

Seed dormancy shapes gene drive dynamics in plants

Received: 24 April 2025

Accepted: 24 February 2026

Published online: 03 April 2026

 Check for updates

Isabel K. Kim ^{1,5}, Leqi Tian ^{1,2,5}, Ryan Chaffee ³, Benjamin C. Haller ¹, Jackson Champer ^{4,6}, Philipp W. Messer ^{1,6} & Jaehee Kim ^{1,6} ✉

Gene drives offer revolutionary potential for the management of problematic plant populations, such as invasive weeds and herbicide-resistant species, by rapidly spreading desired genetic alterations. Two recent studies have provided experimental demonstrations of engineered CRISPR gene drive systems in plants (CAIN and ClvR). However, the successful application of such systems in the field will critically depend on an accurate understanding of plant-specific life-history traits, especially seed dormancy, a ubiquitous yet frequently overlooked eco-evolutionary force. In this study, we develop a comprehensive modelling framework for gene drives in plant populations that incorporates a persistent soil seed bank. We show how the presence of a seed bank can substantially slow gene drive spread but also reduce the genetic load required to achieve population elimination. Furthermore, we show that seed banks substantially increase the required introduction frequency of threshold-dependent gene drives, which could prevent establishment in some cases, yet also provide an intrinsic biosafety mechanism for confining a highly efficient drive to a target population. Our study highlights the need to incorporate seed-bank dynamics into gene drive strategies to ensure realistic predictions and successful field applications.

Weeds are among the major biological threats to global agriculture, causing annual yield losses in major crops that result in substantial economic damage and threaten food security^{1–5}. Globally, weeds are responsible for around 10% of crop losses¹, and in the USA alone, they account for more than US\$26 billion a year in control expenses and lost crop yield⁶. The intensive and widespread use of chemical herbicides has historically provided the primary means of weed control; however, the rapid evolution of herbicide resistance poses an escalating challenge^{7–11}. There is thus an urgent need to develop novel, effective and evolutionarily robust weed management tools. One promising avenue for next-generation weed management is the use of gene drives—selfish genetic elements that can quickly propagate through populations, even

if they confer a fitness cost to individual organisms^{12–14}. Gene drives could be used for either population suppression or population modification of problematic weeds^{15–20}. A suppression drive, for example, could spread deleterious traits (such as sterility or non-viability) to eradicate an invasive weed population, whereas a modification drive could reverse herbicide resistance by propagating susceptibility alleles. This raises the possibility that engineered gene drives could one day provide lasting, inexpensive solutions for controlling weeds that are otherwise difficult or costly to manage.

Recently, two toxin–antidote gene drive constructs were successfully developed and experimentally validated in the model plant *Arabidopsis thaliana*^{21–23}: CRISPR-Assisted Inheritance using *NPG1* (CAIN²⁴;

¹Department of Computational Biology, Cornell University, Ithaca, NY, USA. ²Center for Genetic Epidemiology, Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. ³Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA. ⁴Center for Bioinformatics, School of Life Sciences, Center for Life Sciences, Peking University, Beijing, China. ⁵These authors contributed equally: Isabel K. Kim, Leqi Tian. ⁶These authors jointly supervised this work: Jackson Champer, Philipp W. Messer, Jaehee Kim.

✉ e-mail: jaehee.kim@cornell.edu

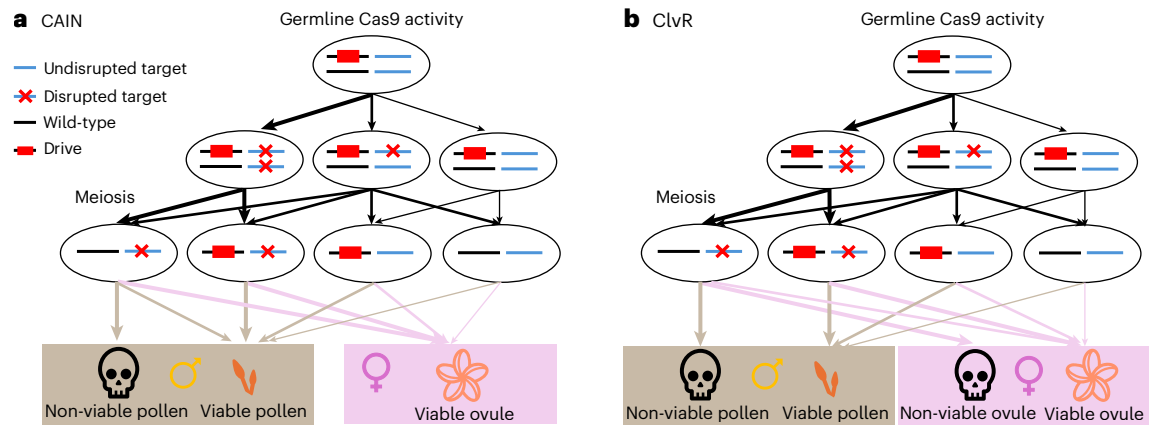


Fig. 1 | Gene drive systems. a, CAIN drive²⁴. In the germline of a drive carrier, the drive cleaves each wild-type target allele at a rate equal to the sex-specific germline cleavage rate. The target gene is essential only for pollen germination; thus, ovules of any genotype are viable. Pollen inheriting the drive construct are always viable, since the drive contains a cleavage-resistant recoded functional target allele. A pollen grain with a wild-type allele and a disrupted target gene (leftmost gamete, bottom row) is non-viable with a probability equal to the penetrance rate of the target gene (estimated at 96%). For suppression, we define the ‘CAIN male suppression’ drive as the CAIN drive inserted into an essential haplosufficient male fertility gene, such that males with two drive

alleles are sterile. **b**, ClvR drive²⁵. ClvR functions similarly to CAIN but targets a gene that is essential for viability in both pollen and ovules. However, ovules inheriting disrupted alleles and no drive allele from drive-carrying females (leftmost gamete, bottom row) remain viable with a probability equal to a maternal carryover rate (estimated at 20.7%); otherwise, they are non-viable. For suppression, we define two drives. ‘ClvR male suppression’ targets an essential haplosufficient male fertility gene, such that homozygous males are sterile, and ‘ClvR female suppression’ targets an essential haplosufficient female fertility gene, such that homozygous females are sterile.

Fig. 1a) and Cleave-and-Rescue gamete killer (ClvR²⁵; Fig. 1b). Unlike traditional gene drives that depend on homology-directed repair (an inefficient mechanism in plants due to their strong tendency towards end-joining DNA repair pathways²⁶), toxin–antidote gene drives avoid the need for homology-directed repair altogether. Instead, these drives spread by linking a toxin (Cas9 and guide RNAs (gRNAs) targeting an essential gene) to an antidote (a tightly linked, cleavage-resistant copy of the target gene), resulting in the elimination of genotypes carrying disrupted alleles but lacking the drive²⁷. In experimental crosses, CAIN targeted *NPG1*, an essential gene required for pollen germination, achieving inheritance rates up to 97% through male gametes²⁴. ClvR targeted *YKT61*, a gene required for viability in both pollen and ovules, achieving near-complete inheritance through males and substantial, though lower, inheritance rates through females²⁵. Preliminary modelling in both studies indicated that such drives could spread rapidly and reliably eliminate populations under panmictic, outcrossing conditions^{24,25}.

However, these initial plant gene drive studies left out a key aspect of plant life history: seed dormancy. Dormancy allows seeds to remain in the soil without germinating for extended periods. This trait leads to a persistent seed bank, which can provide a genetically diverse reservoir of alleles in the soil that can buffer the population against rapid evolutionary or environmental changes^{28–31}. In the presence of a seed bank, gene drive alleles might spread more slowly, extending the time to population modification or suppression^{16,32,33}. Conventional models without seed banks assume that every individual reproduces in each generation, so they cannot capture these age-structured, time-delayed dynamics. Given how common seed banks are among weeds, there is a need to incorporate seed-bank dynamics into gene drive models, not only to improve theoretical accuracy but also to inform feasibility and risk assessment and to identify which ecological parameters most strongly affect spread, timelines and containment prior to any field deployment^{34,35}. In particular, an understanding of seed-bank effects should inform the choice of drive architecture, help determine how many modified individuals should be released and set realistic expectations for the timeline and ultimate efficacy of a gene drive under field conditions. In this study, we address this critical knowledge gap by developing a detailed, individual-based model for the eco-evolutionary dynamics of gene drives in plant populations that includes realistic life cycles and seed-bank features.

Results

We developed a comprehensive simulation model of CAIN²⁴ and ClvR²⁵ (Fig. 1), the only experimentally demonstrated gene drive constructs in plants to date, using the individual-based forward-time simulation framework SLiM v.4.0.1 (ref. 36). Our baseline models focused on dioecious, annual, diploid weed species, but we also explored hermaphroditic species with varying selfing rates. Detailed model descriptions are provided in the Methods, and major symbols, definitions and default values (where applicable) are summarized in Supplementary Tables 1–3.

Briefly, each drive was modelled as either a modification drive or a suppression drive: modification drives carried a desirable allele, and suppression drives induced sterility in drive-homozygous males or females, eliminating the population through pollen or ovule limitation, respectively. Unless noted, drive parameters were fixed at experimentally derived default values (Supplementary Text 1 and Supplementary Table 1). Populations were initialized at carrying capacity (K), equilibrated and then subjected to a single drive-release event. Unless stated otherwise, we implemented releases at the plant stage to set a controlled initial plant-stage drive allele frequency, p_0 . With selection, mating and demographic feedback acting at the plant stage, the realized plant-stage drive frequency provides a standardized and comparable baseline across drive designs and ecological regimes. It also defines the most feasible target for near-term monitoring (for example, plant census and genotyping) without requiring measurement of the below-ground seed bank. To connect this baseline to operational seed-based deployment, we also explicitly modelled seed-drop releases and provide analytical and numerical mappings from seed introduction frequency to the realized plant-stage drive allele frequency.

We modelled the life cycle of dioecious, obligate outcrossing, annual plants with seed-bank dynamics, assuming wind pollination and considering only ‘effective’ gametes that contribute to reproduction (Fig. 2). New seeds enter a seed bank and undergo age-dependent survival and germination, yielding total germination probability γ and average seed-bank duration τ . These metrics depend on seeds’ baseline survival rate (d , first-year survival), baseline germination rate (b , first-year germination conditional on survival) and age-dependent declines in seed survival (q) and germination (m). Seeds older than L are removed from the seed bank.

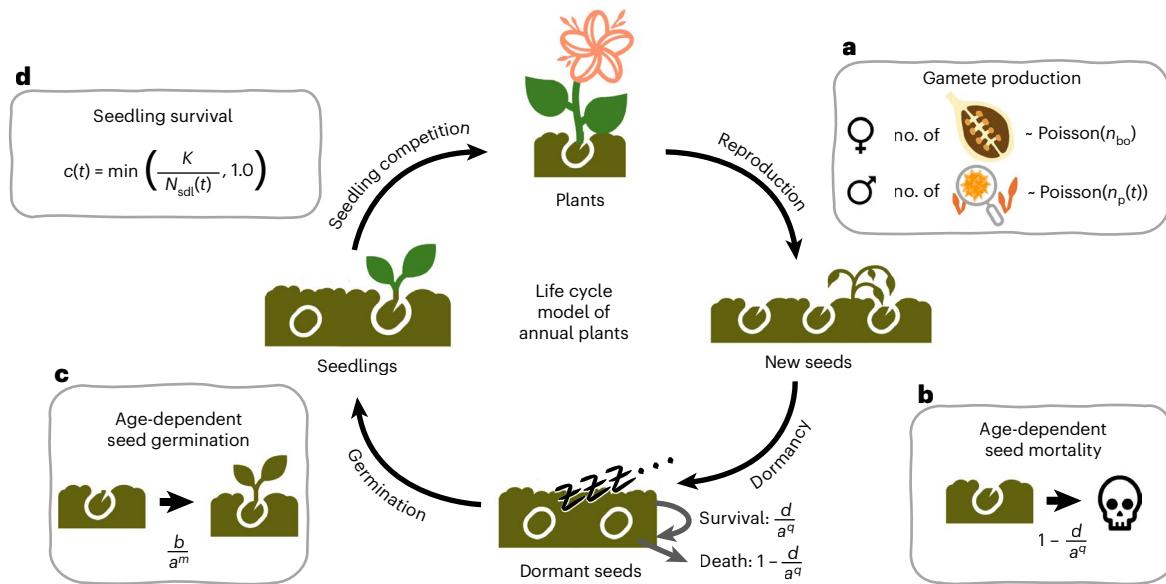


Fig. 2 | Life cycle model of annual plants. Illustration of the life cycle modelled for a dioecious annual plant population. **a**, Plants to seeds: an ‘effective ovule’ is defined as an ovule capable of fertilization and subsequent seed development and an ‘effective pollen grain’ as a pollen grain that can successfully reach a fertile female plant via wind and germinate. Each fertile female plant produces a Poisson-distributed number of effective ovules with baseline mean n_{bo} . Each fertile male plant produces a Poisson-distributed number of effective pollen grains with baseline mean $n_p(t)$, which is proportional to the current number of fertile females. Effective pollen grains are randomly distributed among fertile female plants, and fertilized ovules develop into seeds. After reproduction, all plants are removed from the population. **b, c**, Into and out of the seed bank:

age-dependent seed mortality and germination. A seed survives its first year at rate d , after which survival declines with seed age, such that a seed of age a survives at rate d/a^q , where q modulates age-dependent mortality. Similarly, germination occurs initially at rate b and declines with seed age, such that a seed of age a germinates at rate b/a^m , where m modulates age-dependent germination. **d**, Seed to plants: seedling competition. Germinated seeds become seedlings, which experience density-dependent competition before becoming adult plants. Each seedling at year t survives to adulthood with probability $c(t) = \min(\frac{K}{N_{sdl}(t)}, 1)$, where $N_{sdl}(t)$ is the current number of seedlings, and K is the carrying capacity.

We analysed drive outcomes across life-history parameters—seed-bank durations, total germination probabilities and fecundity (baseline mean effective ovule and pollen counts, n_{bo} and n_{bp})—and gene drive parameters. Suppression drive efficacy was quantified by genetic load (λ), defined as the proportional reduction in effective gamete production relative to wild-type populations, with the required genetic load (λ^*) denoting the threshold necessary for population elimination. Formal definitions of seed-bank metrics and genetic load are provided in equations (1)–(4).

