

Modeling the mutation and reversal of engineered underdominance gene drives

Matthew P. Edgington and Luke S. Alphey

The Pirbright Institute, Ash Road, Woking, Surrey GU24 0NF, UK

Abstract

A range of gene drive systems have been proposed that are predicted to increase their frequency and that of associated desirable genetic material even if they confer a fitness cost on individuals carrying them. Engineered underdominance (UD) is such a system and, in one version, is based on the introduction of two independently segregating transgenic constructs each carrying a lethal gene, a suppressor for the lethal at the other locus and a desirable genetic “cargo”. Under this system individuals carrying at least one copy of each construct (or no copies of either) are viable whilst those that possess just one of the transgenic constructs are non-viable. Previous theoretical work has explored various properties of these systems, concluding that they should persist indefinitely in absence of resistance or mutation. Here we study a population genetics model of UD gene drive that relaxes past assumptions by allowing for loss-of-function mutations in each introduced gene. We demonstrate that mutations are likely to cause UD systems to break down, eventually resulting in the elimination of introduced transgenes. We then go on to investigate the potential of releasing “free suppressor” carrying individuals as a new method for reversing UD gene drives and compare this to the release of wild-types; the only previously proposed reversal strategy for UD. This reveals that while free suppressor carrying individuals may represent an inexpensive reversal strategy due to extremely small release requirements, they are not able to return a fully wild-type population as rapidly as the release of wild-types.

1 Introduction

Gene drive systems have gained much attention in recent years for their predicted ability to increase the frequency of desirable genetic material within a population. These systems have been proposed to have a number of important applications [1]. Our interest in these systems relates to their potential use in preventing the spread of mosquito-borne viruses such as dengue [2]. In this context, refractory genes have been developed that are capable of significantly reducing the ability of *Aedes aegypti* mosquitoes to

29 transmit the virus [3]. However, the incorporation of these genes into the genome resulted in individuals
30 of lower fitness than their wild-type counterparts [4]. It is thus necessary to develop gene drives that are
31 capable of increasing the frequency of such desirable genes in spite of their expected fitness costs.

32 Engineered underdominance (UD) is one class of gene drive that has been proposed for driving desirable
33 genetic traits into a population [5]. This technique is based on the introduction of two transgenic constructs
34 (A and B) at unlinked genetic loci (see Figure 1). These constructs consist of three component genes,
35 namely a lethal effector, a suppressor for the lethal at the other locus and a “cargo” gene conferring
36 a desirable phenotype. Due to the cross suppressing transgenic constructs, individuals carrying one or
37 more copies of either construct will be non-viable if they do not also possess at least one copy of the
38 other construct. Individuals carrying one or more copies of both constructs are viable since each of the
39 lethal effectors are inactivated by suppressors carried at the other locus. These effects combine to create
40 a selection pressure for individuals to either carry both transgenic constructs or neither.

41 Theoretical work on UD has demonstrated that these systems, when introduced above a threshold
42 frequency, are capable of spreading desirable genes thus replacing wild populations with those carrying
43 UD transgenic constructs [5–9]. Whilst the existence of this threshold frequency means that UD systems
44 may be relatively expensive to deploy in the field relative to other classes of gene drive, it should restrict
45 the invasion of transgenes into neighboring populations [10, 11].

46 Previous modeling work has mostly assumed a population genetics framework neglecting the possibility
47 of mutations forming within the introduced transgenes [5–11]. Under this assumption it has been predicted
48 that UD should persist indefinitely when introduced above the threshold frequency.

49 This threshold-dependent nature of UD systems has lead many to suggest that they may be reversed
50 via the introduction of wild-type individuals [12, 13]. This should lower the transgene frequency to sub-
51 threshold levels and result in the elimination of introduced transgenes. To our knowledge this is the only
52 reversal strategy proposed for UD gene drives and has yet to be investigated in detail.

53 Here we extend upon results in the previous literature by formulating a population genetics model of
54 UD that incorporates the effects of loss-of-function mutations in the introduced transgenes. In particular,
55 we investigate the predicted dynamics of UD systems in the presence of a constant rate of mutation
56 for each introduced transgene. We then go on to propose that the release of individuals carrying “free
57 suppressors” could function as a genetics-based reversal strategy and compare this to the introduction of
58 wild-types.

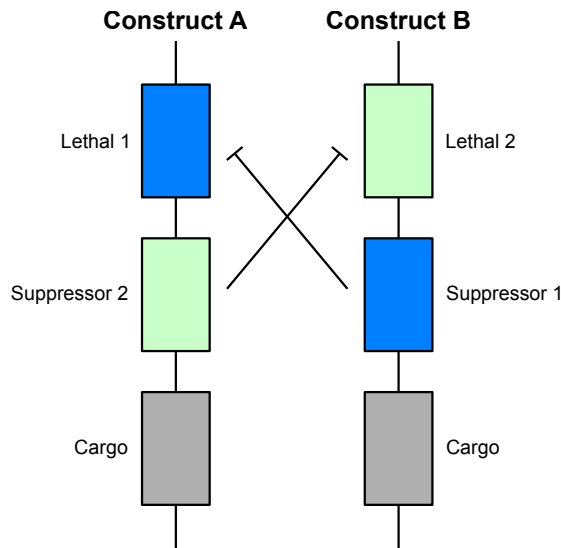


Figure 1: A schematic diagram of the engineered underdominance gene drive system. Each transgenic construct possesses three genes; a lethal, a suppressor for the lethal on the other construct and a desirable genetic cargo. Genotypes possessing one or more copies of both constructs (or none of either) are viable since lethals will be deactivated by the suppressor(s) on the other construct. Those genotypes carrying one or more copies of either construct but none of the other are non-viable since they have non-suppressed lethal genes.

