Evolutionary models of Z-linked synthetic suppression gene drives

Luke Holman*

luke.holman@unimelb.edu.au

Abstract

My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My abstract text. My Keyword1, Keyword2, Keyword3, Keyword4, Keyword5. abstract text. My abstract text. My abstract text.

Keywords:

*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 (**??**).

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 k discrete habitat patches. The model considers the demography and evolution of the population following the release of a number of males carrying a Z-linked allele that is capable of gene drive in females (e.g. through W-shredding) and/or in males (e.g. via gene consersion). Our principle aim is to identify the key factors that determine whether the Z-linked genetic element (hereafter Z*) causes extinction of the population. The model is a stochastic individual-based simulation written in the R scripting language, and was run on the `Spartan` High Performance Computing system at the University of Melbourne. An accompanying website presents the annotated code used to run the model and generate all the figures (link).

Loci and alleles

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

There are three possible Z-linked alleles: a wild-type Z allele (denoted *Z+*) which is vulnerable to gene drive; a gene drive allele (denoted *Z**) which can perform gene drive in females and/or males, and a resistant allele (*Zr*) which is immune to gene conversion. Similarly, there are two possible W-linked alleles: 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 confer resistance to W-shredding and gene conversion respectively. One might call this 'trans-acting resistance', since the resistance is not mediated by a homolog of the gene drive allele, as opposed to the 'cis-acting resistance' conferred by the *Zr* and *Wr* alleles. Specifically, we assume that the *A/a* locus carries alleles *a* and *A*, where the *A* allele is dominant and confers immunity to gene drive

44 (e.g. W-shredding) in females. The B/b autosomal locus carries alleles *b* and *B*, where *B* is
45 dominant and confers immunity to gene drive (e.g. gene conversion) in males.

Gamete production and gene drive

47 We assume that the *A/a* and *B/b* loci segregate independently during meiosis, and they
48 display standard Mendelian inheritance. Inheritance of the sex chromosomes is also Mendelian,
49 except for certain genotypes that carry a single copy of the *Z** gene drive allele. Provided
50 they do no carry a resistant *Wr* chromosome or at least one *A* allele,

51 Firstly, females with the genotypes *Z*W+aaBB*, *Z*W+aaBb*, or *Z*W+aabb* produce a
52 fraction $\frac{1}{2}(1 + p_{shred})$ of Z-bearing gametes and $\frac{1}{2}(1 - p_{shred})$ W-bearing gametes. Therefore,
53 these three female genotypes produce more than 50% male offspring if $p_{shred} > 0$, due to the
54 shortage of *W* chromosomes in their gametes. Note that the gamete frequencies produced by
55 *Z*Wr females*, or by *Z*W+* females carrying at least one *A* allele, conform to the standard
56 Mendelian expectations.

57 Secondly, males with the genotypes *Z*Z+AAbb*, *Z*Z+Aabb*, or *Z*Z+aabb* produce a fraction
58 $\frac{1}{2}(1 + p_{conv} - p_{conv}p_{nhej})$ of gametes carrying the Z* allele, $\frac{1}{2}(1 - p_{conv})$ gametes carrying the
59 *Z+* allele, and $\frac{1}{2}(p_{conv}p_{nhej}))$ gametes carrying the Zr allele. The parameter p_{conv} represents
60 gene conversion, and when $p_{conv} > 0$, the *Z** allele is over-represented in the gametes of these
61 three male genotypes. The parameter p_{nhej} represents the phenomenon of 'non-homologous
62 end joining', in which a gene drive mediated by an endonuclease activity fails to copy itself
63 over to the homologous chromosome, but instead deletes its target site, creating a resistant
64 allele (REFERENCE). As before, the gamete frequencies produced by *Z*Zr* males, or by
65 *Z*Z+ males* carrying at least one *B* allele, conform to the standard Mendelian expectations.

Calculating female fecundity

67 To begin the breeding phase of the lifecycle, we first determine the number of offspring
68 produced by each female in the meta-population. The expected fecundity of each female
69 is determined by three factors: her genotype, the density of males and females in the local
70 patch and/or the meta-population, and global parameters in the model.

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

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

72 where w_i is the relative fitness of female *i* (possible range: 0 to 1, where 1 is the fitness of the
73 wild type *Z+W+aabb* females), D_i is the 'density' experienced by female *i*, *K* is the carrying
74 capacity, and *r* and *c* are constants that scale the maximum possible fecundity and the shape
75 of density-dependence, respectively.

