Evolutionary simulations of Z-linked suppression gene drives

Luke Holman*
**luke.holman@unimelb.edu.au*

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Keywords: *W*- and *X*-shredders, Lepidopteran pests, meiotic drive, population control, schistosomiasis, selfish genes.

*School of BioSciences, The University of Melbourne, Victoria, Australia.

Introduction

Developments in biotechnology will soon make it feasible to control or eliminate populations of disease vectors, pathogens, agricultural pests, and invasive species using 'gene drives' [1–6]. Gene drives assist the propagation of engineered genes through populations using a range of mechanisms including gene conversion, poison-antidote systems, segregation distortion, and genetic incompatibility [7,8]. For example, CRISPR-Cas9 gene editing can be used to create a transgenic insertion that is transmitted to almost 100% of the offspring of heterozygous individuals instead of the usual 50%; this type of gene drive functions by inducing a double-strand break in the wild type allele, which is then repaired using the transgene as a template. Gene drives are often categorised into two types, both of which can be created with CRISPR-Cas9. 'Replacement drives' aim to propagate a human-beneficial allele that would not otherwise spread by selection, e.g. an allele that interferes with the tranmission of malaria by mosquitoes [1,9]. Conversely, 'suppression drives' aim to cause extinction (or at least a reduction in population size), for example by propagating an allele that causes lethality or sterility (MORE HERE [2,5]), or which skews the population sex ratio – typically towards males MORE HERE [10–12].

Recent theoretical papers have investigated the feasibility, efficacy, and potential negative consequencies of emerging gene drive technologies. For example, Noble et al. [6] used models to show that the basic version of a CRISPR gene drive might be highly invasive and could rapidly spread to fixation across whole meta-populations, which will sometimes be undesirable. Conversely, other models have concluded that alleles that are resistant to being cut and replaced by CRISPR gene drives could prevent them from spreading and acheiving their aims [13,14]. The issue of resistance is compounded because the standard implementation of CRISPR drives (but perhaps not updated versions; [4,5,14,15]) tends to create its own resistance alleles by deleting the Cas9 target site via non-homologous end joining (NHEJ; [1–3,14,16]). Given the potential safety, ethical, and sociopolitical concerns surrounding gene drives, some models have focused on gene drives that would go extinct after a time [11,17], would stay confined to particular populations [18,19], and/or could be reversed once they have spread [20].

Here, I focus on the evolutionary dynamics of Z-linked suppression gene drives. The simulation is inspired by proposals for various types of CRISPR-Cas9 Z-linked gene drives from Kevin Esvelt and colleagues (see www.sculptingevolution.org/genedrives/current/schistosomiasis; at the time of writing, these ideas have not been published in a journal or pre-print). Various Z-linked suppression drives proposed by Esvelt and colleagues are shown schematically in Figure XX. The gene drive would enjoy a transmission advantage in ZW females, and optionally also in ZZ males. Esvelt et al. propose that Z-linked drives could be used to control the trematode parasites (*Schistosoma* spp.) responsible for the deadly disease schistosomiasis, though Z-linked drives could theoretically be used to control any organism with female-heterogametic sex determination (such as Lepidopteran crop pests or invasive populations of birds).

A Z-linked gene drive could suppress populations by biasing gametogenesis in females, for example by cutting unique sequences on the W chromosome in order to destroy it; such a gene drive would be a 'W-shredder', similar to the X- and Y-shredders that are being

developed to control XY species [11,12,21–24]. Females carrying the gene drive would thus produce relatively few W-bearing eggs, and therefore produce mainly drive-carrying sons. Esvelt et al. point out that the evolutionary dynamics of the drive will depend on the fitness of drive carriers relative to wold types, the timing of W-shredding (e.g. in pre-meiotic cells vs mature ova), and the ecology of the target species. For example, some W-shredder designs might allow drive females to produce roughly the same number of (mostly-male) offspring as a wild-type female, because the W chromosome is destroyed early enough in oogenesis that lost W-bearing gametes can be replaced (Figure XX). Alternatively, drive-carrying females might produce half the number of offspring (or less), e.g. if the drive works by destroying all ova or offspring that carry a W chromosome, and this loss is not compensated by reduced competition on the surviving offspring. As an alternative to W-shredders, Esvelt et al. also proposed that one could suppress populations using a Z-linked locus that caused sterility or lethality in females. If this female-harming gene was capable of gene drive in males (see below), it could perhaps reach high enough frequencies to suppress the population.

Esvelt and colleagues also note that if the Z-linked locus caused gene drive in *males* in addition to females, it would probably spread through the population faster and be more likely to result in extinction. Male gene drive could be accomplished using 'standard' CRISPR-Cas9 gene conversion (REF), whereby the driving Z allele would convert the wild type locus using homing endonuclease activity followed by DNA repair, causing heterozygous males to produce mostly drive-carrying sperm and offspring.

Here, I present an evolutionary simulation that can accommodate all of these hypothesised types of Z-linked drives. I aimed to test which properties of the gene drive and the ecology of the target species are critical to determining the likelihood and speed with which the gene drive causes extinction. For example, the gene drive will presumably spread faster if it can bias transmission in both sexes, but perhaps a female-only gene drive (which might be easier to engineer) would be perfectly adequate. Also, since the population will become more male-biased as the gene drive invades, there will be eco-evo feedback (REF) that might affect the evolutionary outcome in non-intuitive ways. For example, the altered sex ratio might intensify the fitness advantage accruing to any resistant W chromosomes or autosomal modifiers that prevent W-shredding (due to Fisherian selection for an even sex ratio; [25]), relative to that observed in earlier models focusing on gene drives carried on autosomes [13,14]. Moreover, the change in sex ratio could affect the demographics of the population, particularly if males and females contribute differentially to density-dependent population growth [26], or have different dispersal rates [27]. The model incorporates the possibility that Z-linked resistance alleles are sometimes formed through NHEJ in males that are heterozygous for the drive allele [1,14,16]. It is not clear *a priori* whether the creation of resistant Z-linked alleles by NHEJ is a equally problematic for a Z-linked gene drive as it is for an autosomal drive, because it would only hinder gene conversion in males, assuming that NHEJ does not occur in response to W-shredding (which seems likely, because the W-shredder could be designed to target many repetitive regions of the W chromosome).

