2020

Outsmarting Evolution with Gene Drive Technology: Thinking Beyond the Release

BETH DAVENPORT

TABLE OF CONTENTS

- INTRODUCTION
- BACKGROUND
    1. Molecular Mechanism
    2. History of CRISPR-based Gene Drives
- RESISTANCE: A Barrier to Drives in Wild Populations
    1. Mechanisms of Resistance Production
        a. Homing Process Errors
        b. Natural Genetic Variation
        c. Non-random Mating
    2. Modelling Drive Resistance
    3. Re-designing Drive Constructs
        a. NHEJ Prevention
        b. gRNA Multiplexing
        c. DRA Fitness Cost Maximisation
        d. Alternative Nucleases
        e. Working with Resistance Generation
    4. A Combinatorial Approach
- CONCLUSIONS AND CONTEMPLATION

Word Count: 4,997

ABSTRACT

The recent revolution in gene editing technology has enabled the possibility of solving multiple global health and environmental issues through the novel engineering of CRISPR-based (Clustered Regularly Interspaced Short Palindromic Repeats) gene drives. As drive application on wild populations draws nearer, there is a final obstacle to overcome: the emergence and propagation of gene drive resistance in subject populations. This paper reviews recent developments in gene drive technology, with emphasis on exploring the mechanisms behind, and redesigning of drives to prevent and solve, the resistance impediment.

INTRODUCTION

Gene drive technology (GDT), and more specifically, CRISPR-based gene drive technology (C-GDT), has emerged over the last decade as an exciting young biotechnology field with huge potential for addressing multiple pressing health and environmental issues. Gene drives utilise gene editing technologies to bias a genetic element's inheritance to spread it through a population of interest at super-Mendelian rates (>50% chance of inheritance) (Noble et al. 2017). The idea of suppressing or modifying populations with homing-based endonuclease drives was first expressed by Burt (2003). However, it wasn't until the recent discovery of the bacterial immune system's 'CRISPR' mechanism (containing an RNA-guided Cas endonuclease), that gene drives became an engineering feasibility, thanks to the nuclease being more species transferrable, easy to manipulate (Noble et al. 2017), and evolutionarily stable (Friess et al. 2019, Zhang et al. 2014).

Proposed C-GDT applications include elimination of invasive foreign species and reversal of resistance in crops to herbicides, insects to pesticides, and even pathogenic fungi and viruses to medicines (Shapiro et al. 2018, Walter & Verdin, 2019). The most discussed C-GDT application is the eradication of mosquito-borne diseases through population suppression (Champer et al. 2016): notably malaria, dengue, yellow fever and Zika. C-GDT can also be utilised for population modification, for example, to spread advantageous traits through endangered populations, or spread anti-pathogen genes through disease vectors (Gantz et al. 2015). C-GDT harbours the promise of carrying out these revolutionary solutions in a cost-effective, environmentally-conscious way, and is therefore a field of significant focus and funding in the coming years.

There have been frequent demonstrations of C-GDs achieving suppression and modification effects in multiple species models in the last five years (*Figure 1*), bringing the reality of a successful drive release into wild populations ever closer. However, a substantial impediment has emerged to first-generation C-GD propagation in real-world conditions: the generation and spread of drive-resistant alleles. This has shifted scientific focus away from C-GD design for laboratory trials, to beyond their release: outsmarting resistance evolution to achieve successful propagation of C-GDs on genetically diverse, wild populations. This paper provides a complete review of recent developments in the C-GDT field, with specific emphasis on the re-designing of first-generation C-GDs to battle evolutionary resistance generation.

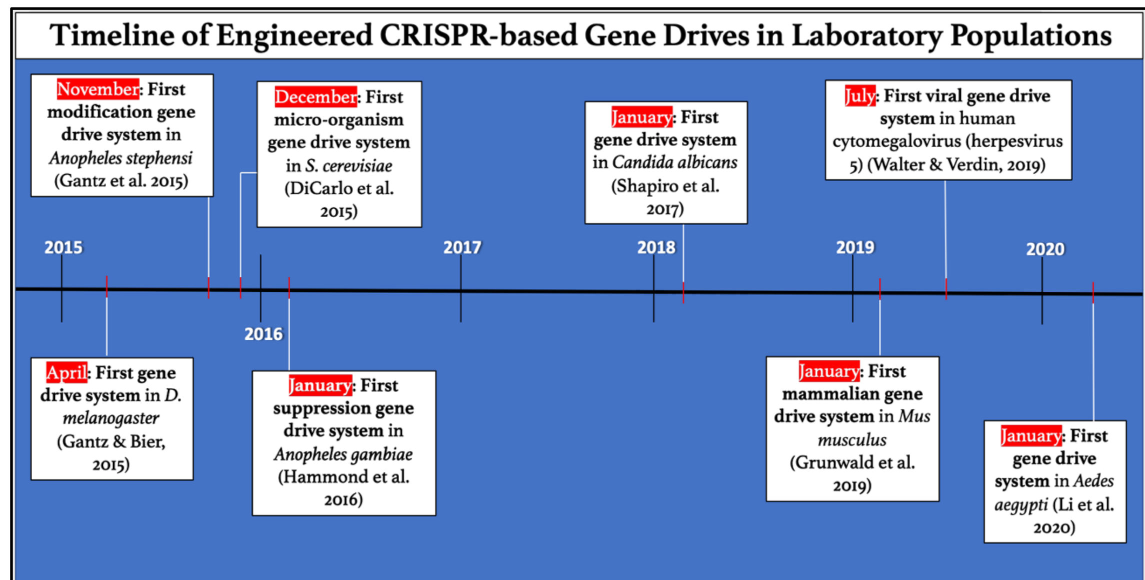

Figure 1*: Developments in laboratory C-GDs, in terms of novel drive types and species used. The majority of C-GDs have been in insect models, but there have also been two developed in yeast, one in a mammalian model and one in a virus.*

