

Genetically Engineered Underdominance for Manipulation of Pest Populations: A Deterministic Model

Krisztian Magori¹ and Fred Gould

Department of Entomology, North Carolina State University, Raleigh, North Carolina 27607

Manuscript received September 29, 2005

Accepted for publication January 10, 2006

ABSTRACT

We theoretically investigate the potential for introgressing a desired engineered gene into a pest population by linking the desired gene to DNA constructs that exhibit underdominance properties. Our deterministic model includes two independently segregating engineered constructs that both carry a lethal gene, but suppress each other. Only genotypes containing both or neither construct are viable. Both constructs also carry the desired gene with an independent regulatory mechanism. We examine the minimal number of individuals of an engineered strain that must be released into a natural population to successfully introgress the desired gene. We compare results for strains carrying single and multiple insertions of the constructs. When there are no fitness costs associated with the inserted constructs (when the lethal sequences are not expressed), the number of individuals that must be released decreases as the number of insertions in the genome of the released strain increases. As fitness costs increase, the number of individuals that must be released increases at a greater rate for release strains with more insertions. Under specific conditions this results in the strain with only a single insertion of each construct being the most efficient for introgressing the desired gene. We discuss practical implications of our findings.

RECENTLY, there has been considerable discussion regarding the potential for decreasing the incidence of mosquito-borne diseases such as malaria and dengue on the basis of spreading engineered genes for refractoriness (*i.e.*, inability to transmit the pathogen) into native mosquito populations (GOULD and SCHLIEKELMAN 2004). Because strains bearing these engineered refractory genes are not expected to have higher fitness than native mosquitoes (CATTERUCCIA *et al.* 2003; IRVIN *et al.* 2004; MOREIRA *et al.* 2004), an engineered gene with Mendelian inheritance will not increase in frequency after being released. Thus, it will be necessary to develop genetic drive systems to spread the genes controlling refractoriness.

Substantial advances have been made in engineering refractory mosquito strains (*e.g.*, ITO *et al.* 2002), but only limited efforts have focused on drive mechanisms and the population genetic factors that could affect their success (RIBEIRO and KIDWELL 1994; KISZEWSKI and SPIELMAN 1998; DAVIS *et al.* 2001; BURT 2003; GOULD and SCHLIEKELMAN 2004).

Underdominance is one mechanism that has been proposed for driving desirable genes into populations (CURTIS 1968a,b; DAVIS *et al.* 2001). Underdominance is classically defined as the genetic condition where the

fitness of heterozygote individuals is lower than the fitness of both of the parental homozygotes (HARTL and CLARK 1989). There are two stable and one unstable equilibria in such a system with two alleles. At the alternate stable equilibria, one of the alleles is typically fixed and the other is lost from the population. The relationship between the initial allelic frequencies and the unstable equilibrium determines which allele becomes fixed. Underdominance is typically considered to involve a single locus, and fitness is considered as survival and reproduction in a single generation. However, if first-generation heterozygotes are as fit as their homozygote parents, but their offspring are less fit than the offspring of homozygotes in a specific population, the same three equilibria can exist. For example, when individuals that are homozygous for alternate forms of a chromosomal translocation are mated, the heterozygotes can be quite fit because their genomes contain one copy of all genes from both parents (ROBINSON 1976). When these heterozygotes mate with each other a fraction of their offspring lack important chromosomal segments and are not viable. Therefore if we expand our view of fitness to multiple generations, this can be considered as a case of underdominance.

A novel form of genetically engineered underdominance has been suggested by DAVIS *et al.* (2001). They used a simple population genetics model to analyze the spread of two engineered constructs in a hypothetical pest population, assuming that the two constructs were inserted into two nonhomologous chromosomes (see

¹Corresponding author: Department of Entomology, North Carolina State University, 840 Method Rd., Unit 1, Raleigh, NC 27607.
E-mail: kmagori@ncsu.edu

