

ORIGINAL ARTICLE

Gene-drive in age-structured insect populationsYunxin Huang¹, Alun L. Lloyd², Mathieu Legros¹ and Fred Gould¹¹ Department of Entomology, North Carolina State University, Raleigh, NC, USA² Department of Mathematics, North Carolina State University, Raleigh, NC, USA**Keywords**

assortative mating, fitness cost, genetic control, mosquito, release threshold.

Correspondence

Yunxin Huang, Department of Entomology, University of Kentucky, S-225 Agricultural Science Center, Lexington, KY 40546-0091, USA.

Tel.: (859) 257 5117; fax: (859) 323 1120;

e-mail: yunxin.huang@gmail.com

Received: 1 July 2008

Accepted: 30 September 2008

First published online: 12 November 2008

doi:10.1111/j.1752-4571.2008.00049.x

Abstract

To date, models of gene-drive mechanisms proposed for replacing wild-type mosquitoes with transgenic strains that cannot transmit diseases have assumed no age or mating structure. We developed a more detailed model to analyze the effects of age and mating-related factors on the number of engineered insects that must be introduced into a wild population to achieve successful gene-drive based on the *Medea* and engineered underdominance mechanisms. We found that models without age-structure and mating details can substantially overestimate or underestimate the numbers of engineered insects that must be introduced. In general, introduction thresholds are lowest when young adults are introduced. When both males and females are introduced, assortative mating by age has little impact on the introduction threshold unless the introduced females have diminished reproductive ability because of their age. However, when only males are introduced, assortative mating by age is generally predicted to increase introduction thresholds. In most cases, introduction thresholds are much higher for male-only introductions than for both-sex introductions, but when mating is nearly random and the introduced insects are adults with *Medea* constructs, male-only introductions can have somewhat lower thresholds than both-sex introductions. Results from this model suggest specific parameters that should be measured in field experiments.

Introduction

Replacement of disease-vectoring native strains of mosquitoes with genotypes that are refractory to disease transmission has been proposed as a potential strategy to control insect-borne diseases such as malaria and dengue (Scott et al. 2002; James 2005; Gould et al. 2006). The strategy involves two key components: an anti-pathogen transgene and a gene-drive mechanism by which the anti-pathogen gene spreads to high frequency or fixation in natural populations. Anti-pathogen candidate genes have been developed and tested in mosquitoes (Ito et al. 2002; Franz et al. 2006). A number of gene-drive mechanisms have been proposed (Sinkins and Gould 2006), and one drive mechanism, *Medea*, has recently been successfully engineered in *Drosophila* (Chen et al. 2007).

Because experimental tests of a gene-drive mechanism can only be conducted after major investment of time

and funds, early assessment with theoretical models is desirable (Scott et al. 2002). Analysis using simple population genetic models that do not consider the ecological complexity of the natural populations is the first step of the theoretical assessment. Simple models have been used to analyze a number of candidate gene-drive systems, such as sex-linked meiotic drive (Huang et al. 2007), homing endonuclease genes (Burt 2003, 2004), *Medea* (Wade and Beeman 1994) and engineered underdominance (EU) (Davis et al. 2001; Magori and Gould 2006). These gene-drive systems involve naturally occurring or artificially constructed genetic elements that spread at the expense of the host (Burt and Trivers, 2006). In ideal situations, models indicate that some of these elements should spread from very low initial frequencies. However, their ability to spread, together with their speed of spread and the final frequency they attain, depends on the fitness costs associated with the inserted genes. Simple models

predict that these gene-drive systems, in the presence of fitness costs, exhibit *introduction thresholds*: a minimum initial frequency must be exceeded in order for the genetic element to spread into a natural population.

In addition to fitness costs, ecologically complex factors such as host age structure and mating are expected to affect gene-drive. Surprisingly little attention has been paid to these factors. A single model with age structure by Rasgon and Scott (2004) examined the spread of *Wolbachia*, a bacterium species that can spread by causing cytoplasmic incompatibility. Although that study examined the impact of the age structure of mosquito populations on the spread of the bacterium, the exploration of the parameter conditions was limited and it is not clear whether the results of that study are applicable to other gene-drive systems. For instance, Rasgon and Scott assumed random mating and considered simultaneous release of males and females. However, it is unlikely that the mating of mosquitoes is completely random, and, because males do not transmit human diseases, it is expected that releases would solely involve males (Klassen and Curtis 2005).

There are a number of age-related factors that may affect the success of a gene-drive system. Of interest are the age-dependent reproductive pattern, the age at which individuals mate, and the degree to which mating is assortative by age. Some of these factors are captured by the classical notion of reproductive value (Fisher 1930; Gotelli 2001), although the applicability of this concept is, as discussed below, limited in this context. As some of these factors may vary from population to population, it is important to parameterize these factors and examine their impact when they vary within a biologically realistic range.

In this paper, we choose two specific gene-drive systems, EU and *Medea*, as cases for examining the potential impact of age structure on gene-drive by numerically analyzing an age-structured population genetic model. These two drive mechanisms were chosen because they differ dramatically in the critical number of engineered insects that need to be introduced. We address three key questions: (i) Does the consideration of age structure significantly change the predictions for the number of engineered insects needed to achieve gene-drive? (ii) What are the best age-specific release strategies for gene-drive? (iii) How does the introduction of only males affect the introduction thresholds?

The population genetics of engineered underdominance and *Medea*

In a gene-drive system, certain types of embryos may be not viable as caused by the gene-drive mechanism,

while those viable embryos with transgenic insertions may have a reduced survival probability compared to those that do not have any transgenic insertions. In the models we use a single mathematical function $f_e(G,J,K)$ to describe the survival probability of an embryo of genotype G , which equals the product of embryo viability and fitness (i.e., the survival probability of viable genotypes). This probability may depend on the parental genotypes J and K . It is also possible that the transgenic insertions reduce the fecundity of adult females. To incorporate this type of fitness cost, we use another function $f_b(J)$ to describe the fecundity of females of genotype J .

Engineered underdominance, originally proposed by Davis et al. (2001), involves the introduction of individuals carrying two co-dependent engineered constructs a and b on separate chromosomes. 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). More generally, there will be fitness costs associated with the EU constructs. We assume that these costs will be expressed as embryonic mortality. Biological justification can be found for assuming inheritance of fitness costs varying from recessive to dominant. To avoid adding additional complexity to our comparison of age-structured and nonage-structured models, we assumed additive inheritance at each locus and multiplicative effects across loci. The fitness and viability are independent of the parental genotypes, so $f_e(G,J,K) = f_e(G)$. Genotypes that only have one of the two EU constructs, i.e., $AABb$, $AAbb$, $AaBB$ and $aaBB$, are not viable and so, for these four genotypes, $f_e(G)=0$. The embryonic fitnesses of the remaining five 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. In this paper, c is allowed to vary from 0 to 0.25. The equations for the basic model describing EU can be found in Appendix A, and further details of the model can be found in Davis et al. (2001) and Magori and Gould (2006).

Medea is a selfish genetic element that has been found in natural *Tribolium* beetle populations (Beeman et al.

1992) and has been successfully engineered in *Drosophila* (Chen et al. 2007). When homozygous *Medea* individuals (*MM*) are introduced and mate with the wild-type individuals (*++*), there are three genotypes in the F_2 generation: *MM*, *M+* and *++*. Nine different mating types are possible when sex and genotype are considered. Among these mating types, those with mothers carrying a *Medea* allele can cause maternal-effect lethality to their offspring that do not inherit the *Medea* allele from the mother or father. Here we assume that this maternal-effect lethality is total and so $f_e(G,J,K) = 0$ if the mother-*J* contains the element while the offspring-*G* is wild-type. Offspring carrying the *Medea* construct are viable, but may have an embryonic fitness cost due to the transgenic insertion.

When there is no associated fitness cost, the *Medea* element is expected to increase in frequency from arbitrarily low initial frequencies (in a deterministic setting). The presence of fitness costs leads to a nonzero introduction threshold, and when these costs are substantial, the *Medea* element will not spread unless a large number of *Medea*-bearing insects are introduced.

We assume that *Medea* imposes an additive fitness cost on viable embryos, so that $f_e(G,J,K) = 1, 1 - c/2$ or $1 - c$ when the genotype of viable embryo is *++*, *M+*, or *MM*, respectively. (As mentioned above, the *++* genotype will only be viable if the mother is also of the *++* genotype.) Here c is the embryonic fitness cost for a homozygous *Medea* individual. We assume that *Medea* also leads to a reduction in the fecundity of females, with this fitness cost again being additive, and so $f_b(J) = 1, 1 - s/2$, or $1 - s$ for female genotypes *++*, *M+*, or *MM*, respectively. Here $0 \leq s \leq 1$ is the fecundity loss of a homozygous *Medea* mother. Throughout this paper, we set $s = 0.1$. This biologically reasonable value results in a baseline introduction threshold for *Medea*, even when there is no embryonic fitness cost. The equations of the basic model are given in Appendix B, and further details can be found in Wade and Beeman (1994) and Chen et al. (2007).

The age-structured model

Engineered underdominance is modeled as two transgenically inserted alleles, with one allele on each of the two independently segregating chromosomes or linkage groups. The *Medea* element is modeled as a single allele.

