A unifying approach to gene drive

Prateek Verma[1], R. Guy Reeves[2], Chaitanya S. Gokhale[1*]

[1]Research Group for Theoretical Models of Eco-evolutionary Dynamics,

[1]Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology,

[2]Department of Evolutionary Genetics,

Max Planck Institute for Evolutionary Biology,

August-Thienemann-Straße 2, 24306 Plön, Germany.

[*] Correspondence. gokhale@evolbio.mpg.de

Abstract

Synthetic gene drive technologies aim to spread transgenic constructs into wild populations even when they impose organismal fitness disadvantages. The properties of gene drive constructs are diverse and depend on their molecular construction, and differential selection pressure they impose in the varied ecological situations they encounter. The extraordinary diversity of conceivable drive mechanisms and the range of selective parameters they may encounter makes it very difficult to convey their relative predicted properties. The sheer number of published manuscripts in this field, experimental and theoretical, is a testament to the possibilities presented by this technology. We evaluate and condense the essential synthetic drive mechanisms from a variety of studies and present a unified mathematical paradigm (and a user-friendly tool `DrMxR` Drive Mixer) describing the properties of a wide variety of single construct gene drives (non-suppression). Within this common framework, we have been able to recapitulate key published results derived using bespoke modelling frameworks. Because a unified framework is employed, it is also possible to seamlessly explore the consequences of combining multiple drive approaches within a single construct. We provide a

method for analytically assessing the measure of invasiveness of a drive construct. As opposed to typical studies of synthetic drives, we explore the resilience of such drives in a spatially explicit manner advancing the connection between realistic spatial dynamics and typical well-mixed populations. Besides a scientific advance, our results and the tools provided an intuitive and objective way for regulators, scientists and NGOs to evaluate the properties and robustness of proposed and future gene drive approaches.

Introduction

Gene drive techniques increase the frequency of a synthetic genetic element in populations in a manner only partially determined by its impact on organismal fitness (and stochastic events). As an example, the natural Segregation Distorter (SD) locus in *Drosophila melanogaster* imposes an enormous organismal fitness cost, in that it is homozygous lethal (and only viable as heterozygotes) [Sandler et al., 1959, Sandler and Golic, 1985, Crow, 1991]. Consequently, in most circumstances, natural selection, at the organismal level would act to eliminate the SD allele. However, because of its capacity to bias the production of SD functional sperm in +/SD heterozygotes, the allele has rapidly increased to an equilibrium frequency of 1-5% in most natural populations around the globe [Hartl, 1975, Hiraizumi and Thomas, 1984, Brand et al., 2015]. This natural drive element illustrates how drive elements can increase in frequency even where there is a substantial cost to (overall) organismal fitness. Since the development of molecular biological techniques, there has been an interest in developing synthetic drive elements used to push linked genes into wild populations in a self-perpetuating manner. This is generally termed replacement drive, to distinguish from suppression drive that aims to reduce or completely eradicate the size of target populations upon release.

As in the case of SD, it does not necessarily follow that any synthetic drive element will likely increase to a frequency to the extent that it displaces all wild-type alleles at its chromosomal location that were initially present in the wild population. This fix-

ation property is dependent on various drive parameters of the developed system. Other such properties of interest are the speed of action, reversibility and potential to be spatially confined to only target populations. The sensitivity of such fundamental properties of drive systems to drive parameters has been a topic of interest of numerous recent theoretical studies.

Developments in the theories and models of gene drive, to some extent, outstripped the experimental approaches. However, the fast-pace of developments in the field of molecular biology allow us to design complicated drive systems which may be substantially better in the properties of interest than their natural counterparts. The need for theoretical sandboxing of such technology with planetary consequences is therefore imperative before field deployment. It is also critically important to provide the stakeholders of such a technology, sufficient understanding to evaluate the basis of crucial projected outcomes. However, the number of publications on theoretical and experimental synthetic gene drive systems is overwhelming and ever-increasing. Generally, the properties of each of the sequentially proposed synthetic drive approaches are described using bespoke modelling frameworks [Davis et al., 2001, Ward et al., 2011, Unckless et al., 2017]. Even with adequately described mathematical models, a recapitulation of crucial results is often beyond all but expert theoreticians. The capacity to quickly compare the relative sensitivity of fundamental properties of different drive scenarios to parameter changes would be of potential value to both experts and non-experts alike.

We constructed a representative literature database on synthetic gene drive system to be cognizant of the current trends in this rapidly growing field of research. The database consists of 48 publications from year 1995 to 2019. The literature is sorted on the basis of gene drive type (replacement or suppression), the model system under study, theoretical methodology, consideration of breakdown of drive, the possibility of gene drive reversibility and public accessibility of the literature. From the analysis of the literature database, we found that studies on replacement drives [Gantz et al., 2015, Marshall and Akbari, 2016] are given no less importance compared to suppression drives [Hammond et al., 2016, Beaghton et al., 2017, Kyrou

et al., 2018]. The complete database and the summary statistic can be found online on GitHub. The majority of research studies have considered resistance evolution in synthetic gene drive system [Unckless et al., 2017, Noble et al., 2017]. Analytical methodologies mainly employed deterministic and stochastic models. The focus of research is now trending to consider spatial features in their models [Huang et al., 2011, Eckhoff et al., 2017, Tanaka et al., 2017, Dhole et al., 2018, Girardin et al., 2019, Bull et al., 2019, Champer et al., 2019]. The model organism on gene drive studies have been chiefly mosquitoes [Windbichler et al., 2011, Gantz et al., 2015, Hammond et al., 2016], fruit files [Larracuente and Presgraves, 2012, Gantz and Bier, 2015, Buchman et al., 2018] and rodent [Lyon, 2003, Lindholm et al., 2013, Grunwald et al., 2019] respectively. Several theoretical studies also use generic organisms that allow for generalized prediction on the spread of drive organism. Analyzing the select literature, we have distilled the primary components of synthetic gene drive models in a succinct theoretical model and a handy, user-friendly tool `DrMxR` - Drive Mixer.

We develop from the principles of standard population genetics, incorporating the processes that subvert the generally dominant role that organismal fitness plays in how natural selection can impact the frequencies of alleles within natural populations.

To develop the model, first, we consider the lifecycle of a generic diploid organism through the various stages of development, from an adult, forming gametes to zygote and then back to an adult. We discuss how the drive can act at any one or all of these stages. We then proceed to combine the knowledge into a single population dynamic model. We test our developed model by extending it in different ways, thus recovering the specific cases of gene drives discussed in previous theoretical and experimental studies. Next, we extend our analysis to the ecological dimension as well. We determine the risk level of losing a wild population, through accidental introduction or migration of drive capable individuals. Further, we test our results in spatially explicit conditions and determines the extended conditions required for the invasiveness of drive elements.

We thus show that a single theoretical approach when minimally extended provides specific cases of different drive systems. This exercise provides us with a uni-

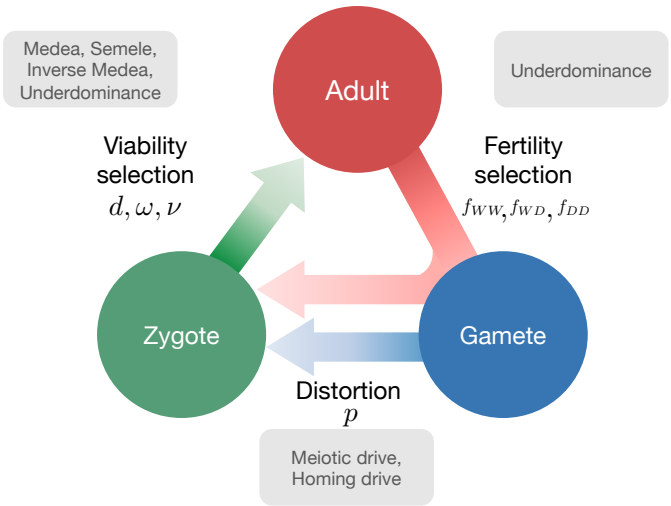

Figure 1: Lifecycle of an individual organism for a generic gene drive model. It is assumed that individuals reproduce sexually and that the lifecycle has three stages, Adult, Gamete and Zygote. Adults produce gametes which combine to form zygotes. Zygotes grow up to become adults. Three factors can act during the life stages of an organism: distortion, viability selection and fertility selection (represented as arrows). Each can influence the probability of inheritance of a gene in the population and so can be potentially manipulated to engineer gene drive constructs. Parameters, described in the text are associated with each of the three arrows. Examples of named drive systems that can be generated are provided associated with the respective arrow.

versal vocabulary as opposed to the invention of new terms for every different drive systems, which makes comparing them tedious. Furthermore, from the regulators perspective, we provide a tool to test specific cases of proposed drive mechanisms. To this end, we begin by detailing the process of theory development in the following section.

Results

One of our main results is the generation of a user-friendly application called `DrMxR` (Drive Mixer) shown in Figure 2 and available on GitHub. With an intuitive interface, both experts and non-experts alike can explore the properties of previously described drive systems across their entire parameter space. Besides, users can combine drive systems to represent the likely properties of largely unexplored combinations. For developing this application, we have assumed an obligate sexually reproducing organism, a likely necessity for successful gene drive where organismal fitness is negatively impacted. The life cycle of the organism is split into three tractable stages; the minimal abstraction required to recover the established results in the field of engineered gene drive systems. Further complications can indeed be added depending on the exact case study in focus.

