Lethal Gene Drive Selects Escape through Inbreeding

J J Bull

ABSTRACT

The use of ‘selfish’ gene drive systems to suppress or even extinguish populations has been proposed on theoretical grounds for almost half a century. Creating these genes has recently become possible with CRISPR technology. One seemingly feasible approach, originally proposed by Burt, is to create a homing endonuclease gene (HEG) that inserts into an essential gene, enabling heterozygote viability but causing homozygote lethality. With 100% segregation distortion in gametes, such genes can cause profound population suppression if resistance does not evolve. Here, population genetic models are used to consider the evolution of inbreeding (selfing) as a possible response to a recessively lethal HEG with complete segregation distortion. Numerical results indicate a rich set of outcomes, but selfing often evolves in response to the HEG. Whether selfing does indeed evolve and its effect in restoring population fitness depends heavily on the magnitude of inbreeding depression. Overall, these results point toward an underappreciated evolutionary response to block the harmful effects of a selfish gene. They raise the possibility that extreme population suppression may be more difficult to achieve than currently imagined.

INTRODUCTION

The proposed use of selfish genes to suppress or extinguish populations is at least half a century old (Hickey and Craig 1966a,b; Hamilton 1967), but the feasibility of actually engineering selfish genes is new. There has thus been much excitement about the possibility of using these approaches to eradicate disease vectors, balanced by concerns about the possibility of unforeseen harm. Perhaps the most tangible approach is one outlined by Burt (2003), of creating a homing endonuclease gene (HEG) that inserts itself into an essential gene. Under the idealized assumptions of 100% segregation distortion in gametes of heterozygotes (germ line only), normal heterozygote viability and fertility but homozygote lethality, such a selfish gene is expected to evolve to such an extreme as to cause a 50% reduction in population fecundity if the segregation distortion is limited to one sex (Bruck 1957; Lewontin 1958). A segregation distortion that operates in both sexes can evolve to fixation and death of all progeny, ensuring extinction (Prout 1953; Lewontin 1958; Burt 2003).

The HEG need not work as completely as expected. Extreme levels of population suppression from the HEG are sensitive to even minor variations in parameter values (Deredec et al. 2008; Unckless et al. 2015). More importantly, an HEG that harms population fitness will select resistance mechanisms. Since HEGs target specific DNA sequences, the most obvious form of resistance to the HEG is a change in the target sequence so that the HEG will no longer duplicate itself in heterozygotes (Burt 2003). Resistance could also take the form of interfering with endonuclease expression or functionality. The problem of target sequence evolution has been countered with the suggestion of deploying multiple HEGs simultaneously (Burt 2003), but other resistance mechanisms would not obviously be thwarted by that approach.

Here I address another possible resistance mechanism to invasion by a lethal HEG: evolution of inbreeding. It is appreciated that inbreeding reduces the population impact of ‘lethal’ HEGs (Burt 2003; Esvelt et al. 2014; Bull 2015). What is not clear is whether inbreeding is actually favored once a lethal HEG has invaded the population. Although a fixed level of inbreeding should reduce the incidence of the recessively lethal HEG, an allele that increases the level of inbreeding will itself suffer increased loss from any excess inviable progeny that it creates, perhaps selecting against inbreeding and even favoring increased outcrossing. It will in fact be shown here that inbreeding does evolve under some conditions, but the extent to which population fitness recovers depends heavily on the magnitude of inbreeding depression. Furthermore, the level of inbreeding that evolves is often not the level that would
maximize population fitness were inbreeding imposed on the population.

THE MODELS

All models assume a population of diploid hermaphrodites capable of a mix of outcrossing and selfing. Each individual produces a fixed amount of sperm and eggs. Eggs can be fertilized either by sperm chosen randomly from an outcross pool or by self sperm, from the individual who produced the ova.

There are two unlinked loci, each with two alleles. An individual’s level of selfing is controlled by its genotype at the A/a locus independently of the genotype at the other locus (Table 1). The D/d locus affects viability and experiences gametic drive. Specifically D is a recessive lethal. Such an allele would normally be lost, but its drive (segregation distortion) more than compensates: Dd is of normal viability and fertility because it remains heterozygous in somatic tissues, but it produces all D gametes in sperm and (in some models) also in ova. The DD genotype dies at conception, so viable genotypes at this locus are limited to dd and Dd.

