (without Allee effect) than in Models $\mathcal{B.A}$ and $\mathcal{D.A}$ (with Allee effect) in the sens that the range of parameters leading to persistence is larger.

Inequality (E.2) also implies (E.3), i.e., in models with Allee effect, for a given a value, the density-dependence constraint on the births (Model $\mathcal{B.A}$) makes the persistence easier than if it is on the deaths (Model $\mathcal{D.A}$).

Note that when $a = -1$, the conditions leading to persistence are equivalent in all four models. When a becomes larger, the range of parameters leading to persistence in models $\mathcal{B.A}$ and $\mathcal{D.A}$ becomes smaller. In case of a strong Allee effect ($a > 0$), the "persistence" regime disappears.

F Comparison of the final densities in case of persistence

F.1 Comparison of the final density in models $\mathcal{B.N}$ and $\mathcal{D.N}$

We compare the final densities for models $\mathcal{B.N}$ and $\mathcal{D.N}$ given in Table 3 in case of persistence.

$$n^{*(\mathcal{D.N})} = 1 - \frac{(r+1)(1-\mathcal{F}(p_{\text{D}}^*))}{r} = 1 - \frac{-r\mathcal{F}(p_{\text{D}}^*) + r + 1 - \mathcal{F}(p_{\text{D}}^*)}{r} = \mathcal{F}(p_{\text{D}}^*) - \frac{1 - \mathcal{F}(p_{\text{D}}^*)}{r} = \mathcal{F}(p_{\text{D}}^*) n^{*(\mathcal{B.N})} \quad (\text{E.1})$$

with $\mathcal{F}(p_{\text{D}}^*) \in [1-s, 1]$. The final density is $\mathcal{F}(p_{\text{D}}^*)$ times lower in model $\mathcal{D.N}$ than in model $\mathcal{B.N}$, in case of persistence. Note that $\mathcal{F}(p_{\text{D}}^*) = 1-s$ in case of a drive invasion, and $\mathcal{F}(p_{\text{D}}^*) \in (1-s, 1)$ in case of coexistence.

F.2 Comparison of the final density in models $\mathcal{B.A}$ and $\mathcal{D.A}$

We compare the final densities for model $\mathcal{B.A}$ and $\mathcal{D.A}$ in case of persistence given in Table 3.

$$n^{*(\mathcal{B.A})} = \frac{1+a+\sqrt{(1+a)^2-4(a+1-n^{*(\mathcal{B.N})})}}{2} \quad \text{and} \quad n^{*(\mathcal{D.A})} = \frac{1+a+\sqrt{(1+a)^2-4(a+1-n^{*(\mathcal{D.N})})}}{2} \quad (\text{E.2})$$

From the previous section F.1, we know that $n^{*(\mathcal{D.N})} = \mathcal{F}(p_{\text{D}}^*) n^{*(\mathcal{B.N})}$ with $\mathcal{F}(p_{\text{D}}^*) \in [1-s, 1]$. Therefore:

$$n^{*(\mathcal{D.A})} \leq n^{*(\mathcal{B.A})} \quad (\text{E.3})$$

This inequality is strict in case of a drive invasion $\mathcal{F}(p_{\text{D}}^*) = 1-s$ or in case of coexistence $\mathcal{F}(p_{\text{D}}^*) \in (1-s, 1)$. We derive both density as functions of a , in case of drive persistence $\mathcal{F}(p_{\text{D}}^*) < 1$:

$$\partial_a n^{*(\mathcal{B.A})} = 1 - \frac{2(1-a)}{2\sqrt{(1-a)^2-4(1-n^{*(\mathcal{B.N})})}} < 0 \quad \text{and} \quad \partial_a n^{*(\mathcal{D.A})} = 1 - \frac{2(1-a)}{2\sqrt{(1-a)^2-4(1-n^{*(\mathcal{D.N})})}} < 0, \quad (\text{E.4})$$

as $n^{*(\mathcal{B.N})} < 1$ and $n^{*(\mathcal{D.N})} < 1$ when $\mathcal{F}(p_{\text{D}}^*) < 1$ (see Table 3). In other words in case of a drive persistence, the stronger the Allee effect, the smaller the final population density in models $\mathcal{B.A}$ and $\mathcal{D.A}$.