Methods

Overview

I model a finite population of dioecious diploids with ZW sex determination, living in j discrete habitat patches that are arranged linearly in a ring. The model considers the demography and evolution of a population into which $n_{release}$ males carrying a Z-linked gene drive are released. The drive allele causes either W-shredding or sterility in females, and optionally also causes gene drive in heterozygous males (e.g. via gene conversion of the non-driving Z). The generations are non-overlapping and each one proceeds as follows: birth, dispersal between patches, breeding with patches, and death of the parental generation. The species has 3 loci with 2-3 alleles each, some of which potentially show non-Mendelian inheritance. The equilibrium population size was roughly 10,000 in all simulations upon release of the gene drive. The model is a stochastic individual-based simulation written in R 3.4.0 and was run on the `Spartan` computer cluster at the University of Melbourne. An accompanying website describes the R scripts used to run, analyse and plot the model (link).

Loci and alleles

Each male in the simulation carries one Z-linked locus and two autosomal loci, each with two alleles. Each female carries a single allele at the Z-linked locus plus a W chromosome, as well as two alleles at both of the autosomal loci.

There are three possible Z-linked alleles: a gene drive allele (denoted Z^*), a wild-type allele ($Z+$) that is vulnerable to gene drive in Z^*Z+ males, and a resistant allele ($Z\hat{}r$) that is immune to gene drive in $Z^*Z\hat{}r$ males. Similarly, there are two possible types of W chromosomes: a wild-type W chromosome ($W+$) that is vulnerable to gene drive by the Z^* allele, and a resistant W chromosome (Wr) that is immune to gene drive.

The two autosomal loci, denoted A/a and B/b, control immunity to W-shredding and gene conversion respectively. A/a and B/b could be called 'trans-acting' resistance loci, since they are at a different locus (indeed, a different chromosome) to the gene drive allele, in contrast to the 'cis-acting' resistance conferred by the $Z\hat{}r$ and Wr alleles. The A allele is dominant to a and confers immunity to Z-linked gene drive (e.g. W-shredding) in females. The B allele is dominant to b and confers immunity to Z-linked gene drive (e.g. gene conversion) in males.

Calculating female and male fitness

I assume that wild-type individuals (i.e. those lacking drive or resistance alleles) have fitness $w = 1$, while other genotypes have $0 \leq w \leq 1$. The fecundity of females carrying the gene drive is reduced by a factor $1 - C_f$. Small C_f implies minimal costs (e.g. because lost gametes/offspring are easily replaced), $C_f = 0.5$ could represent the case where all daughters die and are not replaced, and $C_f = 1$ means that females carrying Z^* are completely sterile

(which is useful for modelling a female-sterilising Z-linked drive as opposed to a W-shredder). Similarly, the fitness of males carrying the gene drive is reduced by a factor $1 - C_m$; male fitness determines mating success (see below). Furthermore, the resistant chromosomes Wr and $Z\hat{}r$ are assumed to reduce fitness by factors of $1 - C_w$ and $1 - C_z$ respectively. For brevity, I assume that the autosomal resistance alleles A and B are cost-free. All costs are multiplicative; for example, a $Z^*Z\hat{}r$ male would have fitness $(1 - C_m)(1 - C_z)$. Additionally, all costs are assumed to be dominant, meaning that having one drive or resistance allele is equally costly as having two.

Gamete production and gene drive

I assume that the A/a and B/b loci segregate independently during meiosis and display standard Mendelian inheritance. Inheritance of the sex chromosomes is also Mendelian, except for certain genotypes carrying one Z^* allele.

Firstly, $Z^*W{+}aaBB$, $Z^*W{+}aaBb$, and $Z^*W{+}aabb$ females produce a fraction $\frac{1}{2}(1 + p_{shred})$ of Z-bearing gametes and $\frac{1}{2}(1 - p_{shred})$ W-bearing gametes. Therefore, these three female genotypes produce >50% sons when $p_{shred} > 0$, due to the shortage of W chromosomes in their gametes. Note that the gamete frequencies of Z^*Wr *females*, or of females carrying at least one A allele, conform to the standard Mendelian expectations due to resistance.

Secondly, $Z^*Z{+}AAbb$, $Z^*Z{+}Aabb$, and $Z^*Z{+}aabb$ males produce a fraction $\frac{1}{2}(1 + p_{conv} - p_{conv}p_{nhej})$ of gametes carrying the Z^* allele, $\frac{1}{2}(1 - p_{conv})$ gametes carrying the $Z+$ allele, and $\frac{1}{2}(p_{conv}p_{nhej}))$ gametes carrying the $Z\hat{}r$ allele. Thus, gene conversion occurs in males if $p_{conv} > 0$, meaning that the Z^* allele is over-represented in the gametes of these three male genotypes. The parameter p_{nhej} represents non-homologous end joining, in which an endonuclease-based gene drive fails to copy itself to the homologous chromosome, and instead deletes its target site and thereby creates a resistant allele. The gamete frequencies of $Z^*Z\hat{}r$ males, or of males carrying at least one B allele, conform to the standard Mendelian expectations due to resistance.

Calculating female fecundity

In the breeding phase of the lifecycle, the model first determines the number of offspring produced by each female. The expected fecundity of female i (F_i) is affected by three factors: the female's genotype, the density of males and females in the local patch and/or in the full population, and some global parameters in the model, as follows:

$$F_i = (1 + w_i r(1 - (D_i/K)^{\alpha}))$$

where D_i is the 'density' experienced by female i, w_i is her fitness ($0 \leq w_i \leq 1$), K is the carrying capacity, and r and α are constants that control the maximum possible fecundity and the shape of density-dependence, respectively [28].