2 Mathematical Modeling

We consider a population genetics model of UD gene drive in a panmictic (randomly mating), isolated (closed) population of infinite size. Mutation is allowed to occur in each transgenic construct at a constant rate (m per gene) and we assume that m is low enough for multiple mutations in a single generation to be neglected. This gives a total of 18 alleles, with nine at each genetic locus (i.e. $a, A, A_L, A_S, A_C, A_{LS}, A_{LC}, A_{SC}, A_{LSC}, b, B, B_L, B_S, B_C, B_{LS}, B_{LC}, B_{SC}$ and B_{LSC}) where lower case denotes wild-type alleles, upper case represents transgenic alleles and subscripts indicate loss-of-function mutations in a given gene (L =lethal, S =suppressor and C =cargo). This results in a total of 2,025 possible genotypes, 819 of which are non-viable.

The fitness of each genotype is expressed relative to wild-type and represents a reduction in survival due to the carrying of transgenic constructs. Many genotypes also suffer a lethal effect from non-suppressed lethal genes. These factors are combined to give the relative fitness of each genotype:

$$\Omega_i = \varepsilon_A^\beta \varepsilon_{A_M}^\phi \varepsilon_B^\mu \varepsilon_{B_M}^\psi (1 - \gamma_i), \quad (1)$$

where ε denotes the relative fitness per construct conferred by non-mutated (A, B) or mutated (A_M, B_M where $M = L, S, C, LS, LC, SC, LSC$) transgenic constructs. We assume relative fitnesses act multiplicatively and that all mutated transgenic constructs confer the same relative fitness regardless of the type or number of mutations (although resulting genotypes may confer a separate lethal effect). Exponents

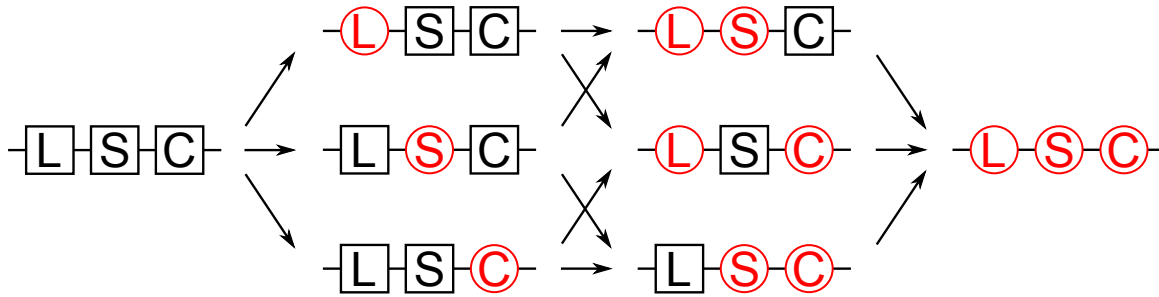


Figure 2: Transgenic constructs are assumed to mutate at a rate m per gene. This is assumed low enough that multiple mutations per generation may be neglected. For example the initial transgenic construct (say, A) mutates at a rate of $3m$, producing mutations in the lethal (giving A_L), suppressor (A_S) and the cargo (A_C) gene each at a rate of m . Then, transgenic constructs possessing one mutated gene (e.g. A_L) mutate at a rate of $2m$ (giving A_{LS} and A_{LC} each at rate m). Transgenic constructs with two mutated genes (e.g. A_{LS}) then mutate at rate m producing constructs with all three genes mutated (i.e. A_{LSC}). Here non-mutated genes are represented by black squares whereas genes with loss-of-function mutations are shown in red circles.

75 β , ϕ , μ and ψ denote the number of each transgenic construct type carried by a given genotype while γ_i
 76 represents the lethality of that genotype ($\gamma_i = 1$ if non-viable or $\gamma_i = 0$ if viable). Finally, we assume
 77 all lethal effectors are 100% effective when unsuppressed and that any number (i.e. one or two) of lethal
 78 effector copies are fully suppressed by any number (one or more) of the relevant suppressor (i.e. the strong
 79 suppression case in [9]).

80 For numerical simulations we used a set of MATLAB (MATLAB R2014b, The MathWorks Inc., Nat-
 81 ick, MA) scripts run in parallel using the MATLAB Parallel Computing Toolbox through a MATLAB
 82 Distributed Computing Server. These allow simulation of the system without a manual formulation of
 83 2,025 difference equations (see Figure 3). Briefly, the process begins by converting each possible genotype
 84 into a numerical form and computing relative fitnesses for each (using equation (1)). Initial conditions are
 85 then calculated according to:

$$G_{WT}^{t=0} = \frac{1}{1 + \alpha}, \quad G_{AABB}^{t=0} = \frac{\alpha}{1 + \alpha}, \quad (2)$$

86 where G_i^t is the i -th genotype frequency t generations after the initial release; α is the release ratio (intro-
 87 duced/wild) of non-mutated transgene double homozygotes ($AABB$) into a wild-type (abb) population;
 88 and all other genotypes have an initial frequency of zero (i.e. $G_i^{t=0} = 0$ for $i \neq WT, AABB$). We as-
 89 sume the absence of mutations in released individuals is feasible due to quality controls at the rearing
 90 facility. Genotype frequencies in subsequent generations are computed iteratively. Expected (Mendelian)
 91 frequencies are first calculated using a matrix of outcomes from every genotype mating pair and then
 92 multiplied by a matrix of mutation rates (giving proportional frequencies G_i^e). Finally, proportional geno-
 93 type frequencies (G_i^e) are normalized using the average fitness of the population ($\bar{\Omega} = \sum \Omega_i G_i^e$) such that
 94 genotype frequencies (G_i^t) sum to one.

