Gene-drive into insect populations with age and spatial structure: a theoretical assessment

Yunxin Huang,1 Alun L. Lloyd,2 Mathieu Legros1 and Fred Gould1

1 Department of Entomology, North Carolina State University, Raleigh, NC, USA
2 Department of Mathematics, North Carolina State University, Raleigh, NC, USA

Introduction

Insect-vectored diseases, such as malaria and dengue fever, continue to pose a substantial impediment to global health. Current approaches for combating insect-vectored diseases are often only partially effective and novel approaches targeting the pathogens and the vectors are being sought (Gates Foundation 2009). One novel approach is to replace existing populations of disease-vectoring insects with genotypes that do not transmit the target pathogen. Early attempts at this approach failed (see Gould and Schliekelman 2004), but recent breakthroughs in molecular genetics have resulted in some optimistic predictions for future success (e.g., James 2005).

Population replacement strategies for controlling disease-vectoring mosquitoes require a genetic system that causes desirable antipathogen transgenes to increase in frequency in the target population even when the transgenes have associated fitness costs. A number of genetic drive systems, such as sex-linked meiotic drive, homing endonuclease genes, Medea, and engineered underdominance (Wade and Beeman 1994; Sinkins and Gould 2006; Huang et al. 2007a,b), have been proposed for achieving this goal. Simple population genetic models predict that for a specific drive system, the critical factors determining if a released transgene will increase in frequency are the initial release frequency and the fitness cost associated with the engineered genes, and migration probability of adults, are examined. Results show that patchy release generally requires the release of fewer engineered insects to achieve success than central release. When the fitness cost is very low, central release covering 25% of the total area can be more effective than a completely uniform release over the whole area. This study demonstrates that to determine the best method of spatial release, and the total number of engineered insects that must be released, it is important to take into account the age and sex of the released insects and spatial structure of the population.

Keywords
age and spatial structure, engineered underdominance, gene drive, release threshold, theoretical assessment.

Abstract

The potential benefits and risks of genetically engineered gene-drive systems for replacing wild pest strains with more benign strains must be assessed prior to any field releases. We develop a computer simulation model to assess the feasibility of using engineered underdominance constructs to drive transgenes into age- and spatially structured mosquito populations. Our practical criterion for success is the achievement of a transgene frequency of at least 0.80 within 3 years of release. The impacts of a number of parameters that may affect the success of gene-drive, such as the release area, release age, density-dependent larval survival, fitness cost of the engineered genes, and migration probability of adults, are examined. Results show that patchy release generally requires the release of fewer engineered insects to achieve success than central release. When the fitness cost is very low, central release covering 25% of the total area can be more effective than a completely uniform release over the whole area. This study demonstrates that to determine the best method of spatial release, and the total number of engineered insects that must be released, it is important to take into account the age and sex of the released insects and spatial structure of the population.
examined the impact of age structure on the efficiency of two specific genetic drive systems, Medea and engineered underdominance (Huang et al. 2009). It was found that variation in the life history and mating characteristics of a population could alter the reproductive values of released males and females, and thereby affect the effort needed to successfully implement a gene-drive project.

In addition to age structure, the spatial structure of a population can impact the spread of a transgene (Barton 1979; Schofield 2002; Rasgon 2003; Soboleva et al. 2003; Yakob et al. 2008a,b; Yakob and Bonsall 2009). The number and initial spatial distribution of the released individuals can affect the ability of the transgene to spread through the population and the speed of the resulting spread depends on how often and how far individuals disperse (Barton 1979; Schofield 2002; Soboleva et al. 2003). Consequently, the design of effective and efficient transgenic mosquito releases must consider spatial structure.

For any specific genetic management program involving gene-drive, one practical question that must be addressed is what spatial release pattern of an engineered insect strain would be most efficient for spreading antipathogen transgenes? The answer to this question is expected to depend on a number of factors such as migration rate, fitness cost of the transgenes, density-dependent mechanisms for population regulation, age, and sex of the released insects, and thus can only be addressed by a model with explicit spatial structure and ecological processes included.

Here, we develop an age- and space-structured deterministic model to analyze gene-drive into the population of a mosquito species with demographic and dispersal characteristics similar to Aedes aegypti, the primary vector of dengue. The gene-drive mechanism chosen for this analysis is engineered underdominance (Davis et al. 2001). We consider a time scale (3 years) relevant to the practical implementation of a transgenic control program. Our primary goal is to identify spatial release strategies that require the least engineered insects, with a particular focus on the potential for interactions among genetic and ecological factors.

The gene-drive mechanism

Engineered underdominance (EU), originally proposed by Davis et al. (2001), involves the introduction of individuals carrying two co-dependent engineered constructs $a$ and $b$ on separate chromosomes. Each construct consists of a desirable antipathogen gene, a lethal gene, and a suppressor of the lethal gene on the other construct. The lethal gene on one construct is expressed when the suppressor on the other construct is absent. Therefore, individuals with only $a$ or $b$ will die. We denote a diploid homozygous individual with the two constructs inserted into two nonhomologous autosomes by $aabb$, and the corresponding wild-type by $AABB$. When the engineered $aabb$ individuals are introduced into a wild population and mate with wild-type individuals, only five of the nine potential $F_2$ genotypes are viable: $AABB$, $AaBb$, $Aabb$, $aaBb$, and $aabb$. The other four genotypes ($AAbb$, $AaBB$, $aaBB$, and $AAbb$) are not viable because they contain only construct $a$ or $b$, but not both.