To ensure that the simulation captures various possible types of life history and ecology, I calculated density D_i in various ways in different simulation runs. First, I define the 'global density' d_g, which acts equally on every female in every patch, as

$$d_g = \sum_{i=1}^{N_f} w_i + \delta N_m$$

where N_f and N_m are the numbers of females and males across all patches, the first term is the summed fitnesses of all these females, and δ is a constant (range: $0 - \infty$) that scales the effect of each male on d_g relative to a female with fitness $w_i = 1$. This formulation means that females with high relative fitness (i.e. fecundity) have a stronger effect on the global density than do low-fitness females. I also assume that each male contributes a fixed amount to the global density, irrespective of his genotype/fitness (since I assume that male fitness only affects male mating success; see below). The parameter δ represents sex differences in ecological niche use and behaviour. For example, we might expect $\delta < 1$ in species where males and females utilise very different environmental niches, or $\delta > 1$ in species where males are harmful to females.

Second, we define the 'local density' d_j, which is experienced by every female in patch j, as

$$d_j = \sum_{i=1}^{n_{f,j}} w_i + \delta n_{m,j}$$

where $n_{f,j}$ and $n_{m,j}$ are the numbers of females and males in patch j. As before, this formulation means that d_j depends on the fitnesses of the females in the patch, as well as the number of males (scaled by the constant δ).

Finally, the overall density experienced by female i in patch j (D_i) is a composite of the global and local densities given by $D_i = \psi d_g + (1 - \psi) d_j$. The parameter ψ scales the importance of global and local density to female fecundity. When $\psi = 0$, only local density matters and selection on females is entirely "soft", while when $\psi = 1$ only global density matters and selection on females is completely "hard" (REFERENCE). Intermediate values of ψ produce a mixture of hard and soft selection on females, and the growth rate of population depends on density at both scales.

After calculating the expected fecundity of each female (F_i), we generate the realised fecundity of the female by randomly sampling from a Poisson distribution with $\lambda = F_i$ (allowing for stochastic variation in fecundity between females with equal F_i). If the resulting number of offspring exceeded the global carrying capacity K, the model randomly selects K surviving offspring.

Competition between males

After determining how many offspring each female produces, we determine the fathers of each of these offspring. We assume that all breeding occurs within patches, such that males

only compete for matings/fertilisations with males in the same patch. If the patch contains k different male genotypes and there are $n_1, n_2, ...n_k$ males of each genotype, the probability that a male of genotype k is the father of any given offspring is

$$p_j = \frac{n_k w_k}{\sum_{i=1}^{k} n_i w_i}$$

such that relatively common and/or high-fitness male genotypes are more likely to sire offspring. This formulation means that both sexes potentially reproduce with multiple partners.

Reproduction, mutation and dispersal

After picking the parents, the model randomly generates each offspring's genotype based on the expected gamete (and thus zygote) frequencies. Offspring are born in the same patch as their parents, and the parental generation is replaced by the offspring generation.

When an offspring is created, each *Z+* allele it carries has a chance μ_Z to mutate to a *Zˆr* allele, and *vice versa* (i.e. mutation in both directions is equally probable). Similarly, each *W+* allele has a chance μ_W to mutate to a *Wr* allele, and *vice versa*.

Female and male offspring disperse to another patch with probabilities x_f and x_m respectively. We model two types of dispersal, in separate simulations: local dispersal, in which offspring move to one of the two neighbouring patches with equal probability (recalling that the patches are arranged in a ring), or global dispersal, in which dispersing offspring can land in any of the other patches.

One compete run of the simulation

The model first initialises a population of 10,000 individuals (the carrying capacity, K) with low or zero frequencies of *Zˆr*, *Wr*, *A* and *B* alleles, higher frequencies of the wild type *Z+*, *W+*, *a*, and *b* alleles, and zero *Z** gene drive alleles. It then runs 50 generations of burn-in to allow the population to reach demographic and genotypic equilibrium. Next, $n_{release}$ males with the genotype *Z*Z*aabb* are added to the population just before fathers are selected, representing the release into the wild of a laboratory-reared strain homozygous for the driving *Z*. In some simulations, all the *Z*Z*aabb* males were released in a single patch, while in others the $n_{release}$ males were randomly and evenly divided across all k patches. The model continued until either A) the driving *Z** allele went extinct, B) the population went extinct, C) the *Wr* chromosome went to fixation (making population suppression impossible), D) the *Z** allele fixed, but failed to cause population extinction, or E) 900 generations had elapsed. The model recorded which of these five outcomes occurred, as well as the allele frequencies, population size, and sex ratio at each generation.

Investigating the parameter space

For each of the parameters in Table 1, I selected two or more possible parameter values (e.g. high versus low rates of W-shredding p_{shred}; many versus few patches k). I then ran the model once for all possible combinations of these parameter values (n = 6,000,000 model runs). The aim was to measure the effect of each parameter across a background of assumptions for the other parameters, as well as to investigate all possible 2-way interactions between the parameters.

Results

Three illustrative simulation runs

Figure 1 shows three contrasting evolutionary outcomes, illustrating some representative evolutionary dynamics from among the 6,000,000 simulation runs. Tables S1-S2 give the relative frequencies of the various possible outcomes (e.g. extinction occurred in 28% of simulations involving W-shredders).

In Figure 1A, the release of 20 Z^*Z^* males at generation 50 (*c.* 0.2% of the population) was followed by the rapid invasion of the Z^* allele, which caused population extinction by reducing the number of females. Figure 1A assumes that the Z^* alleles causes perfect W-shredding ($p_{shred} = 1$), that Z^* has minimal fitness costs, and there is no resistance to W-shredding (Table S3).

