

Modeling the Manipulation of Natural Populations by the Mutagenic Chain Reaction

Robert L. Unckless,^{*,1} Philipp W. Messer,[†] Tim Connallon,[‡] and Andrew G. Clark^{*,†}

^{*}Department of Molecular Biology and Genetics and [†]Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York 14853, and [‡]School of Biological Sciences, Monash University, Clayton, 3800 Victoria, Australia

ORCID IDs: 0000-0001-8586-7137 (R.L.U.); 0000-0001-8453-9377 (P.W.M.); 0000-0002-1962-4951 (T.C.)

ABSTRACT The use of recombinant genetic technologies for population manipulation has mostly remained an abstract idea due to the lack of a suitable means to drive novel gene constructs to high frequency in populations. Recently Gantz and Bier showed that the use of CRISPR/Cas9 technology could provide an artificial drive mechanism, the so-called mutagenic chain reaction (MCR), which could lead to rapid fixation of even a deleterious introduced allele. We establish the near equivalence of this system to other gene drive models and review the results of simple models showing that, when there is a fitness cost to the MCR allele, an internal equilibrium may exist that is usually unstable. In this case, introductions must be at a frequency above this critical point for the successful invasion of the MCR allele. We obtain estimates of fixation and invasion probabilities for the appropriate scenarios. Finally, we discuss how polymorphism in natural populations may introduce sources of natural resistance to MCR invasion. These modeling results have important implications for application of MCR in natural populations.

KEYWORDS mutagenic chain reaction; whole population replacement; CRISPR/Cas9; gene drive

THE prospect of introducing a novel gene into a population and having it spread to high frequency holds great promise for biological control. Such technologies could provide a means of control of highly pesticide-resistant crop pests and disease vectors (Bourtzis and Hendrichs 2014). But the possibility of uncontrolled spread of this artificial genetic material once introduced drives a need to thoroughly understand the population dynamics and behavior of these artificial drive systems (Bohannon 2015). With few exceptions (Hoffmann *et al.* 2011), the practicality of such introductions has been limited by the lack of a means to ensure the spread of the engineered genetic material through a population. In a recent article, Gantz and Bier (2015) describe the mutagenic chain reaction (MCR), an approach that employs the CRISPR/Cas9 system to drive a mutation to high frequency in a population, making gene replacement at the population level practical for any species that can be made to accept a transgene in the laboratory.

There is, in fact, an extensive literature on “gene drive” systems that can transform entire populations (reviewed in Sinkins and Gould 2006, Gould 2008, and Burt 2014 and going back to Curtis 1968 and Foster *et al.* 1972). The work of Burt and colleagues (Burt 2003; Deredec *et al.* 2008; North *et al.* 2013) considering the population genetics of homing endonucleases as a means to transform entire populations is particularly relevant, and much of what we describe below is a specific application of the general principles developed earlier and applied to MCR. Because Gantz and Bier’s CRISPR/Cas9-mediated method is unique at the molecular level and because of the interest received by their recent publication (Gantz and Bier 2015), we use their term “mutagenic chain reaction,” even though formally it is a special case of what is widely referred to as a gene drive system.

We envision three classes of introductions. First, the introduced genetic material could be beneficial to the organism. Imagine a mutation that confers resistance to *Plasmodium* in *Anopheles* mosquitoes without pleiotropic cost or allows an endangered species to tolerate a new environmental stress. Second, the mutation could be neutral in a standard environment. In this case, a mutation that confers insecticide susceptibility could be introduced into an agricultural pest population without the use of that particular pesticide; then when the

Copyright © 2015 by the Genetics Society of America

doi: 10.1534/genetics.115.177592

Manuscript received May 4, 2015; accepted for publication July 30, 2015; published Early Online July 30, 2015.

Supporting information is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.177592/-/DC1.

¹Corresponding author: 227 Biotechnology Bldg., Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853. E-mail: unckless@cornell.edu

mutation is near fixation, the pesticide could be applied, thereby reducing the population severely. Finally, a mutation might be deleterious regardless of environment. This last scenario is similar to sterile male techniques that have been discussed for decades (Foster *et al.* 1972; Prout 1978; Bourtzis and Hendrichs 2014). For example, a life-shortening mutation in mosquitoes that does not allow for a complete incubation period for Dengue virus could severely reduce transmission rates. In this case, the rate of conversion via the MCR process must outweigh the fitness cost to the organism. Here we model all three scenarios, find an internal equilibrium for those mutations with fitness costs, and interpret these results in terms of practical applications.

Modeling MCR

We employ a Wright–Fisher model with a panmictic, random-mating population of infinite size (which we later relax). The mutagenic chain reaction occurs in heterozygotes when the autonomous MCR construct replaces the wild-type allele with a copy of itself (see figure 1 in Gantz and Bier 2015). We set wild-type fitness to 1 and assume the fitness of individuals homozygous for the drive construct have fitness $1 - s$ and those heterozygous for the drive construct have fitness $1 - hs$, where s is the selection coefficient and h reflects the degree of dominance. Fitness effects are assumed to be associated with viability. The wild-type allele in heterozygous individuals is converted to the MCR allele at rate c . We assume this happens in the embryo, a departure from the meiotic drive models, so that individuals that experienced conversion essentially become homozygous for the MCR construct and have fitness $1 - s$. Considering all these processes, we can write the recursion that gives the frequency of the MCR allele in the following generation, given its current frequency q :

$$q' = \frac{q^2(1-s) + q(1-q)[(1-c)(1-hs) + 2c(1-s)]}{\bar{w}}. \quad (1)$$

The terms in the numerator represent individuals homozygous for the MCR allele and their relative fitness $q^2(1-s)$, unconverted heterozygotes and their fitness $q(1-q)(1-c)(1-hs)$, and converted heterozygotes and their fitness $2q(1-q)c(1-s)$, respectively. The denominator, \bar{w} , is the mean fitness that serves to normalize the recursion so that the allele frequencies sum to one. In this case, we have $\bar{w} = q^2(1-s) + 2q(1-q)(1-c)(1-hs) + 2q(1-q)c(1-s) + (1-q)^2$. The fate of the MCR allele is largely determined by its ability to spread when rare, such that $q^2 \approx 0$ and $\bar{w} \approx 1$. In this case, Equation 1 becomes

$$q' \approx q[1 + c - cs(2-h) - hs] = q(1 + s_e), \quad (2)$$

resembling standard exponential growth with an effective selection coefficient $s_e = hs(c-1) + (1-2s)c$. Note that this means the MCR allele can spread when rare even if it is deleterious, as long as $c > hs/(1-2s+hs)$.