Effect of a seed bank on modification drives

We found that both CAIN and ClvR were effective modification drives across broad seed-bank regimes: with baseline germination $b \in [0.05, 1]$ and age-dependent germination $m \in [0, 2]$ ($m = 0$ indicates age-independent; $m = 2$ indicates strong age-dependent decline), both drives consistently fixed. Here we define fixation as 100% drive frequency in both plants and seeds to set a conservative upper bound because seed-bank germination can occur long after the management-relevant end-point (that is, a substantial shift in above-ground composition). With $b = 1$ and the default baseline survival rate $d = 1$, all seeds survive and germinate in the first year, eliminating dormancy beyond one year and making m irrelevant. Though fixation was always achieved, the average time to reach fixation from a 10% introduction frequency varied substantially across seed-bank parameters (Fig. 3 and Supplementary Fig. 1) and increased with the average seed-bank duration τ (Supplementary Fig. 2a and equation (2)). Fixation was fastest at high b and m (low τ , germination dominated by younger seeds) and slowest at low b with weak age dependence (high τ , more older germinants). Because τ sets the mean generation time and gene drive spread is inherently a generational process, increasing τ slows drive spread. This result is consistent with previous studies showing that seed banks reduce the rate of selection^{37,38}.

Across all seed-bank parameter combinations under the default drive parameter values, ClvR spread faster than CAIN, and fixation-time differences were more pronounced at larger τ . Supplementary Fig. 3 shows these differences under default and low-fecundity conditions. Changing fecundity did not affect the general patterns: both drives fixed across all germination rates, fixation time increased with τ and ClvR consistently reached fixation faster than CAIN (Supplementary Fig. 4). To identify drive parameters responsible for performance differences between CAIN and ClvR, we varied each parameter independently within assumed efficient ranges and measured fixation probability and mean time to fixation across $b \in [0.1, 1]$ with $m = 0$; additional detailed results are provided in Supplementary Text 2. Both drives fixed reliably, with fixation times largely insensitive to penetrance rates and to germline cleavage rates in either sex (Supplementary Fig. 5a–c). The maternal carryover rate, however, strongly affected ClvR speed: it spread fastest with no carryover but was still able to fix at 100% carryover because of ClvR’s bias through males (Supplementary Fig. 5d). With matched drive and seed-bank parameters, ClvR consistently reached fixation faster than CAIN, indicating that the speed difference primarily reflects ClvR’s ability to bias inheritance through both sexes, whereas CAIN spreads only through males.

Effect of a seed bank on suppression drives

To evaluate suppression drive performance across seed-bank conditions, we varied b and m while holding the drive parameters at the default values. We tracked the proportion of replicates achieving population elimination and the mean time to elimination after a 10% release in heterozygotes. Here, as with the modification drive, we conservatively define elimination as the complete loss of both above-ground plants and the seed bank. Population elimination occurred only when the drive reduced the effective ovule or pollen pool by a critical amount, attaining the required genetic load λ^* (equation (4);

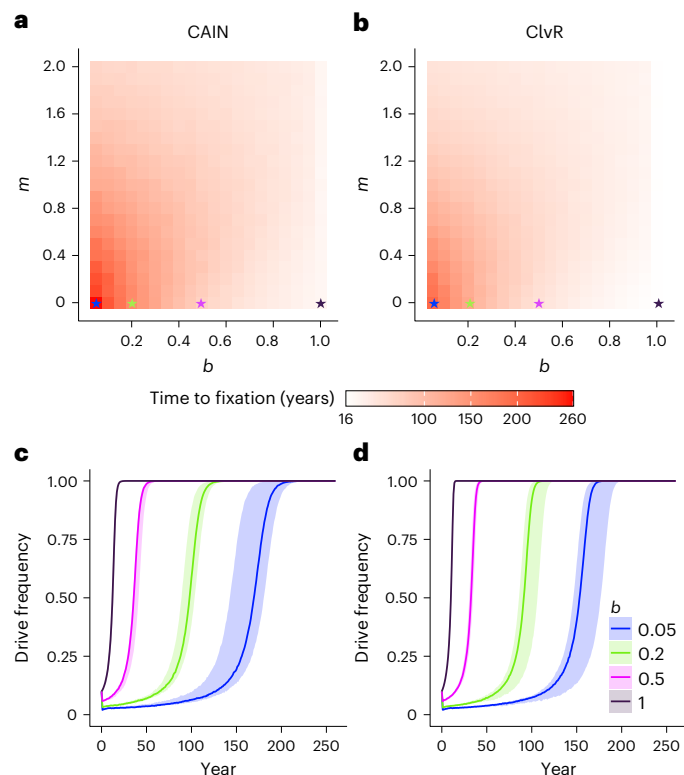


Fig. 3 | Spread of CAIN and ClvR modification drives under varying seed-bank parameters. **a**, Mean fixation time (in years) for the CAIN drive across baseline germination rates (b) and age-dependent germination parameters (m); all other seed-bank parameters were fixed at their default values (Supplementary Table 2). The drive parameters were also set to their experimentally derived values listed in Supplementary Table 1 (see Supplementary Fig. 5 for the joint effect of drive parameters and germination parameters on drive spread). Fixation is defined as 100% drive frequency in both plants and seeds. The colours represent mean fixation times across ten replicates. CAIN achieved fixation under all tested parameter combinations. **b**, Same as **a** but for the ClvR drive, which also achieved fixation under all tested parameter combinations. **c**, CAIN drive frequency trajectories in plants for four baseline germination rates (b) with age-independent germination rates ($m = 0$). The parameter combinations correspond to stars of matching colours in **a**. The solid lines represent median trajectories; the shaded regions indicate the observed range (minimum–maximum) across ten replicates. **d**, Same as **c** but for the ClvR drive, with parameter combinations corresponding to stars of matching colours in **b**.

Fig. 4). This threshold λ^* depends on the total germination probability (γ ; equation (1)) and the mean effective gamete count (n_{bo} under ovule limitation; n_{bp} under pollen limitation). Lower mean effective gamete production reduces the population's low-density growth rate, lowering the required drive strength. Likewise, when γ is low, many seeds fail to reach adulthood, so elimination occurs at a lower genetic load. We defined the genetic load at time t , $\lambda(t)$, as the drive-induced proportional reduction of the effective gamete pool (equation (3)) and evaluated which suppression-drive types and conditions achieved $\lambda(t) > \lambda^*$ for population elimination.

Under the default drive parameter values, the strongest suppression drive was the ClvR male suppression drive, where drive-homozygous males are sterile²⁵. With high cleavage rates and a fully penetrant target gene active in both gametes, it eliminated populations across all germination regimes (Supplementary Figs. 6a and 7a–c). Mean elimination time (Fig. 5a) scaled with mean seed-bank duration (τ ; Supplementary Fig. 2a). Genetic load trajectories are shown in Fig. 5b–d. When there is no seed bank (that is, $b = 1$), ~99% of effective pollen must be removed for the population to decline (Fig. 5d). The ClvR male suppression drive achieves this genetic load, causing

population collapse once the threshold λ^* is reached (at ~33 years). Lowering b (with m fixed) reduces the total germination probability γ (Supplementary Fig. 2b), so fewer seeds emerge from dormancy and λ^* decreases (Fig. 5b,c). Although the drive can always reach this lower λ^* , longer τ delays suppression. Reducing fecundity lowers λ^* across all b and m (Fig. 4), allowing the drive to eliminate the population more quickly with less pollen reduction (Supplementary Figs. 8a and 9a–d).

When ClvR targets a haplosufficient essential female fertility gene, its performance declines. Although cleavage and target penetrance rates remain high, female-mediated inheritance is weakened by maternal carryover: ovules with a disrupted target allele but no drive can remain viable via maternally deposited rescue protein²⁵. Estimated carryover of ~20.7% (Supplementary Text 1) increases the fraction of viable drive-free female gametes, lowering the drive frequency in the effective ovule pool relative to effective pollen. Thus, populations can tolerate greater loss of drive alleles through males than through females, so inducing sterility through drive-homozygous males is more effective than through drive-homozygous females. This asymmetry is evident in the successful suppression parameter regimes (Supplementary Fig. 6).

Unlike the ClvR male suppression drive, the ClvR female suppression drive can only eliminate populations with very low γ (low b , high m). Elimination is faster at higher m (Fig. 5e), since τ decreases. Figure 5h shows λ^* without a seed bank, where genetic load is the relative reduction in effective ovules induced by the drive. In the absence of a seed bank, $\lambda^* \approx 97\%$. However, the ClvR female suppression drive cannot achieve this genetic load; instead, it spreads to a high equilibrium frequency, removing ~90% of effective ovules while still permitting enough seed production for the population to persist (Supplementary Fig. 7d–f). It is not until b is lowered to 0.05 that the ClvR female suppression drive can reach λ^* and eliminate the population around year 345 (Fig. 5f). Here, ~75% of seeds fail to germinate, lowering the required effective ovule pool elimination to ~87%, below the drive's maximum attainable genetic load. Reducing fecundity lowers λ^* across all b and m (Supplementary Fig. 9f–h), enabling population elimination across a broader range of seed-bank parameters (Supplementary Fig. 8b).

The CAIN male suppression drive behaves like the ClvR female suppression drive, lacking sufficient suppressive power to eliminate populations with high γ (Supplementary Fig. 10a). CAIN spreads exclusively through males and targets a gene with incomplete penetrance (96%), so ~4% of pollen carrying disrupted targets remains viable without

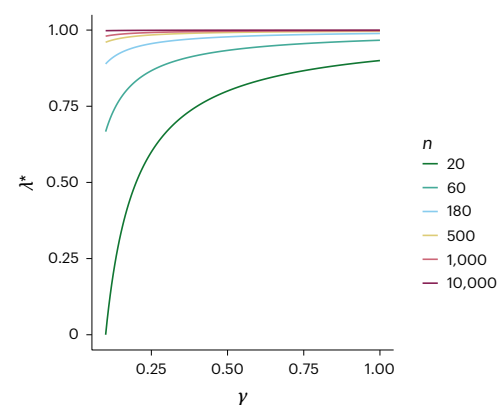


Fig. 4 | Required genetic load for population elimination by a suppression drive. Genetic load is defined as the proportional reduction in the number of effective gametes due to the suppression drive (equation (3)). The required genetic load (λ^*) is a function of the total germination probability (γ) and the effective gamete count (n) of the sex being limited by the drive (equation (4)). For male-targeting drives, n represents the mean number of effective pollen grains per male (n_{bp}), where an effective pollen grain is one capable of reaching a receptive female and germinating. For female-targeting drives, n corresponds to the mean number of effective ovules per female (n_{bo}), where an effective ovule is one that can be fertilized and develop into a seed.

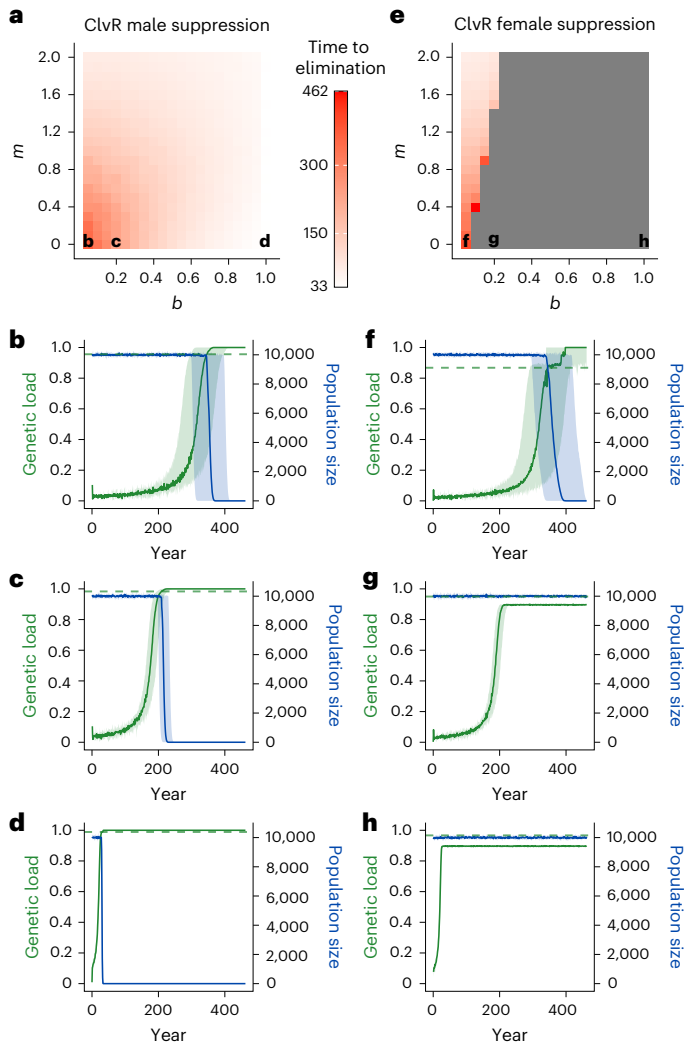


Fig. 5 | Dynamics of ClvR suppression drives under varying seed-bank parameters. **a**, Heat map of the mean time to population elimination for the ClvR male suppression drive across baseline germination rates (b) and age-dependent germination parameters (m); all other parameters were fixed at their default values (Supplementary Tables 1 and 2; see Supplementary Figs. 12–16 for the joint effect of drive parameters and germination rate parameters on drive spread). Population elimination is defined as the complete loss of both seeds and plants due to the drive. The drive succeeded for all tested parameters (Supplementary Fig. 6a). The colours represent mean elimination times across ten replicates; the lettered cells correspond to the subplots below. **b–d**, Trajectories of genetic load (green; left axis) and corresponding population size (blue; right axis) for the ClvR male suppression drive, assuming age-independent germination rates ($m = 0$), with $b = 0.05$ (**b**), $b = 0.2$ (**c**) and $b = 1$ (**d**). Genetic load is defined for this drive as the proportional reduction in total effective pollen grains in the population. The solid lines represent median trajectories; the shaded regions indicate the observed range (minimum–maximum) across ten replicates. The horizontal green dashed lines denote the required genetic load for population elimination (λ^* ; equation (4)). **e**, Heat map as in **a** but for the ClvR female suppression drive. The drive succeeded only for a subset of tested parameter sets (Supplementary Fig. 6b). The colours indicate mean elimination times among successful replicates; the grey cells denote parameter sets where elimination consistently failed. The letter labels in the bottom row indicate parameter sets analysed in the subplots below. **f–h**, Same as **b–d** but for the ClvR female suppression drive, with genetic load defined as the proportional reduction in total effective ovules in the population.

the drive allele. Thus, unlike ClvR, CAIN provides no female-mediated inheritance bias and exhibits slightly weaker male bias. Combined with the loss of drive alleles via sterile drive-homozygous males, elimination is limited to cases where the seed bank removes a large

fraction of seeds (Supplementary Figs. 2b and 10a) or fecundity is low (Supplementary Fig. 10b). Without a seed bank, elimination requires removing ~99% of effective pollen at default fecundity, but the drive’s maximum load is ~96%, resulting in a high equilibrium drive frequency without population decline (Supplementary Fig. 11d). Reducing b to 0.05 lowers λ^* to <96%, enabling population elimination (Supplementary Fig. 11b). Under reduced fecundity, the drive likewise attains λ^* across a broader range of b and m (Supplementary Figs. 10b and 11e–h).

