ABSTRACT

VELLA, MICHAEL RICHARD. Modeling Strategies for Insect Pest Management: Genetic Analyses and Parameter Estimation. (Under the direction of Alun Lloyd and Fred Gould.)

Insect pests cause substantial annual damages in human health as vectors of disease and in agriculture. It has proven challenging to mitigate the damages by reducing insect population size or making the species unable to transmit disease. One of the traditional approaches to reducing population size is using chemical insecticides, but insect resistance to chemicals requires limited use of each insecticide or switching to different chemicals. A variety of genetic strategies exist that involve releasing transgenic strains of a species. Some strategies rely on inundating the population with transgenics that lower the average fitness of the population, sometimes to the point of eliminating the population. An example is the use of female-killing strains, which cause death of the females that inherit the inserted gene. Another set of transgenic strains are examples of gene drives, where the frequency of the transgene would automatically increase after release. Gene drives can spread a gene that lowers fitness and suppresses population size, or a gene that prevents the species from being vectors of disease. In this dissertation, I use mathematical models to analyze several of these approaches.

In Chapter 2, I explore the evolution of a relatively new class of gene drives and several proposed countermeasures that would slow the spread or remove the first drive altogether. These CRISPR-Cas9-based gene drives have the theoretical potential to spread without limit. For this reason, many countermeasures have been theorized and/or developed to halt or reverse a drive if needed, such as an additional gene drive that biases inheritance over the first drive. I develop a discrete-time, allele-frequency model to explore the population genetics of the wild-type, the gene drive, and the countermeasure alleles. I find surprising oscillatory behavior under certain parameter values and that different countermeasures may be best to use in different scenarios.

In Chapter 3, I evaluate the effectiveness of several genetic approaches designed to suppress the population with repeated release. My focus is to compare female-killing strains that have the killing mechanism at a single locus in the genome and those that are spread across two loci. Each of these approaches has been built, but previous modeling efforts have only considered female-killing at one locus. I use a genotype-frequency, differential equation model with age structure

and density-dependent mortality to compare the effectiveness of the approaches. Even though the genetic elements are inherited independently if spread across two loci, the ability to suppress the population is similar compared to the one-locus approach.

In Chapter 4, I analyze data showing the development of resistance to insecticides in mosquitoes in Iquitos, Peru. The resistant allele increased in frequency over the course of many years while there was spraying of insecticides. Of interest are the parameters for the strength of selection against mosquitoes with a single copy and two copies of the susceptible allele. I use a discrete-time, genotype-frequency model as the basis for Bayesian inference. Results from particle Markov chain Monte Carlo suggest that the cost to susceptibles is recessive or additive, meaning that a single resistant allele is enough to substantially lessen the negative effects from the insecticide.

Throughout this dissertation, I show the utility of simple models. While population genetics can sometimes seem comprehendable using reasoning alone, models can illuminate surprising and unintuitive outcomes.

Modeling Strategies for Insect Pest Management: Genetic Analyses and Parameter Estimation

by
Michael Richard Vella

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Biomathematics

Raleigh, North Carolina

2020

APPROVED BY:

Alun Lloyd Co-chair of Advisory Committee	Fred Gould Co-chair of Advisory Committee
Brian Reich	David Rasmussen

Jason Delborne

DEDICATION

To my wife, Tiffany, my parents, John and Kathy, and my brother, Anthony.

BIOGRAPHY

Michael was born in Minneapolis, Minnesota. During high school, Michael became fascinated by applied mathematics, especially in relation to chaos theory and applications to physics. While attending the University of Notre Dame for his bachelor's degree in applied math and statistics, he discovered a passion for biological applications and worked on several biomath research projects throughout his time there. His desire to further his knowledge and continue researching prompted him to pursue a Ph.D. in Biomathematics at North Carolina State University.

ACKNOWLEDGEMENTS

First, I would like to thank my co-advisors Alun Lloyd and Fred Gould for everything they have done to help make my graduate experience enjoyable. It was a true pleasure working so closely with and learning an incredible amount from each of you. I also thank the rest of my committee members, Brian Reich, David Rasmussen, and Jason Delborne, each of whom I have learned from and am thankful to have interacted with.

I was fortunate to be involved in a number of collaborative projects during my Ph.D. I would particularly like to thank Christian Gunning, Sumit Dhole, Max Scott, Sophia Webster-Tostenson, Jennifer Baltzegar, Anthony James, and Thai Pham for their efforts in the projects we collaborated on. Many of those listed spent a tremendous amount of effort generating biological data, and I have appreciated my opportunities to be involved with such interesting research. Additionally, I thank the other members of the lab group for their feedback and questions about my research throughout the years.

More broadly, I would like to thank everybody who I interacted with through IGERT and in the GES Center. I value the many thought-provoking ideas I was exposed to which have expanded the scope of how I think in many areas.

I am grateful to have had such an amazing support system of friends with whom I could relax, laugh, and play intramural sports or board games. My graduate experience would not have been the same without all of you.

Finally, I am extremely thankful for my loving family. To my parents, thank you for your constant support and encouragement. To Anthony, I also thank you for your support and encouragement as well as your help editing this dissertation. To Tiffany, it has been wonderful to have you with me for the latter half of my Ph.D. You keep me on my toes with exciting adventures, and I appreciate all of your support. I love you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Background

Insect pest species pose threats to human health and agricultural yields. Vector-borne diseases cause over 700,000 deaths annually and amount to billions of dollars in costs (World Health Organization, 2017), while agricultural pests are estimated to cause losses of around 20% of global yields (Savary et al., 2019). A variety of approaches have been theorized and/or used to lessen the burden of insect pests. One of the most widespread approaches is the use of chemical insecticides. While insecticides can be effective at reducing population sizes, insects develop resistance, limiting the extent to which insecticides can be effective and requiring careful management to keep resistance levels low (Hemingway and Ranson, 2000; Kelly-Hope et al., 2008).

An alternative set of approaches involve genetic strategies aiming to suppress or eliminate a species. The simplest genetically engineered strains have dominant lethal genes that cause all offspring, or only the females (female-killing) to die and reduce population size over the course of many large, repeated releases (Alphey et al., 2013). Gene drives, rather, are engineered to spread with only a single release (Alphey et al., 2013). Several mechanisms for gene drive exist, but most recently, development of gene drives have utilized the CRISPR-Cas9 system (Hsu et al., 2014; Doudna and Charpentier, 2014) in many different species (e.g., Gantz and Bier, 2015; Champer et al., 2016; Champer et al., 2017).

Mathematical modeling of population genetics is a valuable tool to assist with combating insect pests. Considering genetic approaches, models can be useful during each step of development. First, models can be used to predict expected outcomes of a theoretical idea based on its mechanism and inheritance patterns. Genetic engineering can be extremely time and resource intensive (see, e.g., Gregory et al., 2016), so modeling results at this stage are important to help guide research efforts towards areas that are most likely to be effective in practice. For constructs that have already been developed and are being tested in the lab, models are useful to infer model parameters from lab testing, which can then be used to evaluate theoretical effectiveness outside of the lab. Finally, after release, models can be used to help analyze insect count and genotype data to measure how the release has performed.

Models of population genetics can make a range of assumptions that give them varying degrees of abstraction, even among models that assume spatial homogeneity. For example, models can

consider allele frequency (e.g., Burt, 2003; Deredec et al., 2008) or genotype frequency (e.g., Huang et al., 2009; Robert et al., 2013). Some models include density-dependent effects and age structure, which is important if the population size is being reduced to low numbers (e.g., Robert et al., 2013; Khamis et al., 2018). Additionally, time can be continuous as in differential equation models (e.g., Robert et al., 2013; Khamis et al., 2018) or take discrete steps (e.g., Deredec et al., 2008; Huang et al., 2009; Alphey et al., 2011; Dhole et al., 2018). Each type of model is useful in certain scenarios depending on the goal; generally, simple, abstract models can help to explain expected qualitative behavior, while more detailed models are often needed to quantify effectiveness and to ensure that the qualitative results still apply in scenarios where the simple model assumptions are strongly violated.

1.2 Outline

In this dissertation, I consider a broad range of systems using different models. In Chapter 2, I use a discrete-time, allele-frequency model to evaluate the effectiveness of novel genetic constructs that are designed to reverse CRISPR-Cas9 gene drives. In Chapter 3, I use a genotype-frequency, differential equation model with age structure and density-dependent mortality to compare two genetic alternatives of the female-killing approach. In Chapter 4, I use a genotype-frequency model, discrete-time model as the basis for Bayesian inference with genotype data from Iquitos, Peru, where a resistant allele increased in frequency as insecticide spraying occurred. Throughout each chapter, the common thread is that fairly simple models of population genetics can give valuable insights that can be used to improve approaches for combating insect pests.

Chapter 2

Evaluating strategies for reversing CRISPR-Cas9 gene drives[1]

[1]This chapter is published in *Scientific Reports*: Vella MR, Gunning CE, Lloyd AL & Gould F. (2017) Evaluating strategies for reversing CRISPR-Cas9 gene drives. Scientific Reports. 7: 11038. doi:10.1038/s41598-017-10633-2. I conducted all analyses and wrote the manuscript.

2.1 Abstract

A gene drive biases inheritance of a gene so that it increases in frequency within a population even when the gene confers no fitness benefit. There has been renewed interest in environmental releases of engineered gene drives due to recent proof of principle experiments with the CRISPR-Cas9 system as a drive mechanism. Release of modified organisms, however, is controversial, especially when the drive mechanism could theoretically alter all individuals of a species. Thus, it is desirable to have countermeasures to reverse a drive if a problem arises. Several genetic mechanisms for limiting or eliminating gene drives have been proposed and/or developed, including synthetic resistance, reversal drives, and immunizing reversal drives. While predictions about efficacy of these mechanisms have been optimistic, we lack detailed analyses of their expected dynamics. We develop a discrete time model for population genetics of a drive and proposed genetic countermeasures. Efficacy of drive reversal varies between countermeasures. For some parameter values, the model predicts unexpected behavior including polymorphic equilibria and oscillatory dynamics. The timing and number of released individuals containing a genetic countermeasure can substantially impact outcomes. The choice among countermeasures by researchers and regulators will depend on specific goals and population parameters of target populations.

2.2 Introduction

Recent work has employed the CRISPR-Cas9 system (Hsu et al., 2014; Doudna and Charpentier, 2014) to create homing drives (HD) that increase the frequency of genetic constructs in a population even if they lower the fitness of individuals that carry them (Gantz and Bier, 2015). The drive mechanism exploits homology directed repair (HDR) to replace a targeted, naturally occurring genomic sequence with an engineered construct (Esvelt et al., 2014; Champer et al., 2016). The HD construct codes for Cas9 (or any similar endonuclease, such as Cpf1 (Zetsche et al., 2015) and one or more guide RNAs so that in HD heterozygotes, the combined presence of Cas9 and the guide RNA(s) converts germline cells into HD homozygotes. The engineered construct may also include a novel, expressed gene.

An HD can be used in two different ways: for population suppression (“suppression HD”), where

the drive induces a major genetic load (Burt, 2003), or for population replacement ("replacement HD"), where the expressed gene in the drive construct induces an intended phenotypic alteration, such as blocked transmission of a pathogen (Gantz and Bier, 2015). Despite their promise, HDs carry a number of potential risks, including unforeseen ecological consequences and unintended geographical spread (Esvelt et al., 2014; National Academies of Sciences, Engineering, and Medicine, 2016). The severity of adverse effects could vary widely, with HD individuals and Cas9 remaining in the population. For example, the magnitude of such adverse impacts would likely be affected by the likelihood of undesirable HD migration and by the likelihood of low-probability events such as horizontal gene flow. In some instances, actions to gradually reduce HD frequency may be viewed as sufficient, while in other cases, swift, complete elimination of HD and restoration of the wild-type would be preferred.

It is possible for an HD bearing a fitness cost to naturally go extinct due to evolution against it, such as the spread of drive-resistant alleles developed via non-homologous end joining (NHEJ) (Burt, 2003; Unckless et al., 2015; Bull, 2016). This would likely prevent the HD from reaching fixation but not reduce HD frequencies quickly. HD constructs could also be engineered (i.e., no pre-existing resistant alleles in the population, and multiple guide RNAs to force simultaneous events of NHEJ for resistant alleles to arise) to minimize the likelihood of natural resistance (Deredec et al., 2008; Esvelt et al., 2014). Thus, several countermeasures have been proposed to proactively slow the spread of an HD and/or remove it from a population. In the case of a suppression HD, one option would be to release individuals carrying a synthetic allele of the targeted gene that is resistant to the HD (Burt, 2003; Deredec et al., 2008). In this case, the synthetic resistant (SR) allele would have no substantial fitness advantage over a replacement HD designed to have minimal fitness cost.

A second option that could be useful for stopping either suppression or replacement HDs involves synthetic CRISPR-Cas9 based "overwriting" or "reversal drives" (RD) (Esvelt et al., 2014; DiCarlo et al., 2015). CATCHA (Cas9-triggered chain ablation) and ERACR (elements for reversing the autocatalytic chain reaction) have been proposed as RDs (Gantz and Bier, 2016; Wu et al., 2016). The CATCHA and ERACR constructs contain guide RNAs but do not include the Cas9 gene, depending instead on Cas9 present from the HD. The guide RNAs produced by the RD target the HD construct in the same way that the HD targets the wild-type allele. A third option is using an "immunizing reversal drive" (IRD) that would target both HD and wild-type populations by including both the

Cas9 gene and multiple guide RNAs that target the HD and wild-type sequences (Esvelt et al., 2014). IRDs are designed to replace both HD-bearing and wild-type individuals, with constructs that have active Cas9 and guide RNA production but no intended effect on the organism's phenotype.

The National Academies of Sciences, Engineering, and Medicine report (National Academies of Sciences, Engineering, and Medicine, 2016) recommended the use of mathematical models in evaluating strategies for reducing potential harms of gene drives. An intuitively reasonable expectation, for example, is that RDs could be "employed to eliminate" an HD (Gantz and Bier, 2016). Yet there has been no quantitative assessment to date of the predicted dynamics of reversal and immunizing drives. Here we present a simple, frequency-only population genetics model to elucidate the evolutionary dynamics of genetic strategies for countering HDs. We show that SR alleles and RDs are not guaranteed to eliminate an HD from a population due to the existence, in general, of a stable polymorphic equilibrium in which the countermeasure co-exists with the wild-type and HD. An IRD, on the other hand, is much more likely to eliminate an HD but is also expected to eliminate wildtype alleles and continue production of Cas9.

2.3 Methods

We build on previous deterministic models of HD allelic dynamics that employ non-overlapping generations (i.e., a discrete-time description) and random mating (Deredec et al., 2008; Unckless et al., 2015). We add alleles for SR, RD, and IRD as countermeasures. Alleles for natural resistance are also examined. We assume that Cas9 always produces a double-strand break in wild-type/HD, HD/RD, wild-type/IRD, and HD/IRD heterozygotes. We assume that resistant alleles arise naturally (and only) via NHEJ whenever HDR is unsuccessful, such that the homing rate is equivalent to the probability of HDR. Finally, we assume that fitness costs yield an excess of lethality relative to wild-type at some point prior to reproduction, and that Cas9 is produced only in the germline. Note that, due to drive activity, gamete genotype contribution may differ, but conversion occurs only after somatic mortality via fitness cost is assessed.

We let q_W, q_{HD}, q_C, and q_R be the current generation frequencies of wild-type (W), HD, countermeasure (C), and naturally resistant (R) alleles in the population, respectively. The equations

predicting the next generation frequencies (q') are:

$$q'_{HD} = \big[(1-s_{HD})q_{HD}^2 + (1-h_{HD}s_{HD})q_{HD}q_W(1+e_{HD}) + q_{HD}q_R(1-s_{HD/R}) + (1-s_{HD/C})q_{HD}q_C(1-i_1)\big]\frac{1}{\overline{w}} \tag{2.1}$$

$$q'_C = \big[(1-s_C)q_C^2 + (1-h_C s_C)q_C q_W(1+i_2 e_C) + (1-s_{HD/C})q_{HD}q_C(1+i_1 e_C) + q_C q_R(1-s_{C/R)}\big]\frac{1}{\overline{w}} \tag{2.2}$$

$$q'_R = \big[(1-s_R)q_R^2 + (1-h_R s_R)q_R q_W + (1-h_{HD}s_{HD})q_{HD}q_W(1-e_{HD}) + i_1(1-s_{HD/C})q_{HD}q_C(1-e_C) + (1-s_{HD/R})q_{HD}q_R + (1-s_{C/R})q_C q_R + i_2(1-h_C s_C)q_C q_W(1-e_C)\big]\frac{1}{\overline{w}}, \tag{2.3}$$

where $q_W = 1-q_{HD}-q_C-q_R$ because the frequencies must add to one. The mean population fitness ($\overline{w}$) can be calculated by subtracting fitness cost deaths from one:

$$\overline{w} = 1 - s_{HD}q_{HD}^2 - s_C q_C^2 - s_R q_R^2 - 2\big(h_{HD}s_{HD}q_{HD}q_W + h_C s_C q_C q_W + h_R s_R q_R q_W + s_{HD/C}q_{HD}q_C + s_{HD/R}q_{HD}q_R + s_{C/R}q_C q_R\big). \tag{2.4}$$

Parameters e_{HD} and e_C are the probabilities of successful copying (homing) for the homing drive and countermeasure, respectively. The countermeasure allele represents SR when $i_1 = i_2 = 0$ (no homing), an RD when $i_1 = 1$ and $i_2 = 0$ (homing only in HD/countermeasure heterozygotes), and an IRD when $i_1 = i_2 = 1$ (homing in both HD/countermeasure and wild-type/countermeasure individuals).

We assume wild-type fitness is 1, and define s to be the fitness cost of homozygotes. The degree of dominance, h, gives the fraction of the homozygote fitness cost imposed on a heterozygote with one wild-type allele. We denote fitness costs of HD/R, HD/C, and C/R heterozygotes as $s_{HD/R}$, $s_{HD/C}$, and $s_{C/R}$, respectively. We assume fitness costs are recessive, with heterozygotes bearing the lesser fitness cost of its alleles, unless noted otherwise. Note that the RD and IRD may recode for the gene interrupted by the HD or eliminate an expressed gene in the HD construct such that the countermeasure constructs do not carry the same fitness costs as the HD construct.

2.4 Results

Figure 2.1 shows several examples of countermeasure dynamics that are indicative of behavior over a broad range of parameter values. In these examples, the countermeasures are deployed against a suppression HD, and we assume perfect homing. Figure 2.1a shows the rapid spread of the HD in the absence of countermeasures, where high HD fitness costs would result in population suppression or extinction. Figure 2.1b-g, compares impacts of release of an SR allele (Figure 2.1b-c); release of an RD (Figure 2.1d-e); and release of an IRD (Figure 2.1f-g), with each initiated using a single release of either a 1:1 (Figure 2.1b/d/f) or a 1:10 (Figure 2.1c/e/g) ratio into populations at the end of the 8th generation after the HD release, when the HD frequency has exceeded 0.2. Regardless of release size, the systems with SR and RD releases reach stable, polymorphic equilibria in the long term, whereas the IRD eliminates the HD and reaches fixation. The SR (Figure 2.1b-c) reaches high frequencies and slowly diminishes HD frequencies, though ongoing conversion of wild-type to HD is sufficient to maintain the HD in the population. The larger release of RD (Figure 2.1d) immediately brings the system close to the equilibrium, causing HD frequencies to stay relatively constant. The smaller RD release (Figure 2.1e) allows HD frequencies to initially increase, which may not be desirable. However, the subsequent buildup of RD then reduces HD to very low frequencies, in contrast to what was seen in Figure 2.1d. In this trough of low HD frequency, stochastic loss of HD via drift may occur, with the HD loss probability increasing as population size decreases (Hartl and Clark, 2007). The IRD does not coexist with other alleles because it maintains an advantage over each of the other alleles regardless of its frequency and quickly reaches fixation regardless of release size (Figure 2.1f-g).

Moving to consider replacement HDs, Figure 2.2 shows a set of time series for an HD with lower fitness cost (s_{HD}), but with otherwise identical parameter values as shown in Figure 2.1. The qualitative behavior in the replacement HD setting is similar to the behavior in the suppression HD setting, but the lower HD fitness cost slows dynamics. The difference is most notable for the RD with a small release of countermeasure, for which the system exhibits large, slowly damped oscillations that bring the target HD to low frequencies for many generations (Figure 2.2e). Due to genetic drift, the likelihood of stochastic loss of an allele increases as the time spent with few copies of that allele in the population increases (Hartl and Clark, 2007).

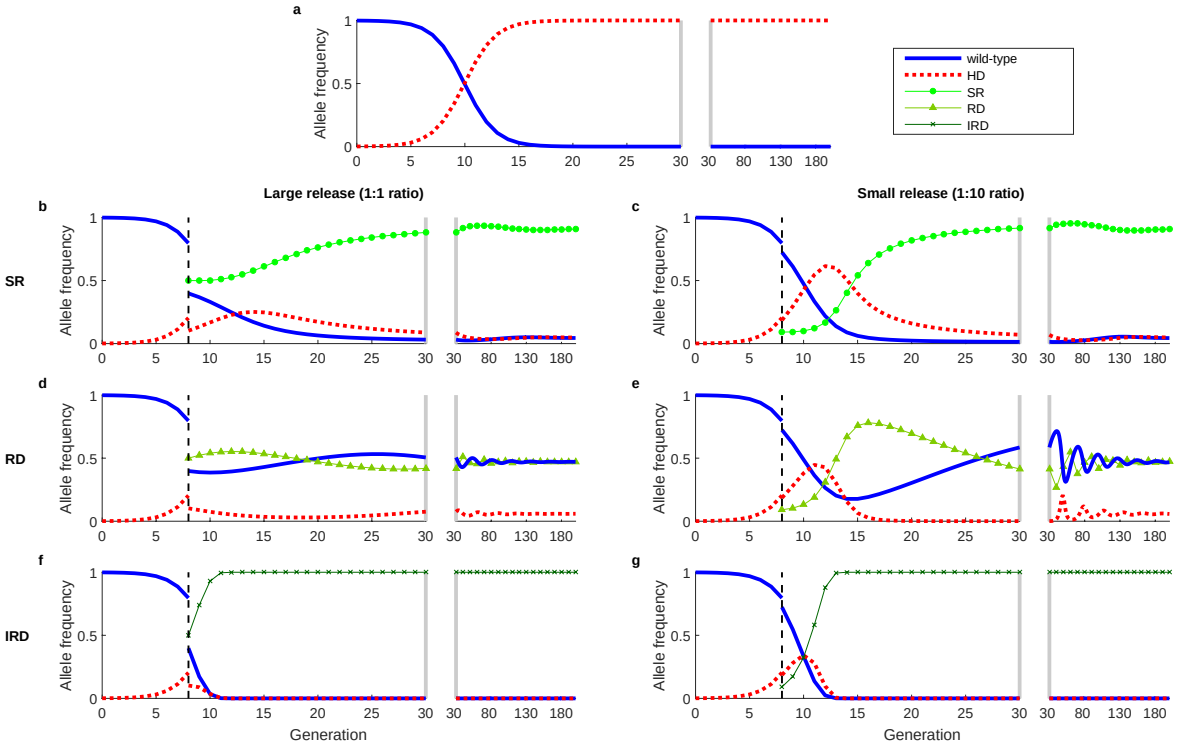

Figure 2.1 Dynamics of a suppression HD, alone (a) and with countermeasures (b-g), which include a synthetic resistant allele (b,c), reversal drive (d,e), and immunizing reversal drive (f,g). Fitness cost (s) is relative to and recessive to wild-type (HD, $s_H = 1$; SR, s_C=0.05; RD/IRD, s_C=0.2.). We use an initial release of 0.1% HD, and assume recessive lethality of the HD allele and perfect homing ($e_{HD} = e_C = 1$). Dashed vertical lines indicate the time of countermeasure release. Large releases (1:1 ratio or countermeasure to pre-countermeasure-release population) are shown in the left column, and small releases (1:10 ratio) are shown in the right column. The split axes with gray bars indicate a change in time scale. **a:** Absent countermeasures, the HD quickly approaches fixation (i.e., would cause population extinction). **b,c:** Release of a SR allele allows a brief increase in HD frequency, followed by a decrease to a low but non-zero equilibrium. **d:** A large RD release yields allelic frequencies after release that are near the stable equilibrium. **e:** A small RD release yields allelic frequencies far from equilibrium, followed by a large transient oscillation, wherein HD frequencies approach zero. **f,g:** Release of IRD results in elimination of HD and wild-type alleles, regardless of release size.

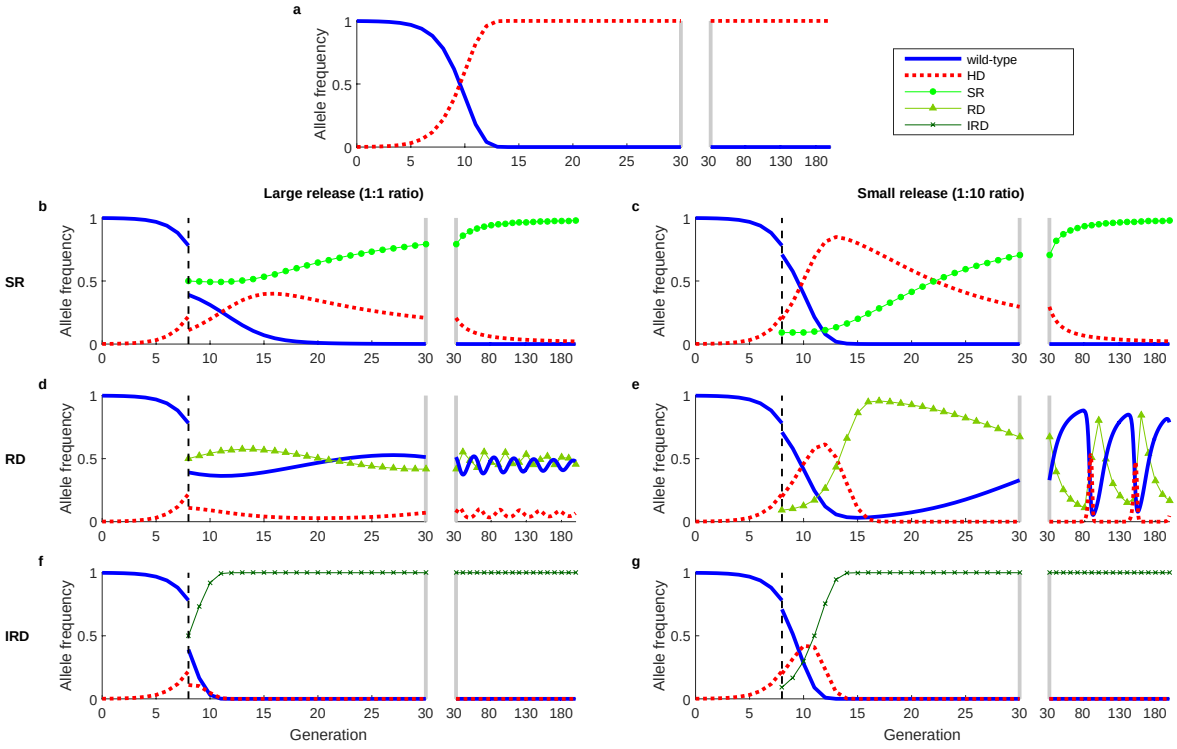

Figure 2.2 Dynamics of a replacement HD, alone (a) and with countermeasures (b-g), which include a synthetic resistant allele (b,c), reversal drive (d,e), and immunizing reversal drive (f,g). The fitness cost of the HD is s_H=0.3. See Figure 2.1 for other details. The behavior is qualitatively similar to Figure 2.1, but the oscillations of the SR and RD are less damped (b-e).

The polymorphic equilibrium and oscillatory dynamics exhibited by the SR and RD systems are due to each allele's frequency-dependent disadvantages relative to other alleles in a "rock-paper-scissors" type fashion. In this case, the disadvantages result from relative fitness costs and the effects of drive (for the RD), but similar dynamics have been recognized in many unrelated systems (Sinervo and Lively, 1996; Durrett and Levin, 1997; Kerr et al., 2002; Cameron et al., 2009). Damped oscillations about a polymorphic equilibrium mean that initial conditions far from the equilibrium result in large fluctuations, temporarily bringing HD frequencies near to zero. Initial conditions close to the equilibrium, on the other hand, do not result in large fluctuations in allelic frequencies, likely allowing the HD to persist (as visualized in a phase plot in Figure A.1a). Initial conditions are not important for determining the fate of the IRD, however, as it does not have frequency-dependent disadvantages to the other alleles.

Relaxing assumptions about fitness but keeping the assumptions of perfect homing and recessive fitness costs in wild-type heterozygotes, we find that a variety of possible stable equilibria may exist for the systems beyond those shown in Figures 2.1 and 2.2 (See Appendix A & Figures A.1-A.3). However, given no fitness cost for heterozygotes containing wild-type alleles, a stable, polymorphic equilibrium exists for the SR and RD countermeasures for most plausible combinations of HD, C, and HD/C fitness costs (e.g., when the HD/C heterozygote fitness cost is between the HD and C homozygote fitness costs). Numerically, we find complex eigenvalues of the Jacobian evaluated at the polymorphic equilibrium, which indicate oscillatory dynamics (see Appendix A). Additionally, assuming additive rather than recessive fitness costs in wild-type heterozygotes changes the regions of parameter space that result in each equilibrium for the SR and RD countermeasures but still results in uncertain removal of the HD for SR and RDs, and likely removal of the HD for IRDs (Figures A.4 and A.5).