Table 1 Genotype control of selfing rate

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Proportion ova selfed</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>$s_0$</td>
</tr>
<tr>
<td>Aa</td>
<td>$s_1$</td>
</tr>
<tr>
<td>AA</td>
<td>$s_2$</td>
</tr>
</tbody>
</table>

Important variations of the model differ in (i) whether drive operates just in sperm or in sperm and ova, and (ii) how selfing decreases an individual’s contribution of sperm to the outcross pool. The first of these properties has well appreciated consequences. Two-sex drive (operating in sperm and ova of the same Dd heterozygote) can evolve to the extreme that all offspring are inviable DD (Prout 1953), ensuring population extinction. Male drive can only evolve to the point that half the offspring are DD (half are Dd) (Bruck 1957; Lewontin 1958). This results in a 50% reduction in population fitness, which may or may not be sufficient for extinction. The two cases have different consequences on the selection of selfing.

The second property, the amount of sperm lost from selfing, also has ramifications for the selection of selfing. If a parent creates selfed offspring with no loss in sperm to the outcrossing pool (sperm constant model), selfing is favored unless the fitness of selfed offspring is less than half that of outcrossed offspring (Lande and Schemske 1985). In a selfed offspring carrying two copies of the parent’s alleles, the genetic benefit to the parent of a selfed offspring is twice that of an outcrossed offspring. If selfed offspring instead cause a loss in the individual’s sperm pool in proportion to the fraction of eggs selfed (sperm reduced model), then selfing is favored only if selfed offspring fitness exceeds that of outcrossed offspring.

Thus four discrete-generation population genetic models will be analyzed (equations in Appendix). To specifically assess the impact of drive on the evolution of selfing, the analyses use ranges of selfed offspring fitness ($\sigma$) that would prevent the evolution of selfing in the absence of drive:

- Male drive, sperm constant. $0 < \sigma < 0.5$
- Male drive, sperm reduced. $0 < \sigma < 0.5$
- Two-sex drive, sperm constant. $0 < \sigma < 0.5$
- Two-sex drive, sperm reduced. $0 < \sigma < 1$
- Two sex drive, sperm reduced. $0 < \sigma < 1$

Model iterations will usually assume that the population is initially approximately 95% genotype addd (fully outcrossed) with the drive genotype aadd at 5%. The A allele raises the level of selfing and is started at 0.015%. This condition would apply if drive has started to invade a population in which selfing alleles are rare.

RESULTS

Fig. 1 shows sample results for the four models, each illustrated for 2 different $\sigma$ levels (fitness of selfed offspring). Each panel uses a solid blue line for mean fitness in the absence of the HEG across different levels of inbreeding (inbreeding rate is specified on the horizontal axis) The line has slope $\sigma - 1$. Each panel also uses a dashed orange curve to show mean fitness in the presence of the HEG for a fixed, uniform level of inbreeding. As selfing rate increases from 0, the yellow dashed curve rises until it intersects the blue line and then declines, coinciding with the blue curve. The rise in the yellow curve is from the reduced impact of the HEG on mean fitness under inbreeding. The HEG is lost for all selfing rates at which the yellow and blue curves coincide, so the decline in that segment is from the loss in fitness from increased exposure to inbreeding depression.

Each panel also shows several black triangles. These triangles are representative outcomes of evolution at the selfing locus; the figures show only the outcomes at or near the highest evolved mean fitnesses observed across several trials. For each set of conditions, some runs were done in which the AA genotype was completely selfed; the level of selfing by Aa was never greater than that for AA.