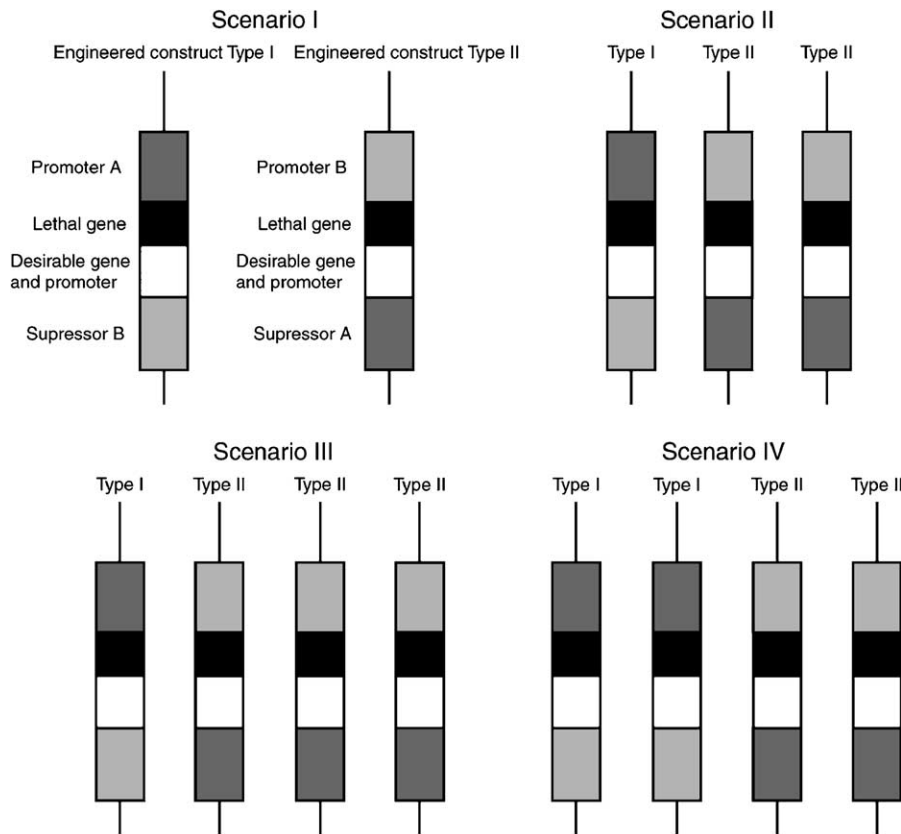


FIGURE 1.—Set of engineered constructs on separate linkage groups in the different scenarios. Individuals carrying at least one type I and one type II construct are viable by the suppression of the lethal gene. The lethal gene is expressed in individuals that have one or more copies of one type of the constructs but lack the other type.

Figure 4 of DAVIS *et al.* 2001, top left diagram). Both engineered constructs contained a lethal gene that was suppressed by a DNA sequence on the alternate construct. In addition to the lethal gene, each construct contained the sequence for a desirable gene and its independent promoter. DAVIS *et al.* (2001) assumed that the engineered constructs had no fitness cost when expression of the lethal gene was suppressed. Individuals containing both or neither of the engineered constructs were assumed to be viable, while all insects having only one of the constructs died due to the expression of the lethal gene. They showed that successful spread of the constructs was possible even with relatively small releases (~27% of the natural population) compared to other genetic control strategies (see, *e.g.*, BUSHLAND *et al.* 1955).

This is less than the 50% threshold for translocations (CURTIS 1968b) because several genotypes that contain three copies of the engineered constructs are viable in the case of engineered underdominance, while the same genotypes are not viable in the case of translocations (see supplemental Figure 1 in supplemental materials, part C, at <http://www.genetics.org/supplemental/>).

In this article, we further investigate and elaborate on the approaches examined by DAVIS *et al.* (2001). First, we investigate the spread of the desirable gene when multiple copies of the engineered constructs are inserted. Second, we consider the effects of fitness costs on the

minimum release numbers needed for the spread of the engineered constructs. Finally, we determine the optimal release method for a particular fitness cost and number of insertions of each construct.

MATERIALS AND METHODS

We modeled a diploid, obligately sexual pest species with each of two to four transgenic constructs inserted on independently assorting linkage groups (*i.e.*, different chromosomes or sometimes distant locations on the same chromosome). Each linkage group could either contain a transgenic construct (α , β , γ , δ) or lack it (A , B , C , D). There are two types of transgenic constructs (Figure 1), each type containing a promoter sequence for a lethal gene, a lethal gene, a desirable gene plus its independent promoter, and a sequence that suppresses the lethal gene promoter on the alternate construct.

Both types of constructs can contain the same lethal and desirable genes. However, the suppressor sequence in the type I construct must suppress only the lethal gene promoter sequence in the type II construct, and vice versa. If a type I construct is present in an individual lacking the type II construct, the expression of the lethal gene is not suppressed, and the individual dies. The same outcome befalls individuals that only have the type II construct. The only viable individuals are those that have none of the transgenic constructs or at least one copy of both types.

The simulation model was developed and built in Visual C++ in the Visual Studio.NET environment, to have up to eight copies of the constructs, 16 gamete types, and 81 genotypes; equations can be found in supplemental materials, parts

A and B, at <http://www.genetics.org/supplemental/>. We assumed Mendelian inheritance and a large panmictic population.

We studied scenarios in which the number and proportion of insertions of the two transgenic constructs differed. In scenario I (Figure 1), only two linkage groups had transgenic constructs, the first a type I and the second a type II, respectively. In scenario II, three linkage groups contained transgenic constructs, with the first linkage group having a type I construct and the second and the third linkage groups having one type II construct each. In scenario III, four linkage groups contained a transgenic construct, but only one had a type I construct. Each of the other three linkage groups contained a type II construct. In scenario IV, two linkage groups had type I and two linkage groups had type II constructs.

Two release strategies were examined: a single temporal release of transgenic insects into the population and releases of transgenic insects in multiple generations. Every simulation started with a native population completely lacking transgenic constructs. In the single-release studies, individuals homozygous for the transgenic constructs were released in specific proportions to native insects at the start of the simulation. In the multiple-release studies, individuals homozygous for the transgenic constructs were released in specific proportions in the first 20 generations. Genotypes that were determined inviable in a particular scenario due to engineered underdominance always had a zero fitness. Each copy of a transgenic construct was assumed to have an equal fitness cost (c) when the lethal gene was suppressed, and the constructs were assumed to interact in a multiplicative manner, unless stated otherwise. The fitnesses of viable genotypes were calculated by multiplying the relative fitnesses of the inserted transgenic constructs they contained, independently of their position,

$$S(k, l, m, n) = (1 - c_1)^k(1 - c_2)^l(1 - c_3)^m(1 - c_4)^n, \quad (1)$$

where $S(k, l, m, n)$ is the fitness of a genotype with up to eight total insertions of the constructs. $k, l, m,$ and n are the numbers of insertions of the constructs on linkage groups 1, 2, 3, and 4, respectively. The values for $k, l, m,$ and n can be 0, homozygous for no insertions; 1, heterozygous for the inserted construct; and 2, homozygous for the inserted construct. c is the fitness cost due to insertion of a single construct in a linkage group (see SCHLIEKELMAN and GOULD 2000).