We model the mosquito life cycle using 35 daily age classes, assuming that the first 10 of these represent immature stages and the remainder are adult stages. All individuals are assumed to mature after 10 days, assuming that they survive to this age. The daily survival rate is assumed to be an age-independent constant, and is taken to equal 0.9 (McDonald 1977). Birth rates of females are age-dependent and described by the function $b(a)$, the

average number of eggs produced per female per day. As the purpose of this study is to address the question of whether age structure affects gene-drive, the reproductive patterns adopted here only aim to capture qualitative features of the real-world reproductive pattern. We employ two patterns, which we call I and II [Fig. 1, the caption of which gives details of the functional form we adopt for $b(a)$]. For pattern I, the female reproductive interval is between age 13 and age 24, with highest fecundity at age 15. In pattern II females have a broader reproductive interval, with highest fecundity at age 20. Based on empirical data (e.g., Harrington et al. 2001), patterns I and II more or less depict extremes in terms of the length of reproductive interval.

For mosquitoes such as *Aedes aegypti*, females usually mate only once, and occasionally can mate more than once in their lifetime (Foster and Lea 1975; Williams and Berger 1980; Young and Downe 1982). In this paper, we assume that females mate only twice in their entire lifetime. The first mating occurs when adults just emerge (i.e., at age 11), while the second mating occurs 10 days after the first mating (i.e., at age 21) when the majority of a cohort has already died. We assume perfect sperm precedence by the last male to mate. As can be seen in

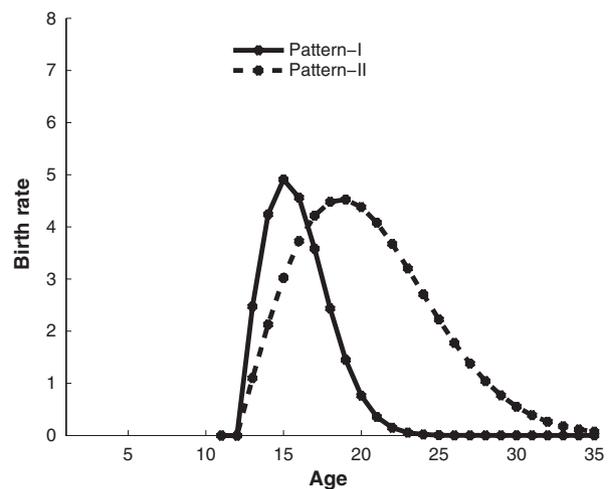


Figure 1 The two age-dependent female reproductive patterns, I and II, employed in our model, shown as the birth rate $b(a)$, i.e., the average number of eggs produced per female per day. For each pattern, insects are classified into 35 age classes: 10 immature age classes and 25 adult age classes. The functional forms of these patterns are provided by truncated Weibull distributions. More precisely, we assume that $b(a) = 0$ for $a \leq 12$ and $b(a) = Sb(a) / \sum_{a'=13}^{35} \hat{b}(a')$ for $a \geq 13$. Here $\hat{b}(a) = 2(a - 12)\lambda^{-2} \exp\{-(a - 12)^2/\lambda^2\}$, for $a \geq 13$. S is the total number of eggs that a female produces in her lifetime if she survives to the maximum age. In reproductive pattern I, $S = 25$ and $\lambda = 4.4$. In reproductive pattern II, $S = 50$ and $\lambda = 9.5$. Note that the larger the parameter λ is, the broader the reproductive interval.

Fig. 1, for reproductive pattern I a female's fecundity is very low by 10 days after her first mating, so the genes of males involved in the second mating are barely represented in the next generation.

Let $N(G,a,t)$ and $N^*(G,a,t)$ be the number of females and the number of males of age a and genotype G at time t , respectively. We assume a 1:1 sex ratio at birth and that survival is sex- and genotype-independent. The age and genotype-specific numbers of the female population can be tracked by the following deterministic recursive equations:

$$N(G, 1, t + 1) = 0.5B(G, t), \quad (1)$$

$$N(G, a + 1, t + 1) = 0.9N(G, a, t), \quad a = 1, \dots, 34. \quad (2)$$

The dynamics of the male population can be tracked by the same set of equations with $N(G,a,t)$ being replaced by $N^*(G,a,t)$. The total number of viable eggs of genotype G produced by adults at time t , $B(G,t)$, can be expressed as

$$B(G, t) = \sum_{J,K} f_e(G, J, K) \Pr(G|J, K) B_1(J, K, t). \quad (3)$$

Here $f_e(G,J,K)$ measures the viability and relative fitness (i.e., survival probability) of a G -egg produced by a J -mother and K -father, as discussed in the previous section. $\Pr(G|J,K)$ is the probability that a zygote has genotype G when the parental genotypes are J and K , respectively. In the case of Mendelian segregation, this probability can be easily calculated by a computer algorithm. $B_1(J,K,t)$ is the total number of eggs produced by females of genotype J who mated with males of genotype K , which can be further expressed as

$$B_1(J, K, t) = \sum_{a=1}^{35} b(a) f_b(J) N(J, a, t) \times \sum_{t'=0}^{m-1} \psi(a, t') \beta(a - t', K, t - t'). \quad (4)$$

Here, $b(a)$ is the average number of eggs produced per day per female of age a , as discussed above. $\beta(a,K,t)$ is the probability that a female of age a at time t mates with a male of genotype K . $\psi(a,t')$ is the probability that a female of age a last mated t' days ago. For the two-occasion mating discussed earlier, we have that $\psi(a, a - 11) = 1$ for $11 \leq a \leq 20$, $\psi(a, a - 21) = 1$ for $a \geq 21$, and $\psi(a, t') = 0$ otherwise. For the multiple/everyday mating we have that $\psi(a, 0) = 1$ and $\psi(a, t') = 0$ for $a \geq 11$ and $t' > 0$. The number of adult age classes is m ($=25$) and so the maximum time t' that can have elapsed since an individual was mated is $m - 1$ days. Notice that

a female of age a at the current time t who last mated t' days ago would then have been of age $a - t'$ and that the time then would have been $t - t'$.

In this paper we assume that mating is random with respect to genotype, but is potentially assortative with respect to age. Females choose their mates with probability weighted by a mating preference function $\phi(x)$ that depends on the age difference x between the female and a given male. In order to model different degrees of assortative mating we take $\phi(x)$ to have the form

$$\phi(x) = q^x / (1 + q + \dots + q^{m-1}), \quad \text{for } x = 0, \dots, m - 1. \quad (5)$$

Here, $0 \leq q \leq 1$ is a parameter by which we can adjust the degree of assortative mating. (For the situation in which $q = 0$, we adopt the notational convention that $0^0 = 1$.) The maximum possible age difference between male and female at mating is $m - 1$.

For a fixed $0 < q < 1$, $\phi(x)$ is a strictly decreasing function of x , which means that the mating preference decreases as the age difference increases. When $q = 0$, $\phi(0) = 1$ and $\phi(x) = 0$ for any $x > 0$, mating is completely assortative. When $q = 1$, $\phi(x) = 1/m$ is constant for all x , mating is random. In this paper we consider five mating preference patterns for which $q = 1, 0.94, 0.875, 0.75$, and 0 , respectively. The five mating patterns (MP), for which the degrees of assortative mating by age vary from being random mating to being completely assortative, are denoted by MP-1 through MP-5 (Fig. 2).

In terms of the mating preference function, we find that the mating function $\beta(a,K,t)$ is given by

$$\beta(a, K, t) = \frac{\sum_x \phi(|a - \alpha|) N^*(K, \alpha, t)}{\sum_{K'} \sum_x \phi(|a - \alpha|) N^*(K', \alpha, t)}, \quad (6)$$

where the sums are only taken over adult age classes.

We consider a one-time introduction of homozygous engineered individuals, i.e., genotype $aabb$ in the case of EU and genotype MM in the case of *Medea*. We consider three specific age-related introduction methods: (i) Single-age introduction, where all introduced engineered individuals have the same age; (ii) Two-age introduction, where equal numbers of engineered individuals from two different age classes are introduced; (iii) All-age introduction, where engineered individuals from all age classes are introduced, but with the same age distribution as the existing wild-type population. We assume that no introduced female has been mated prior to introduction. (We remark that this leads to a minor modification in the above model equations as the mating history of introduced and pre-existing females must be tracked separately until sufficient time has passed for the introduced cohort to

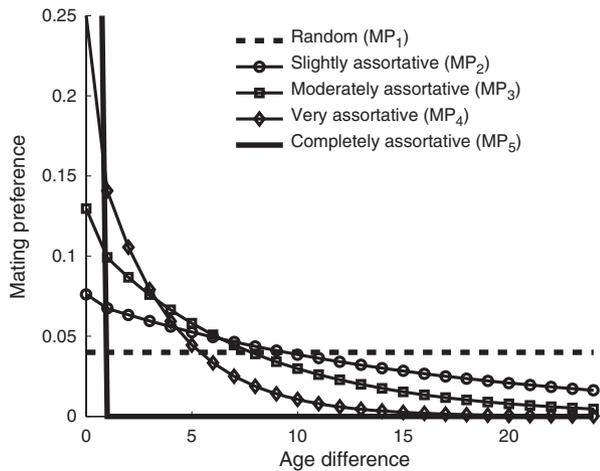


Figure 2 Graphs of the five mating preference patterns, $\phi(x)$, where x is the age difference between a male and a female. The five patterns represent different degrees of mating preference varying from random to completely assortative by age, which are denoted by MP_1 , MP_2 , MP_3 , MP_4 and MP_5 , respectively.

have died.) We assume that introductions are made into a wild-type population that has attained its stable age distribution. In the simulations, we find this distribution by running the model for a sufficiently long time. All results presented in this paper were obtained numerically using a C++ simulation code (available on request from the authors).

