SCIENTIFIC REPORTS
natureresearch

OPEN

Mathematical modeling of self-contained CRISPR gene drive reversal systems

Matthew G. Heffel[1] & Gregory C. Finnigan[2*]

There is a critical need for further research into methods to control biological populations. Numerous challenges to agriculture, ecological systems, and human health could be mitigated by the targeted reduction and management of key species (e.g. pests, parasites, and vectors for pathogens). The discovery and adaptation of the CRISPR/Cas editing platform co-opted from bacteria has provided a mechanism for a means to alter an entire population. A CRISPR-based gene drive system can allow for the forced propagation of a genetic element that bypasses Mendelian inheritance which can be used to bias sex determination, install exogenous information, or remove endogenous DNA within an entire species. Laboratory studies have demonstrated the potency by which gene drives can operate within insects and other organisms. However, continued research and eventual application face serious opposition regarding issues of policy, biosafety, effectiveness, and reversal. Previous mathematical work has suggested the use of modified gene drive designs that are limited in spread such as daisy chain or underdominance drives. However, no system has yet been proposed that allows for an inducible reversal mechanism without requiring the introduction of additional individuals. Here, we study gene drive effectiveness, fitness, and inducible drive systems that could respond to external stimuli expanding from a previous frequency-based population model. We find that programmed modification during gene drive propagation could serve as a potent safeguard to either slow or completely reverse drive systems and allow for a return to the original wild-type population.

The discovery of the CRISPR/Cas system as a powerful genetic editing biotechnology has revolutionized many fields across agriculture and biomedical research. To date, control of populations remains essential to managing of our food supply as well as preserving natural habitats and their diverse ecological composition. For instance, invasive species cause severe damage on a wide scale in many ecosystems[1–4] and insects serve as vectors for transmitting an increasing number of diseases including Zika, dengue, malaria, Lyme disease, and typhus[5–7]. Current mechanisms to mitigate biological populations rely on a variety of strategies ranging from mosquito netting[8] to natural predators[9]. However, no previous methodology has existed that could effectively alter an entire population on a global scale.

The demonstration that a nuclease-based "gene drive" (GD) could artificially propagate genetic information through a population could have profound impacts for human health and the environment. This mechanism utilizes a strategy within a single diploid cell to transform the heterozygous condition (between the pairing of a genetically modified organism (GMO) and a wild-type (WT) individual) to the homozygous condition resulting in *Super*-Mendelian inheritance of the desired genetic cargo and a rapid sweep through a population[10] (Fig. S1).

The potential benefits and applications of such a system are numerous. The eradication of insect populations that serve as disease vectors could have profound impacts globally[11]. Successful gene drive systems have now been demonstrated within fungi, mammals, and insects[12–15]. The intended trait to drive through the species of choice is often to interfere with sex determination to bias the male-female ratio to an extreme to cause populations to crash[16–18]. However, challenges remain in the design and effective propagation of the drive within sample populations and ethical and ecological concerns regarding actual use of this system within wild populations remain.

Methodologies with traditional drive systems allow only two possible outcomes: (i) the GD runs to completion and takes over the population or (ii) the GD is removed from the population (via evolved resistance, sub-optimal fitness, or is destroyed, etc.). Therefore, we envisioned a GD system that could be tuned or modified while it was

[1]Division of Biology, 116 Ackert Hall, Kansas State University, Manhattan, KS 66506, USA. [2]Department of Biochemistry and Molecular Biophysics, 141 Chalmers Hall, Kansas State University, Manhattan, KS 66506, USA. *email: gfinnigan@ksu.edu

actively spreading and already present within a population. The rationale for this type of programmable system includes issues of biosafety, tunability, and customization. In terms of design, any added components would be installed within the GD itself, proximal to the nuclease and guide RNA(s) or at additional loci. Previous studies have highlighted numerous ways that the activity of nucleases could be controlled and have a programmed "failure" rate[13,19] such that the expected propagation through a population was slowed. These include inhibition by added domain fusions[20], direct inhibition by anti-CRISPRs[21–24], restriction on trafficking to and from the nucleus[13], split nuclease systems[25], and regulation of protein levels[13,26,27].

Importantly, others have demonstrated that external cellular cues can be coupled to nuclease localization and/or activity[28–30]. For example, fusions to plasma-membrane localized G-protein coupled receptors (GPCRs) spatially restrict dCas9 until external ligand binding. When activated, a conformational change in the GPCR activates protease-dependent cleavage of the fused dCas protein, allowing it to be shuttled into the nucleus to modulate a downstream response[29]. At the organismal scale, external cues could include a range of stimuli including small molecules (pheromone), environmental changes (temperature, diet), or artificial means (small molecules). A recent study demonstrated that a domain-destabilized Cas9-based gene drive in flies could allow for a titration of drive activity; addition of an external cue stabilized the nuclease in a dose-dependent manner[27]. We envisioned a GD system where sensory information has been converted into an alteration to nuclease function (e_W) and/or overall biological fitness (f).

In this study, we have expanded on a previously developed frequency-based population model[31] to explore mechanisms to reverse traditional gene drives. We focus on use of inducible parameters for GD effectiveness, fitness, and a self-cleaving drive. We demonstrate that these designs could allow for a means of control that does not require the introduction of additional individuals. These findings could aid in development of an effective, safe, and fully reversible gene drive that could restore a native wild-type population.

Results

An inducible gene drive system (drive efficiency). Previous gene drive models[31,32] have examined the effects of drive efficiency (e_W) and fitness (f). Our simulations of this same model[31] illustrated the required initial parameters for both GD efficiency and fitness to allow for successful propagation (Fig. S2). We observed that a range of e_W values still resulted in GD takeover; however, there was a limit to the fitness cost that could be present within GD individuals despite values of e_W that approached one suggesting a much stronger effect of fitness on drive success. These data were also consistent with a previous individualistic GD model[32] used to evaluate the contributions of e_W and f. In that study, the equation $f(e_W + 1) > 1$ represented scenarios where the GD was favored and would take over the population[32].