Given a large degree of uncertainty in CAIN and ClvR parameters in target weed populations, we systematically varied each drive parameter and examined population elimination rates across (b, m) (the detailed results are provided in Supplementary Text 2). Outcomes were highly sensitive to the target gene penetrance rate. Drives targeting genes essential in nearly all gametes eliminated populations far more often (Supplementary Fig. 12). Higher germline cleavage rates, especially in the sex for which sterility was induced, also increased population elimination rates (Supplementary Figs. 13 and 14). Lastly, the maternal carryover rate strongly influenced the success of the ClvR female suppression drive (Supplementary Fig. 15). With zero maternal carryover, the drive eliminated populations across all (b, m), often faster than the ClvR male suppression drive (Supplementary Fig. 16), since female-targeting drives have lower λ^* than male-targeting drives (as $n_{bo} < n_{bp}$).

Effects of selfing and seed banks on drive dynamics

To assess gene drive potential in hermaphroditic weeds, we modelled individuals producing both effective pollen and effective ovules, with a constant selfing rate $\sigma \in [0, 1.0]$. We varied b and m while holding the drive parameters at the default values. Mechanistically, selfing reduces outcrossing, thereby limiting the encounters between drive and wild-type alleles and accelerating the formation of drive homozygotes in self-fertilized seeds. A long-lived seed bank (high τ) further slows the drive since older wild-type seeds often germinate in the years immediately following the release.

As modification drives, CAIN and ClvR reached fixation across all seed-bank regimes for $\sigma < 0.8$. With very frequent selfing ($\sigma > 0.8$), however, fixation within 500 years was not guaranteed in long-lived seed banks (Extended Data Fig. 1a,b). For CAIN, when each ovule of a heterozygote is self-fertilized with probability σ and the baseline effective ovule count is n_{bo} , the expected number of self-fertilized ovules is approximately equal to σn_{bo} (all ovules are viable under CAIN). Because most viable pollen grains carry the drive, whereas only half of ovules do, selfing yields approximately $\frac{\sigma n_{bo}}{2}$ drive heterozygotes and $\frac{\sigma n_{bo}}{2}$ drive homozygotes, thereby modestly accelerating drive spread (Extended Data Fig. 1c). At high σ , however, reduced outcrossing limits drive–wild-type matings, lowering heterozygote formation, thereby slowing spread and sometimes preventing fixation. For ClvR, heterozygotes already lose around half of their effective gametes in both sexes due to target disruption without rescue. With selfing, only around $\frac{\sigma n_{bo}}{2}$ ovules are self-fertilized, and because most viable pollen and ovules carry the drive, the resulting seeds are predominantly drive homozygotes. Thus, relative to CAIN, selfing in ClvR produces fewer drive-carrying seeds and does not increase the production of drive heterozygotes, which are more effective in spreading the drive. As a result, ClvR fixation time increases monotonically with σ , and fixation fails at high σ (Extended Data Fig. 1b,d). Together, high σ and high τ further slow fixation.

For suppression drives, selfing increases the required genetic load λ^* and reduces the drive’s maximum attainable load. At $\sigma = 0$, ClvR male suppression eliminated populations across seed-bank regimes (Extended Data Fig. 2a), whereas the ClvR female and CAIN male suppression drives succeeded only at very low γ (Extended Data Fig. 2f,k). Increasing $\sigma \in \{0.03, 0.05, 0.08, 0.10\}$ progressively constrained the elimination region in (b, m) to lower γ for the ClvR female and CAIN

male suppression drives (Extended Data Fig. 2g–j,l–o). The ClvR male suppression drive showed the same qualitative trend but retained the broadest elimination domain, with stochastic elimination persisting even at high σ and high b or m (Extended Data Fig. 2b–e). Mechanistically, selfing raises λ^* for male sterility drives (equation (7)) because fertile plants generate additional selfed seeds that must be suppressed, while simultaneously lowering the maximal achievable load by reducing mating opportunities between drive-carrying pollen and wild-type ovules. Thus, in many cases, the realized genetic load saturates just below λ^* , with occasional stochastic overshoots beyond λ^* leading to elimination. When b or m is high (low τ), the seed bank is shallow. Few older seeds remain to rescue the population, so once the age-0 cohort falls below the demographic replacement threshold, populations crash rapidly.

Effect of a seed bank on threshold-dependent gene drives

So far, we have assumed no fitness costs for CAIN or ClvR, making them effectively zero-threshold drives: despite slow initial spread, each will increase from low frequency provided it avoids early stochastic loss. With fitness costs, however, these drives become threshold-dependent and require introductions above an invasion threshold to spread^{25,27}. We considered two fitness-cost types: (1) gamete-viability cost (s_g), which reduces seed production in carriers and hinders invasion from low frequency; and (2) seed-survival cost (s_s), which lowers the survival and germination of drive-carrying seeds in the seed bank and thus requires a higher initial release frequency.

We first examined the gamete-viability cost ($s_g > 0$) with $s_s = 0$, where s_g is the probability a drive-carrying gamete is non-viable, imposing a codominant fecundity cost; thus, homozygotes produce half as many viable gametes as heterozygotes. Let \hat{p} be the baseline invasion threshold without a seed bank ($b = 1$) and \hat{p}_e be the effective invasion threshold with a seed bank. For simplicity, we considered a single release of the drive (not multi-year) and restricted this analysis to the dioecious model. Because some seeds die or fail to germinate, we expect $\hat{p}_e > \hat{p}$, with the extent of this difference being dependent on seed-bank parameters. To estimate invasion thresholds for a given drive and fitness cost, we followed the protocol outlined in the Methods. To fully explore plant-stage introduction frequencies p_0 (including $p_0 > 0.5$), we introduced the drive by replacing a fraction p_0 of plants with drive homozygotes.

We decreased the baseline germination rate b and estimated \hat{p}_e for each b . The results (Supplementary Figs. 17 and 18) indicate that \hat{p}_e scales roughly with τ : $\hat{p}_e \approx \hat{p} \times \tau$. When τ is low, younger seeds are more likely to germinate than older ones; hence, the drive can be introduced near \hat{p} because many drive seeds germinate soon after release. When τ is high, a longer temporal lag necessitates a higher introduction frequency. In the initial years following drive release, most germinants are pre-drive-release wild types, causing an initial decline in the plant-level drive frequency. As post-release seeds germinate, the fraction of drive-carrying germinated seeds increases. Once the plant-level drive frequency exceeds \hat{p} , the drive is on a trajectory to fixation, expected to rise in frequency each subsequent year. Thus, τ captures the temporal lag, and larger τ requires a higher initial drive frequency to offset the early influx of older wild-type seeds.

A key implication of this relationship is that with long-lived seed banks (high τ) or drives with high baseline invasion thresholds, $\hat{p} \times \tau$ can exceed 1. In such cases, even a 100% plant-level release cannot prevent loss (for example, Supplementary Figs. 17k,l and 18k,l). Post-release germination is dominated by older wild-type seeds, pulling the drive frequency in plants below \hat{p} . Recovery requires later-germinated drive seeds to lift the drive frequency in plants above \hat{p} ; however, when $\hat{p} \times \tau > 1$, seed-bank dilution makes this impossible, and the drive is ultimately lost. Figure 6 compares predicted ($\hat{p}_e \approx \hat{p} \times \tau$) and observed \hat{p}_e for both CAIN and ClvR modification drives. The fit is good at lower s_g but slightly underestimates thresholds

of the highest s_g explored. The same pattern also holds for suppression drives (Supplementary Fig. 19). For male-sterility drives, we introduced the drive by replacing a fraction p_0 of the population with drive-homozygous females; for female-sterility drives, we used drive-homozygous males. With an equal sex ratio, the maximum attainable p_0 in either case was 0.5. We found that if $\hat{p} \times \tau < 0.5$, the prediction holds, and if $\hat{p} \times \tau > 0.5$, the suppression drive cannot spread.

We next examined a seed-viability cost ($s_s > 0$) with $s_g = 0$, modelling s_s as a codominant reduction of the baseline seed survival rate d . A seed of age a thus survives at rate d/a^q if wild-type, $(d - s_s/2)/a^q$ if drive-heterozygous or $(d - s_s)/a^q$ if drive-homozygous, where q controls the age-dependent decline in d . Importantly, $s_s > 0$ results in genotype-dependent variation in τ and γ . As d decreases, τ also decreases because older seeds are more likely to have died, shifting the distribution of germinated seeds towards younger ages. Similarly, γ declines with d due to reduced survival rates across all seed-age classes. Given these complex genotype-dependent shifts in τ and γ , the simple scaling $\hat{p}_e \approx \hat{p} \times \tau$ did not hold. Nonetheless, we still observed a consistent positive correlation between \hat{p}_e and τ , along with drive failure at high \hat{p} and τ combinations (Supplementary Fig. 20). However, we expect the precise relationship among \hat{p}_e , \hat{p} and seed-bank parameters to be more complex when drive fitness costs manifest in seeds rather than in plants.

Drive release strategy and seed bank

A drive can be released at either the plant or the seed stage. Our baseline plant-stage release directly sets the initial drive allele frequency p_0 in plants while avoiding transient changes in total seed output, providing a standardized and interpretable reference point to compare drive dynamics across parameter regimes. However, plant releases may be operationally impractical. Instead, a seed drop introduction may often better match deployment logistics (seed handling/dispersal) but introduces a germination lag before drive-carrying plants emerge. We therefore analysed both strategies and derived analytical and numerical mappings to establish their equivalence. Although our analyses focused on the CAIN and ClvR modification drives, these results are expected to generalize to suppression drives as well.

In our baseline plant replacement strategy, a fraction of wild-type plants were replaced with drive heterozygotes to set the initial plant-level drive allele frequency $p_0 = 0.1$. Alternatively, we modelled a seed drop of age-0 drive-heterozygous seeds, varying the seed introduction frequency p_s (drive frequency in age-0 seeds post-release) and baseline germination rate b (with no age dependence in germination; $m = 0$), and recorded the resulting p_0 in plants (after seed-bank survival and germination), fixation probability and time to fixation. The observed p_0 matched the analytical predictions (equation (11)); achieving $p_0 = 0.1$ required $p_s \gg 0.1$ at low b (where most drive seeds remain in the seed bank) but $p_s \approx 0.1$ as $b \rightarrow 1$ (Extended Data Fig. 3a,b). Since CAIN and ClvR are both zero-threshold drives at the default drive parameters, once the initial frequency exceeded the stochastic loss barrier, both spread to fixation regardless of p_s or seed-bank parameters. For fixed b (and $m = 0$), the mean fixation time under a seed drop aligned best with the baseline plant replacement results (Fig. 3a,b) when $p_s \approx 0.1$ (Extended Data Fig. 3c–f). Thus, with matched introduction frequency, drive outcomes are similar across introduction modes (plant replacement versus seed drop), provided the release enables drive establishment.

We assessed whether the empirical relationship $\hat{p}_e \approx \hat{p} \times \tau$ holds when $s_g > 0$ under seed drop. Because seed drop temporarily increases total seed output in the release year, we also analysed a constant-output control, ‘seed replacement’, in which a fraction p_s of release-year age-0 wild-type seeds was replaced with drive homozygotes without changing the total seed output. Seed replacement served solely as a control to isolate the effect of seed-bank dilution and does not imply operational replacement of older, pre-existing seed-bank cohorts. We found

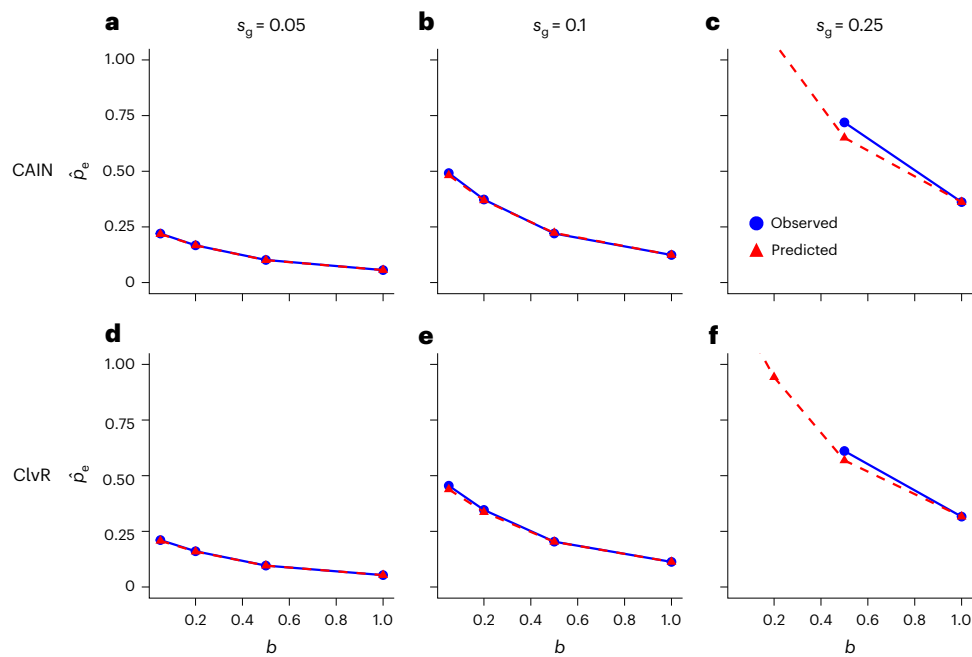


Fig. 6 | Effective invasion thresholds of CAIN and ClvR drives under gamete viability fitness costs. a–f, Baseline germination rate b and drive-associated gamete viability cost s_g were varied, with all other parameters fixed at their default values (Supplementary Table 2) and ten replicates run per parameter combination. The blue points represent observed effective invasion thresholds (\hat{p}_e), defined as the minimum introduction frequency resulting in successful drive spread in more than 50% of simulation replicates (Supplementary Figs. 17

and 18). The red triangles indicate thresholds predicted by $\hat{p}_e \approx \hat{p} \times \tau$. The points at $b = 1$ correspond to baseline invasion thresholds (\hat{p}) in populations without seed banks. Panels a–c show CAIN effective invasion thresholds for $s_g = 0.05, 0.1, 0.25$ and $b = 0.05, 0.2, 0.5, 1$. The absence of blue points at these b values indicates parameter sets where invasion consistently failed, and the absence of red triangles indicates predicted \hat{p}_e exceeding 1. Panels d–f show ClvR effective invasion thresholds, presented as in a–c.

that for the seed drop, observed \hat{p}_e values were below the $\hat{p} \times \tau$ prediction (Extended Data Fig. 4a–c,g–i), but for the seed replacement, the $\hat{p}_e \approx \hat{p} \times \tau$ relationship held (Extended Data Fig. 4d–f,j–l). This is because when drive seeds are dropped into the population, early germinants disproportionately originate from the release year, weakening seed-bank dilution and enabling drive spread even when the initial frequencies are below $\hat{p} \times \tau$, but when seeds are replaced, the realized germination rates remain the same as before the release.