Several points are noteworthy. (1) Given appropriate selfing rates for the Aa and AA genotypes, selfing evolved and increased mean fitness at least slightly, but only if $\sigma$ was large enough [see point (3)]. Mean fitness often did not attain the maximum that could be obtained when selfing was imposed on a population with D. (2) The highest evolved mean fitness exceeded the fitness of selfed offspring in the sperm constant models but not in the sperm reduced models. This difference is likely due to the ‘extra’ male fitness achieved by selfers in the sperm constant models and the resulting depression of the HEG frequency. (3) In the male-drive models, selfing never evolved if the fitness of selfed offspring ($\sigma$) was too low. Invasion analyses indicated that $\sigma$ needed to exceed 0.5 in the sperm reduced model and 0.25 in the sperm constant model. (4) The drive allele (D) was lost only when a selfing rate of 1 could evolve (see below, however).

Selection on the magnitude of selfing exhibited some peculiarities. (5) Alleles (A) with a high selfing rate could be favored when alleles with a lower selfing rate were not. For example, in Fig. 1H, dominant A genotypes with selfing rates of 0.9 and 1.0 were favored whereas those with rates 0.8 and lower were not favored.

(6) Invasion of the A allele did not ensure its ascent to fixation. Again considering Fig. 1H, the leftmost triangle in that panel is for selfing rates of 0.9 for both AA and Aa genotypes, yet the A allele invaded only to the point that the mean population selfing rate was 0.013.

Dynamics anomalies

Initial conditions were sometimes found to drive radical differences in outcomes. For the model and $\sigma$ value in Fig. 1A, an initial population that was 94.8% addd and only 5% aadd (with rare introductions for Aa and AA) led to a temporary rise in D but then lost D completely and lost the selfing allele (if Aa and AA selfing rates
The fitness of selfed offspring ($\sigma$) is given in each panel. The blue line is mean fitness for the selfing rate given on the lower axis in the absence of the HEG (allele $D$). The yellow, dashed line is equilibrium mean fitness when $D$ has evolved to equilibrium, under a uniform selfing rate given on the lower axis. $D$ is lost for all points on the yellow curve that intersect the blue line. The black triangles represent equilibria when selfing was allowed to evolve, showing outcomes at or near the highest mean fitness observed. Except for some runs in Figs 1A and 1C (see text), initial frequencies for the black triangle runs were 0.948 ($aadd$), 0.05 ($aaDd$), 0.001 ($Aadd$, $AaDd$) and 0 for the others.

**Adding loci**

Models restricted to two alleles at one selfing locus may not provide full insight to long term evolutionary dynamics. In particular, once selfing has been introduced into the population, alleles may be favored that refine the level of selfing, even though such alleles are not capable of invading at the start. A 3-locus model was created to investigate this possibility (two loci – $A/a$ and $B/b$ – affected selfing, the third locus encoded the drive). The case of Fig. 1H was evaluated, as the mean fitness evolved under the 2-locus model was substantially below the maximum possible in the presence of the HEG. The initial conditions specified selfing rates of 0, 0.9, 0.9 for the $A$, $A$, and $B$ genotypes, as led to the modest invasion of selfing in Fig. 1H. The $B$ allele forced increases in the lowest level of selfing but maintained 0.9 as the highest level, thus evolution of $B$ would have forced the baseline level of selfing above 0. None of these parameter combinations or the others tested resulted in an increase of $B$, suggesting that a multi-allelic selfing system does not easily evolve to the level of selfing that would maximize mean fitness in the presence of the HEG.

**DISCUSSION**

There is much justified excitement about the possibility of employing gene drive systems to limit wild populations of undesirable species (Sinkins and Gould 2006; Gould et al. 2008; Burt 2014; Esvelt et al. 2014; Unckless et al. 2015). In particular, homing endonuclease genes (HEGs) can now be developed that are recessive lethals but experience close to complete segregation distortion in heterozygotes (Gantz and Bier 2015). Such HEGs can theoretically spread to fixation (extinction) or to the point that the entire viable population is heterozygous (Prout 1953; Bruck 1957; Lewontin 1958), with a major reduction in mean fitness. Evolution of resistance to the
HEG, or to its effects, becomes highly relevant in understanding the possible limitations of these engineered systems.