With a deterministic model, likelihood of stochastic extinction during transient oscillations cannot be measured directly, but the likelihood increases as the minimum frequency decreases. Figures 2.3 and 2.4 show the minimum HD frequency achieved within the first 100 generations after countermeasure release for varying fitness costs and initial conditions, returning to the assumption of recessive fitness costs and that the cost to the HD/C heterozygote is the minimum of the HD and C fitness costs. In general, low countermeasure fitness costs yield the greatest reductions in HD frequencies, both for RD (Figure 2.3) and for SR (Figure 2.4) countermeasures, by lowering the

HD frequency at the polymorphic equilibrium. When HD frequencies are low, an RD released in numbers close to that of the current population (1:1 ratio) causes the system to quickly approach the polymorphic equilibrium instead of exhibiting large transient oscillations that bring the HD frequency near to 0. A very large RD release that immediately limits HD/wild-type mating would likely cause stochastic HD elimination (bottom row) in a randomly mating population, but would require additional time and resources necessary to rear and release a sufficient number of RD individuals. Unlike the RD, SR cannot be effective when its fitness cost exceeds that of the HD (Figure 2.4, top-left corner of each panel). As with the RD, SR releases in size equal to the pre-release population bring the system near to its polymorphic equilibrium, which would prevent the HD frequency from transiently reaching very small frequencies. However, because oscillations occur on a slower time-scale than with the RD (see Figure 2.1b/e), the minimum is not always reached within 100 generations. For an IRD, all panels reach minimum frequencies by generation 100.

Finally, we further relax our assumptions to account for less than perfect homing with the creation of naturally resistant alleles (Figure 2.5). The qualitative behavior found in the case of perfect homing remains, except that the IRD eventually falls out of the population since it has lower fitness than naturally resistant alleles. Given imperfect homing, HD frequencies would fall even in the absence of countermeasures. As with SR, though, the HD is sustained long-term due to a stable, polymorphic equilibrium (Figure 2.5a).

2.5 Discussion

A variety of genetic approaches have been proposed to counter unintended effects of an HD, but there has been limited theoretical evaluation of these approaches. Here we compare the dynamics of SR, RD, and IRD countermeasures upon release into a population prior to HD fixation and find that the long-term behavior of the system differs greatly between countermeasures. In particular, SR and RD countermeasures are not guaranteed to eliminate an HD from a population because these systems often exhibit a stable polymorphic equilibrium. Elimination of the HD via SR or RD becomes less likely with higher countermeasure fitness costs, as the equilibrium HD frequency is further from zero, and oscillations around the polymorphic equilibrium are less likely to cause stochastic loss of the HD. Due to the small magnitude of oscillations with release conditions close to

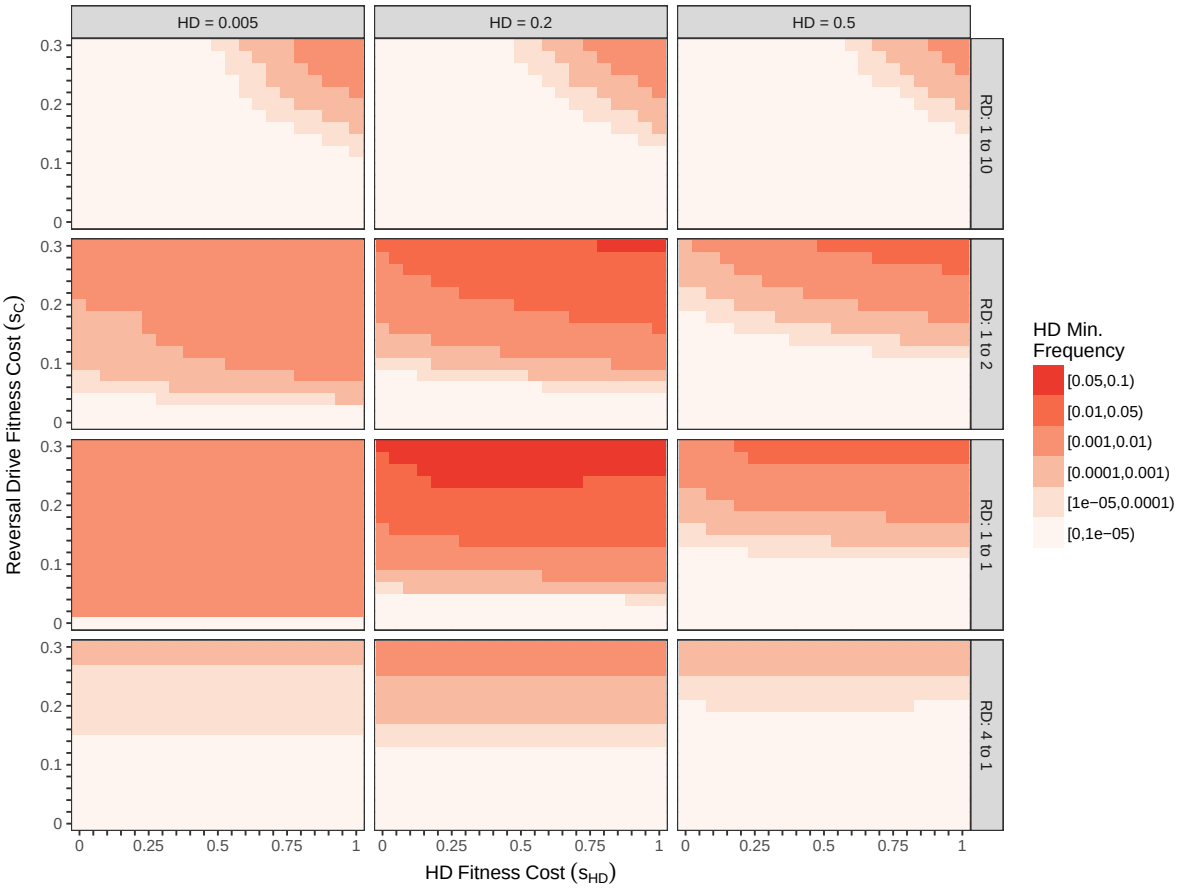

Figure 2.3 Minimum HD allele frequency in the first 100 generations after RD release for various fitness costs, initial conditions, and release ratios. Light shades indicate higher likelihood of stochastic loss of HD, while dark shades highlight instances where removal of the HD is less likely. Axes show fitness costs of the HD (x-axis) and RD (y-axis). Initial conditions vary between panels: columns vary the HD pre-release frequency, and rows vary the RD release size, which is shown as a release ratio (e.g., "4 to 1" releases 4 RD alleles for every pre-release allele). We assume recessive fitness costs and perfect homing. The largest HD fitness cost ($s_{HD} = 1$) corresponds to a suppression HD, whereas small HD fitness costs correspond to a replacement HD. Note that maximum HD frequency varies independently from minimum HD frequency; in small RD releases (top row of panels), the HD frequency can experience large increases before dropping to the low minimum levels show here. Overall, an RD release appears least likely to eliminate a target HD when RD fitness costs are large, and when the RD release yields post-release frequencies near the equilibrium. The higher minimum frequency for larger HD fitness costs in many panels is due to the smaller amplitudes of oscillations compared to systems with lower HD fitness costs, as seen Figures 2.1 and 2.2. Smaller oscillations result in the system tending directly toward the equilibrium.

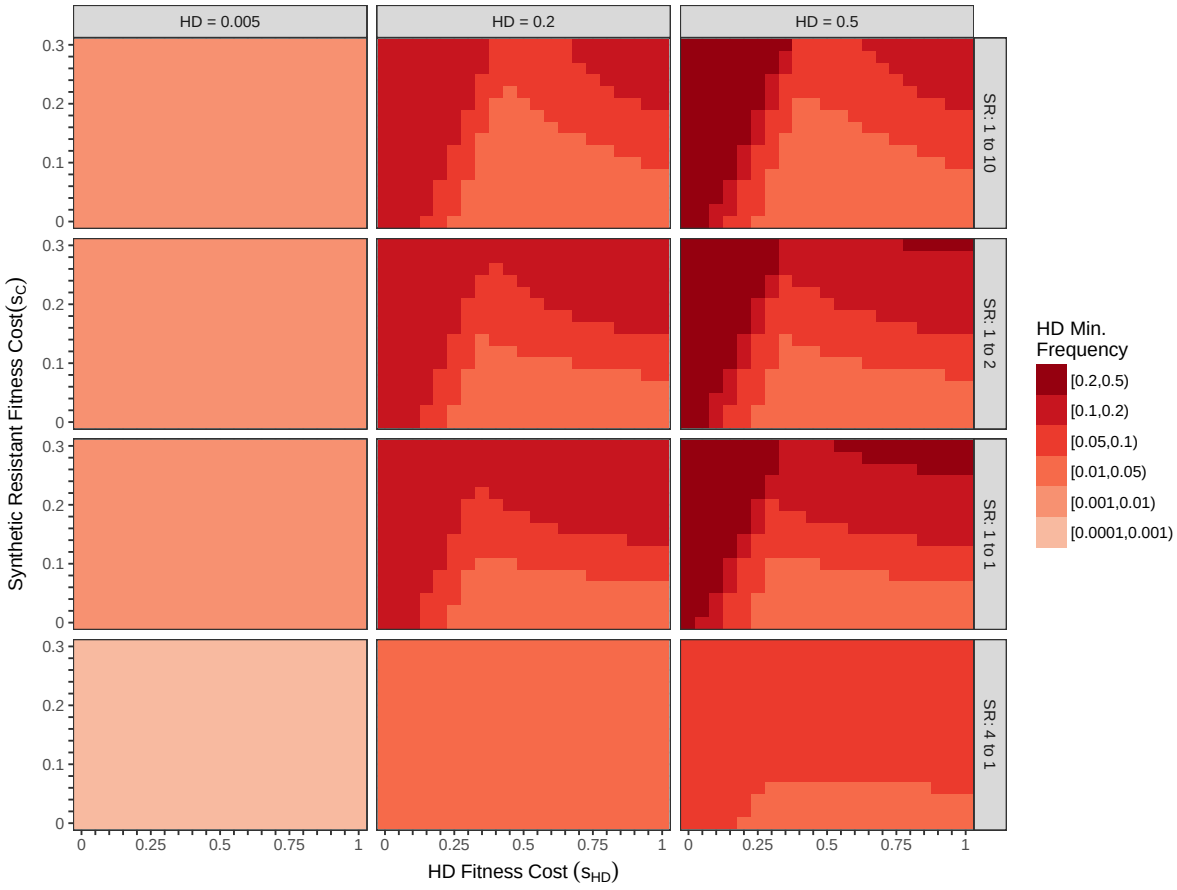

Figure 2.4 Minimum HD allele frequency in the first 100 generations after SR release for various fitness costs, initial conditions, and release ratios. See Figure 2.3 for details, noting that the legend colors refer to different minimum frequencies. Similarly to the RD, the SR is least likely to eliminate a target HD when its fitness costs are large, and when the release yields post-release frequencies near the equilibrium, though equilibrium frequencies are not identical to RDs. In some of the simulations, the system is not yet at equilibrium, and the HD is still decreasing in frequency at 100 generations.

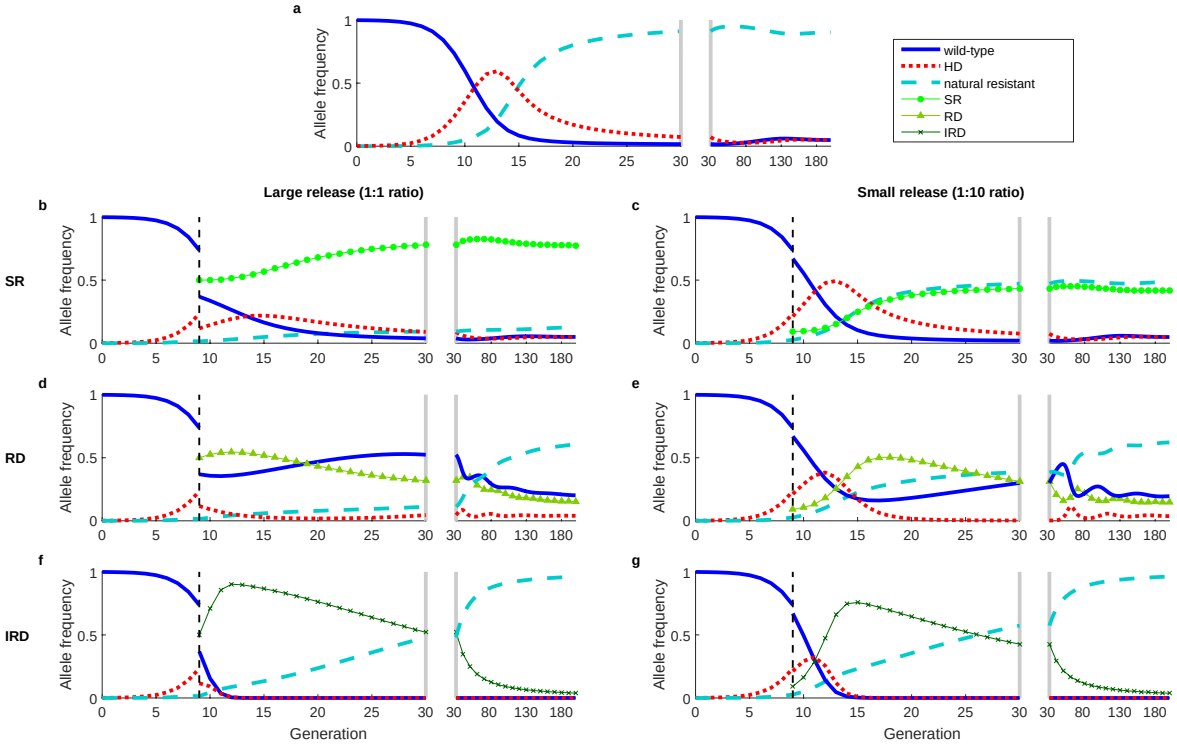

Figure 2.5 Dynamics of an imperfect suppression HD, where the drive fails and produces naturally resistant alleles, alone (a) and with countermeasures (b-g), which include a synthetic resistant allele (b,c), reversal drive (d,e), and immunizing reversal drive (f,g). Homing is imperfect ($e_H = e_C = 0.9$), and unsuccessful homing results in natural resistance via NHEJ with fitness cost s_R=0.05. See Figure 2.1 for other details. The biggest change from accounting for imperfect homing is that the IRD falls out of the population in the long-term (f,g) because of low-fitness cost alleles resistant to cutting.

equilibrium, the frequencies of the HD prior to release and the relative size of the countermeasure release are important factors in determining the likelihood of HD elimination. If either of these countermeasures were to fully eliminate the HD, the wild-type allele would ultimately recover to fixation as long its fitness is higher than the countermeasure.

An IRD that targets both HD and wild-type alleles, on the other hand, would theoretically ensure the rapid removal of the HD from the population, but would also result in the Cas9 gene and guide RNAs remaining in the population. Implications of leaving Cas9 in the population are unclear, such as the likelihood of off-target effects, and future research should seek to evaluate such effects. If any naturally resistant alleles develop, or with the release of an SR allele, the IRD would eventually fall out of the population, provided that the cost of the IRD is greater than the resistant allele. These qualitative differences between countermeasures must be considered when deciding whether they are suitable tools for mitigating adverse effects of an HD.

The model and subsequent analysis presented here yields critical insights into the qualitative behavior of, and differences between, genetic countermeasures. Nonetheless, future work could explore several additional aspects of HD-based countermeasures and provide quantitative risks associated with them. Models that track population size as well as allele frequency, and that incorporate demographic stochasticity, could be used to better assess options for eliminating suppression HDs. For suppression HDs, population size could drastically decrease, and the effects of genetic drift could predominate (Okamoto et al., 2014). Also deserving of increased attention are the effects of spatial heterogeneity. In particular, spatial isolation of small populations could limit an allele's spread, potentially impacting countermeasure success. Incorporating spatial heterogeneity could also be useful in assessing the impact of movement between the target population and nearby populations on the long-term fates of the relevant constructs. Important consequences of movement include the likelihood of HD spillover to nearby populations, and whether immigration of wild-type organisms could sustain a HD in a system where stochastic elimination is otherwise likely. Effects of spatial heterogeneity may be different for RDs and IRDs, so follow-up modeling studies will be needed. Finally, effects of assumptions about natural resistance to homing drives should be explored further. While some work has explored the development of natural resistance to HDs (Burt, 2003; Unckless et al., 2017; Bull, 2016; Noble et al., 2017; Marshall, Buchman, et al., 2017; Champer et al., 2017), these findings should be updated as HD limitations are understood.

Many have proposed countermeasures as emergency tools to mitigate unintended negative effects that might arise after release of an HD. However, to date only limited theoretical analysis has addressed countermeasures' abilities to reverse HDs. Additionally, discussion about countermeasures has often been ambiguous regarding differences between types of countermeasures and expectations of countermeasure outcomes. Depending on the severity of unintended effects, countermeasures may have the goal of simply halting the spread of an HD, or possibly removing an HD from the population and returning the population to its original state. This work is motivated by a desire to more clearly specify differences between various countermeasure strategies, as well as to critically assess potential outcomes. Here we show that the RD does not eliminate the HD for certain release conditions and fitness parameters. The existence of a polymorphic equilibrium with oscillatory dynamics allows for the HD allele frequency to initially increase, to remain constant, or to decrease, depending on the reversal release size. In such cases, larger countermeasure fitness costs decrease the likelihood of long-term eradication of the HD allele. IRDs are expected to effectively eliminate the HD in a timely manner but leave Cas9 present in the population, though any resistant alleles would cause the IRD to eventually fall out of the population. RDs leave only guide RNAs if they successfully eliminate the HD, but given any fitness cost to the RD, the wild-type would be expected to return. Overall, these results show that no single countermeasure, as currently proposed, should be considered a "silver bullet" for mitigating unintended effects of HDs. As such, we recommend careful examination of risks associated with each of the countermeasures' limitations prior to release.

2.6 Acknowledgements

This research was funded by the National Institutes of Health (NIH) grant R01-AI091980, the W. M. Keck Foundation, and the National Science Foundation (RTG/DMS - 1246991 and NSF-IGERT - 1068676). We thank Brandon Hollingsworth, Jaye Sudweeks, Sumit Dhole, Jennifer Baltzegar, Kevin Esvelt, Ethan Bier, Anthony James, and Valentino Gantz for helpful discussion.

Chapter 3

Mathematical modeling of genetic pest management through female-specific lethality: Is one locus better than two?[1]

[1]This chapter is submitted for publication: Vella MR, Gould F & Lloyd AL. Mathematical modeling of genetic pest management through female-specific lethality: Is one locus better than two? I conducted all analyses and wrote the manuscript.

3.1 Abstract

Many novel genetic approaches are under development to combat insect pests. One genetic strategy aims to suppress or locally eliminate a species through large, repeated releases of genetically engineered strains that render female offspring unviable under field conditions. Strains with this female-killing (FK) characteristic have been developed either with all of the molecular components in a single construct or with the components in two constructs inserted at independently assorting loci. Strains with two constructs are typically considered to be only of value as research tools and for producing solely male offspring in rearing factories which are subsequently sterilized by radiation before release. A concern with the two-construct strains is that once released, the two constructs would become separated and therefore non-functional. The only FK strains that have been released in the field without sterilization are single-construct strains. Here, we use a population genetics model with density dependence to evaluate the relative effectiveness of female killing approaches based on single- and two-construct arrangements. We find that, in general, the single-construct arrangement results in slightly faster population suppression, but the two-construct arrangement can eventually cause stronger suppression and cause local elimination with a smaller release size. Based on our results, there is no a priori reason that males carrying two independently segregating constructs need to be sterilized prior to release. In some cases, a fertile release would be more efficient for population suppression.

3.2 Introduction

Insect pests remain a burden to human health and agriculture (World Health Organization, 2017; Deutsch et al., 2018). Genetic pest management aims to reduce this burden by releasing engineered insects that either introduce a desired trait into a natural population or reduce the size of the population. There have historically been several large area-wide inundative releases of male insects that were rendered sterile by exposure to radiation (Gould and Schliekelman, 2004). In these releases, local elimination of the target species was achieved as females increasingly mated with the sterile males rather than the wild-type males with whom they would produce viable offspring. Instead of using radiation to cause sterility, a contemporary alternative is to genetically engineer strains in

which the males cause all their offspring or exclusively their daughters to die or to have low fitness (Alphey, 2002). Genetically engineered strains in a number of species have been cage- or field-tested (Wise De Valdez et al., 2011; Ant et al., 2012; Harris et al., 2012; Lacroix et al., 2012; Leftwich et al., 2014; Carvalho et al., 2015; Harvey-Samuel et al., 2015; Gorman et al., 2016).

One approach to developing these functionally-sterile strains involves inserting a repressible, dominant lethal trait, which can be active in both sexes or in females only (Heinrich and Scott, 2000; Thomas et al., 2000). For either the female-killing (FK, also sometimes referred to as fsRIDL, or female-specific release of insects carrying dominant lethals) or bisex-killing (BK), in order to rear the transgenic strain in the generations prior to release, it must be possible to inactivate the dominant lethal gene. Often this is achieved through a Tet-off system where tetracycline in the diet represses the activator for a lethal gene (Gossen and Bujard, 1992). For an FK strain, the release generation is reared on a diet not containing tetracycline. This results in only males surviving. Further, as the offspring of released FK or BK males would feed on a tetracycline-free diet under field conditions, the lethal gene is turned on and death ensues.

Intuitively, modeling studies have found that FK can be advantageous over BK because it kills females while allowing the transgene to propagate through multiple generations in heterozygote males (Schliekelman and Gould, 2000; Thomas et al., 2000). This would seem especially useful when females but not males transmit pathogens. However, heterozygous males can also serve as a reservoir for wild-type alleles, which can make FK less effective than BK under some conditions (Foster et al., 1988; Gentile et al., 2015). It should be noted that BK strains for mosquito disease vectors typically require sex-sorting because release of females would be considered unacceptable. It can also be advantageous to release only males as females do not contribute to genetic suppression and tend to mate with the released males and thus reduce their efficiency (Rendón et al., 2004) except in some situations where there is age structuring in the population (Huang et al., 2009).

The full molecular design involves two molecular components: 1. the tetracycline-repressible transactivator (tTA) with a promoter, and 2. a lethal gene with an enhancer/promoter consisting of multiple tTA binding sites (tetO) and a core promoter. In the initial two-component systems, tTA was expressed in females by using a female-specific promoter (Heinrich and Scott, 2000; Thomas et al., 2000). The second component was a lethal gene (e.g., proapoptotic) driven by a tetO enhancer-promoter. The two molecular components were built in separate constructs that were inserted

independently. Subsequently, a simpler, two-component system was developed in which tTA acts as both the activator and lethal gene. Here, a single construct includes a tTA coding sequence driven by a tetO enhancer-promoter. In this autoregulated system, high levels of the tTA activator cause lethality in late stage larvae or in pupae. The mortality is possibly due to a general interference in transcription (Gong et al., 2005). FK single-construct strains have included a sex-specifically spliced intron from the transformer or doublesex genes inserted within the tTA gene (Fu et al., 2007). In these FK strains, only the female tTA transcript encodes a functional protein. A different, single-construct approach for FK with Aedes aegypti and Aedes albopictus uses a female-specific indirect flight muscle promoter from the Actin-4 gene (Fu et al., 2010; Labbé et al., 2012). All field trials with transgenic FK or BK strains have been with single-construct strains.

More recently, two-construct FK strains have been made with an early embryo promoter driving tTA expression and a tTA-regulated lethal gene that contains a sex-specifically spliced intron (Yan et al., 2020). An advantage of these strains is that female lethality occurs at the embryo or early larval stages, which produces considerable savings in larval diet costs in a mass rearing facility. Although it should be possible to develop any two-component systems as a single construct (Yan and Scott, 2015), they are typically developed as independently-segregating constructs. Germline transformation in insects is often achieved through the use of transposable elements such as piggyBac, and due to the randomness of the insertion process, a large number of injections can be required to obtain transgenic strains (Gregory et al., 2016). Furthermore, there are often multiple potential choices for one or both of the components. For this reason, it can be advantageous to separately produce strains with different promoters and lethal genes, then produce individuals bearing both components by crossing to test effectiveness of different combinations. The final transgenic insects have the two components located at two, separate loci (Schetelig and Handler, 2012; Ogaugwu et al., 2013; Scott, 2014; Schetelig et al., 2016; Yan et al., 2020).

FK strains with two constructs are generally thought of as useful research tools with potential to be used in rearing facilities so that the final generation before release would only produce males (e.g., Schetelig and Handler, 2012; Ogaugwu et al., 2013; Yan and Scott, 2015). It has been suggested that independent inheritance of the components would cause a breakdown in the female killing in the second generation after release (Ogaugwu et al., 2013; Yan and Scott, 2015). However, previous theoretical studies of FK systems have only modeled the components as being inserted together on

Table 3.1 1-locus genotypes, with associated viabilities and fitnesses. Viability of genotype i, $\gamma_i^{a,S}$, takes the value listed when the approach causes loss of viability in sex S (for female killing, only when $S = \text{F}$; for bisex, when $S = \text{F}$ or $S = \text{M}$) with timing a (for early approaches, when $a = \text{E}$; for late approaches, when $a = \text{L}$), and are 1 otherwise. Fitnesses w_i^x apply for both hatching ($x = \text{H}$) and male mating competitiveness ($x = \text{M}$).

i	**genotype**	$\gamma_i^{a,S}$	w_i^x
1	kk	1	1
2	Kk	0	$1-h\cdot s^x$
3	KK	0	$1-s^x$

a single locus (Thomas et al., 2000; Schliekelman and Gould, 2000; Alphey et al., 2011; Gentile et al., 2015).

Here we evaluate the effectiveness of 1- and 2-locus FK, along with BK for comparison. We use a computational model parameterized for the Aedes aegypti mosquito that is a vector for several human pathogens. We explore the release of strains with killing in either juveniles or adults. We show that under reasonable assumptions about fitness costs of the insertions, there is not a substantial difference between the 1- and 2-locus FK approaches, particularly when compared to the differences between FK and BK. These results demonstrate the release potential of recently developed 2-locus FK constructs.

3.3 Methods

Our computational model implements the genetics of FK and BK by separately tracking the number of individuals in the population of each genotype, with genotype denoted by subscript i. For the single-locus system (Table 3.1), we let the transgenic allele be represented by K and the wild-type allele at that locus be represented by k, with a total of $N = 3$ possible genotypes. For the 2-locus system (Table 3.2), we let A and B represent the transgenic alleles (i.e., tTA and lethal gene) inserted at two separate loci with wild-type alleles a and b, respectively, for a total of $N = 9$ possible diploid genotypes.

We assume complete effectiveness of the constructs, so when there is no gene repression via tetracycline, all individuals bearing the functional BKS system and all females with the functional FK system die, with a (i.e. genotype viability of zero) (see, e.g., Fu et al., 2007; Ogaugwu et al., 2013;

Table 3.2 2-locus genotypes, with associated viabilities and fitnesses. Viability of genotype i, $\gamma_i^{a,S}$, takes the value listed when the approach causes loss of viability in sex S (for female killing, only when $S = \text{F}$; for bisex, when $S = \text{F}$ or $S = \text{M}$) with timing a (for early approaches, when $a = \text{E}$; for late approaches, when $a = \text{L}$), and are 1 otherwise. Fitnesses w_i^x apply for both hatching ($x = \text{H}$) and male mating competitiveness ($x = \text{M}$).

i	**genotype**	$\gamma_i^{a,S}$	w_i^x
1	aabb	1	1
2	aaBb	1	$1-h\cdot s^x(1-c_A)$
3	aaBB	1	$1-s^x(1-c_A)$
4	Aabb	1	$1-h\cdot s^x c_A$
5	AaBb	0	$1-h\cdot s^x$
6	AaBB	0	$1-h\cdot s^x c_A - s^x(1-c_A)$
7	AAbb	1	$1-s^x c_A$
8	AABb	0	$1-s^x c_A - h\cdot s^x(1-c_A)$
9	AABB	0	$1-s^x$

Yan et al., 2020). One copy of K is assumed to be sufficient to induce lethality in the 1-locus system, and only one copy each of A and B is required in the 2-locus system. We consider lethality acting at different points in the lifecycle. In insects that experience strong resource competition during larval stages, having the transgene-induced mortality occur during or shortly after the pupal stage, instead of during the egg or larval stages, can yield stronger population suppression. This is because the transgenic juveniles consume resources and thereby increase wild-type juvenile mortality. We model early mortality (E) as occurring in the embryo and late mortality (L) as occurring in pupal stages or in adults before mating, and we assume these differentiate whether the individual contributes toward density-dependent mortality of all individuals in the population. We let $\gamma_i^{E,S}$ and $\gamma_i^{L,S}$ represent the early (embryonic) and late (adult) expected viabilities for individuals of sex S and genotype i. Tables 3.1 and 3.2 give expected viabilities for individuals with each construct and genotype.

We classify constructs into four different approaches depending on when the dominant lethal gene is active similarly to Gentile et al., 2015: early bisex-killing (E-BK), late bisex-killing (L-BK), early female-killing (E-FK), and late female-killing (L-FK). We assume male transgenic homozygotes are released, so mating with wild-type females will produce offspring that are entirely heterozygous, with a copy of each transgene. If the construct(s) affects both sexes (BK), none of these offspring will survive to mate and pass on their genes, making bisex 1-locus and 2-locus equivalent in terms of both population genetics and population dynamics. Female-specific approaches (FK) allow males

to continue to propagate the transgenes, and thus inheritance differs between 1-locus and 2-locus approaches. In all, we consider the following six approaches: E-BK, L-BK, 1-locus E-FK (E-FK1), 2-locus E-FK (E-FK2), 1-locus L-FK (L-FK1), and 2-locus L-FK (L-FK2).

Separate from the transgenic, toxin-induced lethality, we account for potential fitness costs caused by the genetic insertion itself. We allow the fitness costs of inserting a novel genetic element to manifest at an early stage as a reduction in the ability of a zygote to survive beyond the egg stage, i.e., the fraction of eggs of that genotype which survive and hatch into larvae. We let the genotype's hatching fitness, w_i^H, equal the probability of successfully entering the larval stage, with wild-type hatching fitness $w_1^H = 1$. We also allow for transgenic fitness costs to males in the form of reduced mating competitiveness, w_i^M, as defined below, with wild-type mating competitiveness w_1^M=1.