G Final population density when the strength of the Allee effect varies, in case of drive invasion

The “pure drive” areas described in Figure 2 correspond to the final density expected in case of a drive invasion: persistence ($n^* = n^+$), eradication ($n^* = 0$) or bistability (either $n^* = n^+$ or $n^* = 0$ depending on the initial conditions). These final densities are given in Table 3 with $p_D^* = 1$ (the proportion of drive allele in the final population is one in case of a drive invasion without coexistence). Without Allee effect, there only exists pure drive eradication and persistence areas (Models $\mathcal{B}\mathcal{N}$ and $\mathcal{D}\mathcal{N}$). However, if we consider an Allee effect (Models $\mathcal{B}\mathcal{A}$ and $\mathcal{D}\mathcal{A}$), a pure drive bistable area appears. In Figure S3, we plot the boundary lines delimiting the pure drive areas in models with Allee effect, for different values of a . The first boundary line, delimiting the pure drive persistence area (above it) from the pure drive bistable area (under it), will be referred to as the *upper boundary line* and is represented with a dash-dotted line (----). The second boundary line delimiting the pure drive bistable area (above it) from the pure drive eradication area (under it), will be referred to as the *under boundary line* and drawn with a dashed line (---).

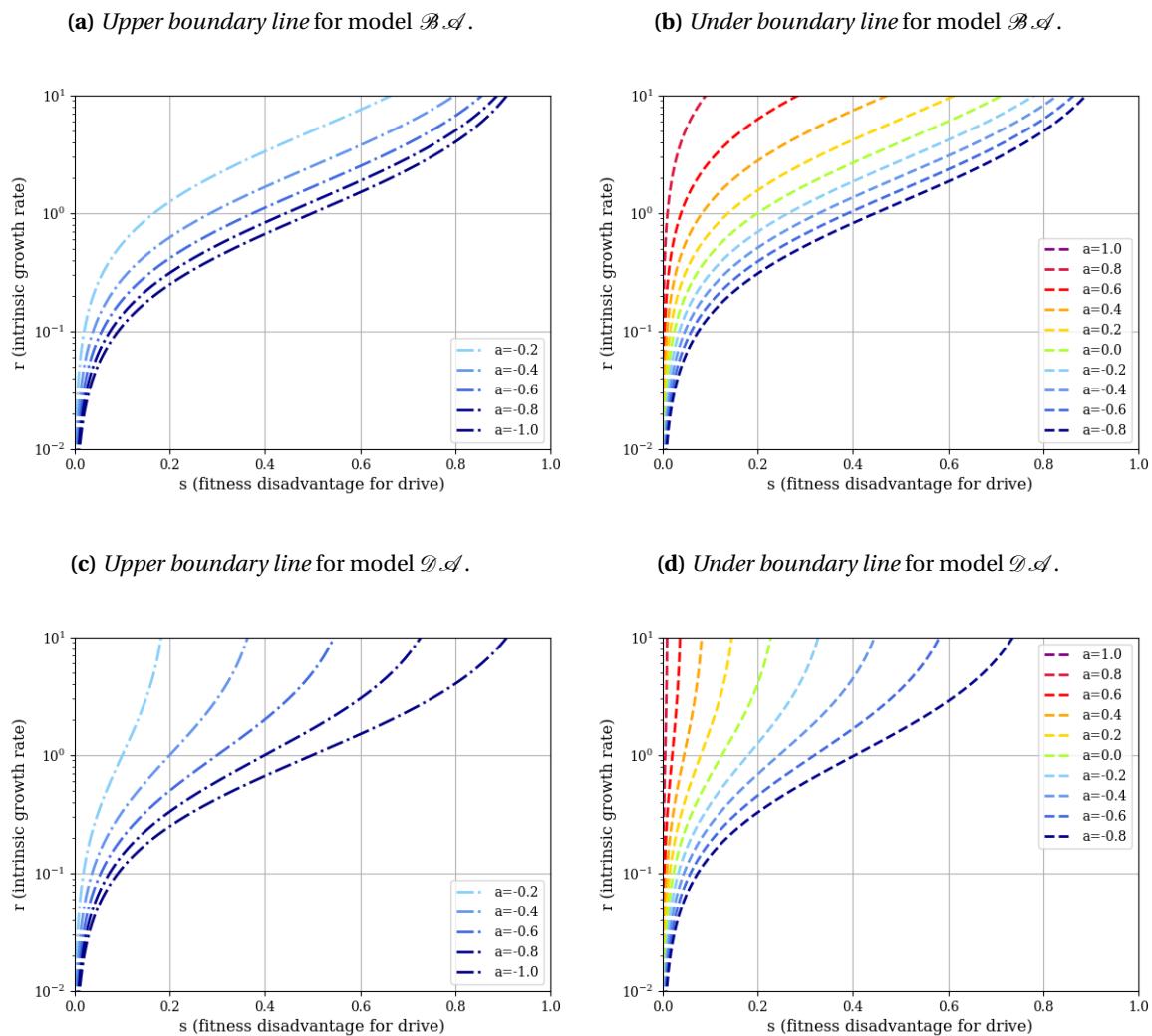


Figure S3: Boundary lines for different values of a in models with Allee effect $\mathcal{B}\mathcal{A}$ and $\mathcal{D}\mathcal{A}$. The *upper boundary line* plotted on panels (a) and (c), delimits the pure drive persistence area (above it) from the pure drive bistable area (under it). It is represented with a dash-dotted line. The *under boundary line* plotted on panels (b) and (d), delimits the pure drive bistable area (above it) from the pure drive eradication area (under it). It is represented with a dashed line. These boundaries were represented for two values of a in Figures 2 and 3 ($a = -0.2$ and $a = 0.2$).