The essential feature of MCR with respect to population frequency dynamics is that heterozygotes produce an excess of gametes bearing the MCR construct. Meiotic drive, homing endonucleases, and gene conversion are three distinct mechanisms that also produce such non-Mendelian gametic counts. The population genetic consequences of meiotic drive (Prout 1953) and of gene conversion (Gutz and Leslie 1976) have both been shown to result in rapid fixation. When balanced by opposing natural selection, both mechanisms can produce internal equilibria (Hiraizumi *et al.* 1960; Walsh 1983). In fact, the model for which MCR is a special case was given by Hartl (1970), and here we reparameterize this model to make clear how it relates to the biological process of MCR and its relevance to biological control. While most of the interesting modeling occurs when s is positive (the allele is costly), our model does allow for the introduction of a beneficial ($s < 0$) or neutral ($s = 0$) allele.

To get some sense of the behavior of this model, we show plots of the predicted frequency of the MCR construct, starting at an initial frequency of 0.001, and a range of selection coefficients and conversion rates (Figure 1). Note that the rate of spread is impressively rapid, even in cases where the MCR construct confers a 25% reduction in fitness, and with a conversion efficiency of only 80%. These simulation results certainly underscore the potential hazard of allowing MCR to run uncontrolled in natural populations.

This model can admit a single internal equilibrium,

$$\hat{q} = \frac{c + cs(h-2) - hs}{s(1-2c-2h+2ch)}, \quad (3)$$

where $q' = q$ and allele frequencies do not change. The stability of this internal equilibrium depends on the values of the three parameters of the model (Supporting Information, Figure S1 and Figure S2, and Appendix A). Briefly, if fitness costs of the MCR allele are dominant ($h = 1$), then the internal equilibrium is unstable. For MCR to successfully invade, there is a constraint on the fitness cost $s < c/(1+c)$. If fitness costs are recessive, the internal equilibrium is stable if $c < 1/2$ and unstable if $c > 1/2$. For additive fitness, the internal equilibrium is always unstable (Figure 2). The role of an unstable equilibrium for population transformation is discussed in detail elsewhere (Altrock *et al.* 2010).

In the case of biological control applications, the goal will be to attain fixation of the MCR allele, and so the construct must be introduced into the wild population at sufficient frequency to exceed the unstable equilibrium frequency. This is reminiscent of the spread of *Wolbachia* through populations (Turelli and Hoffmann 1995). In both cases, insufficient initial frequency will result in loss from the population. From the point of view of biosafety, we would want any escaped version of the MCR to have an elevated unstable equilibrium, so that the population would be able to exceed this threshold only with a concerted effort by researchers consciously working toward that goal. Figure 2 shows the location of this critical equilibrium frequency and shows

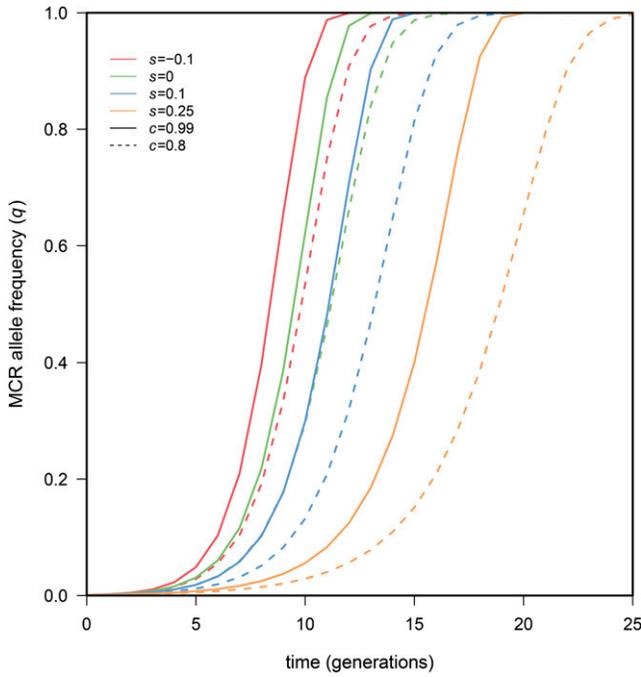


Figure 1 Trajectories of introduced MCR alleles reveal that even deleterious alleles sweep to fixation very quickly. Only parameter sets leading to fixation are presented, and all cases shown assume that fitness costs are recessive ($h = 0$).

that the equilibrium frequency is elevated when fitness costs are high and when conversion rates to the MCR allele are relatively low. One possible route to increase the safety of an introduction would be to engineer means to ensure low conversion frequency and high fitness cost, at the same time yielding an equilibrium frequency that is practical in terms of ability to produce and release the engineered organisms. Of course, one should expect that all parameters in the system would be free to evolve once released into a natural population, necessitating other safeguards (see *Discussion*).