We generally assume that the fitness costs are equal for the homozygotes in the 1-locus and 2-locus systems to facilitate a direct comparison between the two systems. The 2-locus system has the same components as the 1-locus system, which makes equal fitness costs a reasonable base assumption for the purposes of this work (this assumption is relaxed in Figure 3.3 and Figure B.1). We let s^H and s^M be the hatching and mating competitiveness fitness costs, respectively, to the homozygotes KK and AABB, and we allow the two types of costs to vary independently. For simplicity, we assume the degree of dominance for the fitness costs, h, is equal for hatching and mating competitiveness. Unless otherwise noted, we assume costs are additive, with $h = 0.5$, such that each copy of the K allele alone contributes a fitness cost of $0.5s^x$ for the 1-locus system (x here indicates that the fitness cost can either be hatching or mating). For the 2-locus system, we allow for unequal fitness costs between each of the insertions. We let two copies of the A allele contribute a fitness cost of $s^x c_A$, where c_A is the proportion of the total 2-locus fitness cost accounted for by the A allele, and one copy of the A allele contributes a fitness cost of $h \cdot s^x c_A$. A single B allele contributes a fitness cost of $h \cdot s^x(1-c_A)$, while being homozygous for B induces a cost of $s^x(1-c_A)$. Resulting fitness expressions for all genotypes are listed in Tables 3.1 and 3.2.

We model genotype counts over time using a system of ordinary differential equations adapted from Robert et al. (2013). We let $J_i^M(t)$ and $J_i^F(t)$ be the number of juvenile (larvae and pupae) males and females, respectively, of genotype i at time t, and $A_i^M(t)$ and $A_i^F(t)$ be the number of viable adult male and adult female mosquitoes in the population, respectively, of genotype i at time t. This gives a maximum of 12 classes of individuals to track (each with different combinations of the 3 genotypes,

2 sexes, and 2 age classes) for the 1-locus system and 36 classes for the 2-locus system, though lethality from the genetic construct prevents survival of certain classes. For instance, E-BK only has five non-zero classes (wild-type male and female juveniles and adults, and male adult homozygote transgenic, which are released). We also assume a 50:50 sex ratio and equal hatching fitness costs between males and females, allowing a further reduction in the number of unique classes for late acting approaches because $J_i^M(t) = J_i^F(t)$ for all i and t. This results in 7 classes for L-FK1 (after removing 3 juvenile classes and 2 non-viable adult female classes) and 23 classes for L-FK2 (after removing 9 juvenile classes and 4 non-viable adult female classes). These dimensionality reductions can be useful when finding analytical solutions, but for simplicity, we computationally simulated all 12 (1-locus) or 36 (2-locus) equations.

Accounting for fitness costs, adult females produce juveniles of genotype i at time t at rate

$$B_i(t) = w_i^H \sum_m A_m^F(t) \lambda \sum_n P(i|m,n) \frac{w_n^M A_n^M(t)}{\sum_g w_g^M A_g^M(t)}, \tag{3.1}$$

where λ is the per-capita birth rate and $P(i|m,n)$ is the probability that a juvenile produced from a mating between a female and male of genotypes m and n, respectively, will be of genotype i. The fraction gives the probability that a randomly chosen male adult is of genotype n, weighted by mating competitiveness w_n^M. The offspring genotype probabilities are calculated assuming Mendelian inheritance, and for the 2-locus case, independent segregation of genes at each locus. Because of the 50:50 sex ratio, half of the hatching juveniles are male and half are female.

Juveniles of each genotype and sex emerge to adulthood at per-capita rate ν. We assume juveniles, adult males, and adult females have per-capita density-independent mortality rates of μ_J, μ_M, and μ_F, respectively. Juveniles also undergo density-dependent mortality at a per-capita rate $(\alpha J)^{\beta-1}$, where J is the total number of juveniles. The strength of density dependence is adjusted by varying β, with higher β resulting in a faster return to equilibrium population size after a small perturbation. A value of $\beta = 2$ gives the logistic model for population dynamics. By default, we let $\beta = 3$ to model an environment which would be more difficult for successful suppression (e.g., Hibbard et al., 2010). The equilibrium size of an entirely wild-type population varies with α, and to keep simulations with different values of β comparable, we choose the value of α so that the equilibrium number of wild-type females remains the same.

We assume a continuous release of homozygote engineered males at (daily) rate $u_i^M = r \cdot A_1^M(t = 0)/7$, where r is the weekly release ratio (engineered:wild-type) based on the equilibrium number of males prior to the release. By maintaining a constant number of released males, the effective release ratio increases as the population size decreases. The release genotype is KK for 1-locus ($i = 3$) and AABB for 2-locus ($i = 9$), and because we assume that no females are released, $u_i^F = 0$ for all i.

The resulting system of differential equations (with time dependence of $J_i^S(t)$, $A_i^S(t)$, and $B_i(t)$ omitted for simplicity of notation) is

$$\frac{dJ_i^S}{dt} = \frac{1}{2}\gamma_i^{E,S} B_i - J_i^S(\alpha J)^{\beta-1} - \mu_J J_i^S - \upsilon J_i^S \tag{3.2}$$

$$\frac{dA_i^S}{dt} = \upsilon\gamma_i^{L,S} J_i^S - \mu_S A_i^S + \mu_i^S, \tag{3.3}$$

for $i = 1..N$ and $S = F$ or M. All model parameters are listed in Table 3.3 and are based on those used for *Ae. aegypti* by Robert et al. (2013). While the rates of mortality, larval production, and emergence to adulthood apply to *Ae. aegypti*, the resulting population dynamics, simulated with different strengths of density dependence, would likely be similar to many other species. All numerical simulations of the differential equations have initial conditions at the wild-type equilibrium.

In order to explore the effects of demographic stochasticity and genetic drift, we also run simulations using an analogous continuous time Markov chain model (for details, see Appendix B.1).

3.4 Results

Each of the FK (female-killing) and BK (bisex-killing) genetic strategies has the goal of causing the population to decline by reducing the number of reproductive adult females. The strength of density-dependent mortality moderates the reductions in population size because stronger density dependence (higher β) causes the juvenile mortality rates to decrease more quickly as population size decreases from equilibrium. In a system with strong density dependence, the weekly release ratio (r) must be larger to achieve the same amount of population suppression as in a system with weak density dependence. Large r can result in target population extinction. In contrast, small r results in a new, lower equilibrium population density, where the proportion of individuals that die due to bearing the transgene is not high enough to outweigh the increased survival of juveniles due

Table 3.3 Model parameters.

Parameter	Description	Value/range	Reference
μ_J	Density-independent per-capita juvenile mortality rate	0.03 day^{-1}	Rueda et al., 1990
μ_M	Adult male per-capita mortality rate	0.28 day^{-1}	Muir and Kay, 1998, Fouque et al., 2006
μ_F	Adult female per-capita mortality rate	0.10 day^{-1}	Muir and Kay, 1998, Fouque et al., 2006
λ	Female per-capita larval production rate	8 day^{-1}	Harrington et al., 2001, Styer et al., 2007
ν	Per-capita rate of emergence to adulthood	0.14 day^{-1}	Muir and Kay, 1998
β	Strength of density dependence (see text)	2 to 4	
$A_1^F(0)$	Equilibrium number of wild-type females	2000	
r	Weekly release ratio of transgenic: equilibrium wild-type males	0 to 12	
α	Density dependence parameter chosen based on β and $A_1^F(0)$	Dependent on β and $A_1^F(0)$	
w_i^H, w_i^M	Egg hatching and male mating competitiveness fitnesses of genotype i	See Tables 3.1 and 3.2	
s^H, s^M	Egg hatching and male mating competitiveness fitness costs to homozygotes	0 to 1 (0.2 and 0.1 default, respectively)	
h	Fitness cost degree of dominance (percent of s^x incurred in heterozygotes)	0 to 1 (0.5 default)	
c_A	Proportion of s contributed by A (2-locus)	0.5 to 1 (0.55 default)	
$\gamma_i^{E,S}, \gamma_i^{L,S}$	Viabilities of genotype i for early (E; embryonic) and late (L; adult) approaches in sex S	See Tables 3.1 and 3.2	

to decreased density-dependent mortality in the smaller population.

Figure 3.1 demonstrates the outcome of release for each genetic approach at $r = 1$ under the deterministic model (see Figure B.1 for time series with multiple release ratios). At this release ratio, approaches that are late-acting (i.e., mortality in pre-mated adults, indicated by "L-") reduce the number of viable females to a lower number than approaches that are early-acting (i.e., mortality in the embryonic life stage, indicated by "E-"). BK approaches reduce the number of females faster and to lower levels than female-specific approaches with late-acting mortality, but the opposite is true for early-acting mortality. Among FK approaches, 1-locus reduces the number of females more quickly initially, but 2-locus eventually suppresses the population slightly more than 1-locus. Overall, this suggests that L-BK is most effective, followed by L-FK, E-FK, and E-BK, with little difference between 1- and 2-locus FK. Under stronger density dependence (higher values of β), releases result in weaker suppression of the population, with none of the approaches causing extinction of the population if $\beta = 4$.

In general, large release ratios result in extinction (the population goes to an equilibrium size of zero), and small release ratios result in a suppressed but non-zero equilibrium population size. For most sets of parameters, there is a critical release ratio, r_c, above which the release is large enough to cause the population to go extinct. With such a large number of released males, population extinction is the only stable equilibrium, meaning the release will cause extinction regardless of initial population size. For ongoing release at release ratios below r_c, there is a non-zero stable equilibrium for the number of viable adult females, meaning that release will not push a wild-type population to extinction. A population size of zero is also stable, meaning the release could protect against re-invasion of the wild-type if continued after population extinction, but the system will approach the non-zero equilibrium unless starting from very low population sizes, i.e., it is a bistable system with a low invasion threshold (as shown in Figure B.2 with time series starting from multiple initial conditions). Mathematical details on the shift in qualitative behavior at r_c and analysis using Mathematica Version 12.0 (Wolfram Research, Inc., 2019) can be found in Appendix B.2 and Figure B.3.

Figure 3.2A shows how release ratios affect stable equilibria when $\beta = 3$. As a result of the bifurcation, each of the lines is discontinuous, jumping from a non-zero equilibrium to a zero equilibrium at r_c. The smallest release size required to cause extinction for each approach is $r_c^{LBS} =$

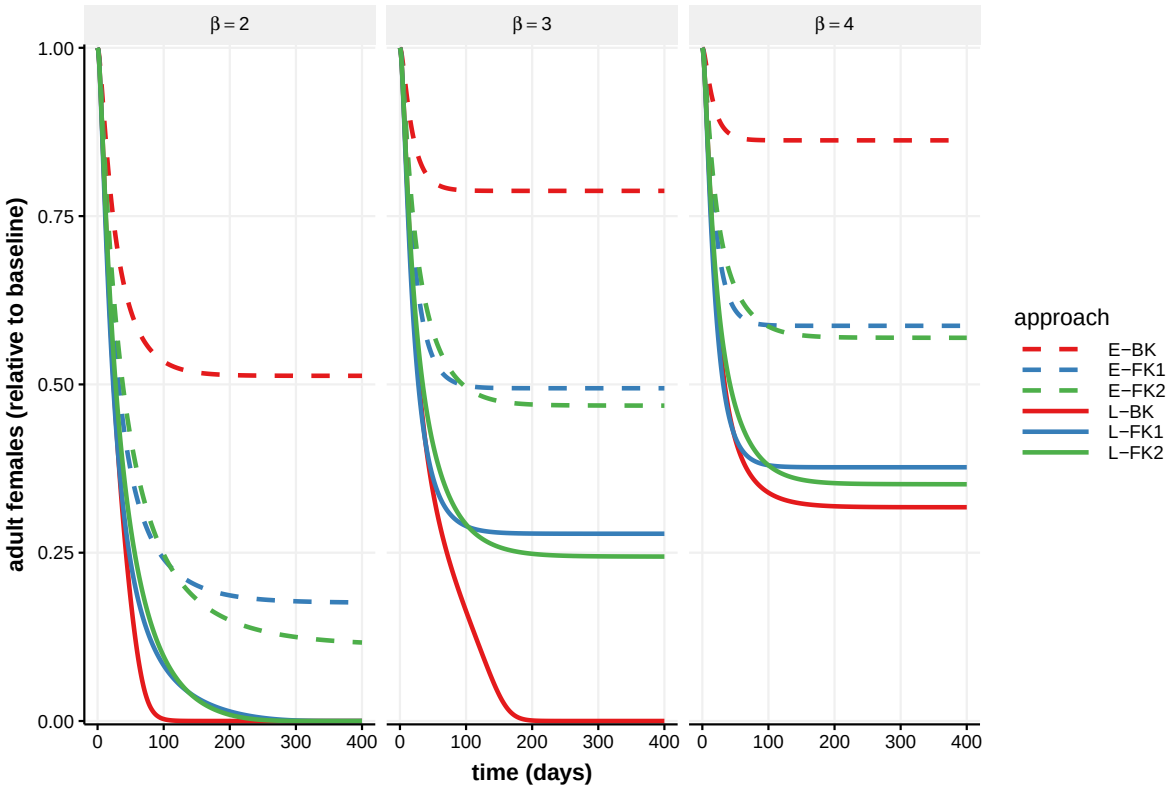

Figure 3.1 Effect of transgenic releases on population size over time for various strengths of density dependence. The number of viable adult females (relative to pre-release equilibrium) over time is plotted for deterministic simulations with adults for each genetic approach released at a continual weekly release ratio of 1:1 transgenic males to the pre-release equilibrium wild-type males ($r = 1$). Line type differentiates embryo (early-acting, dashed line) and adult (late-acting, solid line) mortality, and line color differentiates bisex (red), 1-locus female-killing (blue), and 2-locus female-killing (green) constructs. Releases are less effective as strength of density dependence (β) increases across panels from left to right. Simulations use fitness parameters $s^H = 0.2$, $s^M = 0.1$, and $c_A = 0.55$ and the remaining parameters as listed in Table 3.3.

$0.86 < r_c^{LFK2} = 2.52 < r_c^{LFK1} = 2.86 < r_c^{EFK2} = 4.74 < r_c^{EFK1} = 5.40 < r_c^{EBS} = 5.97$. This order of effectiveness matches that of Figure 1 and mirrors the results for single locus constructs in Gentile et al. 2015. However, when r is smaller than r_c^{LBS} and extinction does not occur, L-BK has a higher equilibrium population size than either of the late FK methods. There is a small relative difference between the different FK approaches, with similar equilibrium sizes when r is below the critical release ratios and 1-locus FK requiring a release ratio less than 15% larger than 2-locus FK to cause extinction.

In settings where release causes the population to go extinct, we can consider the time it takes to reach extinction after starting release (Figure 3.2B). Given that deterministic simulations will only approach extinction asymptotically, we use the time it takes for the number of females to reach less than 0.05% of the pre-control equilibrium, which is suppression to below 1 adult female when starting from a pre-release equilibrium of 2000. The times using this threshold are comparable to the average time to extinction in stochastic simulations (see Figure B.4). Once release ratios are high, L-BK drops the population under 0.05% of the equilibrium faster than late FK methods. Also, both of the 1-locus FK approaches are slightly faster than the respective 2-locus FK approaches, though the differences are small for practical purposes. At $r = 12$, for example, the times are 86 (L-BK), 114 (L-FK1), 119 (L-FK2), 133 (E-BK), 149 (E-FK1), and 153 (E-FK2) days. The results are similar for $\beta = 2$ (see Figure B.5).

Overall, FK1 and FK2 are quite similar. If either of the A or B alleles in 2-locus FK becomes fixed in the population, the 2-locus approach becomes nearly identical to 1-locus FK, where one copy of the unfixed allele causes mortality in females. For example, if all individuals in the population already have the B allele, only one copy of the A component is additionally necessary, just as a single copy of the K allele causes mortality. In this case, the long-term equilibrium can be identical to 1-locus FK, though a fitness cost to the fixed allele decreases the average fitness of the entire population and makes the population size lower for 2-locus FK than 1-locus FK. Whether one of the 2-locus FK alleles become fixed depends on the fitness costs and release ratios of the system (see Appendix B.3, Figure B.6, and Figure B.7).

An observation from Figure 3.2A is that L-BK, L-FK2, and L-FK1 cause a similar level of suppression when the release ratio is near $r = 0.8$. Previous work has suggested that L-FK allows the wild-type allele to propagate in heterozygous males, making it less effective than L-BK (Gentile

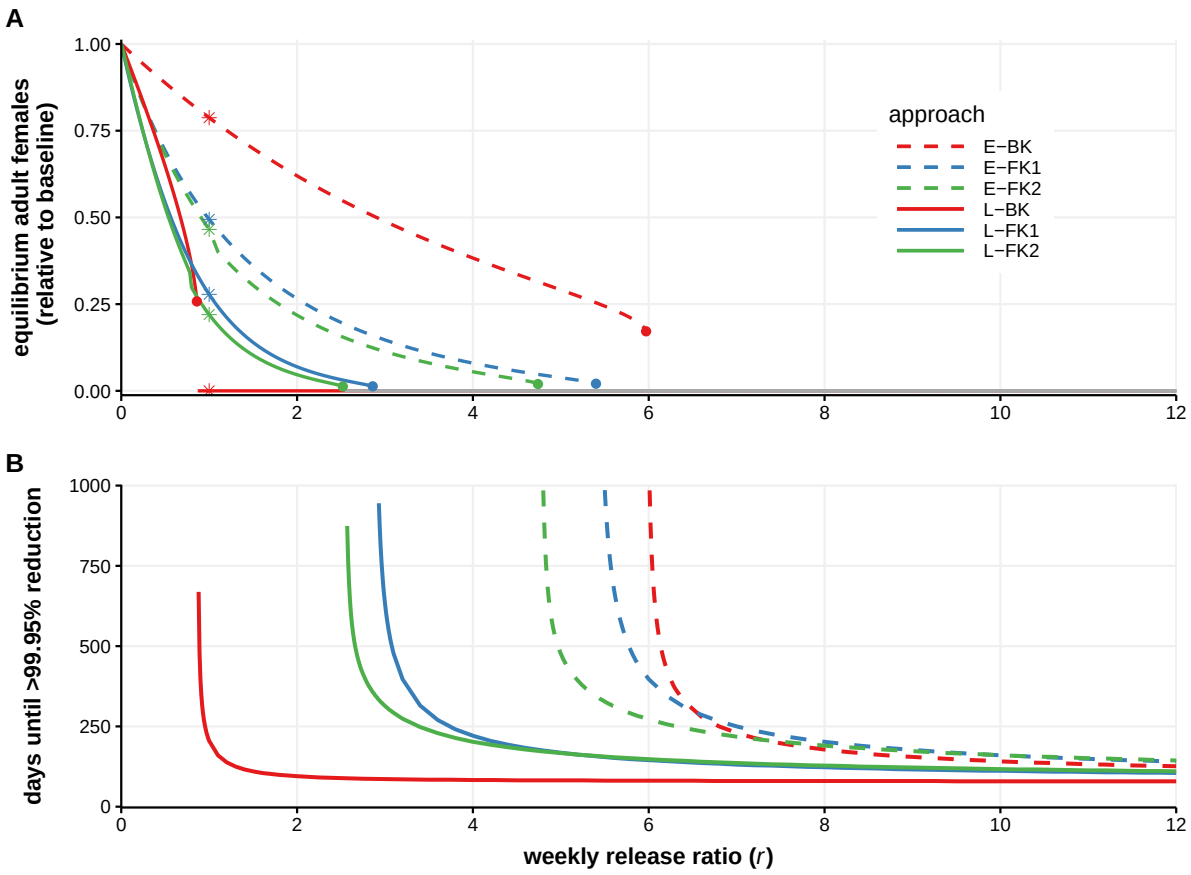

Figure 3.2 Release outcomes across different release ratios. A: Long-term, stable equilibria for number of viable adult females (relative to pre-release equilibrium) for different r, found by simulating the system of differential equations until at steady-state. The asterisks indicate $r = 1$, for which the equilibrium values correspond to the middle panel of Figure 3.1. Each genetic approach exhibits a bifurcation at a critical release ratio, rc, indicated by a solid circle. Above the r_c of each approach, that approach only has an equilibrium at zero. This makes each line have a discontinuity, visible with the red line (L-BK). Above $r = 2.52$, two or more approaches lead to extinction of the population and hence have equilibria at zero: this is indicated using a grey line. Note that below r_c, each system exhibits bistability, with both non-zero and zero count stable equilibria. Only the non-zero equilibrium is plotted, which is the equilibrium reached when starting simulations from the wild-type equilibrium. **B:** Time until the number of viable adult females is under 0.05% of equilibrium in deterministic simulations for different r. Color and line type match that of Figure 3.1. Parameters are $\beta = 3$, $s^H = 0.2$, $s^M = 0.1$, and $c_A = 0.55$, as in Figure 3.1.

et al., 2015). This is true at high release ratios, but not at all release ratios. Without fitness costs, the strategies are equally effective if the number of released males is equal to the number of wild-type males in the population at equilibrium: heterozygotes carry both wildtype and transgenic alleles, and therefore the survival of heterozygous males does not affect allele frequency. At small release ratios, when L-BK release results in low transgenic frequency and thus a large equilibrium population size, survival of heterozygous males would allow the transgenic allele to propagate further and increase in frequency, explaining why L-FK has a lower equilibrium than L-BK in this narrow window of small releases.

The main difference exhibited between FK1 and FK2 can also be explained by their propagation of the transgenic and wild-type alleles. When the components are separated across two loci, the A and B alleles become unlinked, with some individuals only inheriting one allele or the other, while having linked components guarantees inheritance of the transgenic allele and reduces the population size more quickly initially. Eventually, however, the accumulation of transgenic alleles in the 2-locus system causes production of a higher proportion of unviable genotypes and greater population suppression (Figure 3.2).

Results over a wide range of fitness parameters highlight the minimal differences between FK1 and FK2. We use the time it takes to suppress the number of viable females to under 0.05% of the equilibrium when there is a high release ratio of $r = 7$ as a way to measure effectiveness (Figure 3.3). For each approach and degree of dominance (h), there is a region with low fitness costs where there is fairly little variation between the time it takes to suppress the population below the threshold. As one or both of the fitness costs increase, there is a margin of longer times separating the successful suppression region from the unsuccessful region. As evident from Figure 3.2, long times to reduce the population indicate that the release ratio of $r = 7$ is only slightly greater than r_c for those parameter values, and a system with even higher fitness costs have $r_c > 7$, so suppression below 0.05% of equilibrium is not achieved.

There is little difference between FK1 and FK2, particularly for L-FK, since the release ratio is much higher than rc in most of the region with successful release. Even with different fitness costs, FK1 and FK2 do not differ drastically in time until extinction; FK2 would be substantially less effective only if the costs become large enough for the release to become too small to cause extinction (i.e., r becomes smaller than r_c because of the higher costs). Unpacking the differences

between all approaches, for the E-BK approach, neither h nor s^H affect the length of time for reduction; only released males experience costs so the degree of dominance does not affect results, and all offspring that inherit a transgene are inviable so hatching costs do not affect results. For E-BK, increasing the cost to male mating competitiveness drastically decreases release efficacy because doing so effectively reduces the release ratio (fewer of the released males successfully mate and contribute to population reduction), and with no surviving transgenic offspring, the release ratio is the main contributing factor to E-BK efficacy. When $s^H = 1$ and $h = 1$, none of the offspring survive for the E-FK1 and E-FK2 approaches, making the times shown identical to E-BK. Decreasing these parameters makes a difference for FK approaches because the males are subject to fitness costs. Dominant fitness costs (bottom row) actually enable successful population reduction for higher male mating costs. With large, dominant mating costs, FK approaches have few mating adult males, becoming effectively similar to the highly effective L-BK, which has no adult males. The full explanation for increased effectiveness with higher degrees of dominance relates to the propagation of wild-type alleles, similar to the previous description: homozygote males that are being released are affected regardless of h, and when $h > 0$, the fitness cost also reduces the spread of wild-type alleles in heterozygotes. This effect is also evident in other transgenic systems, such as two-locus underdominance, where dominant transgenic allele fitness costs prevent the wild-type from being maintained in the population at small frequencies (Dhole et al., 2018).

3.5 Discussion

The recent literature on FK systems makes the assumption that strains built with constructs inserted at two independent loci will not be as useful for field releases as those built with a single construct. The assumption is that the two constructs will separate from each other in the second generation after a release and will become non-functional. Our modeling results demonstrate that a 2-locus FK (FK2) should behave similarly to a 1-locus FK (FK1) and would not present any significant disadvantages in its ability to suppress a population. We generally made the assumption that the 2-locus and 1-locus approaches would have similar total fitness costs because they have the same components. If the total cost of the 2-locus approach was less than for the 1-locus approach, the 2-locus approach would likely be preferred. The reverse also holds. Importantly, based on our results,

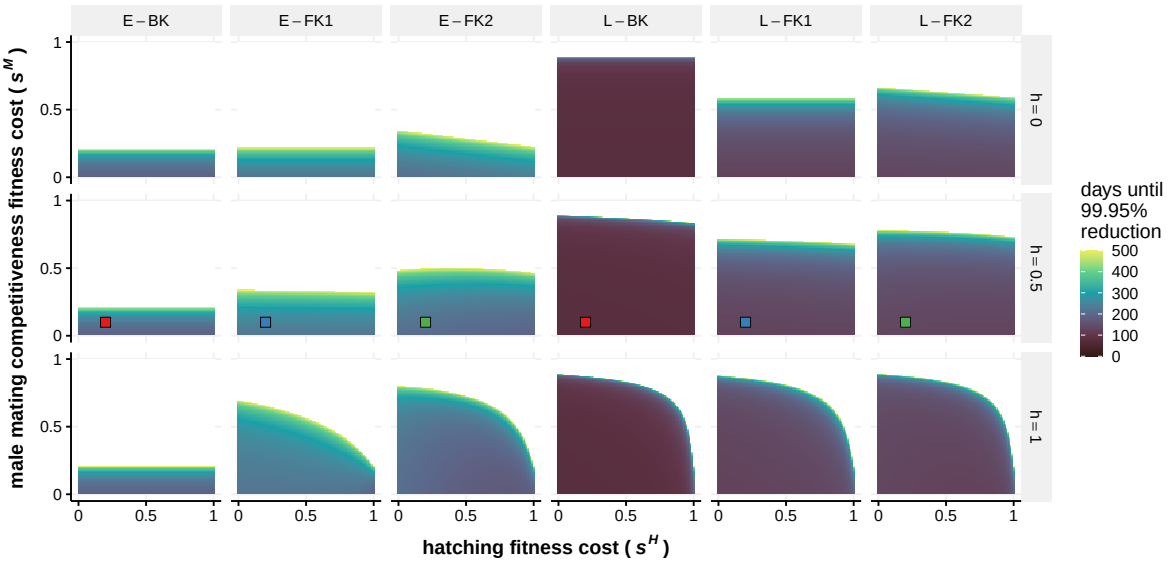

Figure 3.3 Effects of fitness cost variation on release efficacy for weekly release ratio $r = 7$ and $\beta = 3$. Each column of panels shows the results for a different genetic approach, while each row of panels depicts a different degree of dominance, h. Within each individual panel, the hatching fitness cost, s^H, increases from 0 to 1 along the x-axis, and the cost to male mating competitiveness, s^M, increases from 0 to 1 along the y-axis. For every point, a deterministic simulation was run with a unique combination of genetic approach and fitness parameters, and color indicates the number of days until the number of viable adult females is under 0.05% of equilibrium. Darker colors show faster times, with a minimum time of 73 days, and lighter colors show slower times up to 500 days. White areas indicate that the number of females did not fall below the threshold within 500 days. The colored points in the middle row correspond to the times in Figure 3.2B at $r = 7$.

there is no a priori, general reason for genetic engineers to favor a 1-locus system. The choice will likely depend on specific biological and genetic characteristics of the target species.

Assuming equal costs, FK1 is slightly faster at initial population reduction, but FK2 can eventually suppress the population to lower numbers. FK2 also has a slightly lower critical release ratio than FK1, meaning a smaller release size is necessary to guarantee extinction. For many combinations of fitness costs and release ratios, one of the FK2 alleles would be driven to fixation, resulting in a genetic system similar to FK1. The differences between FK1 and FK2 are much smaller than between FK and BK approaches. Comparing FK and BK approaches, our results are generally similar to previous work (Gentile et al., 2015). Late acting approaches cause extinction with a lower release ratio than early acting approaches, with L-BK causing extinction with a lower release ratio than L-FK, and E-FK causing extinction with a lower release ratio than E-BK.

While our modeling results indicate that L-BK outperforms the other methods, there are other considerations that affect which approach may be best suited for a given scenario. In our model, parameterized for mosquitoes, density-dependent mortality during early life-stages was an important factor and caused early-acting approaches to result in less population reduction than late-acting approaches. In species with little density-dependent dynamics in juveniles, the difference in effectiveness between early and late-acting would be minor, though this is not the case for many pest species.

Beyond population dynamics, there are economic and social factors that differ between approaches. For some systems, it will be necessary to engineer constructs into lab strains and then backcross the construct or constructs into a strain that have a genetic makeup similar to the targeted population. In general, it should be easier to do the backcrossing with a one-locus system. Rearing costs are also expected to vary between approaches. With E-FK, juvenile females experience mortality before consuming food, whereas E-BK, L-BK and L-FK require rearing of juveniles of both sexes. Furthermore, BK approaches typically require sexing to remove females prior to release, which increases the total rearing costs and is often difficult to do with complete accuracy. When releasing a species that is a disease vector, sexing accuracy is meaningful from a social perspective as release of females could contribute to disease transmission. Though it would not apply for many species of interest, in a system where releasing females is acceptable because neither the females nor their larvae can cause damage, L-BK could be more effective when releasing females because the released

females' progeny would increase density-dependent competition in larvae, then die before reaching adulthood.