The higher a , the larger the pure drive eradication area and the more the pure drive persistence and bistable areas are restricted to high values of r and small values of s . In case of a strong Allee effect ($0 < a < 1$), the pure drive persistence area even disappears; this is why only negative values of a are plotted in Figure S3 (a) and (c).

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work is funded by ANR-19-CE45-0009-01 TheoGeneDrive. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 865711).

We are grateful to the INRAE MIGALE bioinformatics facility (MIGALE, INRAE, 2020. Migale bioinformatics Facility, doi: 10.15454/1.5572390655343293E12) for providing computing resources

References

- [1] Luke S. Alphey et al. "Standardizing the Definition of Gene Drive". In: *Proceedings of the National Academy of Sciences* 117.49 (Dec. 2020), pp. 30864–30867. DOI: 10.1073/pnas.2020417117.
- [2] Austin Burt. "Site-Specific Selfish Genes as Tools for the Control and Genetic Engineering of Natural Populations." In: *Proceedings of the Royal Society B: Biological Sciences* 270.1518 (May 2003), pp. 921–928. ISSN: 0962-8452. DOI: 10.1098/rspb.2002.2319.
- [3] Austin Burt and Andrea Crisanti. "Gene Drive: Evolved and Synthetic". In: *ACS Chemical Biology* 13.2 (Feb. 2018), pp. 343–346. ISSN: 1554-8929. DOI: 10.1021/acschembio.7b01031.
- [4] James E. DiCarlo et al. "Safeguarding CRISPR-Cas9 Gene Drives in Yeast". In: *Nature Biotechnology* 33.12 (Dec. 2015), pp. 1250–1255. ISSN: 1546-1696. DOI: 10.1038/nbt.3412.