BACKGROUND

I. Molecular Mechanism

The CRISPR genetic editing system involves a complex consisting of a Cas endonuclease (commonly Cas9 from *Streptococcus pyogenes:* SpCas9), and a single-stranded guide RNA (gRNA) sequence. This gRNA is programmable to precisely guide the Cas9 to a desired complementary target sequence within a genome (*Figure 2*), where it cleaves the DNA to create a Double Stranded Break (DSB). Depending on what repair mechanism is utilised at the DSB, a gene knockout or a desired gene insertion can be created within an organism's genome (Zaidi et al. 2017).

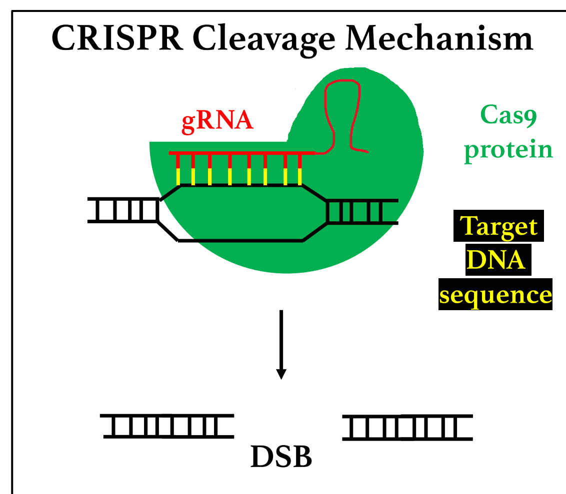

***Figure 2**: Cas9/gRNA complex genetic mechanism; the Cas9 nuclease is guided to the target sequence by its gRNA using Watson-Crick base pairing, where it cleaves the DNA.*

When used in C-GDT, a Drive Allele (DA) cassette is constructed and inserted into an organism's genome. The DA comprises genes encoding (i) Cas9; (ii) a programmed gRNA whose target sequence is within the wildtype allele of the same genomic locus; (iii) a desired gene to be driven through the population (optional); and (iv) flanking homologous sequences to two sites enclosing the gRNA's target (Zentner & Wade, 2017). Cas9 cleaves the DNA sequence targeted by the gRNA in the wildtype allele, activating the cell's DNA repair mechanisms (Cong et al. 2013). Homology Directed Repair (HDR) uses the DA as a template to repair the DSB, resulting in its 'homing' (copying) into the wildtype chromosome (Ran et al. 2013, *Figure 3*).

To ensure super-Mendelian DA transmission rates to following generations, DA homing ideally takes place within the germline, so that all gametes contain the DA; hence, the Cas9-gRNA complex enables HDR to facilitate biased inheritance (theoretically up to 100%) of all components within the flanking homologous arms (Champer et al. 2016, *Figure 4*).

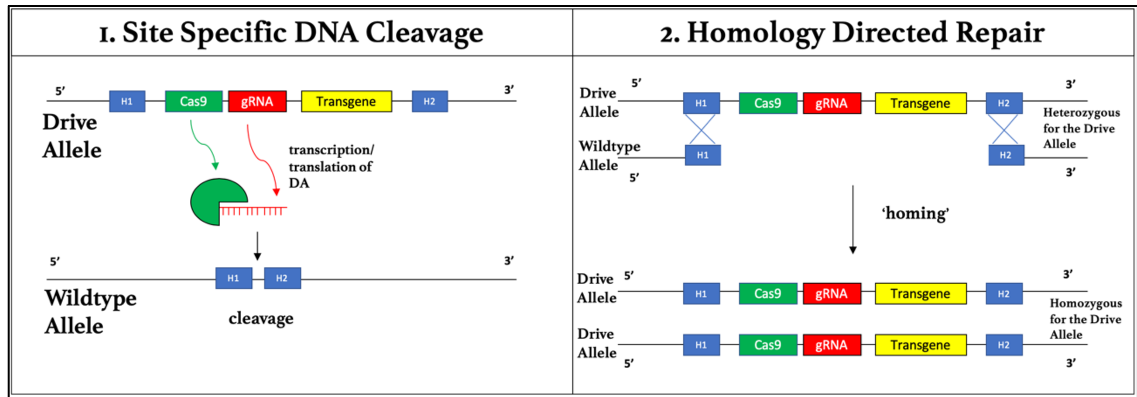

Figure 3*: (1) In a heterozygous DA genome, the translated Cas9 nuclease is guided to the target sequence by the transcribed gRNA where it cleaves the DNA. (2) As the DA flanking arms are homologous to those adjacent to the DSB (labelled H1 and H2), the DA is used by HDR machinery as a template to resynthesise and repair the DSB. Therefore, the DA is homed into the wildtype chromosome, resulting in a homozygous DA genome.*

