Locally Fixed Alleles: A method to localize gene drive to island populations

Jaye Sudweeks [a], Brandon Hollingsworth [b], Dimitri V. Blondel [c], Karl J. Campbell [d], Sumit Dhole [e], John D. Eisemann [f], Owain Edwards [g], John Godwin [c,h], Gregg R. Howald [d], Kevin Oh [f,i], Antoinette J. Piaggio [f], Thomas A. A. Prowse [j], Joshua V. Ross [j], J. Royden Saah [d,h], Aaron B. Shiels [f], Paul Thomas [k], David W. Threadgill [l], Michael R. Vella [b], Fred Gould [e,h] and Alun L. Lloyd [a,b,*]

(a) Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA
(b) Biomathematics Graduate Program, North Carolina State University, Raleigh, NC 27695, USA
(c) Department of Biological Sciences, North Carolina State University, Raleigh, NC 27695, USA
(d) Island Conservation, 2100 Delaware Ave., Suite 1, Santa Cruz, CA 95060 USA
(e) Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA
(f) National Wildlife Research Center, US Department of Agriculture, Fort Collins, CO 80521, USA
(g) CSIRO Land & Water, Centre for Environment and Life Sciences, Floreat, WA, Australia
(h) Genetic Engineering and Society Center, North Carolina State University, NC 27695, USA
(i) Department of Microbiology, Immunology & Pathology, Colorado State University, Fort Collins, CO 80523, USA
(j) School of Mathematical Sciences, The University of Adelaide, Adelaide, SA 5005, Australia
(k) The Robinson Research Institute and School of Medicine, The University of Adelaide, Adelaide, SA 5005, Australia
(l) Department of Molecular and Cellular Medicine, Texas A&M University, College Station, TX 77843, USA

* Corresponding author, email: alun_lloyd@ncsu.edu

Abstract

Invasive species pose a major threat to biodiversity on islands. While successes have been achieved using traditional removal methods, such as toxicants aimed at rodents, these approaches have limitations and various off-target effects on island ecosystems. Gene drive technologies designed to suppress a population provide an alternative approach, but the potential for drive-bearing individuals to escape from the target release area and impact populations elsewhere is a major concern. Here we propose the "Locally Fixed Alleles" approach as a novel means for localizing a suppression drive to an island population that exhibits significant genetic isolation from neighboring populations. Our approach is based on the assumption that in small island populations of rodents, genetic drift will lead to multiple genomic alleles becoming fixed. In contrast, multiple alleles are likely to be maintained in larger populations on mainlands. Utilizing the high degree of genetic specificity achievable using homing drives, for example based on the CRISPR/Cas9 system, our approach aims at employing one or more locally fixed alleles as the target for a gene drive on a particular island. Using mathematical modeling, we explore the feasibility of this approach and the degree of localization that can be achieved. We show that escape of the drive to a neighboring population in which the target allele is not fixed will at most lead to modest transient suppression of the non-target population across a wide range of parameter values.

Introduction

Genetic modification of pest species has been suggested as a means to address a wide variety of pest problems, including those impacting human health, pre- and post-harvest crop losses, and conservation of endangered species (Curtis 1968, Burt 2003, Gould 2008, Esvelt et al. 2014, Champer et al. 2016). One approach to genetic pest management involves introducing into the pest genome a DNA sequence that causes its own over-representation in future generations by inducing super-Mendelian inheritance; generally referred to as gene drive. An engineered gene coding for a desirable trait, e.g. one that renders the pest species less troublesome (e.g. reduction of vector competence of a mosquito species) can be linked to the gene drive sequence in order to increase its frequency in the population. Alternatively, the gene drive sequence can be engineered to disrupt a gene critical to survival or reproduction and thereby suppress a population. Ever since such drives were proposed, concerns have been raised in the peer-reviewed literature and in the popular media about unintended consequences of releases and ethical dimensions of the work. These include failure of the genetic construct being driven, but also the spread of gene drives beyond the region in which spread is intended and approved by the local population and governing authorities (Spielman 1994, Alphey et al. 2002, Burt 2003, Macer 2005, Marshall 2010, Akbari et al. 2015, Esvelt and Gemmell 2017, Marshall and Akbari 2018). This concern is most acute in the case of drives that are designed to suppress and eliminate a population: spread of such a drive could have a risk of leading to global eradication of a species that is a pest in one area but a valuable component of an ecosystem in other areas (e.g. mice and rats). As a result, there is much interest in the ability to design gene drives that exhibit spatial localization, i.e. ones that have the ability to spread in a given region but will not spread globally.

Several approaches have been suggested to achieve localization of gene drives. Drives that exhibit an invasion threshold, such as the engineered underdominance (EU) approach, provide a natural means to achieve localization (Davis et al. 2001, Altrock et al. 2010). These drives exhibit frequency-dependent dynamics where the drive can only spread if its frequency exceeds a particular level—the invasion threshold. Below this level, the frequency of the drive will decrease, leading to its loss from the population. Spread of a threshold drive across a patchy environment is more difficult, and becomes highly unlikely or even impossible when the invasion threshold is 50% or higher (Barton 1979, Barton & Turelli 2011). Other approaches have been suggested to achieve localization, including Killer-Rescue (Gould et al. 2008), Multi-locus assortment (Rasgon 2009) and daisy-chain drive (Noble et al. 2016). Theoretical analysis has suggested the ability of daisy-chain drive to simultaneously achieve spread and localization to a single area is only possible in a limited set of circumstances, and this concern also pertains to some other gene drives developed to be localized (Dhole et al. 2018). While all of these localized drives could change characteristics of pests in a population, their ability to locally suppress populations is questionable (Alphey and Bonsall 2014, Dhole et al. 2018, although see Khamis et al. 2018).

In this paper, we propose a localization method, the “Locally Fixed Alleles” (LFA) approach, that can be utilized for relatively small populations that exhibit a significant degree of genetic isolation from other populations. This method is particularly suited for suppression of pest species from small oceanic islands, where the target population has small effective population size and for which there is naturally limited gene flow with other populations. While multiple alleles are expected to be commonly maintained at loci in large populations, genetic drift in small island populations is predicted to result in fixation of alleles at some loci in the genome (Kimura and Ohta 1968, Frankham 1997).

Utilizing the high degree of genetic specificity of homing drives based on the CRISPR/Cas9 system (Esvelt et al. 2014), our approach aims at employing one or more locally fixed alleles as the target for a gene drive on a particular island. Such a drive can spread to individuals carrying that allele, but individuals that do not have that specific allele are naturally resistant to the drive. By design, we would search for alleles that are fixed in the population on the target island but not fixed in populations beyond that island. Consequently, the drive could be expected to result in only limited transient suppression beyond the island. A special case of this approach, the "private allele" (PA) approach (dubbed "precision drive" by Esvelt et al. 2014; see also Esvelt and Gemmell, 2017), occurs when the target allele is specific to the target population, but absent from other populations. We emphasize that the LFA approach does not require the target allele be a private allele, simply that it not be fixed in non-target populations.

The Locally Fixed Alleles approach can be used with a variety of different gene drives (e.g. standard homing drives, sex-biasing drives, and so on). Here, for simplicity, we illustrate the method using a standard homing drive (Burt 2003, Esvelt et al. 2014) aimed at population suppression. We describe this in the setting of removal of a rodent species, such as the house mouse, *Mus musculus*, from an island. Invasive mice, and other rodents, are a particular concern for species conservation (Howald et al. 2007), having significantly impacted many island ecosystems, including causing extinctions of endemic island vertebrate, invertebrate and plant species (Cuthbert and Hilton 2004, Towns et al. 2006, Campbell et al. 2015). Although an island release would involve procedures that attempt to confine the gene drive mice to the island, unintended escape of these mice must be considered a possibility. Here, we use mathematical modeling to explore the impact that escape of drive individuals would have on mainland populations.

This paper is organized as follows: we first introduce the mathematical model, then briefly describe the single-patch (island-only) dynamics before discussing those seen in a two-patch (island-mainland) setting.

Methods

We employ a continuous-time non age-structured island-mainland model that describes the population dynamics and genetics of two populations. As our primary concern here is the impact of escape from the island, we assume unidirectional migration from the island to the mainland. (We remark that migration from mainland to island would be an important consideration in the period following successful suppression or eradication from the island, but this is not the topic of this study.) Throughout, we assume a 50:50 sex ratio and so track numbers of female individuals. For an *n*-genotype system, denoting genotypes with a subscript and denoting island population numbers with superscript *I* (N^I_i) and mainland population numbers with a superscript *M* (N^M_i), we have

$$\frac{dN_i^I}{dt} = f^I(N_1^I, \cdots, N_n^I) - \mu N_i^I \tag{1}$$

$$\frac{dN_i^M}{dt} = f^M(N_1^M, \cdots, N_n^M) + \mu N_i^M. \tag{2}$$

Here, the functions f^I and f^M describe the population dynamics and genetics that occur on island and mainland, respectively, and we assume unidirectional migration with per-capita migration rate equal to μ.

We assume that island and mainland populations both undergo random mating (i.e. are well-mixed) and exhibit logistic-type population dynamics. Per-capita birth and death rates both change linearly with population size, with different coefficients on mainland and island. Within either the island or mainland, and in the absence of migration, genotype dynamics are described by

$$\frac{dN_i}{dt} = b_i(t) \max\left(1 - q\sum_j N_j, 0\right) - \rho N_i - \alpha N_i \sum_j N_j. \tag{3}$$

Here, superscripts denoting the location have been suppressed for clarity on both state variables and parameters. The functions $b_i(t)$, described below, depict the genotype-specific birth rates in the absence of density dependence. For an entirely wild-type population, $b(t)$ would equal λN, where λ is the per-capita fecundity rate. The $b_i(t)$ are multiplied by a function that describes the linear density-dependent decline in per-capita birth rates with total population size. (Note that the max function is required to ensure that birth rates remain non-negative.) Per-capita death rates are assumed to increase linearly with overall population size but be independent of genotype. The density-independent component of the per-capita death rate (i.e. the reciprocal of the average lifespan when the population is at low density) is written as ρ, while the coefficient α describes the density-dependent linear increase in per-capita mortality. With these population dynamics, a single patch has a carrying capacity of $N = \rho(R_0 - 1)/(\lambda q + \alpha)$, where the basic reproductive number, R_0, (i.e. the average number of female offspring of a female over its lifetime, at low population density) is equal to λ/ρ.