The EU strategy requires that the frequency of engineered insects exceeds a nonzero introduction threshold in order for the constructs to go to fixation even if there are no fitness costs (Davis et al. 2001). In reality, fitness costs are expected to be associated with the EU constructs. Such fitness costs may affect one or more aspects of the insect’s life cycle, for instance causing reduced fecundity, embryonic survival, larval development, and/or survival of adults (Irvin et al. 2004). The precise nature of fitness costs is likely to be strain-specific. In the absence of this information we make a simple assumption, namely that the only fitness cost is a reduction of the embryonic survival probability. Similarly, the dominance of fitness costs is currently unknown. As in an earlier study of a nonspatial age-structured model (Huang et al. 2009), we assume that fitnesses are calculated multiplicatively across the two loci, but that, at each locus, fitness is calculated additively. (Making this assumption allows us to make direct comparisons between corresponding spatial and nonspatial age-structured models.) We denote embryonic fitnesses by $f(G)$, where $G$ denotes the genotype. The embryonic fitnesses of the five viable genotypes, $AABB$, $AaBb$, $Aabb$, $aaBb$, and $aabb$, are $1$, $(1 - c/2)^2$, $(1 - c/2)(1 - c)$, $(1 - c)(1 - c/2)$, and $(1 - c)^2$, respectively, while $f(G) = 0$ for the four inviable genotypes. In this study, we consider fitness costs per locus, $c$, that range between 0 and 0.2. The equations for the basic model describing EU can be found in Davis et al. (2001) and Magori and Gould (2006).

The model

The model we present is quite general, but in order to carry out simulations we must choose a specific set of parameter values. Some of the choices made in this study reflect our interest in transgenic strategies aimed at reducing the transmission of dengue virus by the A. aegypti mosquito (Magori et al. 2009), which is one of the major current targets of genetic pest management.

We consider a metapopulation of mosquitoes living in a habitat that consists of 1024 identical cells or patches that are arranged as a $32 \times 32$ square grid. (In the A. aegypti context, because the mosquito is anthropophilic...
we think of this as being an urban habitat consisting of a collection of houses). The patchy habitat is denoted by \( H = \{(i,j) \mid i,j = 1, \ldots, 32\} \) in which each patch is assigned a two-dimensional integer co-ordinate \((i,j)\). Patches are linked by migration and we assume periodic boundary conditions.

The details of the local population dynamics and genetics as well as the migration scheme are described separately below. Major model parameters and their default value or range are listed in Table 1.

Within-patch dynamics

The mosquito life cycle is modeled using 35 daily age classes. The immature mosquito stages span 10 age classes, while the adult stage has 25 age classes. The numbers of immature and adult female mosquitoes of age \( a \) and genotype \( G \) at time \( t \) are written as \( I(G,a,t) \) and \( A(G,a,t) \), respectively. The corresponding numbers of immature and adult male mosquitoes are denoted by \( I^*(G,a,t) \) and \( A^*(G,a,t) \), respectively.

The birth rate, i.e., the average number of eggs laid per female per day, is assumed to be age-dependent and described by the function \( b(a) \). In this study, rather than choosing a form for \( b(a) \) that is estimated from laboratory or field data for a specific mosquito species (see e.g., Harrington et al. 2001 or Rasgon and Scott 2004), we instead employ a pattern that only aims to capture the qualitative features of mosquito fecundity. We choose a simple functional form in which females have a broad reproductive interval and have highest fecundity at age 10 (see Fig. 1A, the caption of which gives the details of the functional form we adopt for \( b(a) \)).

The daily survival probability of adult mosquitoes, \( S_A \), is assumed to be constant (independent of sex, age, and genotype, but dependent on their density (McDonald 1977; Dye 1984; Legros et al. 2009). Density dependence is poorly characterized for many mosquito species, including A. aegypti (Legros et al. 2009). Typically, a phenomenological model of density dependence is used (see Bellows 1981 for a comparison of a number of models for density-dependent insect survival). For many species, intraspecific competition for nutrients between immature mosquitoes in the larval habitat is believed to be an important component of density dependence. Consequently, we assume that the daily survival probability of immature individuals, \( S(D(t)) \), depends on the density of the immature population, \( D(t) \). Specifically, we adopt a modified form of a two-parameter function, due to Maynard Smith and Slatkin (1973), discussed by Bellows (1981) (see also Yakob and Bonsall 2009):

\[
S(D(t)) = \frac{S_I}{1 + a|D(t)|^b}.
\]

Here, \( a \) and \( b \) are two parameters that characterize the magnitude and shape of density dependence and are discussed below. The density \( D(t) \) is given by \( \sum_{a=1}^{10} w_a \sum_G (I(G,a,t) + I^*(G,a,t)) \), where \( w_a \) is the average weight of an immature of age \( a \) (the values of these coefficients are given in the caption to Fig. 1).

Under this survival function, the maximum daily survival probability for immature mosquitoes, \( S_I \), is assumed to be 0.9. In this study, we set \( a = 0.035 \) and \( b = 1 \) or \( b = 4 \), except where specifically indicated otherwise. The two values of \( b \) are chosen to represent weak \((b = 1)\) and strong density dependence \((b = 4)\). The value of \( a \) is chosen so that the number of adults in a stable isolated patch is approximately 20 when \( b \) is taken to equal one. Figure 1(B) illustrates the daily survival probability as a function of total immature weight for these two values of \( b \).