Apart from engineering and rearing, for most agricultural pests, the juvenile stages of males and females cause damage to crops and livestock. In the first generations of transgenic pest releases, the late acting approaches will leave feeding immatures in the environment, and E-FK will result in male immatures that still cause damage. This may not be favored by farmers even though the overall population could be decreasing rapidly, and E-BK could be preferred. Finally, even if late-acting mortality may be ideal for a given scenario, controlling the timing of mortality at the intended life stage may not always be feasible, for example due to leaky expression of the lethal gene.

The model used here has several limitations. An important factor that could affect population genetics is spatial heterogeneity. For example, in a spatial model of FK2, it would be possible, particularly in small populations, for different patches to have different transgenic alleles reach fixation. A spatial model would also be useful to determine if FK2 has any differences in resilience to wild-type reinvasion. The details of such a spatial model, including rates of release, would depend on species. A species-specific model could also implement different forms of density-dependence, age-structure, and mating parameters. Finally, given its generality, our model does not account for any potential mechanisms for resistance development. Depending on the mechanism of lethality, there may be advantages for having both components for lethality inserted together. While these areas require further investigation, our results indicate that overall, there is little difference in the pest population suppression efficacies of 1-locus FK and 2-locus FK.

3.6 Acknowledgements

This research was funded by the Research Training Group in Mathematical Biology, funded by NSF grant RTG/DMS-1246991 (MRV and ALL), NSF IGERT grant 1068676 (MRV, FG and ALL), NIH grant 1R01AI139085-01 (FG and ALL), and the NC State Drexel Endowment (ALL). We thank M.J. Scott, S. Dhole, B. Hollingsworth, J. Baltzegar, and J. Sudweeks for comments and helpful discussion.

Chapter 4

Inference of selection coefficients for insecticide resistance in Iquitos, Peru[1]

[1]The methods and results of this chapter will form part of a manuscript currently in preparation: Baltzegar J, Vella MR, et al. Evolution of Knockdown Resistance (kdr) Haplotypes in Response to Pyrethroid Selection in *Aedes Aegypti*. I conducted all analyses included here and wrote this chapter.

4.1 Abstract

The use of chemical insecticides is one of the most commonly used methods to suppress insect populations and reduce the spread of vector-borne diseases. However, the effectiveness of insecticides has been hindered by the development of resistance to insecticides. Much research has been conducted in the lab to determine characteristics of resistant alleles. For example, studies have suggested that fitness costs against the susceptible, wild-type allele are dominant, where phenotypic resistance is not induced by a single copy of a resistant allele. The degree of dominance is an important quantity affecting the ability of a resistant allele to increase in frequency in a population. Recently published data shows an increase in frequency of resistant alleles in *Aedes aegypti* in Iquitos, a city in Northern Peru, and analysis of the data suggests that the costs were closer to recessive than dominant, unlike lab experiment results. Here, we construct a hidden Markov model to conduct inference. We model the genotype frequencies in the population, accounting for random genetic drift, and we allow for an overdispersion in sampling by using a Dirichlet-Multinomial distribution. Inference using particle Markov chain Monte Carlo indicates that a single resistant allele is enough to substantially lessen the negative effects from the insecticide, strengthening the claim that costs are not dominant. Given the flexibility in model specification allowed by particle Markov chain Monte Carlo, our approach may be applicable to other problems concerning inference fitness cost.

4.2 Introduction

Vector-borne diseases such as malaria, lymphatic filariasis, and dengue cause substantial risk to human health throughout the world, with 80% of the world at risk from at least 1 disease and over 700,000 deaths annually (World Health Organization, 2017). For dengue alone, the global cost was estimated to be $8.9 billion in 2013, with an estimated 390 million infections, of which 96 million were apparent (Bhatt et al., 2013). It has proven to be difficult to find an effective treatment or develop an effective and safe vaccine for dengue. This is in part due to there being at least four serotypes of the virus that cause infection (Sridhar et al., 2018). In the absence of an effective way to treat or prevent the disease, many efforts to control the spread of the disease have focused on its vector.

The dengue virus is transmitted primarily by the mosquito *Aedes aegypti*, which is also a vector for the Zika, yellow fever, and chikungunya viruses. *Ae. aegypti* is well adapted to urban environments and exhibits several traits that make efforts to suppress the population size challenging. The species has high reproductive capacity, is able to use standing water from small sources (e.g., aluminum cans) as breeding sites, and can quickly adapt to different environmental conditions (Carvalho and Moreira, 2017). While new approaches, including *Wolbachia* and releases of transgenic mosquitoes, show promise as control methods, they are still under development and would likely be used in an integrated approach along with the use of chemical insecticides (Alphey et al., 2013).

Large-scale, chemical-based eradication campaigns have been the primary control method used since at least the 1940s, when DDT was shown to be a powerful tool in reducing insect populations (Camargo, 1967). Major concerns about DDT's negative impact on the environment and human health eventually halted its use as an insecticide. However, the eradication potential of DDT was also limited by mosquitoes resistant to the chemical's killing mechanism, which prevented eradication in areas and also reinfested previously eradicated areas (Camargo, 1967). More recently developed insecticide classes include organophosphates, pyrethroids, chitin synthesis inhibitors, and juvenile hormone analogues (Carvalho and Moreira, 2017). Pyrethroids, a widely used class of insect neurotoxins, function by inhibiting inactivation of voltage-gated sodium channels, leading to paralysis and death (Du et al., 2016). Resistance to pyrethroids is widespread and has been found to be the result of a variety of different single nucleotide polymorphisms in the targeted sodium channel protein, such as the F1534C mutation, named for the position of the mutation on the protein and the amino acids of the mutation (Du et al., 2016). Interestingly, many of these mutations are associated with knock-down resistance (kdr) to both pyrethroids and DDT (Brengues et al., 2003).

Ae. aegypti has been well studied in Iquitos, Peru, where the species was considered eradicated by the late 1950s, had returned by 1984, and was large enough in population size to cause a dengue outbreak in 1990 (Phillips et al., 1992). As part of various entomological and epidemiological research efforts, mosquitoes have been collected from Iquitos since the 1990s, while there were a series of city-wide spraying campaigns of pyrethroid insecticides from 2002-2014 (Getis et al., 2003; Morrison et al., 2004; LaCon et al., 2014; Gunning et al., 2018). Baltzegar (2020) later genotyped collected mosquitoes for specific kdr mutations and among other results, showed a temporal increase in the F1534C allele frequency in the population.

Details about how susceptible alleles are selected against compared to resistant alleles are of broad interest. An important factor for how a resistant allele spreads in a population is the extent that heterozygotes (with one copy of the allele) are resistant to insecticides compared to homozygotes (with two copies). For example, if the heterozygote is mostly susceptible to insecticides, the resistant allele would take much longer to initially increase in frequency than if the heterozygote is mostly resistant (Conner and Hartl, 2004). Bioassays based on the phenotypic response to insecticide exposure have shown that mosquitoes heterozygous for pyrethroid resistance are still mostly susceptible despite having a copy of the resistant allele (e.g. Saavedra-Rodriguez et al., 2007, Fan and Scott, 2020). This means a dominant fitness cost, (i.e., lower relative survival and reproductive success), to the susceptible allele. This could make sense mechanistically if a single copy of the susceptible allele is enough for the neurotoxin to cause death. However, analysis of the Iquitos data by Baltzegar (2020) shows that the cost is closer to being recessive, with one resistant allele being enough to give mosquitoes insecticide resistance. Baltzegar notes that findings from real-world data could differ from findings based on laboratory conditions for several reasons, including complexities of phenotypic resistance that extend beyond kdr mutations (Smith et al., 2019).

There are many existing approaches for the analysis of allele frequency data. Bank et al. (2014) and Tataru et al. (2017) give overviews of methods for inference of fitness costs. The majority of approaches used to analyze allele frequencies are based on the Wright-Fisher model, which makes simplifying assumptions about mating and allele inheritance (Tataru et al., 2017). A natural setup for inference is to treat the system as a hidden Markov model (HMM), where the true allele frequencies (unobserved, hidden states) advance forward in time according to the model, and observations (data) are assumed to be sampled from those hidden states. Parameter estimation with maximum likelihood estimation (MLE) or many Bayesian approaches rely on computing sampling likelihoods integrated over all possible allele frequencies. This calculation is frequently facilitated by approximating the Wright-Fisher model as a diffusion process (e.g., Malaspinas et al., 2012; Mathieson and McVean, 2013; Schraiber et al., 2016; He et al., 2020) and assuming that the samples are distributed binomially (or multinomially in the case of multiple alleles). Alternatively, calculation of the likelihood can be avoided by using simulation-based methods such as approximate Bayesian computation (ABC). Foll et al. (2015) establish one such approach, Wright-Fisher ABC (WFABC), and demonstrate its effectiveness for parameter estimation. Baltzegar (2020) uses WFABC to analyze the

Iquitos data, but given the surprising result of recessive costs, additional analysis with a different approach is warranted.

Here, we perform Bayesian inference on the Iquitos data to understand the fitness parameters of the F1534C mutation, with particular interest in determining whether susceptibility to insecticide spraying is dominant or recessive. We use a genotype frequency model based on Wright-Fisher assumptions to build an HMM. To account for non-random experimental sampling, we allow for sampling variance to be larger than that of a multinomial distribution through the use of a Dirichlet-multinomial distribution. Using particle Markov chain Monte Carlo (pMCMC) to produce parameter posterior distributions, we find that susceptibility is closer to recessive than dominant, supporting the previously conducted analysis of the data. While the computation required for our approach would likely make it ill-suited for data extending over many hundreds of generations, it could potentially be used for inference on many types of short-term experiments.

4.3 Methods

4.3.1 Mathematical model of population genetics

We use a genotype-frequency model as the basis to evaluate the spread of the F1534C mutation in Iquitos, Peru. Our model assumes the population is well-mixed with random mating and operates on discrete (non-overlapping) generations of mosquitoes, with each generation lasting one month. Populations of *Ae. aegypti* are known to have high degrees of spatial clustering and age structure (Getis et al., 2003; LaCon et al., 2014). Some of this structure can be captured in models with more complexity (Magori et al., 2009; Huang et al., 2009), but we use a simple model as it is generally easier to parameterize and use for inference. We also assume the population is contained, without immigration or emigration of mosquitoes.

We let allele R represent the F1534C mutation resistant allele, while allele S represents the wild-type, susceptible allele. The frequencies of each possible genotype are denoted by X_{RR}, X_{SR}, and X_{SS}, with $\vec{X}(t)$ denoting the vector of all three genotype frequencies at generation t. For brevity, we omit the vector notation from all vector variables as specific elements of the vector will include a subscript for genotype i. The frequency of the R allele in the population is expected to be low until spraying begins because the mutation would likely confer a fitness cost (e.g., proportionately less

individuals successfully mate or individuals contribute proportionately fewer offspring) in absence of spraying. Samples from before spraying began indicate that the R allele was maintained at a low frequency, which could be due to mutation/selection balance or pre-existing resistance from DDT when the population was re-established (Baltzegar, 2020). For simplicity, we do not consider de novo mutations and assume the initial genotype frequencies begin at Hardy-Weinberg equilibrium (see Conner and Hartl, 2004 for background). Letting $p(t)$ denote the genotype probabilities of offspring entering generation t, we can calculate $p(1)$ from the starting allele frequency of the resistance allele, R_0:

$$p(1) = \left(p_{RR}(1) = R_0^2,\ p_{SR}(1) = 2R_0(1-R_0),\ p_{SS}(1) = (1-R_0)^2\right). \tag{4.1}$$

While there is spraying, mosquitoes with an S allele suffer a loss in fitness. Specifically, assuming RR is the favored genotype while under selection, SS and SR individuals have proportionately fewer offspring. By using a discrete-generation model, fitness is purely a measure of the mosquitoes' contributions to the mating pool, and it does not matter whether the fitness costs manifest as larval death, adult death, or fewer offspring. The fitness cost to SS individuals relative to RR mosquitoes is denoted by s, and the fitness cost to SR individuals is $h \cdot s$, where h is the degree of dominance of the cost to the S allele.

Using the current frequencies, $X(t)$, the genotype probabilities of the next generation, $p(t+1)$, can be calculated based on the probabilities of each of the six possible matings. The assumption of a well mixed and randomly mating population allows the probability of a given mating to be calculated as the product of the pair of genotype frequencies. Each pairing of different genotypes is multiplied by a factor of two to account for the two possible couplings of males and females. The offspring probabilities from each mating are calculated assuming Mendelian inheritance. Fitness costs, as described above, result in a relative reduction in the SR and SS frequencies. The resulting set of difference equations to calculate $p(t+1)$ from $X(t)$ (omitting dependence on time for brevity) is:

$$\begin{aligned}
p_{RR} &= \left(X_{RR}^2 + 0.25X_{SR}^2 + X_{RR}X_{SR}\right)\frac{1}{\overline{w}} \\
p_{SR} &= (1-hs)\left(2X_{RR}X_{SS} + X_{RR}X_{SR} + X_{SS}X_{SR} + 0.5X_{SR}^2\right)\frac{1}{\overline{w}} \\
p_{SS} &= (1-s)\left(X_{SS}^2 + 0.25X_{SR}^2 + X_{SS}X_{SR}\right)\frac{1}{\overline{w}},
\end{aligned} \tag{4.2}$$

where the mean fitness, $\overline{w}$, is the sum of the un-normalized frequencies and normalizes so that $\sum_i p_i = 1$.

4.3.2 Genetic drift

Regardless of population size, the expected value of $X_i(t+1)$ is equal to $p_i(t)$. However, the relative variance of $X_i(t+1)$ increases with decreasing population size due to randomness involving surviving and mating. This effect is termed genetic drift. The amount of drift is characterized by the effective population size, N_e. N_e is often much smaller than the total (census) population size because it accounts for factors in population structure that decrease genetic diversity, such as non-random mating and variability in number of offspring per female (Saarman et al., 2017). Stochasticity from genetic drift can thus be incorporated into our model by assuming a mosquito count equal to N_e in each generation. We let the genotype counts in the next generation, $C(t+1)$, be distributed multinomially based on the probabilities from (4.2). Then we have:

$$\begin{aligned} C(t+1) &\sim \text{Multinomial}\big(n = N_e, \text{prob} = p(t)\big) \\ X(t+1) &= \frac{C(t+1)}{N_e}. \end{aligned} \tag{4.3}$$

As in Baltzegar (2020), we assume $N_e = 500$ in each generation based on the analysis done by Saarman et al. (2017).

The density function $X(t+1)|X(t), \theta \sim f(X(t+1)|X(t), s, h, N_e)$, for $t \in [1, \ldots, T]$, represents the iteration from generation t to $t+1$ as described in (4.2) and (4.3), where T is the total number of generations. Here, $f(\cdot)$ is a Markov process as it depends only on the current state of the system $X(t)$, and not previous states, as shown in Figure 4.1.

4.3.3 Sampling

The experimental sample size varies between months. We denote the sample size in generation t by $N_s(t)$. We assume the sampling distribution at generation t is independent of sampling in previous generations, meaning there is only dependence on the state in that generation, $X(t)$. Additionally, while sampling removes mosquitoes from the population, we assume the resulting effect on the population genetics is minor enough to ignore.

If we were to assume that samples of individual mosquitoes in a given generation, n_k, $k \in [1,\ldots,N_s(t)]$ are independent and identically distributed (i.i.d.), then the sampling could be approximated by a multinomial distribution, $Y \sim \text{Multinomial}(n = N_s, \text{prob} = X)$ (with time dependence omitted for brevity). The sampling distribution would have the following properties:

$$f\left(y \mid X,\ N_s\right) = \frac{N_s!}{\prod(y_i)!}\prod X_i^{y_i} \tag{4.4}$$

$$\mathbb{E}[Y_i|X,\ N_s] = N_s X_i \tag{4.5}$$

$$\text{Var}[Y_i|X,\ N_s] = N_s X_i(1-X_i)\,, \tag{4.6}$$

where $i=1$, $i=2$, and $i=3$ correspond to genotypes RR, SR, and SS, respectively.

However, random sampling poorly depicts the mosquito collection process. Sampled mosquitoes are not independent because there are often multiple mosquitoes collected from each sampled house. The samples are also not identically distributed because any spatial heterogeneity in genotype frequencies would result in different underlying sampling probabilities at different locations. Thus, we employ an overdispersed sampling distribution with added variance compared to a multinomial distribution. The Dirichlet distribution, which is a multivariate extension of the Beta distribution, can provide a means to such overdispersion. Given parameter $\vec{\alpha}$, where $\alpha_i > 0$, we let $Z \sim \text{Dirichlet}(\alpha)$, with properties:

$$f\left(z \mid \alpha\right) = \frac{\Gamma\left(\sum_i \alpha_i\right)}{\prod_i \Gamma(\alpha_i)}\prod_i z_i^{\alpha_i - 1} \tag{4.7}$$

$$\mathbb{E}[Z_i|\alpha] = \frac{\alpha_i}{\sum_i \alpha_i} \tag{4.8}$$

$$\text{Var}[Z_i|\alpha] = \frac{\frac{\alpha_i}{\sum_i \alpha_i}\left(1-\frac{\alpha_i}{\sum_i \alpha_i}\right)}{\sum_i \alpha_i + 1}\,, \tag{4.9}$$

where $\Gamma(\cdot)$ is the gamma function. The expected values are equal to the proportions of α_i, and the variances decrease with increasing $\sum_i \alpha_i$.

We consider drawing samples $Z(t)$ from a Dirichlet distribution parameterized by $\alpha = A \cdot X(t)$, where A scales $X(t)$, the vector of frequencies in the population at time t. Because α is proportional to $X(t)$, $\mathbb{E}[Z_i(t)] = X_i(t)$. Note that $\sum_i AX(t) = A\sum_i X(t) = A$. If we then let the data $Y(t) \sim$

Multinomial($n = N_s(t)$, prob $= Z(t)$), we can calculate the joint probability density of $Y(t)$ and $Z(t)$ (with time dependence omitted):

$$f\left(y, z \,\middle|\, X, A\right) = f\left(y \,\middle|\, z\right) f\left(z \mid X, A\right) \tag{4.10}$$

$$= \frac{N!}{\prod_i y_i!} \prod_i z_i^{y_i} \cdot \frac{\Gamma(A)}{\prod_i \Gamma(AX_i)} \prod_i z_i^{AX_i - 1} \tag{4.11}$$

$$= \frac{N!\,\Gamma(A)}{\prod_i y_i!\,\Gamma(AX_i)} \prod_i z_i^{y_i + AX_i - 1} \,. \tag{4.12}$$

The second product term in (4.12) is the kernel (i.e., the probability density function up to a multiplicative constant) of a Dirichlet distribution parameterized by $\alpha = y + AX$. We can write the probability density function as an integral over z which equals one, giving us:

$$\int_z \frac{\Gamma\left(\sum_i y_i + AX_i\right)}{\prod_i \Gamma(y_i + AX_i)} \prod_i z_i^{y_i + AX_i - 1} \mathrm{d}z = 1 \tag{4.13}$$

$$\frac{\Gamma(N_s + A)}{\prod_i \Gamma(y_i + AX_i)} \int_z \prod_i z_i^{y_i + AX_i - 1} \mathrm{d}z = 1 \tag{4.14}$$

$$\int_z \prod_i z_i^{y_i + AX_i - 1} \mathrm{d}z = \frac{\prod_i \Gamma(y_i + AX_i)}{\Gamma(N_s + A)} \,. \tag{4.15}$$

Now, writing 4.12 as the marginal probability distribution of Y with respect to Z and using 4.15 to re-write the integral, we get:

$$f\left(y \,\middle|\, X, A\right) = \int_z \frac{N!\,\Gamma(A)}{\prod_i y_i!\,\Gamma(AX_i)} \prod_i z_i^{y_i + AX_i - 1} \mathrm{d}z \tag{4.16}$$

$$= \frac{N!\,\Gamma(A)}{\prod_i y_i!\,\Gamma(AX_i)} \frac{\prod_i \Gamma(y_i + AX_i)}{\Gamma(N_s + A)} \tag{4.17}$$

$$= \frac{N!\,\Gamma(A)}{\Gamma(N_s + A)} \prod_i \frac{\Gamma(y_i + AX_i)}{y_i!\,\Gamma(AX_i)} \,, \tag{4.18}$$

which is the Dirichlet-multinomial distribution parameterized by AX. This distribution commonly arises when using a Dirichlet distribution as the prior for a multinomial distribution and has also been used for inferring genetic structure from genomic samples (e.g., Kitakado et al., 2006; Cowell, 2016). While the mean of the distribution is equal to the multinomial distribution, the variance is

$$\mathrm{Var}[Y_i(t) | X_i(t), A] = N_s(t) X_i(t) (1 - X_i(t)) \frac{N_s(t) + A}{1 + A} \,, \tag{4.19}$$

which is an inflation of the multinomial distribution variance given in (4.6) by the factor $\frac{N_s(t)+A}{1+A}$. Thus, A affects the amount of overdispersion, and the Dirichlet-multinomial distribution approaches a multinomial distribution as A approaches infinity. We let the Dirichlet-multinomial sampling density in (4.18) be given by $Y|X,\theta \sim g(Y|X,A)$. To allow for calculation of likelihoods when $X_i(t)=0$ but $Y_i(t)>0$, in practice we parameterize the distribution with $AX+0.0001$. This has a minimal effect on the properties of the distribution but accounts for the possibility of additional sampling error, e.g., if a sample was mislabeled. Sampling is visualized in context of the full HMM in Figure 4.1.

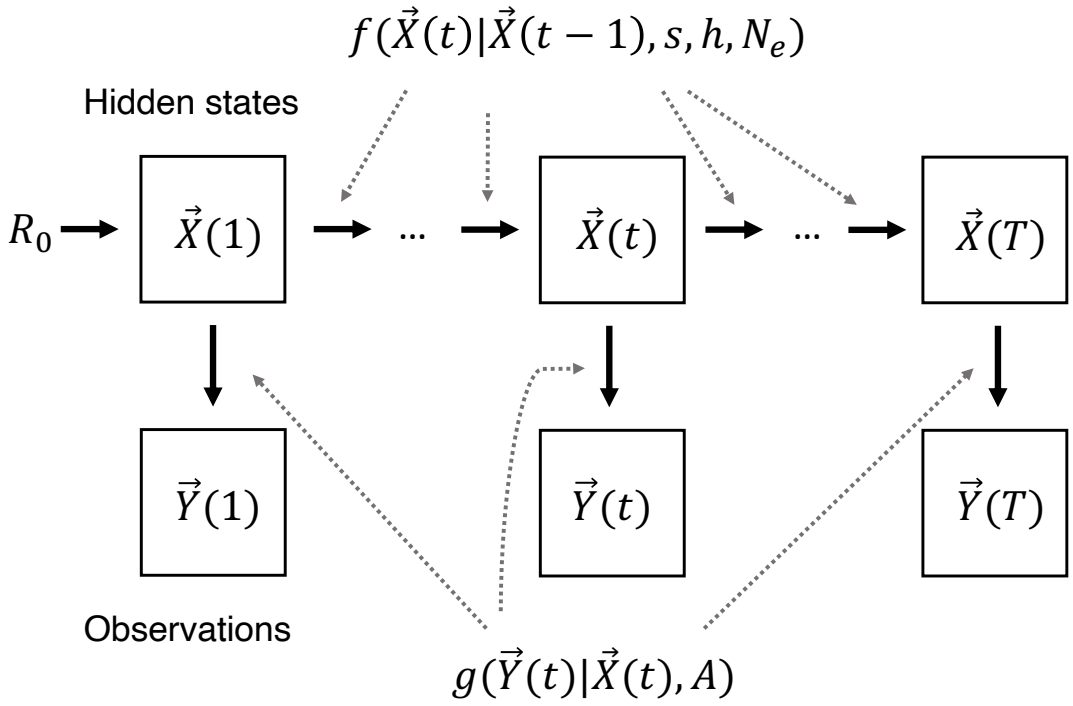

Figure 4.1 Hidden Markov model schematic. The true genotype frequencies in the population at time t, $\vec{X}(t)$, are governed by a Markov model with density $f(\vec{X}(t)|\vec{X}(t-1),s,h,N_e)$. At each generation, genotype observations $\vec{Y}(t)$ are generated by Dirichlet-multinomial sampling based on the true genotype frequencies in the population, with density represented by $g(\vec{Y}(t)|\vec{X}(t),A)$. The genotype frequencies in the first generation are generated based on the initial R allele frequency, R_0.

4.3.4 Model analysis

We conduct Bayesian inference on the HMM using particle Markov chain Monte Carlo (pMCMC). This approach uses a particle filter, running forward simulations to produce unbiased estimates of the marginal likelihood $\hat{L}\big(y(1:T)|\theta\big)$. These estimates are used in the calculation of the Metropolis-Hastings acceptance ratio, and the resulting Markov chain converges to the joint posterior distribution $f(x_{1:T},\theta|y_{1:T})$ (see Appendix C). We implement pMCMC with a multivariate normal proposal distribution and adaptive Metropolis-Hastings acceptance in R (R Core Team, 2019) using the package nimble (Valpine et al., 2017). For parameters in range $[0,1]$, we use uninformative priors $s \sim \text{Beta}(1,1)$, $h \sim \text{Beta}(1,1)$, and $R_0 \sim \text{Beta}(1,1)$. For A, we use the uninformative prior $A \sim \text{Gamma}(0.01, 0.01)$. To improve the time to convergence, we initialize the parameters using their maximum likelihood estimates, ignoring process error (i.e., $X(t+1) = p(t+1)$) and assuming $Y(t)$ is distributed multinomially.

4.3.5 Experimental data

As described above, mosquitoes were collected from Iquitos, Peru, over the course of over two decades and were genotyped for specific kdr mutations by Baltzegar (2020). The F1534C mutation, which we focus on here, was found in few mosquitoes in the year 2000 but in nearly every mosquito by the end of sampling in 2017. For our analysis, we let $t = 1$ correspond to October, 2002, and $T = 87$ correspond to December, 2009. Pyrethroid insecticides were used from 2002 until 2014, but we exclude data from past 2009 because another kdr mutation that begins to increase in frequency after 2009 may affect the fitness costs of the F1534C allele (Baltzegar, 2020). The sample size (of genotyped mosquitoes) in each month, $N_s(t)$, varied. Of the 87 months considered, samples in 35 months consisted of 10 or less mosquitoes, and samples from only 2 months consisted of over 100 mosquitoes.

4.4 Results

We first calibrate the particle filter to determine an appropriate number of particles to use for pMCMC. While the pMCMC algorithm is valid for any number of particles, small numbers of

particles can result in poor mixing, with the Markov chain getting stuck at values that happened to produce a high marginal likelihood estimate. The reason for this is evident in Figure 4.2, which demonstrates how likelihood estimates of the data are affected by the number of particles. If the filter uses few particles, the underlying dynamics of the system are less likely to be captured, resulting in high variances of marginal log likelihood estimates and lower average estimates. The variance of the estimates decreases with increasing numbers of particles but with decreasing marginal benefits of adding particles. The time it takes to run the particle filter increases linearly with the number of particles because each particle requires an additional forward simulation. For this reason, we choose to calibrate at 1000 particles, which offers a balance between keeping the variance of the marginal log likelihood estimates low and the run times low. The outcomes of this calibration are expected to be similar for other parameter sets, though more particles could be needed for parameter values that are not likely and have low marginal log likelihood. Initializing pMCMC at high-likelihood values produced from maximum likelihood estimation lessens the chances of this happening.

With the choice of 1000 particles for the particle filter, we can evaluate the performance of pMCMC on simulated data. We choose an arbitrary set of parameters, $s = 0.3$, $h = 0.2$, $R_0 = 0.2$, and $A = 10$, and simulate from the Markov process and sampling distributions ($f(\cdot)$ and $g(\cdot)$) to produce simulated population genotype frequencies $x_{1:T}$ and data $y_{1:T}$. The simulated frequencies and samples are shown in Figure 4.3. Inference for pMCMC is based on convergence of the Markov chain for each parameter to the appropriate posterior distributions. Visually, based on trace plots (which plot the values of each parameter in their Markov chains) for 60,000 iterations of pMCMC, independent chains each appear to converge fairly quickly, with A taking the most iterations to converge at around 10,000 (Figure 4.4). Results for other simulations and initial parameter values (not shown here) suggest fairly robust convergence, though some scenarios can take many more iterations to converge. We choose to discard the first 30,000 iterations to ensure convergence. We also consider several diagnostics to measure convergence quantitatively. In the 30,000 iterations after burn-in (second half as shown in Figure 4.4), we see there is somewhat substantial autocorrelation, with over 50 lag sometimes needed to make the samples independent (Figure 4.5). The resulting effective sample sizes for each parameter are A : 5367, h : 4460, R_0 : 5377, and s : 4992, suggesting that 60,000 iterations with 30,000 iterations of burn-in produces sufficient sampling of the posterior distributions. Finally, the Gelman-Rubin statistic is 1 for each parameter, indicating that the 3 chains

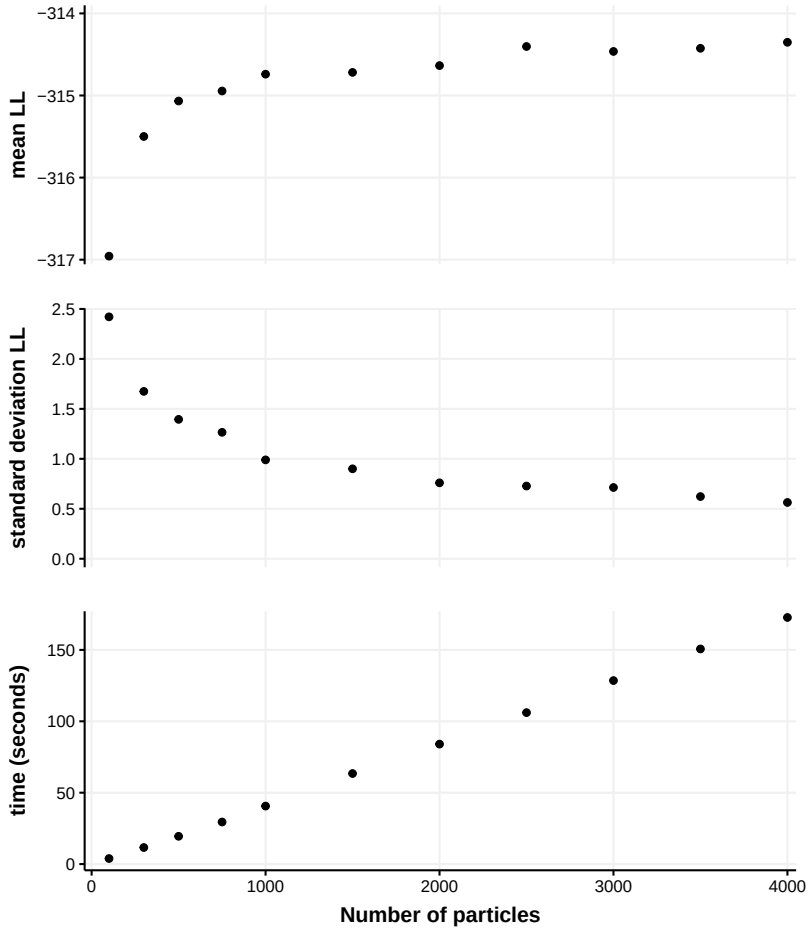

Figure 4.2 Particle filter calibration on experimental data. 100 bootstrap filters were run with each number of particles (x-axis) to allow comparison of time to run and mean and standard deviation of the estimated marginal log likelihood values (LL). The bootstrap filter was run using parameters $s = 0.3$, $h = 0.2$, $R_0 = 0.2$, and $A = 5$.

have converged (Gelman, Rubin, et al., 1992).