Population genetics is determined by the functions $b_i(t)$ which give the genotype-specific birth rates (c.f. Robert et al. 2013) before accounting for the effects of density-dependence

$$b_i(t) = \lambda w_i \sum_j \sum_k \frac{P(i|j,k) N_j N_k}{\sum_l N_l}. \tag{4}$$

Here, λ is the baseline per-capita birth rate for females. $P(i|j,k)$ gives the proportion of offspring from a mating involving individuals of types j and k that will have genotype i. The effects of the gene drive on biasing of inheritance are coded into these quantities. (The 216 entries of $P(i|j,k)$, together with a fully written out set of model equations, appear in the Maple worksheet in the Supplementary Information.) The w_i describe genotype-specific relative fitnesses, which we assume here to act at the embryonic stage and be equal in males and females of a given genotype.

As mentioned above, we use a simple homing-based suppression drive to illustrate the Locally Fixed Alleles method. We consider three alleles: the drive allele (D), the susceptible (S) allele, i.e. the target for the drive, and an allele that is resistant to the drive (R). All individuals in the target (island) population initially have the genotype SS, but those in the non-target (mainland) population can have SS, RS or RR genotypes. We assume only a single resistance allele in the mainland population. A large mainland

population could have a number of different alleles that would be resistant to the drive, but they would all act similarly in being unaffected by the drive. Therefore our 3-allele model is sufficient for capturing dynamics of these cases. Note that here we are considering natural resistance to the drive, rather than drive-resistant alleles that are generated de novo as a result of the drive, e.g. by non-homologous end joining during homing. As a consequence, our model over-estimates the ability of the drive to spread and suppress populations (Prowse et al 2017). Given that we are trying to evaluate the risk and implications of escape of drive, this assumption is conservative for our purposes. Homing is assumed to occur during gametogenesis, meaning that successful homing leads to an SD heterozygote individual giving rise to only D gametes (Homing exclusively in the germline at any point during development would have the same consequences). Successful homing occurs with probability e. The fitness, w_i, of SS, RR, or SR individuals is assumed to be 1. The fitness of DD individuals is (1-s), and the fitness of SD and RD individuals is (1-hs), where s is the fitness cost of the drive and h is the degree of dominance of the fitness cost. For instance, a recessive lethal drive has s = 1 and h = 0. Notice that we assume there is no fitness cost for the RR individuals that occur naturally in the non-target population.

We employ parameters (see Table 1) that are appropriate for *Mus musculus* populations, largely based on those used by Backus & Gross (2016) (see also Nathan et al. 2015). For this set of parameters, the basic reproductive number of a wild-type population is 3.5. Assuming an island of area 6 hectares, these parameters lead to an equilibrium population size of 1000 females. We assume that the mainland has a population size that is 100 times larger than this.

Results

Single Patch Model

The single-patch behavior of the standard homing drive that we use here has been well-studied previously (Burt 2003, Deredec et al. 2008, Alphey and Bonsall 2014, Unckless et al. 2015). Depending on the drive parameters s, h and e, several qualitatively different behaviors can occur following release of drive into an otherwise entirely susceptible population: guaranteed fixation of the drive, guaranteed loss of the drive, co-existence of drive and susceptible alleles at a stable polymorphic equilibrium or invasion threshold behavior (i.e. drive either goes to fixation or is lost, depending on its initial frequency) resulting from the existence of an unstable polymorphic equilibrium. Suppression of the population is possible when the drive remains in the population (going to a stable equilibrium with a positive frequency—either fixation or a polymorphic equilibrium) and imposes a cost that is sufficiently high to bring the basic reproductive number of the population below one.

We illustrate these dynamics using two sets of parameters: one for which the drive can spread through a susceptible population regardless of its frequency (no invasion threshold scenario), and another for which the drive can only spread when its frequency exceeds an invasion threshold (invasion threshold scenario). In both cases we consider a fitness cost of s = 0.8 and a homing probability of e = 0.95. The two scenarios differ only in the dominance, h, of the drive. For the no invasion threshold scenario we take h = 0.3, while for the invasion threshold scenario we take h = 0.8, which leads to an invasion threshold frequency of approximately 0.621 (corresponding to a ratio of approximately 1.64:1 drive:wild-type individuals). Releases occur into a population that is at carrying capacity, and population sizes are assessed relative to this carrying capacity.

Figure 1 shows the population dynamics that result from drive releases in these two scenarios. For the no invasion threshold scenario (green curve), even a small release of drive individuals, so that the initial relative population size is only just above one, leads to spread of the drive allele and hence reduction and eventual elimination of the population. For the invasion threshold scenario (blue and red curves), spread of the drive, and hence the fate of the population, depends on whether the release frequency of the drive is above (blue curve) or below (red curve) the invasion threshold. A sufficiently large release (blue curve) leads to fixation of the drive and elimination of the population, while an insufficient release (red curve) leads to loss of the drive and recovery of the population following a transient period of reduction.

Island-Mainland Model

We now turn to the main question of how migration from an island population on which drive individuals have been released will impact a mainland population that has a mix of susceptible and resistant individuals. To present something approaching a worst-case scenario, we first assume that the frequency of resistant individuals on the mainland is rather low, with a resistance allele frequency of 5% and susceptible allele frequency of 95%. We take our island population to be $1/100^{th}$ the size of the mainland population, with unidirectional migration from the island to the mainland occurring at per-capita rate of .012 per year, corresponding to movement at the rate of one mouse per month at baseline.

For the invasion threshold drive scenario, we consider an island release that is above threshold, so that the drive approaches fixation in the long run (Figure 2, upper panel, red curve) and the island population is successfully eliminated (Figure 2, upper panel, blue curve). Even though drive individuals migrate from the island to the mainland, the large size of the mainland population means that the drive frequency remains small and never exceeds the invasion threshold. Thus, the drive cannot spread on the mainland, leaving the size and genetic composition of the mainland population largely unaffected.

The situation is different in the no invasion threshold scenario. The drive is able to spread on the mainland even when present at the low frequencies that arise due to migration of drive individuals from the island. Individuals with a resistance allele are unaffected by the drive, however, and so spread only occurs through the susceptible individuals. Again, the drive spreads on the island (Figure 3, upper panel, red curve) causing successful elimination of the island population (Figure 3, upper panel, blue curve). On the mainland, drive spreads through the susceptible portion of the population (Figure 3, lower panel, red dashed curve) causing a reduction in the population size (Figure 3, lower panel, blue curve) due to fitness costs incurred by drive-bearing individuals. Because they are unaffected by the drive, resistant individuals benefit from having a larger relative fitness in the presence of drive individuals and so their frequency increases (Figure 3, lower panel, red dot dashed curve), while the drive frequency decreases. As the frequency of resistant individuals increases, the average fitness of the population returns to the level seen initially, and density-dependent dynamics returns the population to its original size. In this setting, we see a transient reduction in the mainland population size before a recovery due to the presence of resistance. The genetic composition of the mainland undergoes a shift during this process, with a reduction in the frequency of susceptible alleles (although not their elimination) and a corresponding increase in the frequency of resistance (see Supplemental Text and Supplemental Figure 3).

Figures 4 and 5 explore how the magnitude of population suppression and the maximum allele frequency of the gene drive observed on the mainland depend on the initial level of resistance on the mainland and the migration rate from the island to the mainland in the no invasion threshold scenario. Naturally, the level of transient suppression (fractional reduction below initial equilibrium) on the mainland depends strongly on the frequency of resistant and susceptible individuals there (Figure 4), although an important observation is that the level of transient suppression is considerably lower than the initial frequency of susceptible individuals. Similarly, the initial composition of the mainland population impacts the peak drive frequency reached on the mainland, while the level of migration again has little impact (Figure 5). For the assumed demographic and drive parameters, 10% or higher levels of resistance on the mainland lead to a transient population suppression of at most 20% and peak drive frequency below 30%. Alternative choices for demographic and drive parameters would impact these numbers. Obviously, in the PA setting, drive would be unable to spread on the mainland, but Figures 4 and 5 reveal that even in situations for which the susceptible allele is not absent, but only fairly uncommon on the mainland, the levels of suppression that would result would be small, and quite likely smaller than natural fluctuations in population size that might result from demographic or environmental stochasticity.

The peak level of suppression and the maximum drive frequency on the mainland occur at roughly similar times. The timing of these events is relatively insensitive to the initial frequency of resistance on the mainland and the migration rate over a wide range of conditions (see Supplementary Text and Supplementary Figures 1 and 2), with dynamics playing out over the course of 10 to 20 years. High initial levels of resistance, however, can lead to much slower spread of the drive on the mainland. In the Supplementary Text we show that drive cannot invade when the initial frequency of the resistant allele is above $1 - hs/\{e(1-hs)\}$. Invasion of drive will occur only slowly (and with a low peak level) when the frequency of resistance is not too far below this level. Supplemental Figures 4 and 5 show that the peak level of suppression on the mainland is almost independent of the size of the release on the island, while the time until this suppression occurs is only weakly dependent on the release size.