In Figure 1B, Z^* invaded but failed to cause extinction, even though it was assumed that $p_{shred} = 1$ and there is no resistance to W-shredding. However, the simulation in Figure 1B assumed the presence of heavy fitness costs to individuals carrying at least one Z^* allele ($c_f = 0.5$ and $c_m = 0.2$), and that there was no gene drive in males ($p_{conv} = 0$). The assumptions $p_{shred} = 1$ and $c_f = 0.5$ could imply that the W-bearing eggs/offspring of Z^*W+ females are destroyed but not replaced, such that W-shredding increases the proportion but not the absolute number of offspring that inherit the Z^* allele. Essentially Z^* spreads via 'spite' [REF], in that it removes W chromosomes from the local population and thereby makes room for more Z^* alleles, creating indirect fitness benefits. However, the net fitness returns of the Z^* allele's 'strategy' (i.e. sacrificing 20% fitness in males in order to remove W chromosomes in females) decline as the W chromosome becomes rarer, allowing the cost in males to greatly slow the spread of Z^*.

Lastly, Figure 1C shows a case where the invasion of Z^* was halted and then reversed by the evolution of autosomal and Z-linked resistance alleles. Following the introduction of the Z^* allele, resistant $Z\hat{}r$ mutants were created via non-homologous end joining, and then $Z\hat{}r$ spread to fixation due to its immunity to gene conversion in males. The autosomal resistance allele A also spread; A confers resistance to W-shredding and was initially present in the population at 5% frequency. The spread of A caused the sex ratio to revert to normal, preventing extinction, and Z^* went extinct due to its direct fitness costs no longer being

outweighed by the benefits of W-shredding and gene conversion. Incidentally, the resistant allele A was favoured over a because the male-biased population sex ratio created by Z^* favours the production of daughters, and AA and Aa females produce more daughters than aa females in populations where Z^* is present.

Effects of each parameter on the evolution of a *W*-shredder

Figure 2 shows the main effects of each model parameter, for models of a Z-linked W-shredder that potentially also benefits from gene drive in Z^*Z males. Figure S1 is similar to Figure 2, but instead shows the number of generations until extinction on the y-axis. Under favourable assumptions, extinction occurred around 20 generations after introduction of the gene drive, though if was often longer (Figure S1).

In Figure 2, the parameters are arranged in approximate order of their importance to extinction probability. By far the most important predictors of extinction were the efficiency of W-shredding in females (p_{shred}) and the existence of resistance against W-shredding: extinction never occurred unless p_{shred} was high and autosomal alleles conferring resistance to W-shredding (allele A in the model) were absent. This makes sense, because a W-shredder cannot cause extinction unless Z^*-carrying females produce a strongly male-biased sex ratio, and resistance to W-shredding cannot readily evolve. Extinction also occurred a little more quickly when p_{shred} was 1 rather than 0.95 (Figure S1).

The strength of gene drive in Z^*Z males (p_{conv}; colours in Figure 2) also predicted extinction probability. However, p_{conv} was not as important as was the strength of W-shredding, and the W-shredder frequently caused extinction even if it did not drive in males, or if resistance to male gene drive was common. The benefit to extinction probability provided by male gene drive depended on other factors in the model (see plots of interactions; Figures XX); for example, male gene drive was at its most beneficial when resistance to it could not evolve (either through natural genetic variation, or the creation of resistant $Z\hat{}r$ alleles through NHEJ). Although it did not strongly determine the probability of extinction probability, male gene drive did considerably speed up extinction (Figure S1). For example, assuming perfect W-shredding, adding male gene drive with $p_{conv} = 0.95$ reduced the expected time to extinction from around 75 to 22 generations.

The cost of the Z^* allele to female fitness also affected extinction probability, and its effect interacted with the strength of gene drive in Z^*Z males. Specifically, assuming that the Z^* allele halves female fitness ($c_f = 0.5$) cancels out the fitness benefits of segregation distortion for the Z^* allele, and so extinction could only occur when $c_f = 0.5$ if there was gene drive in males. Reassuringly, increasing c_f from 0.01 or 0.1 had almost no effect on the likelihood of extinction, meaning that W shredders might be an effective means of population control even if females carrying the gene drive suffer a 10% fitness cost. Similarly, assuming that Z^* was costly to male carriers also had little effect on extinction probability: extinction occurred almost as frequently when the reduction in male mating success was 20% rather than 1%. Both c_f and c_m were positively correlated with the time to extinction, particularly when there was no gene drive in males (Figure S1).

Several of the ecological variables examined also affected the extinction probability. Chief among these was the shape parameter, α, of the density-dependence function. $\alpha < 1$ means that female fecundity declines at a decelerating rate as density increases, such that per-female fecundity only approaches its maximum value when the population is heavily depleted, making extinction more likely. Conversely for $\alpha > 1$, fecundity declines at an accelerating rate with increasing density, making extinction less likely due to the immediate increases in per-female fecundity that manifest once the population begins to shrink due to the spread of the gene drive. Unsurprisngly, I also found that populations in which females have a higher maximum possible fecundity (r) are somewhat more difficult to drive extinct, though the model confirmed that W-shredders can, in principle, drive extinct highly fecund species. Also, extinction was slightly more probable when female fecundity was determined more by local density than global density (ψ in Figure 2). This is because local density can remain high (and thus, per-female fecundity can remain low) even in meta-populations that are declining due to the spread of the Z^* allele in some of their sub-populations.

Extinction probability also increased with δ, the parameter that determines how male density affects female fecundity. When δ is high, female fecundity is constrained from increasing as the drive allele spreads by the ever-increasing proportion of males, contributing to extinction. Conversely, lower values of δ mean that male numbers are relatively unimportant in determining female fecundity, making extinction less likely because the shortage of females created by the gene drive alleviates competition on the remaining females. This result highlights that it is important to consider the ecology and population dynamics of target species when designing gene drives that work by eliminating one sex.

Populations that are split into many semi-isolated patches were more difficult to drive extinct that those that comparatively free of spatial structure, though the effect on extinction rate was small (Figure 2). The likely reason is that a highly-structured population allows for refugia that lack the gene drive allele. The frequency and sex bias in dispersal was relatively unimportant to extinction probability, though there was a slight tendency for higher dispersal rates to stave off extinction, presumably because dispersal allows recolonisation of patches that were cleared by the gene drive. Similarly, it did not matter whether dispersal carried individuals to any patch, or only to neighbouring patches. Finally, there was no effect of the release strategy, suggesting that it may be unnecessary to release a W-shredding gene drive across the species' entire range provided that there is gene flow between patches. An additional implication of this result is that we cannot expect Z-linked gene drives to remain confined to their release sites, as previously found for autosomal drives [6].