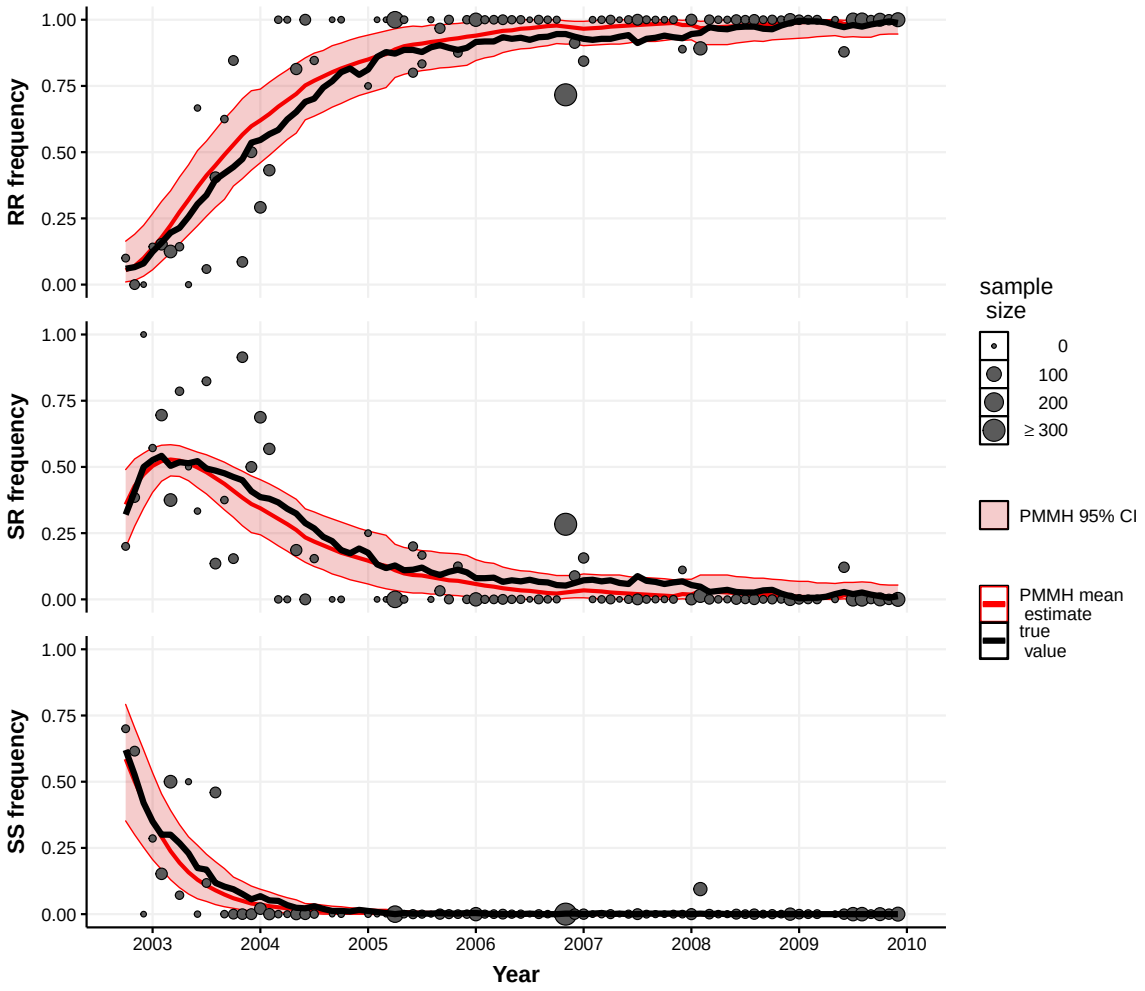

Figure 4.3 Simulated genotype frequencies and data with pMCMC estimates and 95% credible interval (CI). A simulation was run with parameters $s = 0.3$, $h = 0.2, R_0 = 0.2$, and $A = 10$, with sample sizes N_s as in the experiment, producing a time series of true genotype frequencies (black line) and samples (points, scaled in size by sample size). The pMCMC posterior mean genotype frequency estimates are shown by the red line, and the shaded region shows the 95% CI.

Having established convergence, we look at the posterior distributions. The posterior distribution of each parameter contained the true value (used to generate the simulated data) within its 95% credible interval, with mean (95% credible interval) of: $A = 9.22$ $(5.48, 12.25)$, $h = 0.183$ $(0.024, 0.443)$, $R_0 = 0.24$ $(0.115, 0.398)$, and $s = 0.324$ $(0.195, 0.453)$. pMCMC also gives samples from the smoothing distribution, $f(x_{1:T}|y_{1:T})$, which we use to construct 95% credible intervals for the frequencies. In

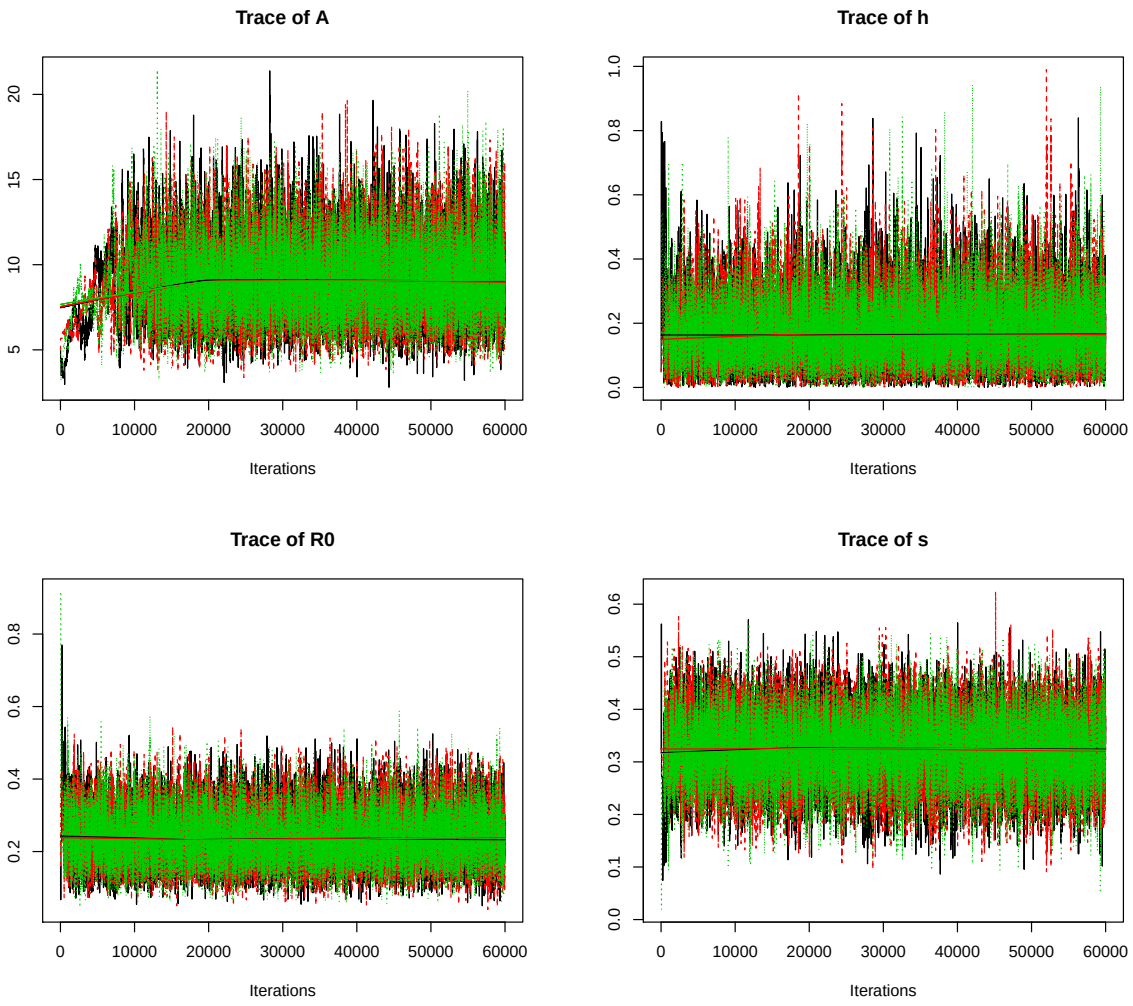

Figure 4.4 Trace plots from pMCMC on the simulated data shown in Figure 4.3. Three independent Markov chains of 60,000 iterations each are shown in different colors. The true parameter values were $s = 0.3$, $h = 0.2$, $R_0 = 0.2$, and $A = 10$, and it is evident that the Markov chains converge to posterior distributions containing these values.

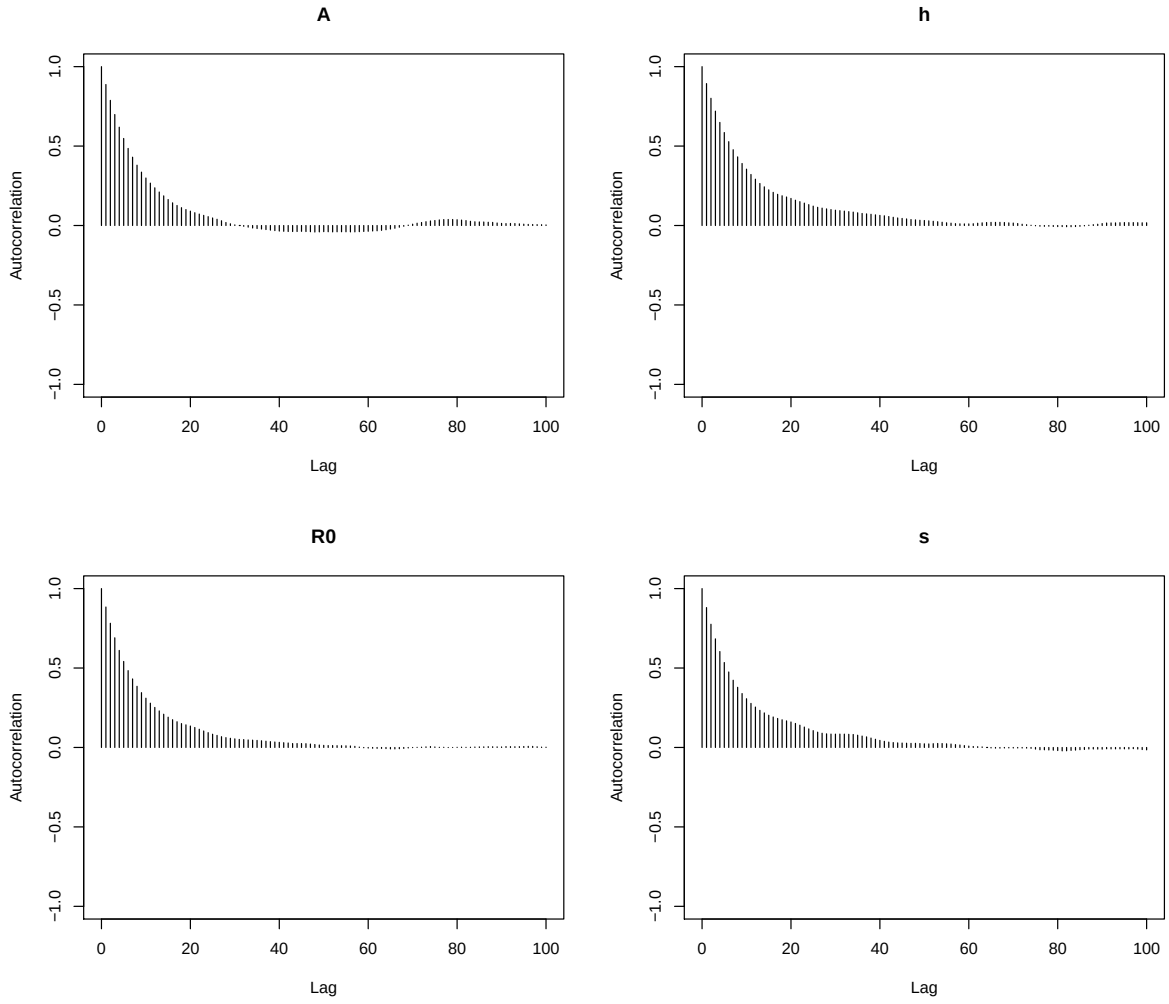

Figure 4.5 Autocorrelation for each parameter in a single Markov chain from Figure 4.4. Each plot shows a different parameter, with bars showing the correlation between Markov chain values of that parameter that are different numbers of samples apart (lag, on the x-axis).

this case, the 95% credible intervals contain the true frequencies for nearly the entire time series (Figure 4.3).

We similarly run pMCMC on the genotype frequency data collected from Iquitos, Peru. The trace plots appear to show that each parameter has converged (Figure 4.6). The Gelman-Rubin statistic is 1 for each parameter, further suggesting convergence, and the effective sample sizes of A : 4307, h : 2880, R_0 : 4251, and s : 4115, while lower than with the simulated data, are still large enough to approximate the posterior distribution.

The estimates of the probability density functions show smooth distributions (diagonal of Figure 4.7). The mean (95% CI) for each parameter is: $A = 4.61$ $(2.88, 6.97)$, $h = 0.178$ $(0.011, 0.560)$, $R_0 = 0.276$ $(0.140, 0.452)$, and $s = 0.185$ $(0.084, 0.281)$. The estimated cost to SS is small, and while the degree of dominance has a high degree of uncertainty, it can be concluded that the fitness cost is more likely to be recessive than dominant. These results mirror those found by Baltzegar (2020). The posterior distributions demonstrate correlation between parameters, with a particularly strong correlation between h and s and between R_0 and s (-0.667 and -0.757, respectively), as well as between h and R_0 (0.513) (Figure 4.7). These correlations make sense intuitively. For example, without large samples in the first several generations, a low initial frequency and large fitness cost may not be possible to distinguish from a higher initial frequency and lower costs.

We can use the samples from the smoothing distribution obtained from pMCMC to construct mean state estimates and 95% CIs (Figure 4.8). Overall, it appears plausible that the 95% credible interval contains the true population frequencies, given that there is high sampling variance. One of the biggest sources of discrepancy is in late 2003 and 2004, where frequencies of sampled heterozygote are consistently less than the state estimates. In later years, the state estimates suggest that the S allele is maintained in the population in heterozygotes. Low estimates for h produce this type of dynamics because there is little difference in fitness between RR and SR mosquitoes.

4.5 Discussion

We employ pMCMC to infer selection parameters of insecticide resistance in Iquitos, Peru. Our approach using an HMM allows a flexible formulation of model and sampling functions, and we can account for increased sample variance that results from non-random sampling of mosquitoes.

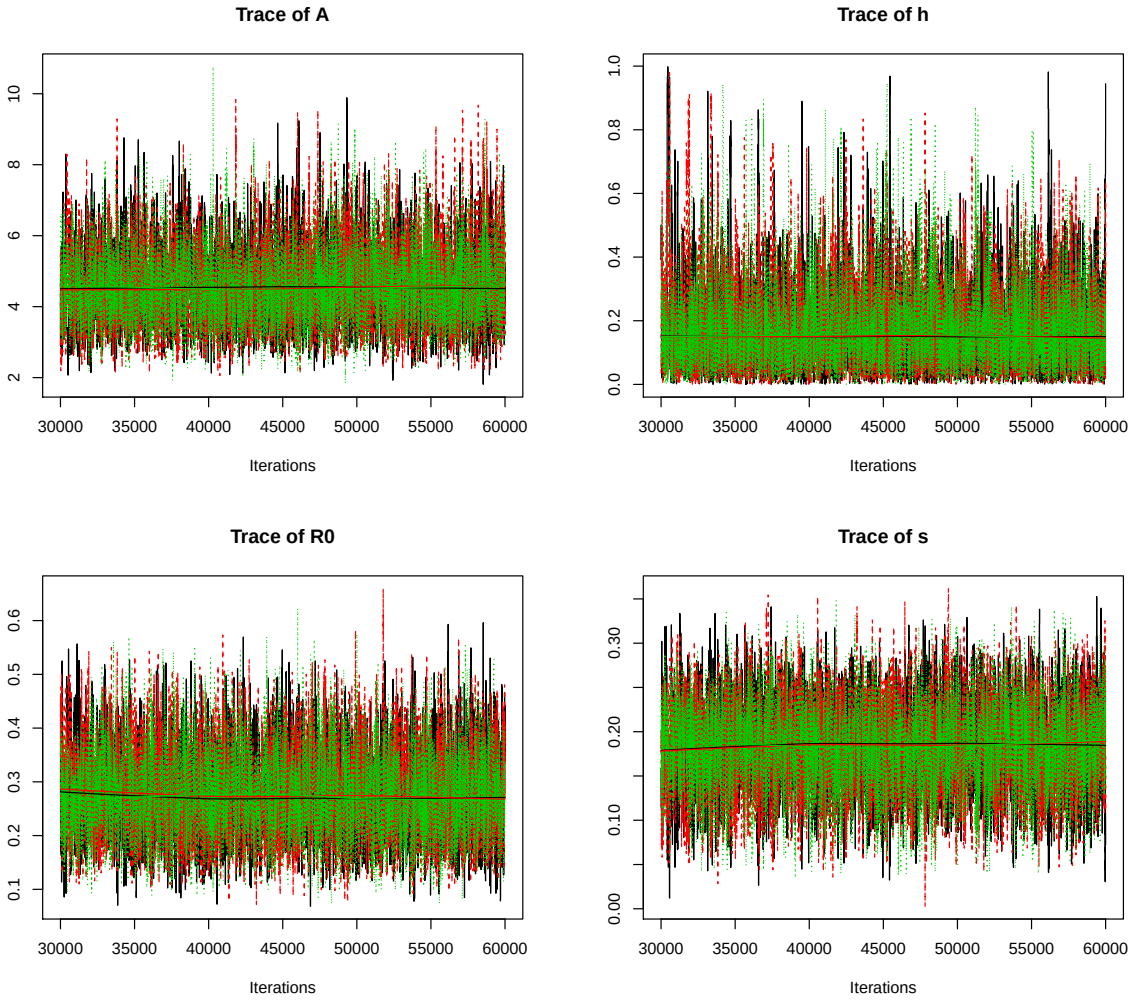

Figure 4.6 Trace plots from pMCMC on experimental data. Three independent Markov chains of 30,000 iterations after 30,000 iterations of burn-in are shown in different colors. The Markov chains appear to have converged.

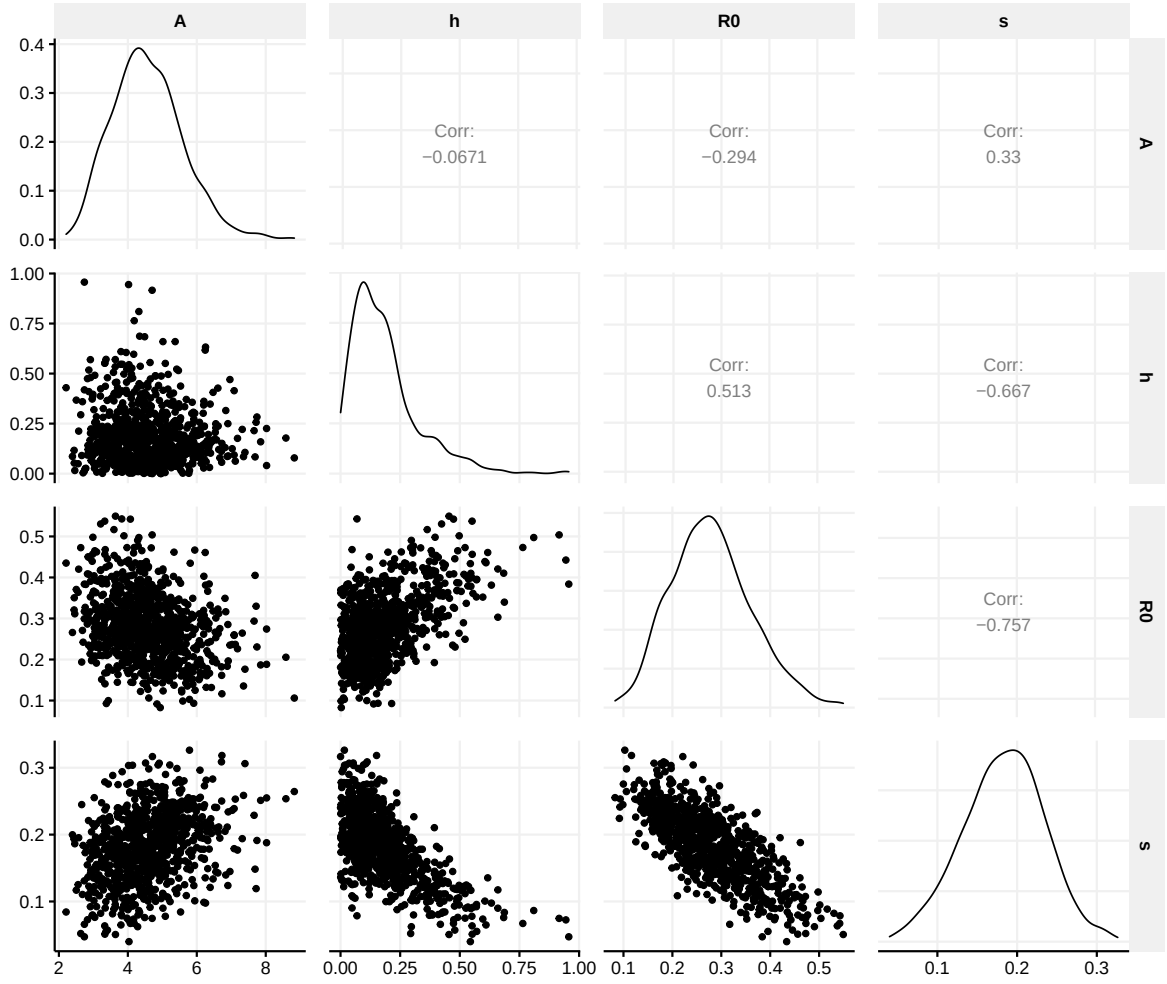

Figure 4.7 Parameter posterior probability density function estimates and correlation between parameters from pMCMC on experimental data in the Markov chains shown in Figure 4.6. The diagonal gives a smoothed estimate of the posterior probability density function for each parameter. The bottom left scatter plots show pairwise relationships between samples of parameters, and the associated correlations between parameters are shown on the top right. The Markov chain samples were thinned by 100 to enable visualization.

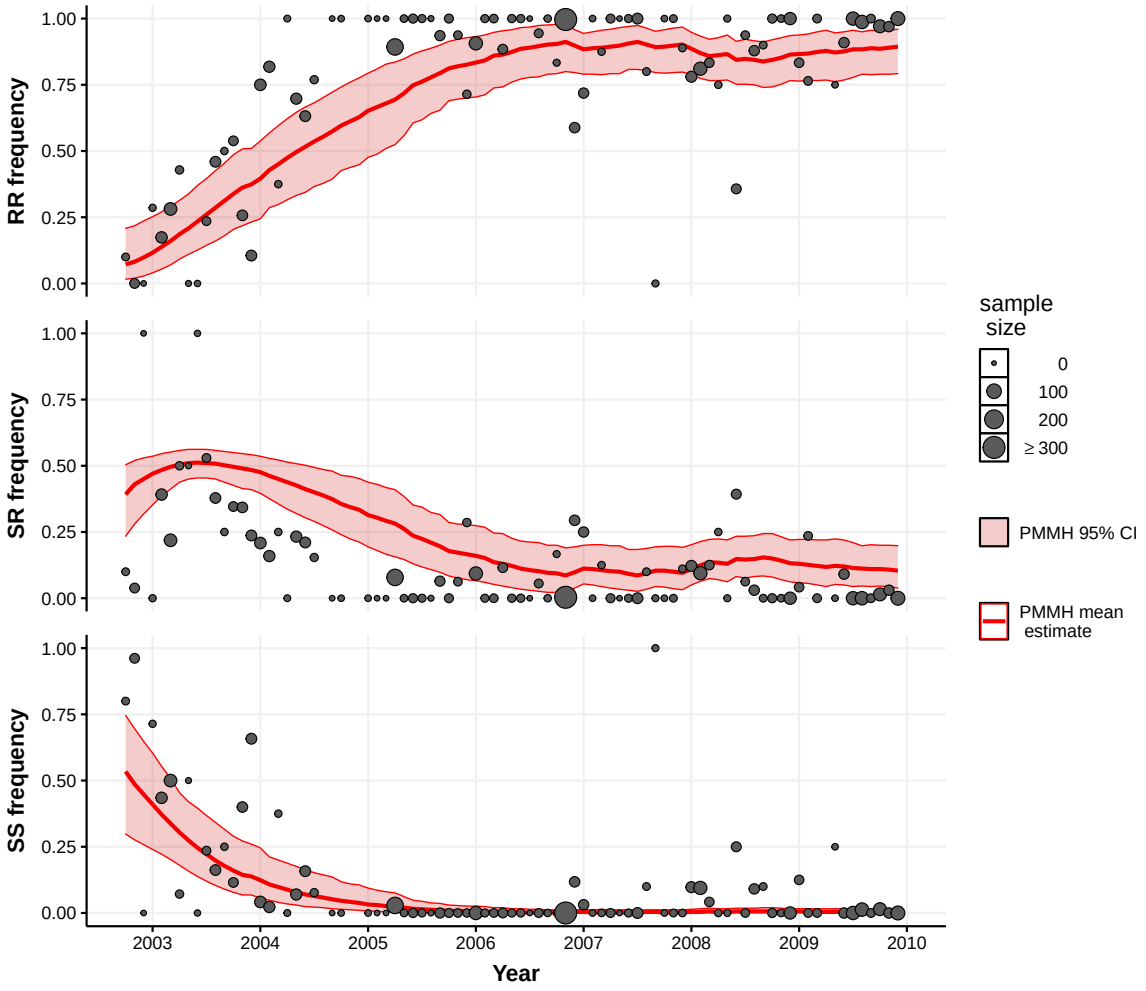

Figure 4.8 Experimental data with pMCMC estimates and 95% credible intervals (CI). Experimental observations are shown by points scaled in size based on that generation's sample size. The pMCMC posterior mean genotype frequencies are shown by the red line, and the shaded region shows the 95% CI.

Our results suggest that the initial frequency of the F1534C mutation in the population was fairly high, the fitness cost to homozygote susceptibles was small, and the degree of dominance of that cost is either recessive or co-dominant. The result of non-dominant fitness costs provide support for the conclusions found by Baltzegar (2020) using WFABC, which contrasted with previous results from lab experiments.

Many assumptions could affect the parameter estimates found here. A major assumption of the model is that of spatial homogeneity. Spatial heterogeneity could significantly alter genotype frequency dynamics, slowing the spread of resistance because the resistant allele would take time to migrate throughout the city. High initial frequencies of the resistant allele, as estimated, would make this factor slightly less impactful because the resistant allele would likely already be present in most sub-populations when spraying begins. However, the spatial distribution of mosquitoes would also affect dynamics. With small populations within each house and limited movement between houses, there would likely be much higher homozygosity in the population because small subpopulations experience high amounts of genetic drift. This could explain the over-estimation of estimated heterozygote frequencies compared to the sampled data.

In addition to spatially heterogeneous mosquito behavior, spraying was highly heterogeneous in reality because certain parts of the city were sprayed at certain times. As suggested by Baltzegar (2020), assuming a constant selection pressure, as we did here, would likely result in under-estimation of the fitness costs; the true effect would be strong when and where spraying occurs, while the overall effect including non-sprayed areas would appear weaker.

The model also assumed discrete-time, non-overlapping generations of one month in time. Given the spatial heterogeneity of the system as discussed above, continuous-time models would not be expected to substantially increase the model's representation of reality and may produce similar results to those here. We also expect a discrete-time model with overlapping generations to produce similar results given that the changes in frequency happen over extended time periods, though if using a model with spatial heterogeneity, overlapping generations would likely be important. Finally, shorter or longer generation times than one month would likely increase or decrease the estimated strength of selection, respectively, because it would correspond to faster or slower changes in a single generation.