Discussion

There is a fundamental tension between the ability of a gene drive to spread locally within a target area and its ability to invade populations beyond that area. For a non-threshold gene drive designed for suppression, the potential impact outside the targeted area is of serious concern even if there is a likelihood that resistance to the drive will evolve before irreversible damage is done. Indeed, such concerns were strongly expressed in many of the seminal early papers on gene drive (e.g. Burt 2003, Gould 2008). Clearly, mechanisms that give some level of control over the spread of a gene drive are necessary prerequisites before the deployment of gene drive for suppression or eradication of a targeted population of a widespread species. Utilizing the high degree of genetic specificity exhibited by homing drives, the LFA approach provides one option to limit the impact of unintended spread of a drive. In the fortunate, although perhaps rather unlikely special case of “private alleles” where all individuals in non-target populations lack the susceptible allele and the susceptible allele is fixed on the island, the drive would be completely confined to the target population.

The LFA approach is related to two-step gene drive approaches that first spread a target allele into a population and then release a second drive that is specific to the target allele. For example, Esvelt *et al.*

(2014) suggested this approach for island populations, arguing that containment would happen provided that release of the second drive occurred before any individuals bearing the first drive escaped the island. The uncertainty involved in this scenario could be problematic for some stakeholders and regulatory authorities. In our approach, naturally occurring targeted alleles are expected to be fixed within small island populations due to genetic drift or founder effects that have occurred in the past. These targeted alleles may be present in other locations, but as long as they are not fixed, impacts on populations in those locations should be transient.

As demonstrated by other mathematical models (e.g. Magori and Gould 2006, Dhole et al. 2018) some spatially limited gene drive mechanisms are unlikely to function well for suppressing populations. In this regard, the LFA approach stands out for its predicted ability to suppress a small isolated population without impacting other populations.

Our analysis employed a deterministic model, which allows for continuous migration from island to mainland. This model describes the average behavior of the system and so gives a good indication of expected dynamics in a large well-mixed mainland population. It describes the magnitude of suppression and its timing, but does not account for the discrete numbers of individuals in populations. Given that migration from the island to mainland is anticipated to occur as relatively infrequent events that involve small numbers of individuals, a stochastic model would be more realistic. Such a model would allow questions about the likelihood of the occurrence of migration events and of resulting transient spread of the drive, and distributions of the magnitude and timing of any transient suppression that results. However, we would not expect the qualitative results of such a model to predict a qualitatively different outcome from the current model in terms of impacts on the overall mainland population.

It is important to recognize that one of the outcomes of mice arriving on the mainland that have a homing drive with no threshold is that the frequency of the resistant allele will increase. If after the mice are eliminated from the island a mouse or a few mice carrying the now more common resistant allele migrate to the island, it is likely that the new population on the island would no longer be a good target for the previously used construct. However, if recolonization was started by one or a few individuals, founder effects and drift would be likely to result in other fixed alleles that could be targeted.

There are a number of technical and societal challenges in moving the LFA from concept to application. While there is empirical evidence of lower polymorphism in small island populations (e.g., Browne 1977, White & Searle 2007, Harradine et al. 2015), not every fixed allele in an island population will be a good target for a CRISPR-based gene drive. Appropriate targets will depend on the intended nature of the drive, impacting the number of available loci. Furthermore, these alleles will optimally have at least two sites that can be targeted by the CRISPR-CAS nuclease complex to decrease the chance of resistant alleles arising through non-homologous end joining. It will be critical that full genomes of individuals be scanned for the best possible targets and those targets will need to be scrutinized for potential problems. Also, before moving ahead with any genetic engineering, island populations will need to be sampled and evaluated extensively to ensure that the probability that a presumed fixed allele is not truly fixed on the target island is below an acceptable threshold. This will not require full genome scanning, but even focused sequencing of the target loci, identified from genome scanning of a subset of individuals, in a large proportion of the target population will require time and effort and will never provide 100% assurance.

At a societal level, lack of an ecological impact of an LFA in a non-target location of the specific target species must be assured. Models that include details of the species biology, population structure and genetics will be critical in moving toward regulatory approval. But assurance of lack of an ecological impact of LFA constructs on the non-targeted populations may not be sufficient for some stakeholders who will have a concern with any number of engineered individuals arriving in their area.

Beyond direct impacts of an LFA, researcher must be cognizant of the fact that in developing the technology for LFA in a problematic species, they are also developing tools that could be used by others to construct unrestricted gene drives in that species. There will always be tradeoffs in developing new technologies and there are unlikely to be simple decisions. Vigilance and input from diverse stakeholders "early and often" (Piaggio et al. 2017) will be critical in coming to decisions.

Acknowledgements

This work was funded by the DARPA Safe Genes Program (Grant DARPA-16-59-SAFE-FP-005), from which several of the authors (DVB, KJC, SD, JDE, OE, JG, GRH, KO, AJP, TAAP, JRS, ABS, PT and ALL) receive support. Support was also received from the Research Training Group in Mathematical Biology, funded by NSF grant RTG/DMS-1246991 (JS, BH, 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). Contributions from other members of the Genetic Biocontrol of Invasive Rodents (GBIRd) consortium (http://www.geneticbiocontrol.org/) are acknowledged and greatly appreciated.

Author Contributions

JS, BH, FG and ALL designed the model. JS, BH and ALL carried out model simulations and analysis. JS, BH, FG and ALL wrote the first draft of the paper. All authors discussed model results and contributed to editing and revision of the manuscript.

References

Akbari, O. S., Bellen, H. J., Bier, E., Bullock, S. L., Burt, A., Church, G. M., Cook, K. R., Duchek, P., Edwards, O. R., Esvelt, K. M., Gantz, V. M., Golic, K. G., Gratz, S. J., Harrison, M. M., Hayes, K. R., James, A. A., Kaufman, T. C., Knoblich, J., Malik, H. S., Matthews, K. A., O'Connor-Giles, K. M., Parks, A. L., Perrimon, N., Port, F., Russell, S., Ueda, R., & Wildonger, J. (2015). Safeguarding gene drive experiments in the laboratory. Science, 349, 927–929.

Alphey, L., Beard, C. B., Billingsley, P., Coetzee, M., Crisanti, A., Curtis, C., Eggleston, P., Godfray, C., Hemingway, J., Jacobs-Lorena, M., James, A. A., Kafatos, F. C., Mukwaya, L. G., Paton, M., Powell, J. R., Schneider, W., Scott, T. W., Sina, B., Sinden, R., Sinkins, S., Spielman, A., Toure, Y., & Collins, F. H. (2002). Malaria control with genetically manipulated insect vectors. Science 298, 119–121.

Alphey, N., & Bonsall, M. B. (2014). Interplay of population genetics and dynamics in the genetic control of mosquitoes. J. R. Soc. Interface, 11(93), 20131071.

Altrock, P. M., Traulsen, A., Reeves, R. G., & Reed, F. A. (2010). Using underdominance to bi-stably transform local populations. J. Theor. Biol. 267, 62-75.

Backus, G. A., & Gross, K. (2016). Genetic engineering to eradicate invasive mice on islands: modeling the efficiency and ecological impacts. Ecosphere *7*(12), e01589.

Barton, N. H. (1979). The dynamics of hybrid zones. Heredity 43, 341-359.

Barton, N. H., & Turelli, M. (2011). Spatial waves of advance with bistable dynamics: Cytoplasmic and genetic analogues of Allee effects. Am. Nat. 178(3), E48-E75.

Beeman, R. W., Friesen, K. S., & Denell, R. E. (1992). Maternal-effect selfish genes in flour beetles. Science *256*(5053), 89-92.

Browne, R A.1977. Genetic variation in island and mainland populations of *Peromyscus leucops*. Am Midl. Nat, 97, 1–9.

Burt, A. (2003). Site-specific selfish genes as tools for the control and genetic engineering of natural populations. Proc. R. Soc. Lond. B 270(1518), 921-928.

Campbell, K. J., Beek, J., Eason, C. T., Glen, A. S., Godwin, J., Gould, F., Holmes, N. D., Howald, G. R., Francine M. Madden, F. M., Ponder, J. B., Threadgill, D. W., Wegmann, A. S., & Baxter, G. S. (2015) The next generation of rodent eradications: Innovative technologies and tools to improve species specificity and increase their feasibility on islands. Biol. Cons. 185, 47-58.

Champer, J., Buchman, A., & Akbari, O. S. (2016). Cheating evolution: engineering gene drives to manipulate the fate of wild populations. Nat. Rev. Genet. 17, 146-159.

Curtis, C. F. (1968). Possible use of translocations to fix desirable genes in insect pest populations. Nature 218(5139), 368–369.

Cuthbert, R., & Hilton, G. (2004). Introduced house mice *Mus musculus*: a significant predator of threatened and endemic birds on Gough Island, South Atlantic Ocean? Biol. Cons. 117, 483-489.

Davis, S., Bax, N., & Grewe, P. (2001). Engineered underdominance allows efficient and economical introgression of traits into pest populations. J. Theor. Biol. 212(1), 83-98.

Deredec, A., Burt, A., & Godfray, C. (2008). The population genetics of using homing endonuclease genes (HEGs) in vector and pest management. Genetics 179, 2013-2026.

Dhole, S., Vella, M. R., Lloyd, A. L., & Gould, F. (2018). Invasion and migration of spatially self-limiting gene drives: A comparative analysis. Evol. Applic. 11(5), 794-808.

Esvelt, K., Smidler, A. L., Catteruccia, F., & Church, G. M. (2014). Concerning RNA-guided gene drives for the alteration of wild populations. eLIFE 3, e03401.

Esvelt, K. M., & Gemmell, N. J. (2017). Conservation demands safe gene drive. PLoS Biology 15(11), e2003850.

Frankham, R. (1997). Do island populations have lower genetic variation than mainland populations? Heredity 78, 311-327.

Gould, F. (2008). Broadening the application of evolutionarily based genetic pest management. Evolution *62*(2), 500-510.