We studied the minimal fraction of released insects to native insects (p_m) necessary to drive the desired gene to high frequency (e.g., >90%) in the population. We systematically searched a large set of single-release fractions in a recursive manner for the minimal fraction. We investigated the minimal fraction of release at different fitness costs of the transgenic constructs for single and multiple releases.

We explored how the recessivity of the fitness costs of the transgenic constructs affects the minimal fraction of release (p_m). The fitness of individuals that were homozygous for an inserted construct on one linkage group was $(1 - c)^2$, while the fitness of individuals heterozygous for the same inserted construct was $(1 - c)^{2d}$, where d is the degree of dominance of the fitness cost of the transgenic construct. We determined the minimal fraction of release at increasing, completely recessive fitness costs, in the case of both single and multiple releases.

RESULTS

Single release in scenario II without fitness cost:

Figure 2a shows the DeFinetti diagram of scenario II with the assumption of zero fitness costs of the transgenic constructs. The DeFinetti diagram (HARTL and CLARK 1989) is a triangular plot used to map the gamete

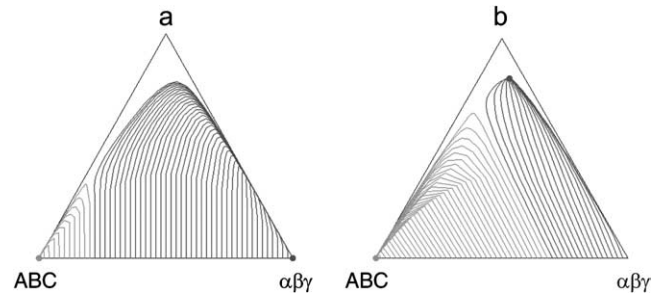


FIGURE 2.—DeFinetti diagrams of scenario II. Diagrams show the trajectories of the composition of gamete frequencies. The 50 trajectories depicted for each scenario all start from the lower boundary representing an initial ratio of homozygote individuals with and without all of the constructs. Light gray trajectories lead to the left-hand corner equilibrium (light gray dot) with the elimination of transgenic constructs. Dark gray trajectories lead to the equilibrium of transgenic constructs (dark gray dot), allowing the persistence of transgenic constructs. (a) No fitness costs of the transgenic constructs, $c = 0$. (b) Fitness costs of the transgenic construct, $c = 0.2$.

dynamics (Figure 2). The left-hand corner corresponds to the frequency of the gamete type completely devoid of transgenic constructs, while the right-hand corner corresponds to the frequency of the gamete type having the maximum possible number of transgenic constructs in the corresponding scenario. The top corner corresponds to the sum of any other possible gamete types for a specific scenario. For example, in scenario II (Figure 2), the gamete frequencies are mapped to the horizontal coordinate $\alpha\beta\gamma D_k - ABCD_k$ and the vertical coordinate $(1 - ABCD_k - \alpha\beta\gamma D_k)\sqrt{3}$, where, e.g., $ABCD_k$ is the frequency of the $ABCD$ gamete type in the k th generation. Each line in the diagram represents the trajectory of gametic composition from a specific initial condition. All trajectories start from the bottom boundary, which represents a pure mixture of homozygote individuals. End results of releases are represented by a large dot, typically on one of the boundaries. Light gray lines represent trajectories of gametic compositions that result in fixation of the gamete type lacking transgenic constructs. Dark gray lines represent trajectories of gametic compositions that result in the persistence of the transgenic constructs and reduction of the frequency of the gamete type lacking transgenic constructs.

After the trajectories move upward due to decreasing linkage disequilibrium, gametic frequencies always converge to one of two stable manifolds (one leading to the left-hand corner and the other leading to the right-hand corner). Both the left- and the right-hand corners are stable equilibrium points for a certain range of initial conditions. The diagram contains a unique release-fraction threshold, which is one point of the separatrix where it meets the horizontal axis. The separatrix is a set of gametic frequencies that are all unstable equilibria, correspond to a saddle point of the dynamics, and separate the basins of attraction of the two stable equilibria.

Since we restricted the initial conditions to the lower boundary, only one unstable equilibrium point of the separatrix is visible.

DeFinetti diagrams of all scenarios can be viewed in supplemental Figure 2 in supplemental materials, part D, at <http://www.genetics.org/supplemental/>. While these are generally similar to the one shown in Figure 2a, there are important differences in the DeFinetti diagrams representing the four scenarios. The most important difference is the different position of the separatrix (minimal fraction of release). With no fitness costs, there is a clear decrease in the minimal fraction of release from scenario I through scenario IV (scenario I, 0.269; scenario II, 0.203; scenario III, 0.172; and scenario IV, 0.154).