Results

In the case of EU, the age-structured model has two stable equilibria E_{s1} and E_{s2} , where the frequency of the wild genotype is 1 (constructs are lost) and at some low level (constructs are at high frequency), respectively. When there is no fitness cost ($c = 0$) the frequency of the wild genotype at E_{s2} is 0 (constructs are fixed). The frequency of wild genotype at E_{s2} increases as c increases. The age-structured model also has an unstable equilibrium, the location of which determines (in a nontrivial way) the introduction threshold. When the initial introduction frequency is below the threshold, the system approaches the E_{s1} equilibrium and the wild-type goes to fixation. When the initial introduction frequency is above the threshold, the system approaches the E_{s2} equilibrium, at which the frequency of the wild genotype is at some low level. The wild-type frequency is lower than 0.05 when the fitness cost c is lower than or equal 0.25.

In the case of *Medea*, the age-structured model has two stable equilibria at which the frequencies of the wild genotype are 1 and 0, respectively. When there are fitness costs associated with the *Medea* construct, the

age-structured model has a unstable equilibrium whose location determines the introduction threshold. When the initial introduction frequency is below the threshold, the wild-type goes to fixation. When the initial introduction frequency is above the threshold, the wild genotype goes extinct.

Equilibrium genotype frequencies in the age-structured model are the same as those in the corresponding non-age-structured model. This result stems from our assumption that genotype-specific fitness effects (i.e., fecundity differences and survival probability differences) only occur at reproduction and at the embryonic stage, but not as individuals move between the various age classes. Consequently, at equilibrium each age class has the same genotype frequency distribution and so these equilibrium frequencies are identical to those in the nonage-structured model.

In the following sections we focus on the introduction thresholds to examine the effects of age structure. Introduction thresholds presented below are found numerically based on whether the long-term frequency of wild genotype is close to 1 or close to 0. Introduction threshold will be calculated as the proportion of released number of insects relative to the total population (including all age classes, both immature and adult) unless mentioned otherwise.

Single-age introduction of both sexes

In the case of EU, the introduction threshold varies considerably with the age of the introduced individuals. Compared with the introduction threshold predicted by the nonage-structured model, the threshold is relatively high for the introduction of a single age class of immature individuals or old adult individuals and relatively low for the introduction of a single age class of young adult individuals (Fig. 3A,B). For example, when there is no fitness cost, the nonage-structured model predicts a threshold of 27%, but the age-structured model with reproductive pattern I predicts a threshold of 13% for the introduction of age class 13 and a threshold of 75% for the introduction of age class 25.

In broad terms, the threshold is inversely related to the reproductive value of the introduced individuals (Fisher 1930; Gotelli 2001). We attempted to more specifically relate the introduction threshold to reproductive value, but found that this was not possible because calculation of reproductive value is based only on females, and males do not have the same reproductive value as females unless mating is completely assortative by age. Furthermore, reproductive value is altered by the introduction itself and the limited number of matings. While it is still useful to analyze the qualitative relations between the

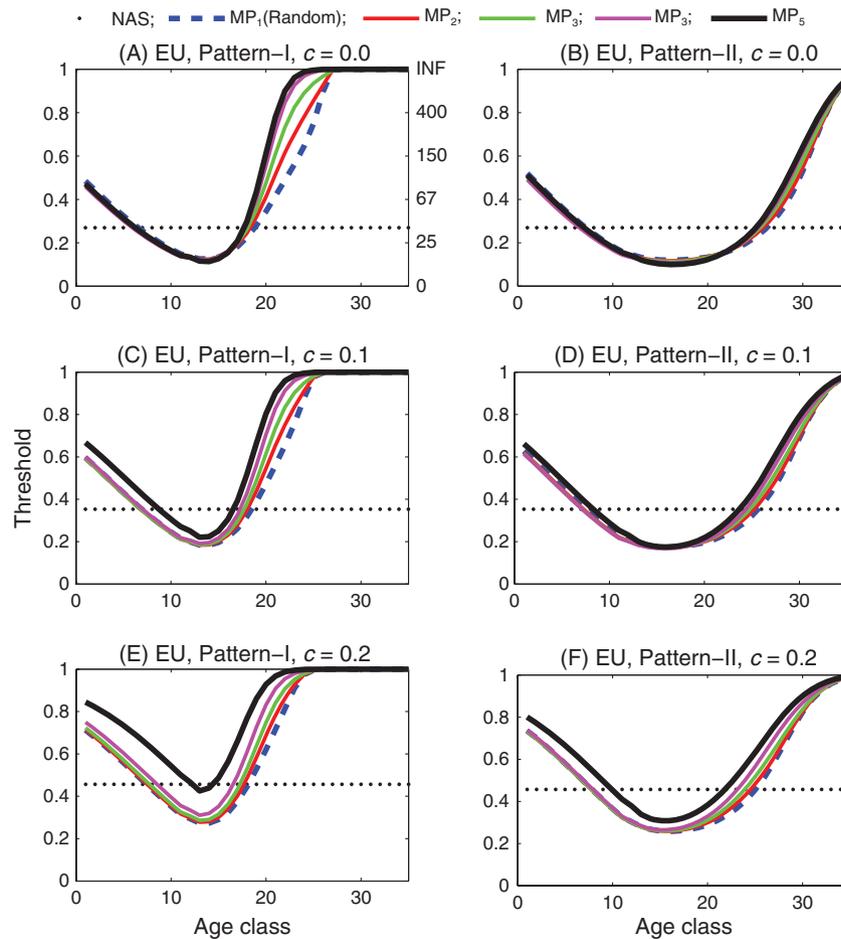


Figure 3 Engineered underdominance: Introduction threshold versus age class for the single-age introduction of both males and females. In panels (A), (C) and (E), the reproductive pattern is I and the fitness cost c is 0, 0.1 and 0.2, respectively. In panels (B), (D) and (F), the reproductive pattern is II and the fitness cost c is 0, 0.1 and 0.2, respectively. In each panel the horizontal dotted line is the threshold predicted by the non-age-structured model (labeled as NAS). The remaining five curves labeled by MP_1 through MP_5 are the age-specific thresholds corresponding to random, slightly assortative, moderately assortative, very assortative and completely assortative mating by age, respectively. In each panel, the left-hand vertical axis gives the threshold in terms of the proportion of introduced insects relative to the total population, whereas the right-hand vertical axis gives the threshold in terms of the number of introduced insects per 100 wild-type insects (INF means infinity).

introduction threshold and reproductive value, a quantitative analysis is not possible. [Further details regarding the application of the reproductive value concept to models such as the one presented here will appear elsewhere (A.L. Lloyd, unpublished data).]

We assumed that females could only mate twice based on empirical data on mosquitoes. Under this assumption, previously emerged wild adult females in the natural population have already mated with wild-type males before the introduction of engineered insects and they do not mate again until up to 10 days after the introduction. In other words, the only wild-type females that can mate at the time of introduction are those newly emerged ones. Compared to the implicit assumption of unlimited multiple mating in the simple nonage-structured model, the

sexual nonreceptivity of wild-type females at the time of introduction limits the production of heterozygous offspring that will contribute to gene-drive in the subsequent generation. Therefore, for the introduction of old adults, this limitation on mating generally causes higher thresholds than unlimited multiple/everyday mating. This has been verified by systematic comparisons between the thresholds for two-occasion and multiple-mating (results not shown). For example, if the reproductive pattern I, there is no fitness cost, and the introduced insects are 22 days old, then the thresholds assuming two-occasion mating and unlimited multiple-mating are 48.8% and 43.5%, respectively. The difference is greater when mating is assortative by age and/or when the introduced insects are older.

Mating preference has little impact on the introduction threshold for immature or young adult age classes. The situation is different, however, for the introduction of an old adult age class because, depending on the MP, older males and females can make quite different contributions to the offspring pool. Older females have very low fecundity, and if mating is assortative then older males correspondingly contribute little to the offspring pool. As the degree of assortativity of mating decreases, older males increasingly mate with younger females and hence benefit from their higher fecundity. As a consequence, when mating is anything other than totally assortative, older males can contribute more to gene-drive than older females, and so introduction thresholds decrease as the degree of assortativity decreases. When mating is random, older males frequently mate with young females and so the

introduction threshold can be considerably lower than under strong assortative mating.

The introduction threshold increases as the fitness cost c increases (Fig. 3C–F). However, even if the fitness cost is as high as 20% the thresholds for the introduction of an immature or a young adult age class do not differ significantly among different mating preference patterns unless mating is almost completely assortative by age.