Effects of each parameter on a female-sterilising *Z* drive

I also used the model to examine the evolution of a Z-linked allele that causes gene drive in males and also causes total sterility in females ($c_f = 1$; Figure S2). This alternative type of gene drive was also effective at causing extinction, but only under the assumption that the population has little or no resistance to gene drive in males. For example, extinction never occurred if even 1% of the progeny of Z^*Z males inherited a resistant $Z\hat{}r$ allele created by non-homologous end joining [14]. Extinction also required that gene drive in males was

strong (high p_{conv}), and that there were no autosomal resistance alleles to male gene drive. The effects of the other parameters in the model were similar as for a W-shredder (Figure S2), and extinction (when it occurred) took a fairly similar number of generations (around 25-30; Figure S3).

Discussion

The model indicates that W-shredders are, in principle, a very effective method for eliminating populations, especially if W-shredding is highly efficient and resistance to it cannot readily evolve. The results of the model have implications for the design and effects of Z-linked W-shredders and female-sterilising suppression drives.

One design consideration is whether to engineer W-shredders that are also capable of gene drive in males, e.g. by including CRISPR/Cas9 guide RNAs that target the Z as well as the W chromosome in the gene drive cassette. In the model, W-shredders very often caused extinction even without male gene drive (i.e. when $p_{conv} = 0$), provided that females carrying the W-shredder had comparable fecundity to wild type females, and that carrier females produce very few daughters. Conversely if W-shredder females had low fecundity (around half that of a wild type, or below) or produced some daughters, male gene drive was essential for the W-shredder to cause extinction, or at least for extinction to occur quickly enough to be useful. Although male gene drive was not always essential to extinction, it did reduce the number of generations until extinction occurred, sometimes substantially. Therefore, I conclude that it would almost certainly be worth the effort to incorporate a male-acting gene drive if developing a W-shredder for species with long generation times, such as invasive birds. However the rate of population decline may be adequate even without male gene drive for species that have multiple generations per year, such as Lepidopteran pests and *Schistosoma* parasites. This could simplify the design of W-shredders since they would only need to target the W, particularly because male-acting CRISPR-based drives are more challenging to develop in at least some taxa [29].

Another aim when designing W-shredders should be to ensure that female carriers produce as few daughters as possible (ideally none), while producing a large number of drive-carrying sons (ideally as many as the total offspring produced by non-carriers). This implies that one should ideally design a construct that cleaves the W chromosome early in gametogenesis or development, to increase the chance that the number of surviving progeny produced by each female is unaffected. For some species, this may mean placing the W-shredder under the control of a promoter that is active in the female germ line [30], such that females are able to replace lost W-bearing oocytes before they are provisioned with limiting maternal resources. For other species, it may be possible to shred the W chromosome in W-bearing ova or embryos using maternally-derived Cas9 and guide RNAs deposited in the egg after being synthesised by the Z^* allele [31]. In Lepidoptera, juvenile density is often strongly negatively correlated with survival, and there are maternally-transmitted endosymbionts that are able to drive through populations by killing males to lessen competition on their infected sisters [32,33]; these observations suggest that W-shredder alleles would invade Lepidopteran

populations even if *Z*W* females produced half as many eggs. Unsurprisngly, models of a female-sterilising *Z*-linked element showed that the element could only spread and cause extinction if it benefitted from gene drive in males. These female-sterilising drives were effective at causing extinction, but were very vulnerable to the evolution of resistance to gene drive in males [14].

The *W*-shredding mechanism should also be designed in a way that makes it difficult for *W*-linked or *trans*-acting resistance to shredding to evolve. One way to do this would be to use a single guide RNA that targets *W*-specific sequences that have high copy number, or to use multiple guide RNAs that target multiple *W*-linked sequences [REF]. This way, multiple changes to the reference sequence would be required for a *W* chromosome to acquire resistance to cleavage by the *W*-shredder. To ensure that the targets of cleavage do not become resistant as a result of indels that are induced by non-homologous end joining (NHEJ), one can ensure that the guide RNA's target lies within an essential gene where an indel would be fatal, preventing resistant alleles from accumulating in the population. This may not be necessary if the *W*-shredder targets many *W*-linked loci, but it is an important design consideration for any male component of the gene drive, because the evolution of *Z*-linked resistance completely nullifies the usefulness of male gene drive [14]. Recent work suggests that it is possible to create arrays which contain many guide RNAs separated by spacers [34]; advances like this suggest that it may soon be less challenging to engineer gene drives that use multiple guide RNAs.

The model also indicated that extinction does not require the release of large numbers of individuals: releasing 20 _Z*Z* males was often enough to eliminate a spatially-structured metapopulation of 10,000 individuals within a few generations. On the one hand, this is an advantageous property because *W*-shredders would be cheap and easy to deploy once they are developed, and they are likely to extirpate most or all of the metapopulation even if gene flow between subpopulations is weak. However, such high invasiveness is not always desirable, because it makes the gene drive more difficult to restrict to one particular site or population. This could limit the usefulness of *W*-shredders to control species like Lepidoptera and birds, where one may wish to eradicate only invasive or agriculturally damaging populations, while leaving other regions untouched. Modifications to gene drive design – such as the self-limiting 'daisy drive' system – are being developed to address this important concern [17,19].

