Evolutionary models of Z-linked synthetic suppression gene drives

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

Abstract

My abstract text.

Keywords: chromosomal shredders, meiotic drive, population suppression, schistosomiasis, selfish genes.

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

Introduction

Here is a reference (e.g. Holman et al. 2018), and here's a link to a figure (**??**).

Here, we use stochastic individual-based simulations to identify parameters determining whether Z-linked meiotic drive elements will spread and/or cause population suppression by reducing female numbers.

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	XXX	1.00
Cost of gene drive allele to male mating success	XXX	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	XXX	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	c	1.00

Aims: - Compare low and high cost shredders ability to suppress the pop. High cost spreads worse since Z*W females have fewer offspring, but it also means fewer offspring are born! - Local vs global release? - butterfly vs worm vs bird - effect of m/f dispersal and patch structure - is the Z conversion needed for spread? - effect of male weighting and softness of selection on females

Methods

Overview

We model a finite population of dioecious diploids with ZW sex determination, living in a landscape containing j discrete habitat patches which are arranged linearly in a ring (preventing 'edge effects'). The model considers the demography and evolution of a population into which a number $n_{release}$ males are released. These released males carry a Z-linked allele that is capable of gene drive in females (e.g. through W-shredding) and optionally also gene drive in males (e.g. via gene consersion). Our principle aim is to identify the key factors that determine whether the Z-linked gene drive allele (termed Z^*) causes extinction of the population. The model is a stochastic individual-based simulation written in R

(REFERENCE), and was run on the `Spartan` high performance computing system at the University of Melbourne. An accompanying website presents and describes the code used to run the model and to generate all the figures (link).

Loci and alleles

Each male individual in the simulation carries a single 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 wild-type allele (denoted $Z+$) which is vulnerable to gene drive; a gene drive allele ($Z*$), and a resistant allele (Zr) which is immune to gene drive. 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 are denoted A/a and B/b, and control immunity to W-shredding and gene conversion respectively. A/a and B/b are '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 Zr and Wr alleles (REFERENCE). The A/a locus carries alleles a and A, where the A allele is dominant and confers immunity to Z-linked gene drive (e.g. W-shredding) in females. The B/b autosomal locus carries alleles b and B, where B is dominant and confers immunity to Z-linked gene drive (e.g. gene conversion) in males.

Calculating female and male fitness

cost_Zdrive_female, cost_Zdrive_male, cost_Wr, cost_Zr, cost_A, cost_B

Gamete production and gene drive

We assume that the A/a and B/b loci segregate independently during meiosis, and they display standard Mendelian inheritance. Inheritance of the sex chromosomes is also Mendelian, except for certain genotypes that carry a single copy of the $Z*$ gene drive allele.

Firstly, females with the genotype $Z*W{+}aaBB$, $Z*W{+}aaBb$, or $Z*W{+}aabb$ 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 more than 50% male offspring if $p_{shred} > 0$, due to the shortage of W chromosomes in their gametes. In contrast, the gamete frequencies of $Z*Wr$ *females*, or of females carrying at least one A allele, conform to the standard Mendelian expectations.

Secondly, males with the genotypes $Z*Z{+}AAbb$, $Z*Z{+}Aabb$, or $Z*Z{+}aabb$ 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 Zr allele. The parameter p_{conv} represents gene conversion, and when $p_{conv} > 0$, 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, thereby creating a resistant allele (REFERENCE). As before, the gamete frequencies of *Z*Zr* males, or of males carrying at least one B allele, conform to the standard Mendelian expectations.

Calculating female fecundity

To begin the breeding phase of the lifecycle, we first determine the number of offspring produced by each female in the population. We first calculate the expected fecundity of each female, which is affected by three factors: the female's genotype, the density of males and females in the local patch and/or the full population, and some global parameters in the model.

Specifically, the expected fecundity of female i (F_i) is calculated as

$$F_i = (1 + w_i r(1 - (D_i/K)^c))$$

where w_i is the relative fitness of female i (possible range: 0 to 1, where 1 is the fitness of the wild type *Z+W+aabb* females), D_i is the 'density' experienced by female i, K is the carrying capacity, and r and c are constants that scale the maximum possible fecundity and the shape of density-dependence, respectively. Thus, we assume that offspring production is density-dependent, and follows the Richards model (REFERENCE).

To ensure that the simulation captures various possible types of life history and ecology (see Introduction), we calculate the density D_i in various ways across different simulation runs. First, we define the 'global density' d_g, which is experienced equally by every female in every patch, as

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

where N_f and N_m is the number of females and males across all patches, the first term is the sum of the 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. We also assume that each male contributes a fixed amount to the global density, irrespective of his genotype/fitness (male fitness is only used to determine male mating success; see below). The parameter δ represents sex differences in ecological niche use and behaviours that affect female fecundity. For example, we might expect $\delta < 1$ in species where males and females utilise very different environmental niches, or $\delta > 1$ in species with strong inter-locus sexual conflict.

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.

Once we have calculated 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 exceeds the global carrying capacity K, we randomly cull the offspring until K are left.

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 mating/fertilisation with males from the same patch (i.e. selection on males is always "soft"; REFERENCE). 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 we assume that both sexes potentially reproduce with multiple different partners.

Reproduction, mutation and dispersal

After picking the parents, we randomly generate each offspring's genotype based on the gamete (and thus zygote) frequencies that are expected from the parental genotypes. Offspring are born in the same patch as their parents, and the parental generation is replaced by the offspring generation (i.e. we assume discrete, non-overlapping generations).

When an offspring is created, each *Z+* allele it carries has a chance μ_Z to mutate to a *Zr* 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 linearly in a ring), or global dispersal, in which dispersing offspring can land in any of the other patches.

Running the simulation

We first initialise the population, with specified (typically low or zero) frequencies for the *Zr*, *Wr*, *A* and *B* alleles, higher frequencies of the 'wild type' *Z+*, *W+*, *a*, and *b* alleles, and no *Z** gene drive alleles. We then iterated the population for 50 generations of burn-in, to allow the population to reach carrying capacity and approach genotypic equilibrium. We then introduce $n_{release}$ males with the genotype *Z*Z*aabb*, representing the release into the wild of a laboratory-reared strain homozygous for the driving Z and for autosomal factors conferring susceptibility to drive. Males are released after density-dependent regulation of female fecundity, but before picking fathers for the offspring. 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. We continued to cycle through the lifecycle (birth, migration, breeding, death) 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. We recorded which of these five outcomes occurred, as well as the allele frequencies, population size, and sex ratio at each generation.

Results

- Note that when females hardly migrate, the Wr is slow to spread across patches. It only has a good invasion probability if Z* is present, otherwise it's neutral or costly

Discussion

Acknowledgements

So long, and thanks for all the fish!

References

Holman, L., D. Stuart Fox, and C. E. Hauser. 2018. The gender gap in science: How long until women are equally represented? PLoS Biology 16:e2004956.

158 # Supporting information