In reproductive pattern I, females have very low fecundity when they are older than 20 days (Fig. 1), so the threshold for the introduction of an age class larger than 20 is very high (Fig. 3A,C,E). In reproductive pattern II, females of age 20 still have rather high fecundity (Fig. 1), so the threshold for the introduction of age class 20 is very low compared with that in reproductive pattern I

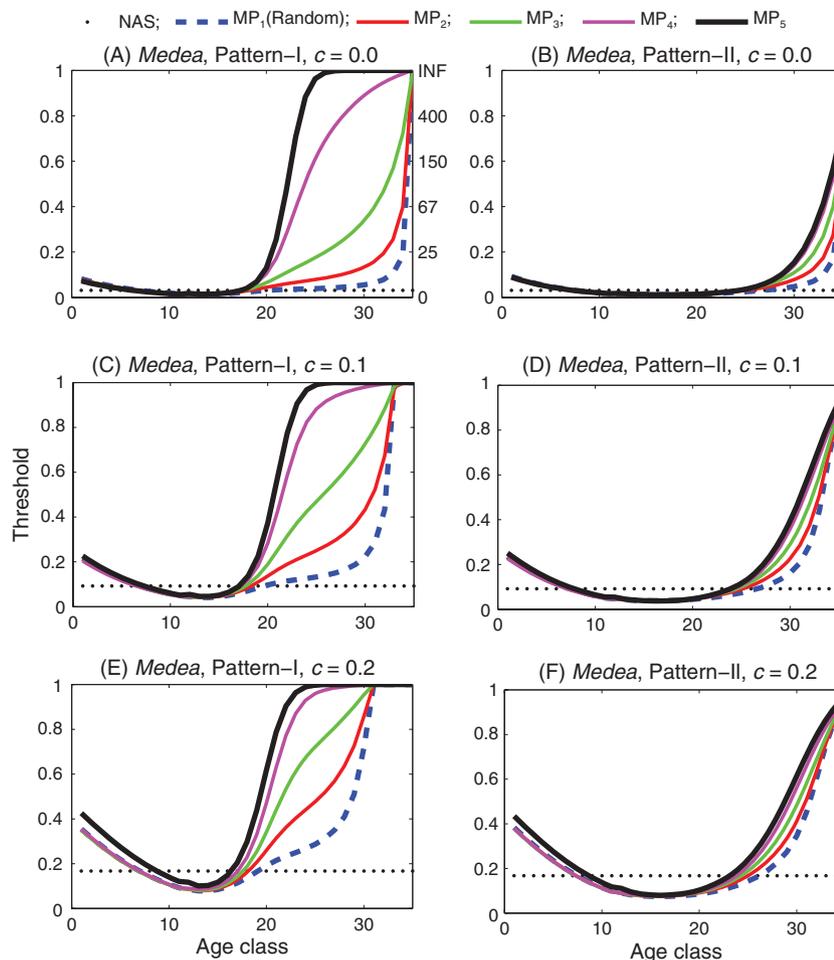


Figure 4 *Medea*: Introduction threshold versus age class for the single-age introduction of both males and females. In this figure, the fecundity loss of a homozygous *Medea* female is $s = 0.1$. In panels (A), (C) and (E), the reproductive pattern is I and the fitness cost c is 0, 0.1 and 0.2, respectively. In panels (B), (D) and (F), the reproductive pattern is II and the fitness cost c is 0, 0.1 and 0.2, respectively. In each panel, the dotted line is the threshold predicted by the nonage-structured model. The remaining five curves labeled by MP_1 through MP_5 are the age-specific thresholds corresponding to random, slightly assortative, moderately assortative, very assortative and completely assortative mating by age, respectively.

Table 1. Comparisons of the introduction thresholds between the nonage-structured model (NAS) and the age-structured model (for the introduction of age class 13). Mating is random. The thresholds are given as the number of introduced insects per 100 wild-type insects.

	EU ($c = 0$)	EU ($c = 0.1$)	EU ($c = 0.2$)
Threshold (NAS)	36.8	54.8	84.2
Threshold (age 13)	14.4	24.7	37.8
Difference	22.4	30.1	46.4
	<i>Medea</i>	<i>Medea</i>	<i>Medea</i>
	($c = 0$)	($c = 0.1$)	($c = 0.2$)
Threshold (NAS)	3.2	10.1	20
Threshold (age 13)	1.4	4.3	8.5
Difference	1.8	5.8	11.5

(Fig. 3B,D,F). The impact of assortative mating on the threshold in reproductive pattern I is stronger than in reproductive pattern II. This is partially due to the different contributions of males between the two reproductive patterns.

In the case of *Medea*, when a single age class of *Medea* males and females are introduced, the degree of assortative mating by age has a similar impact on the introduction thresholds as seen in the case of EU, except that the general threshold levels are much lower (Fig. 4).

In addition to their threshold levels, EU and *Medea* differ in the extent to which the threshold is affected by age structure. To show this we focus on the minimum

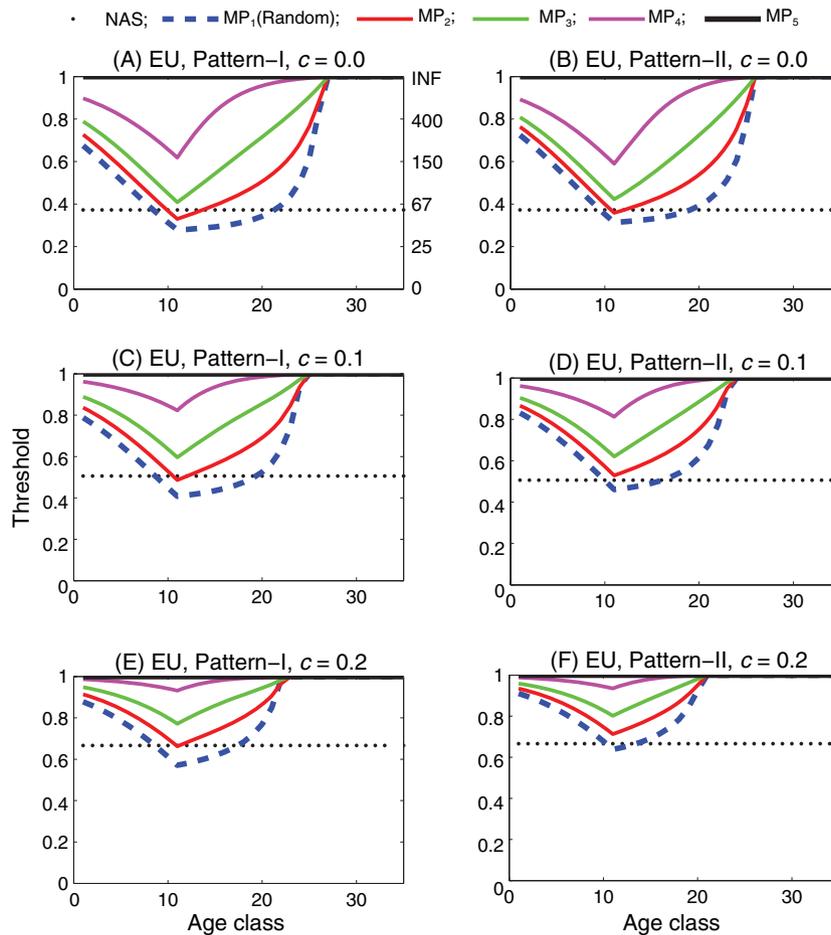


Figure 5 Engineered underdominance: Introduction threshold versus age class for the single-age introduction of only males. In panels (A) through (F), only males are introduced. Reproductive pattern is I in panels (A), (C) and (E), and II in panels (B), (D) and (F). Fitness cost c is 0 in panels (A) and (B), 0.1 in (C) and (D) and 0.2 in (E) and (F). In each panel the horizontal dotted line is the threshold predicted by the nonage-structured model (labeled as NAS). The remaining five curves labeled by MP₁ through MP₅ are the age-specific thresholds corresponding to random, slightly assortative, moderately assortative, very assortative and completely assortative mating by age, respectively. Note that for completely assortative mating by age (MP₅), the introduction thresholds are 1 for any age class, i.e., successful introduction is not possible. (This means that the black solid line is on the top of the frame in each panel.)

numbers of insects that must be introduced per 100 wild-type individuals which is an alternative way to express the threshold. In the case of EU, when mating is random, the absolute differences in the minimum numbers of insects that must be introduced per 100 wild-type individuals between the age-structured model (for the introduction of age class 13) and the simple nonage-structured model are 22.4, 30.1 and 46.4 for $c = 0, 0.1$ and 0.2 , respectively (Table 1). If the target population has one million insects, then for $c = 0.2$, the simple model suggests that 464 000 more insects would have to be introduced than would be predicted by the age-structured model. From a practical perspective, this is a big difference. In the case of *Medea*, the absolute differences in the introduction thresholds between the age-structured model and the simple nonage-structured model are 1.8, 5.8 and 11.5 per 100 wild-type individuals for $c = 0, 0.1$

and 0.2 , respectively. Compared to the case of EU, the differences between the simple model and the age-structured model are much smaller.