Our model for an inducible gene drive expanded upon previous work[31]. We defined e_W' as the GD efficiency after introduction of the inducing agent and defined α as the success rate of the applied stimulus—the fraction of individuals within the population that successfully responded over a single generation (Fig. 1A). The first application of the signal occurred at generation 10 and continued for all subsequent generations. To illustrate one scenario for how this inducible GD could be applied, we modeled a system where the initial conditions (e_W and f) would allow the GD to move to fixation, but when exposed to the cue, the WT would be favored and the GD would be removed from the population (Fig. 1). Not all initial parameters allowed for such a shift—for example, conditions where $e_W = 1$ and $f = 0.95$ that were shifted to $e_W' = 0.2$, the GD was still favored to take over the population, albeit at a slower rate. However, for starting conditions of $e_W = 0.8$, $f = 0.7$, and $e_W' = 0.1$, the GD displayed a shift after application of the signal and this outcome was more potent when $e_W' = 0.01$ (Fig. 1B).

Using this model for inducible drive efficiency, we examined alteration of multiple variables within the simulation to determine which range of conditions allowed for a successful reversion to a wild population (Fig. 1C). In the standard model[31], we illustrated the allele frequency of the gene drive (red) for varying conditions of e_W using a constant initial value for f (Graph i); all scenarios resulted in GD takeover. However, when the drive included an induced e_W' of 0.1 (Graph ii), only lower values of e_W allowed for the population to revert to wild type. These conclusions were based on a given set of initial parameters; these could shift depending on a number of input variables. One of the limitations of initial e_W values was due to the relatively high value of e_W' (0.1). Modeling the population trajectory for various e_W' values demonstrated that in order for the drive to revert to WT, this variable should have a very low value (Graph iii). Moreover, this was confirmed when we explored all combinations of e_W and e_W' using set values for f and α (Fig. S3). Another critical factor for this inducible system was the amount of the population impacted by the shift (α); this system required a high success rate for reversion (Graph iv). For all possible combinations of α and e_W' it was apparent that successful reversion would be achieved by simultaneously maximizing α and minimizing e_W' (Fig. S3). Drive fitness also had an effect; lower values of f allowed the system to remove GD individuals from the population following the inducible signal (Graph v). Finally, there remained some flexibility as to the timing of e_W' induction; application of the signal at generation 10 or 15 still allowed for full reversion to WT (Graph vi). Together, these data demonstrated that an inducible level of drive action could serve as a means to reverse a population to only WT individuals.

An inducible gene drive system (individual fitness). We also modeled an inducible system for altering individual GD fitness (Fig. 2). We envisioned that a reduction in fitness for GD individuals would be empirically determined and pre-programmed either within the drive locus or distant native loci (within split drives). We examined variation of initial conditions (Fig. 2A) similar to GDs with inducible drive efficiencies (Fig. 1C). In the standard model, a variety of initial f values allowed for successful GD takeover (Graph i). Population reversion was maintained for most e_W conditions with the exception of very high values. High e_W values prior to induction resulted in a critical gene drive frequency threshold being reached that, given the initial conditions, could not be reverted to WT (Graph ii). Altering the induced fitness parameter revealed that f' needed to be lowered to approximately 0.4 (for the given set of initial conditions) and this was in stark contrast to recommended levels