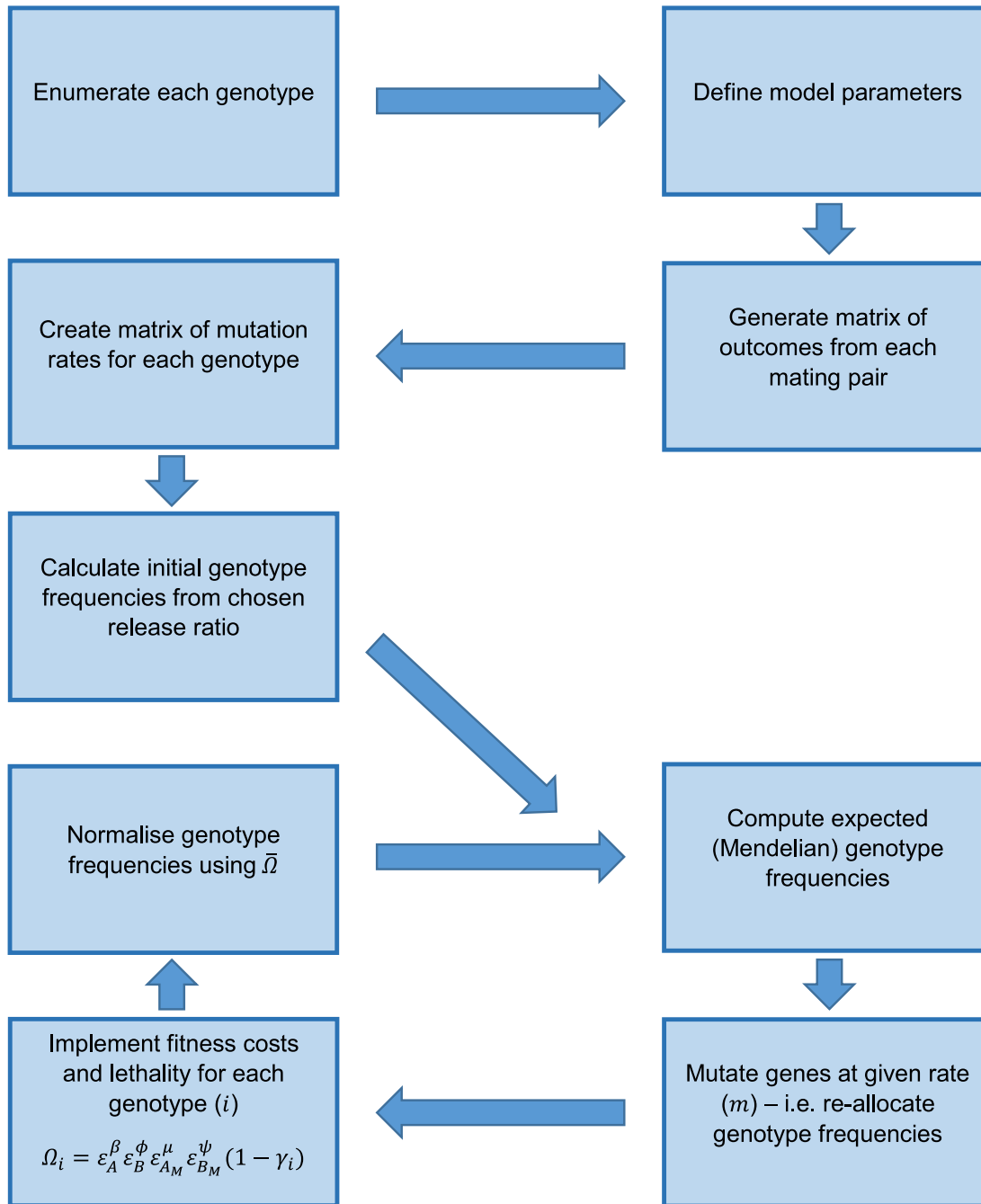


Figure 3: Diagram showing the simulation procedure for an engineered underdominance system with a constant rate of mutation (m) per gene. Parameter symbols used here are: ε , the relative fitness conferred by a given allele (denoted in the subscripts where M represents a mutated allele); β and ϕ , the numbers of non-mutated copies of construct A and B, respectively; μ and ψ , the number of mutated copies of transgenic constructs A and B, respectively; γ_i , the lethality conferred by a given genotype (i); Ω_i , the overall relative fitness of individuals of genotype i ; and $\bar{\Omega}$, the average relative fitness of the entire population.

Table 1: Table of parameter and variable definitions used throughout this study.

Symbol	Definition
G_i	Frequency of genotype (i)
G_i^e	Proportional (i.e. non-normalized) genotype frequency
m	Rate of mutation per gene
α	Release ratio (α =introduced/wild)
$\bar{\Omega}$	Average fitness of the overall population
Ω_i	Relative fitness of genotype (i)
ε_j	Relative fitness conferred by transgenic construct (j)
β	Number of non-mutated copies of transgenic construct A
ϕ	Number of mutated copies of transgenic construct A (i.e. A_M)
μ	Number of non-mutated copies of transgenic construct B
ψ	Number of mutated copies of transgenic construct B (i.e. B_M)
γ_i	Lethality of genotype (i)

3 Results

3.1 Numerical simulations with various rates of mutation per gene

To our knowledge, experimentally measured rates of mutation in key organisms considered as likely targets for UD gene drives are not well known. Thus, here we conduct numerical simulations for rates of mutation (per gene) spanning five orders of magnitude, namely $m = 10^{-4}$, 10^{-5} , 10^{-6} , 10^{-7} and 10^{-8} . Given previous estimates of the mutation rate in *Drosophila melanogaster* (2.8×10^{-9} [14] and 8.4×10^{-9} [15] per nucleotide per generation), the size of gene drive components used in previous studies (1-10kb e.g. [16–18]) and an approximation that 1-10% of nucleotides in these sequences are essential (i.e. causes a loss of gene function when mutated) we anticipate that this range should include rates relevant to a range of proposed target species.