Single-age introduction of only males

In the case of a male-only introduction of EU insects, the introduction thresholds are generally higher than those for the corresponding introduction of both males and females (Fig. 5). As in the case of both-sex introductions, the introduction threshold varies considerably with the age of the introduced individuals, and can be much lower or higher than that predicted by the nonage-structured model. For example, when there is no fitness cost, the nonage-structured model predicts a threshold of 38%, but the age-structured model with reproductive pattern I predicts a threshold of 29% for the introduction of age class

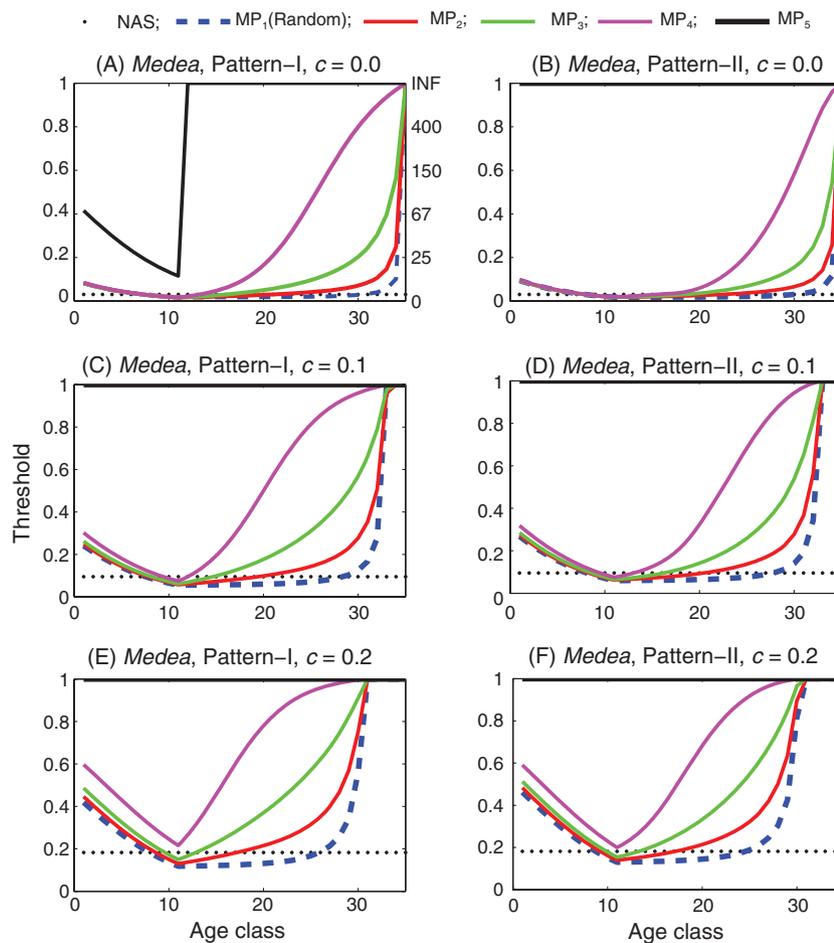


Figure 6 *Medea*: Introduction threshold versus age class for the single-age introduction of only males. In panels (A) through (F), only males are introduced. Reproductive pattern is I in panels (A), (C) and (E), and II in panels (B), (D) and (F). Fitness cost c is 0 in panels (A) and (B), 0.1 in (C) and (D) and 0.2 in (E) and (F). the fecundity loss of a homozygous *Medea* female is $s = 0.1$.

13 and a threshold of 62% for the introduction of age class 25. Note that in reproductive pattern I, males of age 25, as they can mate with younger females, have much larger contributions to gene-drive than females of the same age, so the introduction of only males of age 25 results in a lower threshold than the introduction of both males and females. Thresholds differ between reproductive patterns I and II, but the differences are not as substantial as they were for the introduction of both males and females.

The degree of assortative mating by age has a great impact on the introduction thresholds, especially when the fitness cost is high. This is largely due to the limited mating of the F_0 wild-type adult females with the introduced males, as discussed in the previous subsection. In the case of male-only introduction, the negative impact of this mating limitation imposed by wild-type adult females on gene-drive is even stronger than in the case of both-sex introduction because only the wild-type females that mate with the introduced males can pass the engineered

alleles to the next generation and contribute to gene-drive. In the worst case, i.e., when mating is completely assortative by age, the introduction of only males never achieves gene-drive, regardless of how many males are introduced (Fig. 5A–F).

In the case of a male-only single age-class introduction of *Medea* insects, the degree of assortative mating by age has a similar impact on the introduction thresholds as seen in the case of EU: the stronger the assortative mating is, the higher the threshold (Fig. 6). However, there are important differences between EU and *Medea*: (i) when mating is random, the threshold for the introduction of *Medea* males of age 25 is lower than 0.2 even when the fitness cost is as high as 0.2 (Fig. 6A–F), which is in contrast to the impossibility of achieving gene-drive with the introduction of EU males of the same age; (ii) when mating is random or slightly assortative by age and when the introduced insects are *Medea* adults, the introduction thresholds for the only-male introduction are almost always lower than those for the both-sex introduction

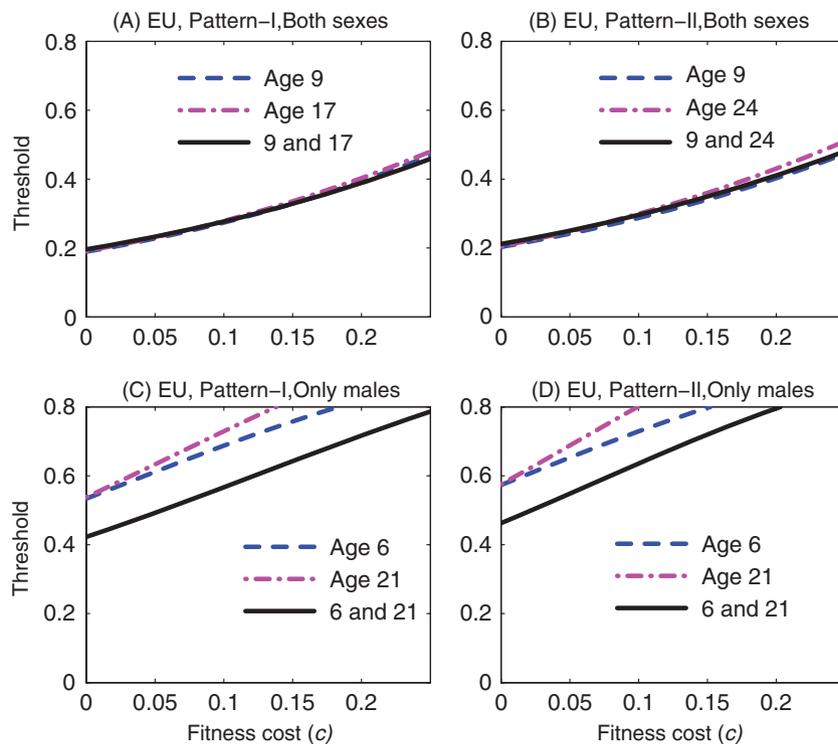


Figure 7 Engineered underdominance: Comparisons between the single-age introductions and two-age introduction. In panels (A) and (B), both males and females are introduced, while in panels (C) and (D), only males are introduced. In panels (A) and (C), the reproductive pattern is I, while in panels (B) and (D), the reproductive pattern is II. In each panel, the dashed line and the dot-dashed line illustrate the thresholds for the single-age introductions, while the solid line gives the thresholds for the two-age introduction. Note that in each panel the two age classes are chosen such that they result in similar introduction thresholds in the case of single-age introduction when there is no fitness cost, so these two age classes are not always the same among the four panels. In all panels, mating is assumed to be slightly assortative by age (i.e., MP_2 illustrated in Fig. 2).

(Figs 4 and 6). This is in contrast to the case of EU where the introduction thresholds for the only-male introduction are almost always higher than those for the both-sex introduction (Figs 3 and 5).

Two-age introduction

In order to determine whether there are differences between the introduction of a single age class and multiple age classes, we examine introductions involving two age classes. For illustration, we choose two age classes that have similar introduction thresholds in the case of single-age introduction when there are no fitness costs. (In order to make sure that the thresholds for the two age classes are as close as possible, the two age classes are chosen differently in some cases.)

In the case of EU, for reproductive pattern I, the introduction of both males and females of either age 9 or 17 results in a similar introduction threshold when there is no fitness cost ($c = 0$) and when mating is slightly assortative by age (with MP-2). When the fitness cost varies from 0 to 0.25, the thresholds for the simultaneous introduction of both age classes just differ slightly from the

thresholds for the single-age introduction (Fig. 7A). Similar results are observed for reproductive pattern II where the two age classes chosen for examination are 9 and 24 (Fig. 7B).

In the case of *Medea*, for reproductive pattern I, the introduction of both males and females of either age 9 or 17 also results in a similar introduction threshold when there is no embryonic fitness cost ($c = 0$) and when mating is slightly assortative by age. The threshold for the introduction of both age classes together is very similar to those for the single-age introductions even when c is 0.25 (Fig. 8A). Similar results are observed for reproductive pattern II where the two age classes chosen for examination are 9 and 23 (Fig. 8B).