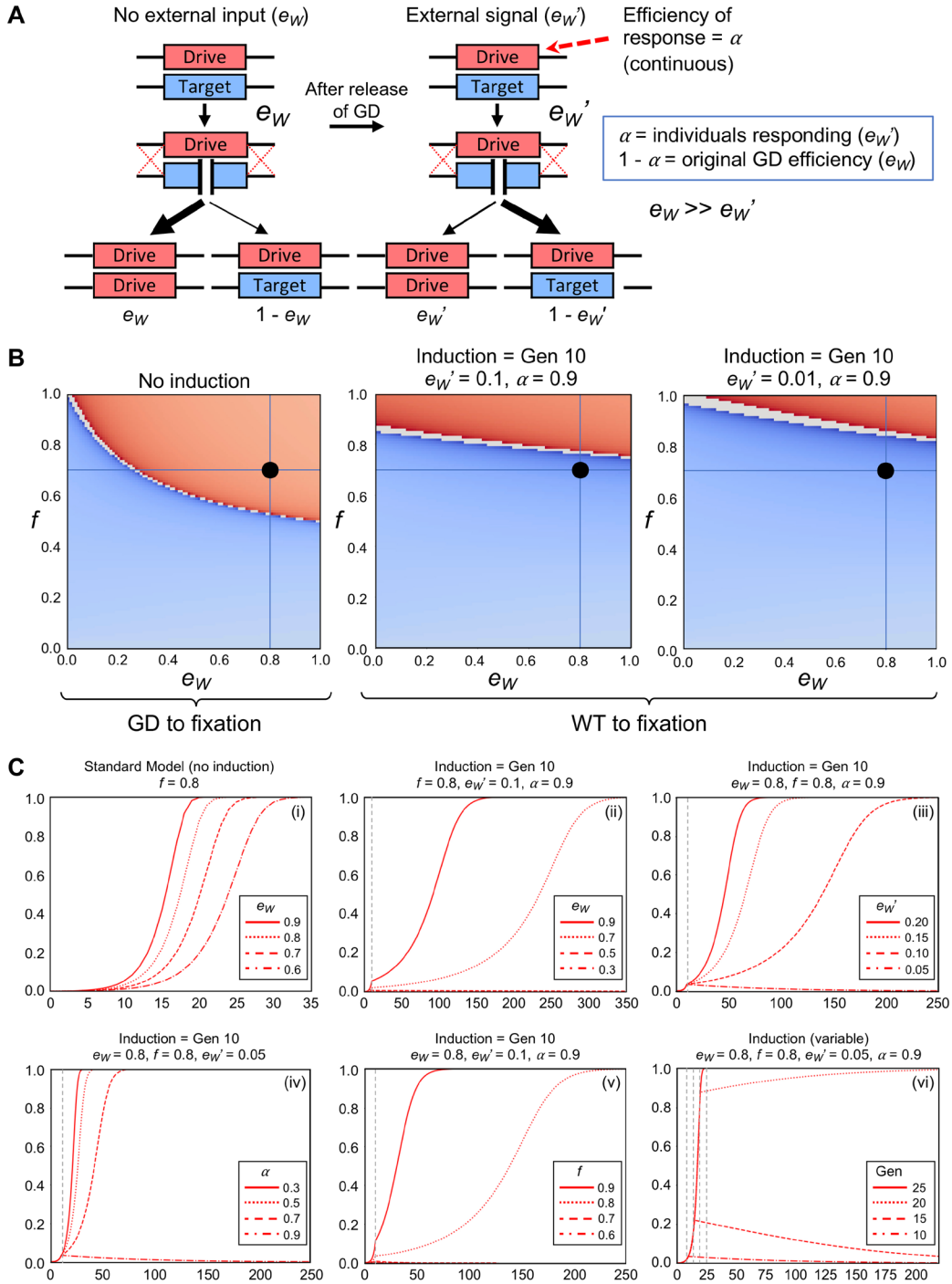

Figure 1. An inducible system to alter gene drive efficiency. (**A**) A theoretical GD system can translate an external signal into a change in GD effectiveness from initial e_W to a desired e_W'. The portion of the population that successfully responded to the external cue was designated α; unaffected population was defined by (1 - α). The example illustrated assumed that the initial e_W was larger than the final e_W' following the shift (population reversion back to WT) and α was set to one. (**B**) A reduced e_W' allowed an initially successful GD to be overtaken by the WT population. Graphs plotting GD e_W and f (*left*, no induction) for an inducible system where e_W' was changed to 0.1 (*middle*) or to 0.01 (*right*) at generation 10 with an effectiveness of $\alpha = 0.9$ are illustrated. A sample data point was highlighted with initial conditions of $e_W = 0.8$ and $f = 0.7$. The coloring scheme was identical to Fig. S2; GD to fixation (red) and WT to fixation (blue). Shading of colors illustrated the length of time required to reach fixation (darker, longer time). (**C**) Examination of varying initial parameters for a GD with inducible e_W. Graphs (i-vi) illustrated GD allele frequencies in red. Population frequency was plotted on the y-axis and generation time was plotted on the x-axis. A single parameter was altered for each graph and the values tested were illustrated with line types (solid, dotted, dashed, and dash/dot). Graph (i) did not include an inducible mechanism. Vertical grey dotted lines indicated the starting point for application of the external signal.

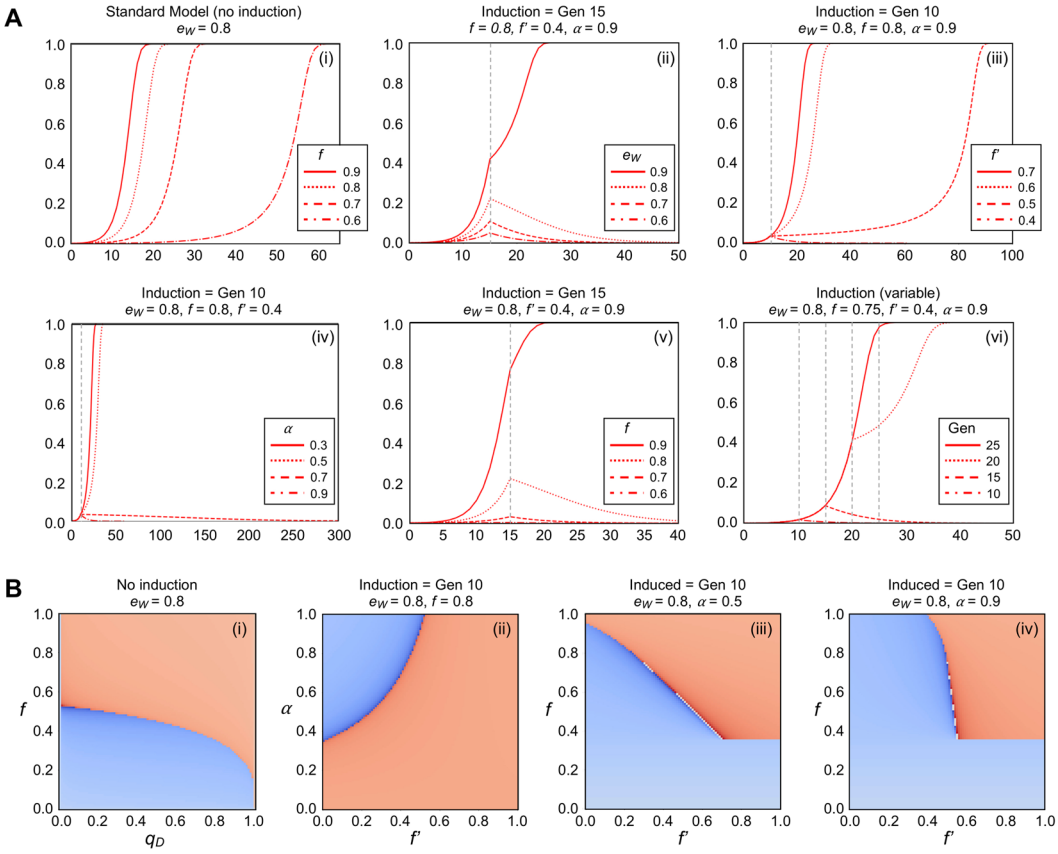

Figure 2. An inducible system to alter gene drive individual fitness. (**A**) Examination of both GD efficiency (e_W) and fitness (f) was explored for a drive system that included an inducible fitness parameter (f'). The induction efficiency was denoted as α. Analysis of varying initial parameters for a GD with inducible f; graphs (i-vi) illustrated GD allele frequencies in red. Frequency was plotted on the y-axis and generation time was plotted on the x-axis. Line types corresponded to four tested values within each parameter. Graph (i) did not include an inducible mechanism. Vertical grey dotted lines highlighted the induction times used. (**B**) Four sets of simulations were performed (i-iv) comparing two parameters for GD systems with inducible f. Graph (i) did not include an inducible GD. The coloring scheme matched the simulations found in Fig. S2.

for e_W' (Graph iii). Along these lines, there was not a strict a requirement for values of α; 0.7 was sufficient to still revert the population to WT unlike for inducible e_W models (Graph iv). Moreover, a wider range of initial f values (0.8 and below) still allowed for this successful population shift (Graph v). Finally, the timing of the external signal needed to take place early within drive propagation (e.g. 10 or 15 generations); scenarios where the drive frequency was too high did not allow for a shift despite a high value of α (Graph vi).