Discussion

In this study, we developed a comprehensive gene drive modelling framework that explicitly integrates plant-specific life-history traits, including seed dormancy, and investigated the eco-evolutionary dynamics of the CAIN and ClvR drives under realistic ecological conditions. Our results reveal the critical roles of seed banks in modulating gene drive efficacy, suppression potential and confinement in plant populations.

In our model with the default drive parameters, the ClvR male suppression drive consistently outperformed both the ClvR female suppression and CAIN male suppression drives, with the latter two achieving population elimination only under sufficiently low population fecundity or total germination probability (γ) (Supplementary Figs. 6, 8 and 10). These results contrast with those of Liu et al.²⁴ and Oberhofer et al.²⁵, who reported successful population elimination by ClvR female suppression and CAIN male suppression, respectively, in models without dormancy. We attribute this discrepancy to the lower low-density growth rates (β) assumed in their studies than in ours. Indeed, decreasing β to their default values reduces the required genetic load, enabling even weaker constructs (ClvR female suppression and CAIN male suppression) to reach these thresholds and collapse the population. Our model employs relatively high β values because we assume complete fertilization of ovules when effective ovules are limiting and pollen is abundant, and complete utilization of pollen

when effective pollen is limiting and ovules are abundant. Consequently, when we evaluate our model without dormancy, all seeds germinate immediately, and at low densities below carrying capacity, all seedlings develop into plants. However, natural populations often experience seedling mortality due to external biotic and abiotic factors—such as interspecific competition, predation, fungal pathogens and physical disturbances—independent of seedling density^{39–41}. Such factors would decrease β below our modelled values, thereby lowering the required genetic load of the drive and potentially allowing weaker drives to achieve population elimination. Our model thus provides a conservative estimate of suppression drive efficacy, and the formula for required genetic load (equation (4)) may need adjustment to incorporate additional population-specific constraints on seedling establishment and survival. However, our gene drive parameter analysis suggests that suppression drives may be able to impose the highest genetic loads when the drive's target gene is essential for the viability of all gametes (Supplementary Fig. 12) and germline cleavage rates are high, particularly in the sex for which the drive induces sterility (Supplementary Figs. 13 and 14).

Our results highlight three life-history metrics critical to the effective deployment of gene drives in weed populations: (1) average seed-bank duration τ , (2) total germination probability γ and (3) fecundity (baseline mean effective ovule count n_{bo} and mean effective pollen grain count n_{bp} per individual). If a candidate population has a long average seed-bank duration under natural conditions and controlled experiments without seed banks indicate a high baseline invasion threshold (\hat{p}), the effective invasion threshold (\hat{p}_e) may exceed practical single-release frequencies (for example, Fig. 6c,f). In such cases, multiple consecutive releases could shorten time to fixation (or elimination) by accelerating establishment and reducing seed-bank dilution, and may also improve confinement within the target population. Furthermore, pre-deployment efforts that reduce the resident wild-type seed bank can increase the effective introduction frequency, thereby

lowering β_e : for example, seed-bank-depletion measures that draw down the seed bank and limit seed rain^{42,43} and/or a large drive-seed drop (Extended Data Fig. 4). Even when establishment is feasible, long-lived seed banks can create a demographic lag that makes suppression by the drive alone impractically slow. In such cases, gene drives are best considered as a part of integrated weed management that also reduces ongoing seed production and accelerates seed-bank depletion⁴³, consistent with ‘sensitizing’ drive concepts intended to improve the efficacy of subsequent interventions¹⁵. Our model is also conservative in its demographic response to genetic load: population size is not substantially reduced until genetic load becomes sufficiently high, after which the population is rapidly eliminated. Real populations may begin to decline much earlier at intermediate genetic loads, depending on the form of density-dependent regulation in the target species. This could yield management-relevant control effects sooner and may also expedite population elimination through Allee effects and demographic stochasticity.

For modification drives, our results suggest that fecundity has minimal impact on drive success; however, long seed-bank durations can substantially delay drive fixation (Fig. 3 and Supplementary Fig. 4). For suppression drives, we found that population elimination is facilitated by reduced fecundity and lower total germination probability, as both decrease the required genetic load of the drive (Fig. 4). Low-fecundity populations also benefit from smaller population sizes, thus requiring fewer drive individuals to achieve necessary release frequencies. Lastly, because selfing can slow or even prevent drive spread (Extended Data Figs. 1 and 2), we recommend prioritizing dioecious species (for example, *Amaranthus palmeri* and *Amaranthus tuberculatus*^{15,44,45}) or strongly self-incompatible species (for example, *Lolium rigidum* and *Alopecurus myosuroides*^{15,46,47}). Our analyses assumed a constant selfing rate, but previous studies indicate that selfing can evolve as a mechanism to avoid drive-imposed fitness costs^{32,48–50}. Thus, for target species capable of selfing with low inbreeding depression, we caution against deploying suppression drives or modification drives with strong fitness costs.

Several limitations of our model present opportunities for further investigation. We assumed a panmictic population, an annual life cycle, constant seed-bank parameters, wind pollination and no spatial dynamics. These assumptions underestimate the complexity of natural populations, where genetic variation, environmental fluctuations and spatial heterogeneity often influence seed survival, germination, pollen dispersal and drive dynamics. We also did not model functional resistance alleles that prevent drive cleavage but retain target gene function. Previous theoretical work shows that if functional resistance alleles confer fitness advantages against the drive, they can eventually outcompete the drive allele^{25,50,51}. However, we expect functional resistance rates to be low for CAIN and ClvR due to their use of multiple gRNAs, increasing target cleavage redundancy^{52–55}. Although resistance alleles may eventually arise in sufficiently large populations, their establishment probability would probably be reduced in plant populations with low germination rates. Another phenomenon known to hinder the success of suppression drives is ‘chasing’, a recolonization–extinction dynamic in spatially structured populations in which wild-type individuals repopulate areas previously cleared by the drive, often preventing population elimination^{50,56,57}. Seed banks may considerably influence chasing dynamics, especially when coupled with other plant-specific life-history traits, such as long-distance dispersal of pollen and seeds and reduced low-density growth rates of populations with low germination rates. Such interactions could be complex and warrant further investigation.

Although CRISPR-based gene drive systems have been studied extensively for controlling disease-vector mosquitoes^{58–62}, their potential for agricultural weed control has remained largely unexplored. One reason for this gap has been the lack of experimentally validated gene drives in plants until the recent development of CAIN

and ClvR^{24,25}. Another is the assumption that key plant life-history traits (for example, high fecundity, self-pollination, polyploidy and seed dormancy) constrain drive efficacy^{15,16}. Our findings challenge this assumption by demonstrating that seed banks can, in some cases, enhance both the efficacy and biosafety of gene drives in plant systems. Specifically, seed banks can lower the required genetic loads for suppression drives and introduce higher invasion thresholds under drive-associated fitness costs, substantially reducing the risks of unintended gene drive spillover into non-target populations. Highly efficient gene drive constructs can thus become inherently safer in populations with seed dormancy. Achieving comparable thresholds in populations without seed banks would require either complex gene drive architectures, which are often challenging to engineer in practice (for example, underdominance drives^{63–66}), or substantially increased fitness costs of toxin–antidote drives, with reduced drive efficacy. Collectively, our work highlights seed dormancy not as a barrier but as an ecological trait that can be strategically leveraged for gene drive deployment, providing opportunities to optimize the feasibility, effectiveness and safety of genetic interventions in weed populations. More broadly, our study underscores the importance of explicitly integrating ecological and life-history complexity into genetic control strategies to enable safer and more targeted biotechnological solutions.

Methods

Gene drive systems

We parameterized our default CAIN and ClvR gene drives using experimentally derived values from their respective experimental studies^{24,25}. Figure 1 illustrates the mechanism of each drive, and Supplementary Table 1 summarizes the corresponding model parameters used in our simulations unless stated otherwise. We varied the drive parameters from these experimental defaults and studied the effects of the drive parameters (see Supplementary Text 2 for further details).

We modelled the germline cleavage rate as the probability that a wild-type target allele—located on a separate chromosome from the drive—is cleaved in the germline of an individual carrying at least one drive allele. This rate may differ depending on the sex of the individual. We assumed the drive carries multiple gRNAs that target conserved sites within the target gene, resulting in an extremely low functional resistance rate⁵⁴; that is, the probability of a target allele becoming uncleavable at every gRNA target site while preserving the gene’s function is minimal. We thus did not model functional resistance; all cleavage events were assumed to produce a non-functional (disrupted) copy of the gene. We also excluded maternal or paternal Cas9 deposition (that is, embryo cleavage) and intrinsic drive-associated fitness costs by default, as these were not reported or explored in the referenced studies. However, we also considered drive fitness costs reducing gamete viability and seed survival when evaluating the effects of seed-bank parameters on threshold-dependent drives.

The CAIN study²⁴ reported germline cleavage rates of 0.984 in males and 0.941 in females. The authors also observed a 96% penetrance rate of the *NPGI* target gene, which we modelled as a 4% probability that a pollen grain lacking a functional copy of the gene remains viable. Because *NPGI* affects only male gametes, ovules with disrupted target alleles are assumed to remain viable. For a modification version of CAIN, we assumed that a cargo gene is attached to the drive construct, such that drive carriers have a desirable phenotype. We refer to the modification construct simply as ‘CAIN’. For suppression, we assumed that the CAIN drive is inserted into a haplosufficient essential male fertility gene in a manner that disrupts the gene, such that males with two drive alleles are sterile. We refer to this construct as ‘CAIN male suppression’.

For ClvR²⁵, exact estimates for male and female cleavage rates are unavailable. We inferred these values from the male and female drive inheritance rates, averaged across experimental crosses and

weighted by sample sizes (Supplementary Text 1). We estimated a male germline cleavage rate of 0.974. In females, drive inheritance rates were lower than in males, a difference attributed to maternal carryover—the partial deposition of the rescue protein (encoded by the drive) into female gametes. Assuming female cleavage occurs at the same rate as in males, we estimated the maternal carryover rate to be approximately 0.207. We modelled this as a 20.7% probability that a female gamete lacking a functional *YKT61* allele remains viable when produced by a drive-carrying female. For population modification, we assumed that a cargo gene is attached to the ClvR drive, such that individuals with a drive allele have some desirable phenotype. We refer to the modification version of the drive as ‘ClvR’. For population suppression, we considered two variants of ClvR: ‘ClvR male suppression’, in which the drive is inserted into a haplosufficient, essential male fertility gene in a disruptive manner such that male homozygotes are sterile, and ‘ClvR female suppression’, in which the drive is inserted into a haplosufficient, essential female fertility gene in a disruptive manner such that female homozygotes are sterile.

Life cycle model of annual plants

Our model is based on wind-pollinated weeds, a major group of potential gene drive targets¹⁵. While some species can be monoecious or hermaphroditic and self-fertilize to varying degrees, we focused on a strictly dioecious system to avoid the potential evolution of selfing as a protective mechanism against gene-drive-induced fitness costs^{48,49}. An extension for hermaphroditic species capable of selfing is detailed in a later Methods section. We also assume a panmictic population with an annual life cycle (Fig. 2). The model employs a stochastic, agent-based framework that explicitly accounts for the plant life cycle and incorporates the CAIN and ClvR drive systems. Our model extends beyond the deterministic homing drive framework of Barrett et al.¹⁶, allowing for a more biologically and ecologically realistic representation of drive dynamics.

Reproduction. In our model, we do not explicitly track all ovules or pollen grains produced in the population; instead, we focus on effective gametes that actually contribute to reproduction (Fig. 2a). We define ‘effective ovules’ as ovules capable of fertilization and subsequent development into viable seeds and ‘effective pollen’ as pollen grains that can reach a receptive female by wind and successfully germinate. Because the probability of pollen deposition on a receptive stigma is extremely low in wind-pollinated systems⁶⁷, the expected effective pollen count is substantially lower than the total pollen production per male. This count varies among males due to differences in pollen production and transfer efficiency^{67–69}. Analogously, the effective ovule count varies among females because of resource allocation, developmental anomalies and other factors affecting ovule viability^{70–73}. We account for this individual-level variation by sampling the number of effective gametes per individual from a Poisson distribution with sex-specific means.

We assume that pollen deposition is uniformly distributed among a female’s flowers, such that only the total number of effective ovules per female determines reproductive output. Let n_{bo} denote the baseline mean effective ovule count per wild-type female, which was held constant throughout the simulation. We draw each fertile female’s maximum number of effective ovules from a Poisson distribution with mean n_{bo} . We then consider drive processes in the germline that created these gametes and whether any such gametes would have been rendered non-viable due to the drive. For the CAIN drive—whose target gene affects only male gametes—all ovules are viable, so a female’s effective ovule count will always equal her maximum effective ovule count drawn from the Poisson distribution. This is also the case for the ClvR drive when females lack a drive allele and a disrupted target allele. However, if a female carries a ClvR drive allele or disrupted target allele, some ovules will be non-viable if they inherit a disrupted target allele with

no rescue (through co-inheriting the drive or through maternal carryover), thereby reducing her effective ovule count below her maximum number drawn. For suppression drives, sterile females are assumed to produce no effective ovules and are excluded from reproduction.

We define n_{bp} as the baseline mean effective pollen count per male in a wild-type population at carrying capacity K . The probability that a pollen grain successfully reaches a receptive stigma at time t is assumed to be proportional to the number of fertile females $N_f(t)$, with all stigmas equally likely to capture pollen. Accordingly, as $N_f(t)$ increases, the per-pollen grain probability of successful deposition increases proportionally.