Figure 1 shows the life cycle of an individual in our model. We focus on two allelic types - wild-type (W) and the driven gene (D). Thus we have adults of three genotypes, wild-type homozygotes WW, heterozygotes WD and drive homozygotes DD. Adults are chosen from the population pool for reproduction. Adults produce gametes which combine to form zygotes. The zygotes grow up to become adults, and the cycle continues. We allow for overlapping generations, a realistic assumption for numerous target species such as mosquitoes, drosophila or rodents [Windbichler et al., 2011, Backus and Gross, 2016, Buchman et al., 2018]. We assume that the alleles during gamete formation are segregated independently according to Mendel's inheritance laws. Hence the total number of alleles in the absence of any evolutionary processes remain conserved over successive generations. Frequencies of genotypes, therefore, reach Hardy-Weinberg equilibrium in the limit of infinite population, random mating and no selection (black parabola connecting WW and DD in the Figure 2).

The essential feature of a gene drive is biasing the chance of inheritance of the desired gene in the population [Champer et al., 2016]. The expected outcome, however, is that the population composition is modifiable in a controlled fashion. Interventions along the lifecycles can accomplish the change via the process of distortion, viability

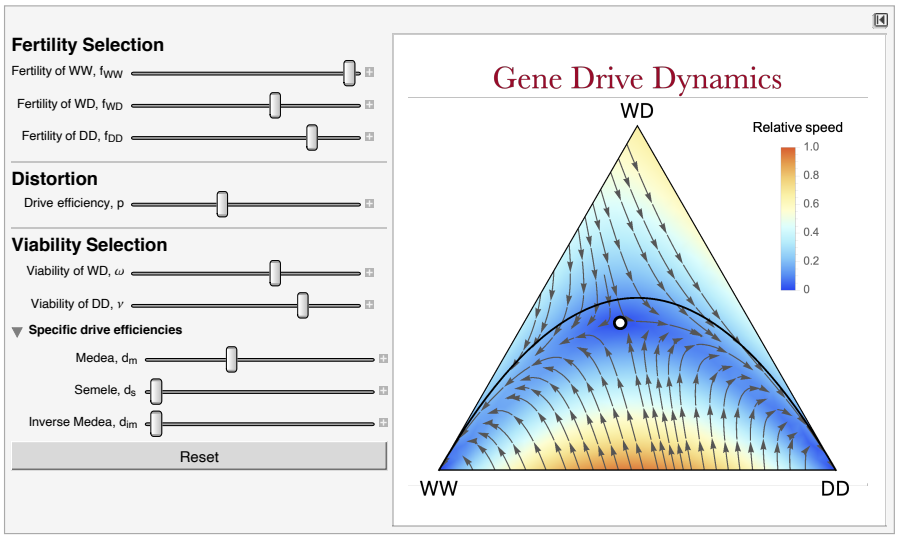

Figure 2: `DrMxR` (Drive Mixer): a handy tool to explore the population level consequences of different drive systems. Different types of drive mechanisms can be designed by biasing the three fundamental phases in the lifecycle of a diploid organism. In `DrMxR` the user can select the magnitude of the impact of the selected drive. Furthermore the consequences of a combination of different types of drives can be visualised in the space of the three genotypes denoted by the wild-type (W) and the Drive (D) homozygotes and the WD heterozygotes. Note that even when parameters embodying multiple types of drive are employed simultaneously the dynamics described by `DrMxR` is always for a single combination drive construct. The tool is available on GitHub.

selection and fertility selection. These processes act at different stages of an individual's life cycle. Distortion acts at the gamete level and biases the transmission of the drive allele in the heterozygote. Gametes combine to form zygotes, but some are non-viable and die. Fertility selection acts at the adult stage when individuals are chosen to reproduce with probability proportional to their fitness. Distortion, viability selection and fertility selection, thus, together or even independently can drive the population aways from the Hardy-Weinberg equilibrium. Synthetic gene drive techniques allow us to engineer such selection pressures.

2.1 Individual dynamics

2.1.1 Viability selection

Viability selection acts during the zygote phase of an individual's lifecycle. The viability fitnesses represent the inherent variation in the fitness of the three genotype WW, WD and DD. The fitness can also capture the payload costs of the drive allele. Viability fitness is defined here as the probability of survival of the zygotes up-to-the adult stage. ω and ν denotes the genotypic viabilities of WD and DD respectively. The above parameters have been normalized with respect to the viability of WW that is kept to 1.

Well described synthetic drive systems that work principally by manipulating viability selection parameters include those using zygotic toxin-antidotes. In these systems, a proportion of zygotes of specific genotypes may become non-viable. Medea (Maternal effect dominant embryonic arrest) is an example of a naturally occurring toxin-antidote gene drive found in flour beetles [Beeman et al., 1992, Wade and Beeman, 1994]. In Medea drive wild-type homozygous offspring of heterozygous mothers are non-viable. Population dynamics of Medea drives have been studied in [Ward et al., 2011, Gokhale et al., 2014]. A synthetically engineered Medea drive first demonstrated in *Drosophila* [Chen et al., 1997] has been extensively studied [Akbari et al., 2014, Buchman et al., 2018]. Similarly, a synthetic viability selection based underdominant population transformation system was developed for *Drosophila melanogaster* in [Reeves et al., 2014]. Figure 3C shows the population dynamics of Medea drive and deviation from Hardy-Weinberg equilibrium parabola.

Using `DrMxR`, Medea and other related synthetic drive systems can be seamlessly modelled including inverse-Medea [Marshall and Hay, 2011], or Semele [Marshall et al., 2011]. The drive efficiencies of Medea, Inverse Medea and Semele drive is represented by parameters d_m, d_m and d_m respectively. A subset of key results of the population dynamics from earlier publications using `DrMxR` is shown in Appendix A.1. The framework used by `DrMxR` is general and applicable to other single construct gene drive system also entirely or partially based on viability selection.

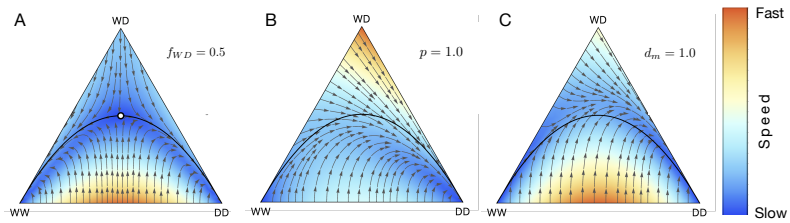

Figure 3: Effects of fertility selection, distortion and viability selection on population dynamics of the three genotypes. Population consist of single genotype at the vertices of a triangle in de Finetti diagram. A point in the interior corresponds to the population composition where all three of the genotypes potentially exist. Their relative abundance is proportional to the distance from the vertices. The black parabola curve represents Hardy-Weinberg equilibrium. The white open point represents the position of the fixed point. Colours exhibit speed of the dynamics inside de Finetti plots. The speed of the dynamics has been normalized for each plot and their absolute values are not directly comparable between diagrams through the flowlines are. (A) Fertility selection for the underdominance case where fertilities of the of the genotypes are $f_{WW} = 1$, $f_{WD} = 0.5$, $f_{DD} = 1$. An unstable point appears in the interior of de Finetti diagram and is denoted by a white circle at $(x_{WW}, x_{WD}, x_{DD}) = (0.25, 0.50, 0.25)$. (B) Distortion when drive heterozygous individuals contribute drive allele with 100% efficiency i.e. $p = 1$. A small release of WD or DD will invade the wild population exclusively consisting of WW. (C) Viability selection for Medea gene drive with drive efficiency $d_m = 1$.

2.1.2 Fertility selection

Specific genotypes may experience fitness advantages because of preference for traits during mating and or because some genotypic pairings are more fertile than others. Both of these fitness components are modelled using the fertility selection parameters. The fact that both mating success and fecundity are considered jointly dictates that the fertility selection arrow on Figure 1 traverses three life stages, rather than the two indicated for the other types of selection. The fertility fitness component arising from mating success is included in the parameter f_{WW}, f_{WW} and f_{WW} for the three genotypes. Fertility selection is an evolutionary phenomenon that drives

the population away from the Hardy-Weinberg equilibrium. In our model, we did not differentiate between sexes of the same genotype which is studied in [Hofbauer and Sigmund, 1998] where the fitness for all possible mating pair is different.

Work by [Feldman and Liberman, 1985, Nagylaki, 1987] shows the rich dynamics that ensue when fertility selection is considered. The population dynamics of two alleles system for different fertilities and sex-dependent viabilities have been extensively studied in [Hofbauer and Sigmund, 1998]. The authors have also accounted for non-random mating between the mating pairs by introducing additional parameters [Hofbauer and Sigmund, 1998]. We have accounted for variable fertility rates by introducing suitable parameters in the context of the gene drive system.