We also examined the existing (or initial) frequency of drive alleles within a population and drive fitness independent of any inducible system (Fig. 2B, Graph i). For our simulations, we included an initial GD frequency of 0.001; this translated into a requirement of GD fitness of greater than approximately 0.5. However, as the frequency of GD alleles increased within the population, the drive fitness could be lowered, and this still resulted in an overall GD takeover. Also, the model with inducible f allowed for a broader range of values for both α (as low as 0.4) and f' (nearly 0.5) compared with e_W' that would still allow for a population reversion to WT (Fig. 2B, Graph ii). Finally, comparison of f and f' at two values for α (Graphs iii and iv) demonstrated that when a larger proportion of individuals responded to the external signal ($\alpha = 0.9$), the initial value of f became less significant (Fig. 2B). For a specific set of initial conditions ($e_W = 0.8$, $f < 0.3$), the simulation assigned the WT because the GD was removed from the population prior to induction at generation 10. Interestingly, our comparisons also highlighted an *increase* in GD fitness for certain combinations of f and f' such that an inducible f gain could allow a GD to propagate despite a low initial fitness <0.5. Together, these data illustrate the potency of alteration of drive fitness across a range of initial conditions compared to inducible drive efficiency.

Modeling an inducible self-cleaving gene drive. While an inducible GD system to modulate e_W or f provides a suite of options, we reasoned that a different regulatory system could be employed to more rapidly remove GD individuals from a population. Therefore, we designed a theoretical GD that would include a self-cleaving module (Fig. 3A). One possible architecture could include an inducible guide RNA cassette that would target the nuclease to sites flanking the drive itself. Alternatively, multiple nucleases might be employed that utilize distinct and non-compatible guide RNAs with the same effect. In our model, we have included only a single nuclease type such that GD efficiency was identical for all GD actions. Here, e_W represented the efficiency

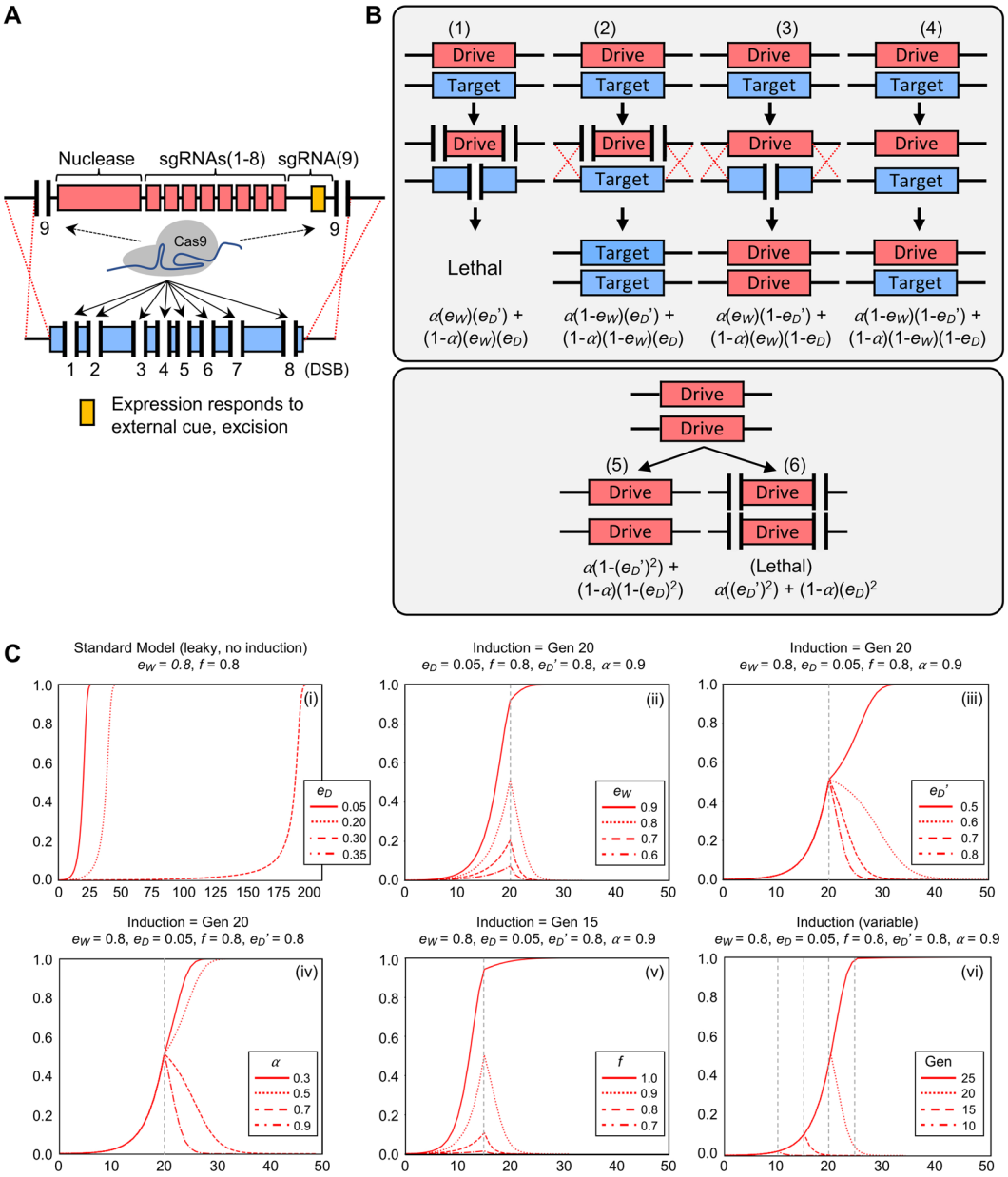

Figure 3. An inducible self-cleaving gene drive system. (**A**) Schematic of a GD arrangement that includes an inducible guide RNA cassette—sgRNA(9)—that would activate self-cleavage of the drive itself. (**B**) Two GD efficiencies were modeled: (i) e_W represented the ability to cleave the intended WT target and (ii) e_D represented the ability to self-target the drive cassette. Four scenarios were outlined (1–4) between pairing of the GD and WT alleles that resulted in a lethal phenotype, homozygous WT, homozygous GD, and heterozygous GD/WT genotypes. Activation of e_D prior to activation of the external cue (e.g. "leaky" activation) was included within the model; e_D' represented self-cleavage after the induction event. The outcomes of a homozygous GD/GD pairing were also modeled (scenarios 5–6). (**C**) Graphs (i-vi) illustrated drive allele frequencies in red. The allele frequency was plotted on the y-axis and the generation time was plotted on the x-axis. Line types (solid, dotted, dashed, and dash/dot) corresponded to four values sampled for each simulation. Graph (i) did not include an inducible element; a basal "leaky" value of e_D was set to various values at generation 0. Vertical grey dotted lines specified application of the external signal.