Of particular interest is the likelihood that an MCR allele would escape stochastic loss and sweep to fixation or to a stable internal equilibrium through either purposeful or accidental introduction. This scenario might be most relevant when considering invasion of the MCR construct into a nontarget species, where initial frequency may be quite low. This type of modeling requires the inclusion of genetic drift and models are therefore stochastic. In cases involving a stable internal equilibrium and those where the MCR allele sweeps to fixation without equilibrium, the question is whether the allele can escape stochastic loss when rare. As discussed above (Equation 2), in most of the parameter space the fate of the MCR allele is determined when it is quite rare and we can ignore terms involving q^2 . Thus, we can approximate selection, using s_e . In this case, the probability of invasion (from Kimura 1962) of the MCR allele can be approximated as

$$\pi(s_e, q_0, N_e) \approx 1 - \exp(-4N_e s_e q_0), \quad (4)$$

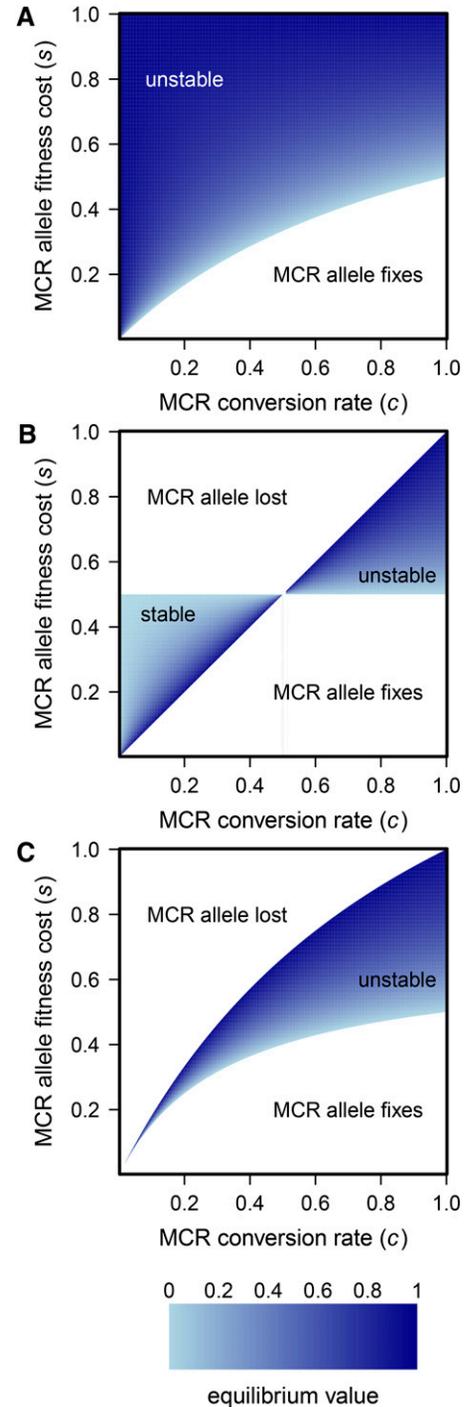


Figure 2 (A–C) Equilibrium frequency of the MCR allele when fitness costs are (A) dominant ($h = 1$), (B) recessive ($h = 0$), and (C) additive ($h = 1/2$). Parameter space leading to fixation, loss, or stable or unstable equilibrium of the MCR allele is shown.

where N_e is the variance effective population size and q_0 is the starting frequency of the MCR allele. The above approximation performs well against simulation, as long as q_0 is small, although the approximation is less accurate as c becomes large (Figure S2 and Figure S3). When a single copy of the mutant is introduced and s_e is small ($s_e < 0.1$), the probability in

Equation 4 can be approximated using the standard branching-process approach (Haldane 1927): $\pi(s_e) \approx 2s_e$. When fitness costs are recessive, this reduces to $\pi(s_e) \approx 2c(1 - 2s)$, proportional to the conversion rate discounted by twice the selection coefficient.

Furthermore, the minimum frequency at which the MCR allele needs to be introduced into the population to escape stochastic loss is $q_{\text{critical}} > 1/(2N_e s_e)$. This means that in the Wright–Fisher scenario, the MCR allele must be introduced in $2N_e q_{\text{critical}} = (1 + s_e)/s_e$ individuals to be relatively certain that it will escape stochastic loss.

When the internal equilibrium is unstable, it is useful to determine the probability that the MCR allele becomes fixed if introduced at or near that unstable point. This requires a more sophisticated approach (see *Appendix B*) (Robertson 1962; Nei and Roychoudhury 1973; Jansen *et al.* 2008). Simulations confirm that this approximation (Equation B5) works quite well (*Figure S4*).

When parameters allow for deterministic fixation of the MCR allele, our approximation for the effective selection coefficient also allows for an estimate of the time it takes for the allele to become prevalent in the population. For example, assuming exponential growth at rate s_e , the MCR allele is expected to reach a frequency of 1/2 after

$$\sim t_{1/2} = \frac{1}{2s_e} \ln \frac{N_e}{q_0} \quad (5)$$

generations (based on a logistic approximation).

Discussion

Genetic manipulation of wild populations is a potentially effective approach for ameliorating the impact of pathogens, disease vectors, and agricultural pests. It also has the potential to do irreparable harm through accidental or purposeful release. An extensive literature exists in this area (Sinkins and Gould 2006; Gould 2008; Burt 2014) and one aim in this work was to merge the recent experimental work with modeling efforts that have been in existence for years. We analyzed a simple model of gene drive or MCR and found that it has the potential to spread very rapidly, but may be slowed by the presence of a stable or unstable internal equilibrium.

We focus on the simplest case and note that several factors not modeled here could alter the spread of an MCR allele. First, the MCR allele could experience a mutation rendering it unable, by itself, to convert the homolog to the MCR state. This could severely limit the spread of the MCR allele when it carries fitness costs ($s > 0$). In fact, this could even be used as a method for controlling the MCR allele. With three alleles: wild type, active MCR, and broken MCR, we imagine that the active MCR converts the wild type but cannot reconvert the broken MCR allele and that the broken MCR allele has higher fitness than the active MCR allele (it could be engineered that way). In this case, the introduction of the broken MCR allele should spread since it is both resistant to conversion from the active MCR

allele and will have higher fitness than the MCR allele. Several proposed means of controlling gene drive take advantage of some of these ideas (see Esvelt *et al.* 2014; Oye *et al.* 2014).