Single release in scenario II with fixed fitness costs in all generations: Figure 2b shows the DeFinetti diagram of scenario II with fitness cost $c = 0.2$ (*i.e.*, a reduction in fitness of 20% for each copy of a construct in an individual). There are significant differences compared to the case without fitness costs. First, the position of the separatrix moves to the right, representing an increase in the minimal fraction of release. Second, all trajectories have an initial tilt to the left, which means a decrease in the frequency of the gamete type that harbors all transgenic constructs and an increase in the frequency of the gamete type that has none. Third, the proportion of the natural gamete type in the stable equilibrium composition of all gamete types increases as the fitness costs of the transgenic constructs increase. This lack of fixation is not discussed in other studies of underdominance. Large proportions of the natural gamete type in the stable equilibrium composition would be detrimental to any effort to drive a refractory gene into a pest population. However, this proportion does not exceed 10% for any fitness cost in any scenario in our simulations, which translates to a maximum of 1% of individuals in a population having no insertions and not expressing the desirable gene.

The saddle point and the stable equilibrium move toward each other with increasing cost. This continues until the two equilibria collide at a certain fitness cost, annihilating each other in a saddle-node bifurcation and leaving the elimination of all transgenic constructs as the only stable equilibrium. Above this critical fitness cost, large selection against the transgenic constructs makes their spread impossible with any fraction of release.

DeFinetti diagrams of scenarios I, III, and IV are presented in supplemental materials, part E, at <http://www.genetics.org/supplemental/>.

Single and multiple releases with multiplicative fitness costs: Figure 3a shows the minimal fraction of release (p_m) at different fitness costs (c) of the transgenic constructs with single release for every scenario. With no cost, scenario IV requires the lowest minimal fraction of release and is the most efficient scenario to drive the desired gene into the hypothetical pest population.

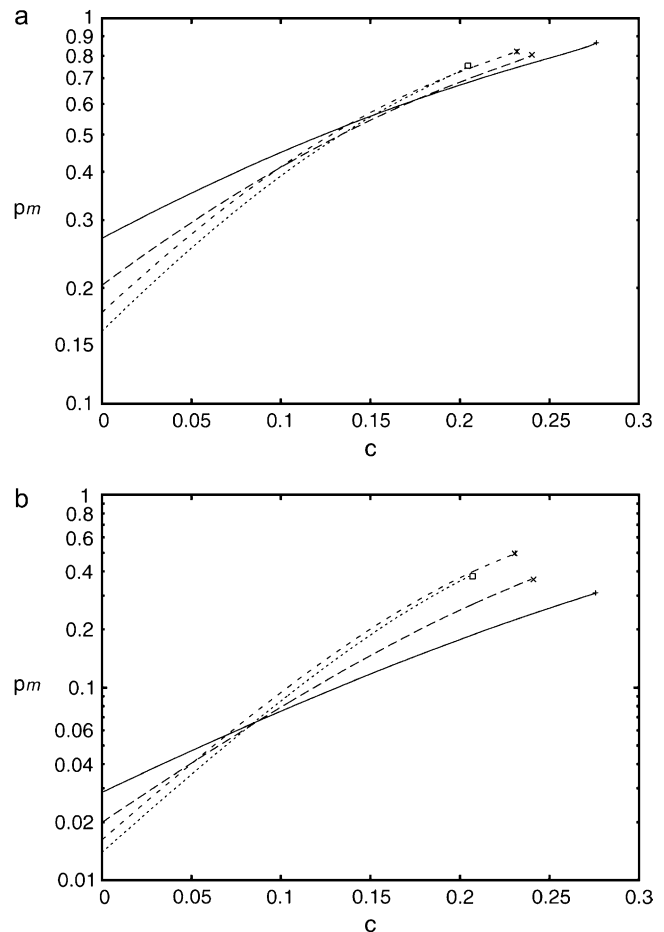


FIGURE 3.—The effect of increasing multiplicative fitness cost (c) per transgenic construct on the minimal fraction of release (p_m). —, scenario I; ---, scenario II; - · -, scenario III; and ···, scenario IV. At fitness costs above the symbols the approach is not successful for any fraction released. (a) Single release. (b) Multiple releases in the first 20 generations (note different logarithmic scales for p_m).

Minimal fraction of release increases dramatically with increasing fitness cost in every scenario. The rate of increase is most rapid in scenario IV and is the slowest in scenario I. At fitness costs of ~ 0.15 , the minimal fractions of release in the different scenarios are almost identical. At high fitness costs the order of the scenarios in efficiency is reversed compared to the results assuming no fitness costs, and the minimal fraction of release is very large in every scenario. This indicates that for the single-release approach, it is imperative to achieve a low fitness cost to efficiently drive the desirable gene into a population.

For each scenario, Figure 3b shows the minimal fraction of release in the case of multiple releases with fitness costs in all generations. With no cost at all, scenario IV requires the lowest minimal fraction of release. The rate of increase in the minimal fraction of release with increasing fitness cost is the most rapid in scenario III, and it is the slowest in scenario I. At fitness costs of ~ 0.075 the minimal fractions of release in the different

scenarios are similar. At high fitness costs the order of the scenarios in efficiency is almost reversed compared to the results assuming no fitness costs. The range of fitness costs in which scenario IV is the most efficient is much smaller than that in either of the single-release cases. In the case of multiple release, the slopes of p_m for scenarios III and IV are even more similar than in the case of single release, and scenario IV is more efficient than scenario III at any fitness cost.

Results of single and multiple releases with multiplicative fitness costs of the transgenic constructs applied only after the release generation are presented in supplemental materials, part H, at <http://www.genetics.org/supplemental/>.