for cleaving the WT target (and copying of the drive via HDR) and e_D represented the efficiency for self-cleaving the GD locus (and subsequent repair) (Fig. 3B). For pairings between the drive and WT, we modeled four separate outcomes: (i) dual cleavage of both the target and drive, (ii) cleavage of only the drive, (iii) cleavage of only the target, and (iv) no cleavage resulting in a heterozygous individual. We reasoned that in our GD design, dual cleavage events to both chromosomes (and removal of any programmed essential gene) would be lethal and that single cleavage of only the GD would present a unique scenario where the WT gene would be doubled to

recreate a homozygous WT individual. For existing GD individuals in the population, successful activation of the self-cleavage module could result in excision of both alleles and a lethal phenotype.

In our model, we assumed that both self-cleaving events occurred at rates e_D or e_D' (after induction) and that HDR-based repair was responsible for repair following DNA cleavage. Finally, we also included calculation of a "leaky" system where an initial value of e_D existed (typically 0.05) prior to activation of the external signal α. This was to test whether a small amount of self-cleavage would affect the initial propagation of the GD system and represented the reasonable biological assumption that even tightly regulated promoters might allow for a small level of transcript (in the case of either the regulated self-targeting guide RNA and/or second nuclease). For a given set of initial conditions, we modeled increasing levels of e_D in a system that did not include any inducible signal (Fig. 3C, Graph i). Interestingly, the system was robust for "leaky" levels of self-cleavage up to 0.3; the overall effect was the shift in the length of time required for GD takeover.

Modeling of the inducible self-cleaving system demonstrated a variety of initial parameters could be accommodated that still resulted in successful population reversion to WT (Fig. 3C). Values of e_W up to 0.8 were still tolerated (Graph ii) as well as induced e_D' values as low as 0.6 (Graph iii). Similar to the application success rate for inducible f, values for α did not need to be extremely high (Graph iv). Furthermore, drives with fitness values up to 0.9 could still be reverted using this system (Graph v). We observed that activation of this system could be relatively late (generation 20) compared with the length of time required for GD takeover for a given simulation and the allele frequency decline was much more rapid than compared with models for inducible e_W or f (Graph vi). We also expanded our analysis of values for e_W, e_D', and induction time which would account for varying values between e_W and e_D/e_D' (Fig. S4). We found that this type of self-cleaving system could still allow for reversion to WT even when the induced e_D' was identical to e_W (for single-nuclease drives) or for scenarios where $e_D' < e_W$ (for dual nuclease drives with differences in efficiency). Together, these data demonstrated an alternative mechanism for inducible GD population suppression that provides a unique set of requirements for GD strength and fitness.

Discussion

Few studies have proposed mechanisms by which removal of GD organisms could be accomplished once already released. Current ideas include gene drives that counter other gene drives (anti-drives)[33,34], drives with built-in limitations (daisy-chain drives)[35,36], or underdominance drives (UD)[35,37–39]. In the case of the anti-drives, a standard reversal system only targets GD individuals; immunizing anti-drives are able to target both GD and WT individuals. For some scenarios, AD systems may not eliminate drives within a population, and, instead, might achieve a stable equilibrium[33]. Moreover, without additional modifications, AD systems require construction at the same locus that the GD was originally installed. This might prove challenging in some scenarios where anti-drives are not already engineered and available for release. Daisy-chain systems can provide local spread of a drive element but they cannot propagate at the same scale as traditional drives; the goal is limited spread of drives, rather than targeted removal of active drives[36]. Finally, in the case of UD systems, proposals for population reversal requires the introduction of either wild-type individuals[40] or "free suppressor" individuals[38]. Therefore, we focused on novel drive systems that might provide a variable level of population control and ultimately allow for reversion to the original WT species without the need for a secondary release of modified (or wild-type) individuals.

Previous theoretical models have included a metric for gene drive efficiency (e, represented in our work as e_W). In our system, we assumed that the two requirements of GD action—cleaving the target(s) and subsequent DNA repair via HDR—were coupled together. Programmed values of e_W less than 1 would result in a predictable frequency of heterozygous individuals without any resistant allele(s). Titration of e_W has the advantage of being generally applicable to other eukaryotic systems. For instance, nucleocytoplasmic transit (of Cas9) is widely conserved across species and expression of the anti-CRISPR peptides would be predicted to function regardless of the cell type or organism. In its simplest form, application of a gene drive where $e_W < 1$ would provide additional time to implement changes or release and distribute countering agents. However, as our model illustrates, use of an inducible e_W to favor return of the WT population after GD release would require an external signal to be recognized by the entire population. While this is certainly feasible when using naturally occurring events (temperature, environmental conditions, etc.), this would present additional challenges for delivery and application, but could certainly be attempted on isolated small-scale field trials or laboratory settings.

We also recognize that the identity of the external signal would have its own challenges and/or potential costs. Ideally, delivery of the signal would have a minimal or negligible effect on the surrounding ecosystem (and human health in the case of agricultural application). We found a much stronger effect of altering the overall fitness of the organism. However, the precise genetic target and mechanism for altering f would likely be species-specific and may be difficult to quantify. However, titration of f allowed for a lower overall sample of the population to respond to an added cue compared to e_W. While others have demonstrated that installation of drives sometimes includes an overall fitness cost, our data demonstrate that there are levels of f that can still allow for GDs to effectively propagate through a population similar to previous findings[41]. Along these lines, previous studies have recommended that "responsible" gene drive systems have a purposefully reduced fitness[42].