The model further showed that *W*-shredders can fail to cause extinction if carrier individuals have low fitness, although extinction was frequently oberved even when these fitness costs were substantial. Populations in which females can become highly fecund as the population shrinks (i.e. low α and high r) are also harder to drive extinct, though this could likely be solved by continually releasing more drive males. The model also highlighted that *W*-shredders, and indeed any gene drive that skews the sex ratio towards males, are most effective in suppressing species in which the density of males is an important determinant of population growth, e.g. because males use resources that females need [35]. By contrast if male density is not very important to population growth (e.g. because females are limited by a resource that is not consumed by males), female fecundity increases as females become rarer, slowing the decline in population size caused by the *W*-shredder and potentially staving off extinction. Interestingly, the sexes are very different in the trematode parasites (*Schistosoma*

sp.) which cause schistosomiasis, which have been proposed as candidates for control using a W-shredders. Female *Schistosoma* live inside the body of the much larger male, who feeds on the host's blood and passes some of it to the female. Presumably, this means that the number of males (not females) is the primary determinant of whether a host/population is full to capacity, making *Schistosoma* a good candidate for control with W-shredders. In Lepidoptera and birds – two other ZW taxa that could potentially be controlled with W-shredders – males and females generally do have similar ecological niches, such that W-shredders should be effective. Other ecological parameters like the patchiness of the population (k), the frequency and sex bias of dispersal (x_f and x_m), and the scale of competition (ψ) had relatively little effect on the probability of extinction.

Finally, I note that W-shredders might tend to be easier to develop than X-shredders. Intital efforts to develop an X-shredder in *Anopheles* mosquitos were hindered because the protein that targets the X was paternally transmitted to the embryo inside both X- and Y-bearing sperm, causing all embryos to die due to cleavage of the maternally-inherited X, resulting in the loss of the drive allele in a single generation. Although this technical issue was later solved using ingenuity, such paternal effects would not be a problem for a W-shredder since the W chromosome is unique to females. Additionally, W-shredders might sometimes be easier to develop than gene drives that work by deleting genes that are essential to female (but not male) fertility or viability [11]. This is because one could design a prototype W-shredder using only sequence data from the sex chromosomes, while identifying genes with female-specific fitness effects requires more detailed data (e.g. expression profiling or knockout studies) that is not always readily available.

Acknowledgements

I thank the organisers (Anna Lindholm and Tom Price), funding bodies (European Society for Evolutionary Biology; Swiss National Science Foundation), and attendees of the 2018 ESEB *Progress Meetings in Evolutionary Biology*, which provided the impetus for this paper. I also thank Kevin Esvelt and colleagues for describing their ongoing research online; their ideas greatly influenced this paper.

References

1. Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, James AA. 2015 Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *PNAS* **112**, E6736–E6743.

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

3. Wang Z, Pan Q, Gendron P, Zhu W, Guo F, Cen S, Wainberg MA, Liang C. 2016 CRISPR/Cas9-derived mutations both inhibit HIV-1 replication and accelerate viral escape.

Cell Reports **15**, 481–489.

4. Prowse TA, Cassey P, Ross JV, Pfitzner C, Wittmann TA, Thomas P. 2017 Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. *Proc. Roy. Soc. B* **284**, 20170799.

5. Kyrou K, Hammond AM, Galizi R, Kranjc N, Burt A, Beaghton AK, Nolan T, Crisanti A. 2018 A CRISPR–Cas9 gene drive targeting *doublesex* causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nature Biotechnology*

6. Noble C, Adlam B, Church GM, Esvelt KM, Nowak MA. 2018 Current CRISPR gene drive systems are likely to be highly invasive in wild populations. *eLife* **7**, e33423.

7. Lindholm AK *et al.* 2016 The ecology and evolutionary dynamics of meiotic drive. *Trends in Ecology & Evolution* **31**, 315–326.

8. Champer J, Buchman A, Akbari OS. 2016 Cheating evolution: Engineering gene drives to manipulate the fate of wild populations. *Nature Reviews Genetics* **17**, 146.

9. Marshall JM, Akbari OS. 2015 Gene drive strategies for population replacement. In *Genetic control of malaria and dengue*, pp. 169–200.

10. Beaghton A, Beaghton PJ, Burt A. 2017 Vector control with driving Y chromosomes: modelling the evolution of resistance. *Malaria Journal* **16**, 286.

11. Burt A, Deredec A. 2018 Self-limiting population genetic control with sex-linked genome editors. *Proc. Roy. Soc. B* **285**, 20180776.

12. Papathanos PA, Windbichler N. 2018 Redkmer: An assembly-free pipeline for the identification of abundant and specific X-chromosome target sequences for X-shredding by CRISPR endonucleases. *The CRISPR Journal* **1**, 88–98.

13. Drury DW, Dapper AL, Siniard DJ, Zentner GE, Wade MJ. 2017 CRISPR/Cas9 gene drives in genetically variable and nonrandomly mating wild populations. *Science Advances* **3**, e1601910.

14. Unckless RL, Clark AG, Messer PW. 2017 Evolution of resistance against CRISPR/Cas9 gene drive. *Genetics* **205**, 827–841.

15. Esvelt KM, Smidler AL, Catteruccia F, Church GM. 2014 Emerging technology: Concerning rna-guided gene drives for the alteration of wild populations. *Elife* **3**, e03401.

16. Gantz VM, Bier E. 2015 The mutagenic chain reaction: A method for converting heterozygous to homozygous mutations. *Science*, aaa5945.

17. Min J, Noble C, Najjar D, Esvelt KM. 2017 Daisyfield gene drive systems harness repeated genomic elements as a generational clock to limit spread. *BioRxiv*, 104877.

18. Maselko M, Heinsch SC, Das S, Smanski MJ. 2018 Genetic incompatibility combined with female-lethality is effective and robust in simulations of *Aedes aegypti* population control. *bioRxiv*, 316406.

19. Noble C *et al.* 2019 Daisy-chain gene drives for the alteration of local populations. *PNAS*, 201716358.

20. Vella MR, Gunning CE, Lloyd AL, Gould F. 2017 Evaluating strategies for reversing CRISPR-Cas9 gene drives. *Scientific Reports* **7**, 11038.

21. Windbichler N, Papathanos PA, Crisanti A. 2008 Targeting the x chromosome during spermatogenesis induces y chromosome transmission ratio distortion and early dominant embryo lethality in anopheles gambiae. *PLoS Genetics* **4**, e1000291.