Single and multiple releases with recessive fitness costs: Our investigation has also shown that more recessive inheritance of fitness costs increases the efficiency of all strategies (see supplemental Figure 5 in supplemental materials, part G, at <http://www.genetics.org/supplemental/>). Recessive fitness cost increases the efficiency of underdominance strategies, because in the transient period of the dynamics most individuals are heterozygous for the transgenic insertions and express small fitness costs. We chose to investigate the effect of increasing, completely recessive fitness costs (c) on the minimal threshold of release (p_m). Figure 4a shows the minimal fraction of release (p_m) at different recessive fitness costs (c) of the transgenic constructs with single release. Scenario IV remains the most efficient approach at all fitness costs. Also, the maximal fitness costs at which all scenarios succeed are greater than in the case of multiplicative fitness costs (the abscissa of Figure 4a is longer than the abscissa of Figure 3a). Figure 4b indicates that recessiveness is less important in the case of multiple release than in the case of single release, due to the repeated introductions of homozygous insects with high fitness costs. Results of single and multiple releases with completely recessive fitness costs of the transgenic constructs after the release generation, but with no fitness costs at all in the release generation, are presented in supplemental materials, part I, at <http://www.genetics.org/supplemental/>.

Effect of fitness costs on the rate of increase in construct frequency: An issue of practical interest is the number of generations necessary to substantially lower the proportion of individuals that are completely non-transgenic (*i.e.*, those that do not express the desired gene). To address this issue, we studied the number of generations necessary to reach a target of 0.1 for the proportion of individuals in a population not containing any transgenic constructs, as a function of the fitness costs (c) (Figure 5). We investigated only the case of single release. The initial release fraction of individuals completely homozygote for all transgenic constructs was set at 0.01 above the minimal fraction of release (p_m) for the corresponding fitness cost (c) of the transgenic constructs. In all scenarios, the number of generations

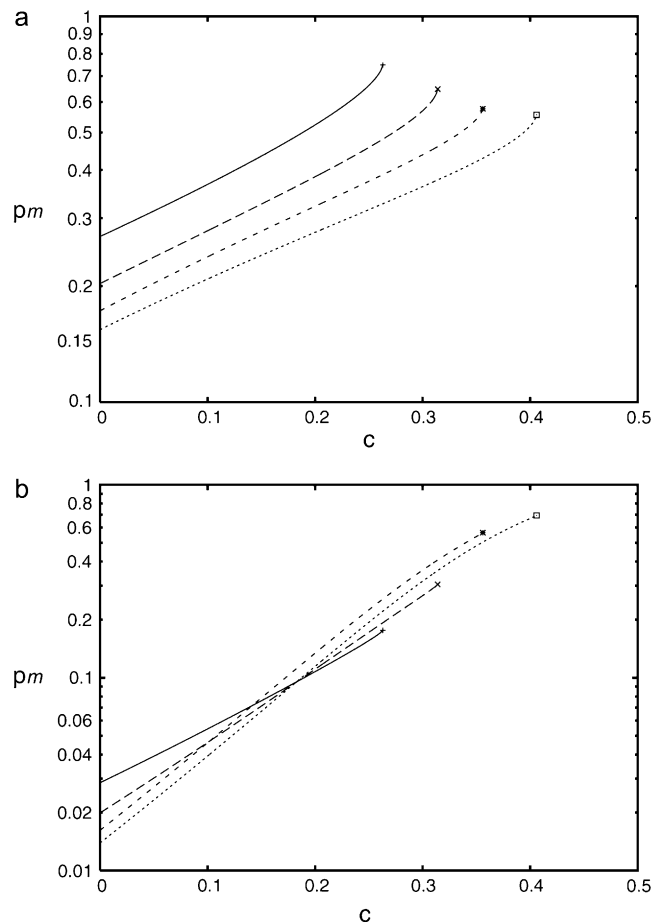


FIGURE 4.—The effect of increasing, completely recessive fitness cost (c) per transgenic construct on the minimal fraction of release (p_m). —, scenario I; ---, scenario II; - - -, scenario III; and ···, scenario IV. At fitness costs above the symbols the approach is not successful for any fraction released. (a) Single release. (b) Multiple releases in the first 20 generations (note different logarithmic scales for p_m).

necessary to reach the target proportion of 0.1 of individuals devoid of any transgenic constructs actually decreases as fitness costs increase. This result is in part due to the increasing minimal fraction of release (p_m) associated with increasing fitness costs. Increasing fitness costs do slightly inhibit the rate of increase of the desirable gene into the population (see supplemental Figure 4 in supplemental materials, part F, at <http://www.genetics.org/supplemental/> for further details), and this can be important when desired target proportions are very low (*e.g.*, 0.02).