Our proposal for an inducible self-cleaving gene drive system would provide multiple benefits in terms of regulation and drive reversal. For one, cleavage of the GD without cleavage of the WT target allele would provide a unique scenario where the WT locus was copied to replace the gene drive. This provides a more potent option than either a reduction of drive efficiency, drive fitness, or even use of an anti-drive because it (i) removes the GD allele directly and (ii) simultaneously increases the proportion of WT alleles within the population (without any supplementation of additional individuals). While our system demonstrated that use of an additional guide RNA cassette could be activated using the same nuclease, two separate CRISPR systems could be employed such that the level of effectiveness for e_W and e_D/e_D' could be tuned accordingly. Moreover, our findings illustrated that this type of drive system required a lower rate of response to an external cue compared to a switch in e_W' and could

Gene Drive Strategy	Description and Outcomes	Pros/Cons[a]
1. Reduced initial e_W and/or f	Slowed drive propagation; sensitized GD population	Requires identification of conserved mechanisms for titration of e_W. Species-specific means to reduce (and quantify) fitness. Cannot alter population once released. Pre-determined outcome for gene drive success.
2. Inducible e_W	Ability to prevent GD take over, revert population to WT	Requires significant induction success rate (α) and very low value for e_W'. Requires continuous application.
3. Inducible f	Ability to prevent GD take over, revert population to WT	Requires lower success rate and less severe reduction in f'. Likely species-specific mechanism. Requires continuous application.
4. Inducible self-cleaving GD	Direct removal of GD alleles without the need for an anti-drive system, reversion to WT population	Allows for differences in e_W and e_D/e_D' for degree of GD removal. Does not require a high induction rate α. System is robust against modest degree of inappropriate activation and self-cleavage. Requires continuous application.

Table 1. Summary of proposed gene drive designs and potential benefits and challenges. [a]The advantages and challenges for each system are discussed. These are not focused on challenges of design and creation of the intended drive, but rather on drive application and potential uses.

still serve as an effective GD despite a high basal level of e_D activation and self-removal of a fraction of the population. Coupling separate inducible systems together (such as self-cleavage, GD fitness, and/or e_W) could provide multiple independent mechanisms to control drive propagation and allow for targeted drive removal should the need arise.

While a complex gene drive that allows for external input may face technical challenges in design and construction, the goal of this study was to explore theoretical systems that would specifically address the issues of containment and reversal/removal (Table 1). We envision that these types of safety mechanisms would provide additional levels of control and programmability that are not currently possible in simple GD setups that are designed with only initial parameters and a single outcome. Furthermore, failsafe systems to protect the original WT species, even if never used in application, are a critical step towards gaining support for future laboratory research and possible application of gene drives within native ecosystems.

Methods

Population models for inducible gene drive efficiency and fitness. We expanded on a previous frequency-based model (illustrated below) of gene drive population dynamics; this model assumed random mating, non-overlapping generations and an infinite population size[31]. Our modifications to this equation introduced new variables for our inducible and self-cleaving designs (e_W', f', e_D, and e_D').

$$q_D^{'} = \frac{f(q_D^2 + 2q_W q_D e_W) + (1 - h + fh)(q_W q_D(1 - e_W))}{1 - (1 - f)(q_D^2 + 2q_W q_D e_W) - (h - fh)2q_W q_D(1 - e_W)}$$

Frequency-based model. We assumed no GD-resistance of any kind (including no NHEJ-based repair and subsequent resistance). For the standard model, we included two allele types, W and D, which denoted wild-type and gene drive with frequencies q_W and q_D, respectively. The summation of these frequencies always totaled to one. This resulted in three distinct classes; homozygous wild-type, homozygous gene drive, and the heterozygote between wild-type and gene drive. The success rate of the GD, the ability to both cleave WT target DNA and copy the drive cassette to the homologous chromosome, was denoted by e_W. While we recognize that alternate scenarios exist (such as DNA cleavage without copying of the drive), these would be accounted for in the overall design of the drive itself—NHEJ-based repair in the absence of copying of the drive would result in individuals that would be sterile or inviable (see Fig. S1). Failure of the drive was denoted by $(1 - e_W)$.

Heterozygote fitness. The gene drive homozygote had an associated fitness relative to wild-type that was represented by the variable f. The fitness of the heterozygote was calculated using a degree of dominance, h; this was represented by $(1 - h) + fh$. For all simulations, the fitness of wild type individuals was set to one and h was set to 0.5. For heterozygous individuals (WT/GD) with an active drive system, the genotype was immediately converted to the homozygous state (GD/GD) and the corresponding fitness of the individual was calculated based on this genotype.

Applied external signal. Our equations for inducible drives also introduced the variable α which represented the fraction of individuals affected by the external inducing agent. This value was shifted from zero to a specific level at a predetermined generation (the shift occurred within one generation). The inducing agent was then continuously applied at the same level for every subsequent generation after its initial release.

Simulation parameters. The initial frequency of the GD was set to 0.1% of the total population. Simulations were run until the population achieved (i) fixation of the GD, (ii) loss of the GD, or (iii) the maximum number of generations (1,000) was reached. We declared the population had reached fixation when any allele had a frequency

>0.99999. Graphics utilized a 2D graphics package[43]. The allele frequencies of the next generation (q_W' and q_D') were calculated by the following equations, where $q_W' = 1 - q_D'$.