22. North A, Burt A, Godfray HCJ. 2013 Modelling the spatial spread of a homing endonuclease gene in a mosquito population. *Journal of Applied Ecology* **50**, 1216–1225.

23. Galizi R, Doyle LA, Menichelli M, Bernardini F, Deredec A, Burt A, Stoddard BL, Windbichler N, Crisanti A. 2014 A synthetic sex ratio distortion system for the control of the human malaria mosquito. *Nature Communications* **5**, 3977.

24. Prowse TA, Adikusuma F, Cassey P, Thomas P, Ross JV. 2019 A Y-chromosome shredding gene drive for controlling pest vertebrate populations. *eLife* **8**, e41873.

25. Holman L, Price TA, Wedell N, Kokko H. 2015 Coevolutionary dynamics of polyandry and sex-linked meiotic drive. *Evolution* **69**, 709–720.

26. Rankin DJ, Kokko H. 2007 Do males matter? The role of males in population dynamics. *Oikos* **116**, 335–348.

27. Li X-Y, Kokko H. 2019 Sex-biased dispersal: A review of the theory. *Biological Reviews* **in press**.

28. Fowler CW. 1981 Density dependence as related to life history strategy. *Ecology* **62**, 602–610.

29. Grunwald HA, Gantz VM, Poplawski G, Xu X-rS, Bier E, Cooper KL. 2019 Super-mendelian inheritance mediated by crispr–cas9 in the female mouse germline. *Nature* **566**, 105.

30. Zhang Z, Niu B, Ji D, Li M, Li K, James AA, Tan A, Huang Y. 2018 Silkworm genetic sexing through w chromosome-linked, targeted gene integration. *Proceedings of the National Academy of Sciences* **115**, 8752–8756.

31. Hay BA, Chen C-H, Ward CM, Huang H, Su JT, Guo M. 2010 Engineering the genomes of wild insect populations: Challenges, and opportunities provided by synthetic medea selfish genetic elements. *Journal of Insect Physiology* **56**, 1402–1413.

32. Jiggins F, Hurst G, Jiggins C, Schulenburg J vd, Majerus M. 2000 The butterfly *Danaus chrysippus* is infected by a male-killing *Spiroplasma* bacterium. *Parasitology* **120**, 439–446.

33. Jiggins FM. 2003 Male-killing *Wolbachia* and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. *Genetics* **164**, 5–12.

34. Kurata M *et al.* 2018 Highly multiplexed genome engineering using crispr/cas9 gRNA arrays. *PloS one* **13**, e0198714.

538 35. Li X-Y, Kokko H. 2019 Intersexual resource competition and the evolution of sex-biased
539 dispersal. *Frontiers in Ecology and Evolution* **7**, 111.

Tables

Table 1: List of variables, and their corresponding parameter(s) in the model, which were varied in order to study their effects on the likelihood of population extinction.

Variable	Parameter(s)	Outcome
Strength of gene drive in females (e.g. W-shredding)	p_{shred}	1.00
Strength of gene drive in males (e.g. gene conversion)	p_{conv}	1.00
Cost of gene drive allele to female fecundity	c_f	1.00
Cost of gene drive allele to male mating success	c_m	1.00
Frequency of W-linked resistance mutations	μ_W	1.00
Frequency of Z-linked resistance mutations and NHEJ	μ_Z and p_{nhej}	1.00
Frequency of autosomal resistance alleles	μ_A and μ_B	1.00
Patchiness of the population	k	1.00
Dispersal rate of males and females	x_m and x_f	1.00
Global versus local density-dependence of female fecundity	ψ	1.00
Contribution of males relative to females in density-dependence	δ	1.00
Number of gene drive carrier males released	$n_{release}$	1.00
Release strategy: all in one patch, or global	-	1.00
Fecundity of females at low population densities	r	1.00
Shape of density dependence	α	1.00