- [5] Silke Fuchs et al. "Resistance to a CRISPR-based Gene Drive at an Evolutionarily Conserved Site Is Revealed by Mimicking Genotype Fixation". In: *PLOS Genetics* 17.10 (Oct. 2021), e1009740. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1009740.
- [6] Emily Yang et al. "A Homing Suppression Gene Drive with Multiplexed gRNAs Maintains High Drive Conversion Efficiency and Avoids Functional Resistance Alleles". In: *G3 Genes|Genomes|Genetics* 12.6 (June 2022), jkac081. ISSN: 2160-1836. DOI: 10.1093/g3journal/jkac081.
- [7] Jackson Champer et al. "Novel CRISPR/Cas9 Gene Drive Constructs Reveal Insights into Mechanisms of Resistance Allele Formation and Drive Efficiency in Genetically Diverse Populations". In: *PLOS Genetics* 13.7 (July 2017), e1006796. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1006796.
- [8] Kyros Kyrou et al. "A CRISPR–Cas9 Gene Drive Targeting Doublesex Causes Complete Population Suppression in Caged *Anopheles Gambiae* Mosquitoes". In: *Nature Biotechnology* 36.11 (Nov. 2018), pp. 1062–1066. ISSN: 1546-1696. DOI: 10.1038/nbt.4245.
- [9] Andrew Hammond et al. "Gene-Drive Suppression of Mosquito Populations in Large Cages as a Bridge between Lab and Field". In: *Nature Communications* 12.1 (July 2021), p. 4589. ISSN: 2041-1723. DOI: 10.1038/s41467-021-24790-6.
- [10] H. Charles J. Godfray, Ace North, and Austin Burt. "How Driving Endonuclease Genes Can Be Used to Combat Pests and Disease Vectors". In: *BMC Biology* 15.1 (Sept. 2017), p. 81. ISSN: 1741-7007. DOI: 10.1186/s12915-017-0420-4.
- [11] Léo Girardin and Florence Débarre. "Demographic Feedbacks Can Hamper the Spatial Spread of a Gene Drive". In: *Journal of Mathematical Biology* 83.6 (Dec. 2021), p. 67. ISSN: 1432-1416. DOI: 10.1007/s00285-021-01702-2.
- [12] Ethan Bier. "Gene Drives Gaining Speed". In: *Nature Reviews Genetics* 23.1 (Jan. 2022), pp. 5–22. ISSN: 1471-0064. DOI: 10.1038/s41576-021-00386-0.