2.1.3 Distortion

Gametic distortion alters the transmission of drive alleles in heterozygotes, so they substantially exceed the Mendelian expectation of a half and is controlled by the single parameter p in our model. Biologically such distortion happens in natural meiotic drives where meiosis is subverted due to intra-genomic conflict [Sandler and Novitski, 1957, Palopoli and Wu, 1996, Lindholm et al., 2016]. Examples of naturally occurring gene drive elements based on distortion are segregation distorter and t-haplotype in heterozygous fruit fly and mice respectively [Lyon, 2003, Larracuente and Presgraves, 2012]. These drive elements bias their transmission during spermatogenesis by killing sperm carrying non-driving alleles (W). Though the killing of non-carrier sperm also has the potential to reduce fertility [Price and Wedell, 2008, Lindholm et al., 2016], 'distortion' can be conceived as an independent evolutionary force responsible for biased transmission of drive allele. The synthetic homing drive also distorts the transmission of alleles in heterozygotes. To keep the model tractable, both analytically and in terms of user comprehension, `DrMxR` does not currently consider sex-ratio gene drives (Y-driving, X-Shredder) [Burt, 2003, Burt and Deredec, 2018]. Figure 3B shows the effect of distortion on the population dynamics of the three genotypes: WW, WD, DD. Previously published evolutionary dynamics of homing drive using CRISPR are recovered in Appendix A.1.

All the above methods of biasing the inheritance pattern of a gene can be captured by the means of our generic model. In Methods and Materials we first derive the mathematical formulations of the processes independently and then combine them in a single dynamical model system. To demonstrate the generality of our approach we recover the results of [Noble et al., 2017], [Marshall and Hay, 2011], [Marshall et al., 2011] and [Gokhale et al., 2014] as special cases of our model formulation. Ecologically it is important to characterise the spread of a genetic construct. We do this in panmictic as well as spatially constrained populations. We provide an analytical form for calculating the zone of refraction (the safe amount of drive heterozygotes and homozygotes), when released, the population recovers the wild-type state. For spatially constrained systems we show the exact form in which the probability of invasion and fixation of a drive element depends on the connectivity of the network.

2.2 Combined dynamics

The three evolutionary forces viz. distortion, viability and fertility selection have the potential to act during the three stages of an organism's lifecycle. The individual impact of the forces on the population dynamics is illustrated in Figure 3 by varying parameters using our application `DrMxR`. The equilibrium dynamic changes in different ways relative to the Hardy-Weinberg equilibrium line in Figure 3. Besides individual impact, our application allows intuitive exploration of scenarios when more than one of these three evolutionary forces acts in combination. Realistically, such scenarios arise when a drive element impacts simultaneously both distortion and fertility selection [Price and Wedell, 2008, Lindholm et al., 2016]. In the Drosophila segregation distorter, sperm carrying wild-type allele in heterozygous males is selectively killed biasing the transmission of drive allele and also potentially reducing the fertility of the males. Homing Endonuclease gene drive based on CRISPR Cas9 has also been mathematically modelled to bias transmission and reduce the fertility of the genotype carrying payload gene [Noble et al., 2017].

Our approach recovers the result of [Noble et al., 2017] showing the combined

effect of distortion and fertility selection on population dynamics. Additionally, our application allows us to study various drive combinations as well. Appendix A.1 Fig. A.2 recovers the result of [Gokhale et al., 2014] shows the combined effect of fertility selection (underdominance) and viability selection (Medea gene drive). Similarly, population dynamics of other drive combinations across their entire parameter range can be intuitively explored using the `DrMxR`, for example, Medea (viability selection) and homing endonuclease (distortion) can be studied.

2.3 Ecological factors

2.3.1 Invasion threshold

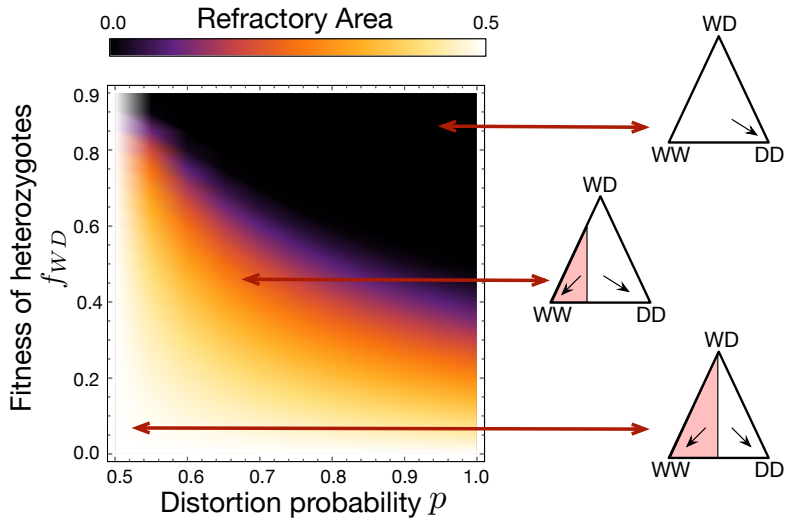

Figure 4: Heat-map showing the refractory zone with variation in distortion probability and fertility fitness of heterozygotes, p and f_{WD}. Illustration of refractory zone for specific values of p and f_{WD} of the heat-map. Trajectories of de Finetti diagram when $2pf_{WD} > f_{WW}$, drive individuals invade the wild population. Refractory zone is zero and is shown by black colour in the heatmap. $p = 0.5$ corresponds to 'no distortion' case. The values of other parameter is fixed to $f_{WW} = 1$, $f_{DD} = 1$.

The unintended spread of certain types of drive to non-target populations has

been a significant concern ever since the conception of synthetic gene drives. This interest is particularly the case for replacement drives (not intended to alter the size of populations) since the negative selection costs (fertility and viability) imposed by replacement-drive constructs are generally much smaller than for suppression drives. In this context, making the replacement gene drive localized to target populations has been a significant focus for both scientists and regulators [Backus and Delborne, 2019]. A mechanism for localizing the driven construct is the imposition of a suitably high invasion threshold. The invasion threshold is the property of the drive system that quantifies the minimum frequency of drive organisms necessary to be released to replace the wild target population. If the invasion threshold is high, the drive is more spatially restricted because the invasion of the non-target populations will require a large number of introduced individuals. Similarly as high threshold drives theoretically limit their spatial spread, they also limit the spread of drives taxonomically into partially interfertile species (or subspecies) that they may encounter. Accidental release of a few drive organisms may completely transform wild populations for gene drives with low or no threshold [Noble et al., 2018]. A recent review of different types of gene drives based on a quantitative analysis of their invasiveness can be found in [Frieß et al., 2019].

A relevant quantity of interest is the possible combination of heterozygotes and homozygotes release required for the successful invasion (if possible) of a wild population. In our model, the invasion threshold can be quantified based on the direction of the flow lines in the de Finetti diagram. We define the refractory zone as the area of the flow lines towards the fully wild population in the de Finetti diagrams. Thus we quantify the amount of accidental release or migratory influx that a population can sustain and still revert to the wild-type. Simply, we quantify the basin of attraction of the wild-type vertex.

We calculated the refractory zone by analytically computing the equation of invariant manifold separating the flow lines through approximations. Details of the calculation are in the Material and Method section. The refractory zone quantifies the minimal number of drive heterozygotes (WD) and homozygotes (DD), either released

or migrating into the target population and capable of transforming the wild-type population (WW).

Variation in the drive efficiency and fitness of different genotypes affects the refractory zone of a gene drive system. Using the insight provided from the simplified Fig. 1 we consider the case of distortion based gene drive along with fertility selection. Figure 4 shows the heat-map of the refractory zone with variation in distortion probability p and fertility fitness of heterozygotes f_{WD}. When both the drive efficiency and fitness of heterozygous is high, the refractory zone for the distortion drive is zero. Hence an accidental release of only a small frequency of drive organism would lead to complete replacement of the wild population. The gene drive system is, therefore, absolutely non-localized. Low distortion drive efficiency and fitness of heterozygotes make the drive system localized, so a significant release of drive organism is required for a successful transformation of the wild population. For intermediate values of p and f_{WD}, the gene drive system is localized and does not require a massive release.