To model this dependency, we scale n_{bp} by the ratio of $N_f(t)$ to the expected number of fertile females at carrying capacity, $K/2$, yielding the time-dependent mean number of effective pollen grains per male, $n_p(t)$:

$$n_p(t) = n_{bp} \times \frac{N_f(t)}{K/2}.$$

At carrying capacity, $n_p(t) = n_{bp}$; however, as the number of fertile females declines, a greater fraction of pollen grains fails to reach receptive females, thereby reducing the effective pollen count. This scaling is especially relevant for suppression drives that reduce the number of fertile females. A similar procedure is conducted in males as in females: for each fertile male, we first sample the maximum effective pollen count from a Poisson distribution with mean $n_p(t)$, then model drive processes in the germline that produced these pollen grains and lastly discard any pollen grains that were rendered non-viable due to the drive. Because of drive-associated loss of gamete viability, males carrying a CAIN or ClvR drive allele or disrupted target allele ultimately produce fewer effective pollen grains than wild-type males. For suppression drives targeting male fertility genes, sterile males are assumed to produce no effective pollen and are excluded from reproduction.

Effective pollen grains are then collected from the $N_m(t)$ fertile males and randomly allocated to the $N_f(t)$ fertile females. The total number of effective pollen ($N_p^{tot}(t)$) has expectation $N_m(t) \times n_p(t)$. We model pollen allocation as a multinomial random variable with $N_p^{tot}(t)$ trials and uniform category probabilities ($\frac{1}{N_f(t)}, \dots, \frac{1}{N_f(t)}$), assuming each fertile female is equally likely to receive a pollen grain. The expected number of pollen grains per female is $N_p^{tot}(t)/N_f(t) = \frac{N_m(t) \times n_{bp}}{K/2}$, though stochastic variation may lead to deviations from this mean.

Fertile females typically receive enough effective pollen to fertilize all their effective ovules, in which case each effective ovule randomly samples an effective pollen grain from the female’s effective pollen pool, assuming all effective pollen grains are equally likely to be chosen. However, if a suppression drive reduces pollen availability (for example, CAIN male suppression or the ClvR male suppression drive), some females may not receive enough effective pollen to fertilize all effective ovules. In such a case, each effective pollen grain fertilizes a randomly chosen effective ovule from the female, assuming all effective ovules have an equal probability of being fertilized. Fertilized effective ovules become seeds, which have an equal probability of being male or female. Female plants that receive no effective pollen produce no seeds. Similarly, if a suppression drive reduces the number of effective ovules, fewer fertilization events occur, thereby reducing seed production.

Movement out of the seed bank via age-dependent seed mortality and germination. Newly produced seeds enter the seed bank and survive their first year with probability d (Fig. 2b,c). Thereafter, survival rates decline with seed age, such that the probability of an age- a seed surviving, given its persistence in the seed bank, is d/a^q , where the parameter q determines the rate at which survival rates decrease with age. Surviving seeds have an opportunity to germinate each year, with age-dependent germination probabilities modelled similarly to survival probabilities. Newly produced seeds germinate with probability b , conditional on surviving their first year. For seeds of age $a > 1$,

germination rates decline with age, such that the probability of an age- a seed germinating, given its persistence in the seed bank, is b/a^m , where m determines the rate at which germination rates decrease with age. Germination rates are assumed to be genotype-independent. Seeds can persist in the seed bank for a maximum of L years^{74,75}, after which they are assumed to be non-viable and are removed from the model. To simulate populations without seed banks, we set $d = 1$ and $b = 1$, such that all seeds survive and immediately germinate. In such a case, seed age never exceeds one year, making the parameters m , q and L have no effect.

Seedling competition. Germinated seeds develop into seedlings that compete for limited resources (Fig. 2d)^{41,76–78}. We assume that seedling competition regulates the population size, maintaining the number of plants near the carrying capacity K at equilibrium. We denote the number of seedlings at year t as $N_{\text{sdl}}(t)$. When $N_{\text{sdl}}(t) \leq K$, all seedlings survive to adulthood. When $N_{\text{sdl}}(t) > K$, density-dependent competition reduces survival probability proportionally. The probability of a seedling surviving to adulthood at time t is given by:

$$c(t) = \min\left(\frac{K}{N_{\text{sdl}}(t)}, 1\right).$$

Surviving seedlings mature into adult plants, completing the life cycle. Each model time step corresponds to one year.

Average seed-bank duration and total germination probability

To obtain the average seed-bank duration (τ), we first derive the probability of germinating at age a (g_a) for seeds of age $a \in \{1, \dots, L\}$, where L represents the maximum viable seed age, following previous approaches^{30,37,79}. For a seed to germinate at age a , it must (1) remain dormant through the preceding $a-1$ years without dying or germinating and (2) both survive and germinate in year a . Thus, g_a is given by

$$g_a = \frac{d}{a^q} \frac{b}{a^m} \prod_{k=1}^{a-1} \left[\frac{d}{k^q} \left(1 - \frac{b}{k^m}\right) \right],$$

where d and b are the baseline survival and germination rates, and q and m are shape parameters governing the age-dependent decline in survival and germination, respectively, as defined in the previous section.

We define the total germination probability γ as the cumulative probability of germination across all seed ages up to L :

$$\gamma = \sum_{a=1}^L g_a. \quad (1)$$

With this, we derive τ , equivalent to both the expected age of germinated seeds and the generation time of the population:

$$\tau = \frac{1}{\gamma} \sum_{a=1}^L (a \times g_a). \quad (2)$$

Implementation of gene drive simulation model with a seed bank

We implemented our model in SLiM v.4.0.1 (ref. 36), an individual-based, forward-time population genetic simulation framework. We initialized each simulation with K wild-type plants and allowed the plant and seed populations to equilibrate for $2L$ years (where L represents the maximum seed age) before introducing the drive. This burn-in period was sufficient for the empirical average seed-bank duration to match its analytical expectation derived in the previous section. We then introduced the drive at a frequency of p_0 by replacing a fraction $2p_0$ of plants with drive heterozygotes. The drive was introduced only once per simulation. The simulation was run until one of the following outcomes occurred: drive fixation, drive loss, population suppression or 500 years elapsed after drive introduction.

Low-density growth rate

For modification drives, we evaluated the probability of fixation and time to fixation. For suppression drives, we evaluated the probability and timing of population elimination. In the latter case, the efficacy of suppression depends largely on the density-dependent dynamics of the population^{80,81}. As the suppression drive spreads, gamete production declines, but the few remaining gametes face reduced competition and an increased probability of successful fertilization. If the maximum per capita growth rate at low density, denoted β , is sufficiently high, population elimination may be prevented. To derive β for our suppression constructs, we first note that maintaining the population at carrying capacity K requires, on average, each female to produce at least two fertilized effective ovules (assuming abundant effective pollen grains) or each male to pollinate two effective ovules (assuming abundant effective ovules). These fertilized ovules become seeds, which then must survive the seed bank and germinate. The minimum required seed production per male or female is therefore $n_{\text{min}} = 2/\gamma$, where γ is the total germination probability, as defined in equation (1).

For suppression drives that generate sterile males, the male low-density growth rate β_m can be derived by considering the limiting case in which a single fertile male remains, while the number of fertile females remains at $K/2$. This male produces, on average, n_{bp} viable pollen grains. Given an excess of effective ovules and no competition from other males, all effective pollen is assumed to fertilize effective ovules successfully, producing n_{bp} seeds per male at low density. The male low-density growth rate is thus $\beta_m = n_{\text{bp}}/n_{\text{min}} = n_{\text{bp}}\gamma/2$. Alternatively, if the suppression drive generates sterile females, the female low-density growth rate β_f is derived by considering the limiting case in which a single fertile female remains, while the number of fertile males is $K/2$. This female produces n_{bo} effective ovules on average. With abundant effective pollen and no competition among females, all of her effective ovules are successfully fertilized, resulting in n_{bo} seeds produced per female. The corresponding female low-density growth rate is $\beta_f = n_{\text{bo}}/n_{\text{min}} = n_{\text{bo}}\gamma/2$.

Required genetic load

We focus on the low-density growth rate (β) that the suppression drive must overcome to achieve suppression. Let n denote the baseline number of effective gametes per wild-type individual. Because the parameters determining reproductive success differ according to which sex is rendered sterile by the drive, we set these parameters sex-specifically: for drives inducing sterile males, we set $\beta = \beta_m$ and $n = n_{\text{bp}}$, and for drives inducing sterile females, we set $\beta = \beta_f$ and $n = n_{\text{bo}}$. The suppressive effect of the drive can be quantified by its genetic load, which measures the reduction in population fitness induced by the drive^{80,82–85}. We measure population fitness by the total number of effective gametes produced in a given year, relative to the number expected in a wild-type population of the same size. We define the genetic load $\lambda(t)$ at time t as:

$$\begin{aligned} \lambda(t) &= 1 - \frac{\text{total number of effective gametes}(t)}{\text{expected total number of effective gametes}} \\ &= 1 - \frac{\text{total number of effective gametes}(t)}{nK/2}, \end{aligned} \quad (3)$$

where $nK/2$ represents the expected number of effective gametes, assuming the population is at carrying capacity.

For the drive to suppress the population, it must reduce the effective gamete pool below the threshold necessary to produce n_{min} seeds per individual. The critical genetic load required for population decline is at least $1 - 1/\beta$ (ref. 83), which we define as the required genetic load λ^* :

$$\lambda^* = 1 - \frac{1}{\beta} = 1 - \frac{2}{n\gamma}. \quad (4)$$

Fitness cost of gene drive

Our base model assumes no fitness cost associated with CAIN or ClvR. To investigate the impact of drive-associated fitness costs, we extended the model to incorporate drive-induced fitness costs on gamete viability and seed survival. We define s_g as the probability that a drive-carrying gamete is non-viable, effectively representing a codominant fitness cost reducing gamete production. Thus, $s_g/2$ of drive-heterozygote gametes and s_g of drive-homozygote gametes are non-viable in expectation. Similarly, we define s_s as a codominant fitness cost reducing seed survival, modelled as a reduction in the baseline survival rate (d). Drive-heterozygous seeds have a baseline survival rate of $d - s_s/2$, while drive-homozygous seeds have $d - s_s$. Seed survival rates also decline with seed age (a), following the general form d/a^q for wild types, $(d - s_s/2)/a^q$ for drive heterozygotes and $(d - s_s)/a^q$ for drive homozygotes, where q controls the rate at which survival probabilities decline with age.

We examined each fitness cost separately to quantify its impact on the effective invasion threshold (\hat{p}_c) of the drive. For a given set of seed-bank and drive parameters, we varied the introduction frequency (p_0), ran ten replicates per p_0 and recorded the proportion of runs in which the final drive frequency in plants exceeded p_0 . We then fit a logistic response of spread rate to p_0 using the `drm` R package⁸⁶ and defined \hat{p}_c as the p_0 at which the fitted curve equals 0.5 (that is, spread in half of replicates).

Gene drive parameter scan

Since CAIN and ClvR drive parameters may differ from experimental estimates in target weed populations, we systematically analysed how each drive parameter affects each drive's performance across a range of seed-bank parameters. We assumed efficient drives and, unless varied, set the target-gene penetrance to 1 (essential target), the male germline cleavage rate to 0.97, the female germline cleavage rate to 0.97 and the maternal carryover rate (for ClvR) to 0.207. We imposed no drive fitness costs ($s_g = s_s = 0$) and released the drive at $p_0 = 0.1$ frequency in drive-heterozygous plants. We ran ten replicates per parameter combination.

To analyse the effect of the target gene penetrance rate, we set it to 0.96, 0.98 or 1 and recorded (1) the fraction of replicates achieving fixation (modification drives) or population elimination (suppression drives) within 500 years and (2) the average time to fixation or population elimination among successful replicates. We define fixation as the drive reaching 100% frequency in all plants and seeds, and elimination as the drive causing the loss of all plants and seeds. We repeated these numerical experiments over baseline germination rates (b in [0.1, 1] with step 0.1) and age-dependence germination parameters (m in [0, 2] with step 0.2) for all CAIN and ClvR modification drives and suppression drives, with ten replicates per parameter combination. All other seed-bank parameters were set to their default values in Supplementary Table 2.

We applied the same procedure to the remaining drive parameters but with parameter-specific ranges. We varied the male and female germline cleavage rates over {0.9, 0.95, 1}, holding the opposite-sex rate at the default, and varied the maternal carryover rate (ClvR only) over {0, 0.207, 0.5, 1}. For each drive and parameter combination, we recorded the proportion of replicates achieving fixation (modification drive) or population elimination (suppression drive) within 500 years, and the mean time to fixation or elimination, conditional on success.

Hermaphrodite model with selfing

To examine the effect of self-pollination on the spread of modification and suppression drives, we constructed a hermaphroditic model in which each individual produces an average of n_{bo} effective ovules and n_{bp} effective pollen grains at equilibrium. Additional model parameters are listed in Supplementary Table 3. To match the total effective gamete counts to those in the dioecious model, we set the average effective

gamete counts to half of their previous defaults: $n_{bo} = 30$ and $n_{bp} = 90$. Seed-bank dynamics remained the same across models, and each drive was parameterized using its experimentally derived default values (Supplementary Table 1).

In the hermaphrodite model, selfing occurs at a constant rate $\sigma \in [0, 1]$ specified at initialization. An effective ovule is fertilized by self pollen with probability σ , independently across ovules. We assume that self pollen is not limited; that is, an effective ovule samples from the plant's own pollen pool with replacement. After accounting for selfing, all remaining receptive ovules and outcross pollen are pooled and paired at random (panmictic outcrossing), as in the dioecious model. However, to ensure that self-pollination does not occur at a higher rate than intended, we remove an individual's own pollen from its outcross pool.

CAIN and ClvR suppression drives act analogously to the dioecious model. For the CAIN male and ClvR male suppression drives, drive homozygotes produce no effective pollen (male sterile), only ovules. For the ClvR female suppression drive, drive homozygotes produce no effective ovules (female sterile) and lack functional female reproductive structures; hence, they cannot receive pollen. Consequently, any drive homozygote lacking one gamete class cannot self-pollinate.

Since we define 'effective pollen' as pollen grains that reach receptive ovules and germinate, the mean effective pollen available for outcrossing at time t is reduced by individuals lacking functional female reproductive structures (that is, homozygotes for the ClvR female suppression drive) and ovules already fertilized via selfing. Let $N_{\text{recv}}(t)$ be the number of plants capable of receiving pollen (the total population minus ClvR-female homozygotes) and $N_{\text{self}}(t) (\leq N_{\text{recv}}(t))$ be the subset of receivers that are capable of selfing (the total population minus all suppression-drive homozygotes).