Gould, F., Huang, Y., Legros, M., & Lloyd, A. L. (2008). A Killer–Rescue system for self-limiting gene drive of anti-pathogen constructs. Proc. R. Soc. Lond. B 275(1653), 2823-2829.

Harradine, E., How, R. A., Schmitt, L. H., & Spencer, P. B. S. (2015). Island size and remoteness have major conservation significance for how spatial diversity is partitioned in skinks. Biodiv. Cons. 24, 2011–2029.

Howald, G., Donlan, C. J., Galvan, J. P., Russell, J. C., Parkes, J., Samaniego, A., Wang, Y., Veitch, D., Genovesi, P., Pascal, M., Saunders, A., & Tershy, B. (2007). Invasive rodent eradication on islands. Conserv. Biol. 21(5), 1258-1268.

Khamis, D., El Mouden, C., Kura, K., & Bonsall, M. B. (2018). Ecological effects on underdominance threshold drives for vector control. J. Theor. Biol. 456, 1-15.

Kimura, M., & Ohta, T. (1968). The average number of generations until fixation of a mutant gene in a finite population. Genetics 61, 763–771.

Macer, D. (2005). Ethical, legal and social issues of genetically modifying insect vectors for public health. Insect Biotech. Mol. Biol. 35, 649–660.

Magori, K., & Gould, F. (2006). Genetically engineered underdominance for manipulation of pest populations: A deterministic model. Genetics 172, 2613-2620.

Marshall, J. M. (2010). The Cartagena Protocol and genetically modified mosquitoes. Nature Biotech. 28, 896–897.

Marshall, J. M. & Akbari, O. S. (**2018**) Can CRISPR-Based Gene Drive Be Confined in the Wild? A Question for Molecular and Population Biology. ACS Chem. Biol. 13(2), 424– 430.

Nathan, H. W., Clout, M. N., MacKay, J. W., Murphy, E. C., & Russell, J. C. (2015). Experimental island invasion of house mice. Popul. Ecol. 57(2), 363-371.

Noble, C., Min, J., Olejarz, J., Buchthal, J., Chavez, A., Smidler, A., DeBenedictis, E., Church, G., Nowak, M., & Esvelt, K. (2016). Daisy-chain gene drives for the alteration of local populations. bioRxiv 057307. doi:10.1101/057307.

Piaggio, A. J., Segelbacher, G., Seddon, P. J., Alphey, L., Bennett, E. L., Carlson, R. H., Friedman, R. M., Kanavy, D., Phelan, R., Redford, K. H., Rosales, M., Slobodian, L., & Wheeler, K. (2017) Is it time for synthetic biodiversity conservation? Trends Ecol. Evol. 32, 97-107.

Prowse, T. A. A., Cassey, P., Ross, J. V., Pfitzner, C., Wittmann, T. A., & Thomas, P. (2017) Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. Proc. R. Soc. Lond. B 284, 20170799.

Rasgon, J. L. (2009). Multi-locus assortment (MLA) for transgene dispersal and elimination in mosquito populations. PLoS ONE, 4, e5833.

Robert, M. A., Okamoto, K., Lloyd, A. L., & Gould, F. (2013). A reduce and replace strategy for suppressing vector-borne diseases: insights from a deterministic model. PLoS One 8(9), e73233.

Spielman, A. (1994). Why entomological antimalaria research should not focus on transgenic mosquitoes. Parasitol. Today 10,374–376.

Towns, D. R., Atkinson, I. A. E., & Daughtery, C. H. (2006). Have the harmful effects of introduced rats on islands been exaggerated? Biol. Invasions 8, 863-891.

Unckless, R. L., Messer, P. W., Connallon, T., & Clark, A. G. (2015). Modeling the manipulation of natural populations by the mutagenic chain reaction. Genetics 201(4), 425-531.

White, T. A., & Searle, J. B. (2007). Genetic diversity and population size: island populations of the common shrew, *Sorex araneus*. Mol. Ecol. 16, 2005-2016.

Table 1: List of Parameters and Their Values

	Parameter	Value	Source
λ	Female fecundity rate	8.4 $(\text{year})^{-1}$	Backus & Gross (2016)
q	Coefficient quantifying density-dependent decline in birth rate	Island: $6.427\text{x}10^{-3}$ $(\text{mouse})^{-1}$ Mainland: $6.427\text{x}10^{-3}$ $(\text{mouse})^{-1}$	Backus & Gross (2016)
ρ	Density independent per-capita death rate	2.4 $(\text{year})^{-1}$	Backus & Gross (2016)
α	Coefficient quantifying density-dependent increase in death rate	Island: $5.000\text{x}10^{-5}(\text{mouse})^{-1}(\text{year})^{-1}$ Mainland: $5.000\text{x}10^{-7}(\text{mouse})^{-1}(\text{year})^{-1}$	Backus & Gross (2016)
w_i	Genotype specific fitness	Dependent on *s* and *h*	
μ	Per-capita migration rate from island to mainland	Varied (10^{-5}-10^{-1} $(\text{year})^{-1}$)	
s	Drive allele fitness cost	Non-threshold drive: 0.3 Threshold drive: 0.8	
h	Degree of dominance of fitness cost of drive allele	0.8	
e	Homing probability	0.95	

Figure Captions

Figure 1: Island population dynamics for the no invasion threshold and invasion threshold scenarios. In both scenarios the drive has an 80% fitness cost (s = 0.8) but have differing degrees of dominance: for the no invasion threshold scenario (green curve) we take h = 0.3, while for the invasion threshold scenario we take h = 0.8. Population sizes are plotted relative to pre-release equilibrium population numbers. In the no invasion threshold scenario, arbitrarily small releases of drive individuals lead to invasion and fixation of the drive allele, leading to suppression of the population (green curve). In the invasion threshold scenario, invasion of the drive depends on whether the initial release exceeds the invasion threshold (for this choice of parameters, the invasion threshold frequency for drive is 0.621, corresponding to approximately 1.64 drive individuals for each wild-type individual). Red curve depicts a sub-threshold release (1.63 homozygous drive females per wild-type female), leading to loss of the drive allele and only temporary suppression of the population before its return to carrying capacity. Blue curve depicts a successful release (1.65 homozygous drive females per wild-type female), for which the drive invades and reaches fixation, leading to elimination of the population. Values of other parameters are given in Table 1.

Figure 2: Island and mainland population dynamics (blue curves) and genetics (red curves) in the invasion threshold scenario (s = 0.8 and h = 0.8), following an above-threshold release of drive individuals on the island (1650 drive homozygotes released at t = 0). Left axis (blue) denotes population size relative to pre-release equilibrium, right axis (red) denotes allele frequency. The drive spreads on the island (dashed red curve shows allele frequency for drive), suppressing its population (solid blue curve shows relative population size). Migration from the mainland to the island (on average, one island individual travels to the mainland a month) leads to a low drive frequency on the mainland that does not exceed the invasion threshold there. Consequently, the mainland population is largely unaffected. (Dot-dashed line denotes frequency of resistance allele.)

Figure 3: Island and mainland population dynamics and genetics in the no invasion threshold scenario (s = 0.8 and h = 0.3). As in Figure 2, blue curves and axes denote population sizes, measured relative to pre-release equilibria, while red curves and axes denote allele frequencies. A small release (five homozygous drive individuals) occurs at time t = 0. We assume that resistance is very low on the mainland (allele frequency of just 5%), representing a rather pessimistic scenario in terms of the susceptibility of the mainland population to the drive. The drive spreads to fixation and suppresses the island population. Migration to the mainland (on average, one island individual travels to the mainland a month) means that the drive is introduced to the mainland, where it can spread through the susceptible population, but not the resistant population. The total population undergoes a temporary suppression as the drive spreads through the susceptible population. The frequency of resistant alleles increases as a result of drive, and density dependent population regulation returns the mainland population to the pre-release equilibrium level.

Figure 4: Maximum level of transient suppression seen on the mainland following an island release of 100 homozygous drive individuals in the no invasion threshold scenario (s = 0.8 and h = 0.3) across combinations of different migration rates and initial frequencies of resistant alleles on the mainland. Different levels of suppression are denoted by different colors (see color key below figure). Low initial resistance frequencies depict pessimistic scenarios (mainland is almost entirely susceptible to the drive), while high initial resistance frequencies approach the “private allele” scenario discussed in the text.

Figure 5: Maximum drive frequency observed on the mainland following an island release. Maximum drive frequency is indicated by color. All other details as in Figure 4.