2.4 Spatial organisation within a population

Recent work has highlighted the need for realistic spatial modelling for more accurately predicting the outcome of gene drive release, especially for suppression drives [Champer et al., 2019, Girardin et al., 2019]. Assuming random mating may in some circumstances give an incorrect prediction about the invasion condition of the gene drive. Here we derive the condition for a gene drive to invade a single wild population in a continuous landscape environment. We tune the spatial structuring between individuals within a population using the parameter k (where k tending to infinity corresponds to complete mixing, a simplifying assumption common to many models including `DrMxR`). Consequently, we have developed a framework to explore the consequences of relaxing this assumption. In this derivation, we use the framework of evolutionary game theory and track the allele frequency instead of genotype. The link between games and gene drive have been previously explored for the meiotic drives [Haig, 2010, Traulsen and Reed, 2012]. Under suitable assumptions, the payoff matrix

for the meiotic drive, i.e. with distortion and selection is given by:

$$\begin{array}{cc} & \begin{array}{cc} W & D \end{array} \\ \begin{array}{c} W \\ D \end{array} & \begin{pmatrix} f_{WW} & 2f_{WD}(1-p) \\ 2f_{WD}p & f_{DD} \end{pmatrix} \end{array}.$$

The equation that governs the population dynamics at allele level is then given by the standard selection equation [Crow and Kimura, 1970, Hofbauer and Sigmund, 1998]:

$$\dot{x}_D = x_D(f_{DD}x_D + 2f_{WD}p(1-x_D) - \phi) \tag{1}$$

where $\phi = f_{DD}^2 x_D^2 + 2f_{WD}px_D(1-x_D) + f_{WW}^2(1-x_D)^2$ is the average fitness of W and D alleles. The drive allele can invade if $2f_{WD}p < f_{WW}$ (as derived in [Noble et al., 2017] and fix in the population if $p > 1 - \frac{f_{DD}}{2f_{WD}}$. Describing the dynamics using selection equation allows us to write the population dynamics of the gene drive on a regular graph specifically infinitely large Bethe lattices of degree k using pair-approximation method. Incidentally, this equation is the replicator equation with transformed payoff matrix used in studying evolutionary games on networks [Ohtsuki and Nowak, 2006]. The payoff matrix transformation is different for different update rule. We will use birth death update rule in our analysis. In birth death update rule, first an individual is selected proportional to its fitness which then replaces one of its randomly chosen neighbour. When the payoff matrix of the game is $A = [a_{ij}]$. The payoff matrix for the birth death update rule is transformed to $A' = [a_{ij}] + [b_{ij}]$ [Ohtsuki and Nowak, 2006] where,

$$b_{ij} = \frac{a_{ii} + a_{ij} - a_{ji} - a_{jj}}{k-2} \tag{2}$$

Driven gene will invade and fix in the population if $a_{21} + b_{21} > a_{11} + b_{11}$ and $a_{22} + b_{22} > a_{12} + b_{12}$. The conditions for invasion for the case of distortion and fertility selection is:

$$\begin{aligned} & a_{21} + b_{21} > a_{11} + b_{11} \\ \Rightarrow\ & p > \left(\frac{f_{WW}}{2f_{WD}}\right) + \frac{1}{k}\left(\frac{2f_{WD} - f_{DD} - f_{WW}}{2f_{WD}}\right) \end{aligned} \tag{3}$$

If $2f_{WD} > f_{DD} + f_{WW}$, the critical p required for invasion increases relative to the mixed population scenario. Lower the network degree higher is the critical p_c required

for invasion. If $2f_{WD} < f_{DD} + f_{WW}$, the critical p required for invasion decreases. The condition obtained for the mixed population regime is recovered in the limit of $k \to \infty$. The additional condition for the fixation of the gene drive is:

$$\begin{aligned} & a_{22} + b_{22} > a_{12} + b_{12} \\ \Rightarrow p > & \left(1 - \frac{f_{DD}}{2f_{WD}}\right) - \frac{1}{k}\left(\frac{2f_{WD} - f_{DD} - f_{WW}}{2f_{WD}}\right) \end{aligned} \tag{4}$$

Here also the condition for fixation can be recovered for the mixed population regime in the limit of $k \to \infty$. It is also worth noting that the condition for invasion and fixation remain intact with variation in k if $2f_{WD} = f_{DD} + f_{WW}$. But a constraint is also put on the invasion and fixation conditions.

Figure 5 shows the invasion and fixation outcomes within a single population vary depending on the degree of spatial mixing and the efficiency of distortion. Increasing network degree can move a population where the drive cannot invade or fix to a situation where the drive can fix but cannot invade from rare for lower to moderate values of p ($p = 0.65$ to 0.80). Fixation but no-invasion case corresponds to the introduction of invasion threshold that can help in localization of gene drive. Interestingly, one can move to this regime by regulating the degree of the network. For higher values of $p > 0.80$ when the drive can both invade and fix in the population, increasing network degree again can introduce invasion threshold. A similar trend ensues in Figure 5B but here increasing network degree may allow the drive to invade the wild population but does not allow it to get fixed in the population. This scenario corresponds to the over-dominance case, and mathematically, the dynamics correspond to a stable fixed point in the interior of the simplex. The condition for the fixation and the invasion are expected to tend towards a well-mixed population regime for higher k.

Discussion

We have developed a minimalist modelling framework and identified three forces/factors responsible for the propagation of gene drive in the presence of an organismal fitness cost. These forces act during different stages of target organisms lifecycle and relate

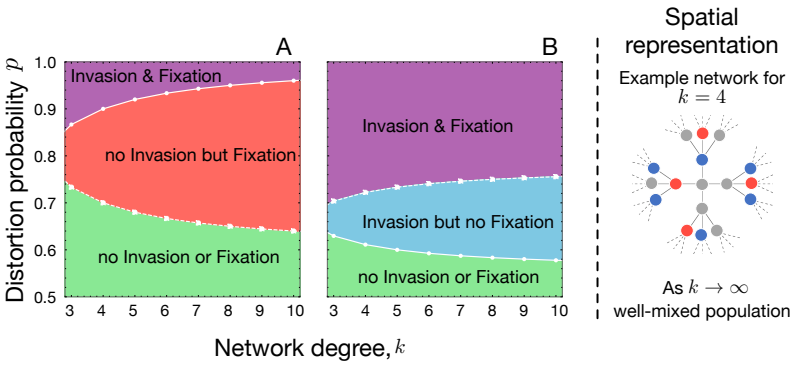

Figure 5: Spatial structure affects the condition for the invasion and fixation of the gene drive. (A) Variation in invasion (full line with circles) and fixation (dashed line with squares) conditions with respect to network degree (k) and distortion parameter (p) for $f_{WD} = 0.5$ and **(B)** $f_{WD} = 0.9$. The values of other parameters are fixed to $f_{WW} = 1$, $f_{DD} = 0.4$. A drive might show certain population dynamic result (Invasion/Fixation and combinations of them, such as no invasion from rare but fixation if sufficient drive individuals are released/migrate) for well-mixed populations ($k \to \infty$). However, for more and more structured populations, the properties can change qualitatively.

the gene driving mechanism with the organism's biology. Such a type of approach is arguably missing in earlier works on gene drive. For example, [Noble et al., 2017] studied the population dynamics of CRISPR gene drive without explicitly stating that the fitness they incorporated belongs to fertility selection parameters. In other models fitness costs have been introduced through viability fitness parameter [Marshall and Hay, 2011, Marshall et al., 2011, Gokhale et al., 2014]. One can demonstrate using `DrMxR` that the evolutionary outcome for the two cases (drive acting through viability or fertility but leading to similar costs) differs substantially. Thus our work stresses the biology of the target organism and knowing the exact phases of the lifecycle where the synthetic construct will be expressed. The current modelling approach also provides a classification of drives based on the biology of how the drive is designed (out of the three constituent forces) and avoids unnecessarily new and confusing terminology.

The eventual aim of several synthetic gene drive constructs is potential deploy-

ment in the field. At this point, the drive technology needs to be compared with other available techniques, not by experts of the particular system but regulators who need a broader perspective. Our work employs standard population genetics methods while keeping our model as generic and minimal as possible. The resulting model allows us to provide a birds-eye view of the dynamics over the space of different drive mechanisms. Teachers and regulators would benefit from using our `DrMxR` for studying the population dynamics of the gene drive. A user can choose the driving factor for the drive and its corresponding effect on the biology of the target organism by tweaking the various parameters provided in the `DrMxR`. In the absence of any selective or demographic effect, the population converges to the null Hardy-Weinberg equilibrium (which users, especially educators, can develop an intuitive understanding via `DrMxR`. Deviations from the equilibrium may be studied via the effect of the three driving factors, individually or combined, using the `DrMxR`). Conditions for invasion and fixation of the drive and its tolerance to fitness cost, that is of interest to regulators, can also be investigated. As use-cases of our unified approach, we have recovered the results of various drives such as CRISPR homing endonuclease drive, Medea, Underdominance, Inverse Medea and Semele in the Appendix A.1 [Marshall et al., 2011, Marshall and Hay, 2011, Gokhale et al., 2014, Noble et al., 2017]. The results can be recovered by the readers and explored further using `DrMxR`.

Empirical studies have shown that the selfish genetic elements based on transmission distortion can reduce both fertility (offspring production) [Dyer and Hall, 2019, Larner et al., 2019] and viability (egg to adult ratio) [Finnegan et al., 2019] of the target species. In order to correctly predict the evolutionary outcome, we have allowed to jointly vary the factors influencing the propagation of such gene drives. Flexibility to see the combined effect for various evolutionary factors influencing the spread of gene drive on the population dynamics is an essential feature of the `DrMxR`. We believe that analytical results for the evaluation of refractory zone would be useful for the regulators in determining the invasiveness of the studied drive. Methodologically, the calculation of the refractory zone is general enough to allow an interdisciplinary dialogue, e.g. with evolutionary games and population genetics [Altrock et al., 2010,

Traulsen and Reed, 2012].