Following the classic mass-action "pollen-pool" mating framework⁸⁷, outcrossing arises from random pairing between a population-level pool of effective pollen and a population-level pool of receptive ovules. We treat selfing as a per-ovule process that removes a fraction of ovules from the outcross pool on plants capable of selfing^{88,89}. A pollen grain succeeds in outcrossing only if it (1) lands on a pollen-receiving plant and (2) encounters an ovule that has not been pre-empted by selfing. The first event occurs with probability $N_{\text{recv}}(t)/K$, where K denotes the population's carrying capacity. Among receivers, the fraction that can self is $N_{\text{self}}(t)/N_{\text{recv}}(t)$; on those plants, a fraction $1 - \sigma$ of ovules remain available to outcross pollen, whereas receivers that cannot self have all ovules available. The mean fraction of ovules on receivers available to outcross pollen is thus $1 - \sigma N_{\text{self}}(t)/N_{\text{recv}}(t)$. Multiplying these factors by the baseline mean effective pollen n_{bp} gives the mean number of effective pollen grains per individual available for outcrossing:

$$n_p(t) = n_{bp} \frac{N_{\text{recv}}(t)}{K} \left(1 - \sigma \frac{N_{\text{self}}(t)}{N_{\text{recv}}(t)}\right) = \frac{n_{bp}}{K} (N_{\text{recv}}(t) - \sigma N_{\text{self}}(t)).$$

When all receivers can self ($N_{\text{recv}}(t) = N_{\text{self}}(t)$), the above expression reduces to $n_p(t) = n_{bp}(1 - \sigma)N_{\text{recv}}(t)/K$. If $\sigma = 1$, all ovules of self-capable receivers are pre-empted, and $n_p(t) = 0$.

Unlike the dioecious model, where individuals of the limiting sex had to produce a minimum of two seeds that survive the seed bank for the population to sustain itself at K , a hermaphrodite must replace itself with one surviving seed on average; thus, $n_{\text{min}} = 1/\gamma$, where γ denotes the total germination probability of each seed.

If the suppression drive limits pollen production (for example, CAIN or ClvR male suppression), we derive the corresponding β_m by considering the limiting case in which all but one individual are pollen-sterile (but ovule-fertile). Because self pollen is not limiting, the individual contributes on average σn_{bo} seeds through selfing and n_{bp} seeds through outcrossing. The total seed output in this limiting case is thus $n_{bp} + \sigma n_{bo}$, giving a 'male-analogous' low-density growth rate:

$$\beta_m = \gamma(n_{bp} + \sigma n_{bo}). \quad (5)$$

For ClvR female suppression, where ovules are limiting, consider a scenario in which all individuals lack ovules (and thus cannot self) except for one. That individual produces n_{bo} seeds, which are all fertilized (by self or outcross pollen), giving a ‘female analogous’ low-density growth rate:

$$\beta_f = \gamma n_{bo}. \quad (6)$$

Genetic load is again defined as the proportional reduction in total effective ovules or pollen relative to a wild-type population of the same size. Because wild types produce both ovules and pollen grains in this hermaphrodite model, the expected numbers of effective ovules and pollen grains at equilibrium are Kn_{bo} ovules and Kn_{bp} pollen. Using the critical-load threshold $1 - 1/\beta$ (ref. 83), the required genetic loads are:

$$\lambda_m^* = 1 - \frac{1}{\gamma(n_{bp} + \sigma n_{bo})} \quad (\text{ClvR male suppression and CAIN male suppression}), \quad (7)$$

$$\lambda_f^* = 1 - \frac{1}{\gamma n_{bo}} \quad (\text{ClvR female suppression}). \quad (8)$$

We set each drive to its experimental defaults (Supplementary Table 1) and, for a given selfing rate, varied b in $[0, 1]$ (step 0.1) and m in $[0, 2]$ (step 0.2), with ten replicates per parameter combination. For modification drives, we recorded the fraction of replicates reaching fixation (in both plants and seeds) within 500 years and the mean time to fixation among successful replicates. For suppression drives, we tracked the fraction of replicates achieving population elimination (of both plants and seeds) within 500 years.

Drive frequency in plants following a seed introduction

By default, in our dioecious model, we introduced the drive by replacing sufficient wild-type plants with drive heterozygotes, setting the initial drive allele frequency in the plant population $p_0 = 0.1$. As a practical alternative, we modelled a seed introduction, in which drive-carrying age-0 seeds enter the seed bank and germinate alongside wild-type seeds. The resulting plant-level drive frequency after seed-bank survival, germination and density-dependent seedling competition can be obtained analytically as follows.

Assuming complete pollination of effective ovules at wild-type equilibrium, a baseline mean effective ovule count of n_{bo} and a carrying capacity of K , the dioecious model produces $S = n_{bo} \times \frac{K}{2}$ age-0 seeds per year. Let S_0 denote the number of age-0 seeds in the release year. For a seed replacement in the release year (not a drop), $S_0 = S$ and the age-0 seed drive allele frequency is p_s . However, if the drive is released as a drop of seeds, $S_0 > S$ during the release year. To achieve seed-level allele frequency p_s , the required number of dropped seeds (S_d) is:

$$S_d = \begin{cases} \frac{2Sp_s}{1-2p_s} & \text{for a drive-heterozygote drop } (0 < p_s < \frac{1}{2}), \quad (9) \\ \frac{Sp_s}{1-p_s} & \text{for a drive-homozygote drop } (0 < p_s < 1), \quad (10) \end{cases}$$

such that $S_0 = S + S_d$.

Age-0 seeds survive the seed bank and germinate at rates d and b , respectively, so age-0 seeds enter the seedling pool at expected count S_0db . At wild-type equilibrium, older seed-bank cohorts contribute $S(\gamma - db)$ seedlings. The total number of seedling competitors after the drive release is thus $S_0db + S(\gamma - db)$. Under density-dependent regulation, the per-seedling survival rate to the plant stage is $\frac{K}{S_0db + S(\gamma - db)}$; hence, allele frequencies are preserved from the seedling pool to adults. The resulting initial drive allele frequency in the plant population is:

$$p_0 = \frac{p_s S_0 db}{S_0 db + S(\gamma - db)}, \quad (\text{seed drop}) \quad (11)$$

which reduces under seed replacement ($S_0 = S$) to:

$$p_0 = \frac{p_s db}{\gamma}, \quad (\text{seed replacement}). \quad (12)$$

Descriptions of the variables introduced in this section are provided in Supplementary Table 3.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data supporting the simulation studies reported in this manuscript are publicly available via Zenodo at <https://doi.org/10.5281/zenodo.15110717> (ref. 90).

Code availability

The code used for the simulation studies reported in this manuscript is publicly available via Zenodo at <https://doi.org/10.5281/zenodo.15110717> (ref. 90).

References

- Oerke, E. C. Crop losses to pests. *J. Agric. Sci.* **144**, 31–43 (2006).
- Chauhan, B. S. Grand challenges in weed management. *Front. Agron.* **1**, 3 (2020).
- Sharma, G., Barney, J. N., Westwood, J. H. & Haak, D. C. Into the weeds: new insights in plant stress. *Trends Plant Sci.* **26**, 1050–1060 (2021).
- Clements, D. R., Upadhyaya, M. K., Joshi, S. & Shrestha, A. *Global Plant Invasions on the Rise 1–28* (Springer International, 2022).
- Scott, D. & Freckleton, R. P. Crop diversification and parasitic weed abundance: a global meta-analysis. *Sci. Rep.* **12**, 19413 (2022).
- Pimentel, D., Lach, L., Zuniga, R. & Morrison, D. Environmental and economic costs of nonindigenous species in the United States. *BioScience* **50**, 53–65 (2000).
- Heap, I. M. The occurrence of herbicide-resistant weeds worldwide. *Pestic. Sci.* **51**, 235–243 (1997).
- Mithila, J., Hall, J. C., Johnson, W. G., Kelley, K. B. & Riechers, D. E. Evolution of resistance to auxinic herbicides: historical perspectives, mechanisms of resistance, and implications for broadleaf weed management in agronomic crops. *Weed Sci.* **59**, 445–457 (2011).
- Heap, I. & Duke, S. O. Overview of glyphosate-resistant weeds worldwide. *Pest Manage. Sci.* **74**, 1040–1049 (2018).
- Travlos, I., de Prado, R., Chachalis, D. & Bilalis, D. J. Herbicide resistance in weeds: early detection, mechanisms, dispersal, new insights and management issues. *Front. Ecol. Evol.* **8**, 213 (2020).
- Kersten, S. Standing genetic variation fuels rapid evolution of herbicide resistance in blackgrass. *Proc. Natl Acad. Sci. USA* **120**, e2206808120 (2023).
- Burt, A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proc. R. Soc. Lond. B* **270**, 921–928 (2003).
- Champer, J., Buchman, A. & Akbari, O. S. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nat. Rev. Genet.* **17**, 146–159 (2016).
- Bier, E. Gene drives gaining speed. *Nat. Rev. Genet.* **23**, 5–22 (2022).
- Neve, P. Gene drive systems: do they have a place in agricultural weed management?. *Pest Manage. Sci.* **74**, 2671–2679 (2018).
- Barrett, L. G. Gene drives in plants: opportunities and challenges for weed control and engineered resilience. *Proc. R. Soc. B* **286**, 20191515 (2019).

17. Pixley, K. V. Genome editing, gene drives, and synthetic biology: will they contribute to disease-resistant crops, and who will benefit? *Annu. Rev. Phytopathol.* **57**, 165–188 (2019).
18. Legros, M. Gene drive strategies of pest control in agricultural systems: challenges and opportunities. *Evol. Appl.* **14**, 2162–2178 (2021).
19. Wong, A. C. S., Massel, K., Lam, Y., Hintzschke, J. & Chauhan, B. S. Biotechnological road map for innovative weed management. *Front. Plant Sci.* **13**, 887723 (2022).
20. Yin, K., Chung, M. Y., Lan, B., Du, F. K. & Chung, M. G. Plant conservation in the age of genome editing: opportunities and challenges. *Genome Biol.* **25**, 279 (2024).
21. Neve, P. & Barrett, L. Driving evolution in wild plants. *Nat. Plants* **10**, 840–841 (2024).
22. Stokstad, E. Two teams supercharge gene spread in plants. *Science* **384**, 1394–1395 (2024).
23. Hay, B. A. Applications and status of gene drive in plants. *PLoS Biol.* **23**, e3003148 (2025).
24. Liu, Y., Jiao, B., Champer, J. & Qian, W. Overriding Mendelian inheritance in *Arabidopsis* with a CRISPR toxin–antidote gene drive that impairs pollen germination. *Nat. Plants* **10**, 910–922 (2024).
25. Oberhofer, G., Johnson, M. L., Ivy, T., Antoshechkin, I. & Hay, B. A. *Cleave and Rescue* gamete killers create conditions for gene drive in plants. *Nat. Plants* **10**, 936–953 (2024).
26. Mao, Y. Application of the CRISPR–Cas system for efficient genome engineering in plants. *Mol. Plant* **6**, 2008–2011 (2013).
27. Champer, J., Kim, I. K., Champer, S. E., Clark, A. G. & Messer, P. W. Performance analysis of novel toxin–antidote CRISPR gene drive systems. *BMC Biol.* **18**, 27 (2020).
28. Brown, J. S. & Venable, D. L. Evolutionary ecology of seed-bank annuals in temporally varying environments. *Am. Nat.* **127**, 31–47 (1986).
29. Ellner, S. & Hairston, N. G. Role of overlapping generations in maintaining genetic variation in a fluctuating environment. *Am. Nat.* **143**, 403–417 (1994).
30. Tellier, A. Persistent seed banking as eco-evolutionary determinant of plant nucleotide diversity: novel population genetics insights. *New Phytol.* **221**, 725–730 (2019).
31. Lennon, J. T., den Hollander, F., Wilke-Berenguer, M. & Blath, J. Principles of seed banks and the emergence of complexity from dormancy. *Nat. Commun.* **12**, 4807 (2021).
32. Legros, M. & Barrett, L. G. DriverSEAT: a spatially-explicit stochastic modelling framework for the evaluation of gene drives in novel target species. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.06.13.496025> (2022).
33. Kim, J. Incorporating ecology into gene drive modelling. *Ecol. Lett.* **26**, S62–S80 (2023).
34. National Academies of Sciences, Engineering, and Medicine. *Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values* (National Academies Press, 2016).
35. EFSA Panel on Genetically Modified Organisms (GMO). Adequacy and sufficiency evaluation of existing EFSA guidelines for the molecular characterisation, environmental risk assessment and post-market environmental monitoring of genetically modified insects containing engineered gene drives. *EFSA J.* **18**, e06297 (2020).
36. Haller, B. C. & Messer, P. W. SLiM 4: multispecies eco-evolutionary modeling. *Am. Nat.* **201**, E127–E139 (2023).
37. Templeton, A. R. & Levin, D. A. Evolutionary consequences of seed pools. *Am. Nat.* **114**, 232–249 (1979).
38. Hairston Jr, N. G. & De Stasio Jr, B. T. Rate of evolution slowed by a dormant propagule pool. *Nature* **336**, 239–242 (1988).
39. Wenny, D. G. Seed dispersal, seed predation, and seedling recruitment of a neotropical montane tree. *Ecol. Monogr.* **70**, 331–351 (2000).
40. Bell, T., Freckleton, R. P. & Lewis, O. T. Plant pathogens drive density-dependent seedling mortality in a tropical tree. *Ecol. Lett.* **9**, 569–574 (2006).
41. Comita, L. S. Testing predictions of the Janzen–Connell hypothesis: a meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. *J. Ecol.* **102**, 845–856 (2014).
42. Norsworthy, J. K. Reducing the risks of herbicide resistance: best management practices and recommendations. *Weed Sci.* **60**, 31–62 (2012).
43. Walsh, M., Newman, P. & Powles, S. Targeting weed seeds in-crop: a new weed control paradigm for global agriculture. *Weed Technol.* **27**, 431–436 (2013).
44. Roberts, J. & Florentine, S. A review of the biology, distribution patterns and management of the invasive species *Amaranthus palmeri* S. Watson (Palmer amaranth): current and future management challenges. *Weed Res.* **62**, 113–122 (2022).
45. Tranel, P. J. Herbicide resistance in *Amaranthus tuberculatus*. *Pest Manage. Sci.* **77**, 43–54 (2021).
46. McCraw, J. M. & Spoor, W. Self-incompatibility in *Lolium* species. *Heredity* **50**, 29–33 (1983).
47. Chauvel, B. & Gasquez, J. Relationships between genetic polymorphism and herbicide resistance within *Alopecurus myosuroides* Huds. *Heredity* **72**, 336–344 (1994).
48. Bull, J. J. Lethal gene drive selects inbreeding. *Evol. Med. Public Health* **2017**, 1–16 (2017).
49. Bull, J. J., Remien, C. H. & Krone, S. M. Gene-drive-mediated extinction is thwarted by population structure and evolution of sib mating. *Evol. Med. Public Health* **2019**, 66–81 (2019).
50. Champer, J., Kim, I. K., Champer, S. E., Clark, A. G. & Messer, P. W. Suppression gene drive in continuous space can result in unstable persistence of both drive and wild-type alleles. *Mol. Ecol.* **30**, 1086–1101 (2021).
51. Unckless, R. L., Clark, A. G. & Messer, P. W. Evolution of resistance against CRISPR/Cas9 gene drive. *Genetics* **205**, 827–841 (2017).
52. Champer, J. Reducing resistance allele formation in CRISPR gene drive. *Proc. Natl Acad. Sci. USA* **115**, 5522–5527 (2018).
53. Oberhofer, G., Ivy, T. & Hay, B. A. Behavior of homing endonuclease gene drives targeting genes required for viability or female fertility with multiplexed guide RNAs. *Proc. Natl Acad. Sci. USA* **115**, E9343–E9352 (2018).
54. Champer, S. E. Computational and experimental performance of CRISPR homing gene drive strategies with multiplexed gRNAs. *Sci. Adv.* **6**, eaaz0525 (2020).
55. Yang, E. A homing suppression gene drive with multiplexed gRNAs maintains high drive conversion efficiency and avoids functional resistance alleles. *G3 (Bethesda)* **12**, jkac081 (2022).
56. Pail, J. F. & Phillips, B. L. Slow and steady wins the race: spatial and stochastic processes and the failure of suppression gene drives. *Mol. Ecol.* **31**, 4451–4464 (2022).
57. Zhang, X., Sun, W., Kim, I. K., Messer, P. W. & Champer, J. Population dynamics in spatial suppression gene drive models and the effect of resistance, density dependence, and life history. Preprint at *bioRxiv* <https://doi.org/10.1101/2024.08.14.607913> (2024).
58. Kyrou, K. A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nat. Biotechnol.* **36**, 1062–1066 (2018).
59. Simoni, A. A male-biased sex-distorter gene drive for the human malaria vector *Anopheles gambiae*. *Nat. Biotechnol.* **38**, 1054–1060 (2020).