The total number of viable eggs of genotype \( G \) produced by adults at time \( t \), \( B(G,t) \), is calculated by finding the number of eggs arising from each parental genotype pair, allocating the eggs according to the offspring genotype distribution that results from each parental genotype pair, and then accounting for the viability and embryonic fitness costs of each type of egg. \( B(G,t) \) is calculated as follows:

\[
B(G,t) = f_c(G) \times \sum_{j,K} \Pr(G \mid j,K) \left( \sum_{a=1}^{25} b(a)A(J,a,t)\gamma(a,J,K,t) \right).
\]

The parenthetic term in this equation calculates the total number of eggs produced by matings between females of genotype \( J \) and males of genotype \( K \). The probability that, at time \( t \), a female of age \( a \) and genotype \( J \) mates with a male of genotype \( K \) is denoted \( \gamma(a,J,K,t) \). In this study, we assume that mating is random and occurs

| Table 1. Model constants and parameters. |
|-----------------------------------------|
| **Meaning**                             | **Notation** | **Value/range** |
| Life-time egg production of a female     | \( E \)      | 50             |
| Maximum immature survival probability    | \( S_I \)    | 0.9            |
| Adult survival probability               | \( S_A \)    | 0.9            |
| Magnitude parameter of density-dependence| \( \alpha \) | 0.035          |
| Shape parameter of density-dependence   | \( \beta \)  | 1, 4           |
| Daily migration probability              | \( m \)      | [0.05, 0.5]    |
| Maximum migration distance               | \( d_m \)    | 5              |
| Fitness cost per locus                   | \( c \)      | [0, 0.2]       |

© 2010 Blackwell Publishing Ltd 4 (2011) 415–428
we assume that every day, so \( \gamma(a, J, K, t) \) is simply the fraction of adult males of genotype \( K \) at that time, i.e.,

\[
\gamma(a, J, K, t) = \frac{\sum_{a'} A^*(K, a', t)}{\sum_{a'} \sum_{a''} A^*(K, a'', t)}, \tag{3}
\]

where the sums are taken over the 25 adult age classes. The formula for \( \gamma(a, J, K, t) \) can be easily extended to consider different mating patterns, such as assortative mating by age (as given in Huang et al. 2009) or genotype, or mating that only occurs at certain times in an adult female’s life, such as a single mating event soon after emergence.

\( \Pr(G \mid J, K) \) is the probability that a zygote has genotype \( G \) when the maternal and paternal genotypes are \( J \) and \( K \), respectively. In the case of Mendelian segregation, this probability can be easily calculated by a computer algorithm (available from the authors).

Finally, \( f_t(G) \) measures the relative fitness of a type \( G \)-egg, expressed as embryonic survival relative to that of the wildtype. (As discussed earlier, other types of fitness cost are possible; these are easily accommodated by making simple modifications to this equation.)

Incorporating the assumptions for daily survival and reproduction as given in eqns (1)–(3), and assuming a 1:1 sex ratio at birth, we have the following deterministic recursive equations for the numbers of females of given genotypes and ages on successive days:

\[
L(G, 1, t + 1) = 0.5B(G, t), \tag{4}
\]

\[
L(G, a + 1, t + 1) = S(D(t))L(G, a, t), \quad a = 1, \ldots, 9, \tag{5}
\]

\[
A(G, 1, t + 1) = S(D(t))L(G, 10, t), \tag{6}
\]

\[
A(G, a + 1, t + 1) = S_A(A(G, a, t)), \quad a = 1, \ldots, 24. \tag{7}
\]

The dynamics of the male population can be tracked by the corresponding set of equations, with \( L(G, a, t) \) and \( A(G, a, t) \) being replaced by \( L^*(G, a, t) \) and \( A^*(G, a, t) \), respectively.

**Migration**

Immature mosquitoes are assumed to be sedentary, while adult mosquitoes are allowed to move between patches. Many factors such as age, sex, season, location of feeding, and oviposition sites can affect movement of mosquitoes. However, for \( A. aegypti \), a review of 21 mark-recapture experiments (Harrington et al. 2005) indicated that movement was typically limited to neighboring houses and that age, sex, season, and geographical location did not significantly affect the probability of movement or mean distance moved. One more recent study did find a difference in the mean distance moved between males and females,
but only when the larvae were raised under stressful conditions (Maciel-De-Freitas et al. 2007). We therefore assume that the daily migration probability, denoted by $m$, is constant.

Because of the relatively short-range dispersal of A. aegypti, we assume the probability, $\phi_d$, that a migrating mosquito moves to a patch $d$ units away from its current location decreases geometrically with $d$, as given by:

$$\phi_d = \frac{s^{d-1}}{1 + s + \ldots + s^{d_m-1}}, \quad d = 1, \ldots, d_m$$

Here, $d_m$ is the maximum distance that the mosquitoes can move in a day, and $0 < s < 1$ is a parameter for adjusting the mean daily dispersal distance ($d = [1 - s^{d_m} - d_m s^d (1 - s)] / [(1 - s) (1 - s^{d_m})]$). We calculated distances by adding the number of horizontal and vertical steps needed to move between two patches on our grid using the shortest path (the so-called taxicab, or Manhattan, distance, with appropriate modification to account for the periodic boundary conditions).

For simplicity, when comparing across differing levels of dispersal in this study, we choose to do so by varying the value of $m$; more generally, we could also vary $s$ and/or $d_m$. In this study, we choose to fix $s = 0.5$ and $d_m = 5$, giving a mean daily dispersal distance of approximately 1.84. A geometric dispersal kernel is not the only possible choice for $\phi_d$, a more commonly-used alternative is a Gaussian kernel. Many mosquito species have a looser association with people than do A. aegypti and have much longer dispersal distances (Service 1993): few of their movements are between adjacent houses. For such species, a Gaussian kernel might be more appropriate. We make some comments on the likely impact of different dispersal kernels in the Discussion.