Figure 1

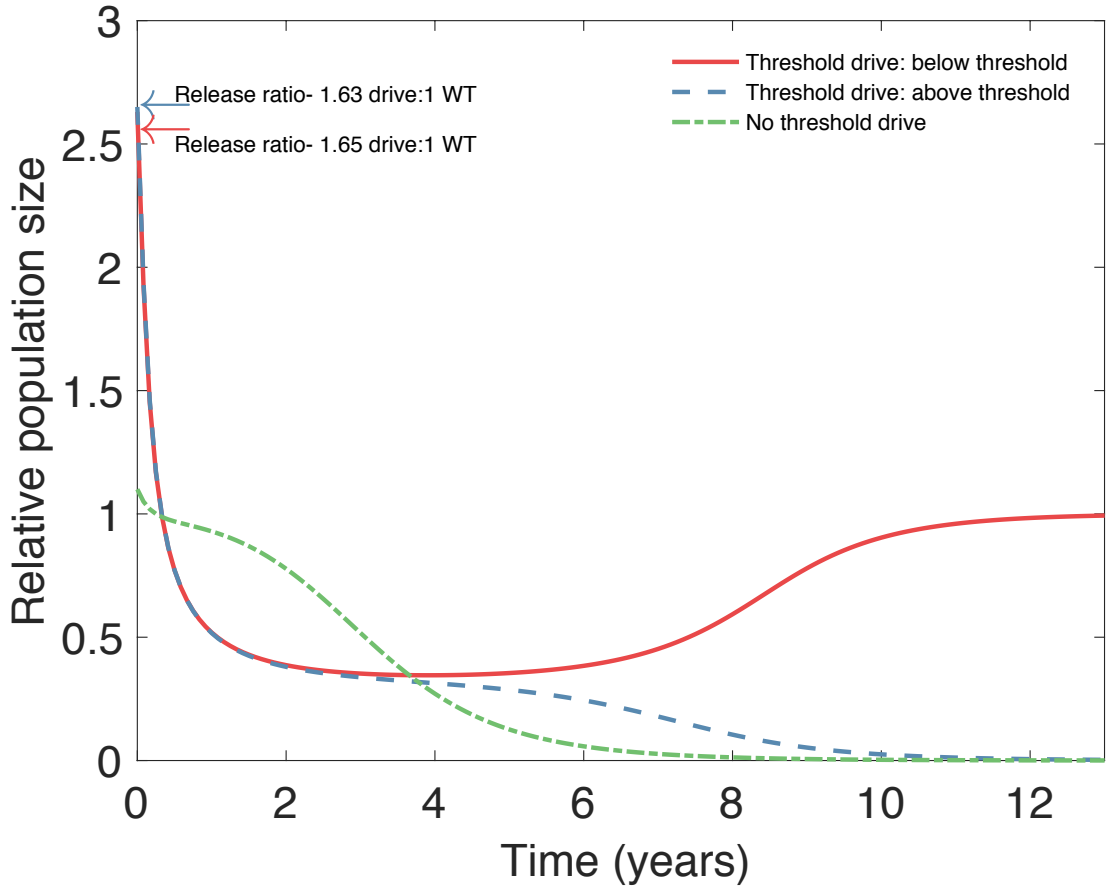

Threshold drive: below threshold
Threshold drive: above threshold
No threshold drive
Release ratio- 1.63 drive:1 WT
Release ratio- 1.65 drive:1 WT
Relative population size
Time (years)
0
0.5
1
1.5
2
2.5
3
0
2
4
6
8
10
12

Figure 2

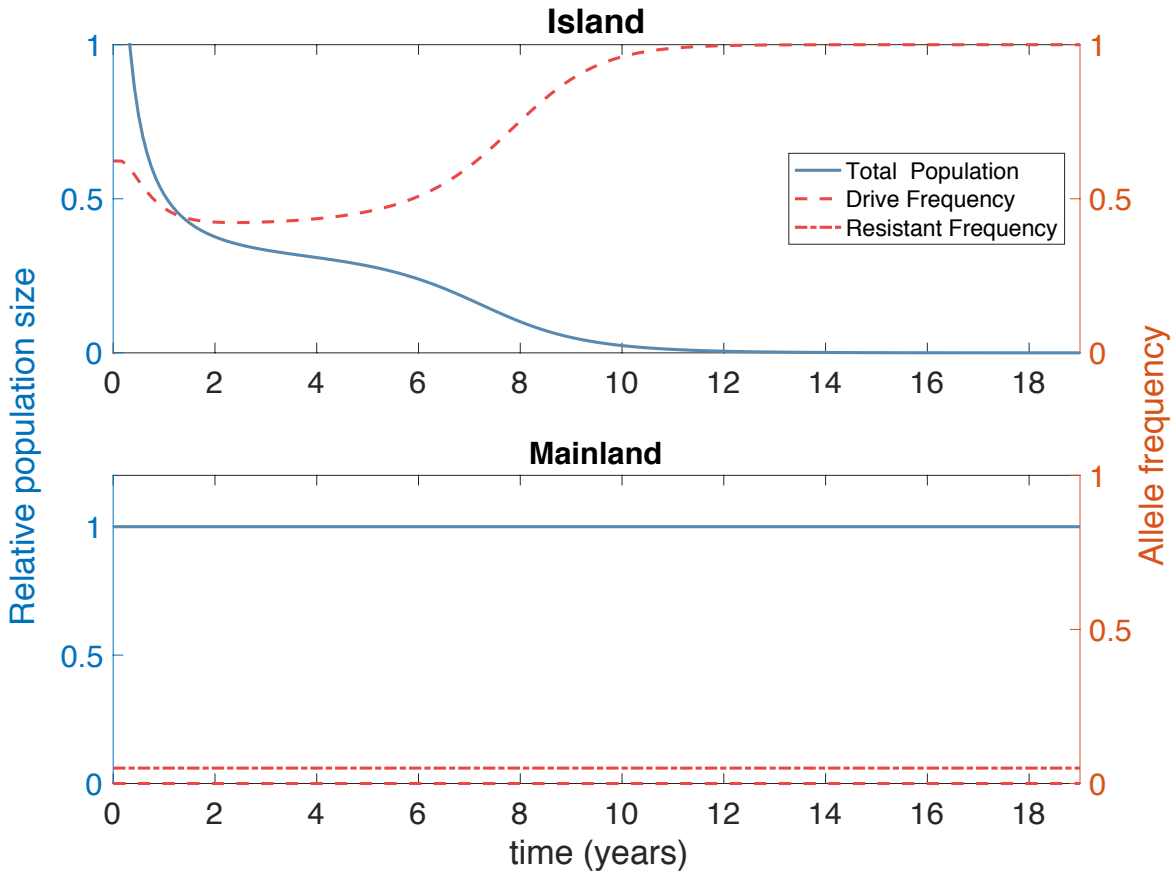

Island
Mainland
Total Population
Drive Frequency
Resistant Frequency
Relative population size
Allele frequency
time (years)
0
0.5
1
0
2
4
6
8
10
12
14
16
18

Figure 3

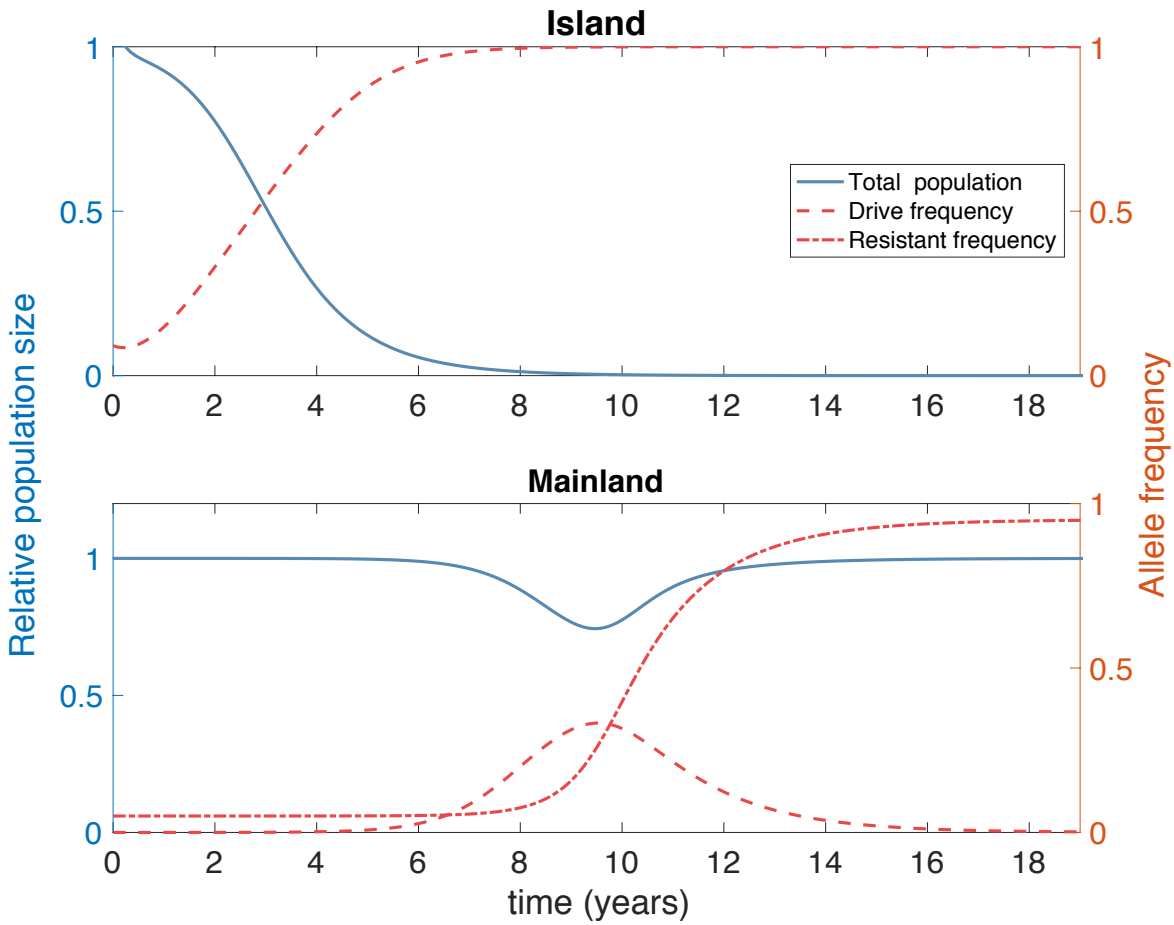

Island
Mainland
Total population
Drive frequency
Resistant frequency
Relative population size
Allele frequency
time (years)
0
0.5
1
0
2
4
6
8
10
12
14
16
18