Another possibility is that genetic variation for conversion susceptibility may segregate in populations. It is likely that the target site of the guide RNA would be variable in natural populations. In this case, some naturally occurring wild-type alleles would both be resistant to MCR insertion and not carry the engineered cost of the construct and prevent invasion of the MCR allele (reviewed in Esvelt *et al.* 2014). Population geneticists have modeled this situation with a second locus that serves to modify the strength of meiotic drive at the driven locus (Prout *et al.* 1973; Huang *et al.* 2007; Pauwels *et al.* 2014). The conclusion from this analysis is that modifier loci can easily select alleles that affect the population dynamics, and when they occur in stable equilibria, they generally are in linkage disequilibrium with the driven locus. The very rapid fixation of MCR may preclude the invasion of such modifiers, but in any case there is already machinery for analysis of such modifiers (Rasgon and Gould 2005; Magori *et al.* 2009).

These results show that the MCR may provide an effective means for population replacement and that the speed of the process presents reason for considerable caution before considering a field release of such a construct. In fact, there are conditions in which accidental introductions of a single individual can lead to fixation of the MCR allele even with significant deleterious fitness consequences to the individual. For example, an allele with perfect conversion efficiency and a selective cost of 0.41 is still expected to escape stochastic loss and sweep to fixation nearly 30% of the time in a Wright–Fisher population. When parameters allow for an unstable equilibrium, the MCR construct would spread in a population only if that unstable equilibrium point is sufficiently small (*e.g.*, Barton and Turelli 2011). Thorough quantitative modeling of MCR population dynamics is strongly warranted, not only to put bounds on the frequency trajectories expected from release of an MCR, but also for possible choke points for controlling and preventing the expansion of an escaped or mutated MCR allele in a natural population.

Several additional lines of inquiry will be the subject of future work. For example, in a structured population a balance between conversion and migration might facilitate the spread of the MCR allele. Also, one potential application of MCR is to purposefully drive pest populations to extinction. In this case it would be useful to predict the expected time to extinction and the likelihood that the population can rescue itself before extinction. A population might be rescued by an allele in the standing genetic variation or a new mutation that is resistant to conversion (Orr and Unckless 2014).

Acknowledgments

We thank Brian Lazzaro, Tony Long, Graham Coop, Ethan Bier, Kevin Esvelt, and two anonymous reviewers for thoughtful criticisms and comments on an earlier version of this

manuscript. This work was supported by National Institutes of Health (NIH) grant R01-GM064590 (to A.G.C.), NIH grant F32-HD071703 (to R.L.U.), and an Australian Research Council discovery grant (to T.C.).

Literature Cited

- Altrock, P. M., A. Traulsen, R. G. Reeves, and F. A. Reed, 2010 Using underdominance to bi-stably transform local populations. *J. Theor. Biol.* 267: 62–75.
- Barton, N. H., and M. Turelli, 2011 Spatial waves of advance with bistable dynamics: cytoplasmic and genetic analogues of allele effects. *Am. Nat.* 178: E48–E75.
- Bohannon, J., 2015 Biologists devise invasion plan for mutations. *Science* 347: 1300.
- Bourtzis, K., and J. Hendrichs, 2014 Preface: development and evaluation of improved strains of insect pests for sterile insect technique (sit) applications. *BMC Genet.* 15(Suppl 2): 11.
- Burt, A., 2003 Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proc. Biol. Sci.* 270: 921–928.
- Burt, A., 2014 Heritable strategies for controlling insect vectors of disease. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369: 20130432.
- Curtis, C. F., 1968 Possible use of translocations to fix desirable genes in insect pest populations. *Nature* 218: 368–369.
- Deredec, A., A. Burt, and H. C. J. Godfray, 2008 The population genetics of using homing endonuclease genes in vector and pest management. *Genetics* 179: 2013–2026.
- Esvelt, K. M., A. L. Smidler, F. Catteruccia, and G. M. Church, 2014 Concerning rna-guided gene drives for the alteration of wild populations. *eLife* DOI: <http://dx.doi.org/10.7554/eLife.03401>.
- Foster, G. G., M. J. Whitten, T. Prout, and R. Gill, 1972 Chromosome rearrangements for the control of insect pests. *Science* 176: 875–80.
- Gantz, V. M., and E. Bier, 2015 The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science* 348: 442–444.
- Gould, F., 2008 Broadening the application of evolutionarily based genetic pest management. *Evolution* 62: 500–510.
- Gutz, H., and J. F. Leslie, 1976 Gene conversion: a hitherto overlooked parameter in population genetics. *Genetics* 83: 861–866.
- Haldane, J. B. S., 1927 A mathematical theory of natural and artificial selection, part v: selection and mutation. *Math. Proc. Camb. Philos. Soc.* 23: 838–844.
- Hartl, D. L., 1970 Analysis of a general population genetic model of meiotic drive. *Evolution* 24: 538–545.
- Hiraizumi, Y., L. Sandler, and J. F. Crow, 1960 Meiotic drive in natural populations of *Drosophila melanogaster*. iii. Populational implications of the segregation-distorter locus. *Evolution* 14: 433–444.
- Hoffmann, A. A., B. L. Montgomery, J. Popovici, I. Iturbe-Ormaetxe, P. H. Johnson *et al.*, 2011 Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. *Nature* 476: 454–457.
- 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.
- Jansen, V. A., M. Turelli, and H. C. J. Godfray, 2008 Stochastic spread of *Wolbachia*. *Proc. R. Soc. Lond. B Biol. Sci.* 275: 2769–2776.
- Kimura, M., 1962 On the probability of fixation of mutant genes in a population. *Genetics* 47: 713–719.
- Magori, K., M. Legros, M. E. Puente, D. A. Focks, T. W. Scott *et al.*, 2009 Skeeter buster: a stochastic, spatially explicit modeling tool for studying *Aedes aegypti* population replacement and population suppression strategies. *PLoS Negl. Trop. Dis.* 3: e508–e508.
- Nei, M., and A. K. Roychoudhury, 1973 Probability of fixation and mean fixation time of an overdominant mutation. *Genetics* 74: 371–380.
- North, A., A. Burt, and H. C. J. Godfray, 2013 Modelling the spatial spread of a homing endonuclease gene in a mosquito population. *J. Appl. Ecol.* 50: 1216–1225.
- Orr, H. A., and R. L. Unckless, 2014 The population genetics of evolutionary rescue. *PLoS Genet.* 10: e1004551.
- Oye, K. A., K. Esvelt, E. Appleton, F. Catteruccia, G. Church *et al.*, 2014 Regulating gene drives. *Science* 345: 626–628.
- Pauwels, K., N. Podevin, D. Breyer, D. Carroll, and P. Herman, 2014 Engineering nucleases for gene targeting: safety and regulatory considerations. *N. Biotechnol.* 31: 18–27.
- Prout, T., 1953 Some effects of variations in the segregation ratio and of selection on the frequency of alleles under random mating. *Acta Genet. Stat. Med.* 4: 148–151.
- Prout, T., 1978 The joint effects of the release of sterile males and immigration of fertilized females on a density regulated population. *Theor. Popul. Biol.* 13: 40–71.
- Prout, T., J. Bundgaard, and S. Bryant, 1973 Population genetics of modifiers of meiotic drive i. The solution of a special case and some general implications. *Theor. Popul. Biol.* 4: 446–465.
- Rasgon, J., and F. Gould, 2005 Transposable element insertion location bias and the dynamics of gene drive in mosquito populations. *Insect Mol. Biol.* 14: 493–500.
- Robertson, A., 1962 Selection for heterozygotes in small populations. *Genetics* 47: 1291–1300.
- Sinkins, S. P., and F. Gould, 2006 Gene drive systems for insect disease vectors *Nat. Rev. Genet.* 7: 427–435.
- Turelli, M., and A. A. Hoffmann, 1995 Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. *Genetics* 140: 1319–1338.
- Walsh, J. B., 1983 Role of biased gene conversion in one-locus neutral theory and genome evolution. *Genetics* 105: 461–468.