Figure 4 shows results for 1:1 (introduced:wild) releases of double homozygote (i.e. $AABB$) individuals. We assume each non-mutated transgenic construct confers a fitness load of either 5% or 10% (i.e. $\varepsilon_A = \varepsilon_B = 0.95$ or 0.90) and mutated constructs 4% or 8% (i.e. $\varepsilon_{A_M} = \varepsilon_{B_M} = 0.96$ or 0.92), respectively. Thus, mutated transgenic constructs have a small fitness advantage over non-mutated ones, yet still a deficit relative to wild-type. In examples conducted with mutated transgenic alleles conferring a greater fitness cost than non-mutated alleles, the UD system progressed without significant accumulation of mutated alleles, thus we do not consider these cases any further.

In Figure 4 the non-mutated UD system initially reaches a high frequency, as previously modeled [5–11]. However, each type of mutated construct begins to accumulate in the population, with large frequencies reached by constructs carrying a single mutation in either the lethal or cargo gene. This results in a concurrent decrease in the frequency of non-mutated constructs. Once these mutated constructs have reached a high frequency in the population, they begin to be replaced by the remaining wild-type alleles due to the relative fitness advantage of wild-type alleles over mutated transgenic alleles. This eventually

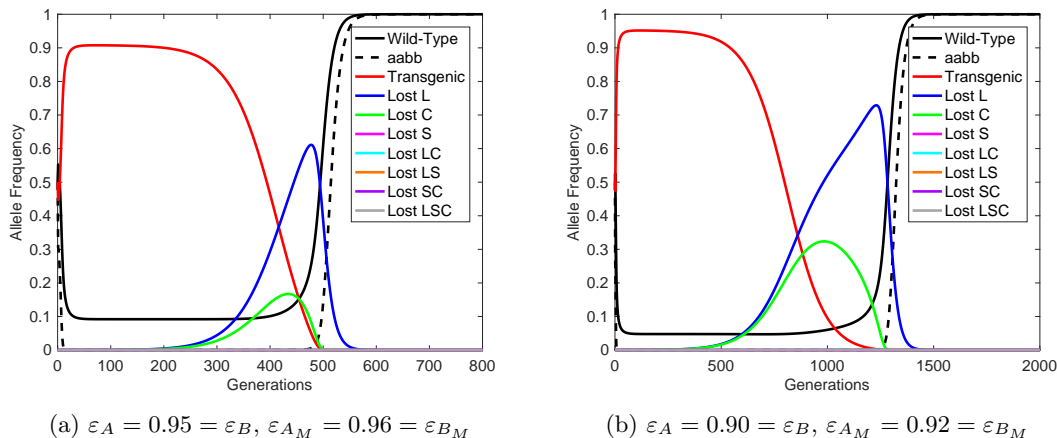


Figure 4: Mutation of transgenic constructs is predicted to return a fully wild-type population. Results here are presented for 1:1 (introduced:wild) introductions of non-mutated double homozygote (*AABB*) individuals into a wild-type (*aabb*) population. Here, transgenic constructs are assumed to mutate at a rate $m = 10^{-6}$ per gene, neglecting the possibility of multiple genes mutating within a single generation. Solid lines represent the allele frequency of wild-type, transgenic and each variety of mutated allele with details given in the figure legend whilst black dashed lines denote the genotype frequency of double wild-type homozygotes. Here a number of alleles only reach very low maximum frequencies (see Figure 5(a)) and thus they appear to overlie one another along the horizontal axis. Panel (a) shows results for $\epsilon_A = 0.95 = \epsilon_B$ and $\epsilon_{A_M} = 0.96 = \epsilon_{B_M}$ whereas (b) is for $\epsilon_A = 0.90 = \epsilon_B$ and $\epsilon_{A_M} = 0.92 = \epsilon_{B_M}$. Note the difference in time-scales between these panels.

118 returns the population to a fully wild-type state.

119 For the full range of mutation rates considered here we observe similar dynamics to those in Figure 4
 120 except that increasing (decreasing) mutation rates lead to faster (slower) accumulation and higher (lower)
 121 maximum frequencies of most types of mutated transgene allele (see Figure 5(a)). This inevitably means

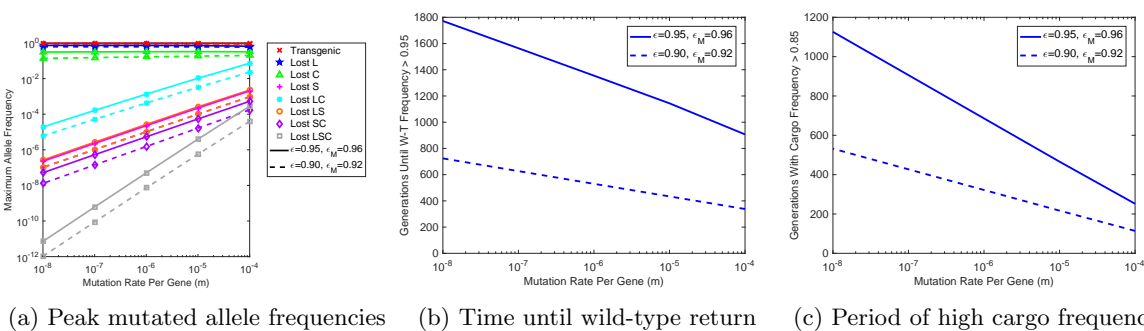


Figure 5: Rates of mutation in introduced engineered underdominance (UD) transgenes affect mutated allele frequencies and the efficacious period of the system. (a) Maximum frequencies attained by each of the non-mutated and mutated transgene alleles. Colors represent each different type of allele with details given in the figure legend. (b) The number of generations from the time of initial transgenic release until wild-type alleles return to a frequency greater than 0.95. (c) The number of generations that the UD system is able maintain a high frequency (>0.85) of transgenes with a functional copy of the cargo gene. In all panels solid lines represent cases with relative fitness parameters $\epsilon_A = 0.95 = \epsilon_B$ and $\epsilon_{A_M} = 0.96 = \epsilon_{B_M}$ whilst dashed lines are for cases with $\epsilon_A = 0.90 = \epsilon_B$ and $\epsilon_{A_M} = 0.92 = \epsilon_{B_M}$. These timings should represent a reasonable proxy for the period over which the desired phenotype conferred by the UD system would be effective.