Our result on the spatial model reveals that the inclusion of landscape feature changes the invasion and fixation condition of the gene drive relative to the mixed population model. We found that for lower values of network degree, the region of phase space in figure 5 for invasion & fixation and no invasion or fixation increases. Hence introducing spatial feature during interaction makes the drive either highly invasive or redundant. These results are informative to the regulators who might decide on gene drive invasiveness based on mixed population models and countries with very different geographies as opposed to where the drives are tested.

In this study, we develop a mathematical model that encompasses a variety of synthetic (and natural) gene drive techniques. Currently, we limit our study to replacement drives – spreading the drive gene along with its trait to the entire population by leading it to fixation. Suppression drives – intended to eradicate or reduce the target population or 'reversal drives' – intended to reverse the genetic alteration brought by the first gene drive [Esvelt et al., 2014, DiCarlo et al., 2015, Vella et al., 2017, Edgington and Alphey, 2019] are not included. Self-exhausting drives that first rapidly spread in the population and then self-exhaust after limited generations are also excluded in this study [Noble et al., 2019]. Our work focuses on a single locus and highlights the complexities that single locus drives can generate.

We have extended our analysis to spatial systems as per [Ohtsuki and Nowak, 2006]. Studying density-dependent migrations between patches [Altrock et al., 2011] could be included to understand the spread of different drive systems. Inclusion of ecological parameters such as seasonality and environmental disturbances would be necessary [Eckhoff et al., 2017] utilising the theory to model a specific target species. For specific species, considering detailed life history and influences in the lifecycle of the organism would be a valid extension. For example, a mosquitoes lifecycle consist of egg, larva, pupae and adult stage. Density-dependent effects due to larval competition become relevant and changing fitness components such as viability and fertility may have a small effect. Hence adding appropriate lifecycle history depending on the model organism is necessary for accurate prediction of gene drive spread. The

theoretical framework that inspires our study allows for such complications [Hofbauer and Sigmund, 1998] to be added. However, we emphasise the disparity between the theoretical developments in simple synthetic drive scenarios and urge towards a unified understanding at the primary level.

Materials and methods

We consider diploid individuals of single locus with two alleles: wild-type (W) and drive allele (D). The possible genotypes are WW, WD and DD.

4.1 Base model

We start with the simplest case assuming infinitely large population, random mating, random segregation of alleles during meiosis and no distinction between male and female genotype. We will relax some of these assumptions as we proceed. Considering all mating pairs in table 1(a), the rate of production of each genotype can be written as:

$$\begin{aligned} F_{WW} &= x_{WW}^2 + x_{WW}x_{DD} + \frac{x_{WD}^2}{4} \\ F_{WD} &= x_{WW}x_{WD} + x_{WD}x_{DD} + 2x_{WW}x_{DD} + \frac{x_{WD}^2}{2} \\ F_{DD} &= x_{DD}^2 + x_{WD}x_{DD} + \frac{x_{WD}^2}{4} \end{aligned} \tag{5}$$

where x_α and F_α are the frequency and rate of genotype production respectively and $\alpha \in$ (WW, WD, DD). The population dynamics of the genotypes in continuous time is governed by the following set of differential equation:

$$\dot{x}_\alpha = F_\alpha - x_\alpha \bar{F}. \tag{6}$$

Here, $\bar{F}$ is the average fitness of the three genotype:

$$\bar{F} = \sum_\alpha F_\alpha. \tag{7}$$

Table 1: Offspring proportions when alleles are segregated randomly during meiosis.

Parents		Offspring		
♂	♀	WW	WD	DD
WW	WW	1.0		
WW	WD	0.5	0.5	
WW	DD		1.0	
WD	WW	0.5	0.5	
WD	WD	0.25	0.5	0.25
WD	DD		0.5	0.5
DD	WW		1.0	
DD	WD		0.5	0.5
DD	DD			1.0

The total population remains constant hence the frequencies of all genotypes sum to unity.

$$x_{WW} + x_{WD} + x_{DD} = 1. \tag{8}$$

Constrains on frequencies allows us to represent the dynamics of Eq. (6) on a de Finetti diagram. Population consist of single genotype at the vertices of the de Finetti diagram. Each point in the interior corresponds to the population composition where all three of the genotypes exist. The relative abundance of the genotypes is proportional to the distance from the vertices. We will now derive the population dynamics equations when the three factors namely viability selection, fertility selection and distortion are added to the system one by one.

4.2 Viability selection

Viability selection is observed in many toxin-antidote gene drive constructs. These drives adhere to Mendel's inheritance laws and do not distort the transmission of alleles at the gamete level. In such systems, particular offsprings become non-viable dur-

ing zygote stage of the life cycle. Examples include Medea, Inverse Medea, Semele and engineered under-dominance drive etc [Beeman et al., 1992, Marshall and Hay, 2011, Marshall et al., 2011]. Depending on the type of gene drive construct one can write the rate of genotypes formation from Table 2 and Appendix A.1 Table 3. Independent of the toxin-antidote construct, variation at the genotype level may also give rise to variation in the viabilities, that is, the probability of survival of zygote up to the adult stage. Here ω and ν are the genotypic viabilities of the drive heterozygotes (WD) and homozygotes (DD) respectively. The rate of zygote production in the next generation for Medea, Inverse Medea and Semele gene drive can be written as:

$$\begin{aligned} F_{WW} &= \Bigg(x_{WW}^2 + (1-0.5d_s)\frac{x_{WW}x_{DD}}{2} \\ &\quad + (1-d_m)(1-0.5d_s)\frac{x_{WW}x_{DD}}{2} + (1-d_m)\frac{x_{WD}^2}{4} \Bigg) \\ F_{WD} &= \omega\Bigg((1-0.5d_s)(1-0.5d_{im})\frac{x_{WW}x_{WD}}{2} + x_{WD}x_{DD} + (2-d_s) \\ &\quad \times (2-d_{im})\frac{x_{WW}x_{DD}}{2} + \frac{x_{WD}^2}{2} \Bigg) \\ F_{DD} &= \nu\Bigg(x_{DD}^2 + x_{WD}x_{DD} + \frac{x_{WD}^2}{4} \Bigg) \end{aligned} \tag{9}$$

Here d_m, d_{im} and d_s measures the drive efficiency of Medea, Inverse Medea and Semele drives respectively.

4.3 Fertility selection

The relative number of offsprings produced from reproduction may differ because of the variation in the fertilities of the adult mating pairs. Assuming no distinction between the fertilities of different sexes of same genotype, the fitness component due to differential fertilities can be incorporated in the parameters f_α where $\alpha \in$ (WW, WD, DD). The rate of the offspring production for the three genotypes because of fertility

Table 2: Effect of fertility selection, distortion and viability selection on mating rates and offspring proportions. (a) Fertility selection changes the mating rate of genotypes at adult stage. (b) Distortion biases the transmission of drive allele from heterozygous individual by probability $p > 0.5$. Each entry in offspring's column gives the proportion of genotype produced from the mating pair in the corresponding row. (c) Viability selection effects offspring proportions as some may become non-viable. To illustrate an example, we consider Medea gene drive where wild-type homozygous offspring of heterozygous mother are non-viable.

Fertility Selection			Distortion			Viability Selection (e.g. Medea)		
♂	♀	Fertilities	WW	WD	DD	WW	WD	DD
WW	WW	$f_{WW}.f_{WW}$	1.0			1.0		
WW	WD	$f_{WW}.f_{WD}$	$(1-p)$	p		$0.5(1-d_m)$	0.5ω	
WW	DD	$f_{WW}.f_{DD}$		1.0			1.0ω	
WD	WW	$f_{WD}.f_{WW}$	$(1-p)$	p		0.5	0.5ω	
WD	WD	$f_{WD}.f_{WD}$	$(1-p)^2$	$2p(1-p)$	p^2	$0.25(1-d_m)$	0.5ω	0.25ν
WD	DD	$f_{WD}.f_{DD}$		$(1-p)$	p		0.5ω	0.5ν
DD	WW	$f_{DD}.f_{WW}$		1.0			1.0ω	
DD	WD	$f_{DD}.f_{WD}$		$(1-p)$	p		0.5ω	0.5ν
DD	DD	$f_{DD}.f_{WD}$			1.0			1.0ν

selection changes to

$$
\begin{aligned}
F_{WW} &= f_{WW}^2 x_{WW}^2 + f_{WW} f_{DD} x_{WW} x_{DD} + f_{WD}^2 \frac{x_{WD}^2}{4} \\
F_{WD} &= f_{WW} f_{WD} x_{WW} x_{WD} + f_{WD} f_{DD} x_{WD} x_{DD} + 2 f_{WW} f_{DD} x_{WW} x_{DD} + \\
&\quad + f_{WD}^2 \frac{x_{WD}^2}{2} \\
F_{DD} &= f_{DD}^2 x_{DD}^2 + f_{WD} f_{DD} x_{WD} x_{DD} + f_{WD}^2 \frac{x_{WD}^2}{4}.
\end{aligned}
\tag{10}
$$

The population dynamics is again given by equation 6. We assume in equation 10 that all the offsprings have equal viabilities $\omega = \nu = 1$ and no toxin-antidote drive

is present hence $d_m = d_{im} = d_s = 0$. A generalized version of the above equation includes differential mating choice (non-random mating) and distinction in the fertility of different sexes [Hofbauer and Sigmund, 1998].