Communicating editor: N. H. Barton

Appendix A

To determine the stability of the equilibrium in Equation 2, we first write an expression for the change in MCR allele frequency in a single generation:

$$\begin{aligned}\Delta q &= q' - q \\ &= \frac{-q(1-q)[c - s(c(2q(h-1) - h + 2) - 2hq + h + q)]}{qs(2c(h-1)(q-1) - 2h(q-1) + q) - 1}.\end{aligned}\quad (\text{A1})$$

At equilibrium (\hat{q} , Equation 2), the eigenvalue for Equation A1 is

$$\begin{aligned}\lambda &= \frac{d}{dq}(\Delta q)\Big|_{q=\hat{q}} \\ &= \frac{[c(h-2)s + c - hs][c(hs-1) - hs + s]}{c^2 - (c-1)hs^2[c(h-2) - h] - 2(c-1)s(c+h) - s}.\end{aligned}\quad (\text{A2})$$

For an MCR allele with dominant fitness costs ($h = 1$), the equilibrium in Equation 2 falls between zero and one only if $s > c/(1+c)$; otherwise the MCR allele always fixes. However, when $s > c/(1+c)$, the eigenvalue in Equation A2 is positive, and therefore the equilibrium is unstable. This also means that MCR alleles with a dominant fitness cost $> s = 0.5$ cannot successfully invade, since no interior equilibrium exists, and the equilibrium at $q = 1$ is unstable.

If the MCR allele carries recessive fitness costs ($h = 0$), the equilibrium in Equation 2 is between zero and one, if one of two sets of conditions is met. First, if $c < 1/2$, the following must be true: $c < s < 1/2$. Alternatively, if $c > 1/2$, then we must have $1/2 < s < c$. The internal equilibrium is stable when $c < 1/2$, because for the eigenvalue in Equation 2 to be positive, either $s > 1/2$ or $c > s$, both conditions would not permit an internal equilibrium. If $c > 1/2$, then the internal equilibrium is always unstable.

Finally, in the case of additive fitness ($h = 1/2$), an internal equilibrium exists if $2c/(1+3c) < s < 2c/(1+c)$, and this equilibrium is always unstable. The general case, where h is a free parameter between zero and one, is solvable but complex and therefore not presented here.

Appendix B

To find the probability that an MCR allele introduced near its unstable equilibrium point escapes stochastic loss and sweeps to fixation, we follow the approach of Jansen *et al.* (2008). We begin by rewriting our expression for the change in frequency per generation as

$$\Delta q = \frac{s[1 - 2h - 2c(1-h)]q(1-q)(\hat{q} - q)}{\bar{w}}, \quad (\text{B1})$$

where \hat{q} and \bar{w} are as defined above. The diffusion machinery requires the average change in frequency ($M = \Delta q$) and sample variance ($V = q(1-q)/2N_e$). Next we find a linear approximation of M/V in the vicinity of \hat{q} ,

$$M/V = \frac{2N_e}{q(1-q)} \frac{s[1 - 2h - 2c(1-h)]q(1-q)(\hat{q} - q)}{\bar{w}} \approx (\hat{q} - q)\Psi, \quad (\text{B2})$$

where $\Psi = 2N_e s[1 - 2h - 2c(1-h)]/(\hat{q}^2(1-s) + 2\hat{q}(1-\hat{q})(1-c)(1-hs) + 2\hat{q}(1-\hat{q})c(1-s) + (1-\hat{q})^2)$. The general diffusion approximation for fixation probability is

$$\prod(q_0) = \frac{\int_0^{q_0} G(x)dx}{\int_0^1 G(x)dx}, \quad (\text{B3})$$

where

$$G(x) = \exp\left(-2 \int_0^x \frac{M}{V} dx\right) \approx \exp(-\Psi x(x - 2\hat{q})). \quad (\text{B4})$$

Plugging Equation B4 into Equation B3, we get an approximation for the probability of fixation given an unstable equilibrium,

$$\Pi(q_0) \approx \frac{\int_0^{q_0} \exp(-\Psi x(x - 2\hat{q})) dx}{\int_0^1 \exp(-\Psi x(x - 2\hat{q})) dx} = \frac{\operatorname{erf}(\hat{q}\sqrt{\Psi}) - \operatorname{erf}((\hat{q} - q_0)\sqrt{\Psi})}{\operatorname{erf}(\hat{q}\sqrt{\Psi}) - \operatorname{erf}((\hat{q} - 1)\sqrt{\Psi})}, \quad (\text{B5})$$

where $\operatorname{erf}(z) = (2/\sqrt{\pi}) \int_0^z \exp(-t^2) dt$ is the error function of z .