In contrast, when only males are introduced there are significant differences between the single-age and two-age introductions. In the case of EU, the age-structured model for the two-age introduction predicts much lower thresholds than for the single-age introductions (Fig. 7C,D). In the case of *Medea*, the age-structured model for the two-age introduction also predicts lower thresholds than for the single-age introduction, but the differences are not so great as in the case of EU

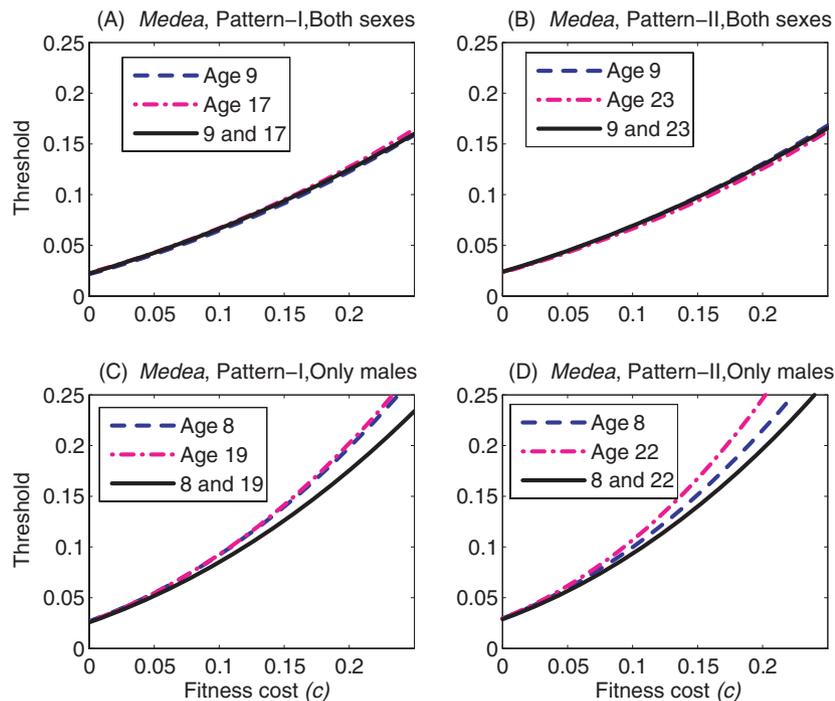


Figure 8 *Medea*: Comparisons between the single-age introductions and two-age introduction. In panels (A) and (B), both males and females are introduced, while in panels (C) and (D), only males are introduced. In panels (A) and (C), the reproductive pattern is I, while in panels (C) and (D), the reproductive pattern is II. In each panel, the dashed line and the dot-dashed line illustrate the thresholds for the single-age introductions, while the solid line gives the thresholds for the two-age introduction. In all panels, mating is assumed to be slightly assortative by age (i.e., MP₂ illustrated in Fig. 2).

(Fig. 8C,D). These differences between the thresholds are largely due to the differences in the chances of mating between the wild-type adult females and the introduced males at the time of introduction. When the introduced males are of the same age, they compete for the same wild-type females that are available for mating at the time of introduction. When the introduced males are of two distinct ages, the degree of competition is reduced. The degree of competition increases as the number of introduced males increases, and that is why differences are small when the thresholds are lower, as seen in the case of *Medea*.

All-age introduction

In order to compare the age-structured model with the simple nonage-structured model under the same initial conditions, we examined the age-structured model in the case of all-age introduction (where the number of introduced insects is proportional to the number of wild insects with the same age). When both males and females

are introduced, similar results are observed for EU and *Medea*: the introduction thresholds predicted by the age-structured model are slightly higher than those predicted by the simple model with the same value of the fitness cost. The differences in the thresholds become somewhat larger as c increases (Figs 9A,B and 10A,B). The higher thresholds in the age-structured model are largely due to the limited mating of the F_0 wild-type adult females, as discussed earlier.

When only males are introduced, the two-occasion mating of females has similar effects: during the initial period following the introduction, only those newly emerged adult wild-type females can mate with the introduced males and produce heterozygous offspring. Therefore, in the cases of EU and *Medea* the age-structured model predicts higher introduction thresholds than the nonage-structured model as long as the fitness cost is not very high (Figs 9C,D and 10C,D). However, when c is high and mating is random the age-structured model predicts lower introduction thresholds than the nonage-structured model in the case of EU (Fig. 9C,D).

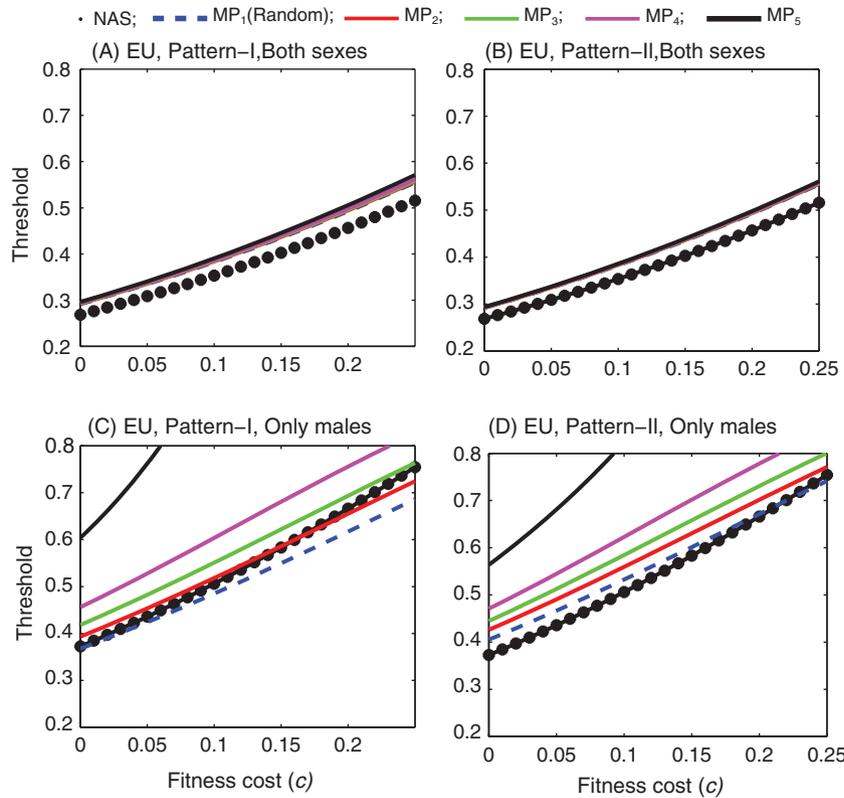


Figure 9 Engineered underdominance: Introduction thresholds as functions of fitness cost in the case of all-age introduction. In panels (A) and (B) both males and females are introduced. In panels (C) and (D) only males are introduced. In each panel, the thick dotted line is predicted by the simple nonage-structured model. The remaining five curves labeled by MP₁ through MP₅ correspond to the five mating preference patterns: random mating through completely assortative mating.

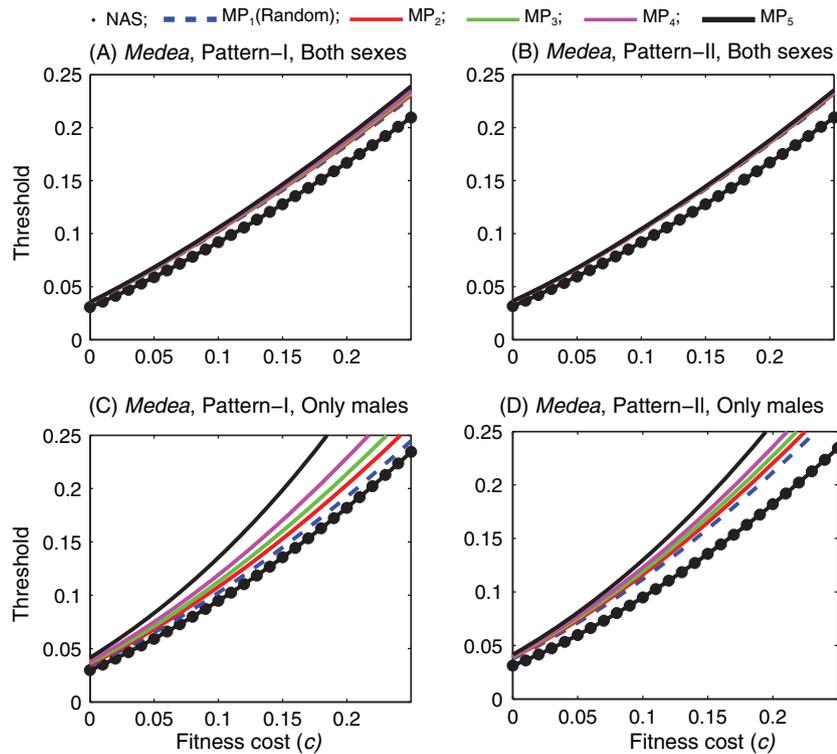


Figure 10 *Medea*: Introduction thresholds as functions of fitness cost in the case of all-age introduction. In panels (A) and (B) both males and females are introduced. In panels (C) and (D) only males are introduced. In each panel, the thick dotted line is predicted by the simple nonage-structured model. The remaining five curves labeled by MP_1 through MP_5 correspond to the five mating preference patterns: random mating through completely assortative mating.

When mating is random and only males are introduced, the age-structured model always predicts lower thresholds than the nonage-structured model as long as females mate every day and the introduction includes all age classes (Fig. 11). For example, when there is no fitness cost the threshold predicted by the age-structured model is about 33%, which is lower than the 38% predicted by the nonage-structured model. This is because in the nonage-structured model, males can only mate with females that are from the same discrete, nonoverlapping generation, while in the age-structured model with random mating, some males from the release generation survive long enough to mate with young females from the next generation. When mating occurs only twice instead of every day, the introduction thresholds predicted by the age-structured model increase, but they are still lower than those predicted by the nonage-structured model in cases where the fitness cost is high. Note that the fitness cost causes greater increases in thresholds in a nonage-structured population than in an age-structured population. It is also interesting that the age-structured model with complete assortative mating by age predicts almost the same thresholds as the nonage-structured model does.