Figures

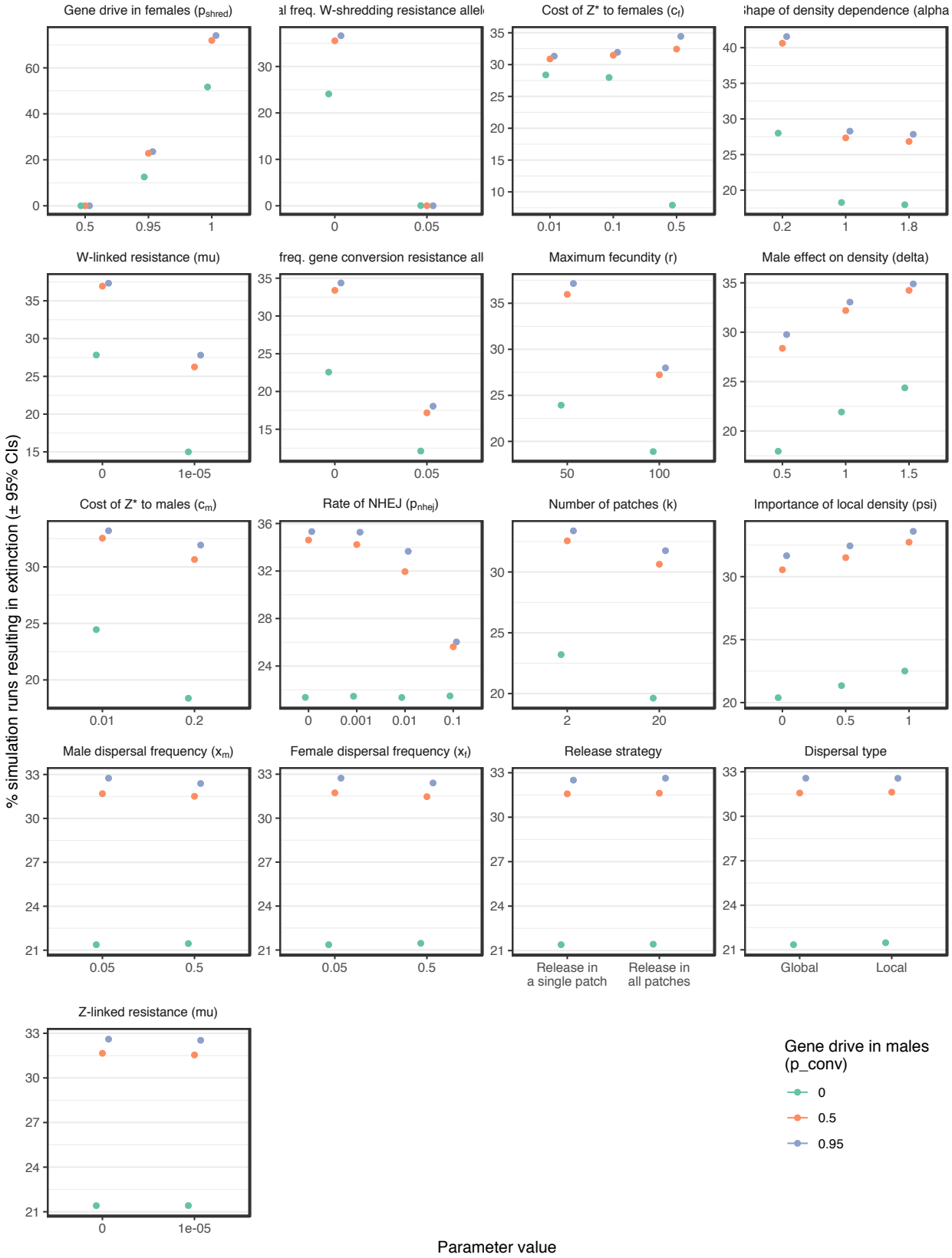

Figure 1: The percentage of simulations of a W-shredder that ended in extinction, for all runs with a particular value (shown on the x-axis) for a given parameter (shown in the panels). For example, there were no extinctions in any of the thousands of runs for which I assumed $p_{shred} = 0.5$, while 60% of runs where $p_{shred} = 1$ resulted in extinction. The panels are ordered by the range of the x-axis, which indicates

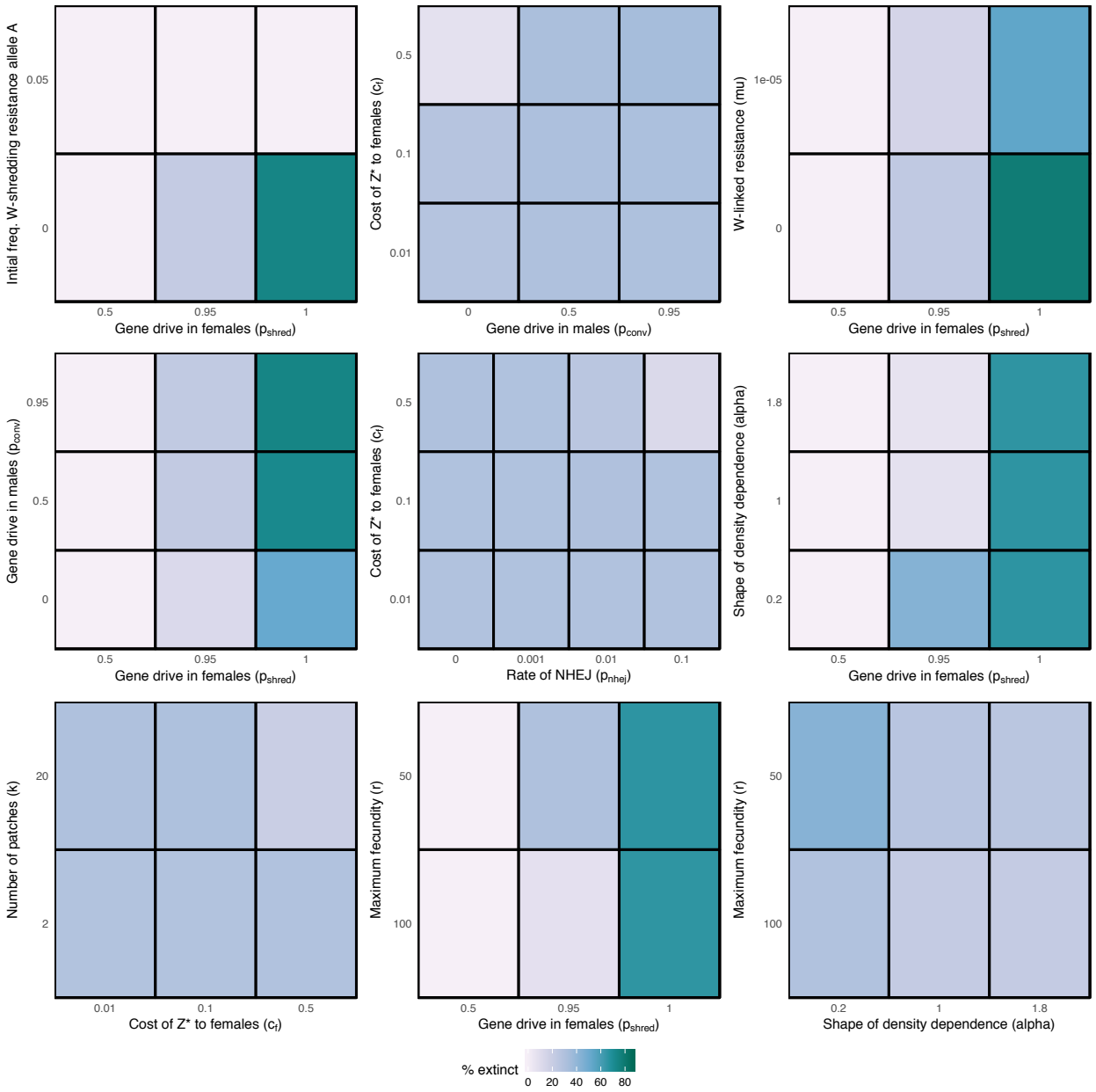

Figure 2: Heatmap showing the nine strongest interactions between pairs of parameters in the model, as determined by the GLM plotted in Figure SX.