GENETICS

Supporting Information

www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.177592/-/DC1

Modeling the Manipulation of Natural Populations by the Mutagenic Chain Reaction

Robert L. Unckless, Philipp W. Messer, Tim Connallon, and Andrew G. Clark

Supplemental figures

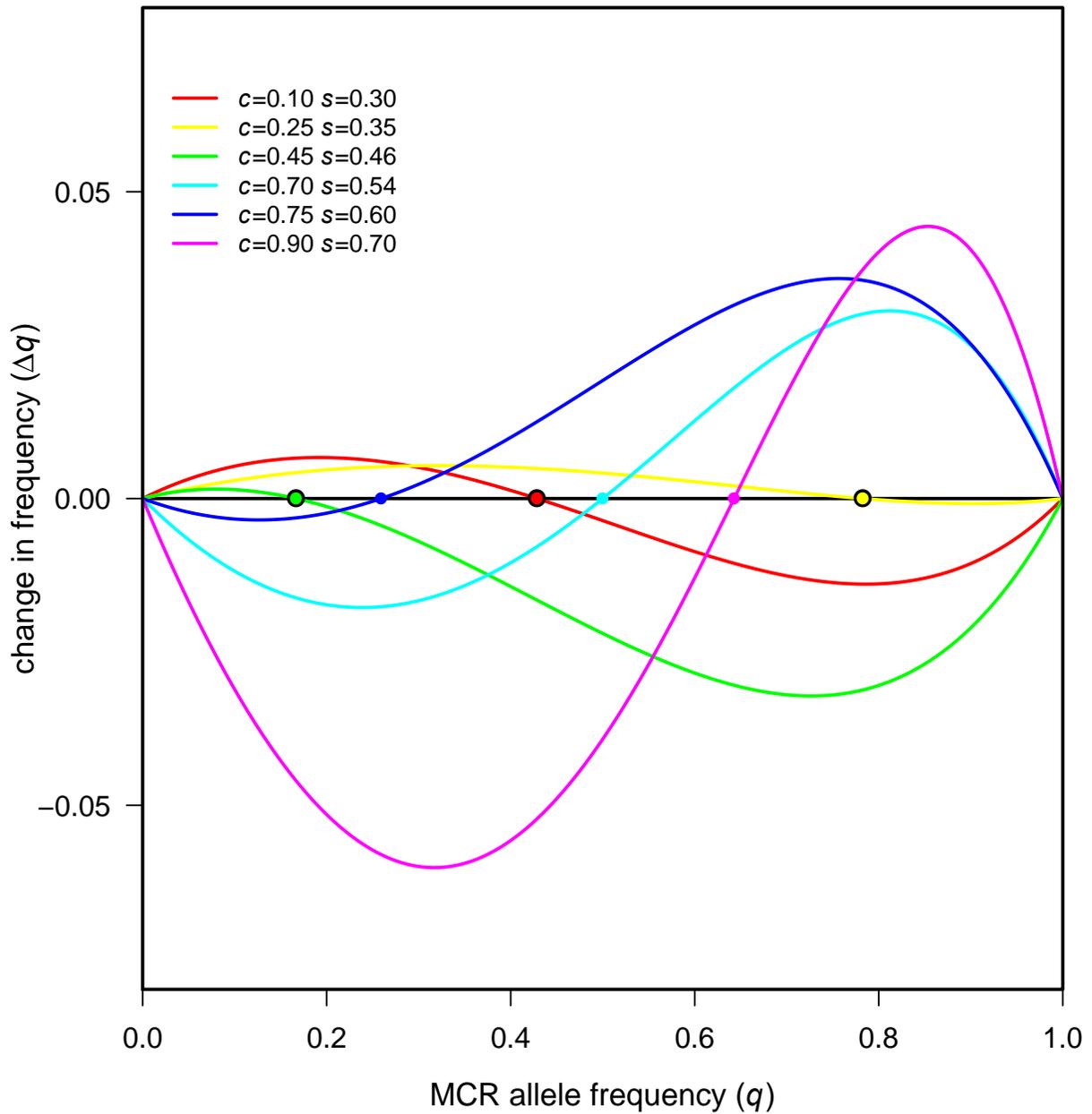


Figure S1 The change in MCR allele frequency (Δq) near equilibrium reveals the local stability of the equilibrium. Positive slopes are associated with unstable equilibria and negative slopes are associated with stable equilibria (black borders on circles). Fitness costs are recessive ($h = 0$).