Equation for inducible efficiency

$$q_D' = \frac{f(q_D^2 + 2q_Wq_D(\alpha e_W' + (1-\alpha)e_W)) + (1 - h + fh)(q_Wq_D(1 - (\alpha e_W' + (1-\alpha)e_W)))}{1 - (1-f)(q_D^2 + 2q_Wq_D(\alpha e_W' + (1-\alpha)e_W)) - (h - fh)2q_Wq_D(1 - (\alpha e_W' + (1-\alpha)e_W))} \quad (1)$$

Equation for inducible fitness

$$q_D' = \frac{(\alpha f' + (1-\alpha)f)(q_D^2 + 2q_Wq_De_W) + (1 - h + (\alpha f' + (1-\alpha)f)h)(q_Wq_D(1 - e_W))}{1 - (1 - (\alpha f' + (1-\alpha)f))(q_D^2 + 2q_Wq_De_W) - (h - (\alpha f' + (1-\alpha)f)h)2q_Wq_D(1 - e_W)} \quad (2)$$

Model for self-cleaving gene drive system

The model for the inducible self-excising drive introduced new variables e_D and e_D'. The presence of e_D (initial self-cleavage of the gene drive) existed to allow for any possible activation prior to exposure to the inducing agent (e.g. "leaky" inducible system) while e_D' represented the intentional self-cleaving activation.

$$q_D' = \frac{f(q_D^2(1 - ((1-\alpha)e_D + \alpha e_D')^2) + 2q_Wq_D((1-\alpha)e_D + \alpha e_D')(1 - e_W)) + (1 - h + fh)q_Wq_D(1 - ((1-\alpha)e_D + \alpha e_D'))(1 - e_W)}{1 - (1-f)(q_D^2(1 - ((1-\alpha)e_D + \alpha e_D')^2) + 2q_Wq_D(1 - ((1-\alpha)e_D + \alpha e_D'))e_W) - (h - fh)2q_Wq_D(1 - ((1-\alpha)e_D + \alpha e_D'))(1 - e_W) - q_D^2((1-\alpha)e_D + \alpha e_D')^2 - 2q_Wq_De_W((1-\alpha)e_D + \alpha e_D')} \quad (3)$$

Animal and human subject statement. This study does not use any animals or human subjects.

Data availability

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request for research purposes.

Received: 5 January 2019; Accepted: 19 November 2019;
Published online: 27 December 2019

References

1. Card, D. C. *et al.* Novel ecological and climatic conditions drive rapid adaptation in invasive Florida Burmese pythons. *Molecular ecology*. https://doi.org/10.1111/mec.14885 (2018).
2. Marshak, A. R., Heck, K. L. Jr. & Jud, Z. R. Ecological interactions between Gulf of Mexico snappers (Teleostei: Lutjanidae) and invasive red lionfish (Pterois volitans). *PloS one* **13**, e0206749, https://doi.org/10.1371/journal.pone.0206749 (2018).
3. Goodell, K. & Parker, I. M. Invasion of a dominant floral resource: effects on the floral community and pollination of native plants. *Ecology* **98**, 57–69, https://doi.org/10.1002/ecy.1639 (2017).
4. Westbrook, C. E., Ringang, R. R., Cantero, S. M. & Toonen, R. J. Survivorship and feeding preferences among size classes of outplanted sea urchins, Tripneustes gratilla, and possible use as biocontrol for invasive alien algae. *PeerJ* **3**, e1235, https://doi.org/10.7717/peerj.1235 (2015).
5. Berthet, N. *et al.* Molecular characterization of three Zika flaviviruses obtained from sylvatic mosquitoes in the Central African Republic. *Vector borne and zoonotic diseases (Larchmont, N.Y.)* **14**, 862–865, https://doi.org/10.1089/vbz.2014.1607 (2014).
6. Nelder, M. P. *et al.* Human pathogens associated with the blacklegged tick Ixodes scapularis: a systematic review. *Parasites & vectors* **9**, 265, https://doi.org/10.1186/s13071-016-1529-y (2016).
7. Rezza, G. Aedes albopictus and the reemergence of Dengue. *BMC public health* **12**, 72, https://doi.org/10.1186/1471-2458-12-72 (2012).
8. Kampango, A., Braganca, M., Sousa, B. & Charlwood, J. D. Netting barriers to prevent mosquito entry into houses in southern Mozambique: a pilot study. *Malaria journal* **12**, 99, https://doi.org/10.1186/1475-2875-12-99 (2013).
9. Vu, S. N. *et al.* Elimination of dengue by community programs using Mesocyclops(Copepoda) against Aedes aegypti in central Vietnam. *The American journal of tropical medicine and hygiene* **72**, 67–73 (2005).
10. Godfray, H. C. J., North, A. & Burt, A. How driving endonuclease genes can be used to combat pests and disease vectors. *BMC biology* **15**, 81, https://doi.org/10.1186/s12915-017-0420-4 (2017).
11. Hammond, A. M. & Galizi, R. Gene drives to fight malaria: current state and future directions. *Pathogens and global health* **111**, 412–423, https://doi.org/10.1080/20477724.2018.1438880 (2017).
12. DiCarlo, J. E., Chavez, A., Dietz, S. L., Esvelt, K. M. & Church, G. M. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nature biotechnology* **33**, 1250–1255, https://doi.org/10.1038/nbt.3412 (2015).
13. Roggenkamp, E. *et al.* Tuning CRISPR-Cas9 Gene Drives in Saccharomyces cerevisiae. *G3 (Bethesda, Md.)* **8**, 999–1018, https://doi.org/10.1534/g3.117.300557 (2018).
14. Grunwald, H. A. *et al.* Super-Mendelian inheritance mediated by CRISPR/Cas9 in the female mouse germline. *bioRxiv*, https://doi.org/10.1101/362558 (2018).
15. Gantz, V. M. *et al.* Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E6736–6743, https://doi.org/10.1073/pnas.1521077112 (2015).
16. Kyrou, K. *et al.* A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nature biotechnology* **36**, 1062–1066, https://doi.org/10.1038/nbt.4245 (2018).
17. KaramiNejadRanjbar, M. *et al.* Consequences of resistance evolution in a Cas9-based sex conversion-suppression gene drive for insect pest management. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 6189–6194, https://doi.org/10.1073/pnas.1713825115 (2018).
18. Prowse, T. A., Adikusuma, F., Cassey, P., Thomas, P. & Ross, J. V. A Y-chromosome shredding gene drive for controlling pest vertebrate populations. *eLife* **8**, https://doi.org/10.7554/eLife.41873 (2019).
19. Basgall, E. M. *et al.* Gene drive inhibition by the anti-CRISPR proteins AcrIIA2 and AcrIIA4 in Saccharomyces cerevisiae. *Microbiology (Reading, England)* **164**, 464–474, https://doi.org/10.1099/mic.0.000635 (2018).
20. Liu, K. I. *et al.* A chemical-inducible CRISPR-Cas9 system for rapid control of genome editing. *Nature chemical biology* **12**, 980–987, https://doi.org/10.1038/nchembio.2179 (2016).