60. Carballar-Lejarazú, R. Next-generation gene drive for population modification of the malaria vector mosquito, *Anopheles gambiae*. *Proc. Natl Acad. Sci. USA* **117**, 22805–22814 (2020).
61. Adolphi, A. Efficient population modification gene-drive rescue system in the malaria mosquito *Anopheles stephensi*. *Nat. Commun.* **11**, 5553 (2020).
62. Hammond, A. Gene-drive suppression of mosquito populations in large cages as a bridge between lab and field. *Nat. Commun.* **12**, 4589 (2021).
63. Altrock, P. M., Traulsen, A., Reeves, R. G. & Reed, F. A. Using underdominance to bi-stably transform local populations. *J. Theor. Biol.* **267**, 62–75 (2010).
64. Edgington, M. P. & Alphey, L. S. Conditions for success of engineered underdominance gene drive systems. *J. Theor. Biol.* **430**, 128–140 (2017).
65. Champer, J., Zhao, J., Champer, S. E., Liu, J. & Messer, P. W. Population dynamics of underdominance gene drive systems in continuous space. *ACS Synth. Biol.* **9**, 779–792 (2020).
66. Champer, J., Champer, S. E., Kim, I. K., Clark, A. G. & Messer, P. W. Design and analysis of CRISPR-based underdominance toxin–antidote gene drives. *Evol. Appl.* **14**, 1052–1069 (2021).
67. Friedman, J. & Barrett, S. C. H. Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. *Ann. Bot.* **103**, 1515–1527 (2009).
68. Allison, T. D. Pollen production and plant density affect pollination and seed production in *Taxus canadensis*. *Ecology* **71**, 516–522 (1990).
69. Delph, L. F., Johannsson, M. H. & Stephenson, A. G. How environmental factors affect pollen performance: ecological and evolutionary perspectives. *Ecology* **78**, 1632–1639 (1997).
70. Cooper, D. & Brink, R. Partial self-incompatibility and the collapse of fertile ovules as factors affecting seed formation in alfalfa. *J. Agric. Res.* **60**, 453–472 (1940).
71. Mogensen, H. L. Ovule abortion in *Quercus* (Fagaceae). *Am. J. Bot.* **62**, 160–165 (1975).
72. Lee, T. & Bazzaz, F. Maternal regulation of fecundity: non-random ovule abortion in *Cassia fasciculata* Michx. *Oecologia* **68**, 459–465 (1986).
73. Field, D. L., Pickup, M. & Barrett, S. C. The influence of pollination intensity on fertilization success, progeny sex ratio, and fitness in a wind-pollinated, dioecious plant. *Int. J. Plant Sci.* **173**, 184–191 (2012).
74. Crocker, W. Life-span of seeds. *Bot. Rev.* **4**, 235–274 (1938).
75. Nagel, M. & Börner, A. The longevity of crop seeds stored under ambient conditions. *Seed Sci. Res.* **20**, 1–12 (2009).
76. Cook, R. E. *Patterns of Juvenile Mortality and Recruitment in Plants* 207–231 (Macmillan Education UK, 1979).
77. Ellner, S. Competition and dormancy: a reanalysis and review. *Am. Nat.* **130**, 798–803 (1987).
78. Tielborger, K. & Valleriani, A. Can seeds predict their future? Germination strategies of density-regulated desert annuals. *Oikos* **111**, 235–244 (2005).
79. Kaj, I., Krone, S. M. & Lascoux, M. Coalescent theory for seed bank models. *J. Appl. Probab.* **38**, 285–300 (2001).
80. Dhole, S., Lloyd, A. L. & Gould, F. Gene drive dynamics in natural populations: the importance of density dependence, space, and sex. *Annu. Rev. Ecol. Evol. Syst.* **51**, 505–531 (2020).
81. Xu, Z. & Bonsall, M. B. Ecology of gene drives: the role of density-dependent feedbacks on the efficacy and dynamics of two-locus underdominance gene drive systems. *Evol. Appl.* **18**, e70079 (2025).
82. Deredec, A., Burt, A. & Godfray, H. C. J. The population genetics of using homing endonuclease genes in vector and pest management. *Genetics* **179**, 2013–2026 (2008).
83. Deredec, A., Godfray, H. C. J. & Burt, A. Requirements for effective malaria control with homing endonuclease genes. *Proc. Natl Acad. Sci. USA* **108**, E874–E880 (2011).
84. Champer, S. E., Kim, I. K., Clark, A. G., Messer, P. W. & Champer, J. *Anopheles* homing suppression drive candidates exhibit unexpected performance differences in simulations with spatial structure. *eLife* **11**, e79121 (2022).
85. Han, Y. & Champer, J. A comparative assessment of self-limiting genetic control strategies for population suppression. *Mol. Biol. Evol.* **42**, msaf048 (2025).
86. Ritz, C., Baty, F., Streibig, J. C. & Gerhard, D. Dose–response analysis using R. *PLoS ONE* **10**, e0146021 (2015).
87. Holsinger, K. E. Mass-action models of plant mating systems: the evolutionary stability of mixed mating systems. *Am. Nat.* **138**, 606–622 (1991).
88. Lloyd, D. G. & Schoen, D. J. Self- and cross-fertilization in plants. I. Functional dimensions. *Int. J. Plant Sci.* **153**, 358–369 (1992).
89. Lloyd, D. G. Self- and cross-fertilization in plants. II. The selection of self-fertilization. *Int. J. Plant Sci.* **153**, 370–380 (1992).
90. Kim, I. K. et al. Gene drive dynamics in plants: the role of seedbanks: data. *Zenodo* <https://doi.org/10.5281/zenodo.15110717> (2025).

Acknowledgements

This work was supported by the National Institutes of Health (grant nos R35GM152242 to P.W.M. and R35GM156957 to J.K.), the National Science Foundation Graduate Research Fellowship Program (grant no. DGE2139899 to R.C.), the Center for Life Sciences (J.C.) and the National Natural Science Foundation of China (grant nos 32270672 and W2432018 to J.C.).

Author contributions

J.K. conceived the study. I.K.K. and L.T. developed the models with input from R.C., B.C.H., J.C., P.W.M. and J.K. I.K.K. performed the analyses. I.K.K., P.W.M. and J.K. wrote the original draft of the manuscript. All authors reviewed and edited the manuscript. J.C., P.W.M. and J.K. supervised the study.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41477-026-02256-1>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41477-026-02256-1>.

Correspondence and requests for materials should be addressed to Jaehee Kim.

Peer review information *Nature Plants* thanks Wenfeng Qian and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

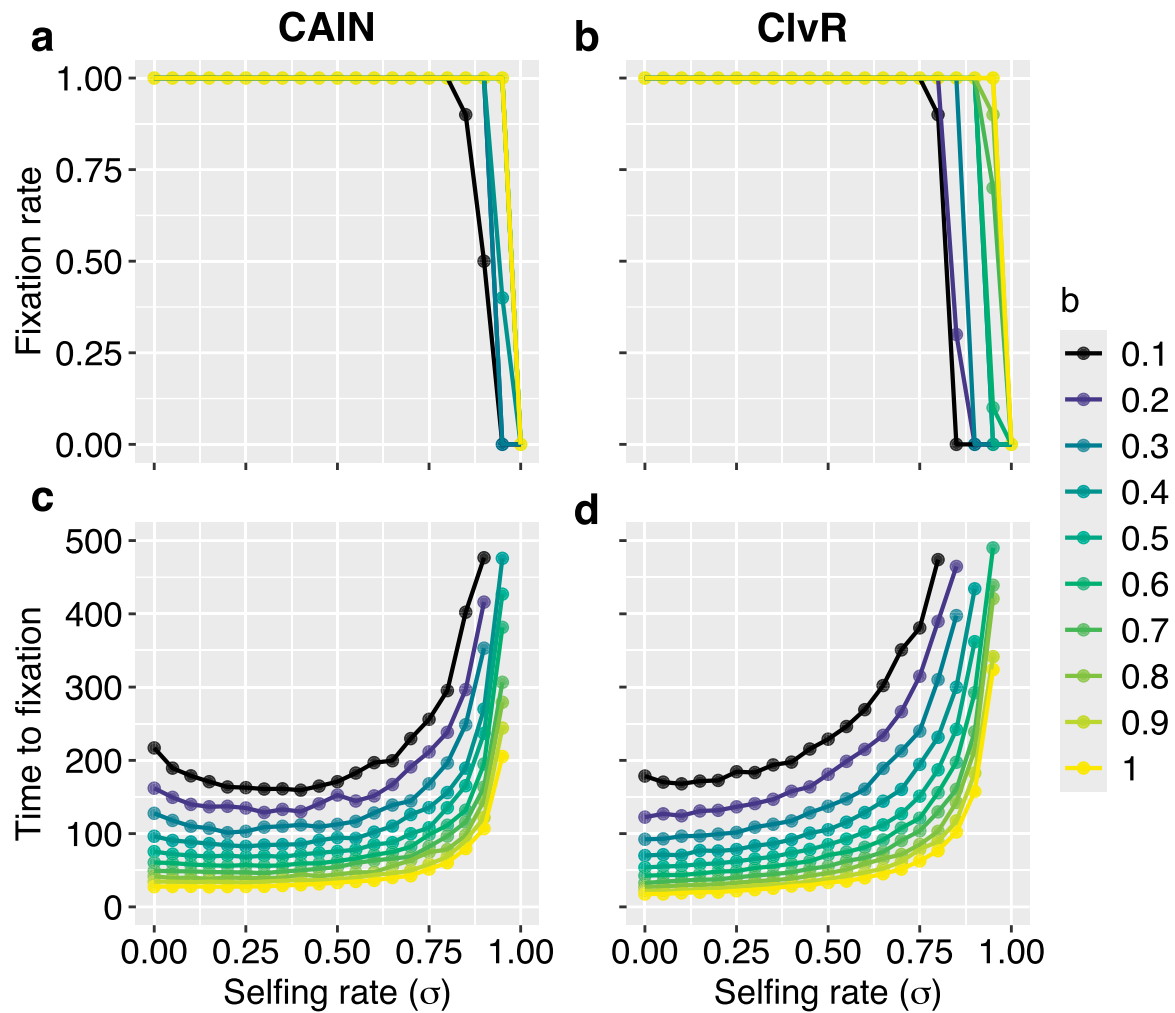
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit

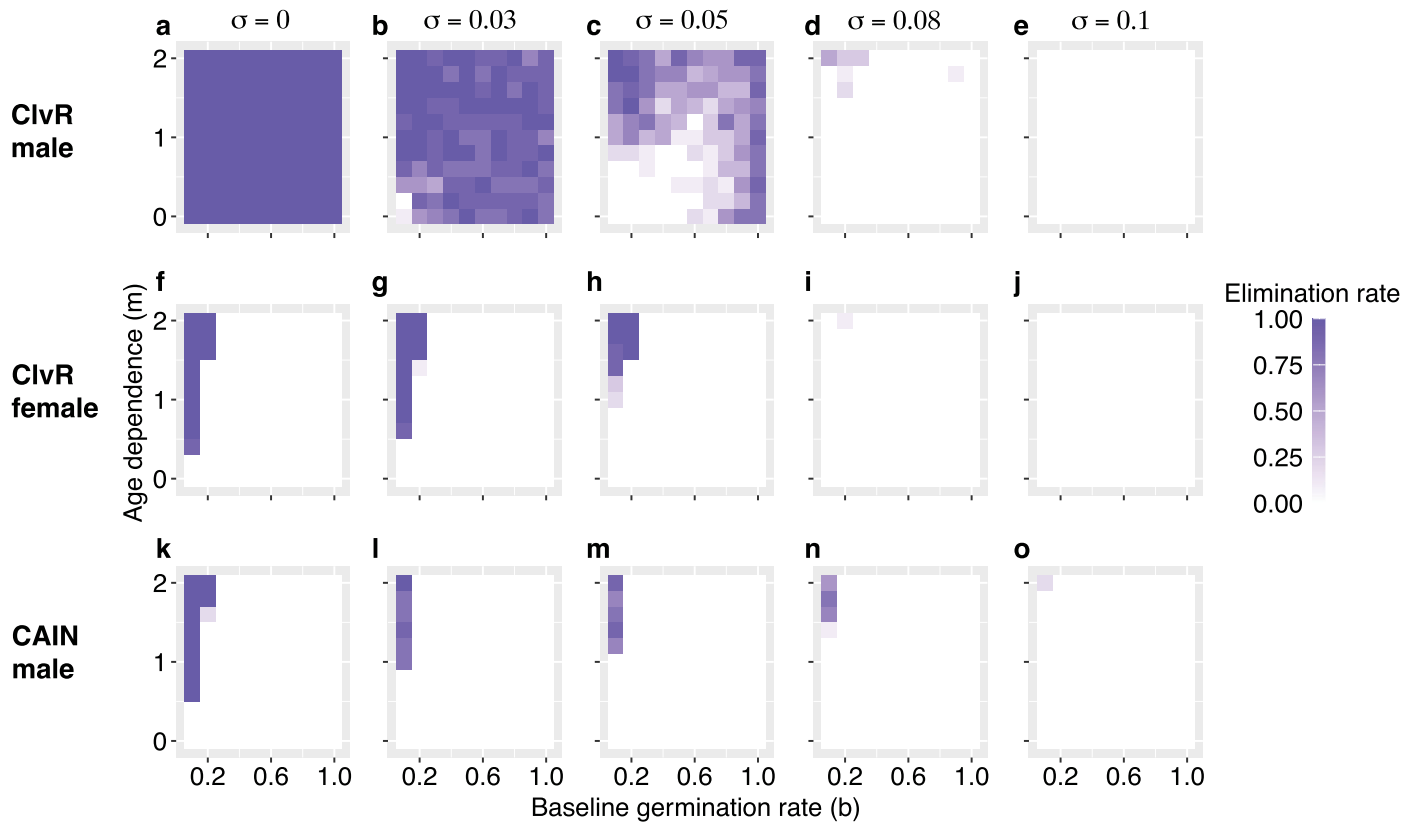
line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026