Let $\hat{A}_{ij}(t)$ and $\hat{A}_{ij}^*(t)$ be the total number of female and male adults in patch $(i,j)$, respectively. The net daily change in the number of female adults in this patch due to dispersal is the number of immigrants minus the number of emigrants:

$$\Delta \hat{A}_{ij}(t) = -m\hat{A}_{ij}(t) + \sum_{d=1}^{d_m} \frac{1}{4d} \phi_d \sum_{(k,l) \in \Omega_d(i,j)} \hat{A}_{kl}(t). \quad (9)$$

Here, we denote the neighborhood of patches that are exactly $d$ units away from a specific patch $(i,j)$ as $\Omega_d(i,j)$. There are $4d$ patches that are exactly $d$ units away from the patch $(i,j)$ (note that the periodic boundary conditions ensure that this is true for all patches). The equation for males can be written similarly. Note that the daily migration probability, $m$, is assumed to be independent of sex, age, and genotype.

**Release methods**

In the previous study (Huang et al. 2009), we considered release of insects of different ages. Here, we mostly consider a one-time release of engineered teneral adults into selected patches. We examine releases of either both sexes or only males. For mosquitoes, only females transmit diseases, therefore safety concerns favor the release of only males (Klassen and Curtis 2005). The released individuals are of genotype $aabb$, i.e., homozygous for both EU constructs. We consider two specific release methods: (i) central release where the release area is a square of varied size (i.e., number of patches) at the center of the habitat (Fig. 2), (ii) patchy release where the habitat is divided into $8 \times 8$ equal-sized square blocks (with each block consisting of 16 patches) and the release area is a square

![Figure 2](image-url)
of varied size in each block (Fig. 2). We have chosen the 8 × 8 division as just one way to consider patchy release. Other scales of division such as 16 × 16 are also possible. Both central release and patchy release have a limiting case, which we refer to as full release, where there are (identical) releases into every patch. Given our assumptions of homogeneous environment and initial conditions, and deterministic dynamics, each patch will exhibit identical dynamics and so the impact of full release could be assessed using a nonspatial model.

Release threshold

As described in a previous section, engineered underdominance requires the release of a minimum number of transgenic individuals per wild-type individual to achieve gene-drive, i.e., after a transient period the transgenes increase in frequency and eventually reach a stable high frequency or fixation. From a practical perspective, gene-drive also has a requirement for the speed at which the transgenes increase in frequency across the entire habitat. In this study, we focus on a practical criterion for success of a gene-drive project as one in which the frequency of transgenic individuals (summed across all transgenic genotypes) reaches 80% in the entire habitat within 3 years (i.e., the wild-type frequency falls below 20% within 1095 days). We refer to the minimum number of mosquitoes that must be released per wild-type mosquito to achieve this criterion as the release threshold throughout the remainder of the manuscript. We emphasize that the release threshold is the ratio of the number of released mosquitoes to the number of wild-type mosquitoes of all ages (both adults and immatures) in the entire spatial area, even when the engineered mosquitoes are released into only a fraction of the total area.

Numerical simulations

Because the model takes genetic, age, and spatial structure into account, simulation of the model is not particularly fast; calculation of the release threshold is fairly slow even for a single set of parameters. This problem is more severe when examining variation in the release threshold as a parameter value is changed. As a result, we assess release thresholds using coarser increments in parameter values than is optimally desirable. When examining variations that result from changes in migration probability, m, we vary m between 0.05 and 0.5 in increments of 0.05. This restricts our ability to precisely locate features on graphs of functions of m (for example, the values of m that minimize release thresholds in Fig. 3A): this point should be kept in mind when interpreting some of the statements below. When examining variations that result from changes in the fitness cost, c, we vary the parameter from 0 to 0.2 in increments of 0.01.

Releases are always made into populations that are at equilibrium: the populations of each of the 1024 patches are equal in number and are each at the stable age distribution. (This initial state was determined by running the single-patch model for 600 time steps).

Results

Effects of release area on release threshold

Central release

In the case of central release of both males and females when there is no fitness cost associated with the EU constructs (c = 0), there is an optimal migration probability at which the release threshold is lowest for releases into 1/16 and 1/4 area of the total area (Fig. 3A). When the migration probability is smaller than the optimal value, the outward expansion of the transgenes is slow, which makes it relatively difficult for the transgenes to reach a frequency of 80% across all patches in 3 years (as required by our practical criterion). When the migration probability is greater than the optimal value, the transgenes spread quickly into many patches, which causes the ratio of released individuals to wild individuals in the release area to drop quickly after the release and makes it relatively difficult for transgenes to increase in frequency in the release area.

The release of only males is similar to the release of both males and females in terms of the qualitative impact of migration probability and spatial release area on the release thresholds (Fig. 3B). The quantitative differences between the only-male release and both-sex release is that the former generally has higher release thresholds than the latter. As explained in Huang et al. (2009), this is because an only-male release leads to an excess of males in the population with the consequence that some of the introduced males will be unable to find mates. Compared to a both-sex release that leaves the population’s sex ratio unaffected, fewer engineered alleles get passed to the next generation under an only-male release. We remark that this effect can be overturned if transgenic individuals incur fitness costs (see Huang et al. (2009) for more details).

When there is no fitness cost, for each specific migration probability, there is an optimal proportion of the total area for which the release threshold is lowest (Fig. 4A). The optimal proportion exists because release into a relatively small area favors local establishment but it takes relatively long for the transgenic alleles to spread throughout the entire habitat, while release into a relatively large area favors a fast spread of the transgenic alleles once they are established but makes it hard for local establishment in the first place. If both the release
area and migration probability are small, the transgenic individuals may spread over only a part of the entire area in 3 years and may never reach a frequency of 80% even if a large number of insects are released. If the release area is large, the transgenic individuals can easily spread over the entire area once the number of released insects allows for local establishment and so, the release threshold is almost independent of the migration probability.