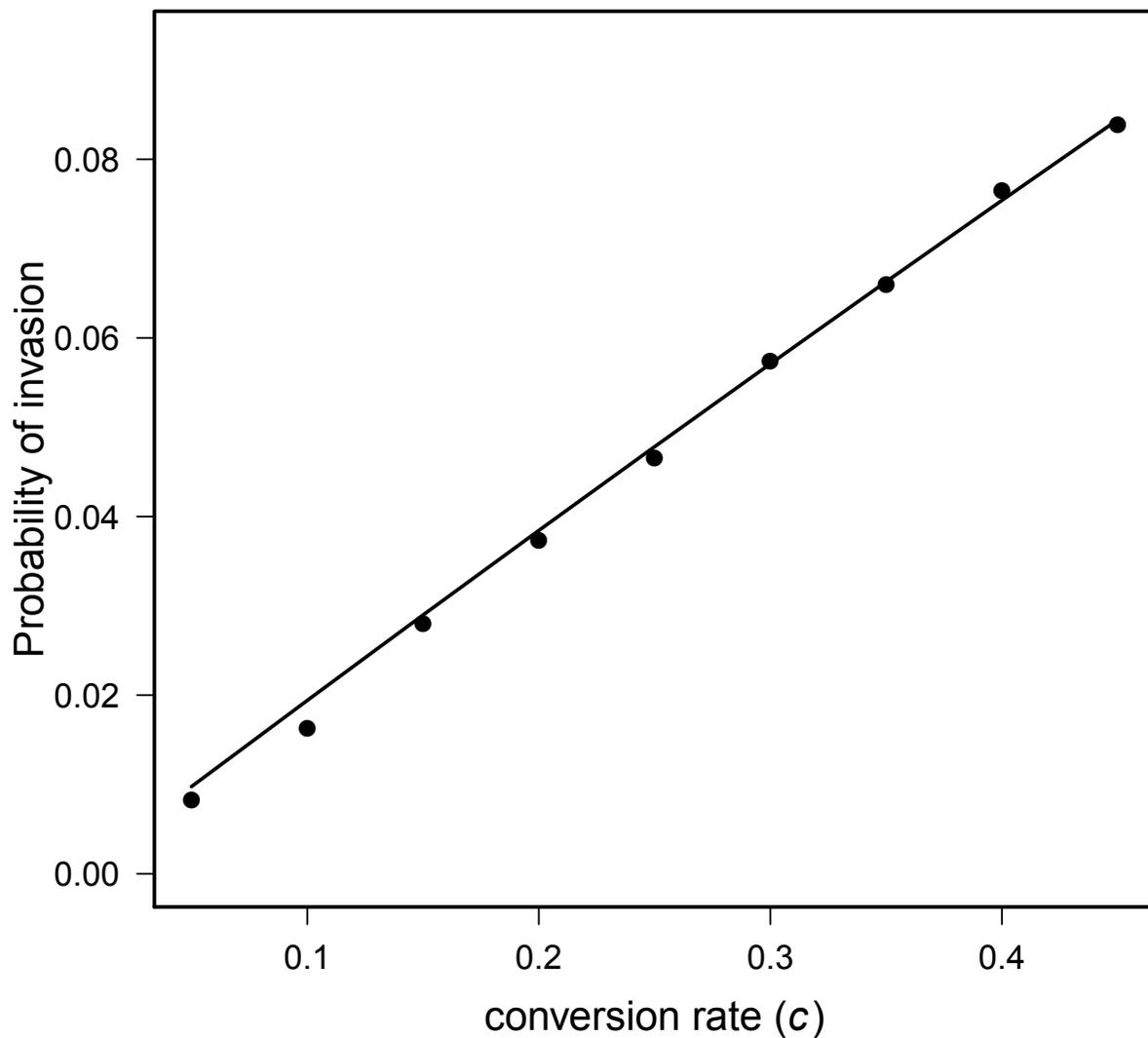


Figure S2 The probability that an MCR allele escapes stochastic loss when rare, when parameters allow for a stable internal equilibrium. Dark line represents empirical approximation (equation 4) and points represent the proportion of 100,000 simulation realizations resulting in invasion. The starting frequency of the MCR allele in each realization is $q_0 = \frac{1}{2N_e}$, the population size is $N_e = 10,000$, the fitness cost is $s = 0.451$ and fitness costs are recessive ($h = 0$).