4.4 Distortion

Let us now consider the case of distorted allele transmission, a violation of Mendel's standard segregation law. The gene drives engineered for distortion are in true sense non-Mendelian or super-Mendelian [Goddard and Burt, 1999]. If a drive allele is transmitted from heterozygous parents with probability p, the fitness of the three genotypes can be written as in table 2b. The rate of genotype production then changes to

$$\begin{aligned} F_{WW} &= x_{WW}^2 + 2(1-p)x_{WW}x_{WD} + (1-p)^2 x_{WD}^2 \\ F_{WD} &= 2px_{WW}x_{WD} + 2(1-p)x_{WD}x_{DD} + 2x_{WW}x_{DD} + 2p(1-p)x_{WD}^2 \\ F_{DD} &= x_{DD}^2 + 2px_{WW}x_{DD} + p^2 x_{WD}^2 \end{aligned} \tag{11}$$

Again the population dynamics for the distorted case is given by Eqs. (6), but the effective genotype production rate changes. While deriving Eqs. (11) we assume that there is no variation in intrinsic viabilities of the genotypes ($\omega = \nu = 1$), no toxin-antidote drive is present ($d_m = d_{im} = d_s = 0$) and no fertility selection ($f_{WW} = f_{WD} = f_{DD} = 1$). We can recover back the standard dynamics for $p = 0.5$ when there is no distortion in transmission probabilities of alleles. If $p > 0.5$, allele transmission from a heterozygote is biased in favour of the driven allele. Heterozygous individuals transmit only the drive allele for $p = 1$. This distortion is also the case of 'homing drive' with 100% drive efficiency.

4.4.1 Combined Dynamics

The rate of the production for the three genotypes because of viability selection, fertility selection and distortion is given by

$$
\begin{aligned}
F_{WW} &= \Big(f_{WW}^2 x_{WW}^2 + (1-p)(1-0.5d_s)(1-0.5d_m) f_{WW} f_{DD} x_{WW} x_{DD} \\
&\quad + (1-p)^2 (1-d_m) x_{WD}^2 \Big) \\
F_{WD} &= \omega \Big(p(1-0.5d_s)(1-0.5d_{im}) f_{WW} f_{WD} x_{WW} x_{WD} + 2(1-p) f_{WD} f_{DD} \\
&\quad \times x_{WD} x_{DD} + (2-d_s)(2-d_{im}) f_{WW} f_{DD} \frac{x_{WW} x_{DD}}{2} + 2p(1-p) x_{WD}^2 \Big) \\
F_{DD} &= \nu \Big(f_{DD}^2 x_{DD}^2 + 2p f_{WW} f_{DD} x_{WW} x_{DD} + p^2 f_{WD}^2 x_{WD}^2 \Big)
\end{aligned}
\tag{12}
$$

The population dynamics for the combined case is then given by including the above F_i's in Eqs. (6).

4.5 Refractory zone

For estimating the refractory zone, we analytically approximated the equation of unstable manifold when distortion and fertility selection both acts at the same time. Setting viability parameters to $\omega = \nu = 1$, no toxin-antidote based drive $d_m = d_{im} = d_s = 0$, the rate of offspring production in the next generation is given by:

$$
\begin{aligned}
F_{WW} &= f_{WW}^2 x_{WW}^2 + 2(1-p) f_{WW} f_{DD} x_{WW} x_{DD} + (1-p)^2 f_{WD}^2 x_{WD}^2 \\
F_{WD} &= 2p f_{WW} f_{WD} x_{WW} x_{WD} + 2(1-p) f_{WD} f_{DD} x_{WD} x_{DD} \\
&\quad + 2 f_{WW} f_{DD} x_{WW} x_{DD} + 2p(1-p) f_{WD}^2 x_{WD}^2 \\
F_{DD} &= f_{DD}^2 x_{DD}^2 + 2p f_{WW} f_{DD} x_{WW} x_{DD} + p^2 f_{WD}^2 \frac{x_{WD}^2}{4}
\end{aligned}
\tag{13}
$$

Using the fact that $x_{WD} = 1 - x_{WW} - x_{DD}$, the three population dynamic Eqs. (6) for the three genotype can be reduced to two. Keeping all other parameters fixed but p and f_{WD}, we found that an unstable fixed point exist in the interior of the simplex

at $(x^*_{WW}, x^*_{DD}) = \left(\frac{(1-2f_{WD}(1-p)^2}{4(1-f_{WD})^2}, \frac{(1-2f_{WD}p^2}{4(1-f_{WD})^2}\right)$. From the chain rule of derivatives, we can write

$$\dot{x}_{WW} = \frac{dx_{WW}}{dx_{DD}}\dot{x}_{DD} \tag{14}$$

Now, we approximate x_{DD} by a polynomial of single indeterminate x_{WW} keeping other parameters constant.

$$x_{DD} = \sum_{k=0}^{n} a_k x_{WW}^k \tag{15}$$

where a_k are the coefficients of the polynomial and n has a finite value. Substituting Eq. (15) in Eq. (14) and comparing the coefficients on both sides gives us many solutions for (15). The correct solution can be filtered by imposing an additional condition that the polynomial passes through the unstable fixed point in the interior of the simplex. Incidentally, the approximated polynomial is a line equation. Finally, the refractory area can be calculated by obtaining the coordinates of the points intersecting the vertex of the simplex. The appropriate codes for the calculations are available on GitHub.

Acknowledgments

The authors gratefully acknowledge funding from the Federal Agency forNature Conservation (BfN) and the Max Planck Society.

A.1 Appendix 1

In this section, we will demonstrate the flexibility of our generic modelling approach by recovering the results of earlier work on different gene drive systems. Here we present population dynamics of the three genotypes WW, WD and DD for some special cases using our generic model. Please note that the results shown here are only a subset of the work done in the original studies.

1.1 Recovering *Noble et al.* (2017)

Noble et al. [2017] studied the population dynamics of CRISPR based homing endonuclease gene drive. These gene drive construct induces double strand break at the target sequence (wild-type allele). The drive is then copied at the break site using homologous recombination. If resistance evolution is ignored, final consequence is that the heterozygous individuals only transmit drive allele during recombination. Looking from the perspective of our generic model, the drive acts in the gamete stage and uses distortion for propagating drive allele in the population. The authors in their study also accounted for the variation in the fertility rates of genotypes due to drive construct. Hence every individual undergoes both distortion and fertility selection during its life cycle. We can recover the population dynamics equations for the case using information provided in Table 2 for distortion and fertility selection. The authors derived the following condition which lead to the invasion of wild-type by the gene drive:

$$2pf_{WD} > f_{WW}$$

The above invasion condition of Noble paper is demonstrated in Fig. A.1. The original study also analysed the implication of resistance evolution and utility of multiple guide RNAs construct on the evolutionary dynamics. These features can also be included in our model and would entail addition of more genotypes and their corresponding dynamics.

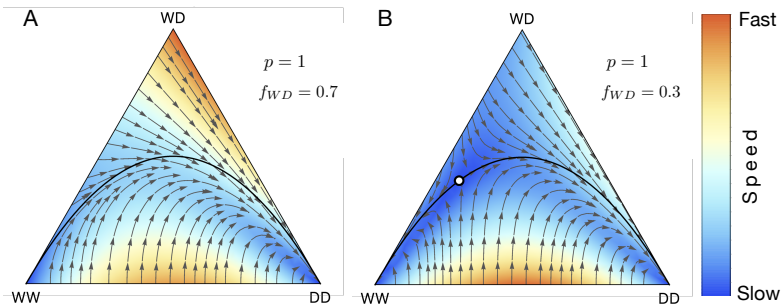

Figure A.1: Population dynamics of CRISPR based homing endonuclease gene drive. (A) When the fertility rate of heterozygous adults is 0.7 and drive efficiency is 100%, we have $2pf_{WD} > f_{WW}$. A small release of WD/DD will invade the population consisting entirely of WW. (B) When the fertility rate of heterozygous adults is 0.3, we have $2pf_{WD} < f_{WW}$. Successful invasion by gene drive would require threshold release of WD/DD in the population. The position of unstable fixed point is $(WW, DD) = (0.286, 0.354)$. Other parameters are fixed to $f_{WW} = 1, f_{DD} = 1$ for both A and B.

1.2 Recovering *Gokhale et al.* (2014)

Gokhale et al. [2014] analysed the synergistic effect of combined Medea and underdominance in a single transgenic construct. Medea gene drive utilize viability selection which acts during zygote stage of an organism. In Medea construct, wild-type homozygous offspring of heterozygous mother becomes non-viable (See Table 2C). In underdominance, the heterozygotes are less fit than both wild and drive homozygotes. Population dynamics of Medea and underdominance can be recovered from Eqs. (6) and (9). Fig. A.2 recovers the results of [Gokhale et al., 2014] for special parameter set.