Conclusions and discussion

In this study we used an age-structured genetic model to analyze the effects of age and mating-related factors on the number of engineered insects that must be introduced into a wild population to achieve successful gene-drive. In general, the introduction thresholds are lowest when young adults are introduced because they have the highest reproductive potential. When both males and females are introduced, assortative mating by age has little impact on the introduction threshold unless the introduced females are old, with diminished reproductive ability. Because males do not transmit disease, it has been recommended that only males be used in any introduction of engineered mosquitoes (Klassen and Curtis 2005). In general, male-only introductions increase introduction thresholds. For such introductions, even slight assortative mating by age can cause this increase to be dramatic. Even when there are no fitness costs and mating is random, the introduction of males 15 days after their emergence as adults cannot result in success of the EU approach. Previous models of EU and *Medea* drive mechanisms have not accounted for age structure and age-related mating

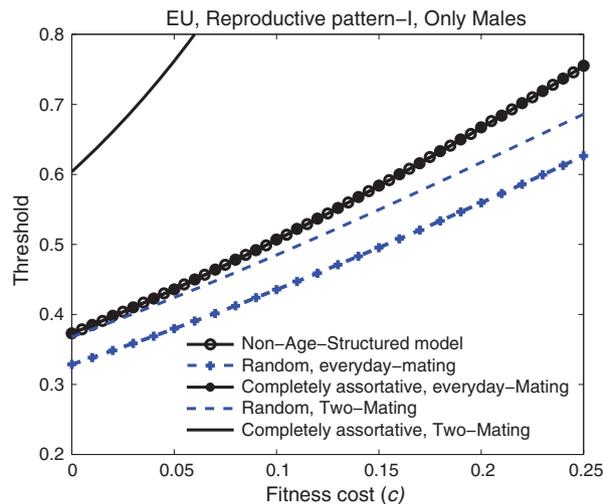


Figure 11 Engineered underdominance: Comparisons of the introduction thresholds among the nonage-structured model, the age-structured model with double-mating and the age-structured model with everyday-mating. In both the age-structured model and the nonage-structured model, only males are introduced. In the age-structured model, the introduction method is all-age introduction and the reproductive pattern is I. The two thicker lines with markers are the thresholds under the assumption of everyday-mating, while the two thinner lines with no markers are the thresholds under the assumption of two-mating. Note that the line corresponding to the nonage-structured model overlaps the line corresponding to the age-structured model when mating is completely assortative by age and occurs everyday.

details, and can overestimate or underestimate the numbers of engineered insects that must be introduced.

In this paper we have selected two specific gene-drive mechanisms as examples to examine the impact of age structure and age-related factors: *Medea* and EU. *Medea* achieves gene-drive by maternal-effect lethality to the offspring of *Medea* mothers that do not inherit a *Medea* allele from their mother or father. For EU, lethality occurs when an embryo inherits just one of the two co-dependent constructs, but involves no maternal effects. Although these two mechanisms are different, most results concerning the impact of age structure and age-related factors for the two mechanisms are similar, except that the general introduction threshold levels for EU are much higher than for *Medea*.

For both gene-drive mechanisms, *Medea* and EU, there are no deaths of first generation offspring from a cross between homozygous engineered insects and wild-type insects, so the gene-drive typically does not start until the second generation. Any population factors that decrease the percentage of matings between the engineered strain and the wild-type strain in the release generation will slow the onset of drive. When

there are fitness costs, this will decrease the frequency of engineered alleles that are present in the second generation. Compared with random mating, assortative mating by age directly reduces the percentage of the engineered insects that mate with wild-type insects in the case of single-age introduction and so increases the thresholds. Compared with unlimited multiple mating, limited two-occasion mating and the fact that the wild-type adult females have mated with wild-type males before release also reduce the percentage of matings between the engineered insects and the wild-type insects, and this causes higher thresholds in the age-structured model than in the nonage-structured model.

Introduction of only transgenic male mosquitoes is generally considered more socially acceptable than the introduction of males and females because males do not bite humans and therefore do not transmit disease (Klassen and Curtis 2005). Male-only introductions (and introductions involving unmated sexually mature females) require the ability to sex insects on a large scale. There have been successful examples of separating mosquitoes by sex, e.g., according to pupal size or using genetic techniques (Klassen and Curtis 2005). However, as demonstrated in this paper, male-only introduction generally cause higher thresholds than both-sex introduction. It must also be understood that unless there is a way of causing death of female eggs in the last generation of factory-reared transgenic insects, the rearing cost associated with release of one transgenic male is expected to approach the cost of releasing one male and one female. The benefit of achieving the more socially acceptable approach of male-only release may far outweigh issues associated with critical thresholds and rearing costs, but these factors should at least be considered.

While we have demonstrated that age structure, mating behavior and fitness costs could work in combination to affect the population dynamics of gene-drive, we lack detailed field studies in these aspects. Based on the results of our model, it would be useful, but not critical to study the extent of assortative mating by age if both males and females are expected to be used in the introductions. However, if a male-only introduction is anticipated, knowledge of the degree of assortative mating would be essential. In our simulations, there was only a minor impact from the second mating (data not shown). However, if the lifespan and reproductive period of females were longer, the second mating could be more important. Given that these parameters vary between species and even from population to population, our results emphasize the need for detailed field studies of these parameters before predictions can be made about the numbers and ages of engineered insects to introduce in a specific situation.

Acknowledgements

We thank Catherine Ward for helpful discussion and comments. We are indebted to two anonymous reviewers who gave useful comments. This work is funded by 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 and a NIH grant R01-AI54954-01A2.

Literature cited

- Beeman, R. W., K. S. Friesen, and R. E. Denell. 1992. Maternal-effect, selfish genes in flour beetles. *Science* **256**:89–92.
- Burt, A. 2003. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proceedings. Biological Sciences/The Royal Society* **270**:921–928.
- Burt, A. 2004. Homing endonuclease genes: the rise and fall and rise again of a selfish element. *Current Opinion in Genetics and Development* **14**:609–615.
- Burt, A., and R. Trivers. 2006. *Genes in Conflict: The Biology of Selfish Genetic Elements*. Harvard University Press, Cambridge, U.K.
- Chen, C. H., H. Huang, C. M. Ward, J. T. Su, L. V. Schaeffer, M. Guo, and B. A. Hay. 2007. A synthetic maternal-effect selfish genetic element drives population replacement in *Drosophila*. *Science* **316**:597–600.
- Davis, S., N. Bax, and P. Grewe. 2001. Engineered underdominance allows efficient and economical introgression of traits into pest populations. *Journal of Theoretical Biology* **212**:83–98.
- Fisher, R. 1930. *The Genetical Theory of Natural Selection*. Dover, New York [reprinted 1958].
- Foster, W. A., and A. O. Lea. 1975. Renewable fecundity of male *Aedes aegypti* following replenishment of seminal vesicles and accessory glands. *Journal of Insect Physiology* **21**:1085–1090.
- Franz, A. W. E., I. Sanchez-Vargas, Z. N. Adelman, C. D. Blair, B. J. Beaty, A. A. James, and K. E. Olson. 2006. Engineering RNA interference- based resistance to dengue virus type-2 in genetically-modified *Aedes aegypti*. *Proceedings of the National Academy of Sciences of the United States of America* **103**:4198–4203.
- Gotelli, N. J. 2001. *A Primer Of Ecology*. Sinauer Associates, Sunderland, MA.
- Gould, F., K. Magori, and Y. Huang. 2006. Genetic strategies for controlling mosquito-borne diseases. *American Scientist* **94**:238–246.
- Harrington, L. C., J. D. Edman, and T. W. Scott. 2001. Why do female *Aedes aegypti* (Diptera: Culicidae) feed preferentially and frequently on human blood? *Journal of Medical Entomology* **38**:411–422.
- Huang, Y., K. Magori, A. L. Lloyd, and F. Gould. 2007. Introducing desirable transgenes into insect populations using Y-linked meiotic drive: a theoretical assessment. *Evolution* **61**:717–726.
- Ito, J., A. Ghosh, L. A. Moreira, E. A. Wimmer, and M. Jacobs-Lorena. 2002. Transgenic *anopheline* mosquitoes impaired in transmission of a malaria parasite. *Nature* **417**:452–455.
- 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.
- Magori, K., and F. Gould. 2006. Genetically engineered underdominance for manipulation of pest populations: a deterministic model. *Genetics* **172**:2613–2620.
- McDonald, P. T. 1977. Population characteristics of domestic *Aedes aegypti* (Diptera: Culicidae) in villages on the Kenya Coast I. Adult survivorship and population size. *Journal of Medical Entomology* **14**:42–48.
- 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.
- Scott, T. W., W. Takken, B. G. J. Knols, and C. Boete. 2002. The ecology of genetically modified mosquitoes. *Science* **298**:117–119.
- Sinkins, S. P., and F. Gould. 2006. Gene-drive systems for insect disease vectors. *Nature Reviews. Genetics* **7**:427–435.
- Wade, M. J., and R. W. Beeman. 1994. The population dynamics of maternal-effect selfish genes. *Genetics* **138**:1309–1314.
- Williams, R. W., and A. Berger. 1980. The relation of female polygamy to gonotrophic activity in the ROCK strain of *Aedes aegypti*. *Mosquito News* **40**:597–604.
- Young, A. D. M., and A. E. R. Downe. 1982. Renewal of sexual receptivity in mated female mosquitoes, *Aedes aegypti*. *Physiological Entomology* **7**:467–471.

Appendix A

The basic population genetic model of engineered underdominance

We consider introduction of engineered insects homozygous for two co-dependent engineered constructs inserted into two nonhomologous autosomes, a and b . The introduced genotype and the wild-type are denoted as $aabb$ and $AABB$, respectively. We assume that (i) generations are nonoverlapping; (ii) mating is random; (iii) sex ratio in F_1 and later generations is 1:1.