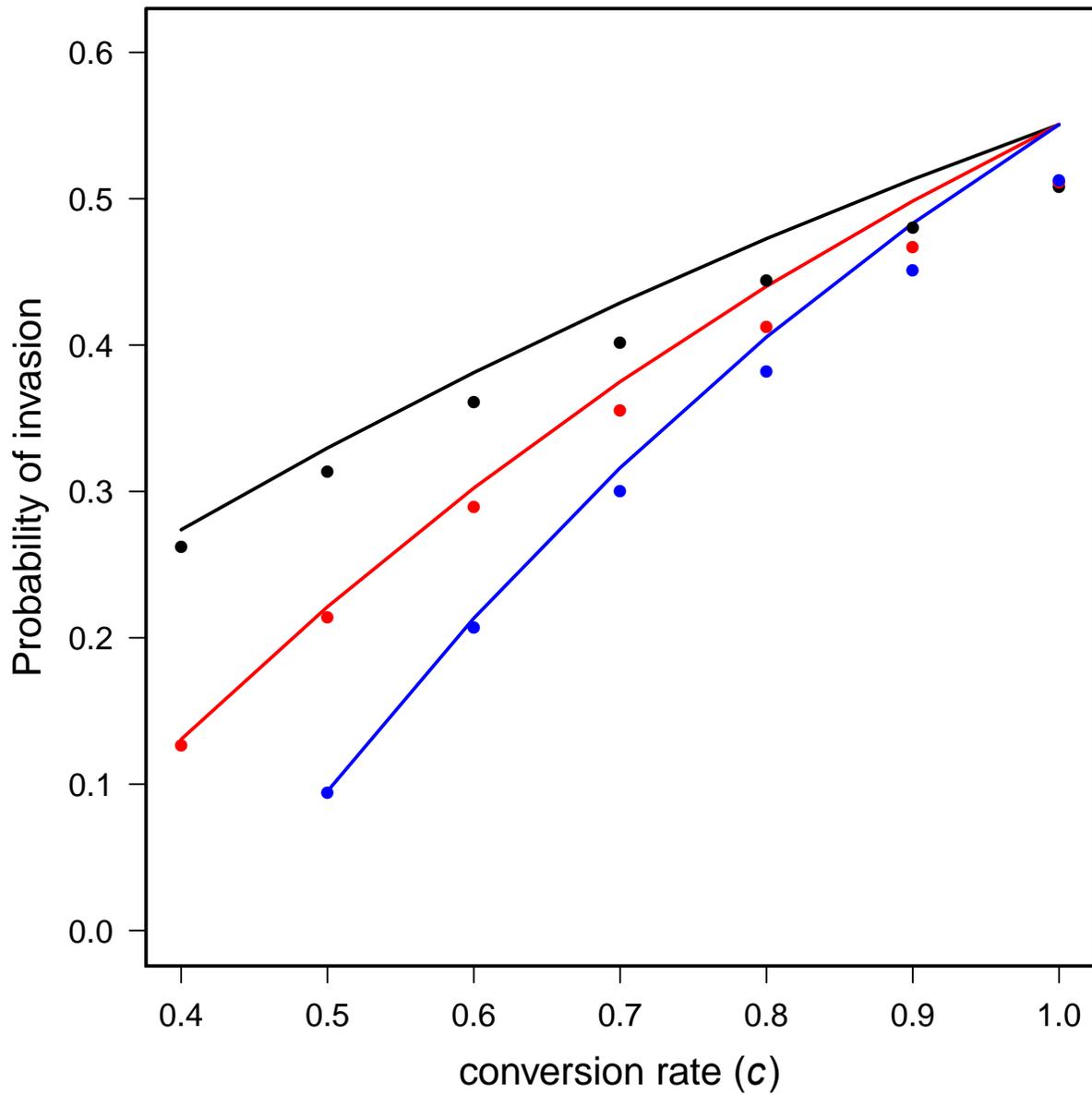


Figure S3 The probability that an MCR allele escapes stochastic loss when rare and becomes fixed when parameters allow for fixation without internal equilibrium. Lines represent the empirical approximation (equation 4) and points represent the proportion of 100,000 simulation realizations resulting in invasion. The starting frequency of the MCR allele in each realization is $q_0 = \frac{1}{2N_e}$, the population size is $N_e = 10,000$, the fitness cost is $s = 0.4$ and blue represents dominant ($h = 1$) fitness costs, red represents additive ($h = 0.5$) fitness costs and black represents recessive ($h = 0$) fitness costs.

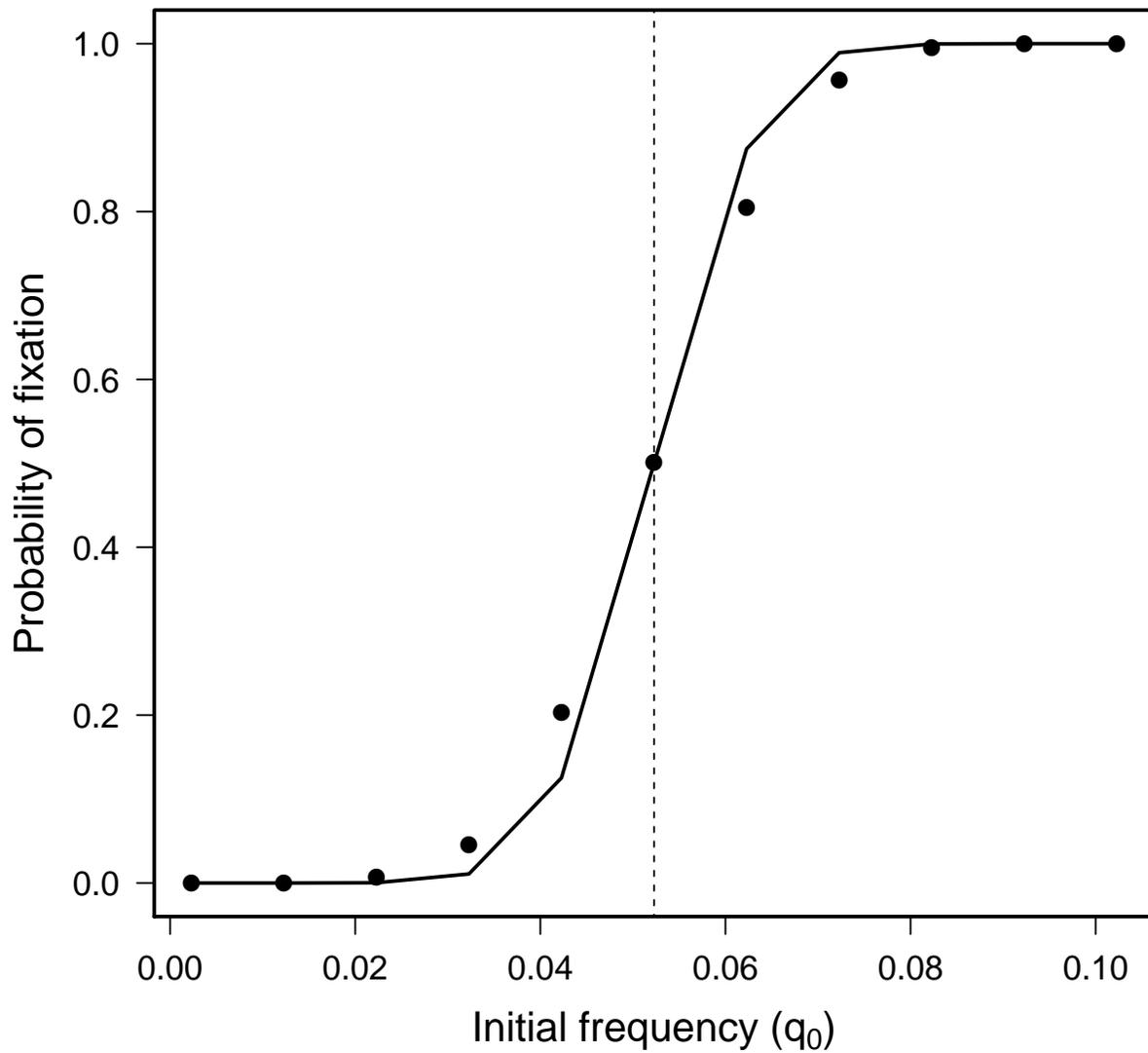


Figure S4 The probability that an MCR allele escapes stochastic loss when introduced near its equilibrium and becomes fixed when parameters allow for an unstable internal equilibrium. Line represents the empirical approximation (equation A2.5) and points represent the proportion of 100,000 simulation realizations resulting in invasion. The conversion efficiency is $c = 0.8$, the population size is $N_e = 10,000$, the fitness cost is $s = 0.51$ and fitness effects are recessive ($h = 0$). Dashed line represents the unstable equilibrium frequency.