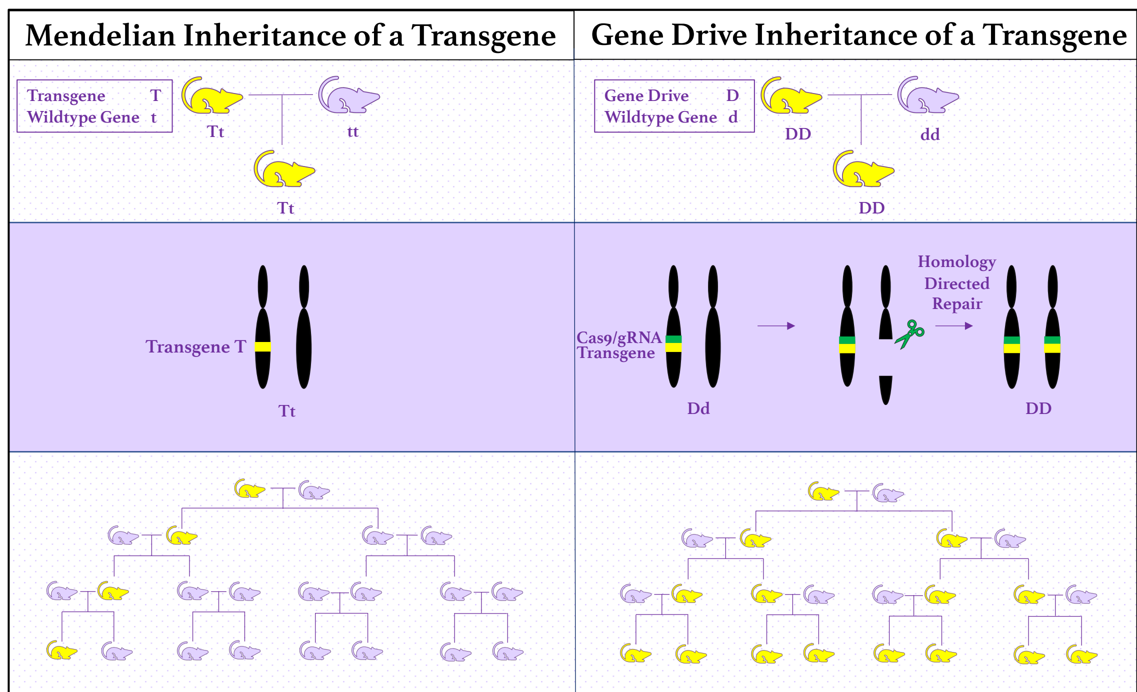

Figure 4*: The theoretical pattern of transgene spread by Mendelian inheritance throughout four generations in a large population, compared to transgene spread by C-GD.*

2. History of CRISPR-based Gene Drives

The first demonstration of a successful C-GD was in the insect model *Drosophila melanogaster*; when Gantz & Bier (2015) drove a DA within the *yellow* gene (changing body colour) through the laboratory population, exhibiting remarkable efficiency (from 95% to 100% DA phenotype within three generations). This evidenced a potent drive system for further use. Shortly following this, a drive changing cell colour in *Saccharomyces cerevisiae* yeast exhibited 99% Cas9 homing efficiency into wildtype alleles (DiCarlo et al. 2015), and successful modification and suppression drives in *Anopheles stephensi* and *Anopheles gambiae* mosquito models respectively were demonstrated (Gantz et al. 2015, Hammond et al. 2016).

RESISTANCE: A Barrier to Drives in Wild Populations

For several years after these four novel gene drive demonstrations (*Figure 1*), there was a research focus on employing previously developed drive platforms to identify limitations and optimise practical efficiency through redesign. This was largely a consequence of studies drawing attention to observed generation of Drive-Resistant Alleles (DRAs) in early laboratory trials despite high homing rates (Gantz et al. 2015, Hammond et al. 2016), and to their predicted generation in modelled wild populations when subject to original DA designs; Unckless et al. (2017)'s model implicated resistance as the biggest impediment to successful drive spread. Though DRA generation rate would not prevent a DA's immediate spread through wild populations, in time out-competition through natural selection would occur, purging the DA from the population (Champer et al. 2018). The duration of time before the DA purge is dependent on multiple factors, though largely the difference in fitness cost between the DA and DRA (Unckless et al. 2017).