- [13] Bruce A. Hay, Georg Oberhofer, and Ming Guo. "Engineering the Composition and Fate of Wild Populations with Gene Drive". In: *Annual Review of Entomology* 66.1 (2021), pp. 407–434. DOI: 10.1146/annurev-ento-020117-043154.
- [14] Tony Nolan. "Control of Malaria-Transmitting Mosquitoes Using Gene Drives". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 376.1818 (Dec. 2020), p. 20190803. DOI: 10.1098/rstb.2019.0803.
- [15] Valentino Gantz et al. "Highly Efficient Cas9-mediated Gene Drive for Population Modification of the Malaria Vector Mosquito *Anopheles Stephensi*". In: *Proceedings of the National Academy of Sciences* 112 (Nov. 2015). DOI: 10.1073/pnas.1521077112.
- [16] Andrew Hammond et al. "A CRISPR-Cas9 Gene Drive System Targeting Female Reproduction in the Malaria Mosquito Vector *Anopheles Gambiae*". In: *Nature biotechnology* 34 (Dec. 2015). DOI: 10.1038/nbt.3439.
- [17] Kevin M Esvelt et al. "Concerning RNA-guided Gene Drives for the Alteration of Wild Populations". In: *eLife* 3 (July 2014). Ed. by Diethard Tautz, e03401. ISSN: 2050-084X. DOI: 10.7554/eLife.03401.
- [18] Nicolas O. Rode et al. "Population Management Using Gene Drive: Molecular Design, Models of Spread Dynamics and Assessment of Ecological Risks". In: *Conservation Genetics* 20.4 (Aug. 2019), pp. 671–690. ISSN: 1572-9737. DOI: 10.1007/s10592-019-01165-5.
- [19] Bhagyashree Kaduskar et al. "Reversing Insecticide Resistance with Allelic-Drive in *Drosophila Melanogaster*". In: *Nature Communications* 13.1 (Jan. 2022), p. 291. ISSN: 2041-1723. DOI: 10.1038/s41467-021-27654-1.
- [20] Paul Neve. "Gene Drive Systems: Do They Have a Place in Agricultural Weed Management?" In: *Pest Management Science* 74.12 (2018), pp. 2671–2679. ISSN: 1526-4998. DOI: 10.1002/ps.5137.
- [21] Sumit Dhole, Alun L. Lloyd, and Fred Gould. "Gene Drive Dynamics in Natural Populations: The Importance of Density Dependence, Space, and Sex". In: *Annual Review of Ecology, Evolution, and Systematics* 51.1 (Nov. 2020), pp. 505–531. ISSN: 1543-592X. DOI: 10.1146/annurev-ecolsys-031120-101013.
- [22] Anne Deredec, Austin Burt, and H. C. J. Godfray. "The Population Genetics of Using Homing Endonuclease Genes in Vector and Pest Management". In: *Genetics* 179.4 (Aug. 2008), pp. 2013–2026. ISSN: 0016-6731. DOI: 10.1534/genetics.108.089037.
- [23] Robert L Unckless et al. "Modeling the Manipulation of Natural Populations by the Mutagenic Chain Reaction". In: *Genetics* 201.2 (Oct. 2015), pp. 425–431. ISSN: 1943-2631. DOI: 10.1534/genetics.115.177592.
- [24] Tom J. de Jong. "Gene Drives Do Not Always Increase in Frequency: From Genetic Models to Risk Assessment". In: *Journal of Consumer Protection and Food Safety* 12.4 (Dec. 2017), pp. 299–307. ISSN: 1661-5867. DOI: 10.1007/s00003-017-1131-z.
- [25] Hidenori Tanaka, Howard A. Stone, and David R. Nelson. "Spatial Gene Drives and Pushed Genetic Waves". In: *Proceedings of the National Academy of Sciences of the United States of America* 114.32 (Aug. 2017), pp. 8452–8457. ISSN: 1091-6490. DOI: 10.1073/pnas.1705868114.
- [26] Philip T. Leftwich et al. "Recent Advances in Threshold-Dependent Gene Drives for Mosquitoes". In: *Biochemical Society Transactions* 46.5 (Sept. 2018), pp. 1203–1212. ISSN: 0300-5127. DOI: 10.1042/BST20180076.
- [27] Jackson Champer, Anna Buchman, and Omar S. Akbari. "Cheating Evolution: Engineering Gene Drives to Manipulate the Fate of Wild Populations". In: *Nature Reviews Genetics* 17.3 (Mar. 2016), pp. 146–159. ISSN: 1471-0064. DOI: 10.1038/nrg.2015.34.
- [28] Léna Kläy et al. "Pulled, Pushed or Failed: The Demographic Impact of a Gene Drive Can Change the Nature of Its Spatial Spread". In: *Journal of Mathematical Biology* 87.2 (July 2023), p. 30. ISSN: 1432-1416. DOI: 10.1007/s00285-023-01926-4.
- [29] Claus Rueffler, Martijn Egas, and Johan A. J. Metz. "Evolutionary Predictions Should Be Based on Individual-Level Traits." In: *The American Naturalist* 168.5 (Nov. 2006), E148–E162. ISSN: 0003-0147. DOI: 10.1086/508618.