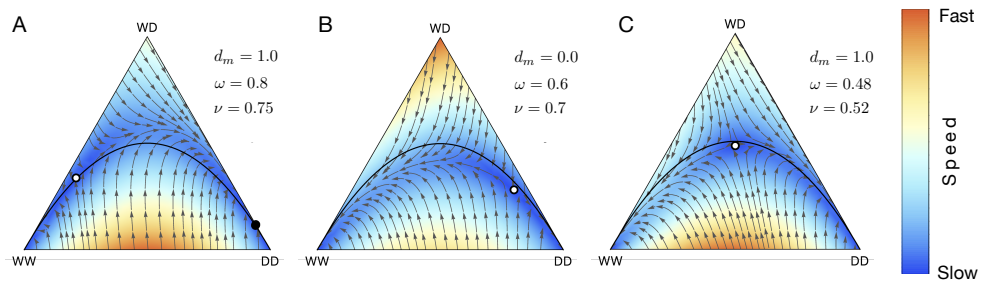

Figure A.2: de Finetti diagram showing the population dynamics of Medea, underdominace and their combined effect. (**A**) Medea only (**B**) Underdominance only (**C**) Combined effect of Medea and underdominance

1.3 Recovering *Marshall and Hay* (2011)

In Marshall and Hay [2011], Inverse Medea, homozygous offspring of wild-type mother are non-viable (see table 3).

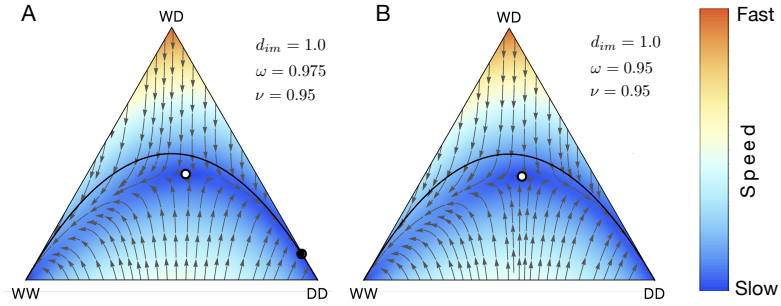

Figure A.3: Population dynamics of Inverse Medea. (**A**) For $h = 0.5$ if transgenic individuals are released above a threshold, population converges to a stable point consisting of 99.7% of DD and WD. (**B**) For $h = 1$ above a threshold release, drive homozygous (DD) invades the whole population. $d_{im} = 1$

1.4 Recovering *Marshall et al.* (2011)

Semele drive was first proposed in Marshall et al. [2011] and is based on toxin-antidote system. Transgenic males carry toxin and transgenic females carry corresponding antidote. Offspring of transgenic male carrying toxin and wild-type female with no antidotes are non-viable. The proportions of offspring of different genotypes is given in table 3. *Semele* drive like Medea and Inverse Medea utilise viability selection and acts during zygotes stage. The dynamical equation for the minimal case can be recovered using table 3 are visualised in Fig. A.4.

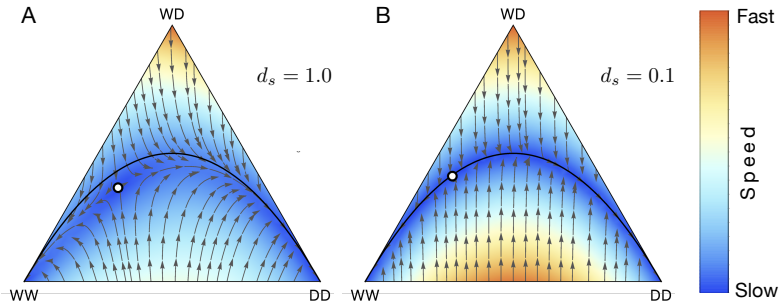

Figure A.4: Population dynamics of *Semele* drive when there is no fitness cost. (**A**) Drive efficiency is 100% (**B**) Drive efficiency is 10%

Table 3: Offspring proportions for Inverse Medea and *Semele* gene drive. Inverse medea is a two- component system with a with a zygotic toxin and a maternal antidote. Thus the heterozygous offspring of the wild-type mothers have a reduced viability $(1 - d_{im})$. In *Semele* the males carry the toxin while the females carry the antidote. Thus offspring inheriting no antidote from their mothers have reduced viability $(1 - d_s)$.

Parents		Inverse Medea			*Semele*		
♂	♀	WW	WD	DD	WW	WD	DD
WW	WW	1.0			1.0		
WW	WD	0.5	0.5		0.5	0.5	
WW	DD		1.0			1.0	
WD	WW	0.5	$0.5(1 - d_{im})$		0.5	$0.5(1 - d_s)$	
WD	WD	0.25	0.5	0.25	0.25	0.5	0.25
WD	DD		0.5	0.5		0.5	0.5
DD	WW		$(1 - d_{im})$			$(1 - d_s)$	
DD	WD		0.5	0.5		$0.5(1 - d_s)$	0.5
DD	DD			1.0			1.0

References

Omar S Akbari, Chun-Hong Chen, John M Marshall, Haixia Huang, Igor Antoshechkin, and Bruce A Hay. Novel synthetic medea selfish genetic elements drive population replacement in drosophila; a theoretical exploration of medea-dependent population suppression. *ACS synthetic biology*, 3(12):915–928, 2014.

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

P. M. Altrock, A. Traulsen, and F. A. Reed. Stability properties of underdominance in finite subdivided populations. *PLoS Computational Biology*, 7:e1002260, 2011.

Gregory A Backus and Jason A Delborne. Threshold-Dependent Gene Drives in the Wild: Spread, Controllability, and Ecological Uncertainty. *BioScience*, 69(11):900–907, 2019.

Gregory A Backus and Kevin Gross. Genetic engineering to eradicate invasive mice on islands: modeling the efficiency and ecological impacts. *Ecosphere*, 7(12):116, 2016.

Andrea Beaghton, Pantelis John Beaghton, and Austin Burt. Vector control with driving y chromosomes: modelling the evolution of resistance. *Malaria journal*, 16(1): 286, 2017.

R. W. Beeman, K. S. Friesen, and R. E. Denell. Maternal-effect selfish genes in flour beetles. *Science*, 256:89–92, 1992.

Cara L. Brand, Amanda M. Larracuente, and Daven C. Presgraves. Origin, evolution, and population genetics of the selfish segregation distorter gene duplication in european and african populations of drosophila melanogaster. *Evolution*, 69(5): 1271–1283, 2015.

Anna Buchman, John M Marshall, Dennis Ostrovski, Ting Yang, and Omar S Akbari. Synthetically engineered Medea gene drive system in the worldwide crop pest Drosophila suzukii. *Proceedings of the National Academy of Sciences of the United States of America*, 115(18):4725–4730, 2018.

James J Bull, Christopher H Remien, Richard Gomulkiewicz, and Stephen M Krone. Spatial structure undermines parasite suppression by gene drive cargo. *PeerJ*, 7: e7921, 2019.

Austin Burt. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proceedings of the Royal Society B: Biological Sciences*, 270(1518):921–928, 2003.

Austin Burt and Anne Deredec. Self-limiting population genetic control with sex-linked genome editors. *Proceedings. Biological sciences / The Royal Society*, 285(1883): 20180776, 2018.

Jackson Champer, Anna Buchman, and Omar S Akbari. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nature Reviews Genetics*, 17(3):146–159, 2016.

Jackson Champer, Isabel Kim, Samuel E Champer, Andrew G Clark, and Philipp W Messer. Suppression gene drive in continuous space can result in unstable persistence of both drive and wild-type alleles. *bioRxiv*, 28:769810, 2019.

Chun-Hong Chen, Haixia Huang, C. M. Ward, J. T. Su, L. V. Schaeffer, M. Guo, and B. Hay. A synthetic maternal-effect selfish genetic element drives population replacement in drosophila. *Science*, 316:597–600, 1997.

J. F. Crow. Why is mendelian segregation so exact? *BioEssays*, 13:305–312, 1991.

J. F. Crow and M. Kimura. *An Introduction to Population Genetics Theory*. Harper and Row, New York, 1970.

S. Davis, N. Bax, and P. Grewe. Engineered underdominance allows efficient and economical introgression of traits into pest populations. *Journal of Theoretical Biology*, 7:83–98, 2001.

Sumit Dhole, Michael R Vella, Alun L Lloyd, and Fred Gould. Invasion and migration of spatially self-limiting gene drives: A comparative analysis. *Evolutionary Applications*, 11(5):794–808, 2018.

James E DiCarlo, Alejandro Chavez, Sven L Dietz, Kevin M Esvelt, and George M Church. Safeguarding crispr-cas9 gene drives in yeast. *Nature biotechnology*, 33 (12):1250, 2015.

Kelly A Dyer and David W Hall. Fitness consequences of a non-recombining sex-ratio drive chromosome can explain its prevalence in the wild. *Proceedings of the Royal Society B*, 286(1917):20192529, 2019.

Philip A Eckhoff, Edward A Wenger, H Charles J Godfray, and Austin Burt. Impact of mosquito gene drive on malaria elimination in a computational model with explicit spatial and temporal dynamics. *Proceedings of the National Academy of Sciences*, 114(2):E255–E264, 2017.