Figure 4

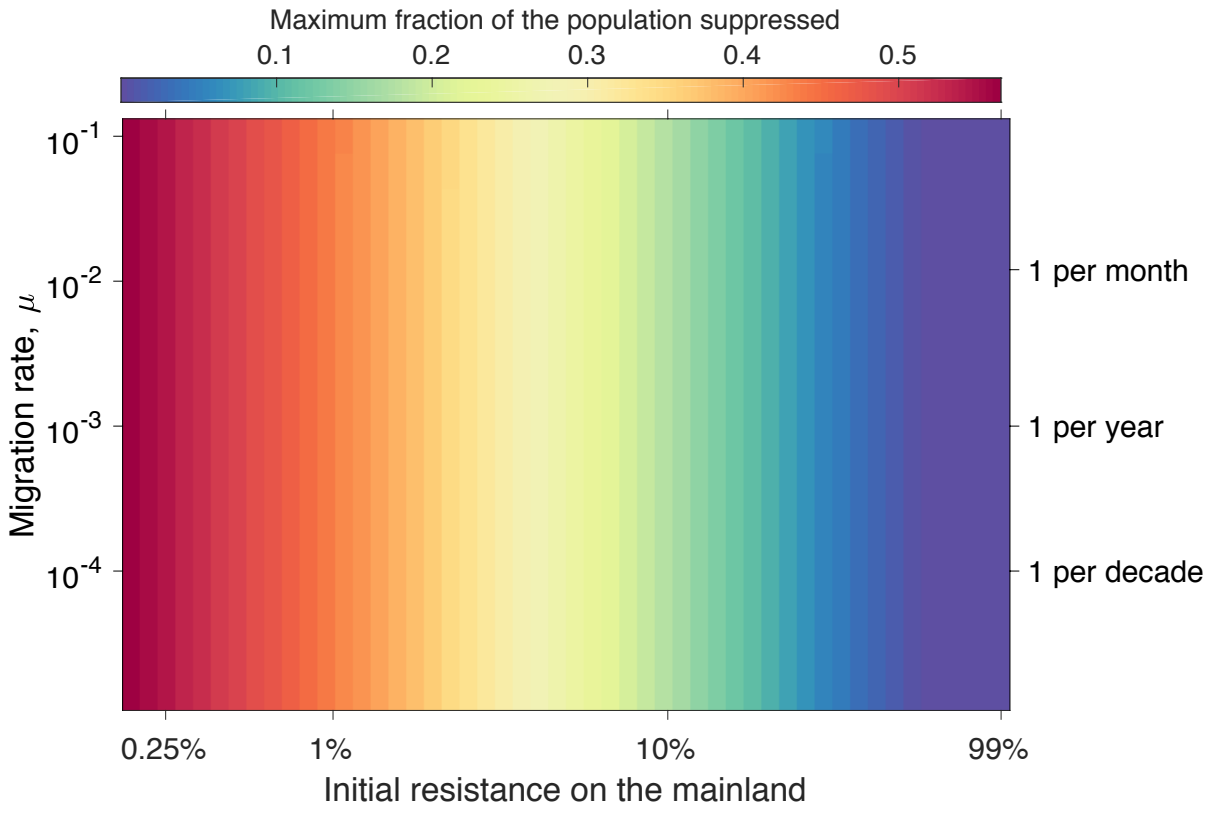
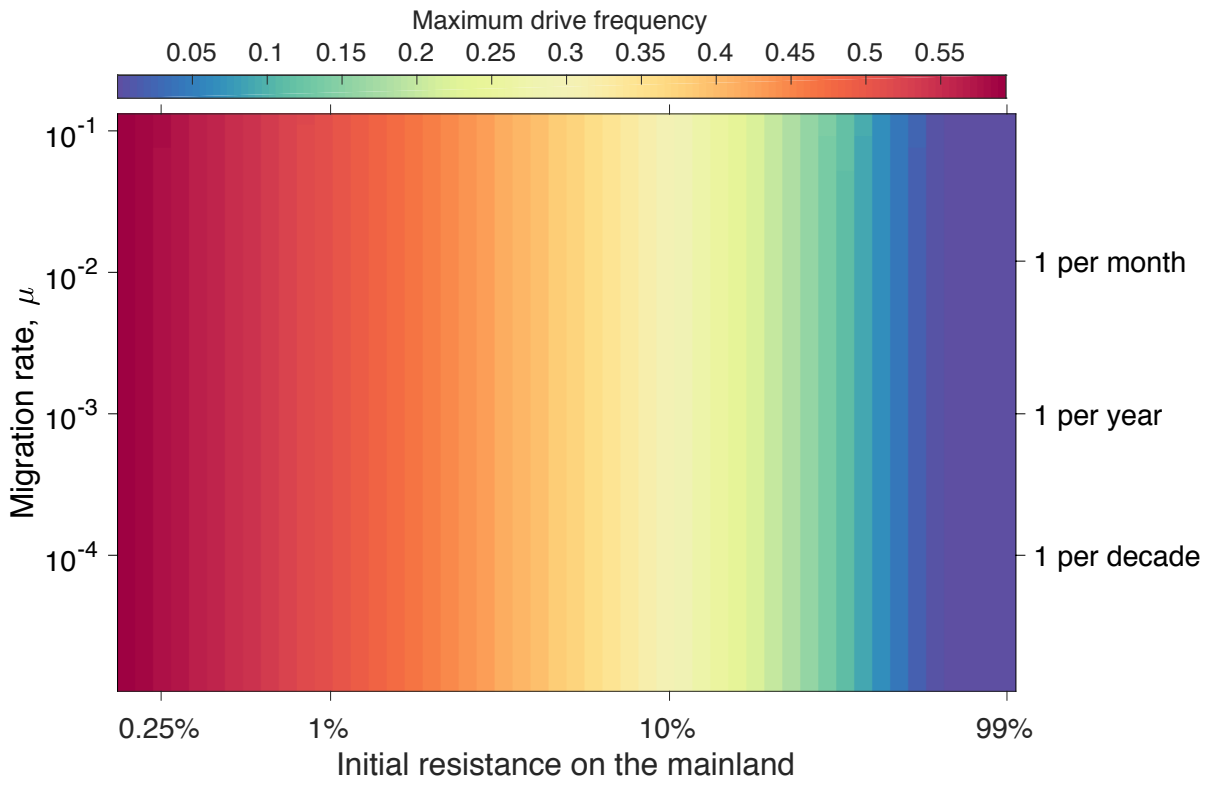

Maximum fraction of the population suppressed
0.1
0.2
0.3
0.4
0.5
Migration rate, μ
10^-1
10^-2
10^-3
10^-4
1 per month
1 per year
1 per decade
0.25%
1%
10%
99%
Initial resistance on the mainland

Figure 5.


Maximum drive frequency
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
Migration rate, μ
10^-1
10^-2
10^-3
10^-4
1 per month
1 per year
1 per decade
0.25%
1%
10%
99%
Initial resistance on the mainland

Supplemental Text and Figures

Supplemental Figures 1 and 2 explore the timing of (1) suppression on the mainland and (2) the peak level of drive seen on the mainland. These two events occur at similar times (but not at exactly the same time). We note that over a wide range of initial levels of resistance and migration rates, there is only weak dependence of either time on these quantities. For high levels of resistance, however, we see that these times are much longer. This is a consequence of there being a threshold level of resistance above which drive cannot invade. This threshold can be derived by examining the stability of a drive-free equilibrium state of the single patch model, using the standard approach of linearizing the model about the equilibrium (see, for example, Strogatz 2014). This process involves calculation of the eigenvalues of the Jacobian matrix (i.e. the matrix of the partial derivatives of the right-hand sides of the differential equations (eqns 3 in the main text) with respect to the state variables of the model) evaluated at the equilibrium of interest. The 6x6 Jacobian matrix and its eigenvalues are easily calculated with the assistance of a computer algebra package (e.g. Maple [Maplesoft, a division of Waterloo Maple, Inc., Waterloo, Ontario] or Mathematica [Wolfram Research, Inc., Champaign, IL]). A Maple worksheet providing details of the calculations is provided as a supplementary file.

Given that we assume equal fitnesses for susceptible and resistant types, in the absence of drive there is no selection between susceptible and resistant types and so we have non-unique drive-free equilibria:

$$N_{\mathrm{SS}} = (1-q_R)^2 N^*,\ N_{\mathrm{SR}} = 2q_R(1-q_R)N^*,\ N_{\mathrm{RR}} = q_R^2 N^*,\ N_{\mathrm{SD}} = N_{\mathrm{RD}} = N_{\mathrm{DD}} = 0$$

Here q_R is the frequency of the resistant allele and N^* is the drive-free equilibrium population size (equal to $(\lambda-\rho)/(\lambda q+\alpha)$) . We note that these equilibria form a line in phase space.

The linear stability analysis gives six eigenvalues, one of which is zero (corresponding to neutral stability along the line of drive-free equilibria), one equal to $\rho-\lambda$ and three equal to $-\lambda(q\rho+\alpha)/(\lambda q+\alpha)$. We see that these last four eigenvalues are negative (note that λ must be greater than ρ in order for N^* to be positive). The ability of drive to invade (from low initial frequency) is determined by the sign of the final eigenvalue, whose value is equal to $\lambda[(1-q_R)(1-hs)e-hs](q\rho+\alpha)/(\lambda q+\alpha)$. Invasion is only possible if this quantity is positive, and so consideration of the quantity in square brackets leads to the condition

$$q_R < 1 - hs/\{e(1-hs)\}.$$

Another way to derive this threshold (and one that immediately applies to a number of previous studies in the literature) is by considering a discrete-time (discrete generations) description framed in terms of allele frequencies (see, for example Deredec et al. 2008 or Vella et al. 2017).

If the allele frequencies for Susceptible, Drive and Resistant alleles in the current generation are written as q_S, q_D and q_R, it can be shown that the frequency of the drive allele in the next generation, q_D', will be given by

$$q_D' = \frac{1}{\overline{w}}\left(\{1-s\}q_D^2 + \{1-hs\}q_S\ q_D\,\{1+e\} + \{1-hs\}q_R\ q_D\right),$$

where $\overline{w}$ is the mean fitness, which equals

$\bar{w} = 1 - sq_D^2 - 2hsq_D(q_S + q_R)$.