76 To ensure that the simulation captures variation in life history and ecology between species
77 (see Introduction), we calculate the density D_i in various ways across different simulation

78 runs. First, we define the 'global density' d_g, which is experienced equally by every female in
79 every patch, as

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

80 where N_f and N_m is the number of females and males across all patches, the first term is
81 the sum of the fitnesses of all these females, and δ is a constant (range: $0 - \infty$) that scales
82 the effect of each male on d_g, relative to a female with fitness $w_i = 1$. This formulation
83 means that females with high relative fitness (i.e. fecundity) have a stronger effect on the
84 global density than do low-fitness females. We also assume that each male contributes a fixed
85 amount to the global density, irrespective of his genotype/fitness (male fitness is only used
86 to determine male mating success; see below). The parameter δ represents sex differences
87 in ecological niche use and behaviours that affect female fecundity. For example, we might
88 expect $\delta < 1$ in species where males and females utilise very different environmental niches,
89 or $\delta > 1$ in species with strong inter-locus sexual conflict.

90 Second, we define the 'local density' d_l, which is experienced by every female in a particular
91 patch, as

$$d_l = \sum_{i=1}^{n_f} w_i + \delta n_m$$

92 where n_f and n_m are the numbers of females and males in the patch. As before, this
93 formulation means that d_l depends on the fitness of the females in the patch, but only on the
94 number of males (multiplied by the constant δ).

95 Finally, the overall density D_i is a composite of the global and local densities given by
96 $D_i = \psi d_g + (1 - \psi)d_l$. The parameter ψ scales the importance of global and local density to
97 female fecundity. When $\psi = 0$, only local density matters and selection on females is entirely
98 "soft", while when $\psi = 1$ only global density matters and selection on females is completely
99 "hard" (REFERENCE). Intermediate values of ψ produce a mixture of hard and soft selection
100 on females, and the growth rate of population depends on density at both scales.

101 Once we have calculated the expected fecundity of each female (F_i), we generate the realised
102 fecundity of the female by randomly sampling from a Poisson distribution with $\lambda = F_i$
103 (allowing for stochastic as well as deterministic variation in fecundity between females). If
104 the resulting number of offspring exceeds the global carrying capacity K, we randomly cull
105 the offspring until K are left.

Competition between males

107 Next, we select fathers for each of the offspring. We assume that all breeding occurs within
108 patches, such that males are only in competition with males from the same patch (i.e. selection

¹⁰⁹ on males is always "soft"; REFERENCE). If the patch contains k different male genotypes
¹¹⁰ and $n_1, n_2, ...n_k$ males of each genotype, the probability that a male of genotype k is the
¹¹¹ father of any given offspring is

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

¹¹² such that relatively common and/or high-fitness male genotypes are more likely to sire
¹¹³ offspring. This formulation means that we assume that both sexes potentially reproduce with
¹¹⁴ multiple different partners.

Reproduction, mutation and dispersal

¹¹⁶ After picking both parents for each offspring, we randomly generate each offspring's genotype
¹¹⁷ based on the gamete (and thus zygote) frequencies that are expected given the parental
¹¹⁸ genotypes. Offspring are born in the same patch as their parents, and the parental generation
¹¹⁹ is removed from the population after reproducing (i.e. we assume discrete, non-overlapping
¹²⁰ generations).

¹²¹ When an offspring is created, each Z+ allele it carries has a chance μ_Z to mutate to a Zr
¹²² allele, and *vice versa* (i.e. mutation is equally probable in both directions). Similarly, each
¹²³ W+ allele has a chance μ_W to mutate to a Wr allele, and *vice versa.*

¹²⁴ Female and male offspring disperse to another patch with probabilities x_f and x_m respectively.
¹²⁵ We model two types of dispersal, in separate simulations: local dispersal, in which offspring
¹²⁶ move to one of the two neighbouring patches with equal probability (recalling that the patches
¹²⁷ are arranged linearly in a ring), or global dispersal, in which dispersing offspring can land in
¹²⁸ any of the other patches.

Running the simulation

¹³⁰ We first initialise the population, with specified (typically low or zero) frequencies for the
¹³¹ Zr, Wr, A and B alleles. We then iterated the population for 100 generations of burn-in,
¹³² to allow the population to reach its equilibrium size and genotypic composition. We then
¹³³ introduce $m_{release}$ males with the genotype Z*Z*aabb, representing the release into the wild of
¹³⁴ a laboratory-reared strain homozygous for the driving Z and for autosomal factors conferring
¹³⁵ susceptibility to drive. In some simulations, all the Z*Z*aabb males were released in a single
¹³⁶ patch, while in others the $m_{release}$ males were randomly and evenly divided across all k
¹³⁷ patches. We continued to cycle through the lifecycle (birth, migration, breeding, death)
¹³⁸ until either A) the driving Z* allele went extinct, B) the population went extinct, C) the
¹³⁹ Wr chromosome went to fixation (making population suppression impossible), D) the Z*
¹⁴⁰ allele fixed, but failed to cause population extinction, or E) 400 generations had elapsed. We
¹⁴¹ recorded which of these five outcomes occurred, as well as the allele frequencies, population
¹⁴² 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.

152 # Supporting information