21. Rauch, B. J. *et al*. Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell* **168**, 150–158.e110, https://doi.org/10.1016/j.cell.2016.12.009 (2017).
22. Pawluk, A. *et al*. Naturally Occurring Off-Switches for CRISPR-Cas9. *Cell* **167**, 1829–1838.e1829, https://doi.org/10.1016/j.cell.2016.11.017 (2016).
23. Marino, N. D. *et al*. Discovery of widespread type I and type V CRISPR-Cas inhibitors. *Science (New York, N.Y.)* **362**, 240–242, https://doi.org/10.1126/science.aau5174 (2018).
24. Watters, K. E., Fellmann, C., Bai, H. B., Ren, S. M. & Doudna, J. A. Systematic discovery of natural CRISPR-Cas12a inhibitors. *Science (New York, N.Y.)* **362**, 236–239, https://doi.org/10.1126/science.aau5138 (2018).
25. Zetsche, B., Volz, S. E. & Zhang, F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nature biotechnology* **33**, 139–142, https://doi.org/10.1038/nbt.3149 (2015).
26. Maji, B. *et al*. Multidimensional chemical control of CRISPR-Cas9. *Nature chemical biology* **13**, 9–11, https://doi.org/10.1038/nchembio.2224 (2017).
27. López Del Amo, V. *et al*. Small-molecule control of super-Mendelian inheritance in gene drives. *bioRxiv*, 665620, https://doi.org/10.1101/665620 (2019).
28. Polstein, L. R. & Gersbach, C. A. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nature chemical biology* **11**, 198–200, https://doi.org/10.1038/nchembio.1753 (2015).
29. Kipniss, N. H. *et al*. Engineering cell sensing and responses using a GPCR-coupled CRISPR-Cas system. *Nature communications* **8**, 2212, https://doi.org/10.1038/s41467-017-02075-1 (2017).
30. Baeumler, T. A., Ahmed, A. A. & Fulga, T. A. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. *Cell reports* **20**, 2639–2653, https://doi.org/10.1016/j.celrep.2017.08.044 (2017).
31. Deredec, A., Burt, A. & Godfray, H. C. The population genetics of using homing endonuclease genes in vector and pest management. *Genetics* **179**, 2013–2026, https://doi.org/10.1534/genetics.108.089037 (2008).
32. Noble, C., Adlam, B., Church, G. M., Esvelt, K. M. & Nowak, M. A. Current CRISPR gene drive systems are likely to be highly invasive in wild populations. *eLife* **7**, https://doi.org/10.7554/eLife.33423 (2018).
33. Vella, M. R., Gunning, C. E., Lloyd, A. L. & Gould, F. Evaluating strategies for reversing CRISPR-Cas9 gene drives. *Scientific reports* **7**, 11038, https://doi.org/10.1038/s41598-017-10633-2 (2017).
34. Wu, B., Luo, L. & Gao, X. J. Cas9-triggered chain ablation of cas9 as a gene drive brake. *Nature biotechnology* **34**, 137–138, https://doi.org/10.1038/nbt.3444 (2016).
35. Dhole, S., Vella, M. R., Lloyd, A. L. & Gould, F. Invasion and migration of spatially self-limiting gene drives: A comparative analysis. *Evolutionary applications* **11**, 794–808, https://doi.org/10.1111/eva.12583 (2018).
36. Noble, C. *et al*. Daisy-chain gene drives for the alteration of local populations. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 8275–8282, https://doi.org/10.1073/pnas.1716358116 (2019).
37. Edgington, M. P. & Alphey, L. S. Conditions for success of engineered underdominance gene drive systems. *Journal of theoretical biology* **430**, 128–140, https://doi.org/10.1016/j.jtbi.2017.07.014 (2017).
38. Edgington, M. P. & Alphey, L. S. Modeling the mutation and reversal of engineered underdominance gene drives. *Journal of theoretical biology* **479**, 14–21, https://doi.org/10.1016/j.jtbi.2019.06.024 (2019).
39. Dhole, S., Lloyd, A. L. & Gould, F. Tethered homing gene drives: A new design for spatially restricted population replacement and suppression. *Evolutionary applications* **12**, 1688–1702, https://doi.org/10.1111/eva.12827 (2019).
40. Champer, J., Buchman, A. & Akbari, O. S. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nature reviews. Genetics* **17**, 146–159, https://doi.org/10.1038/nrg.2015.34 (2016).
41. Unckless, R. L., Clark, A. G. & Messer, P. W. Evolution of Resistance Against CRISPR/Cas9 Gene Drive. *Genetics* **205**, 827–841, https://doi.org/10.1534/genetics.116.197285 (2017).
42. Tanaka, H., Stone, H. A. & Nelson, D. R. Spatial gene drives and pushed genetic waves. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 8452–8457, https://doi.org/10.1073/pnas.1705868114 (2017).
43. Hunter, J. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**, 90–95, https://doi.org/10.1109/MCSE.2007.55 (2007).

Acknowledgements

We thank Caterina Scoglio (Kansas State University) for useful suggestions and comments. This work was supported by the USDA National Institute of Food and Agriculture, Hatch Project 1013520 to G.C.F. The content is solely the responsibility of the authors and does not necessarily represent the official views of USDA.

Author contributions

M.G.H. generated *in silico* constructs and simulations, M.G.H. and G.C.F. performed all data analyses and figure preparation. G.C.F. wrote the manuscript.

Competing interests

G.C.F. has applied for CRISPR-related patents: (i) April 20, 2017 by the Univ. of California, Berkeley, "Methods and Compositions for Genomic Editing" International Application No. PCT/US2017/028676 and published as No. WO 2017/189336 A1 on November 2, 2017, (ii) September 29, 2017 (U.S. Provisional Patent Application Serial No. 62/565,651, "Programmed modulation of CRISPR/Cas9 activity") followed by patent filing on January 31, 2018 and (iii) a PCT application "Multi-Locus Gene Drive System" No. PCT/US2019/041538 on July 12, 2019. M.G.H. has private investments in companies that may utilize various CRISPR-related biotechnologies. Authors declare no non-financial conflict of interest of any kind.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-54805-8.

Correspondence and requests for materials should be addressed to G.C.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019