122 that higher rates of mutation reduce the period over which the UD system persists (see Figure 5(b)) and
123 also the period for which functional cargo genes are present at high frequency (see Figure 5(c)). This
124 should form a reasonable proxy for the efficacious period of a released UD system.

125 3.2 Reversal strategies

126 To our knowledge, the only reversal strategy previously proposed for UD systems is the release of wild-
127 type individuals in sufficient numbers to push the transgene frequency to sub-threshold levels [12, 13].
128 When successful, wild-type alleles should recover and eliminate transgenes from the population. Whilst
129 this mechanism has been mentioned a number of times previously it has yet to be explored in any detail.

130 Based on results presented in Figure 4 here we propose an alternative, genetics-based, reversal strategy
131 for UD systems. In this strategy transgenic individuals carrying only the suppressor genes of the original
132 UD system would be released. These “free suppressor” carrying individuals are assumed to be of greater
133 fitness than those carrying the original UD constructs since they carry less genetic material and will also
134 suffer no lethal effect.

135 We conduct numerical simulations in order to compare the different reversal strategies discussed here.
136 Similar to the examples above, we consider the relative fitness conferred by the initial UD system to
137 be $\varepsilon_A = 0.95 = \varepsilon_B$ or $\varepsilon_A = 0.90 = \varepsilon_B$ and the free suppressor elements to confer a relative fitness of
138 $\varepsilon_{\text{Rev}} = 0.96$ or $\varepsilon_{\text{Rev}} = 0.92$, giving them a small fitness advantage over the original UD constructs and
139 a deficit relative to wild-type. Results are shown in Figure 6 for a 2:1 (introduced:wild (all genotypes))
140 release of wild-type individuals and individuals carrying free suppressors at both loci. These releases are
141 assumed to occur 100 generations after an initial 1:1 (introduced:wild) UD release so that the original
142 system is at high frequency and mutations would not have accumulated to significant frequencies. As such
143 we initially assume that no mutation occurs (i.e. $m = 0$) to ensure that free suppressors work in absence
144 of other types of allele.

145 Figure 6(b) demonstrates the feasibility of reversing of UD systems via the introduction of free sup-
146 pressors in the absence of mutated alleles. We then relax this assumption to consider the behavior of
147 a reversal drive in the presence of mutation (here we assume $A_{\text{Rev}} = A_{LC}$ and $B_{\text{Rev}} = B_{LC}$). These
148 numerical simulations demonstrate that this reversal strategy can function even with small releases (i.e.
149 $\alpha = 0.01$ and $\alpha = 0.1$; see Figure 7). However, one would need to be wary of stochastic effects when mak-
150 ing extremely small releases. These results imply that the release of free suppressor carrying individuals
151 could represent a very cost effective reversal strategy for UD systems.

152 It is clear that release of wild-type individuals provides a much faster option for returning a fully
153 wild-type population (Figure 6). However, releases of wild-type individuals must be significantly larger
154 (~ 1.75 times the wild population when $\varepsilon_A = 0.95 = \varepsilon_B$) in order to achieve reversal of the initial UD

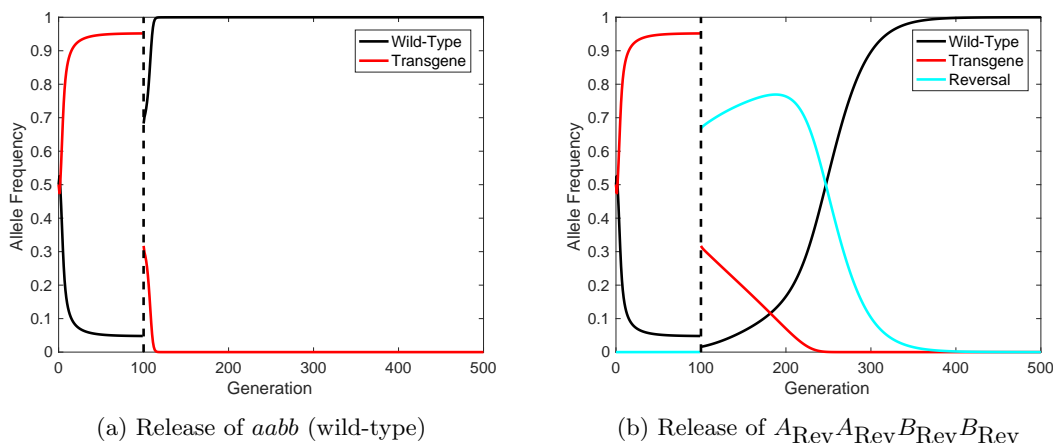


Figure 6: Comparison of two different reversal strategies for engineered underdominance (UD) gene drive systems. (a) An example of reversal through the release of wild-type ($aabb$) individuals. (b) Release of individuals carrying free suppressors at both loci (i.e. $A_{Rev}A_{Rev}B_{Rev}B_{Rev}$). Here line colors denote each type of allele with black representing wild-type (a, b); red denoting non-mutated transgenes (A, B); and cyan showing suppressor-only alleles (A_{Rev}, B_{Rev}). In each example the initial UD release is made at a 1:1 (introduced:wild) ratio with relative fitness parameters $\varepsilon_A = 0.95 = \varepsilon_B$. After 100 generations one of the reversal strategies is released at a ratio of 2:1 (introduced:wild).