1. Mechanisms of Resistance Production

Molecular mechanisms behind DRA generation are errors in the homing process itself or natural genetic variation within the target population. Alternatively, resistance to drive propagation can occur through selection of genes favouring non-random mating behaviours (Zentner & Wade, 2017).

a. Homing Process Errors

Observed errors in the homing process that generate DRAs occur more frequently than those arising from natural genetic variation (Unckless et al. 2017). DRAs can be produced from incomplete HDR copying and Non-Homologous End Joining (NHEJ) events (*Figure 5*). HDR is the preferred method of DSB repair in germline cells, hence many DAs utilise germline-specific promoters (however, exact ratios of HDR:NHEJ is dependent on species, cell-cycle phase, developmental stage and individual genetic identity (Friess et al. 2019)). Hence, excluding natural genetic variation, DRA generation in germline cells is largely a result of incomplete HDR events.

i. Incomplete HDR

Incomplete HDR has been observed during C-GD trials (Hammond et al. 2016), and occurs due to replication errors from HDR machinery, resulting in indels or point mutations in the newly synthesised DA. The likelihood of an incomplete HDR event on homing is positively correlated with DA size (Marshall et al. 2017). However, incomplete HDR homing events occur infrequently enough that they are modelled to rarely impede DA spread to fixation in wild populations, unlike NHEJ events (Unckless et al. 2017).

ii. NHEJ

NHEJ occurs instead of HDR to mend a DSB if there is no long homologous sequence present to that which is broken (McVey & Lee, 2008) – which often takes precedence

immediately post-fertilisation (Friess et al. 2019), and during G1 cell-cycle stage in early embryogenesis (Lin et al. 2014). If NHEJ is utilised after Cas9 cleavage, the DA is not copied as template, and it can create mutant (and possibly drive-resistant) alleles through internal deletions. (Noble et al. 2017, *Figure 5*).

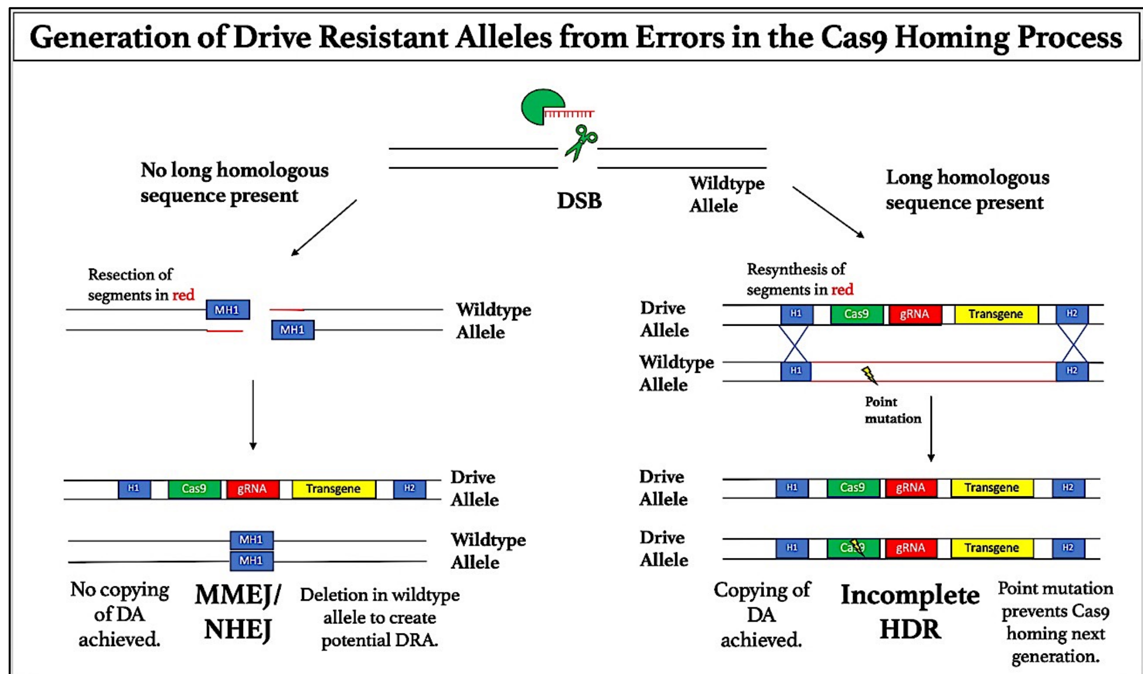

***Figure 5**: Microhomology-mediated end joining (MMEJ: a form of NHEJ) does not use long homologous strands as template, and instead excises away from DSBs until microhomologies on either strand are found. These are then ligated together, producing small deletions in Cas9 cleavage sites. Incomplete HDR utilises DAs as long homologous templates. However, replication errors result in point mutations/indels which knock out functioning components necessary for efficient homing (here, it has knocked out Cas9's cleavage function).*

Many studies have observed NHEJ DRA generation post-fertilisation in their drive trials due to Cas9 cleavage in the zygote/early embryo (Gantz et al. 2015, Hammond et al. 2016, Champer et al. 2017). This has been proposed as the result of (i) leaky germline-specific promoters which allow some Cas9 transcription post-fertilisation (Champer et al. 2017), or (ii) the 'dominant maternal effect' – on maternal drive inheritance, Cas9 protein and gRNA deposits from the oocyte cleave the sperm's genome immediately after fusion,

before the egg's genome can be utilised for homology by HDR machinery (Friess et al. 2019).