When there is a fitness cost to the transgenes there is a general expectation that an increased number of transgenic insects will need to be released, but our results indicate that the level of this increase is greatly affected by the release area and migration probability. When there is a fitness cost of 0.2 per locus associated with the EU constructs, a central release 1/16 of the total area results in a very high release threshold (more than 100 times that of the total wild population). A central release 1/4 of the total area also results in a very high release threshold for small migration probabilities (Fig. 3C,D). For a fixed migration probability, the release threshold increases as the size of release area decreases (Fig. 4B). When the transgenic alleles carry fitness cost, central release into a smaller area results in more matings between the transgenic types in the first few generations than release into a larger area. Because more matings between the transgenic types cause more loss of transgenic alleles, more insects must be released into a smaller area than a larger area to achieve the goal of gene drive.

Patchy release
In the case of patchy release, when the release area is either 1/16 or 1/4 of the total area, the release thresholds decrease as the migration probability increases (Fig. 3; gray lines). When the migration probability is small, there are some differences in the release thresholds between 1/16-, 1/4-, and full release (Fig. 3A,B). For example,
when \( m = 0.05 \) and both sexes are released, the difference between 1/16-release and full release is approximately 0.2 (0.148 vs 0.35). When the migration probability is large (e.g., \( m > 0.3 \)) the differences in the release threshold for the three proportions of release area are very small. When there is a fitness cost of 0.2, the release thresholds are higher than when there is no fitness cost, but there is less of an interaction between the area of release and migration probability than was seen for the case of central release (Fig. 3C,D). No matter whether there is fitness cost or not, the differences in the release thresholds between the three proportions of release area become smaller as the migration probability increases.

Comparison between central release and patchy release
For small migration probabilities (\( m < 0.1 \)), when there is no fitness cost associated with the EU constructs, a patchy release into 1/4 of the total area results in lower release threshold than a central release into 1/4 of the total area. However, for high migration probabilities, the central release has the lowest release thresholds (Fig. 3A,B). This is at least partially because at low migration probabilities the patchy releases cause a faster spread of the transgenic alleles into the nonrelease patches, but when the migration probabilities are high the invasion into the nonrelease patches is not a major limitation. When the fitness cost is 0.2 the patchy release results in much lower release thresholds than the central release. This is at least in part because in early generations after releases, more matings between transgenic males and females happen when they are released into a single contiguous area rather than into a number of discrete locations (Fig. 3C,D). The mating between transgenic individuals maintains stronger linkage disequilibrium among transgenic alleles and therefore a higher effective fitness cost.

Effects of fitness cost and migration on release threshold
In the case of releasing into 1/4 of the total area, for a fixed migration probability there is a critical fitness cost above which patchy release results in a lower release threshold than central release. For example, when the migration probability is 0.2, the critical fitness cost is between 0.03 and 0.04 for both-sex release (Fig. 5A), and between 0.01 and 0.02 for only-male release (Fig. 5B). Numerical simulation at different migration probability and fitness cost pairs, \((m,c)\), shows that there is a critical fitness cost above which patchy release results in a lower release threshold than central release for any migration probability. This critical fitness cost is smaller than 0.04 for both-sex release (Fig. 5C) and smaller than 0.02 for only-male release (Fig. 5D).

Effects of age, density dependence and migration on release threshold
The parameters \( x \) and \( \beta \) measure the strength of density-dependent effects on the survival of immature insects.
Until this point in our analysis, we set \( \alpha = 0.035 \) and \( \beta = 1 \), parameters for which density dependence is weak. We now relax these assumptions to analyze the impact of density dependence on the release efforts required for achieving gene-drive. As a simple approach to this analysis, we contrast the release thresholds for two different values of \( \beta \), 1 and 4, which represent a relatively weak and strong density dependence, respectively. To look into the possibility of an interaction between age structure and spatial structure, we consider and compare the releases of three specific age classes: newly emerged adults, 5-day-old adults and 10-day-old adults. We focus our analysis on the case of central release into 1/4 of the total area.

For any of the three age classes and any positive migration probability, the release thresholds for \( \beta = 1 \) are higher than those for \( \beta = 4 \) (Fig. 6). These results appear to be qualitatively similar to those for a single isolated patch: the higher the \( \beta \), the lower the release threshold (Fig. 7). There are also qualitatively similar results for \( \alpha \), the other parameter that adjusts the magnitude of density-dependence (Fig. 7). While there are differences in the release threshold between \( \beta = 1 \) and \( \beta = 4 \), the differences appear to be not significant except when one day old adults are released and fitness cost is high. The differences are partially due to stronger density-dependence increasing the reproductive values of the released adults. In all of our analysis we define the release threshold in terms of the number of released mosquitoes relative to all mosquitoes in the natural population. This means that, we value each egg and young adult equally. As described in Huang et al. (2009), this can be misleading especially if the reproductive values of eggs and young adults vary as we change model parameters. When there is high density-dependent mortality at the larval stage, as for \( \beta = 4 \), the reproductive value of young adults is higher than when there is low larval mortality, as for \( \beta = 1 \). The higher the reproductive value of the released insects, the lower the

---

Figure 5 Comparison of central release with patchy release (both into 1/4 of the total area) when the fitness cost of the drive gene and migration probability are varied. Graphs (A) and (B) show the release threshold as a function of the fitness cost when the migration probability is 0.2. The solid lines correspond to the central release (labeled as RT_C), while the dashed lines correspond to the patchy release (labeled as RT_P). Graphs (C) and (D) are parameter space diagrams, with the unshaded areas showing regions of the \((m, c)\) parameter space in which the release threshold is higher under central release than under patchy release. The shaded areas depict regions in which the release threshold is higher under patchy release than under central release. For each point on a grid of \((m, c)\) values, the more effective release mechanism was determined. The smooth curves depicting the boundaries of the shaded regions were then determined by polynomial interpolation.
release threshold. As the result, release of fewer transgenic adults is required to achieve the same goal of gene-drive.