155 system. This likely makes the release of wild-type individuals a far more costly reversal strategy than the
156 use of free-suppressor carrying individuals.

157 4 Discussion

158 Since gene drive systems were first proposed they have gained much attention for their potential to
159 help fight a number of important global issues (e.g. [1,2,19,20]). However, various genetic control measures
160 have been shown to be limited by the potential generation of mutation or resistance [18,21–25]. A number
161 of studies have also begun to investigate mechanisms capable of reversing gene drives in case they produce
162 unexpected consequences [26–28]. Here we showed that mutations will likely lead to the break down of
163 UD systems and the eventual elimination of introduced transgenes although under the assumptions used,
164 the UD system is likely to persist at high allele frequency for hundreds of generations. We then went on
165 to demonstrate that reversal of UD systems is feasible via the release of either wild-type individuals or
166 individuals carrying free suppressors.

167 As with all mathematical models, the work presented here relies on a number of simplifying assump-
168 tions, the majority of which are common in this type of study [5,6,9,11]. Since these assumptions have
169 been discussed previously, we do not consider them any further here. There are however a few areas specific
170 to this study where more detailed modeling would be useful to further elucidate the effects of mutation
171 in UD systems. Firstly, we assumed that mutations completely eliminate gene function whereas in reality
172 they may produce only a partial loss of function. We also assumed that all mutated transgenic constructs

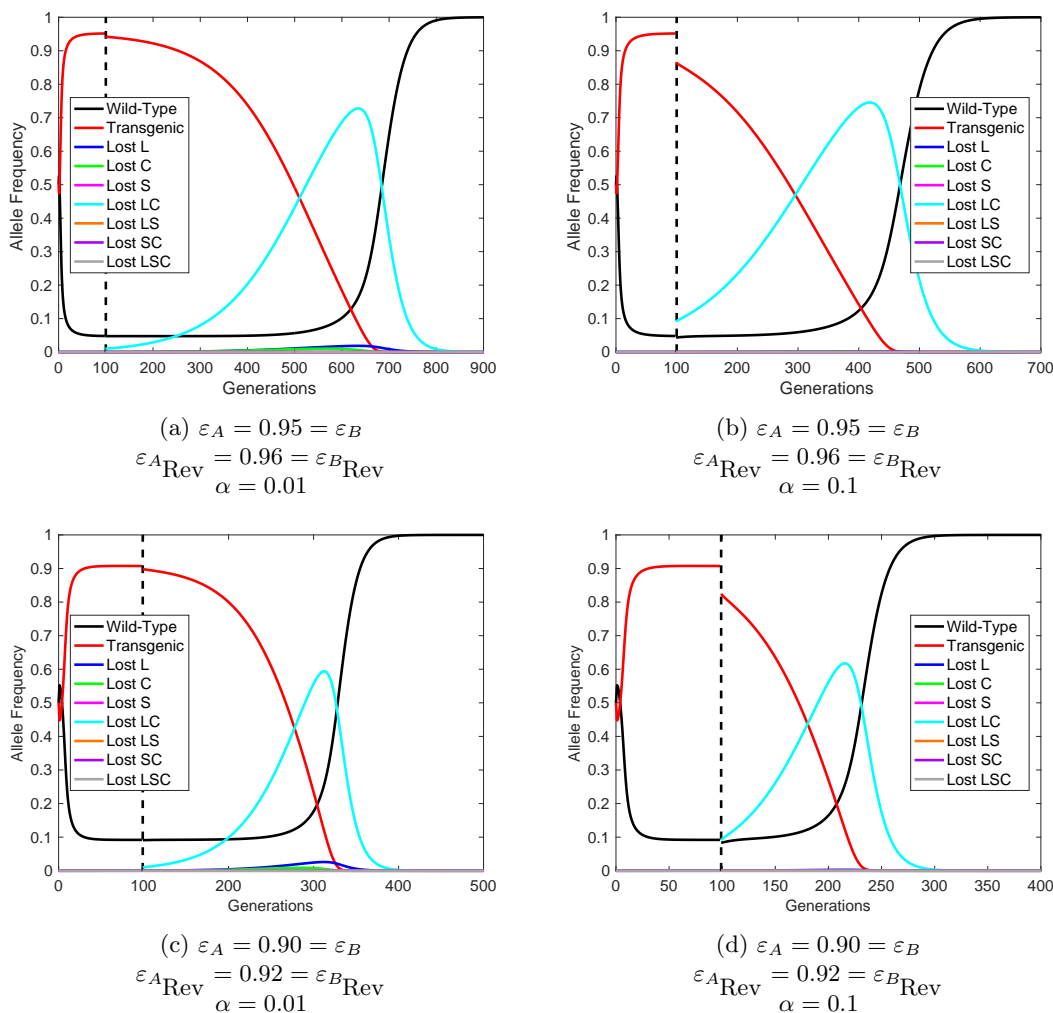


Figure 7: Free suppressor constructs can reverse engineered underdominance (UD) gene drives with small releases. Here, releases at generation 100 of free suppressor constructs (assuming $A_{\text{Rev}} = A_{LC}$ and $B_{\text{Rev}} = B_{LC}$) are numerically simulated for release ratios of $\alpha = 0.01$ (panels (a) and (c)) and $\alpha = 0.1$ (panels (b) and (d)) - far smaller than those required for reversal with wild-type. Lines represent the possible alleles with colors denoting the nature of each (details given in figure legends).