Champer et al. (2017) tested two DAs in *D. melanogaster* with *vasa* (used by Gantz & Bier (2015)) and *nanos* (supposedly more germline-specific) promoters, to observe whether DRA formation occurs in germline, in early embryo due to leaky promoter expression, or early embryo due to maternal deposition. While both constructs produced DRAs at high rates post-fertilisation, homing was observed to occur exclusively in the germline, indicating DRA generation as due to the dominant maternal effect, and not leaky promoters. In concurrence, Gantz et al. (2015) showed a >77% rate in DRA generation post-fertilisation in *A. stephensi* individuals that inherited the drive maternally. Kandul et al. (2019) investigated the effect of maternal deposition with a split drive system in *D. melanogaster* (a drive lacking the Cas9 gene): as there was still a homing rate of 59 to 72% (the rest repaired by NHEJ) observed in early embryos despite no means of Cas9 expression, maternal deposition of Cas9 protein can be concluded as sufficient for homing (and therefore DRA generating) alone, regardless of promoter germline-specificity.

b. Natural Genetic Variation

Natural genetic variation in a population can generate DRAs, either prior to drive application (i.e. a pre-existing allele whose target site cannot be recognised/cleaved by the drive due to single nucleotide polymorphisms), or post-drive application, due to de novo mutations within the DA, target site or trans-acting elements necessary for drive homing. Genetic diversity in both the drive location, and genetic background of individuals within a wild population, can dictate whether a C-GD successfully achieves fixation, or is outcompeted by DRAs soon after release (Champer et al. 2017).

i. Direct DRA generation

Drury et al. (2017) used insect population genetic data to investigate the likelihood of pre-existing DRA existence subject to a standard C-GD; most populations contained genetic variants within Cas9 target sites, some of which would infer homing immunity. As one pre-existing or post-drive application de novo generated low-fitness cost DRA in a population would be enough to outcompete a DA it is subject to (Champer et al. 2017), it is important to characterise genetic variability within the target population's wildtype allele before DA design. However, de novo mutations occurring in the target locus or DA post-release cannot be predicted, and therefore remain an unsolved risk to successful gene drive propagation (Marshall et al. 2017).

ii. Indirect DRA generation

Genetically diverse backgrounds can affect homing-error resistance emergence rates. Champer et al. (2017) tested two DAs in genetically diverse fruitfly populations - though homing efficiency and germline DRA rates showed little variance, early embryo DRA formation exhibited high levels of variation. DRA generation rate changes due to genetic elements not within the DA or target locus has been termed 'trans-resistance' (Unckless et al. 2017). Proposals for trans-resistance mechanisms include differing Cas9 expression timings/levels, differing Cas9 degradation rates, differing HDR:NHEJ in the early embryo (Champer et al. 2017), and RNA interference of homing machinery (Unckless et al. 2017). Hence, a genetically diverse population presents an issue for DRA generation regulation: just one individual with a high DRA generation rate could change the likelihood and timing of DRA emergence for an entire breeding population (Champer et al. 2017).

To better understand how genetic diversity contributes towards DRA generation rates post-fertilisation, Champer et al. (2019) drove a DA through >200 genetically diverse fruitfly lines. Early embryo DRA generation rates ranged from 7 to 79%. A genome wide association study (GWAS) was conducted to search for common alleles in lines displaying highest rates, but only several genetic polymorphisms with weak association were identified – their manipulation to increase the efficacy of fruitfly drives would create limited improvements (Champer et al. 2019). Clearly, before DA design, studies into genetic diversity extent in target populations, both within target loci and genome wide, is a necessity to prevent DRA generation from natural genetic variation.

c. Non-random Mating

The only non-molecular resistance mechanism briefly discussed in literature is a population's potential for inbreeding behaviour (Zentner & Wade, 2017). Inbreeding would reduce chances of wildtype individuals mating with drive individuals, diminishing the frequency of drive heterozygotes, and therefore could emerge as an evolutionary response to blocking suppression drives (Drury et al. 2017). Bull (2017) and Bull et al. (2019) observed high rates of inbreeding in response to laboratory suppression drives, and Zentner & Wade (2017) modelled that moderate inbreeding can rapidly remove a C-GD from wild populations. This is a realistic impediment to drive success; high inbreeding rates have been observed in relevant wild populations (for example, mice and mosquitos) (Zentner & Wade, 2017), and no feasible genome engineering has been proposed to solve it.

2. Modelling Drive Resistance

C-GD modelling in wild populations allows resistance emergence to be quantified, identifying shortcomings in DA design. Models show evolutionary instability of applied first-generation C-GDs, with drive-resistant populations rebounding after a few years due to DRA emergence (Marshall et al. 2017). DRA generation rates must be orders of magnitude smaller than those observed in past laboratory suppression trials for a drive to achieve successful population suppression; mosquito DRA generation rates must lower from 10^{-1} (Hammond et al. 2016) to 10^{-5} to prevent a population of 10,000 rebounding on drive subjugation (lowering further for larger populations) (Noble et al. 2017, Marshall et al. 2017). Unckless et al. (2017) modelled pre-existing natural variation, de novo mutation and NHEJ generated DRAs in insect populations, highlighting successful drive application if (i) NHEJ is suppressed, or (ii) DA fitness cost is less than DRA fitness cost. These findings fuelled a second wave of DA design targeting NHEJ and DRA fitness costs, discussed in sections 3.a. and 3.c.

However, assumptions and limitations to C-GD models thus far include exclusion of inbreeding subpopulations and trans-resistance effects due to natural genetic diversity (models being restricted to cage trial data) (Unckless et al. 2017, Drury et al. 2017). As discussed, these aspects can greatly influence the fate of released C-GDs, hence their incorporation into future models is vital for accurate prognostication.