When there is no fitness cost and when both sexes are released, the release of 5-day-old adults results in the lowest release threshold among the three age classes because this age class has the highest reproductive values (Fig. 6A). This is different from the case of release of only males, where release of newly emerged adults results in the lowest release threshold (Fig. 6B). The difference is partially a consequence of the differing reproductive values of males and females of the same age when mating is random (Huang et al. 2009). When there is a fitness cost of 0.2, release of newly emerged adults results in the lowest release threshold among the three age classes for either both-sex release or male-only release (Fig. 6C,D).

Discussion

Over 40 years ago Richard Levins (Levins 1966) pointed out that model building is constrained by ‘the conflicting desiderata of generality, realism, and precision’. He explained that no single model can attain all of these goals, but that the alignment of predictions from sets of diverse models will provide robust theorems in population biology. His statements are as relevant today as they were when they were written. One can argue the pros and cons of very simple general models, and equally, the pros and cons of more realistic complex models, but it is the overall knowledge gained by comparing the results of these diverse models that provides a measure of confidence in their predictions (also see Gerber et al. 2003).

In modeling gene drive systems there is a wide gap between simple models and species specific models that can make comparative assessment difficult. For example, simple population genetics models cannot act as a qualitative check on some predictions of a complex model built for one mosquito species in one city if the simple models do not include males and females, or differential reproduction by old and young individuals. Models of intermediate complexity such as the one described here serve as a bridge between models aimed at extreme generality and those aimed at realism.

Figure 6 Release threshold for central release into 1/4 of the total area as a function of migration probability for two \( \beta \) values and three different release ages. (A) Release of both sexes, \( c = 0 \); (B) Release of only males, \( c = 0 \); (C) Release of both sexes, \( c = 0.2 \); (D) Release of only males, \( c = 0.2 \). In each graph, the solid lines correspond to \( \beta = 1 \) (weak density-dependent competition between immatures), while the dashed lines correspond to \( \beta = 4 \) (strong density-dependent competition). Lines plotted as black, blue and red correspond to the release of 1-, 5-, and 10-day-old adults, respectively. In graph (D), the release thresholds corresponding to the release of 10-day-old adults (red lines) are not shown in the figure because they are higher than 100.
As indicated in the Introduction section, the first models of engineered underdominance (Davis et al. 2001; Magori and Gould 2006) predicted efficacy of a release strategy simply based on the minimum allelic frequencies of transgenes that resulted in future increases in frequency of those transgenes in a panmictic population. Some population genetic models have analyzed the general spatial spread of underdominant alleles (e.g., Barton 1979; Schofield 2002; Soboleva et al. 2003; Turelli 2010), but they include very simple or even no population dynamics. We felt that it would be difficult to use such models alone as a check on a complex, species, and location specific set of models that we have been building for assessing the utility of gene drive systems for the suppression of dengue virus that is vectored by the mosquito, *A. aegypti* (Magori et al. 2009). This motivated us to develop the age-structured model (Huang et al. 2009), and now this model that includes explicit spatial structure. While these two models are not realistic enough to optimize the release strategy for a particular transgenic strain, they make qualitative predictions about mechanistic factors that could affect predictions of more complex and realistic models.

If we are to use models of intermediate complexity as a bridge, we must also connect and relate them to the more simple models mentioned above (Barton 1979; Turelli and Hoffmann 1991; Schofield 2002; Soboleva et al. 2003; Turelli 2010). The behavior of these population genetic models is seen to have much in common with population dynamic models for the invasion of a species that is subject to a (strong) Allee effect (e.g., Lewis and Kareiva 1993; Kot et al. 1996; Wang et al. 2002). Analytic results can be derived that provide threshold conditions determining whether invasion is or is not possible, and expressions for the speed of the invasion wave that results.

A simple setting in which such analyses can be carried out assumes a population that is distributed continuously along a line and for which movement occurs by diffusion (corresponding, in some sense, to a Gaussian dispersal kernel; Baeumer et al. 2007). Furthermore, it is assumed that the (local) genetics can be specified by a single gene frequency that evolves continuously in time, governed by a cubic function. Initially, all individuals at negative values of the spatial co-ordinate are taken to be transgenic and those at positive values wild-type. Under these conditions it is found (Barton 1979; Turelli and Hoffmann 1991; Lewis and Kareiva 1993) that the transgene will invade provided that the local invasion threshold, \( \rho \), is below 0.5 and that the resulting invasion wave has (asymptotic) speed proportional to the product of the mutation or selection coefficient and the density-dependent survival probability of immature individuals. The density-dependent survival probability depends on two specific parameters, \( a \) and \( b \). In panels (A) and (B), \( b = 1 \) is fixed and the release threshold is plotted as a function of the parameter \( a \), with either no fitness cost (solid line) or with fitness cost 0.2 (dashed line). In panels (C) and (D), \( a = 0.035 \) is fixed and the release threshold is plotted as a function of the parameter \( b \), with either no fitness cost (solid line) or with fitness cost 0.2 (dashed line). In panel (A) and (C) both males and females are released. In panels (B) and (D) only males are released. In all instances, 1-day-old adults are released.
(1 − 2p)√D, where D is the diffusion coefficient. Because $D = \sigma^2/2$, where \( \sigma \) is the standard deviation of the corresponding Gaussian dispersal kernel, the wave speed is seen to be proportional to \( \sigma \).