173 confer the same fitness cost regardless of the number and type of mutations carried. These assumptions
 174 could be relaxed in future models by expressing gene efficacy as a function of mutations carried. Finally, it
 175 is possible that natural genetic polymorphism in a target population could mean that certain individuals
 176 would be, at least partially, resistant to the effects of introduced genes. In such a case, the UD system
 177 may place resistance alleles under a selective pressure causing them to increase in frequency and reduce
 178 the efficacy of UD systems.

179 Results presented here assume that individuals carrying mutated transgenes are fitter than those car-
 180 rying non-mutated transgenic constructs. This is a common assumption when modeling the formation of
 181 resistance/mutations in gene drive systems (e.g. [21, 24]). The fitness advantage of individuals carrying
 182 specific mutations in each transgene (or combination of transgenes) is likely to determine the exact dynam-

183 ics and time scales that would be observed in experimental work. Results presented here are intended to
184 give a guide as to the type of behavior we would expect to emerge rather than giving definitive predictions
185 for specific applications.

186 Previous literature has estimated the mutation rate in *Drosophila melanogaster* to be $\sim 5.6 \times 10^{-9}$ per
187 nucleotide per generation (mean of estimates in [14] and [15]). Given the size of gene drive components
188 used in previous experimental studies (1-10kb e.g. [16–18]) and an approximation that 1-10% of nucleotides
189 in these sequences are essential (i.e. causes a loss of gene function when mutated), we believe the mutation
190 rates explored here to be feasible for a range of proposed target organisms. Using these mutation rates,
191 we assumed only one mutation was able to emerge in a single generation. In practice, deletions could
192 remove two adjacent components simultaneously. However, assuming the ordering of components shown
193 in Figure 1 such deletions would leave individuals without suppressor elements, thus making them non-
194 viable in presence the of the construct at the other locus. Coupling this with the extremely low probability
195 of multiple mutations occurring simultaneously, we do not anticipate that relaxing this assumption would
196 lead to major differences from the results presented here.

197 A number of recent studies have discussed the need to explore reversal strategies for gene drives in case
198 they have unexpected consequences. While the majority of work has focused on the reversal of CRISPR-
199 Cas9 gene drive systems (e.g. [26–29]), it has been suggested that the release of wild-type individuals could
200 reverse UD systems [12,13]. The release of wild-type individuals represents a threshold dependent reversal
201 strategy with the precise number of wild-types required to reverse a released UD system depending on
202 the fitness of UD carrying individuals and the size of the wild population, both of which are difficult to
203 measure. What is known however, is that the size of these wild-type releases would need to be very large.
204 For example, when transgenic constructs each confer a 5% fitness cost (i.e. $\varepsilon_A = 0.95 = \varepsilon_B$), the wild-type
205 release would need to be ~ 1.75 times the size of the entire wild population. Importantly, insufficiently
206 sized releases of wild-type individuals would fail to reverse a given UD system. While it would be feasible
207 to make further wild-type releases, this would have obvious cost implications. To help overcome these
208 issues, here we proposed a genetics-based alternative (release of free suppressor carriers) that appears
209 to be threshold independent (i.e. it is predicted to be effective from very small releases, leaving aside
210 stochastic effects) but will take longer to return a fully wild-type population than release of wild-types.
211 It may also be difficult to convince the public that releasing further transgenics is an acceptable method
212 of eliminating a gene drive that produced unexpected consequences. Here we do not wish to draw a firm
213 conclusion on which reversal strategy should be pursued, since this may depend on case-specific social,
214 economic and technical factors that are beyond the scope of this study.

215 In spite of demonstrating that UD gene drives will likely break down over time with the emergence
216 of mutations, this study provides reason to be optimistic about the prospect of using UD gene drives to

217 spread desirable genes through a target population. In particular, even though introduced transgenes are
218 likely to be eliminated, the desirable genetic cargo is expected to reach and maintain a high frequency
219 in the population for hundreds of generations. This is likely long enough for them to have produced
220 their desired effect. Here we have given the first theoretical examination of this UD break down due to
221 mutation. We anticipate that future modeling studies will be able to produce more application specific
222 models, thus refining the predictions presented here.

223 Acknowledgements

224 MPE is funded through a Wellcome Trust Investigator Award [110117/Z/15/Z] made to LSA. LSA is
225 supported by core funding from the UK Biotechnology and Biological Sciences Research Council (BBSRC)
226 to The Pirbright Institute [BBS/E/I/00007033].

227 References

- 228 [1] F. Gould. Broadening the application of evolutionarily based genetic pest management. *Evolution*,
229 62(2):500–510, 2008.
- 230 [2] L. Alphey. Genetic control of mosquitoes. *Annual Review of Entomology*, 59:205–224, 2014.
- 231 [3] A.W.E. Franz, I. Sanchez-Vargas, Z.N. Adelman, C.D. Blair, B.J. Beaty, A.A. James, and K.E. Olson.
232 Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes*
233 *aegypti*. *Proceedings of the National Academy of Sciences*, 103(11):4198–4203, 2006.
- 234 [4] A.W.E. Franz, I. Sanchez-Vargas, R.R. Raban, W.C. Black IV, A.A. James, and K.E. Olson. Fitness
235 impact and stability of a transgene conferring resistance to dengue-2 virus following introgression into
236 a genetically diverse *Aedes aegypti* strain. *PLoS Neglected Tropical Diseases*, 8(5):e2833, 2014.
- 237 [5] S. Davis, N. Bax, and P. Grewe. Engineered underdominance allows efficient and economical intro-
238 gression of traits into pest populations. *Journal of Theoretical Biology*, 212(1):83–98, 2001.
- 239 [6] K. Magori and F. Gould. Genetically engineered underdominance for manipulation of pest popula-
240 tions: A deterministic model. *Genetics*, 172(4):2613–2620, 2006.
- 241 [7] Y. Huang, A.L. Lloyd, M. Legros, and F. Gould. Gene-drive in age-structured insect populations.
242 *Evolutionary Applications*, 2(2):143–159, 2009.
- 243 [8] Y. Huang, A.L. Lloyd, M. Legros, and F. Gould. Gene-drive into insect populations with age and
244 spatial structure: A theoretical assessment. *Evolutionary Applications*, 4(3):415–428, 2011.