The methodology described here could potentially be extended to other problems. Part of the

motivation for our choices of model and sampling functions is that they could be altered, in some cases minimally, to apply to a variety of other problems. For example, caged experiments or field trials can often be described by well-defined dynamical models and sampling functions, but properly accounting for the sources of stochasticity can be challenging. The model here could be expanded to account for overlapping generations, multiple sub-populations, additional alleles or loci, or non-Mendelian inheritance such as gene drives. However, there are several possible issues that could arise. First, for any more complex model than that shown here, the computational burden could quickly become too high, especially if many more particles are needed to consistently estimate the likelihood. Also, after adding additional model parameters, all parameters may no longer be identifiable, i.e., two or more parameters could affect the model output similarly enough that each could take a range of values. Correlation between parameters was already evident in the simple model here, which resulted in high degrees of uncertainty around parameter estimates but did not interfere with convergence. In scenarios with many parameters and limited sampling, convergence could take substantially longer. In such cases, some of the parameters would likely need to be set to a constant to enable estimation of the remaining parameters (by nature of the inference problem and not this methodology). A related issue is that the posterior distribution of the states could be multimodal. For example, if modeling a population as two meta-populations, the overall genotype frequencies are the additive combination of frequencies of the meta-populations, which would thus require appropriate data to accurately determine the composition of the individual meta-populations.

While our model has several limitations, the parameter estimates here suggest that susceptibility is not dominant to the F1534C mutation. This is an important contribution to the understanding of insecticide resistance and adds to the previous analysis done by Baltzegar (2020). These results suggest the need for further analysis of the data using other models and investigation of dominance with lab experiments. While there are many potential issues, our approach using pMCMC may also be possible to extend to other inference problems.

REFERENCES

Alphey, L. (2002). “Re-engineering the sterile insect technique”. In: *Insect biochemistry and molecular biology* 32.10, pp. 1243–1247.

Alphey, L., A. McKemey, D. Nimmo, M. Neira Oviedo, R. Lacroix, K. Matzen, and C. Beech (2013). “Genetic control of Aedes mosquitoes”. In: *Pathogens and global health* 107.4, pp. 170–179.

Alphey, N., M. B. Bonsall, and L. Alphey (2011). “Modeling resistance to genetic control of insects”. In: *Journal of Theoretical Biology* 270.1, pp. 42–55.

Andrieu, C., A. Doucet, and R. Holenstein (2010). “Particle markov chain monte carlo methods”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72.3, pp. 269–342.

Ant, T., M. Koukidou, P. Rempoulakis, H.-F. Gong, A. Economopoulos, J. Vontas, and L. Alphey (2012). “Control of the olive fruit fly using genetics-enhanced sterile insect technique”. In: *BMC biology* 10.1, p. 51.

Baltzegar, J. (2020). “Population Genetics of Two Insect Pest Species of the Poor, the Mosquito, Aedes aegypti, and the Maize Weevil, Sitophilus zeamais.” PhD thesis. North Carolina State University.

Bank, C., G. B. Ewing, A. Ferrer-Admettla, M. Foll, and J. D. Jensen (2014). “Thinking too positive? Revisiting current methods of population genetic selection inference”. In: *Trends in Genetics* 30.12, pp. 540–546.

Bhatt, S., P. W. Gething, O. J. Brady, J. P. Messina, A. W. Farlow, C. L. Moyes, J. M. Drake, J. S. Brownstein, A. G. Hoen, O. Sankoh, et al. (2013). “The global distribution and burden of dengue”. In: *Nature* 496.7446, pp. 504–507.

Brengues, C., N. J. Hawkes, F. Chandre, L McCarroll, S. Duchon, P. Guillet, S. Manguin, J. Morgan, and J. Hemingway (2003). “Pyrethroid and DDT cross-resistance in Aedes aegypti is correlated with novel mutations in the voltage-gated sodium channel gene”. In: *Medical and veterinary entomology* 17.1, pp. 87–94.

Bull, J. J. (2016). “OUP: lethal gene drive selects inbreeding”. In: *Evolution, Medicine, and Public Health* 2017.1, pp. 1–16.

Burt, A. (2003). “Site-specific selfish genes as tools for the control and genetic engineering of natural populations”. In: *Proceedings of the Royal Society of London B: Biological Sciences* 270.1518, pp. 921–928.

Camargo, S. d. (1967). “History of Aedes aegypti eradication in the Americas.” In: *Bulletin of the World Health Organization* 36.4, p. 602.

Cameron, D. D., A. White, and J. Antonovics (2009). “Parasite–grass–forb interactions and rock–paper–scissor dynamics: predicting the effects of the parasitic plant Rhinanthus minor on host plant communities”. In: *Journal of Ecology* 97.6, pp. 1311–1319.

Cao, Y., D. T. Gillespie, and L. R. Petzold (2007). “Adaptive explicit-implicit tau-leaping method with automatic tau selection”. In: *The Journal of chemical physics* 126.22, p. 224101.

Carvalho, D. O., A. R. McKemey, L. Garziera, R. Lacroix, C. A. Donnelly, L. Alphey, A. Malavasi, and M. L. Capurro (2015). “Suppression of a field population of Aedes aegypti in Brazil by sustained release of transgenic male mosquitoes”. In: *PLoS neglected tropical diseases* 9.7, e0003864.

Carvalho, F. D. and L. A. Moreira (2017). “Why is Aedes aegypti Linnaeus so successful as a species?” In: *Neotropical entomology* 46.3, pp. 243–255.

Champer, J., A. Buchman, and O. S. Akbari (2016). “Cheating evolution: engineering gene drives to manipulate the fate of wild populations”. In: *Nature Reviews Genetics* 17.3, pp. 146–159.

Champer, J., R. Reeves, S. Y. Oh, C. Liu, J. Liu, A. G. Clark, and P. W. Messer (2017). “Novel CRISPR/Cas9 gene drive constructs in Drosophila reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations”. In: *bioRxiv*, p. 112011.

Conner, J. K. and D. L. Hartl (2004). *A primer of ecological genetics.* Sinauer Associates Incorporated.

Cowell, R. (2016). “Combining allele frequency uncertainty and population substructure corrections in forensic DNA calculations”. In: *Forensic Science International: Genetics* 23, pp. 210–216.

Deredec, A., A. Burt, and H. C. J. Godfray (2008). “The population genetics of using homing endonuclease genes in vector and pest management”. In: *Genetics* 179.4, pp. 2013–2026.

Deutsch, C. A., J. J. Tewksbury, M. Tigchelaar, D. S. Battisti, S. C. Merrill, R. B. Huey, and R. L. Naylor (2018). “Increase in crop losses to insect pests in a warming climate”. In: *Science* 361.6405, pp. 916–919.

Dhole, S., M. R. Vella, A. L. Lloyd, and F. Gould (2018). “Invasion and migration of spatially self-limiting gene drives: A comparative analysis”. In: *Evolutionary applications* 11.5, pp. 794–808.

DiCarlo, J. E., A. Chavez, S. L. Dietz, K. M. Esvelt, and G. M. Church (2015). “Safeguarding CRISPR-Cas9 gene drives in yeast”. In: *Nature biotechnology* 33.12, pp. 1250–1255.

Doudna, J. A. and E. Charpentier (2014). “The new frontier of genome engineering with CRISPR-Cas9”. In: *Science* 346.6213, p. 1258096.

Du, Y., Y. Nomura, B. S. Zhorov, and K. Dong (2016). “Sodium channel mutations and pyrethroid resistance in Aedes aegypti”. In: *Insects* 7.4, p. 60.

Durrett, R. and S. Levin (1997). “Allelopathy in spatially distributed populations”. In: *Journal of theoretical biology* 185.2, pp. 165–171.

Esvelt, K. M., A. L. Smidler, F. Catteruccia, and G. M. Church (2014). “Concerning RNA-guided gene drives for the alteration of wild populations”. In: *Elife* 3, e03401.

Fan, Y. and J. G. Scott (2020). “The F1534C voltage-sensitive sodium channel mutation confers 7-to 16-fold resistance to pyrethroid insecticides in Aedes aegypti”. In: *Pest Management Science* 76.6, pp. 2251–2259.

Foll, M., H. Shim, and J. D. Jensen (2015). “WFABC: a W right–F isher ABC-based approach for inferring effective population sizes and selection coefficients from time-sampled data”. In: *Molecular ecology resources* 15.1, pp. 87–98.

Foster, G., W. Vogt, T. Woodburn, and P. Smith (1988). “Computer simulation of genetic control. Comparison of sterile males and field-female killing systems”. In: *Theoretical and Applied Genetics* 76.6, pp. 870–879.

Fouque, F., R. Carinci, P. Gaborit, J. Issaly, D. J. Bicout, and P. Sabatier (2006). “Aedes aegypti survival and dengue transmission patterns in French Guiana”. In: *Journal of Vector Ecology* 31.2, pp. 390–399.

Fu, G., K. C. Condon, M. J. Epton, P. Gong, L. Jin, G. C. Condon, N. I. Morrison, T. H. Dafa’alla, and L. Alphey (2007). “Female-specific insect lethality engineered using alternative splicing”. In: *Nature biotechnology* 25.3, pp. 353–357.

Fu, G., R. S. Lees, D. Nimmo, D. Aw, L. Jin, P. Gray, T. U. Berendonk, H. White-Cooper, S. Scaife, H. K. Phuc, et al. (2010). “Female-specific flightless phenotype for mosquito control”. In: *Proceedings of the National Academy of Sciences* 107.10, pp. 4550–4554.

Gantz, V. M. and E. Bier (2015). “The mutagenic chain reaction: A method for converting heterozygous to homozygous mutations”. In: *Science* 348.6233, pp. 442–444.

— (2016). “The dawn of active genetics”. In: *BioEssays* 38.1, pp. 50–63.

Gelman, A., D. B. Rubin, et al. (1992). “Inference from iterative simulation using multiple sequences”. In: *Statistical science* 7.4, pp. 457–472.

Gentile, J. E., S. S. Rund, and G. R. Madey (2015). “Modelling sterile insect technique to control the population of Anopheles gambiae”. In: *Malaria journal* 14.1, p. 92.

Getis, A., A. C. Morrison, K. Gray, and T. W. Scott (2003). “Characteristics of the spatial pattern of the dengue vector, Aedes aegypti, in Iquitos, Peru”. In: *The American journal of tropical medicine and hygiene* 69.5, pp. 494–505.

Gong, P., M. J. Epton, G. Fu, S. Scaife, A. Hiscox, K. C. Condon, G. C. Condon, N. I. Morrison, D. W. Kelly, T. Dafa’alla, et al. (2005). “A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly”. In: *Nature biotechnology* 23.4, pp. 453–456.

Gorman, K., J. Young, L. Pineda, R. Márquez, N. Sosa, D. Bernal, R. Torres, Y. Soto, R. Lacroix, N. Naish, et al. (2016). “Short-term suppression of Aedes aegypti using genetic control does not facilitate Aedes albopictus”. In: *Pest management science* 72.3, pp. 618–628.

Gossen, M. and H. Bujard (1992). “Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.” In: *Proceedings of the National Academy of Sciences* 89.12, pp. 5547–5551.

Gould, F. and P. Schliekelman (2004). “Population genetics of autocidal control and strain replacement”. In: *Annual Review of Entomology* 49.1, pp. 193–217.

Gregory, M., L. Alphey, N. I. Morrison, and S. M. Shimeld (2016). “Insect transformation with piggyBac: getting the number of injections just right”. In: *Insect molecular biology* 25.3, pp. 259–271.

Gunning, C. E., K. W. Okamoto, H. Astete, G. M. Vasquez, E. Erhardt, C. Del Aguila, R. Pinedo, R. Cardenas, C. Pacheco, E. Chalco, et al. (2018). “Efficacy of Aedes aegypti control by indoor Ultra Low Volume (ULV) insecticide spraying in Iquitos, Peru”. In: *PLoS neglected tropical diseases* 12.4, e0006378.

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?” In: *Journal of medical entomology* 38.3, pp. 411–422.

Harris, A. F., A. R. McKemey, D. Nimmo, Z. Curtis, I. Black, S. A. Morgan, M. N. Oviedo, R. Lacroix, N. Naish, N. I. Morrison, et al. (2012). “Successful suppression of a field mosquito population by sustained release of engineered male mosquitoes”. In: *Nature biotechnology* 30.9, pp. 828–830.

Hartl, D. and A. Clark (2007). *Principles of population genetics. Sunderland (MA)*.

Harvey-Samuel, T., N. I. Morrison, A. S. Walker, T. Marubbi, J. Yao, H. L. Collins, K. Gorman, T. E. Davies, N. Alphey, S. Warner, et al. (2015). “Pest control and resistance management through release of insects carrying a male-selecting transgene”. In: *BMC biology* 13.1, p. 49.

He, Z., X. Dai, M. A. Beaumont, and F. Yu (2020). “Maximum likelihood estimation of natural selection and allele age from time series data of allele frequencies”. In: *BioRxiv*, p. 837310.

Heinrich, J. C. and M. J. Scott (2000). “A repressible female-specific lethal genetic system for making transgenic insect strains suitable for a sterile-release program”. In: *Proceedings of the National Academy of Sciences* 97.15, pp. 8229–8232.

Hemingway, J. and H. Ranson (2000). “Insecticide resistance in insect vectors of human disease”. In: *Annual review of entomology* 45.1, pp. 371–391.

Hibbard, B. E., L. N. Meihls, M. R. Ellersieck, and D. W. Onstad (2010). “Density-dependent and density-independent mortality of the western corn rootworm: impact on dose calculations of rootworm-resistant Bt corn”. In: *Journal of Economic Entomology* 103.1, pp. 77–84.

Hoppensteadt, F. C. and C. Peskin (2013). *Mathematics in medicine and the life sciences*. Vol. 10. Springer Science & Business Media.

Hsu, P. D., E. S. Lander, and F. Zhang (2014). “Development and applications of CRISPR-Cas9 for genome engineering”. In: *Cell* 157.6, pp. 1262–1278.

Huang, Y., A. L. Lloyd, M. Legros, and F. Gould (2009). “Gene-drive in age-structured insect populations”. In: *Evolutionary Applications* 2.2, pp. 143–159.

Johnson, P. (2019). *adaptivetau: Tau-Leaping Stochastic Simulation*. R package version 2.2-3.

Kelly-Hope, L., H. Ranson, and J. Hemingway (2008). “Lessons from the past: managing insecticide resistance in malaria control and eradication programmes”. In: *The Lancet infectious diseases* 8.6, pp. 387–389.

Kerr, B., M. A. Riley, M. W. Feldman, and B. J. Bohannan (2002). “Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors”. In: *Nature* 418.6894, pp. 171–174.

Khamis, D., C. El Mouden, K. Kura, and M. B. Bonsall (2018). “Ecological effects on underdominance threshold drives for vector control”. In: *Journal of Theoretical Biology* 456, pp. 1–15.

Kitakado, T., S. Kitada, H. Kishino, and H. J. Skaug (2006). “An integrated-likelihood method for estimating genetic differentiation between populations”. In: *Genetics* 173.4, pp. 2073–2082.

Labbé, G. M., S. Scaife, S. A. Morgan, Z. H. Curtis, and L. Alphey (2012). “Female-specific flightless (fsRIDL) phenotype for control of Aedes albopictus”. In: *PLoS Negl Trop Dis* 6.7, e1724.

LaCon, G., A. C. Morrison, H. Astete, S. T. Stoddard, V. A. Paz-Soldan, J. P. Elder, E. S. Halsey, T. W. Scott, U. Kitron, and G. M. Vazquez-Prokopec (2014). “Shifting patterns of Aedes aegypti fine scale spatial clustering in Iquitos, Peru”. In: *PLoS Negl Trop Dis* 8.8, e3038.

Lacroix, R., A. R. McKemey, N. Raduan, L. K. Wee, W. H. Ming, T. G. Ney, S. R. AA, S. Salman, S. Subramaniam, O. Nordin, et al. (2012). "Open field release of genetically engineered sterile male Aedes aegypti in Malaysia". In: *PloS one* 7.8, e42771.

Leftwich, P. T., M. Koukidou, P. Rempoulakis, H.-F. Gong, A. Zacharopoulou, G. Fu, T. Chapman, A. Economopoulos, J. Vontas, and L. Alphey (2014). "Genetic elimination of field-cage populations of Mediterranean fruit flies". In: *Proceedings of the Royal Society B: Biological Sciences* 281.1792, p. 20141372.

Magori, K., M. Legros, M. E. Puente, D. A. Focks, T. W. Scott, A. L. Lloyd, and F. Gould (2009). "Skeeter Buster: a stochastic, spatially explicit modeling tool for studying Aedes aegypti population replacement and population suppression strategies". In: *PLoS Negl Trop Dis* 3.9, e508.

Malaspinas, A.-S., O. Malaspinas, S. N. Evans, and M. Slatkin (2012). "Estimating allele age and selection coefficient from time-serial data". In: *Genetics* 192.2, pp. 599–607.

Marshall, J. M., A. Buchman, et al. (2017). "Overcoming evolved resistance to population-suppressing homing-based gene drives". In: *Scientific Reports* 7.

Mathieson, I. and G. McVean (2013). "Estimating selection coefficients in spatially structured populations from time series data of allele frequencies". In: *Genetics* 193.3, pp. 973–984.

Michaud, N., P. de Valpine, D. Turek, C. J. Paciorek, and D. Nguyen (2017). *Sequential Monte Carlo Methods in the nimble R Package*. arXiv: 1703.06206 [stat.CO].

Morrison, A. C., K. Gray, A. Getis, H. Astete, M. Sihuincha, D. Focks, D. Watts, J. D. Stancil, J. G. Olson, P. Blair, et al. (2004). "Temporal and geographic patterns of Aedes aegypti (Diptera: Culicidae) production in Iquitos, Peru". In: *Journal of medical entomology* 41.6, pp. 1123–1142.

Muir, L. E. and B. H. Kay (1998). "Aedes aegypti survival and dispersal estimated by mark-release-recapture in northern Australia." In: *The American journal of tropical medicine and hygiene* 58.3, pp. 277–282.

National Academies of Sciences, Engineering, and Medicine (2016). *Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values*. National Academies Press.

Noble, C., J. Olejarz, K. M. Esvelt, G. M. Church, and M. A. Nowak (2017). “Evolutionary dynamics of CRISPR gene drives”. In: *Science Advances* 3.4, e1601964.

Ogaugwu, C. E., M. F. Schetelig, and E. A. Wimmer (2013). “Transgenic sexing system for Ceratitis capitata (Diptera: Tephritidae) based on female-specific embryonic lethality”. In: *Insect biochemistry and molecular biology* 43.1, pp. 1–8.

Okamoto, K. W., M. A. Robert, F. Gould, and A. L. Lloyd (2014). “Feasible introgression of an anti-pathogen transgene into an urban mosquito population without using gene-drive”. In: *PLoS Negl Trop Dis* 8.7, e2827.

Phillips, I., J. Need, J. Escamilla, E. Colán, S. Sánchez, M. Rodríguez, L. Vásquez, J. Seminario, T. Betz, and A. Travassos da Rosa (1992). “First documented outbreak of dengue in the Peruvian Amazon region”. In: *Bulletin of the Pan American Health Organization (PAHO); 26 (3), 1992*.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Rendón, P, D. McInnis, D Lance, and J Stewart (2004). “Medfly (Diptera: Tephritidae) genetic sexing: large-scale field comparison of males-only and bisexual sterile fly releases in Guatemala”. In: *Journal of economic entomology* 97.5, pp. 1547–1553.

Robert, M. A., K. Okamoto, A. L. Lloyd, and F. Gould (2013). “A reduce and replace strategy for suppressing vector-borne diseases: insights from a deterministic model”. In: *PLoS One* 8.9, e73233.

Rueda, L., K. Patel, R. Axtell, and R. Stinner (1990). “Temperature-dependent development and survival rates of Culex quinquefasciatus and Aedes aegypti (Diptera: Culicidae)”. In: *Journal of medical entomology* 27.5, pp. 892–898.

Saarman, N. P., A. Gloria-Soria, E. C. Anderson, B. R. Evans, E. Pless, L. V. Cosme, C. Gonzalez-Acosta, B. Kamgang, D. M. Wesson, and J. R. Powell (2017). “Effective population sizes of a major vector of human diseases, Aedes aegypti”. In: *Evolutionary Applications* 10.10, pp. 1031–1039.

Saavedra-Rodriguez, K, L Urdaneta-Marquez, S. Rajatileka, M Moulton, A. Flores, I Fernandez-Salas, J Bisset, M Rodriguez, P. McCall, M. Donnelly, et al. (2007). “A mutation in the voltage-gated

sodium channel gene associated with pyrethroid resistance in Latin American Aedes aegypti". In: *Insect molecular biology* 16.6, pp. 785–798.

Savary, S., L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson (2019). "The global burden of pathogens and pests on major food crops". In: *Nature ecology & evolution* 3.3, pp. 430–439.

Schetelig, M. F. and A. M. Handler (2012). "A transgenic embryonic sexing system for Anastrepha suspensa (Diptera: Tephritidae)". In: *Insect biochemistry and molecular biology* 42.10, pp. 790–795.

Schetelig, M., A Targovska, J. Meza, K Bourtzis, and A. Handler (2016). "Tetracycline-suppressible female lethality and sterility in the Mexican fruit fly, Anastrepha ludens". In: *Insect molecular biology* 25.4, pp. 500–508.

Schliekelman, P. and F. Gould (2000). "Pest control by the release of insects carrying a female-killing allele on multiple loci". In: *Journal of Economic Entomology* 93.6, pp. 1566–1579.

Schraiber, J. G., S. N. Evans, and M. Slatkin (2016). "Bayesian inference of natural selection from allele frequency time series". In: *Genetics* 203.1, pp. 493–511.

Scott, M. J. (2014). "Development and evaluation of male-only strains of the Australian sheep blowfly, Lucilia cuprina". In: *BMC genetics* 15.S2, S3.

Sinervo, B. and C. M. Lively (1996). "The rock-paper-scissors game and the evolution of alternative male strategies". In: *Nature* 380.6571, p. 240.

Smith, L. B., C. Sears, H. Sun, R. W. Mertz, S. Kasai, and J. G. Scott (2019). "CYP-mediated resistance and cross-resistance to pyrethroids and organophosphates in Aedes aegypti in the presence and absence of kdr". In: *Pesticide Biochemistry and Physiology* 160, pp. 119 –126.

Sridhar, S., A. Luedtke, E. Langevin, M. Zhu, M. Bonaparte, T. Machabert, S. Savarino, B. Zambrano, A. Moureau, A. Khromava, et al. (2018). "Effect of dengue serostatus on dengue vaccine safety and efficacy". In: *New England Journal of Medicine* 379.4, pp. 327–340.

Strogatz, S. (2001). *Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering (Studies in nonlinearity)*. Westview Press.

Styer, L. M., S. L. Minnick, A. K. Sun, and T. W. Scott (2007). "Mortality and reproductive dynamics of Aedes aegypti (Diptera: Culicidae) fed human blood". In: *Vector-borne and zoonotic diseases* 7.1, pp. 86–98.

Tataru, P., M. Simonsen, T. Bataillon, and A. Hobolth (2017). "Statistical inference in the Wright–Fisher model using allele frequency data". In: *Systematic biology* 66.1, e30–e46.

Thomas, D. D., C. A. Donnelly, R. J. Wood, and L. S. Alphey (2000). "Insect population control using a dominant, repressible, lethal genetic system". In: *Science* 287.5462, pp. 2474–2476.

Unckless, R. L., A. G. Clark, and P. W. Messer (2017). "Evolution of resistance against CRISPR/Cas9 gene drive". In: *Genetics* 205.2, pp. 827–841.

Unckless, R. L., P. W. Messer, T. Connallon, and A. G. Clark (2015). "Modeling the Manipulation of Natural Populations by the Mutagenic Chain Reaction". In: *Genetics*, genetics–115.

Valpine, P. de, D. Turek, C. J. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik (2017). "Programming with models: writing statistical algorithms for general model structures with NIMBLE". In: *Journal of Computational and Graphical Statistics* 26.2, pp. 403–413.

Wise De Valdez, M. R., D. Nimmo, J. Betz, H.-F. Gong, A. A. James, L. Alphey, and W. C. Black (2011). "Genetic elimination of dengue vector mosquitoes". In: *Proceedings of the National Academy of Sciences* 108.12, pp. 4772–4775.

Wolfram Research, Inc. (2019). *Mathematica, Version 12.0*. Champaign, IL.

World Health Organization (2017). "Global vector control response 2017-2030". In:

Wu, B., L. Luo, and X. J. Gao (2016). "Cas9-triggered chain ablation of cas9 as a gene drive brake". In: *Nature biotechnology* 34.2, pp. 137–138.

Yan, Y. and M. J. Scott (2015). "A transgenic embryonic sexing system for the Australian sheep blow fly Lucilia cuprina". In: *Scientific reports* 5, p. 16090.

Yan, Y., M. E. Williamson, R. J. Davis, A. A. Andere, C. J. Picard, and M. J. Scott (2020). “Improved transgenic sexing strains for genetic control of the Australian sheep blow fly Lucilia cuprina using embryo-specific gene promoters”. In: *Molecular Genetics and Genomics* 295.2, pp. 287–298.

Zetsche, B., J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, et al. (2015). “Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system”. In: *Cell* 163.3, pp. 759–771.

APPENDICES

Appendix A

Mathematical analysis of reversing gene drives[1]

We examine the qualitative behavior of the system for various fitness costs and assumptions. Figure A.1 shows different qualitative behavior present with specific sets of parameter values by using phase plots, which illustrate the change in allele frequencies in one generation for a grid of initial conditions. The allele frequencies are displayed in DeFinetti diagrams (Hoppensteadt and Peskin, 2013), where the sum of the frequencies at each point adds to 1. The qualitative difference between the RD and IRD seen in Figures 2.1 and 2.2 is visible in Figure A.1a (RD) and Figure A.1c (IRD), which have parameter values identical to those in Figure 2.2. Increasing the fitness cost to individuals with a copy of both the HD and countermeasure ($s_{HD/C}$) changes the qualitative behavior of the system. For the RD, with $s_{HD/C} = 0.6$, trajectories spiral outward, toward the edges of the system (Figure A.1b). For the IRD, increasing $s_{HD/C}$ to 0.9 demonstrates bistability, where the system can reach fixation of either the IRD or the HD, depending on initial conditions (Figure A.1d).

To consider a range of values of s_{HD}, s_C, and $s_{HD/C}$, equilibria and their stabilities were computed numerically. Stability was calculated by finding the eigenvalues (λ_i) of the Jacobian of the system evaluated at that equilibrium, and checking if $|\lambda_i| \leq 1$ for all i. We assumed perfect homing ($e_{HD} = e_C = 1$) such that $q_R = 0$ at all times, and recessive fitness costs in wild-type heterozygotes ($h_{HD} =$

[1]This appendix is included as supplementary online material for the publication: Vella MR, Gunning CE, Lloyd AL & Gould F. (2017) Evaluating strategies for reversing CRISPR-Cas9 gene drives. Scientific Reports. 7: 11038. doi:10.1038/s41598-017-10633-2.

$h_C = 0$) in Figures A.2 and A.3 and additive fitness costs in wild-type heterozygotes ($h_{HD} = h_C = 0.5$) in Figures A.4 and A.5.

Different equilibria are stable for different regions of parameter values, as indicated by color in Figures A.2 and A.4. For each of the three countermeasures, Figures A.2 and A.4 illustrate the long-term behavior seen for different values of the HD fitness cost (s_{HD}), HD/C heterozygote fitness cost ($s_{HD/C}$) and the countermeasure fitness cost (s_C). Each row of the figure depicts a different countermeasure (SR, RD or IRD), and different panels across a row depict different countermeasure fitness costs. Within each panel, the HD fitness cost is shown on the horizontal axis and the HD/C heterozygote fitness cost on the vertical. Using HD, C, and W as abbrevations for homing drive, countermeasure, and wild-type alleles, respectively, the colors indicate the stable equilibrium as follows: red for HD fixation, blue for C fixation, green for an equilibrium that consists of a combination of C and W (i.e., no HD at equilibrium), purple for an equilibrium with a combination of HD and C, orange for an equilibrium with a combination of HD, C, and W, and yellow for bistability, when two of the previously listed equilibria are both stable. For SR and RD countermeasures, the bistable regions have both HD fixation and a combination of C and W as stable equilibria, with initial conditions determining the long-term behavior of the system. For the IRD, the bistable regions have both HD and IRD fixation as stable equilibria, as shown in the example of Figure A.1d. The difference for the IRD is because unlike the SR and RD, the IRD maintains a relative fitness advantage over W even in the absence of the HD. Finally, the brown regions indicate that those parameter values do not result in any stable equilibria with valid frequencies (i.e., each frequency in range [0,1]). Such cases often result in oscillatory dynamics away from an unstable, polymorphic equilibrium, as depicted in Figure A.1b.

Figures A.3 and A.5 correspond to Figures A.2 and A.4, respectively, and show the HD frequency at equilibrium. The empty areas of the figures indicate that the system does not always reach the same equilibrium (i.e., bistable or no stable). The SR and RD countermeasures only have stable equilibria without HD present when the countermeasure fitness cost is 0 ($s_C = 0$). The IRD, on the other hand, often has a stable equilibrium without any HD. In regions where there are intermediate frequencies of HD at equilibrium, the equilibrium HD frequency tends to increase as countermeasure fitness decreases. This explains the decrease in minimum HD frequency with decreasing countermeasure fitness costs shown in Figures 2.3 and 2.4, since oscillations around smaller frequencies of HD must

reach small HD frequencies.