This study indicates that the final frequency of a recessive lethal affecting complete segregation distortion can be reduced by the evolution of inbreeding, with a consequent increase in mean fitness above that which evolves in the absence of selfing. The models here specifically assumed selfing in hermaphrodites, but the principle will no doubt extend to other forms of inbreeding, such as sib mating in species with separate sexes. The fitness mitigation achieved by selfing is limited largely by the magnitude of inbreeding depression: evolution of inbreeding brings population fitness close to its maximum of 1 only if inbreeding depression is miniscule. Furthermore, if drive is limited to one sex, mean fitness from an unchallenged HEG is 1/2, and inbreeding depression can be too extreme for inbreeding to be favored at all.

There also appear to be limited conditions in which selfing can invade and drive out the invading HEG, ultimately restoring the population to a state of full outcrossing and no HEG. These cases required a high fitness of selfed offspring and were specific to certain initial genotype frequencies. A more extensive analysis will be required to identify the full parameter space and initial conditions allowing such an outcome, but it is not clear that they are of general interest. For now, a conservative approach is to accept that the HEG will invade and persist, except when complete selfing evolves.

The impact of inbreeding on selfish, gene drive systems has long been appreciated. Hamilton (1967) in describing the evolution of distorting sex chromosomes and their potential use as population extinction mechanisms showed that inbreeding reduced the level of sex chromosome drive that would be favored. Burt (2003) and Esvelt et al. (2014) both commented that inbreeding would retard the impact of otherwise harmful gene drive systems. The outcome of this work has been to show that inbreeding is actually favored under many scenarios in which a gene drive system hampers the population.

The empirical feasibility of evolving high levels of inbreeding during assault by a recessive lethal HEG remains to be seen. We have essentially no field experience with gene drive systems. There is, however, over half a century of experience with various forms of sterile insect applications (Bushland et al. 1955; Whitten 1971; Dyck et al. 2005). Sterile insect techniques almost universally rely on the release of lab-reared insects that, when mated with wild insects, cause death or sterility of the progeny (Sinkins and Gould 2006). (Those lab-reared insects may be irradiated or mutated in other ways or may carry chromosomal rearrangements that are incompatible with the wild population.) The assault from sterilizing, lab-reared insects should also favor inbreeding as one of several mechanisms that avoid matings that produce sterile progeny. Yet assortative mating by wild populations subjected to the sterile insect technique has rarely been reported, despite many applications and successes (Bull 2015). In this comparison, it may be important that the sterile insect technique relies on inundation of the wild population with lab-reared insects, the wild individuals becoming increasingly overwhelmed with sterility-inducing matings as the population declines. Gene drive systems may have the opposite effect, encouraging de facto consanguinity as the population density declines.

ACKNOWLEDGMENTS

Supported by NIH GM 57756.

LITERATURE CITED


APPENDIX

The basic recursion equations are common to all models. Assuming discrete non-overlapping generations:

\[
D'_{2,0} = S_{2,0} + M_{1,0}F_{1,0} \\
D'_{2,1} = S_{2,1} + M_{1,1}F_{1,0} + M_{1,0}F_{1,1} \\
D'_{0,0} = S_{0,0} + M_{0,0}F_{0,0} \\
D'_{1,0} = S_{1,0} + M_{0,1}F_{0,0} + M_{0,0}F_{0,1} \\
D'_{0,1} = S_{0,1} + M_{0,1}F_{0,0} + M_{0,0}F_{1,1} \\
D'_{1,1} = S_{1,1} + M_{1,1}F_{0,0} + M_{1,0}F_{0,1} + M_{0,1}F_{1,0} + M_{0,0}F_{1,1}
\]

Here, \(D_{i,j}\) represents a diploid adult, with \(i\) \(A\) alleles and \(j\) \(D\) alleles. Diploids may have 0, 1 or 2 \(A\) alleles but only 0 or 1 \(D\) alleles because \(D\) is a recessive lethal. \(M_{k,n}\) (\(F_{k,n}\)) represents outcrossed male (female) gametes with \(k\) \(A\) alleles and \(n\) \(D\) alleles. \(S_{i,j}\) represents diploid offspring by selfing, with similar subscripting as for \(D_{i,j}\). Recursion equations for the gametes and selfed offspring are specific to each model, as follows. It is assumed that all ova are fertilized, and sperm frequencies are normalized. Although the body of the paper restricts the analyses to complete drive, these equations use an unsubscripted \(D\) to represent the fraction of gametes of a \(D\) heterozygote carrying the \(D\) allele, facilitating the derivation. Selfing rates are \(s_0\) for \(aa\), \(s_1\) for \(Aa\), and \(s_2\) for \(AA\), regardless of the other locus.