When thinking about invasion of drive into a population that initially consists of susceptible and resistant alleles, q_D will be small, and so the following linear equation can be derived for the change in the frequency of the drive allele from one generation to the next:

$$q_D' - q_D \approx q_D\left(\{1 - hs\}q_S\{1 + e\} + \{1 - hs\}q_R - 1\right).$$

(This equation is correct to first order in q_D.)

We see that the drive frequency can only increase if the term in parentheses is positive, meaning that

$q_R < \frac{e(1-hs)-hs}{e(1-hs)}$,

which may be written as

$$q_R < 1 - hs/\{e(1 - hs)\}.$$

Supplemental Figure 3 explores the composition of the mainland population that remains after the transient spread of drive, showing the long-term frequency of the susceptible allele. We note that, while the transient spread of drive leads to the reduction of susceptible alleles on the mainland, it does not lead to their elimination.

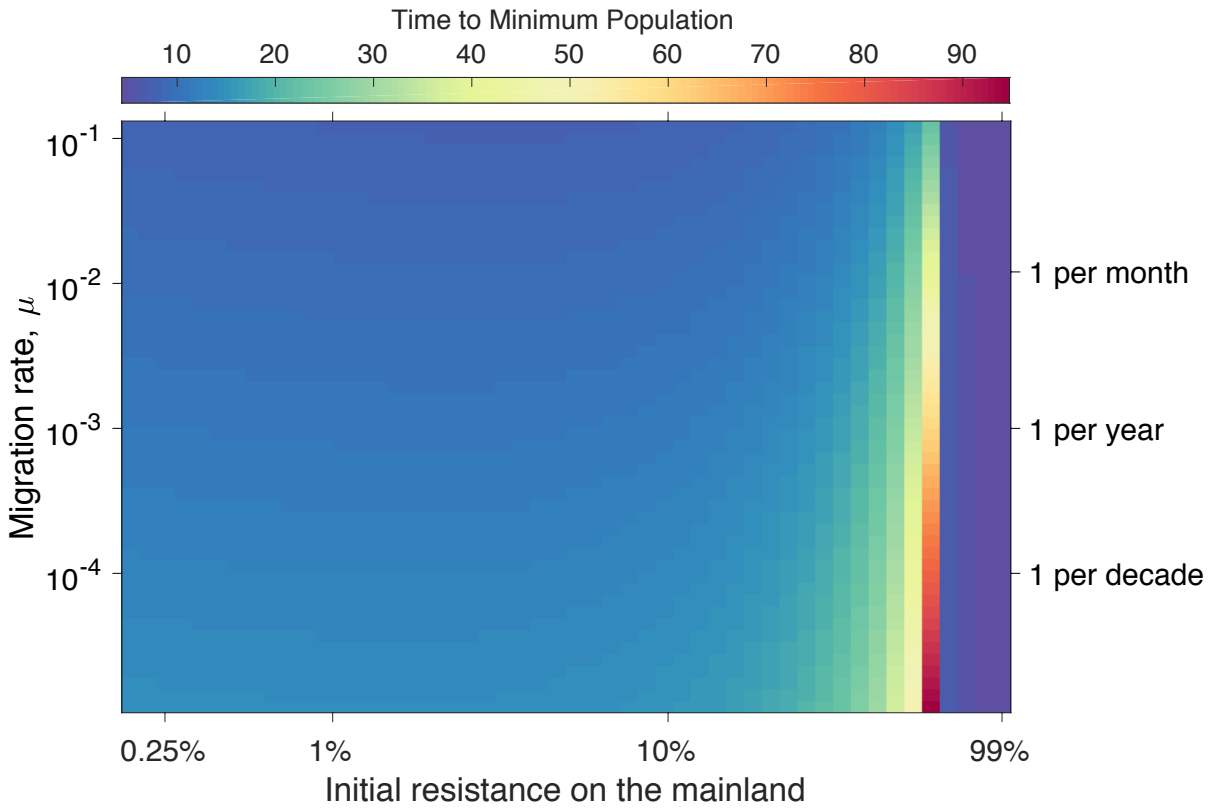
Supplemental Figures 4 and 5 show that the peak level of suppression on the mainland is almost independent of the size of the release on the island, while the time until this suppression occurs is only weakly dependent on the release size.

References:

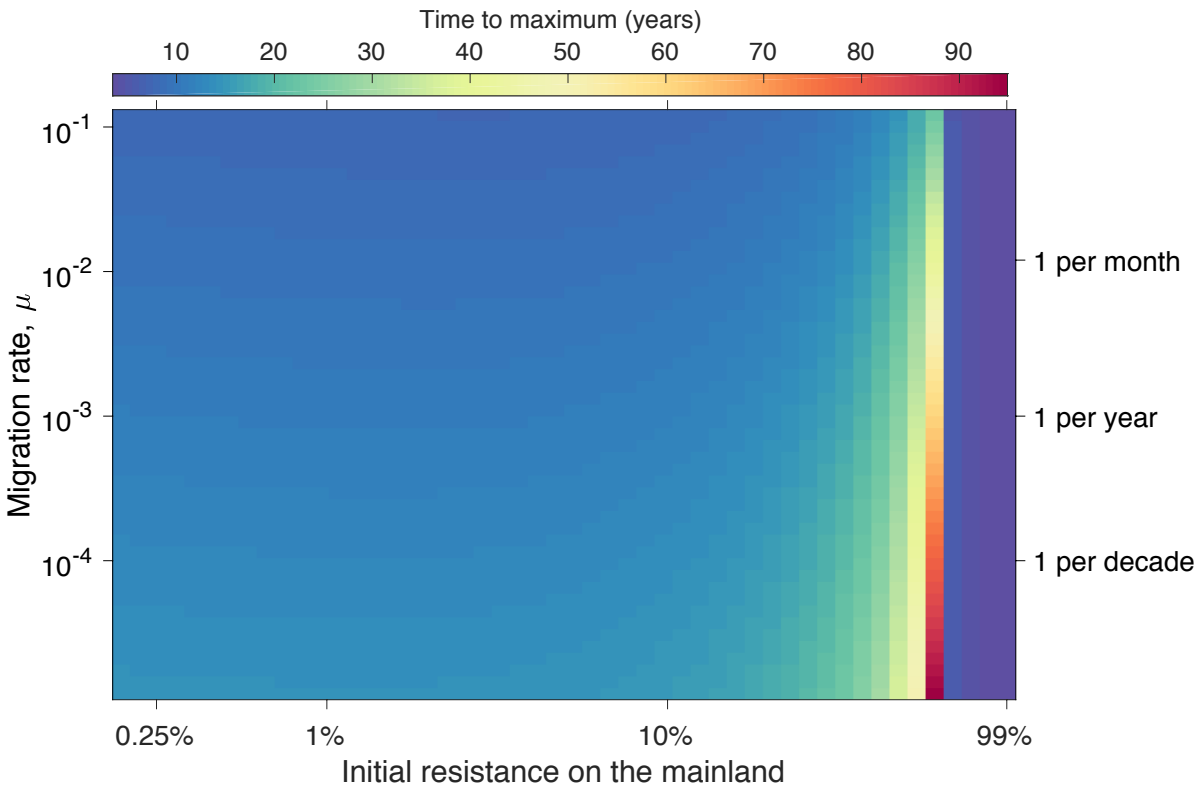
Strogatz, S. H. (2014) Nonlinear Dynamics and Chaos, Second Edition. Westview Press.

Vella, M. R., Gunning, C. E., Lloyd, A. L., & Gould, F. (2017) Evaluating strategies for reversing CRISPR-Cas9 gene drives. Scientific Reports 7, 11038.

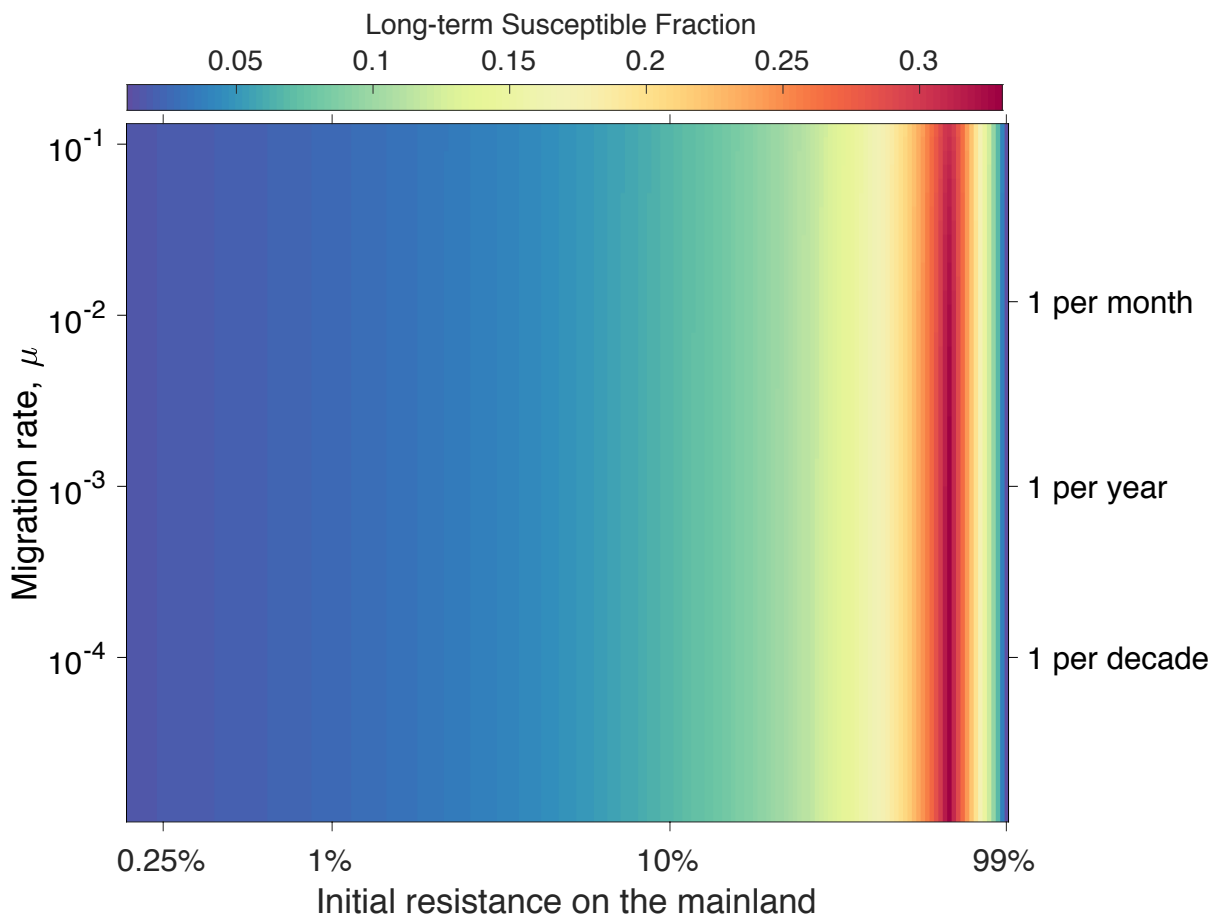
Supplemental Figure 1. Dependence of the time until the mainland population achieves its minimum on the initial level of resistance on the mainland and the migration rate from the island. All other details are as in Figure 4 of the main text.