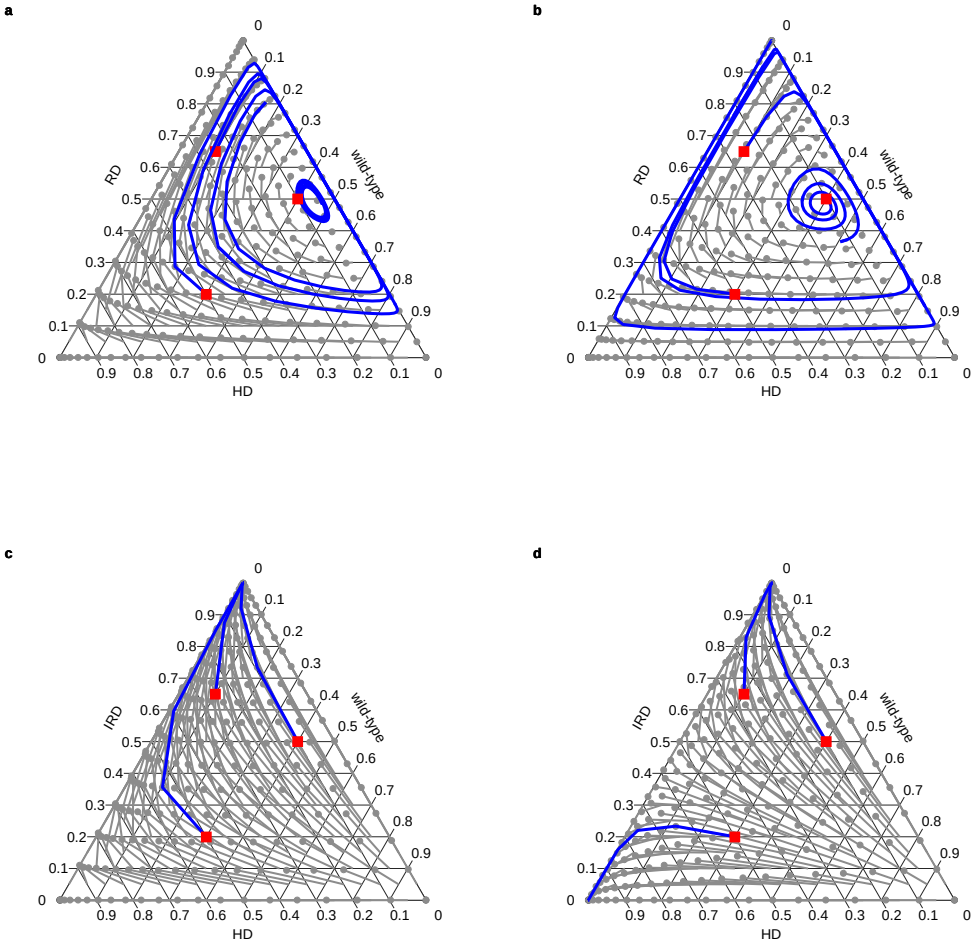

Figure A.1 Phase plots, showing change in allele frequencies in one generation for a grid of initial conditions, for RD (a,b) and IRD (c,d). DeFinetti diagrams (Hoppensteadt and Peskin, 2013) show allele frequencies with gray lines showing the allele frequency change in one generation, going toward the dot. Blue lines follow separate trajectories for 100 generations, starting from the red squares. **a/c)** parameter values from Figure 2.2 ($s_{HD} = 0.3$, $s_C = s_{HD/C} = 0.2$) demonstrate the stable polymorphic equilibrium for the RD and countermeasure fixation for the IRD. **b)** RD system with increased heterozygote fitness cost ($s_{HD/C} = 0.6$), resulting in no stable equilibria with frequencies in range [0,1] (brown region in Figure A.2). Trajectories oscillate away from the unstable, polymorphic equilibrium. **d)** IRD system with increased heterozygote fitness cost ($s_{HD/C} = 0.9$), resulting in a bistable system (yellow region in Figure A.2). Some initial conditions lead to HD fixation and others lead to countermeasure fixation. Note that near the edge of the triangle, stochastic loss of the allele with small frequency becomes likely in a finite, randomly mating population.

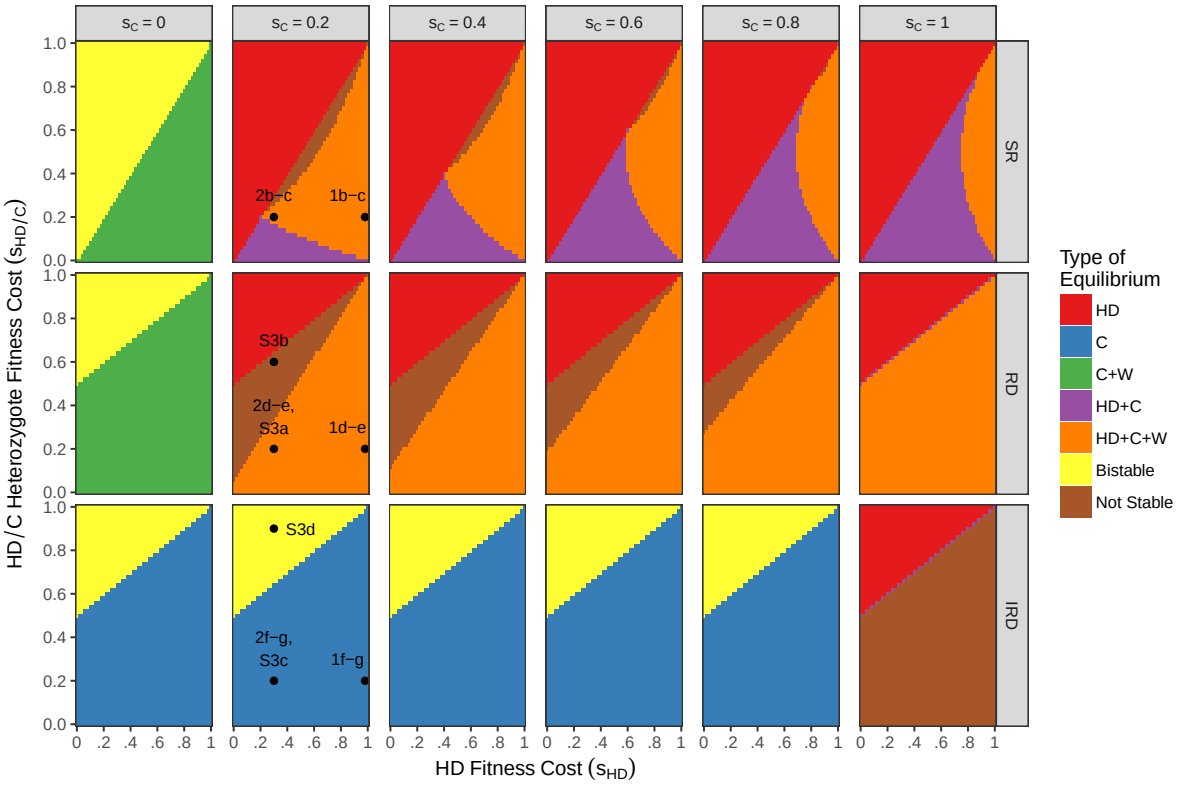

Figure A.2 Possible long-term behaviors of the system for each countermeasure and various fitness costs, assuming perfect homing ($e_{HD} = e_C = 1$) and no fitness costs in wild-type heterozygotes ($h_{HD} = h_C = 0$). Colors indicate which alleles are present in the single stable equilibrium of the system (e.g., green indicating that only countermeasure and wild-type are present at equilibrium), if there are multiple stable solutions (yellow), or if there are no stable solutions where all frequencies exist in range [0,1] (brown). The axes show fitness costs of the HD (x-axis) and HD/C heterozygote (y-axis), and the fitness cost of the countermeasure varies across columns. Type of countermeasure varies across rows. Black points indicate parameter combinations used in other figures.

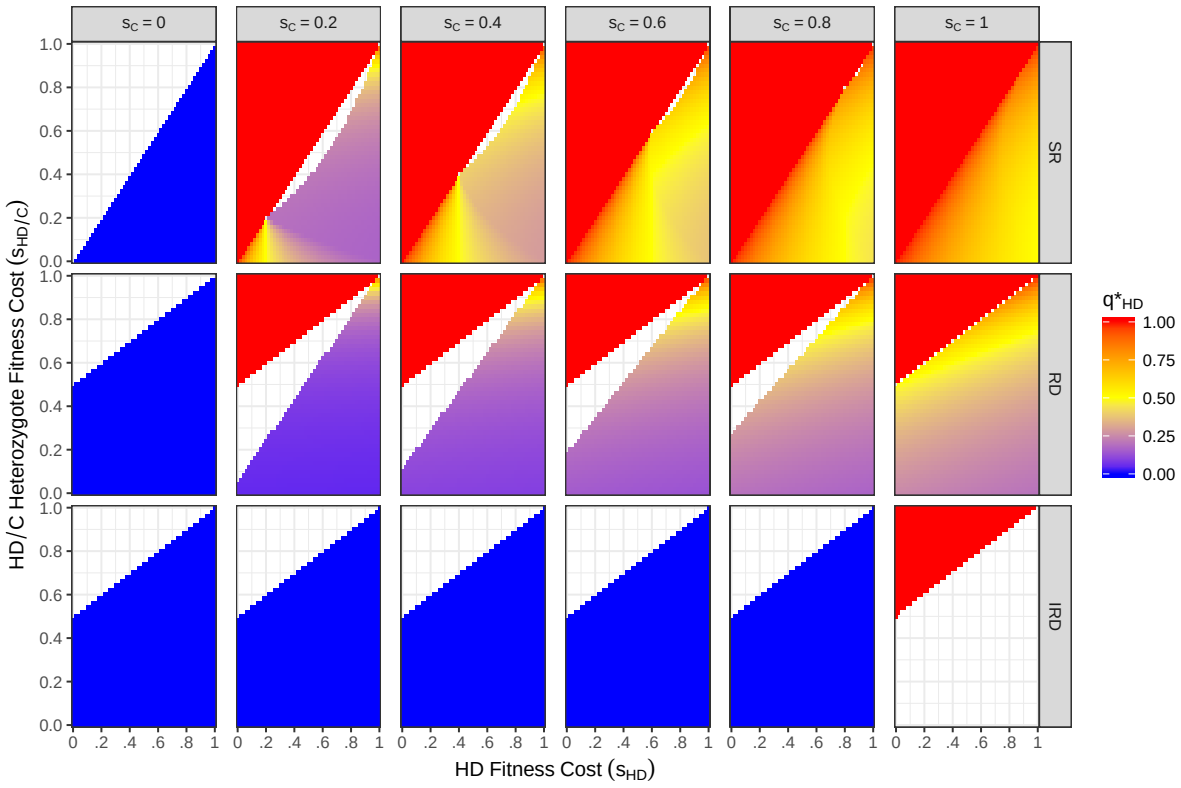

Figure A.3 Frequency of the HD at equilibrium for each countermeasure and varying fitness costs, assuming perfect homing ($e_{HD} = e_C = 1$) and recessive fitness costs in wild-type heterozygotes ($h_{HD} = h_C = 0$). Blue and red areas indicate regions of countermeasure and HD fixation, respectively, from Figure A.2, and intermediate shades show the HD frequency for stable polymorphic equilibria (orange in Figure A.2). Plot details are otherwise identical to Figure A.2.

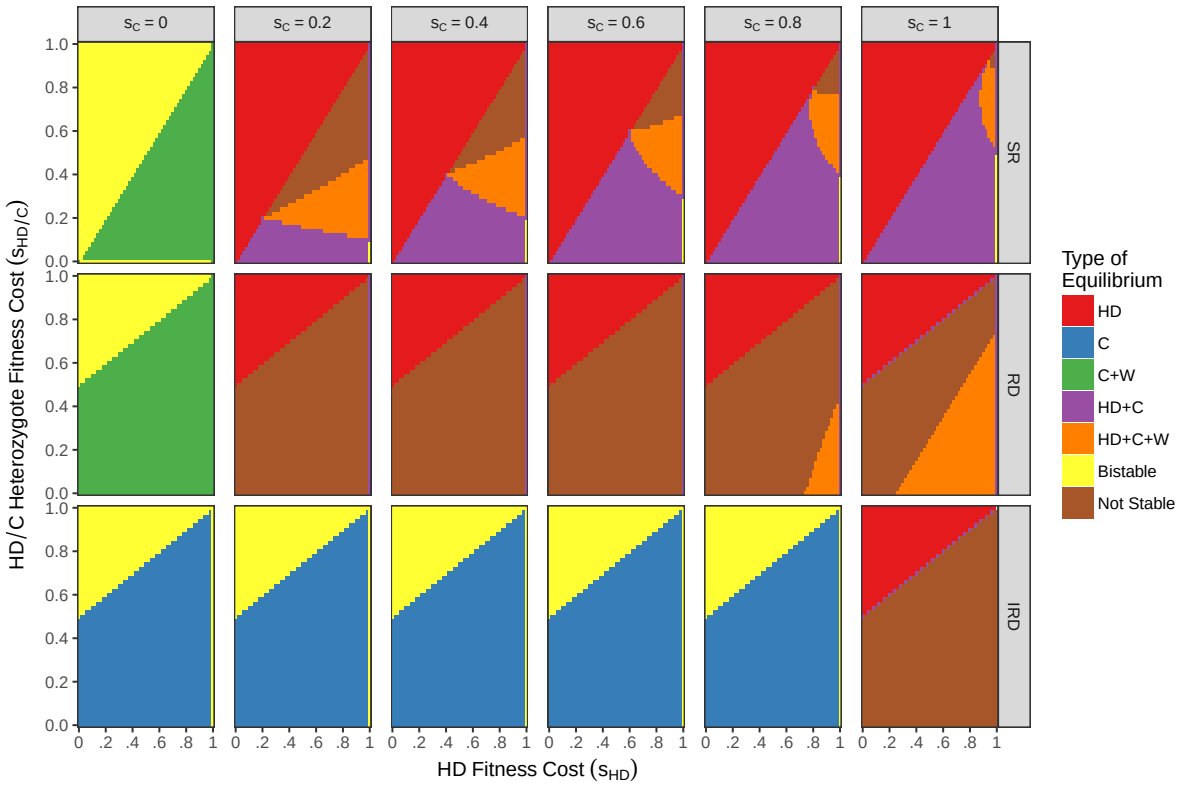

Figure A.4 Possible long-term behaviors of the system for each countermeasure and various fitness costs, assuming perfect homing ($e_{HD} = e_C = 1$) and additive fitness costs in wild-type heterozygotes ($h_{HD} = h_C = 0.5$). The axes show fitness costs of the HD (x-axis) and HD/C heterozygote (y-axis), and the fitness cost of the countermeasure varies across columns, and colors indicate the same types of behavior as in Figure A.2. Compared with recessive fitness costs in wild-type heterozygotes (Figure A.2), the SR countermeasure becomes more likely to result in removal of wild-type, and the RD countermeasure becomes more likely to have no stable equilibria.

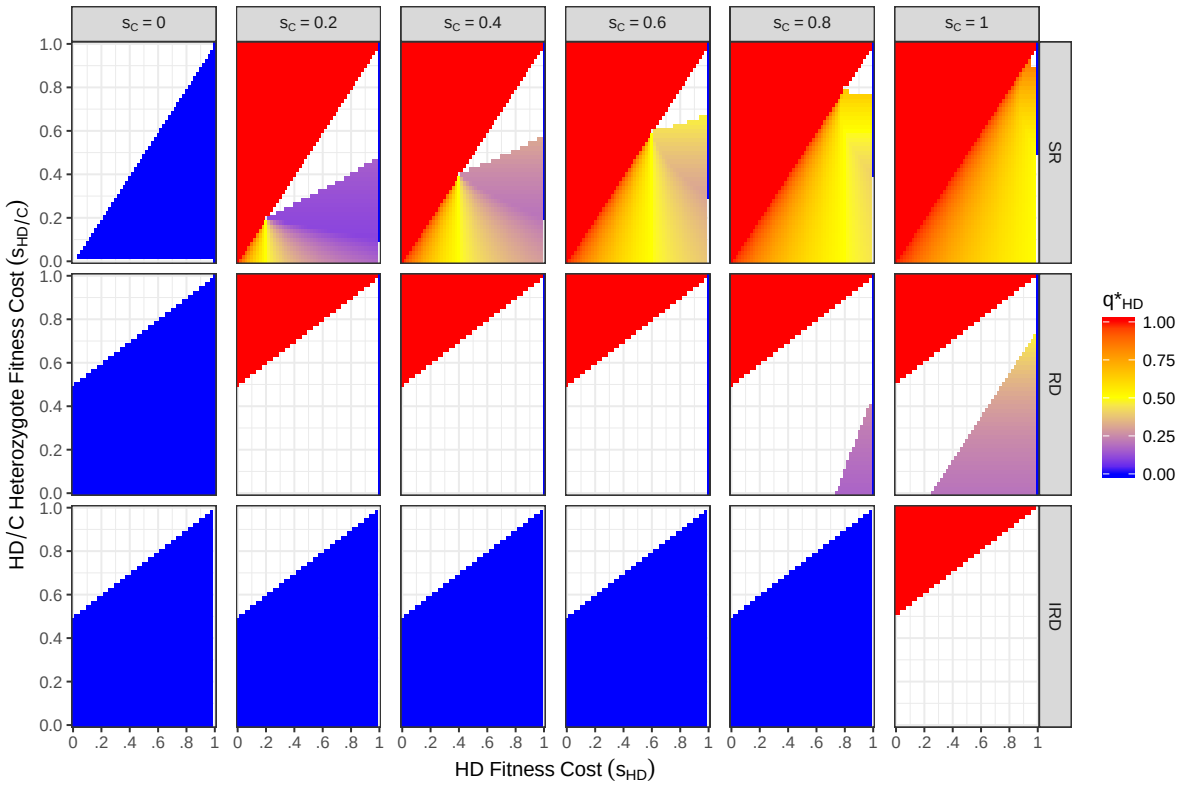

Figure A.5 Frequency of the HD at equilibrium for each countermeasure and varying fitness costs, assuming perfect homing ($e_{HD} = e_C = 1$) and additive fitness costs in wild-type heterozygotes ($h_{HD} = h_C = 0.5$). Blue and red areas indicate regions of countermeasure and HD fixation, respectively, from Figure A.4, and intermediate shades show the HD frequency for stable polymorphic equilibria (orange in Figure A.4). Plot details are otherwise identical to Figure A.4.

Appendix B

Further exploration of 1- and 2-locus female-killing[1]

B.1 Stochastic simulations

In order to understand the importance of demographic stochasticity (e.g. to simulate extinction events) and genetic drift, we formulated a stochastic model that is analogous to our deterministic model. Specifically, the rates of each process (birth, maturation and death) that appear in the deterministic model were taken to be rates of a continuous-time, discrete-state Markov process model. This model was simulated using a tau-leaping approach. The steps are as follows:

1. calculate each of the rates $k_j(t)$ in the system at time t
2. advance the time step by the time step τ
3. approximate the number of times each event occurred using a Poisson distribution with mean $\tau k_j(t)$
4. adjust the states accordingly before repeating all steps

We ran simulations in R (R Core Team, 2019) using the adaptivetau package (Johnson, 2019), which uses automatic selection of τ with adaptive explicit-implicit tau-leaping (Cao et al., 2007).

[1]This appendix will be included as supplementary online material for the manuscript: Vella MR, Gould F & Lloyd AL. Mathematical modeling of genetic pest management through female-specific lethality: Is one locus better than two?

For each approach, we ran 300 stochastic simulations for each value in a range of release ratios with the default parameters from the main text. Simulations were stopped after 1000 days. The mean time until there were no adult females in the stochastic simulations (Figure B.4) is similar to the time until 99.95% reduction in deterministic simulations (as shown in Figure 3.2B). In stochastic simulations, random fluctuations can cause the population to go extinct even when r is below r_c. As r approaches r_c and the equilibrium population size becomes smaller, extinction occurs within 1000 days in a higher percentage of stochastic simulations. Because the mean is taken from the simulations that go extinct, the mean time until extinction is biased downwards, most notably when extinction times frequently exceed 1000 days. For this reason, Figure B.4 only shows outcomes when at least 200 of the 300 simulations reached extinction. While smaller wild-type populations would increase the variance in stochastic simulations, the results suggest the deterministic simulations represent the overall dynamics well.

B.2 Equilibrium analysis

For the 1-locus system, we analytically solved for equilibria and calculated their stabilities using the eigenvalues of the Jacobian - a standard analytic approach for dynamical systems (Strogatz, 2001). The system exhibits a saddle-node bifurcation as r increases. Here this means that when r is small, there is a non-zero stable equilibrium for number of viable females, but past a critical release ratio, r_c, the only stable equilibrium is a population size of zero. The bifurcation diagram is shown in Figure B.3, which shows the same stable equilibria as Figure 3.2a but also shows unstable equilibria. For the 2-locus system, we could not find an analytical solution, but simulation results indicate similar dynamics are present.

Below r_c, the system is bistable, with dynamics bringing the population to extinction if beginning below the unstable equilibrium and bringing the population to a non-zero, stable equilibrium when beginning above the unstable equilibrium. While the unstable equilibrium which serves as the threshold for whether the system goes to zero or goes to the non-zero equilibrium is in multiple dimensions, there is also a threshold when beginning from a wild-type only population. The bistability in a 1-locus L-FK system is illustrated in Figure B.2. Simulations beginning with low numbers of wild-type individuals (the initial number of juveniles and males are reduced equally to

the number of females) go to extinction instead of the non-zero, stable equilibrium. This means that even when the release ratio is below r_c in a system, release could bring the population size to zero following population reduction by another method, such as spraying pesticides. In practice, however, it could be difficult to achieve enough suppression to bring the system below the unstable equilibrium. Additionally, in an area already without any wild-type individuals, ongoing release below r_c could prevent (small-scale) immigration of wild-type from re-establishing a population.

B.3 2-locus population genetics with additive fitness costs

As noted in the main text, if either the A or B alleles are at fixation in a population, the population genetics of 2-locus FK becomes effectively equivalent to 1-locus FK. For example, if B is at fixation, all that is needed is a single copy of A to cause lethality. For some sets of fitness parameters, either A or B is driven to fixation, while other parameters result in intermediate frequencies of both alleles. This effect is illustrated in Figure B.6A, which shows allele frequencies over time for both 1-locus and 2-locus L-FK with parameters $c_A = 0.75$, $s^M = 0.25$, and $h = 0.5$, and several different values of s^H. When $s^H = 0.1$ (left column), the B allele goes to fixation in most simulations, but the A allele can go to fixation in stochastic simulations despite having higher costs than B. The A allele would also go to fixation in deterministic simulations if starting from a much higher frequency than the B allele. At moderate costs (middle column), the B allele always goes to fixation, and at high costs (right column), both alleles reach intermediate frequencies. EFK exhibits similar behavior.

The possible allele outcomes for 2-locus L-FK with different combinations of fitness parameters are shown in Figure B.7. Here, simulations were conducted with additional transgenic adult male and females in the system at time 0. One set of simulations began with the A allele at 0.5 frequency in adults by adding AAbb adult males and females (in number equal to the wild-type adult equilibria) to the population, while another set similarly began with extra aaBB adults. When fitness costs are equal between the A and B alleles ($c_A = 0.5$, left column), the inheritance of A and B is completely symmetric in deterministic simulations. At high release ratios, the population goes to extinction regardless of initial condition (grey areas), but at lower release ratios, there are two possible outcomes. The A and B alleles are attracted toward equal intermediate frequencies when s^H is high compared to the release ratio (with the intermediate frequency indicated by shades of green). When s^H is small

and the release ratio is large enough, the system is attracted toward an equilibrium with whichever of the alleles begins at higher frequency reaching fixation (dark red areas), while the other allele reaches an intermediate frequency. In these cases, the hatching fitness cost is small enough to be outweighed by the influx of AA or BB in released adult males.

When $c_A > 0.5$ (middle and right columns), some parameter sets result in the B allele reaching fixation even in simulations beginning with a higher frequency of A (black areas). As c_A increases, the B allele is pushed to fixation in systems with larger values of s^H because the cost to the B allele is smaller. Likewise, the higher cost of the A allele decreases the (dark red) region where it reaches fixation. In a small region of parameter space, the A allele increases in frequency and imposes a large enough genetic load to cause population extinction, while the B allele would reach fixation without causing extinction (light right areas). As s^M increases, the primary effect is to reduce the effective release size, and thus larger releases are needed to drive an allele to fixation.

In terms of population suppression, the efficacy of 2-locus FK does not significantly depend on the allele frequency dynamics. The amount of suppression is slightly greater when either the A or B allele reaches fixation as long as there are fitness costs; the costs to the fixed allele impose a genetic load on the entire population. This phenomenon is responsible for sudden, minor decreases in the 2-locus equilibria in Figure 3.2A and Figure B.5A, which happens when the release ratios are large enough to cause one allele to become fixed. However, compared to 1-locus FK, 2-locus FK results in greater suppression even when neither allele is fixed because of the additional propagation of transgenic alleles, as described in the main text. As shown in Figure B.6B, the number of adult females over time reaches a smaller number in 2-locus L-FK than in 1-locus L-FK regardless of whether an allele reaches fixation, with a larger difference in suppression when s^H is larger.

B.4 Additional figures

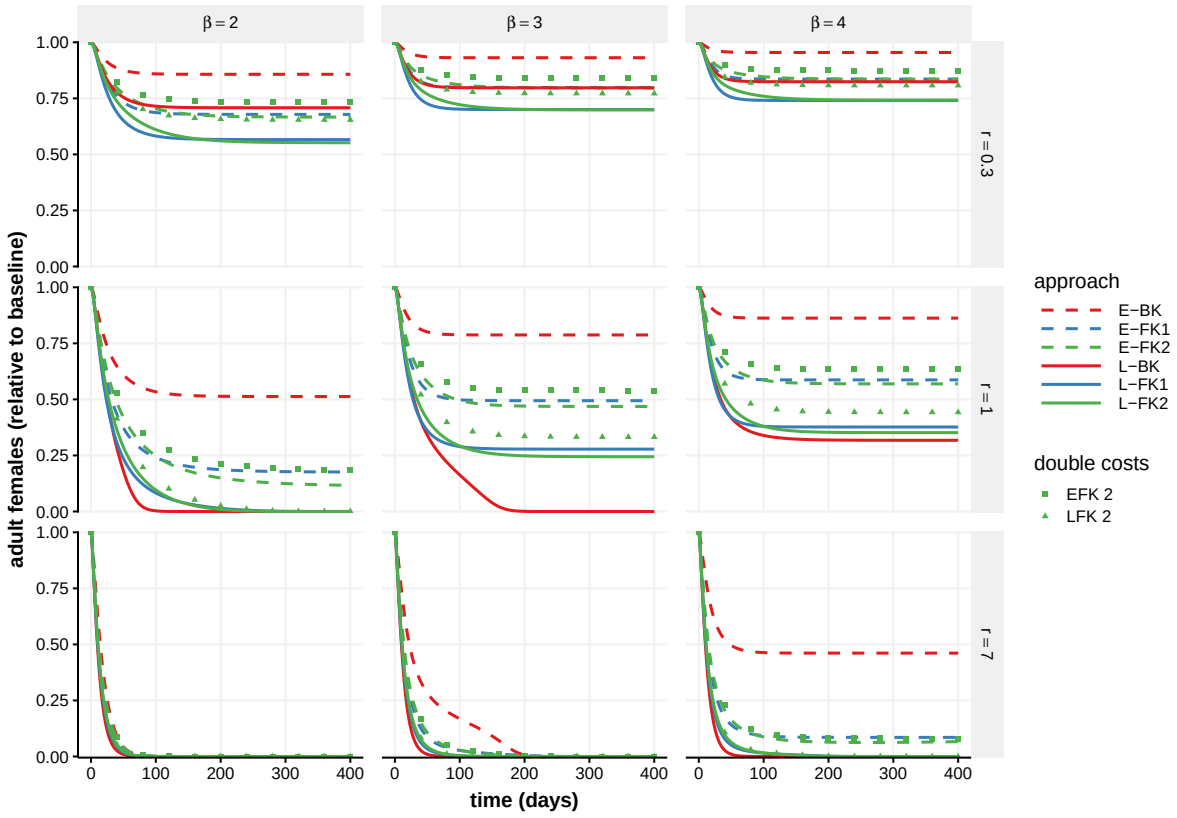

Figure B.1 Effect of transgenic releases on population size over time for various strengths of density dependence and release ratios. The number of viable adult females (relative to pre-release equilibrium) over time is plotted in deterministic simulations. Release ratios vary across rows, with adults for each genetic approach released at a continual weekly release ratio of 3:10, 1:1, and 7:1 transgenic males to the pre-release equilibrium wild-type males (middle row is equivalent to Figure 3.1). Line type and colors vary by approach as in the main text, with the addition of squares and triangles to show 2-locus E-FK and 2-locus L-FK with double the total fitness costs ($s^H = 0.4$, and $s^M = 0.2$, with $c_A = 0.55$). The remaining simulations use the default parameters from the main text: $s^H = 0.2$, $s^M = 0.1$, $c_A = 0.55$, and the remaining parameters as listed in Table 3.3. As strength of density dependence (β) increases and release size (r) decreases, the releases are less effective. If 2-locus FK has higher costs (symbols), it can become less effective than 1-locus FK, but at high release ratios, there is little difference in the number of total females over time.