3. Re-designing Drive Constructs

Efforts to resolve limitations of first-generation C-GDs in real-world application have been made in recent DA design. These include (i) attempting to suppress NHEJ occurrence with higher promoter specificity or repression of NHEJ machinery genes, (ii)

multiplexing gRNAs to render NHEJ events as unproblematic, (iii) DRA evolutionary stability reduction with highly conserved gene targeting, and (iv) overcoming target site genetic diversity with alternative nucleases.

a. NHEJ Prevention

i. New promoters with higher temporal specificity

The standard *vasa D. melanogaster* promoter has induced expression of Cas9 post-fertilisation in multiple studies (Gantz & Bier, 2015, Hammond et al. 2016, Champer et al. 2018). Therefore, a search for more stringent germline-specific promoters that drive expression only in a window of HDR:NHEJ peak has been undertaken since (Champer et al. 2018). Three further *D. melanogaster* promoters - *nanos*, *BiC* and *63U* - have been observed to lower NHEJ-generated DRA rates through somatic Cas9 expression reduction (Champer et al. 2018, Kandul et al. 2019). As promoter behaviour can vary between species, Hammond et al. (2018) investigated *A. gambiae* promoters; *zpg* driven constructs showed most significant NHEJ-generated DRA rate reduction, suggesting its desirable temporal specificity for future mosquito drives.

Cas9 protein maternal deposition from the oocyte is sufficient for somatic homing (Champer et al. 2017, Kandul et al. 2019). Champer et al. (2018) proposed a strategy for dominant maternal effect prevention as an exclusively male-germline homing DA; as sperm do not transmit Cas9 protein in significant quantities to zygotes, this should drastically reduce early-embryonic DRA generation, but is yet to be designed. Hammond et al. (2018) exemplified Cas9 maternal deposition control through promoter choice: while *nanos*::Cas9 and *vasa2*::Cas9 show maternal deposition, *zpg*::Cas9 does not. A narrower spatiotemporal window for Cas9 translation and cleavage in the germline, in which HDR:NHEJ is highest, is a proposed cause (Hammond et al. 2018). Other

promoters with similar desirable windows may exist (perhaps species-specific); henceforth, future research should be directed to their search.

ii. NHEJ machinery gene repression

Genes involved in NHEJ machinery, or associated with decreasing HDR:NHEJ ratios, have been suggested as repression targets in drive organisms; this would be complimented by enhancement of genes involved in HDR machinery, or generation of long overhangs at DSBs to increase HDR:NHEJ (Friess et al. 2019). This could involve a dCas9 activation system in the DA that targets and regulates expression of such genes. Champer et al. (2019) identified associated genetic loci with lower HDR:NHEJ in *D. melanogaster* drive organisms – further GWAS research may identify common elements for future targeting to reduce resistance rates from trans-resistance effectors.

b. gRNA Multiplexing

Multiple gRNAs in DAs to target adjacent sequences within the same wildtype allele (*Figure 6*) has been a major proposal for tackling resistance emergence in C-GDT literature. Modelling shows its promise for resistance reduction in wild populations (Unckless et al. 2017, KaramiNejidRanjibar et al. 2018), with DRA fixation exponentially decreasing with DA gRNA number: calculations suggest a necessary four gRNAs to successfully suppress an insect population of 10 billion, despite DRA-generation rates of $>10^{-1}$ (Marshall et al. 2017). However, Marshall et al. (2017)'s *D. melanogaster* multiplex gene editing with two gRNAs yielded efficiency results approaching 100% for gRNA1, but 60 to 86% for gRNA2; if efficiency rates decrease for every DA incorporated gRNA, initial modelling is inaccurate.