**Male-female drive, sperm constant**

**Sperm**

\[
TM_{1,0} = D_{2,1}(1 - s_2)(1 - D) + D_{1,1}(1 - s_1)(1 - D)/2 \\
+ D_{2,0}(1 - s_2) + D_{1,0}(1 - s_1)/2
\]

\[
TM_{0,0} = D_{1,1}(1 - s_1)(1 - D)/2 + D_{0,1}(1 - s_0)(1 - D) \\
+ D_{1,0}(1 - s_1)/2 + D_{0,0}(1 - s_0)
\]

\[
TM_{1,1} = D_{2,1}(1 - s_2)D + D_{1,1}(1 - s_1)D/2
\]

\[
TM_{0,1} = D_{1,1}(1 - s_1)D/2 + D_{0,1}(1 - s_0)D
\]

\[
T = M_{1,0} + M_{0,0} + M_{1,1} + M_{0,1}
\]

**Ova**

\[
F_{1,0} = D_{2,1}(1 - s_2)(1 - D) + D_{1,1}(1 - s_1)(1 - D)/2 \\
+ D_{2,0}(1 - s_2) + D_{1,0}(1 - s_1)/2
\]

\[
F_{0,0} = D_{1,1}(1 - s_1)(1 - D)/2 + D_{0,1}(1 - s_0)(1 - D) \\
+ D_{1,0}(1 - s_1)/2 + D_{0,0}(1 - s_0)
\]

\[
F_{1,1} = D_{2,1}(1 - s_2)D + D_{1,1}(1 - s_1)D/2
\]

\[
F_{0,1} = D_{1,1}(1 - s_1)D/2 + D_{0,1}(1 - s_0)D
\]

**Selfed**

\[
S_{0,0} = [D_{1,0}s_1/4 + D_{0,0}s_0 + D_{1,1}s_1(1 - D)(1 - D)/4 + D_{0,1}s_0(1 - D)(1 - D)]s_0
\]

\[
S_{2,0} = [D_{2,0}s_2 + D_{1,0}s_0s_2(1 - D)(1 - D) + D_{1,1}s_1s_0(1 - D)(1 - D)]s_0
\]

\[
S_{1,0} = [D_{1,0}s_1^2 + D_{1,1}s_1(1 - D)(1 - D)/2]s_0
\]

\[
S_{2,1} = [2D_{2,1}s_2D(1 - D) + D_{1,1}s_1D(1 - D)/4]s_0
\]

\[
S_{0,1} = [D_{1,1}s_1D(1 - D)/2 + 2D_{0,1}s_0D(1 - D)]s_0
\]

\[
S_{1,1} = [D_{1,1}s_1D(1 - D)]s_0
\]

**Male drive, sperm constant**

**Sperm**

\[
TM_{1,0} = D_{2,1}(1 - D) + D_{1,1}(1 - D)/2 + D_{2,0} + D_{1,0}/2
\]

\[
TM_{0,0} = D_{1,1}(1 - D)/2 + D_{0,1}(1 - D) + D_{1,0}/2 + D_{0,0}
\]

\[
TM_{1,1} = D_{2,1}D + D_{1,1}D/2
\]

\[
TM_{0,1} = D_{1,1}D/2 + D_{0,1}D
\]

\[
T = M_{1,0} + M_{0,0} + M_{1,1} + M_{0,1}
\]