DISCUSSION

Our results demonstrate that when fitness costs are low, increasing the number of inserts of the transgenic constructs decreases the minimal fraction of release. This is due to the positive relationship between the number of engineered inserts in the release strain and the total number of insects produced over time that

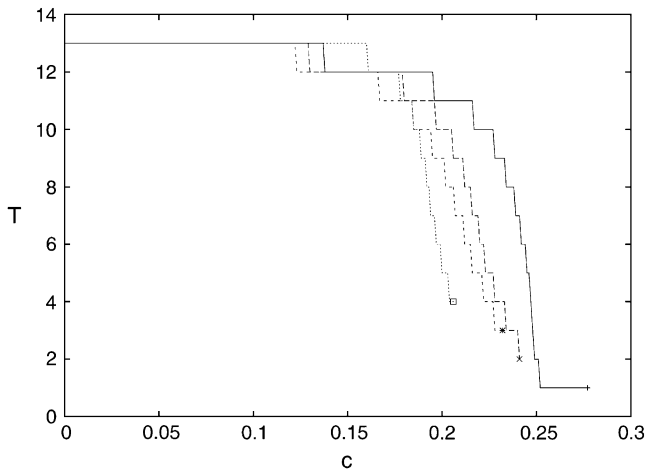


FIGURE 5.—The effect of increasing, multiplicative fitness cost per transgenic construct on the number of generations necessary (T) to reach the target proportion of 0.1 nontransgenic insects in the case of single release. Initial release fraction of individuals completely homozygous for all transgenic constructs is 0.01 above the minimal fraction of release at the corresponding fitness cost. At fitness costs above the symbols the approach is not successful for any fraction released. —, scenario I; ---, scenario II; ···, scenario III; and ···, scenario IV.

have only one of the two types of constructs. Our results also demonstrate that the minimal release fraction increases with increasing fitness costs. The relationship between minimal fraction of release and fitness cost differed among scenarios, and high nonrecessive fitness costs led to the complete reversal in the order of efficiency of the scenarios for single releases and a partial reversal for multiple releases, when compared to results under the assumption of no fitness cost.

Although the transgenic constructs required for establishing underdominance have not yet been produced in a laboratory, steps in this direction have been made (e.g., HEINRICH and SCOTT 2000; THOMAS *et al.* 2000; HORN and WIMMER 2003; GONG *et al.* 2005). Our results stress the potential of such a system for autocidal control and the need to move ahead in developing constructs with these properties. The most often discussed pest targets include several mosquito species (e.g., *Aedes aegypti*, *Anopheles gambiae*), which are important vectors of human diseases, but this approach could also be used with insects that vector crop diseases. Development of appropriate DNA constructs will require a major research effort, and it is clear that unless both constructs can be added simultaneously, an exogenous source of repressor will be needed to inhibit expression of the toxin-coding gene (e.g., GONG *et al.* 2005). The technological hurdles may be high, but so may be the payoff.

The assumption of zero fitness cost for transgenic insertions in a natural environment is unrealistic. However, we show that if each inserted copy of the constructs had a 5% fitness cost and there were multiplicative interactions among copies, a desirable gene could be driven into the natural population with realistic initial

release ratios. Our general result is that in the case of small fitness costs or recessive inheritance of fitness costs, the release of individuals with four transgenic constructs is the most effective approach of the scenarios examined. Although we are aware of the complications that the design of such an engineered insect line would face (GOULD and SCHLIEKELMAN 2004), our results suggest that it is worth consideration, because even a 10% reduction in the number of released insects could result in large financial savings at the implementation stage of a project. For example, scenario IV with fitness costs $c = 0.05$ necessitates the release of only 25% of the wild individuals, compared to 35% for scenario I. Wild populations of *Ae. aegypti* generally have a census size in the magnitude of 10,000–20,000 adults per village (SCOTT *et al.* 2000). This would require the release of 2500–5000 transgenic individuals per village in the case of scenario IV and 3500–7000 in the case of scenario I, respectively. Such release sizes are small compared to releases in classical sterile male release efforts. BUSHLAND *et al.* (1955) released 68,000 screw-worm flies weekly, and KNIPLING *et al.* (1968) stated that the weekly production of 1,000,000 *Ae. aegypti* mosquitoes was available on a routine basis in the 1960s.

The assumption of the fitness costs being completely recessive makes scenario IV the most efficient approach for all feasible fitness costs and widens the range of fitness costs at which there is successful introgression of desirable genes into the theoretical pest population.

The use of a multiple-release approach for 20 generations (~ 2 years for an insect species with 10 generations per year) offers a potentially more feasible possibility for strain replacement, when large numbers of release insects cannot be produced simultaneously. For example, insertions on four linkage groups according to scenario IV with fitness costs $c = 0.05$ require the release of only $\approx 3.5\%$ of the wild individuals in each generation. The assumption of completely recessive fitness costs allows success with the release of $\approx 2.2\%$ in each generation. However, overall numbers of released insects over the span of a project are typically larger in the case of smaller multiple releases compared to a larger single release (e.g., 3.5% releases for 20 generations equals a total of 7000–14,000 insects assuming a wild population of 10,000–20,000).

As fitness costs increase, the value of using multiple insertions often decreases when measured strictly according to the number of insects that must be released. However, there is another potential benefit from the use of multiple insertions. Our preliminary experiments with slightly unequal fitness costs of the transgenic constructs show that when there is more than one copy of each type of transgenic construct, all but the one copy of each construct with the lowest fitness cost will be eliminated. When we add even minute differences to the fitness costs associated with the individual transgenic constructs, the equilibrium points in all scenarios

and cases revert back to the two insert equilibrium frequencies seen in scenario I. When researchers develop insect strains with these constructs, they will presumably measure the fitness costs associated with each insertion in laboratory and field cages tests. However, the history of genetic manipulation of pests indicates that these lab and cage tests are often not good predictors of fitness costs in the natural environment (GOULD and SCHLIEKELMAN 2004). If a strain is released that has only one insertion of each construct, and in the field one of these constructs has a large, unanticipated fitness cost, the drive system will fail. If the release strain has two insertions of each construct then one costly insertion of a construct may be lost from the system while the second insertion is fixed in the population, achieving the desired goal. However, differences in fitness costs of the individual constructs also affect the position of the separatrix even when the sum of all fitness costs remains constant. The analysis of the model with nonequal fitness costs is extremely complicated and will be the subject of future investigation.