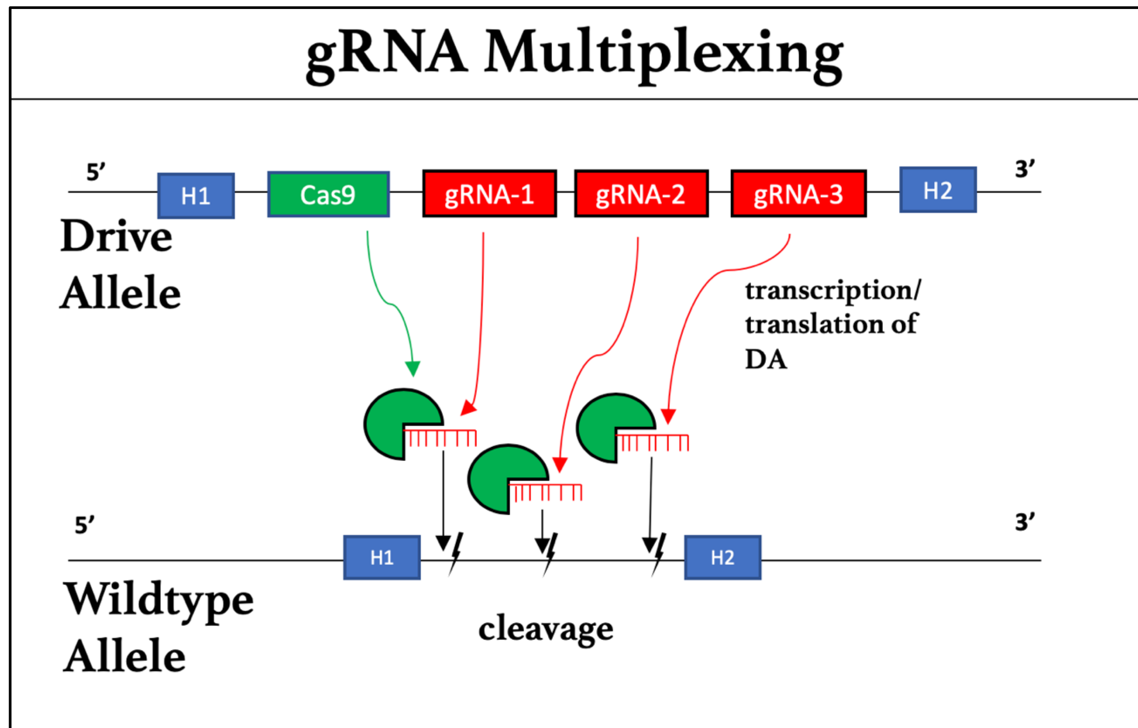

Figure 6: *gRNA multiplexing within a DA. It offers two benefits compared to utilising single gRNAs: (i) if the DNA between two target sites is lost due to simultaneous cleavage, it will create a very probable gene knock-out, dramatically increasing the fitness cost of an NHEJ-repaired wildtype, and (ii) all gRNA sites targeted must mutate for DRA generation, otherwise successful DA homing will occur at any remaining wildtype cleavage sites (Noble et al. 2017).*

Champer et al. (2018) demonstrated the first 2-gRNA multiplexing drive in *D. melanogaster*, which exhibited higher conversion rate and lower DRA-generation rate than a 1-gRNA DA. However, both improvements were lower than expected when assuming independent rates at each cleavage site in modelling, perhaps due to Cas9 protein saturation in the immediate environment, or homologous arms being further apart (homologous sequences sandwich all gRNA targets) (Champer et al. 2018).

Other issues with the multiplexing approach to tackling resistance rates in drive populations include (i) difficulty in finding multiple, genome unique gRNA target sites within a limited sequence size in genetically diverse populations (for example, *A.*

gambiae) (Champer et al. 2018), and (ii) increase in incomplete HDR rate with increasing DA size (Marshall et al. 2017).

c. DRA Fitness Cost Maximisation

By targeting conserved sequences in homing, DRAs are less likely to propagate, due to increased chance a mutated target sequence will be deleterious to the organism (Esvelt et al. 2014). Kyrou et al. (2018) drove a DA targeting the highly conserved *doublesex* gene in order to suppress a *D. melanogaster* laboratory population via female infertility, which reached 100% prevalence within 11 generations, causing population collapse before selection of DRAs occurred. However, if a drive's purpose is to target a less conserved/essential gene (e.g. a facilitating pathogen vector gene), then its predetermination rules out the possibility of utilising this resistance control mechanism (Unckless et al. 2017).

A theoretical drive proposal to overcome this issue is one which includes a complete genetically-recoded version of the cleaved gene (Esvelt et al. 2014). In 2019, Oberhofer et al. designed a novel 'Cleave and Rescue' (ClvR) DA that utilises this proposal by cleaving an essential, highly conserved gene, and providing a recoded, fully-functional sequence of the gene resistant to further cleavage. This was exampled by targeting *tko* in *D. melanogaster*, a lethal cleavage event unless the break is repaired by HDR using the DA as template (containing a *Drosophila vitilis tko* 'rescue' gene whose target site cannot be recognised by the drive gRNA-Cas9) (Oberhofer et al. 2019).

This ClvR system can be incorporated into a standard DA, so it targets both an essential gene, and an effector gene dependent on the drive's purpose (Esvelt et al. 2014). It involves utilising two gRNAs targeting two genes, with DA homing occurring at the essential gene locus (Esvelt et al. 2014, *Figure 7 (A)*). This design was first demonstrated

in two non-essential genes with 100% penetrance in *D. melanogaster;* however, as neither target was essential, DRA generation was observed in following generations, limiting this study's practicability (Kandul et al. 2019). Noble et al. (2017) modelled the design which exhibited vastly improved evolutionary stability of the DA in wild populations, highlighting the need for future studies to trial its full purpose, by targeting one essential gene and observing the decrease in NHEJ-generated DRA rate.