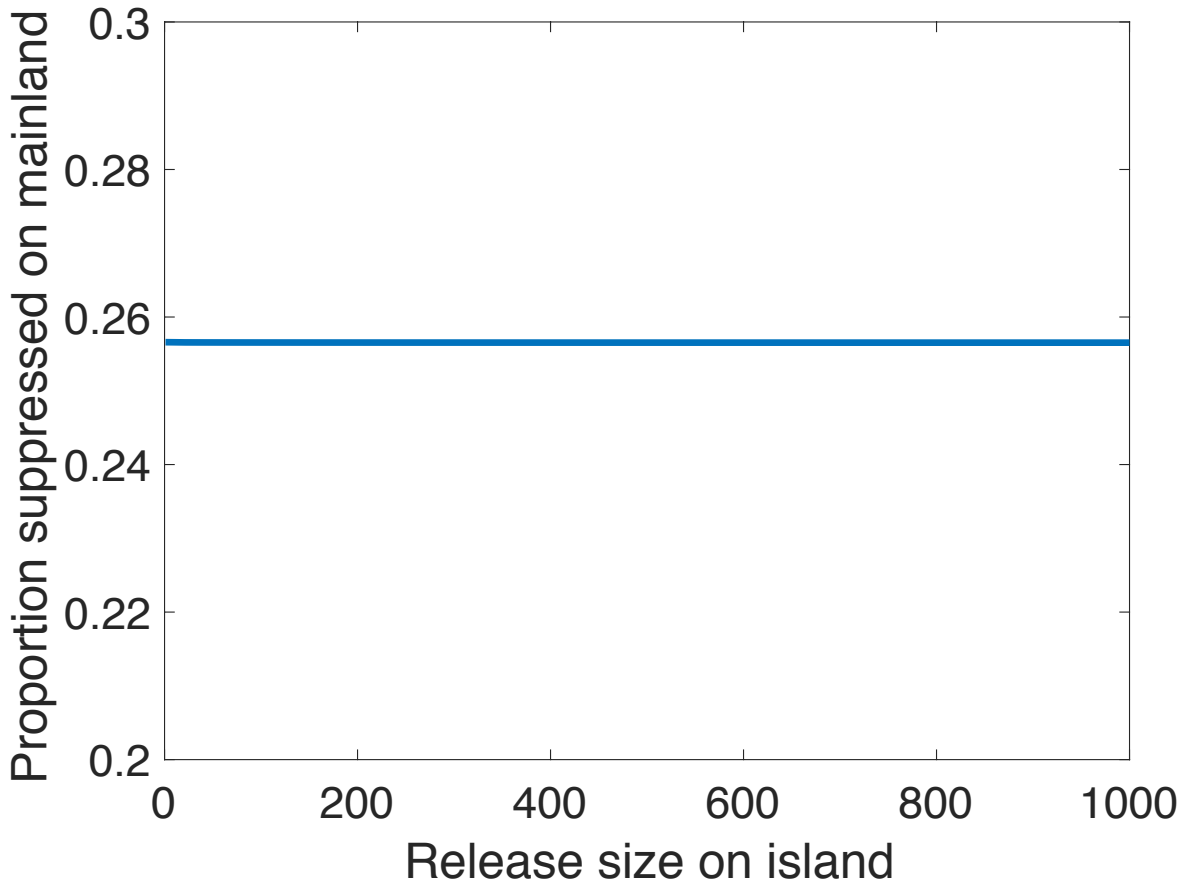
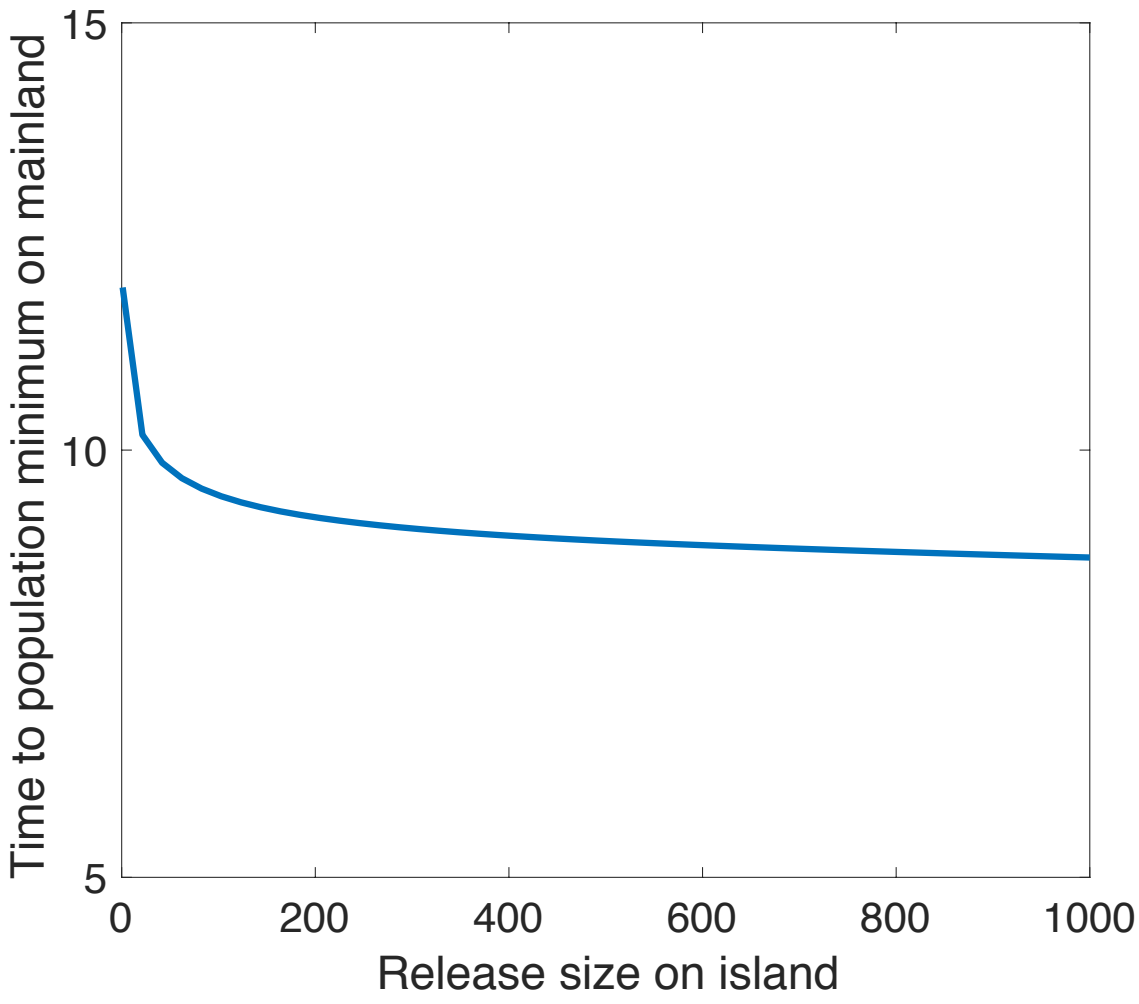
Supplemental Figure 2. Dependence of the time until the drive frequency achieves its maximum on the mainland on the initial level of resistance on the mainland and the migration rate from the island. All other details are as in Figure 4 of the main text.



Supplemental Figure 3. Frequency of the susceptible allele on the mainland years (> 500) after an island release in the no invasion threshold scenario (*s* = 0.8 and *h* = 0.3) across combinations of different migration rates and initial frequencies of resistant alleles on the mainland. All other details are as in Figure 4 of the main text. Note that susceptible alleles remain in the population after the transient spread and loss of drive on the mainland.



Supplemental Figure 4. Maximum level of transient suppression seen on mainland following releases of various sizes on the island under the no invasion threshold scenario. Initial frequency of resistance on the mainland is equal to 0.05 and migration occurs at a per-capita rate of 0.012 per year.



Supplemental Figure 5. Time until the maximum level of transient suppression occurs on the mainland following releases of various sizes on the island in the no invasion threshold scenario. Initial frequency of resistance on the mainland is equal to 0.05 and migration occurs at a per-capita rate of 0.012 per year.