An important observation is that the forward advance of invasion waves in such systems relies on the weight of numbers behind the wave front: invasion requires a sufficiently high frequency over a sufficiently large area, sometimes termed a critical aggregation (or ‘critical bubble’). These so-called ‘pushed’ (or Bartonian) invasion waves are fundamentally different to the more familiar ‘pulled’ (or Fisherian) invasion waves (Fisher 1937) seen in systems that lack an invasion threshold, for which advance is governed by the wave’s leading edge and for which invasion is possible from arbitrarily small localized initial introductions. The gene frequency distribution corresponding to this critical bubble can be calculated (Soboleva et al. 2003). Typically, these calculations have been presented for release over a single contiguous area and assuming a simple geometry, such as radial symmetry. As such, they talk to our central release but not to the patchy release. There does not appear to be an analytic description corresponding to our patchy release, although Soboleva et al. describe (using numerical simulation) a situation in which two nearby releases, neither of which would lead to invasion by themselves, can merge and form a structure that can lead to invasion (Soboleva et al. 2003).

Calculation of the changes in gene frequencies of the EU system requires knowledge of the frequencies of the five viable genotypes. Although this system is more complex than the one-dimensional idealization of the Bartonian model (even ignoring density dependence), numerical results suggest that it may be possible to approximate the dynamics of EU by a one-dimensional model. The potential for such a description is indicated by fig. 6 of Davis et al. (2001) and fig. 2 of Magori and Gould (2006), in which the trajectories of gamete frequency on DeFinetti diagrams are seen to rapidly approach and then move along a one-dimensional curve—the unstable manifold of a saddle equilibrium—on which two stable equilibria fall, with an unstable equilibrium in between. These dynamics suggest that the EU system, following an initial transient lasting a few generations, may, to a good approximation, be described by in terms of a co-ordinate describing position on this curve.

The exact magnitude of the local invasion threshold that prevents establishment of a Bartonian wave depends on the details of the local gene frequency dynamics. While this value need not be at the \( \hat{p} = 0.5 \) of the cubic model described above, this result explains why central release for EU was much less effective in the face of the \( c = 0.2 \) fitness cost and the resulting high local release threshold. The theoretical results just described also show that even if the Bartonian wave can be established, its expansion occurs more slowly when the local release threshold is high. Consequently, it may fail to meet our practical ‘80% in 3 years’ criterion. In such situations, patchy release, with its greater initial spatial extent, may be better placed to achieve the desired result. In addition, when transgenes carry fitness costs, the offspring produced by the homozygous transgenic males and females suffer higher mortality than those produced by other types of mating. Because more matings between transgenic males and females happen when they are released into a single contiguous area rather than into a number of discrete locations, central releases must be larger to compensate for high fitness costs.

Analytic and numerical results (Schofield 2002; Wang et al. 2002) show that the shape of the dispersal kernel has a major effect on the speed of an invading wave, with larger wave speeds resulting from dispersal kernels with heavier tails, but has relatively little effect on the ability of a given introduction to invade. Schofield also describes how similar invasion wave speeds can be obtained from dispersal kernels of different shapes by changing parameters (e.g., increasing dispersal distance in a Gaussian kernel when comparing with a leptokurtic one). As a result, we argue that the shape of the dispersal kernel is likely to have a quantitative, rather than qualitative, impact on the results presented here.

Because female, but not male, mosquitoes transmit human pathogens, male-only releases are likely to be more acceptable to local residents and health officials (Klassen and Curtis 2005), even though the released mosquitoes are engineered to be refractory to disease transmission. Our model indicates that release of males and females can result in a considerably lower release threshold than release of only males. However, as the strength of density-dependence increases (e.g., when \( \beta \) increases from 1 to 4), the differences between male-only and both-sex release become smaller. For example, when there is no fitness cost associated with the transgene, the migration probability is 0.3 and 1-day-old adults are released, male-only release requires approximately 110% more engineered insects than both-sex release when \( \beta = 1 \), while it requires 80% more engineered insects than both-sex release when \( \beta = 4 \). We know very little about the patterns of intraspecific competition in mosquitoes (Legros et al. 2009), but these results indicate that ecological parameters associated with this process could impact the relative efficiency of different gene drive strategies.

Any specific population replacement program will need to consider the economics of a release, both in terms of the cost of rearing the required number of insects and the cost of implementing the release (Atkinson et al. 2007). It is possible that the cost per insect released into one
central area could be much smaller than releasing them into a large number of patches across the habitat because of the transportation and labour cost of visiting many sites. If the relative cost per insect for patchy release is high, then under many ecologically relevant parameter values, the central release could be economically favorable even though it requires more insects. Fitness costs could, however, affect these considerations in an important way because, as we saw above, the comparative utility of a central release diminished considerably in the face of significant fitness costs. Clearly, it will be necessary to link ecological and population genetic models to economic models in order to design optimized release strategies.

Acknowledgements

We thank the associate editor and three anonymous reviewers for their comments. This work is funded by a NIH grant R01-A154954-01A2 and a grant to the Regents of the University of California from the Foundation for the National Institutes of Health through the Grand Challenges in Global Health initiative.

Literature cited

Atkinson, M. P., Z. Su, N. Alphay, L. S. Alphay, P. G. Coleman, and L. M. Wein. 2007. Analyzing the control of mosquito-borne diseases by a dominant lethal genetic system. Proceedings of the National Academy of Sciences of the United States of America 104:9540–9545.