When the engineered individuals cross with the wild-type individuals, there are four types of mating in the F_0 generation: female $AABB \times$ male $AABB$, female $AABB \times$

$$(ab)' = \frac{[f_{22}(aabb) + f_{12}(1/2)(Aabb) + f_{21}(1/2)(aaBb) + f_{11}(1/4)(AaBb)]}{W}. \quad (A.8)$$

male $aabb$, female $aabb \times$ male $AABB$, female $aabb \times$ male $aabb$. These four types of matings lead to three genotypes in the F_1 generations: $AABB$, $AaBb$ and $aabb$ which are all viable. In the case of random mating the frequencies of the three viable genotypes in the F_1 generation, $(AABB)_1$, $(AaBb)_1$ and $(aabb)_1$, can be calculated by the following equations:

$$(AABB)_1 = \frac{f_{00}^2 [(AABB)_{m0} (AABB)_{f0}]}{W_0} \quad (A.1)$$

$$(AaBb)_1 = \frac{f_{00} f_{22} [(AABB)_{m0} (aabb)_{f0} + (aabb)_{m0} (AABB)_{f0}]}{W_0} \quad (A.2)$$

$$(aabb)_1 = \frac{f_{22}^2 [(aabb)_{m0} (aabb)_{f0}]}{W_0}. \quad (A.3)$$

Here,

$$W_0 = f_{00}^2 (AABB)_{m0} (AABB)_{f0} + f_{00} f_{22} (AABB)_{m0} (aabb)_{f0} + f_{22} f_{00} (aabb)_{m0} (AABB)_{f0} + f_{22}^2 (aabb)_{m0} (aabb)_{f0} \quad (A.4)$$

is the mean fitness. The subscripts $m0$ and $f0$ represent the genotype frequencies in males and females, respectively. $f_{00} = 1$ and $f_{22} = (1 - c)^2$ are the fitnesses of the genotypes $AABB$ and $aabb$, respectively. Note that because we assume a 1:1 sex ratio for offspring, in the F_1 and later generations genotype frequencies in males and females are the same. Therefore, we will drop the sex-specific subscripts from the genotype notation from F_1 generation.

In the F_2 and later generations there are nine genotypes: $AABB$, $AABb$, $AaBB$, $AaBb$, $AAbb$, $aaBB$, $Aabb$,

$aaBb$ and $aabb$. Among them only five are viable: $AABB$, $AaBb$, $Aabb$, $aaBb$ and $aabb$. The other four genotypes are not viable because they contain only construct a or b . In Mendelian segregation the frequencies of gametes (AB), (Ab), (aB) and (ab) in the next generation (represented by a prime) can be calculated by the frequencies of genotypes of the current (F_1 or later) generation as follows

$$(AB)' = \frac{[f_{00}(AABB) + f_{11}(1/4)(AaBb)]}{W} \quad (A.5)$$

$$(Ab)' = \frac{[f_{12}(1/2)(Aabb) + f_{11}(1/4)(AaBb)]}{W} \quad (A.6)$$

$$(aB)' = \frac{[f_{21}(1/2)(aaBb) + f_{11}(1/4)(AaBb)]}{W} \quad (A.7)$$

Here $f_{00} = 1$, $f_{11} = (1 - c/2)^2$, $f_{12} = f_{21} = (1 - c/2)(1 - c)$, $f_{22} = (1 - c)^2$ are the fitnesses of the five genotypes $AABB$, $AaBb$, $Aabb$, $aaBb$ and $aabb$, respectively. Note that the fitness costs of constructs a and b are assumed to be the same. W is the sum of all numerators on the right-hand side of Eqns (A.5)–(A.8). In the case of random mating, the relative frequencies of the five viable genotypes can be calculated based on the frequencies of gametes:

$$(AABB) = (AB)^2 \quad (A.9)$$

$$(AaBb) = 2(AB)(ab) + 2(Ab)(aB) \quad (A.10)$$

$$(Aabb) = 2(Ab)(ab) \quad (A.11)$$

$$(aaBb) = 2(ab)(aB) \quad (A.12)$$

$$(aabb) = (ab)^2. \quad (A.13)$$

Substitution of Eqns (A.9)–(A.13) into Eqns (A.5)–(A.8) eliminates the genotype frequencies and leaves a system of four equations for the gamete frequencies.

Because we consider the introduction of homozygous genotype $aabb$ to a wild population with genotype $AABB$, initially $(Ab) = (aB) = 0$. As long as this relation initially holds, it is easy to check that $(Ab) = (aB)$ always holds.

Using this symmetry between a and b and the condition $(AB) + (aB) + (Ab) + (ab) = 1$, this set of four equations reduces to just two

$$\bar{W}G'_{++} = (\hat{G}_{++} + \hat{G}_{M+}/2)(G_{++} + (1 - t_0)s_1G_{M+}/2) \quad (B.3)$$

$$(AB)' = \frac{(AB)^2 + f_{11}[(AB)(ab) + (Ab)^2]/2}{(AB)^2 + f_{22}(ab)^2 + 2f_{12}(Ab)(ab) + f_{11}[(AB)(ab) + (Ab)^2]} \quad (A.14)$$

$$(ab)' = \frac{f_{22}(ab)^2 + 2f_{12}(Ab)(ab) + f_{11}[(AB)(ab) + (Ab)^2]/2}{(AB)^2 + f_{22}(ab)^2 + 2f_{12}(Ab)(ab) + f_{11}[(AB)(ab) + (Ab)^2]} \quad (A.15)$$

Here $(Ab) = (aB) = \frac{1}{2}[1 - (AB) - (ab)]$. When $c = 0$, the model has three equilibria, two locally stable ones: $E_{s1}: (AB, ab) = (0, 1)$ (constructs are fixed) and $E_{s2}: (AB, ab) = (1, 0)$ (constructs are lost) and an unstable one $E_u: (AB, ab) \approx (0.63, 0.17)$. When $c > 0$ the equilibrium values at E_{s2} are the same as for $c=0$, but the equilibrium values at E_{s1} and E_u depend on c . Generally it is hard to derive analytical expressions for these two equilibria. Numerically we find that at E_{s1} , AB increases as c increases, but ab decreases as c increases. When $c = 0.25$ the equilibrium values $(AB, ab) \approx (0.02, 0.76)$ at E_{s1} .

All numerical results concerning the introduction thresholds in the nonage-structured model are based on numerical simulation of the model.

Appendix B

The basic population genetic model of *Medea*

Let G_{MM} , G_{M+} and G_{++} be the frequencies of females of genotypes MM , $M+$ and $++$, respectively. Let \hat{G}_{MM} , \hat{G}_{M+} and \hat{G}_{++} be the frequencies of males of genotypes MM , $M+$ and $++$, respectively. Based on Wade and Beeman (1994), the genotype frequencies in the offspring generation after selection (represented by a prime) are

$$\bar{W}G'_{MM} = c_2(\hat{G}_{MM} + \hat{G}_{M+}/2)(s_2G_{MM} + s_1G_{M+}/2) \quad (B.1)$$

$$\begin{aligned} \bar{W}G'_{M+} = c_1[& (\hat{G}_{MM} + \hat{G}_{M+}/2) + (G_{++} + s_1G_{M+}/2) \\ & + (\hat{G}_{++} + \hat{G}_{M+}/2)(s_2G_{MM} + s_1G_{M+}/2)] \end{aligned} \quad (B.2)$$

$$\begin{aligned} \bar{W} = c_2(\hat{G}_{MM} + \hat{G}_{M+}/2)(s_2G_{MM} + s_1G_{M+}/2) \\ + c_1[& (\hat{G}_{MM} + \hat{G}_{M+}/2)(G_{++} + s_1G_{M+}/2) \\ & + (\hat{G}_{++} + \hat{G}_{M+}/2)(s_2G_{MM} + s_1G_{M+}/2)] \\ & + (\hat{G}_{++} + \hat{G}_{M+}/2)(G_{++} + (1 - t_0)s_1G_{M+}/2) \end{aligned} \quad (B.4)$$

Here $c_1 = 1 - c/2$ and $c_2 = 1 - c$ are the embryonic fitnesses of $M+$ and MM , respectively. While $s_1 = 1 - s/2$ and $s_2 = 1 - s$ are the fecundity fitnesses of the $M+$ and MM adult females, respectively. t_0 is the mortality of wild-type offspring produced by females carrying the *Medea* allele(s). In our model, the maternal fecundity loss $s = 0.1$, and the maternal-effect mortality $t_0 = 1$ are fixed. c is allowed to vary from 0 to 0.25.

After the F_0 generation the sex ratio is assumed to be 1:1, so the genotype frequencies in males and females are the same. This means that we do not have to distinguish between \hat{G} and G . Noticing that $G_{MM} + G_{M+} + G_{++} = 1$ the set of three Eqns (B.1)–(B.3) can be reduced to only two equations.

When $s = c = 0$ we find that the model has two equilibria: an unstable one $E_0: (G_{MM}, G_{M+}, G_{++}) = (0, 0, 1)$, where the entire population is wild-type, and a stable one $E_1: (G_{MM}, G_{M+}, G_{++}) = (1, 0, 0)$, where the wild genotype is eliminated. When $s > 0$ and/or $c > 0$ the model has three equilibria: two stable ones E_0 and E_1 , and an unstable one E_u . The equilibrium E_0 is $(0, 0, 1)$ which is independent of s and c . The equilibrium E_1 has the form $(1 - q, q, 0)$ in which $0 < q < 1$ is a function of s and c . In general it is difficult to derive the expressions of q and the entries of E_u .