- 245 [9] M.P. Edgington and L.S. Alphey. Conditions for success of engineered underdominance gene drive
246 systems. *Journal of Theoretical Biology*, 430:128–140, 2017.
- 247 [10] J.M. Marshall. The effect of gene drive on containment of transgenic mosquitoes. *Journal of Theo-*
248 *retical Biology*, 258(2):250–265, 2009.
- 249 [11] J.M. Marshall and B.A. Hay. Confinement of gene drive systems to local populations: A comparative
250 analysis. *Journal of Theoretical Biology*, 294:153–171, 2012.
- 251 [12] J. Champer, A. Buchman, and O.S. Akbari. Cheating evolution: engineering gene drives to manipu-
252 late the fate of wild populations. *Nature Reviews Genetics*, 17(3):146–159, 2016.
- 253 [13] J.M. Marshall and O.S. Akbari. Gene drive strategies for population replacement. In *Genetic Control*
254 *of Malaria and Dengue*, pages 169 – 200. Academic Press, Boston, 2016.
- 255 [14] P.D. Keightley, R.W. Ness, D.L. Halligan, and P.R. Haddrill. Estimation of the spontaneous mutation
256 rate per nucleotide site in a *drosophila melanogaster* full-sib family. *Genetics*, 196(1):313–320, 2014.
- 257 [15] C. Haag-Liautard, M. Dorris, X. Maside, S. Macaskill, D.L. Halligan, B. Charlesworth, and P.D.
258 Keightley. Direct estimation of per nucleotide and genomic deleterious mutation rates in *drosophila*.
259 *Nature*, 445(7123):82–85, 2007.
- 260 [16] R.G. Reeves, J. Bryk, P.M. Altrock, J.A. Denton, and F.A. Reed. First steps towards underdominant
261 genetic transformation of insect populations. *PLoS One*, 9(5):e97557, 2014.
- 262 [17] N. Windbichler, M. Menichelli, P.A. Papathanos, S.B. Thyme, H. Li, U.Y. Ulge, B.T. Hovde,
263 D. Baker, R.J. Monnat, A. Burt, and A. Crisanti. A synthetic homing endonuclease-based gene
264 drive system in the human malaria mosquito. *Nature*, 473(7346):212–215, 2011.
- 265 [18] J. Champer, R. Reeves, S.Y. Oh, C. Liu, J. Liu, A.G. Clark, and P.W. Messer. Novel CRISPR/Cas9
266 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency
267 in genetically diverse populations. *PLoS Genetics*, 13(7):e1006796, 2017.
- 268 [19] T.A.A. Prowse, P. Cassey, J.V. Ross, C. Pfitzner, T.A. Wittmann, and P. Thomas. Dodging silver
269 bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. *Proc. R. Soc.*
270 *B*, 284(1860), 2017.
- 271 [20] R.E. Thresher, K. Hayes, N.J. Bax, J. Teem, T.J. Benfey, and F. Gould. Genetic control of inva-
272 sive fish: technological options and its role in integrated pest management. *Biological invasions*,
273 16(6):1201–1216, 2014.

- 274 [21] N. Alphey, M.B. Bonsall, and L. Alphey. Modeling resistance to genetic control of insects. *Journal*
275 *of Theoretical Biology*, 270(1):42–55, 2011.
- 276 [22] B. Watkinson-Powell and N. Alphey. Resistance to genetic insect control: Modelling the effects of
277 space. *Journal of theoretical biology*, 413:72–85, 2017.
- 278 [23] A.M. Hammond, K. Kyrou, M. Bruttini, A. North, R. Galizi, X. Karlsson, N. Kranjc, F.M. Carpi,
279 R. D’Aurizio, A. Crisanti, and T. Nolan. The creation and selection of mutations resistant to a gene
280 drive over multiple generations in the malaria mosquito. *PLOS Genetics*, 13(10):1–16, 10 2017.
- 281 [24] R.L. Unckless, A.G. Clark, and P.W. Messer. Evolution of resistance against CRISPR/Cas9 gene
282 drive. *Genetics*, 205(2):827–841, 2017.
- 283 [25] J. Champer, J. Liu, S.Y. Oh, R. Reeves, A. Luthra, N. Oakes, A.G. Clark, and P.W. Messer. Reducing
284 resistance allele formation in CRISPR gene drives. *bioRxiv*, page 150276, 2017.
- 285 [26] M.R. Vella, C.E. Gunning, A.L. Lloyd, and F. Gould. Evaluating strategies for reversing CRISPR-
286 Cas9 gene drives. *Scientific Reports*, 7(1):11038, 2017.
- 287 [27] J.E. DiCarlo, A. Chavez, S.L. Dietz, K.M. Esvelt, and G.M. Church. Safeguarding CRISPR-Cas9
288 gene drives in yeast. *Nature Biotechnology*, 33:1250–1255, 2017.
- 289 [28] B. Wu, L. Luo, and X.J. Gao. Cas9-triggered chain ablation of Cas9 as a gene drive brake. *Nature*
290 *Biotechnology*, 34:137–138, 2016.
- 291 [29] K.A. Oye, K. Esvelt, E. Appleton, F. Catteruccia, G. Church, T. Kuiken, S.B-Y. Lightfoot, J. Mc-
292 Namara, A. Smidler, and J.P. Collins. Regulating gene drives. *Science*, 345(6197):626–628, 2014.

Construct A

Construct B