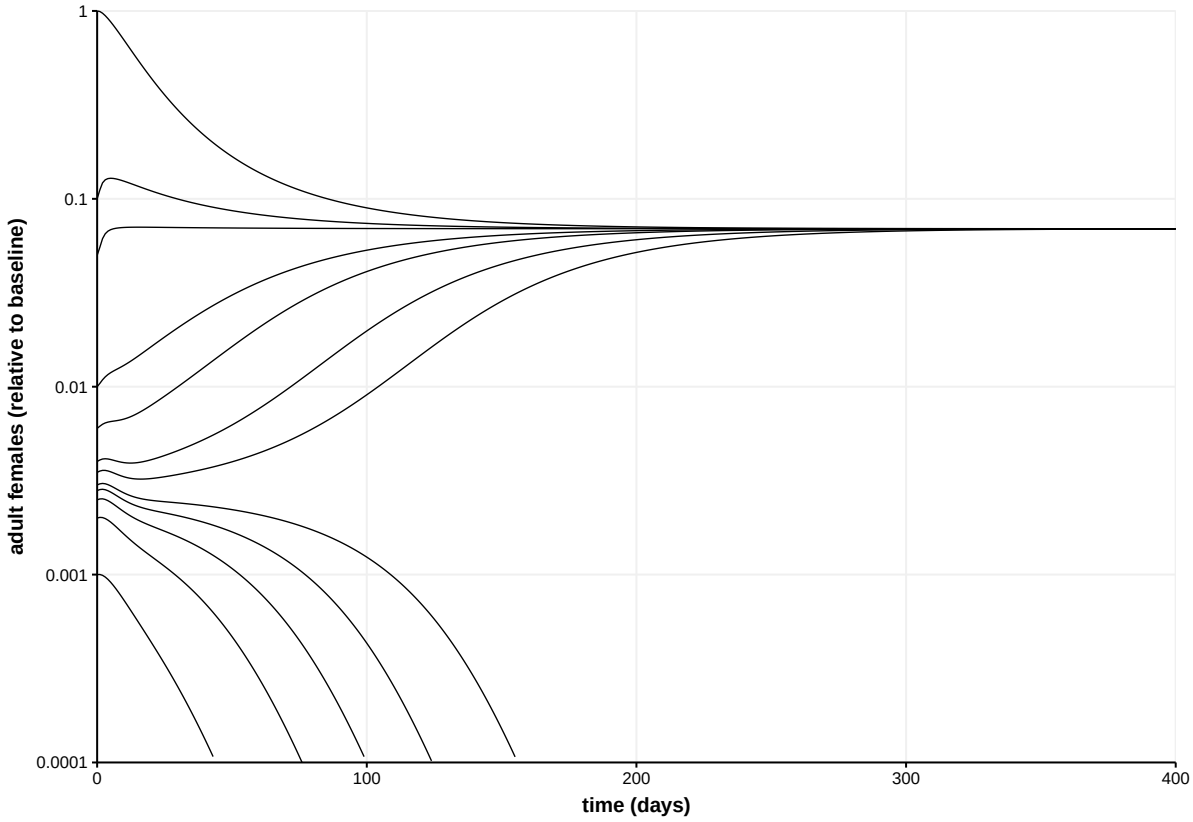

Figure B.2 Bistability of a 1-locus L-FK system with release below r_c. Each line represents a different starting number of wild-type females, with the initial number of adult males and juveniles are reduced by the same amount as the females (0.01 begins with 1/100 of the wild-type equilibrium count of each class). Lower initial counts could represent suppression from a method such as spraying of insecticides, or new immigration into a population that had already been eliminated. Each simulation has a release ratio of $r = 2$ beginning at time 0, and the number of females (relative to the wild-type equilibrium) over time is shown. A log scale for the number of adult females is used in order to illustrate behavior at small values. Simulations starting from low counts of wild-type result in population extinction, whereas higher initial counts result in the population reaching a non-zero equilibrium. Note that only a single dimension of the system is plotted, which means the time-series can exhibit different dynamics at the same value of relative adult females. Fitness parameters are equal to the default values from the main text: $s^H = 0.2$, $c_A = 0.55$, and $s^M = 0.1$.

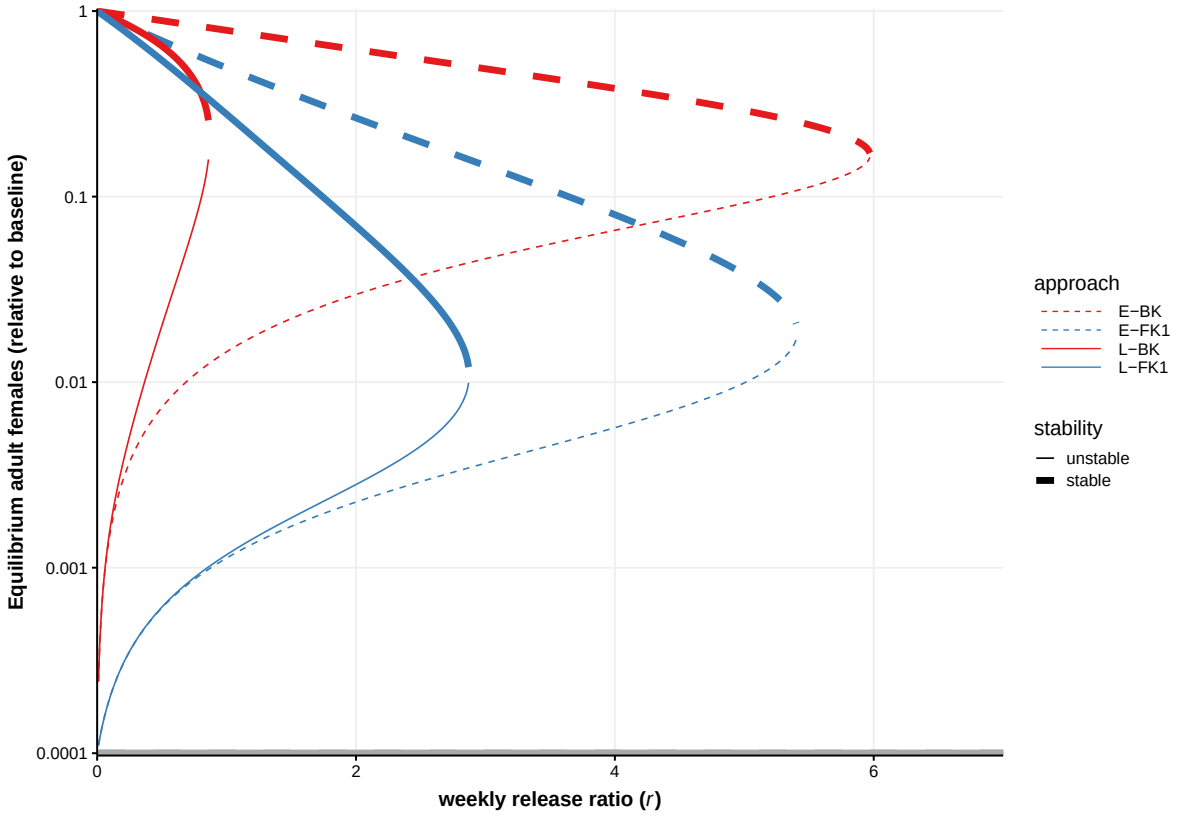

Figure B.3 Bifurcation diagram for 1-locus approaches, with the default parameter values from the main text. Analytical solutions and their corresponding eigenvalues were calculated. Here, the number of adult females in real, positive solutions is plotted for varying r. A log scale for the number of adult females is used in order to illustrate small equilibria. The grey line at the x-axis indicates the stable solution of 0 for all approaches (an extinct population with only released adult males remaining). Otherwise, approach varies with line color and type as in previous figures. The systems exhibit a saddle-node bifurcation, where at small values of r, there are two stable solutions (thick lines, one at zero and one positive) with an unstable solution (thin lines) in between. When r reaches r_c, the positive stable and unstable solutions collide. For larger r, the only equilibrium is the stable equilibrium at 0.

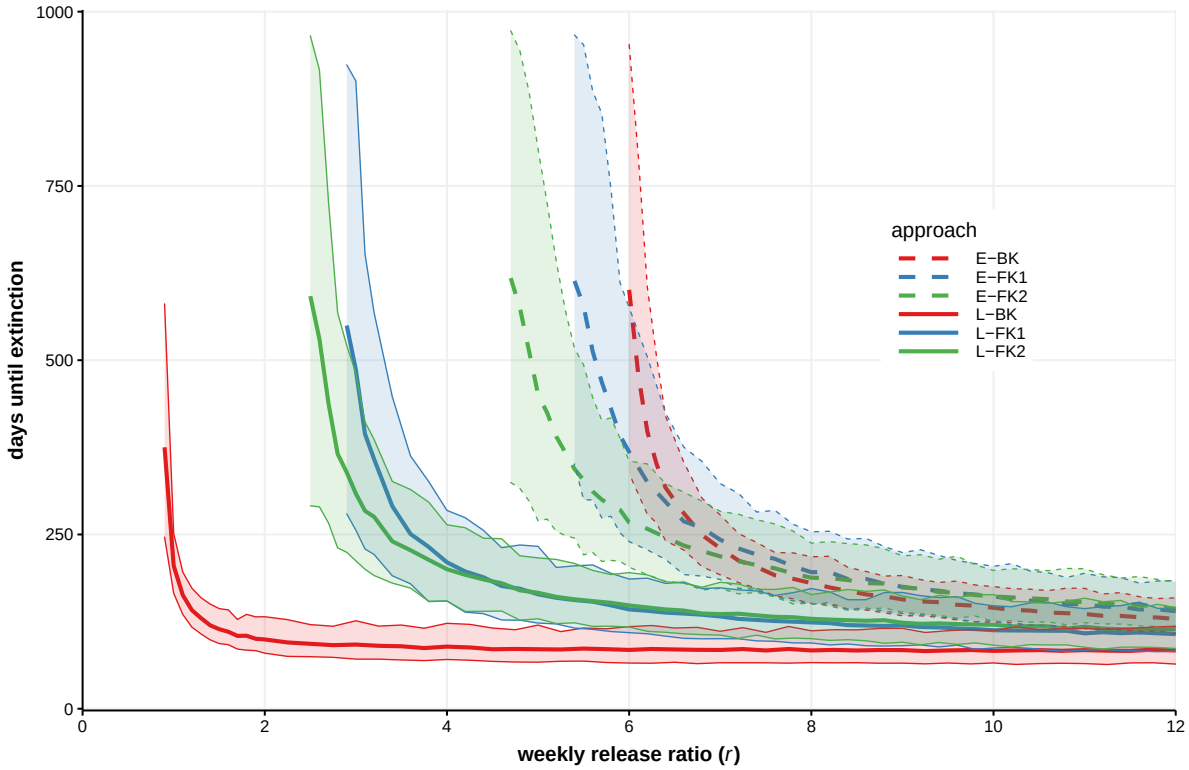

Figure B.4 Time until no adult females remaining in stochastic simulations. For each approach and value of r, 300 stochastic simulations were conducted for up to 1000 days. Thick lines show the mean time until there were zero adult females left in the population, and the ribbons show the 2.5% and 97.5% quantiles. Data is only plotted for values of r where at least 200 simulations resulted in 0 adult females by day 1000. Color and line type indicate approaches as in previous figures. Parameters are the default in the main text: $\beta = 3$, $s^H = 0.2$, $s^M = 0.1$, $c_A = 0.55$, and $h = 0.5$. The mean times are similar to Figure 3.2B in the main text, which shows that the time until falling below 1 adult female (from 2000 wild-type at equilibrium) in deterministic simulations is a good approximation of the stochastic simulations.

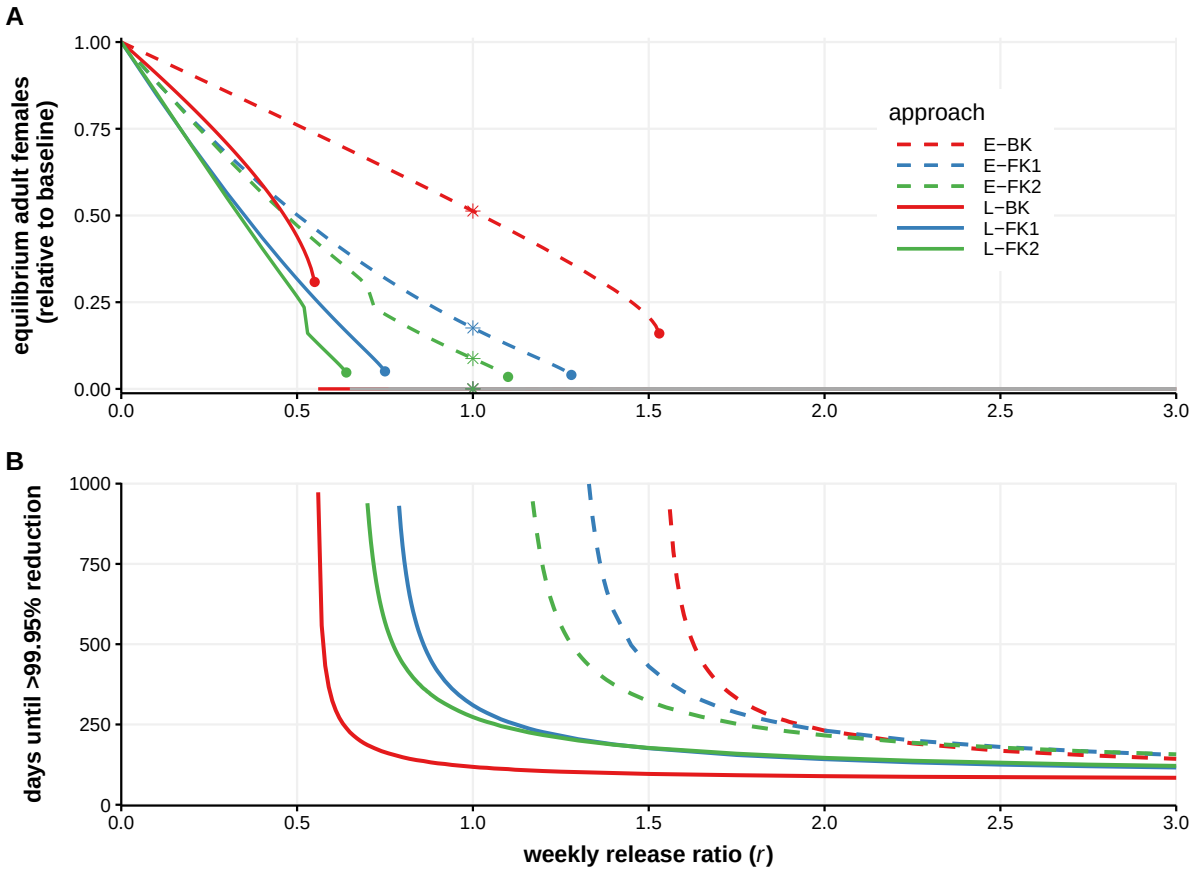

Figure B.5 Release outcomes across different release ratios with $\beta = 2$. A: Long-term, stable equilibria for number of viable adult females (relative to equilibrium) for different r, found by simulating the system of differential equations until at steady-state. The asterisks indicate $r = 1$, for which the equilibrium values correspond to the simulations in the left-column, center-row panel of Figure B.1. Each genetic approach exhibits a bifurcation at a critical release ratio, r_c, indicated by a solid circle. Above the r_c of each approach, that approach has an equilibrium at zero, making a discontinuity as seen with the red line for L-BK. Above $r = 0.64$, two or more approaches lead to extinction of the population and hence have equilibria at zero: this is indicated using a grey line. Each of the 2-locus FK (green lines) approaches exhibit a discontinuity where r becomes large enough to drive one of the alleles to fixation, decreasing the equilibrium population size (as explained in Appendix B.3 and illustrated in Figure B.6 and B.7). **B:** Time until the number of viable adult females is under 0.05% of equilibrium in deterministic simulations for different r. Color and line type match that of previous figures. Fitness parameters are equal to the default values from the main text: $s^H = 0.2$, $c_A = 0.55$, and $s^M = 0.1$.

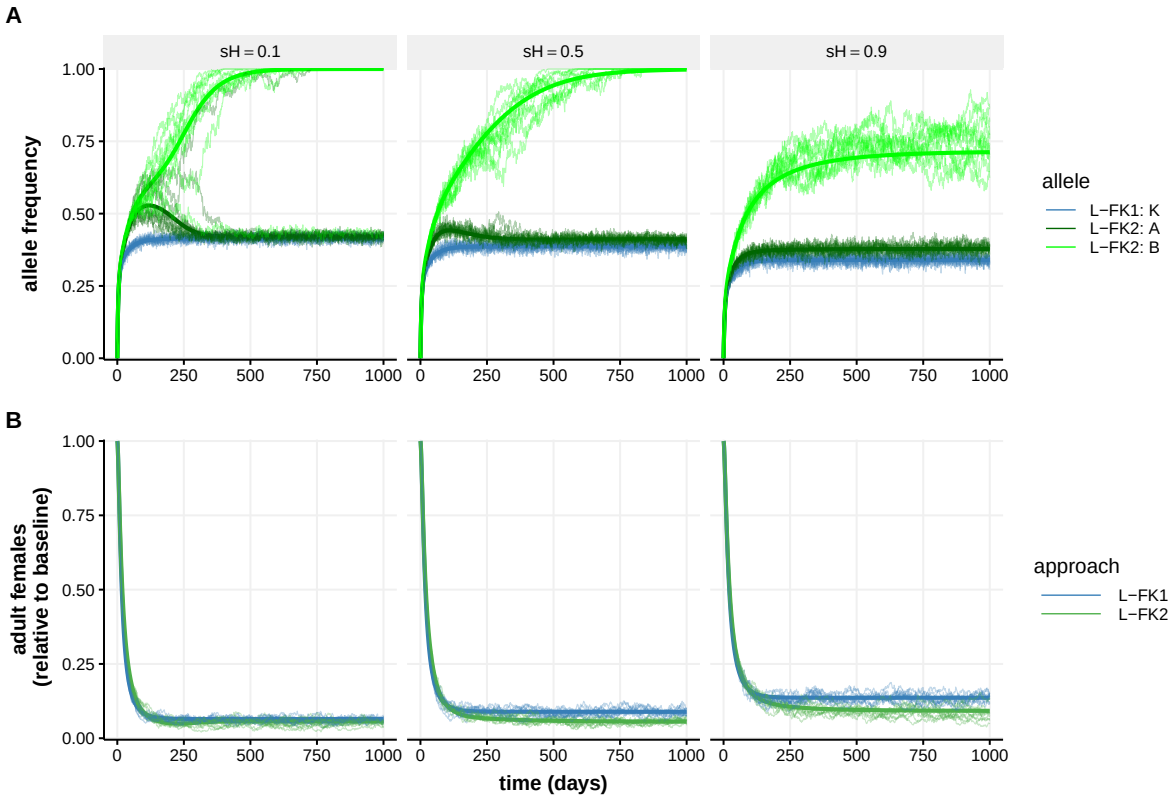

Figure B.6 2-locus L-FK juvenile allele frequencies over time at $r = 2$ for different hatching fitness costs. **A:** Allele frequency in juveniles shows different possible outcomes of A (dark green) and B (light green) allele frequencies for 2-locus LFK, and the K allele frequency in 1-locus LFK. The maximum possible allele frequency for 1-locus (or 2-locus when one allele is at fixation) is 0.5, when all individuals are heterozygous. **B:** Corresponding relative adult females in the population over time. Fitness parameters are $c_A = 0.75$, $s^M = 0.25$, and $h = 0.5$. Deterministic simulations (thick lines) and 10 stochastic simulations (thin lines) are shown. At low hatching fitness costs (left column), the B allele goes to fixation in most simulations, but the A allele can also go to fixation despite having higher costs than B. At moderate costs (middle column), the B allele always goes to fixation, and at high costs (right), both alleles reach intermediate frequencies. See Figure B.7 for the outcomes of allele frequencies across parameter space.

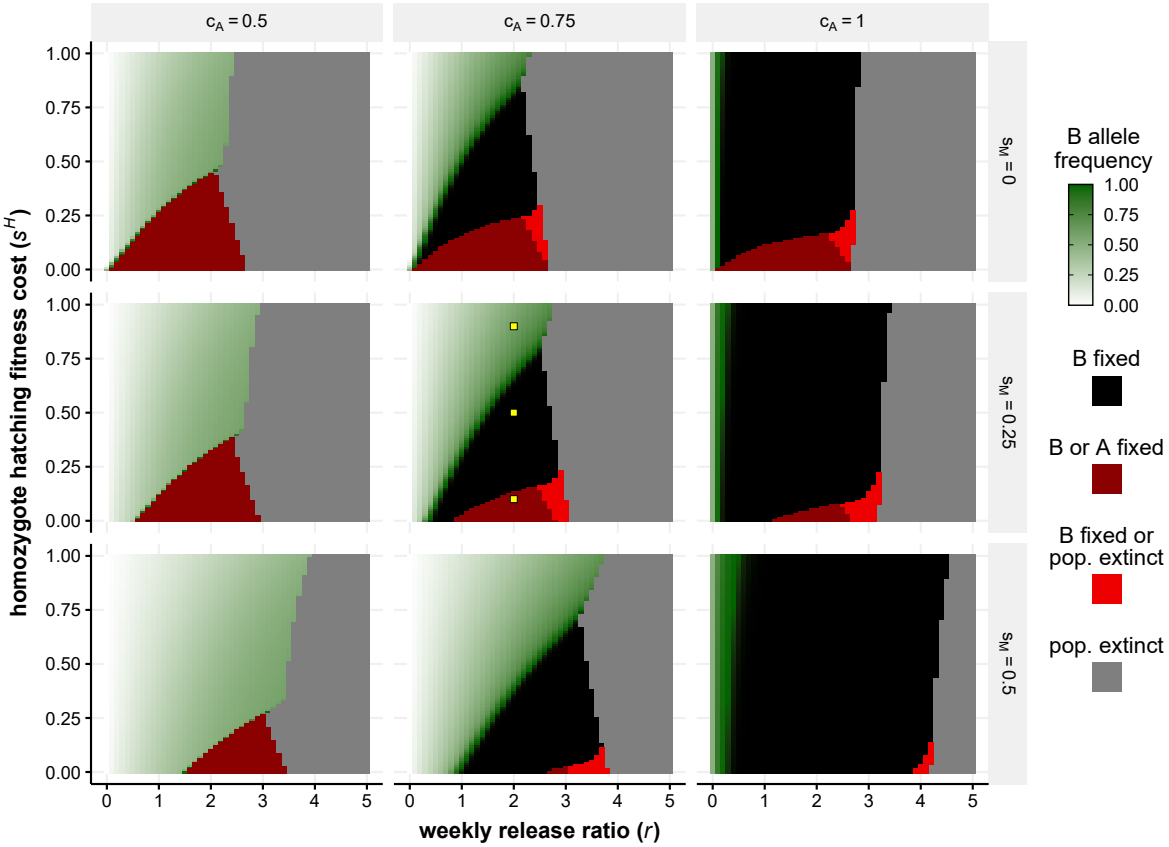

Figure B.7 Allele frequency outcomes from deterministic simulations of 2-locus L-FK for various fitness parameters and release ratios when degree of dominance $h = 0.5$. At each unique set of fitness parameters (male mating competitiveness fitness cost, s^M, differing between rows of panels; proportion of the total cost accounted for by the A allele, c_A, differing between columns of panels; hatching fitness cost, s^H, differing across the y-axis of each panel) and weekly release ratio (r, differing across the x-axis of each panel), long-term outcomes of two simulations with different initial conditions were compared. The systems were perturbed by adding transgenic adult male and females to the system at time 0. One simulation began with aaBB and the other AAbb males and females in equal numbers to the wild-type equilibrium number of males and females. In some regions of parameter space, r is large enough (toward the right side of each panel) to cause the system to go extinct in both simulations (grey regions). At lower r, there are four potential outcomes, three of which are illustrated in Figure B.6 (panel parameters indicated by yellow points): 1. the A and B alleles end at an intermediate frequency in both simulations (B frequency is indicated by shade of green), 2. the B allele goes to fixation in both simulations (black regions), 3. whichever allele was at higher frequency at time 0 goes to fixation (dark red), or 4. the B allele goes to fixation if starting at a higher frequency, while the system goes extinct if beginning with additional A alleles in the population (light red).

Appendix C

Particle Markov chain Monte Carlo

Simplifying notation from the main text, we let X_t and Y_t be random variables representing the vectors of true genotype frequencies of the population and the data, respectively, at generation $t \in [1, \ldots, T]$. The aim of particle Markov chain Monte Carlo (pMCMC) is to infer the joint posterior distribution of the states and parameters (represented by the random variable θ) given the data:

$$f(x_{1:T}, \theta | y_{1:T}) \propto g(y_{1:T} | x_{1:T}, \theta)\, \pi(\theta)\, f(x_{1:T} | \theta) \tag{C.1}$$

$$= \left(\prod_{t=1}^{T} g(y_t | x_t, \theta) \right) \pi(\theta)\, f(x_1 | \theta) \left(\prod_{t=2}^{T} f(x_t | x_{t-1}, \theta) \right), \tag{C.2}$$

where $g(y_t | x_t, \theta)$ is the likelihood of the data given the states of the system and the parameters at each generation, $\pi(\theta)$ is the prior for the parameters, $f(x_1 | \theta)$ is the density function of the initial frequencies, and $f(x_t | x_{t-1}, \theta)$ is the frequency transition density function. The posterior distribution is not possible to calculate directly, so we instead turn to MCMC approaches to sample from the posterior. Here we give a brief overview of the particle filter algorithm and particle marginal Metropolis-Hastings (PMMH) (Andrieu et al., 2010), which we implement via the nimble R package (Michaud et al., 2017).

C.1 Bootstrap filter

A particle filter can be used to approximate the marginal likelihood $\hat{L}\left(y_{1:T} | \theta\right)$ and to sample from $f(x_{1:T} | y_{1:T}, \theta)$. Broadly, the particle filter operates by using $f(x_t | y_{1:t}, \theta)$ to sequentially sample from

$f(x_{t+1}|y_{1:t+1},\theta)$ using importance resampling. In the simple case of the bootstrap filter, this is done by progagating particles forward according to the Markov model and weighting each particle based on the sampling likelihood. The steps of a bootstrap particle filter given θ^* (where each step is repeated for all $j = 1 \dots J$ particles) are as follows:

1. sample $x_1^{(j)} \sim f(x_1|\theta^*)$
2. calculate initial weights $W_1^j = g\left(y_1|x_1^{(j)},\theta^*\right)$

for $t = 2 \dots T$:

3. resample particles $x_{t-1}^{(j)}$ based on weights $W_{t-1}^{(j)}$
4. sample $x_t^{(j)} \sim f(x_t^{(j)}|x_{t-1}^{(j)},\theta^*)$
5. calculate weights $W_t^{(j)} = g(y_t|x_t^{(j)},\theta^*)$.

The marginal likelihoods at each time step $\hat{L}(y_t|y_{1:t-1},\theta^*) = \frac{1}{J}\sum_{j=1}^{J} W_t^{(j)}$ can be used to obtain an estimate of the marginal likelihood $\hat{L}\left(y_{1:T}|\theta^*\right) = \prod_{t=1}^{T} \frac{1}{N}\sum_{j=1}^{J} W_t^{(j)}$. Importantly, $\hat{L}$ is an unbiased estimate of the marginal likelihood. After completing the bootstrap filter, a randomly selected particle sequence can also be selected, giving a sample from the smoothing distribution $f(x_{1:T}|y_{1:T},\theta^*)$.

Resampling in Step 3 is important as it reduces particle degeneracy, a condition where very few particles have most of the weight. There are several options for resampling particles, the simplest of which is multinomial resampling, which uses weighted sampling based on particle weights, thus making high-weight particles likely to be replicated. In the steps above, resampling occurs at every time step, which allows the particles to be treated as being of equal weight. However, it can be advantageous to resample only when particle degeneracy becomes high. The amount of degeneracy is related to the effective sample size of particles, $ESS_t = \frac{1}{\sum_j \left(w_t^{(j)}\right)^2}$, where $w_t^{(j)}$ is the normalized weight of particle j at time t.

C.2 Particle marginal Metropolis-Hastings

The Metropolis-Hastings (MH) algorithm is used to stochastically sample from a posterior distribution when direct sampling is not possible. In the MH algorithm, a Markov chain is generated by proposing candidate parameters at random from a proposal distribution around the current

parameter value. The candidate parameter is accepted with probability equal to the ratio of the likelihood of the candidate value given data to the likelihood of the current value. Intuitively, this allows the random walk parameter values to explore parameter space while spending more time in areas with high probability density, and the Markov chain converges to the targeted posterior distribution.

PMMH uses the estimates of the marginal likelihood from the particle filter in the MH acceptence ratio in order to jointly sample from the joint posterior distribution from (C.1). Andrieu et al. (2010) show that unbiased estimates of the marginal likelihood obtained from a particle filter, even with high variance due to few particles, produce samples from the appropriate target distribution and converges to the desired joint posterior. The PMMH algorithm, run for M iterations, is as follows:

1. choose initial parameter values $\theta^{(0)}$
2. obtain the marginal likelihood estimate $\hat{L}^{(0)}\left(y_{1:T}|\theta^{(0)}\right)$ using the bootstrap filter

for $m = 1 \dots M$:

3. generate proposal $\theta^* \sim q(\theta|\theta^{(m-1)})$
4. obtain the marginal likelihood estimate $\hat{L}^{(m)}\left(y_{1:T}|\theta^*\right)$ using the bootstrap filter
5. evaluate $a = \frac{\hat{L}^{(m)}\left(y_{1:T}|\theta^*\right)\pi(\theta^*)}{\hat{L}^{(m-1)}\left(y_{1:T}|\theta^{(m-1)}\right)\pi\left(\theta^{(m-1)}\right)} \frac{q\left(\theta^{(m-1)}|\theta^*\right)}{q\left(\theta^*|\theta^{(m-1)}\right)}$
6. generate a random number $u \sim \text{Uniform}(0,1)$
7. if $a > u$ set $\theta^{(m)} = \theta^*$, else set $\theta^{(m)} = \theta^{(m-1)}$.

The proposal distribution, $q(\theta|\theta^{(m-1)})$, can take a variety of forms, but the multivariate normal distribution is standard. The covariance matrix and scale of the proposal distribution can be difficult to tune in some cases, though adaptive approaches can be employed to handle this. Additionally, at each iteration, after running the bootstrap filter, a particle can be sampled with its full state history to obtain a proposal sample from the smoothing distribution, i.e., $f(x_{1:T}|y_{1:T}, \theta^*)$. (Note that this sample is relevant to the convergence of PMMH.)

After enough iterations, the samples will be drawn from the stationary distribution, which is the posterior distribution. Thus, it is important to ensure convergence of the Markov chains. It is

standard to discard early values in the chain before convergence (burn-in), and the chains must then be run for many iterations (especially if there is high autocorrelation, which reduces the effective length of the chain because each sample is not independent) to achieve adequate sampling of the posterior distribution. Multiple, independent chains are also typically run, and the extent that the chains converge can be quantified by using the Gelman-Rubin statistic (Gelman, Rubin, et al., 1992).