- [30] A. Tsouliaris and J. Wallace. "Analysis of Logistic Growth Models". In: *Mathematical Biosciences* 179.1 (July 2002), pp. 21–55. ISSN: 0025-5564. DOI: 10.1016/S0025-5564(02)00096-2.
- [31] Nicolas O Rode, Virginie Courtier-Orgogozo, and Florence Débarre. "Can a Population Targeted by a CRISPR-Based Homing Gene Drive Be Rescued?" In: *G3 Genes|Genomes|Genetics* 10.9 (Sept. 2020), pp. 3403–3415. ISSN: 2160-1836. DOI: 10.1534/g3.120.401484.
- [32] Gloria M. Luque et al. "The Genetic Allee Effect: A Unified Framework for the Genetics and Demography of Small Populations". In: *Ecosphere* 7.7 (2016), e01413. ISSN: 2150-8925. DOI: 10.1002/ecs2.1413.
- [33] Franck Courchamp, Ludek Berec, and Joanna Gascoigne. *Allee Effects in Ecology and Conservation*. OUP Oxford, Feb. 2008. ISBN: 978-0-19-152466-0.
- [34] Peter Armbruster, Robert A. Hutchinson, and Trevor Linvell. "Equivalent Inbreeding Depression under Laboratory and Field Conditions in a Tree-Hole-Breeding Mosquito". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 267.1456 (Oct. 2000), pp. 1939–1945. DOI: 10.1098/rspb.2000.1233.
- [35] Rowida Baeshen et al. "Differential Effects of Inbreeding and Selection on Male Reproductive Phenotype Associated with the Colonization and Laboratory Maintenance of Anopheles Gambiae". In: *Malaria Journal* 13.1 (Jan. 2014), p. 19. ISSN: 1475-2875. DOI: 10.1186/1475-2875-13-19.
- [36] Perran A. Ross, Nancy M. Endersby-Harshman, and Ary A. Hoffmann. "A Comprehensive Assessment of Inbreeding and Laboratory Adaptation in Aedes Aegypti Mosquitoes". In: *Evolutionary Applications* 12.3 (2019), pp. 572–586. ISSN: 1752-4571. DOI: 10.1111/eva.12740.
- [37] Samuel E Champer et al. "Anopheles Homing Suppression Drive Candidates Exhibit Unexpected Performance Differences in Simulations with Spatial Structure". In: *eLife* 11 (Oct. 2022). Ed. by George H Perry, Sebald ANR Verkuijl, and Jim Bull, e79121. ISSN: 2050-084X. DOI: 10.7554/eLife.79121.
- [38] N. H. Barton and Michael Turelli. "Spatial Waves of Advance with Bistable Dynamics: Cytoplasmic and Genetic Analogues of Allee Effects." In: *The American Naturalist* 178.3 (Sept. 2011), E48–E75. ISSN: 0003-0147. DOI: 10.1086/661246.
- [39] R. A. Fisher. "The Wave of Advance of Advantageous Genes". In: *Annals of Eugenics* 7.4 (1937), pp. 355–369. ISSN: 2050-1439. DOI: 10.1111/j.1469-1809.1937.tb02153.x.
- [40] A.N. Kolmogorov, I.G Petrovsky, and N.S. Piskunov. "Étude de l'équation de La Diffusion Avec Croissance de La Quantité de Matière et Son Application à Un Problème Biologique". In: *Moscow Univ. Bull. Ser. Internat. Sect. A* 1 (1937), p. 1.
- [41] D. G. Aronson and Hans F Weinberger. "Multidimensional Nonlinear Diffusion Arising in Population Genetics". In: *Advances in Mathematics* 30.1 (Oct. 1978), pp. 33–76. ISSN: 0001-8708. DOI: 10.1016/0001-8708(78)90130-5.
- [42] Frederik J. H. de Haas et al. "Modelling daisy quorum drive: A short-term bridge across engineered fitness valleys". In: *PLOS Genetics* 20.5 (May 2024), pp. 1–26. DOI: 10.1371/journal.pgen.1011262.
- [43] Ace North, Austin Burt, and H. Charles J. Godfray. "Modelling the Spatial Spread of a Homing Endonuclease Gene in a Mosquito Population". In: *Journal of Applied Ecology* 50.5 (2013), pp. 1216–1225. ISSN: 1365-2664. DOI: 10.1111/1365-2664.12133.
- [44] Philip A. Eckhoff et al. "Impact of Mosquito Gene Drive on Malaria Elimination in a Computational Model with Explicit Spatial and Temporal Dynamics". In: *Proceedings of the National Academy of Sciences* 114.2 (Jan. 2017), E255–E264. DOI: 10.1073/pnas.1611064114.
- [45] Jackson Champer et al. "Suppression Gene Drive in Continuous Space Can Result in Unstable Persistence of Both Drive and Wild-Type Alleles". In: *Molecular Ecology* 30.4 (Feb. 2021), pp. 1086–1101. ISSN: 1365-294X. DOI: 10.1111/mec.15788.
- [46] Ace R. North, Austin Burt, and H. Charles J. Godfray. "Modelling the Potential of Genetic Control of Malaria Mosquitoes at National Scale". In: *BMC Biology* 17.1 (Mar. 2019), p. 26. ISSN: 1741-7007. DOI: 10.1186/s12915-019-0645-5.
- [47] Ace R. North, Austin Burt, and H. Charles J. Godfray. "Modelling the Suppression of a Malaria Vector Using a CRISPR-Cas9 Gene Drive to Reduce Female Fertility". In: *BMC Biology* 18 (Aug. 2020), p. 98. ISSN: 1741-7007. DOI: 10.1186/s12915-020-00834-z.