There has been criticism of a number of genetic drive systems based on concern that as the engineered construct is reproduced in nature it will mutate (or recombine) in a way that will eliminate expression of the desirable gene (CARARETO *et al.* 1997; HAHN and NUZHIDIN 2004). In the case of engineered underdominance, mutations could occur in several parts of the constructs with different results. While mutations in either of the suppressors of the lethal gene would be lethal and eliminated from the population, changes in the desirable gene could be selectively neutral. Mutations that debilitate the lethal gene could hinder the mechanism by which the engineered constructs can spread early in the release. However, it is highly unlikely that such a mutation would appear in even one of the constructs before the engineered constructs would increase to high frequency in the population. Mutations that occur in the promoter of the desirable gene, most notably the ones that shut down the production of the desirable gene, could indeed be selected for, especially if the physiological fitness costs due to the desirable gene are significant. In this case, after a successful release, where initially almost no individuals were able to transmit a pathogen, infectious insects would start to reappear. However, for this to happen, independent mutations would have to occur in the promoters of the desirable genes in both of the engineered constructs harbored in the individuals. Another advantage of the engineered underdominance approach over systems such as the transposon drive mechanism is that the engineered constructs are stable insertions. Transposons that lose the desirable gene or contain a mutated version of the desirable gene that causes less fitness cost to the insect will spread faster than the original transposon that contains the intact desirable gene. By the end of the release, most individuals will contain ineffective trans-

posons (HAHN and NUZHIDIN 2004). However, while selection of the transposons happens even inside of the individual mosquitoes during transposition, selection operates on the individuals only in the case of engineered underdominance constructs.

Our results clearly demonstrate that constructs that have recessive expression of fitness costs are expected to be more useful for introgressing desired genes compared to constructs with more dominant expression of fitness costs. It is typically expected that an insertion that disrupts an existing gene on one chromosome may have a low fitness cost as long as the homologous gene on the complementary chromosome is still active. Thus, if the fitness cost is due to gene disruption, the expression may be recessive. In contrast, if the fitness cost is due to incomplete repression of the lethal gene in a construct, expression of the cost may be additive.

When there are fitness costs associated with the insertions of the constructs, our model predicts that in very large populations a small number of insects may remain in the population that have no transgenic constructs and could therefore transmit disease organisms. For malaria and dengue it has often been shown that reducing the density of a mosquito population by 70–80% is sufficient for reducing disease (*e.g.*, YAPABANDARA and CURTIS 2004). Therefore, a strain replacement approach that decreases the percentage of pathogen-transmitting mosquitoes by >90% should generally be effective.

The ability of transgenic constructs with engineered underdominance to spread spatially is pivotal for the practical application of this approach. Even when the transgenic constructs maintain their frequency locally, their spatial spread is not ensured. The frequencies of the transgenic constructs during dispersal may still drop below the minimal fraction of release (p_m) in local populations, and the constructs could eventually be lost. PIALEK and BARTON (1997) concluded that a sufficiently strong physical or genetic barrier to gene flow can prevent the spread of an advantageous allele, when there is heterozygote disadvantage and the allele cannot increase from low frequency. In such cases, genetic drift may be able to free the allele from the local population it is trapped in. SOBOLEVA *et al.* (2003) showed that in general underdominant diploid systems a critical gene carrier aggregation (an analogy of critical nuclei in phase transition theory) has to be achieved to ensure the spatial spread of the alleles. This aggregation depends on both the number of the individuals carrying the alleles and their localization. Eventually, the spread of alleles from a cluster of gene carrier individuals is decided on the basis of the boundaries of that cluster (GANDHI *et al.* 1999). If spatial spread is not expected, all releases will need to be evenly spaced throughout the target area.

In conclusion, we found that over a broad range of conditions engineered underdominance can be an

effective drive mechanism, and multiple copies of engineered constructs increase its effectiveness at realistic fitness costs. However, our theoretical results should also caution empirical researchers developing such transgenic constructs. Fitness costs associated with the engineered constructs have to be low to result in spread of the desirable gene into the wild population. Generally, insertions with a fitness cost below $c = 0.15$ have a substantial chance of being effective.

We thank T. W. Scott, A. Lloyd, J. Rasgon, and D. O'Brochta for useful discussion of theoretical and practical aspects of this study. A. Lloyd, P. Schliekelman, and J. Rasgon offered helpful comments on the manuscript. T. Magori designed Figure 1. This research was in part supported by National Institutes of Health (NIH) grant R01-AI54954-01A2 to F.G. and funded in part by a grant from the Foundation for the NIH through the Grand Challenges in Global Health initiative.