Matthew P Edgington and Luke S Alphey. Modeling the mutation and reversal of engineered underdominance gene drives. *Journal of theoretical biology*, 479:14–21, 2019.

Kevin M Esvelt, Andrea L Smidler, Flaminia Catteruccia, and George M Church. Concerning RNA-guided gene drives for the alteration of wild populations. *eLife*, 3: 20131071, 2014.

Marcus W Feldman and Uri Liberman. A symmetric two-locus fertility model. *Genetics*, 109(1):229–253, 1985.

Sam Ronan Finnegan, Nathan Joseph White, Dixon Koh, M Florencia Camus, Kevin Fowler, and Andrew Pomiankowski. Meiotic drive reduces egg-to-adult viability in stalk-eyed flies. *Proceedings of the Royal Society B*, 286(1910):20191414, 2019.

Johannes L Frieß, Arnim von Gleich, and Bernd Giese. Gene drives as a new quality in gmo releases—a comparative technology characterization. *PeerJ*, 7:e6793, 2019.

V M Gantz, N Jasinskiene, O Tatarenkova, A Fazekas, V M Macias, E Bier, and A A James. Highly efficient cas9-mediated gene drive for population modification of the malaria vector mosquito anopheles stephensi. *Proceedings of the National Academy of Sciences*, 112(49):E6736–E6743, 2015.

Valentino M Gantz and Ethan Bier. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science*, 348(6233):442–444, 2015.

Léo Girardin, Vincent Calvez, and Florence Débarre. Catch Me If You Can: A Spatial Model for a Brake-Driven Gene Drive Reversal. *Bulletin of Mathematical Biology*, 81(12):5054–5088, 2019.

Matthew R Goddard and Austin Burt. Recurrent invasion and extinction of a selfish gene. *Proceedings of the National Academy of Sciences*, 96(24):13880–13885, 1999.

Chaitanya S Gokhale, Richard G Reeves, and F A Reed. Dynamics of a combined medea-underdominant population transformation system. *BMC Evolutionary Biology*, 14(1):98, 2014.

Hannah A Grunwald, Valentino M Gantz, Gunnar Poplawski, Xiang-Ru S Xu, Ethan Bier, and Kimberly L Cooper. Super-mendelian inheritance mediated by crispr–cas9 in the female mouse germline. *Nature*, 566(7742):105, 2019.

David Haig. Games in Tetrads: Segregation, Recombination, and Meiotic Drive. *The American Naturalist*, 176(4):404–413, 2010.

A Hammond, R Galizi, K Kyrou, A Simoni, C Siniscalchi, D Katsanos, M Gribble, D Baker, E Marois, S Russell, A Burt, N Windbichler, A Crisanti, and T Nolan.

A crispr-cas9 gene drive system targeting female reproduction in the malaria mosquito vector anopheles gambiae. *Nature Biotechnology*, 34:78–83, 2016.

D L Hartl. Genetic dissection of segregation distortion ii. mechanism of suppression of distortion by certain inversions. *Genetics*, 80(3):539–547, 1975.

Yuichiro Hiraizumi and Anita M Thomas. Suppressor systems of segregation distorter (sd) chromosomes in natural populations of drosophila melanogaster. *Genetics*, 106(2):279–292, 1984.

J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, UK, 1998.

Yunxin Huang, Alun L Lloyd, Mathieu Legros, and Fred Gould. Gene-drive into insect populations with age and spatial structure: A theoretical assessment. *Evolutionary applications*, 4(3):415–428, 2011.

Kyros Kyrou, Andrew M Hammond, Roberto Galizi, Nace Kranjc, Austin Burt, Andrea K Beaghton, Tony Nolan, and Andrea Crisanti. A crispr–cas9 gene drive targeting doublesex causes complete population suppression in caged anopheles gambiae mosquitoes. *Nature biotechnology*, 36(11):1062, 2018.

William Larner, Tom Price, Luke Holman, and Nina Wedell. An x-linked meiotic drive allele has strong, recessive fitness costs in female drosophila pseudoobscura. *Proceedings of the Royal Society B*, 286(1916):20192038, 2019.

Amanda M Larracuente and Daven C Presgraves. The selfish segregation distorter gene complex of drosophila melanogaster. *Genetics*, 192(1):33–53, 2012.

Anna K Lindholm, Kerstin Musolf, Andrea Weidt, and Barbara König. Mate choice for genetic compatibility in the house mouse. *Ecology and evolution*, 3(5):1231–1247, 2013.

Anna K Lindholm, Kelly A Dyer, Renée C Firman, Lila Fishman, Wolfgang Forstmeier, Luke Holman, Hanna Johannesson, Ulrich Knief, Hanna Kokko, Amanda M Lar-

racuente, et al. The ecology and evolutionary dynamics of meiotic drive. *Trends in ecology & evolution*, 31(4):315–326, 2016.

Mary F Lyon. Transmission ratio distortion in mice. *Annual review of genetics*, 37(1): 393–408, 2003.

John M Marshall and Omar S Akbari. Gene drive strategies for population replacement. In *Genetic control of malaria and dengue*, chapter 9, pages 169–200. Elsevier, 2016.

John M Marshall and Bruce A Hay. Inverse Medea as a Novel Gene Drive System for Local Population Replacement A Theoretical Analysis. *Journal of Heredity*, 103(3): 336–341, 2011.

John M Marshall, Geoffrey W Pittman, Anna B Buchman, and Bruce A Hay. Semele: a killer-male, rescue-female system for suppression and replacement of insect disease vector populations. *Genetics*, 187(2):535–551, 2011.

Thomas Nagylaki. Evolution under fertility and viability selection. *Genetics*, 115(2): 367–375, 1987.

Charleston Noble, Jason Olejarz, Kevin M Esvelt, George M Church, and Martin A Nowak. Evolutionary dynamics of CRISPR gene drives. *Science Advances*, 3(4), 2017.

Charleston Noble, Ben Adlam, George M Church, Kevin M Esvelt, and Martin A Nowak. Current crispr gene drive systems are likely to be highly invasive in wild populations. *Elife*, 7:e33423, 2018.

Charleston Noble, John Min, Jason Olejarz, Joanna Buchthal, Alejandro Chavez, Andrea L Smidler, Erika A DeBenedictis, George M Church, Martin A Nowak, and Kevin M Esvelt. Daisy-chain gene drives for the alteration of local populations. *Proceedings of the National Academy of Sciences*, 116(17):8275–8282, 2019.

H. Ohtsuki and M. A. Nowak. The replicator equation on graphs. *Journal of Theoretical Biology*, 243:86–97, 2006.

M. F. Palopoli and C. I. Wu. Rapid evolution of a coadapted gene complex: evidence from the segregation distorter (sd) system of meiotic drive in drosophila melanogaster. *Genetics*, 143:1675–1688, 1996.

Tom AR Price and Nina Wedell. Selfish genetic elements and sexual selection: their impact on male fertility. *Genetica*, 132(3):295, 2008.

R. G. Reeves, J. Bryk, P. M. Altrock, J. A. Denton, and F A Reed. First steps towards underdominant genetic transformation of insect populations. *PLoS ONE*, 9 (5), 2014.

L. Sandler and Kent Golic. Segregation distortion in drosophila. *Trends in Genetics*, 1(C):181–185, 1985.

L. Sandler and E. Novitski. Meiotic drive as an evolutionary force. *The American Naturalist*, 91:105–110, 1957.

L Sandler, Yuichiro Hiraizumi, and Iris Sandler. Meiotic drive in natural populations of drosophila melanogaster. i. the cytogenetic basis of segregation-distortion. *Genetics*, 44(2):233, 1959.

Hidenori Tanaka, Howard A Stone, and David R Nelson. Spatial gene drives and pushed genetic waves. *Proceedings of the National Academy of Sciences*, 114 (32):8452–8457, 2017.

A. Traulsen and F. A. Reed. From genes to games: Cooperation and cyclic dominance in meiotic drive. *Journal of Theoretical Biology*, 299:120–125, 2012.

Robert L Unckless, Andrew G Clark, and Philipp W Messer. Evolution of Resistance Against CRISPR/Cas9 Gene Drive. *Genetics*, 205(2):827–841, 2017.

Michael R Vella, Christian E Gunning, Alun L Lloyd, and Fred Gould. Evaluating strategies for reversing crispr-cas9 gene drives. *Scientific reports*, 7(1):11038, 2017.

M. J. Wade and R. W. Beeman. The population dynamics of maternal-effect selfish genes. *Genetics*, 138:1309–1314, 1994.

Catherine M Ward, Jessica T Su, Yunxin Huang, Alun L Lloyd, Fred Gould, and Bruce A Hay. Medea selfish genetic elements as tools for altering traits of wild populations: a theoretical analysis. *Evolution*, 65(4):1149–1162, 2011.

Nikolai Windbichler, Miriam Menichelli, Philippos Aris Papathanos, Summer B Thyme, Hui Li, Umut Y Ulge, Blake T Hovde, David Baker, Raymond J Monnat, Austin Burt, and Andrea Crisanti. A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. *Nature*, 473(7):212–215, 2011.