- [48] Andrea K. Beaghton et al. “Gene Drive for Population Genetic Control: Non-Functional Resistance and Parental Effects”. In: *Proceedings of the Royal Society B: Biological Sciences* 286.1914 (Oct. 2019), p. 20191586. DOI: 10.1098/rspb.2019.1586.
- [49] Andrew M. Hammond et al. “The Creation and Selection of Mutations Resistant to a Gene Drive over Multiple Generations in the Malaria Mosquito”. In: *PLOS Genetics* 13.10 (Oct. 2017), e1007039. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1007039.
- [50] Tom A. R. Price et al. “Resistance to Natural and Synthetic Gene Drive Systems”. In: *Journal of Evolutionary Biology* 33.10 (Oct. 2020), pp. 1345–1360. ISSN: 1010-061X. DOI: 10.1111/jeb.13693.
- [51] Andri Manser et al. “Controlling Invasive Rodents via Synthetic Gene Drive and the Role of Polyandry”. In: *Proceedings of the Royal Society B: Biological Sciences* 286.1909 (Aug. 2019), p. 20190852. DOI: 10.1098/rspb.2019.0852.
- [52] Andri Manser, Barbara König, and Anna K. Lindholm. “Polyandry Blocks Gene Drive in a Wild House Mouse Population”. In: *Nature Communications* 11.1 (Nov. 2020), p. 5590. ISSN: 2041-1723. DOI: 10.1038/s41467-020-18967-8.
- [53] Aysegul Birand et al. “Gene Drives for Vertebrate Pest Control: Realistic Spatial Modelling of Eradication Probabilities and Times for Island Mouse Populations”. In: *Molecular Ecology* 31.6 (Mar. 2022), pp. 1907–1923. ISSN: 0962-1083. DOI: 10.1111/mec.16361.
- [54] John M. Marshall and Bruce A. Hay. “Medusa: A Novel Gene Drive System for Confined Suppression of Insect Populations”. In: *PLOS ONE* 9.7 (July 2014), e102694. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0102694.
- [55] Héctor M. Sánchez C. et al. “MGDrivE: A Modular Simulation Framework for the Spread of Gene Drives through Spatially Explicit Mosquito Populations”. In: *Methods in Ecology and Evolution* 11.2 (2020), pp. 229–239. ISSN: 2041-210X. DOI: 10.1111/2041-210X.13318.
- [56] Samuel E. Champer et al. *Finding the Strongest Gene Drive: Simulations Reveal Unexpected Performance Differences between Anopheles Homing Suppression Drive Candidates*. Mar. 2022. DOI: 10.1101/2022.03.28.486009.
- [57] Yiran Liu et al. “Adversarial Interspecies Relationships Facilitate Population Suppression by Gene Drive in Spatially Explicit Models”. In: *Ecology Letters* 26.7 (2023), pp. 1174–1185. ISSN: 1461-0248. DOI: 10.1111/ele.14232.
- [58] Jun Li et al. “Can CRISPR Gene Drive Work in Pest and Beneficial Haplodiploid Species?” In: *Evolutionary Applications* 13.9 (2020), pp. 2392–2403. ISSN: 1752-4571. DOI: 10.1111/eva.13032.
- [59] Andrea Beaghton, Pantelis John Beaghton, and Austin Burt. “Gene Drive through a Landscape: Reaction–Diffusion Models of Population Suppression and Elimination by a Sex Ratio Distorter”. In: *Theoretical Population Biology* 108 (Apr. 2016), pp. 51–69. ISSN: 0040-5809. DOI: 10.1016/j.tpb.2015.11.005.
- [60] John B. Connolly et al. “Gene Drive in Species Complexes: Defining Target Organisms”. In: *Trends in Biotechnology* 41.2 (Feb. 2023), pp. 154–164. ISSN: 0167-7799. DOI: 10.1016/j.tibtech.2022.06.013.
- [61] National Academies of Sciences, Engineering, and Medicine. *Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values*. Washington (DC): National Academies Press (US), 2016. ISBN: 978-0-309-43787-5.
- [62] Martin Strugarek and Nicolas Vauchelet. “Reduction to a Single Closed Equation for 2-by-2 Reaction–Diffusion Systems of Lotka–Volterra Type”. In: *SIAM Journal on Applied Mathematics* 76.5 (Jan. 2016), pp. 2060–2080. ISSN: 0036-1399. DOI: 10.1137/16M1059217.