LITERATURE CITED

- BURT, A., 2003 Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **270**: 921–928.
- BUSHLAND, R. C., A. W. LINDQUIST and E. F. KNIPLING, 1955 Eradication of screw-worms through release of sterilized males. *Science* **122**: 287–288.
- CARARETO, C. M. A., W. KIM, M. F. WOJCIECHOWSKI, F. MARTIN, P. O'GRADY *et al.*, 1997 Testing transposable elements as genetic drive mechanisms using *Drosophila* Pelement constructs as a model system. *Genetica* **101**: 13–33.
- CATTERUCCIA, F., H. C. GODFRAY and A. CRISANTI, 2003 Impact of genetic manipulation on the fitness of *Anopheles stephensi* mosquitoes. *Science* **299**: 1225–1227.
- CURTIS, C. F., 1968a A possible genetic method for the control of insect pests, with special reference to tsetse flies (*Glossina* spp.). *Bull. Entomol. Res.* **57**: 509–523.
- CURTIS, C. F., 1968b Possible use of translocations to fix desirable genes in insect pest populations. *Nature* **218**: 368–369.
- DAVIS, S., N. BAX and P. GREWE, 2001 Engineered underdominance allows efficient and economical introgression of traits into pest populations. *J. Theor. Biol.* **212**: 83–98.
- GANDHI, A., S. LEVIN and S. ORSZAG, 1999 Nucleation and relaxation from metastability in spatial ecological models. *J. Theor. Biol.* **200**: 121–146.
- GONG, P., M. J. EPTON, G. FU, S. SCAIFE, A. HISCOX *et al.*, 2005 A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly. *Nat. Biotechnol.* **23**: 453–456.
- GOULD, F., and P. SCHLIEKELMAN, 2004 Population genetics of autocidal control and strain replacement. *Annu. Rev. Entomol.* **49**: 193–217.
- HAHN, M. V., and S. V. NUZHIDIN, 2004 The fixation of malaria refractoriness in mosquitoes. *Curr. Biol.* **14**: R264–R265.
- HARTL, D. L., and A. G. CLARK, 1989 *Principles of Population Genetics*. Sinauer Associates, Sunderland, MA.
- HEINRICH, J., and M. J. SCOTT, 2000 A repressible female-specific lethal genetic system for making transgenic insect strains suitable for a sterile-release program. *Proc. Natl. Acad. Sci. USA* **97**: 8229–8232.
- HORN, C., and E. WIMMER, 2003 A transgene-based, embryo-specific lethality system for insect pest management. *Nat. Biotechnol.* **21**: 64–70.
- IRVIN, N., M. S. HODDLE, D. A. O'BROCHTA, B. CAREY and P. W. ATKINSON, 2004 Assessing fitness costs for transgenic *Aedes aegypti* expressing the gfp marker and transposase genes. *Proc. Natl. Acad. Sci. USA* **101**: 891–896.
- ITO, J., A. GHOSH, L. A. MOREIRA, E. A. WIMMER and M. JACOBS-LORENA, 2002 Transgenic anopheline mosquitoes impaired in transmission of malaria parasite. *Nature* **417**: 452–455.
- KISZEWSKI, A. E., and A. SPIELMAN, 1998 Spatially explicit model of transposon-based genetic drive mechanisms for displacing fluctuating populations of anopheline vector mosquitoes. *J. Med. Entomol.* **35**: 584–590.
- KNIPLING, E. F., H. LAVEN, G. B. CRAIG, R. PAL, J. B. KITZMILLER *et al.*, 1968 Genetic control of insects of public health importance. *Bull. WHO* **38**: 421–438.
- MOREIRA, L. A., J. WANG, F. H. COLLINS and M. JACOBS-LORENA, 2004 Fitness of Anopheline mosquitoes expressing transgenes that inhibit plasmodium development. *Genetics* **166**: 1337–1341.
- PIALEK, J., and N. H. BARTON, 1997 The spread of an advantageous allele across a barrier: the effects of random drift and selection against heterozygotes. *Genetics* **145**: 493–504.
- RIBEIRO, J. M. C., and M. G. KIDWELL, 1994 Transposable elements as population drive mechanisms: specification of critical parameter values. *J. Med. Entomol.* **31**: 10–16.
- ROBINSON, A. S., 1976 Progress in use of chromosomal translocations in control of insect pests. *Biol. Rev. Camb. Philos. Soc.* **51**: 1–24.
- SCHLIEKELMAN, P., and F. GOULD, 2000 Pest control by the introduction of a conditional lethal trait on multiple loci: potential, limitations and optimal strategies. *J. Econ. Entomol.* **93**: 1543–1565.
- SCOTT, T. W., A. C. MORRISON, L. H. LORENZ, G. G. CLARK, D. STRICKMAN *et al.*, 2000 Longitudinal studies of *Aedes aegypti* (Diptera: Culicidae) in Thailand and Puerto Rico: population dynamics. *J. Med. Entomol.* **37**: 77–88.
- SOBOLEVA, T. K., P. R. SHORTEN, A. B. PLEASANTS and A. L. RAE, 2003 Qualitative theory of the spread of a new gene into a resident population. *Ecol. Modell.* **163**: 33–44.
- THOMAS, D. T., C. A. DONNELLY, R. J. WOOD and L. S. ALPHEY, 2000 Insect population control using a dominant, repressible, lethal genetic system. *Science* **287**: 2474–2476.
- YAPABANDARA, A. M. G. M., and C. F. CURTIS, 2004 Control of vectors and incidence of malaria in an irrigated settlement scheme in Sri Lanka by using the insect growth regulator pyriproxifen. *J. Am. Mosq. Control Assoc.* **20**: 395–400.

Communicating editor: M. W. FELDMAN