Baumer, B., M. Kovics, and M. M. Meerschaert. 2007. Fractional reproduction-dispersal equations and heavy tail dispersal kernels. Bulletin of Mathematical Biology 69:2281–2297.

Barton, N. H. 1979. The dynamics of hybrid zones. Heredity 43:341–359.

Bellows, T. S. Jr. 1981. The descriptive properties of some models for density dependence. The Journal of Animal Ecology 50:139–156.

Davis, S., N. Bax, and P. Grewe. 2001. Engineered underdominance for manipulation of pest populations. Proceedings of the National Academy of Sciences of the United States of America 72:209–220.

Dye, C. 1984. Models for the population dynamics of the yellow fever mosquito, Aedes aegypti. The Journal of Animal Ecology 53:247–268.

Fisher, R. 1937. The wave of advance of advantageous genes. Annals of Eugenics 7:355–369.

Gates Foundation. 2009. Grand Challenges in Global Health. http://www.grandchallenges.org/Pages/BrowseByDisease.aspx.

Gerber, L. R., L. W. Botsford, A. Hastings, H. P. Possingham, S. D. Gaines, S. R. Palumbi, and S. Andelman. 2003. Population models for marine reserve design: a retrospective and prospective synthesis. Ecological Applications 13:547–564.

Harrington, L. C., T. W. Scott, K. Lerdhusnee, R. C. Coleman, A. Costero, G. G. Clark, J. J. Jones et al. 2005. Dispersal of the dengue vector Aedes aegypti within and between rural communities. The American Journal of Tropical Medicine and Hygiene 72:209–220.

Huang, Y., K. Magori, A. L. Lloyd, and F. Gould. 2007a. Introducing desirable transgenes into insect populations using Y-linked meiotic drive: a theoretical assessment. Evolution 61:717–726.

Huang, Y., K. Magori, A. L. Lloyd, and F. Gould. 2007b. Introducing transgenes into insect populations using combined gene-drive strategies: modeling and analysis. Insect Biochemistry and Molecular Biology 37:1054–1063.

James, A. A. 2005. Gene drive systems in mosquitoes: rules of the road. Trends in Parasitology 21:64–67.

Klassen, W., and C. F. Curtis. 2005. History of sterile insect technique. In V. A. Dyck, J. Hendrichs, and A. S. Robinson, eds. Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management, pp. 3–36. Springer, The Netherlands.

Kot, M., M. A. Lewis, and P. van den Driessche. 1996. Dispersal data and the spread of invading organisms. Ecology 77:2027–2042.

Lewis, M. A. Lewis, and P. Kareiva. 1993. Allee dynamics and the spread of invading organisms. Theoretical Population Biology 43:141–158.

Magori, K., and F. Gould. 2006. Genetically engineered underdominance for manipulation of pest populations: a deterministic model. Genetics 172:2613–2620.

Magori, K., M. Legros, M. Puente, D. A. Focks, T. W. Scott, A. L. Lloyd, and F. Gould. 2009. Skeeter Buster: a stochastic, spatially-explicit modeling tool for studying Aedes aegypti population replacement and population suppression strategies. PLoS Neglected Tropical Diseases 3:e508.

Rasgon, J. L. 2003. Population subdivision can help or hinder Wolbachia introductions into vector populations. Proceedings of the Mosquito Vector Control Association, CA 71:80–86.

Rasgon, J. L., and T. W. Scott. 2004. Impact of population age structure on Wolbachia transgene driver efficacy: ecologically complex factors and release of genetically-modified mosquitoes. Insect Biochemistry and Molecular Biology 34:707–713.
Rasgon, J. L., L. M. Styer, and T. W. Scott. 2003. Wolbachia-induced mortality as a mechanism to modulate pathogen transmission by vector arthropods. Journal of Medical Entomology 40:125–132.
Schofield, P. 2002. Spatially explicit models of Turelli-Hoffmann Wolbachia invasive wave fronts. Journal of Theoretical Biology 215:121–131.
Scott, T. W., W. Takken, B. G. J. Knols, and C. Boete. 2002. The ecology of genetically modified mosquitoes. Science 298:117–119.
Service, M. W. 1993. Mosquito Ecology. Field Sampling Methods. Elsevier, London.
Sinkins, S. P., and F. Gould. 2006. Gene-drive systems for insect disease vectors. Nature Reviews. Genetics 7:427–435.
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. Ecological Modelling 163:33–44.
Turelli, M. 2010. Cytoplasmic incompatibility in populations with overlapping generations. Evolution 64:232–241.
Turelli, M., and A. Hoffmann. 1991. Rapid spread of an inherited incompatibility factor in California Drosophila. Nature 353:440–442.
Wade, M. J., and R. W. Beeman. 1994. The population dynamics of maternal-effect selfish genes. Genetics 138:1309–1314.
Wang, M.-H., M. Kot, and M. G. Neubert. 2002. Integrodifference equations, Allee effects, and invasions. Journal of Mathematical Biology 44:150–168.
Yakob, L., and M. B. Bonsall. 2009. Importance of space and competition in optimizing genetic control strategies. Journal of Economic Entomology 102:50–57.
Yakob, L., L. Alphey, and M. B. Bonsall. 2008a. Aedes aegypti control: the concomitant role of competition, space and transgenic technologies. The Journal of Applied Ecology 45:1258–1265.
Yakob, L., I. Z. Kiss, and M. B. Bonsall. 2008b. A network approach to modeling population aggregation and genetic control of pest insects. Theoretical Population Biology 74:324–331.