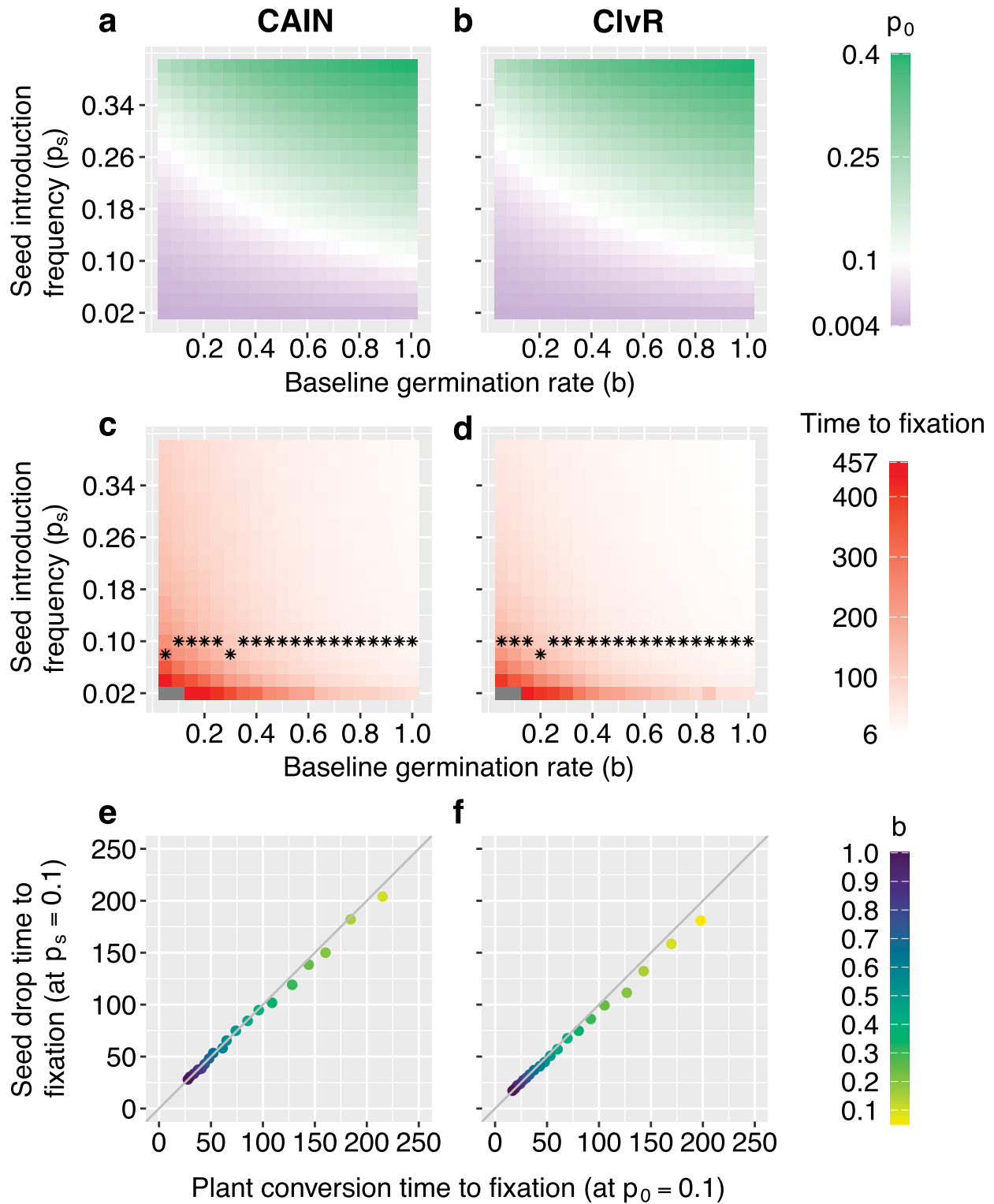
Extended Data Fig. 1 | Spread of CAIN and ClvR modification drives for different selfing rates. In the hermaphrodite model, we varied the selfing rate σ while holding each drive's parameters at their experimental defaults and assuming no age-dependence in germination ($m = 0$). Curves are colored by the

baseline germination rate b . **a, b**, Fixation rate of CAIN and ClvR, respectively, across (b, σ), estimated over 10 replicates per parameter set. **c, d**, Mean time to fixation for CAIN and ClvR, respectively, among successful replicates.



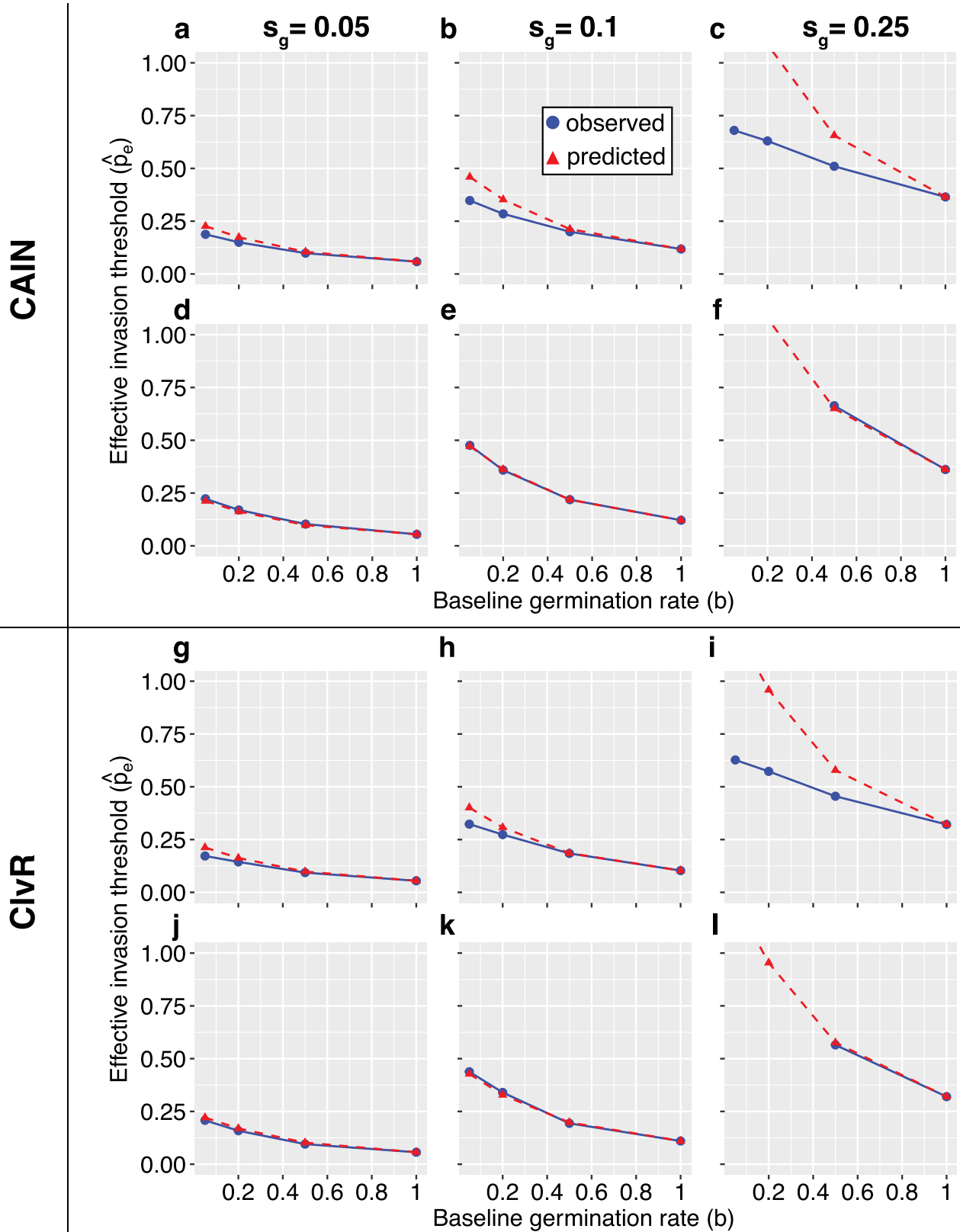
Extended Data Fig. 2 | Effect of selfing on the success of ClvR and CAIN suppression drives. In the hermaphrodite model, we varied the selfing rate (σ), baseline germination rate (b), and age-dependent germination parameter (m), keeping drive parameters at experimentally derived defaults (Supplementary Table 1). For each parameter set, the elimination rate is defined as the fraction

of 10 replicates in which all plants and seeds were eradicated. Heatmaps display elimination rate (color scale) across b and m . **a–e**, ClvR male suppression at selfing rates of 0, 0.03, 0.05, 0.08, and 0.10, respectively. **f–j**, ClvR female suppression (as in **a–e**). **k–o**, CAIN male suppression (as in **a–e**).



Extended Data Fig. 3 | Spread of CAIN and ClvR modification drives when introduced via seed drops. Instead of the default plant replacement (replacing a fraction of wild-type plants with drive heterozygotes), the drive is introduced by dropping drive-heterozygous age-0 seeds into the seed bank. We vary the drive-carrying seed introduction frequency (p_s) and the baseline germination rate (b), assuming germination rates do not decline with age ($m = 0$), and with 10 replicates per parameter combination. **a**, CAIN: mean initial plant-stage drive frequency p_0 after drive-carrying seeds have survived, germinated, and

successfully established as seedlings in the first year (color scale at right). **b**, Same as **a** for ClvR. **c**, CAIN: average time to fixation among successful replicates (≤ 500 years; color scale at right). For each **b** (with $m = 0$), the star marks the p_s whose average time to fixation most closely matches that of the plant replacement at the same b and m . **d**, As in **c** for ClvR. **e**, CAIN: average time to fixation for the plant replacement with introduction frequency $p_0 = 0.1$ in plants versus that of the seed drop at introduction frequency $p_s = 0.1$ in age-0 seeds; points are colored by b , the gray line is $y = x$. **f**, As in **e** but for ClvR.



Extended Data Fig. 4 | Effective invasion thresholds for CAIN and ClvR under seed drop versus seed replacement. We varied the baseline germination rate b and drive-associated gamete viability cost s_g (probability a drive-carrying gamete is nonviable), holding all other drive and seed-bank parameters at their defaults (Supplementary Tables 1 and 2). Blue points denote observed effective invasion thresholds \hat{p}_e , defined as the minimum seed introduction frequency (p_i) yielding drive spread in more than 50% of replicates; red triangles are predictions from

$\hat{p}_e \approx \hat{p} \times \tau$. Points at $b=1$ give the baseline invasion threshold \hat{p} without seed banks. **a–c**, CAIN released by dropping drive-homozygous seeds (seed drop); invasion thresholds for $s_g \in \{0.05, 0.1, 0.25\}$ and $b \in \{0.05, 0.2, 0.5, 1\}$. **d–f**, CAIN released by replacing a fraction p_i of age-0 seeds with drive homozygotes (seed replacement). Missing blue points indicate consistent failure to invade; missing red triangles indicate predicted $\hat{p}_e > 1$. **g–i**, As in **a–c** for ClvR. **j–l**, As in **d–f** for ClvR.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Simulations were performed using SLiM (version 4.0.1), an open source individual-based simulation framework. SLiM can be downloaded at <https://messerlab.org/slim/>

Data analysis Figures were created in R (version 4.3.3). All code associated with this project can be found on Zenodo at <https://doi.org/10.5281/zenodo.15110717>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data and code used for the simulation studies presented in this manuscript is publicly available on Zenodo at <https://doi.org/10.5281/zenodo.15110717>

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study investigates the ecological and evolutionary dynamics of CRISPR-based gene drive systems (CAIN and ClvR) in annual, wind-pollinated plants by incorporating seed dormancy into individual-based simulations. It investigates how seedbank characteristics influence gene drive efficacy, suppression potential, and invasion thresholds, highlighting seedbanks as critical modulators for field applications in weed management.
Research sample	N/A
Sampling strategy	N/A
Data collection	Data were collected through individual-based forward-time simulations implemented in SLiM, parameterized with experimentally derived gene drive values. Summary statistics, including mean drive fixation time, were calculated by averaging across replicates for each parameter set.
Timing and spatial scale	N/A
Data exclusions	No data were excluded from the analysis.
Reproducibility	All simulation codes, parameters, and instructions necessary to reproduce experimental results are provided on Zenodo (doi: 10.5281/zenodo.15110717). For each parameter set, 10 replicate simulations were performed, and all attempts to repeat these simulations were successful.
Randomization	N/A
Blinding	N/A

Did the study involve field work? Yes No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- n/a | Involved in the study
- Antibodies
 - Eukaryotic cell lines
 - Palaeontology and archaeology
 - Animals and other organisms
 - Clinical data
 - Dual use research of concern
 - Plants

Methods

- n/a | Involved in the study
- ChIP-seq
 - Flow cytometry
 - MRI-based neuroimaging

Plants

Seed stocks

N/A

Novel plant genotypes

N/A

Authentication

N/A