**Ova**

\[
F_{1,0} = D_{2,1}(1 - s_2)(1 - D) + D_{1,1}(1 - s_1)(1 - D)/2 \\
+ D_{2,0}(1 - s_2) + D_{1,0}(1 - s_1)/2
\]

\[
F_{0,0} = D_{1,1}(1 - s_1)(1 - D)/2 + D_{0,1}(1 - s_0)(1 - D) \\
+ D_{1,0}(1 - s_1)/2 + D_{0,0}(1 - s_0)
\]

\[
F_{1,1} = D_{2,1}(1 - s_2)D + D_{1,1}(1 - s_1)D/2
\]

\[
F_{0,1} = D_{1,1}(1 - s_1)D/2 + D_{0,1}(1 - s_0)D
\]

**Selfed**

\[
S_{0,0} = [D_{1,0}s_1/4 + D_{0,0}s_0 + D_{1,1}s_1(1 - D)(1 - D)/2]s_0
\]

\[
S_{2,0} = [D_{2,0}s_2 + D_{1,0}s_0s_2(1 - D)(1 - D) + D_{1,1}s_1s_0(1 - D)(1 - D)]s_0
\]

\[
S_{1,0} = [D_{1,0}s_1^2 + D_{1,1}s_1(1 - D)(1 - D)/4]s_0
\]

\[
S_{2,1} = [2D_{2,1}s_2D(1 - D) + D_{1,1}s_1D(1 - D)/4]s_0
\]

\[
S_{0,1} = [D_{1,1}s_1D(1 - D)/2 + 2D_{0,1}s_0D(1 - D)]s_0
\]

\[
S_{1,1} = [D_{1,1}s_1D(1 - D)]s_0
\]
Male drive, sperm reduced

**Sperm**

\[
TM_{1,0} = D_{2,1}(1 - s_2)(1 - D) + D_{1,1}(1 - s_1)(1 - D)/2 \tag{11}
\]
\[
+ D_{2,0}(1 - s_2) + D_{1,0}(1 - s_1)/2
\]
\[
TM_{0,0} = D_{1,1}(1 - s_1)(1 - D)/2 + D_{0,1}(1 - s_0)(1 - D)
+ D_{1,0}(1 - s_1)/2 + D_{0,0}(1 - s_0)
\]
\[
TM_{1,1} = D_{2,1}(1 - s_2)D + D_{1,1}(1 - s_1)D/2
\]
\[
TM_{0,1} = D_{1,1}(1 - s_1)D/2 + D_{0,1}(1 - s_0)D
\]
\[
T = M_{1,0} + M_{0,0} + M_{1,1} + M_{0,1}
\]

**Ova**

\[
F_{1,0} = D_{2,1}(1 - s_2)/2 + D_{1,1}(1 - s_1)/4
\]
\[
+ D_{2,0}(1 - s_2) + D_{1,0}(1 - s_1)/2
\]
\[
F_{0,0} = D_{1,1}(1 - s_1)/4 + D_{0,1}(1 - s_0)/2
+ D_{1,0}(1 - s_1)/2 + D_{0,0}(1 - s_0)
\]
\[
F_{1,1} = D_{2,1}(1 - s_2)/2 + D_{1,1}(1 - s_1)/4
\]
\[
F_{0,1} = D_{1,1}(1 - s_1)/4 + D_{0,1}(1 - s_0)/2
\]

**Selfed**

\[
S_{0,0} = [D_{1,0}s_1/4 + D_{0,0}s_0
\]
\[
+ D_{1,1}s_1(1 - D)/8 + D_{0,1}s_0(1 - D)/2]\sigma
\]
\[
S_{2,0} = [D_{2,0}s_2 + D_{1,0}s_1/4 + D_{2,1}s_2(1 - D)/2 + D_{1,1}s_1(1 - D)/8]\sigma
\]
\[
S_{1,0} = [D_{1,0}s_1/2 + D_{1,1}s_1(1 - D)/4]\sigma
\]
\[
S_{2,1} = [D_{2,1}s_2/2 + D_{1,1}s_1/8]\sigma
\]
\[
S_{0,1} = [D_{1,1}s_1/8 + D_{0,1}s_0/2]\sigma
\]
\[
S_{1,1} = [D_{1,1}s_1/4]\sigma
\]