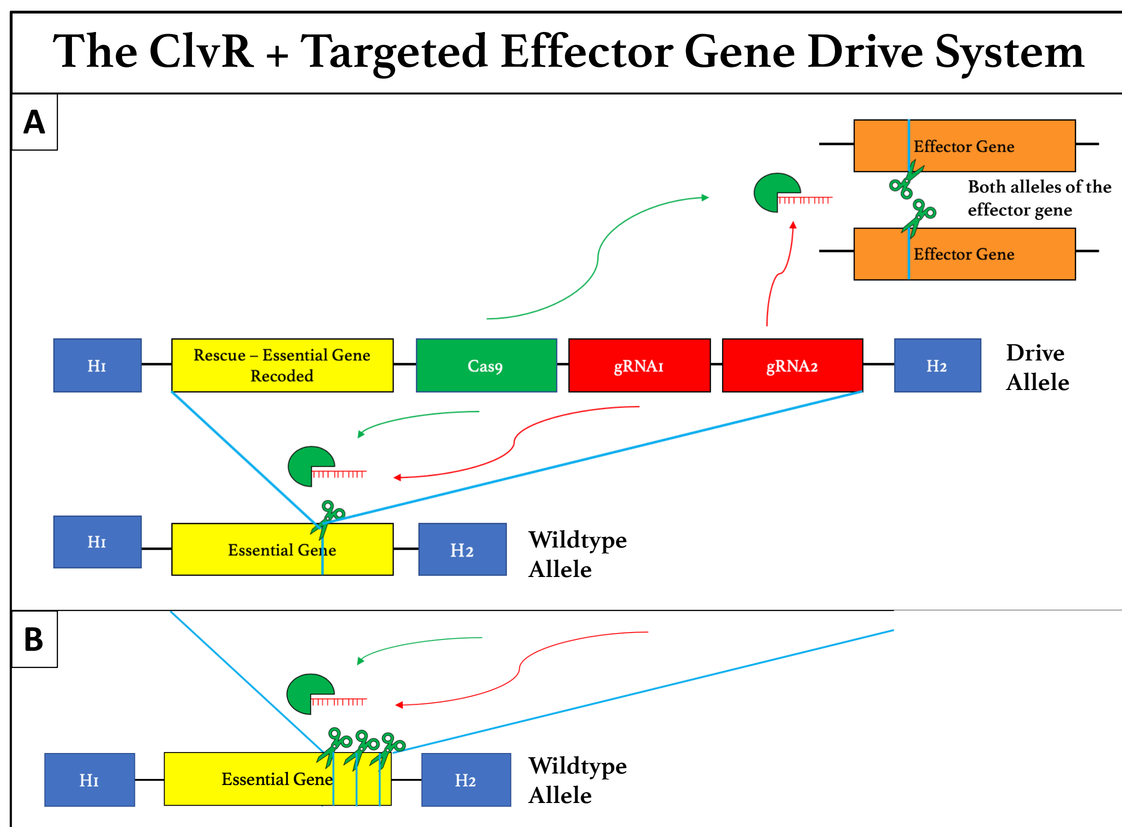

Figure 7*: (A) The ClvR + targeted effector gene drive system. Two separate gRNAs guide Cas9 to (i) an effector gene in which the drive's purpose is to knockout, and (ii) an essential gene whose function would be rescued by homing of the DA into the cleavage site via HDR. It could also include a payload-gene to be inserted at an effector site, for drives wishing to propagate a certain trait through a population. (B) The ClvR + targeted effector gene drive system combined with multiplexing gRNAs for adjacent target sequences in the essential gene.*

While the ClvR approach can significantly reduce DRA generation, pre-existing DRAs and de novo point mutations that confer resistance of essential genes to cleavage (be them

rare) would impede the rescue system's success (Oberhofer et al. 2019). Therefore, drive trials conducted on more genetically diverse populations are necessary to understand its full effect on DRA generation rates.

d. Alternative Nucleases

Recent developments in nuclease engineering for use within the CRISPR system as alternatives to the canonical SpCas9 may offer solutions to resistance emergence due to natural genetic variation within a population's target sequence (Hammond & Galizi, 2017). While SpCas9 cleaves only if a PAM (protospacer adjacent motif) sequence of 5'-NGG-3' is immediately after the gRNA complimentary sequence, alternative nucleases have been engineered to recognise different PAMs; this has enabled gene editing at a wider set of genomic locations (Fonfara et al. 2016, Hirano et al. 2016). By studying the genetic variation existing at a target sequence in the relevant wild population, a nuclease most suitable (whose PAM sequence is conserved in all existing variants) to successful homing can be chosen (Drury et al. 2017).

e. Working with Resistance Generation

As DRAs outcompete high fitness cost DAs through natural selection, there is a lag between the DA spreading to high frequencies at super-Mendelian rates, to when a DRA purges it from the population (Zentner & Wade, 2017). This lag period (variable and dependent on target population features, such as life cycle and mating behaviours) could be incorporated into a drive intended for short-term population transformations (Zentner & Wade, 2017). For example, spreading anti-plasmodium effector genes through *A. stephensi* (Gantz et al. 2015) only needs to reach high frequency for a low number of generations in order to accommodate local virus elimination (Li et al. 2020). In addition, if DRAs expel the DA from a target population before its purpose is achieved, a genetic

survey of DRAs evolved could be used to design an altered drive with Cas9 re-recognition of the mutated target site (Unckless et al. 2017).

Alternatively, the evolution of resistance could be embraced as a mechanism for controlling propagation; for example, intentional release of DRAs into a population harbouring a DA at high frequencies could provide regulation of a drive's spatiotemporal transmission (Unckless et al. 2017).

4. A Combinatorial Approach

A combinatorial approach incorporating methods discussed to reduce, and/or work with the evolution of resistance to C-GDs released into wild populations will prove most effective in achieving C-GDT potential. Combining multiplexed gRNAs and ClvR in the same DA has been evidenced for maximal drive effectiveness (Esvelt et al. 2014, Champer et al. 2016, Noble et al. 2017, *Figure 7 (B)*); this would reduce both NHEJ-generated DRAs (largely due to targeting essential genes), and DRA generation due to natural genetic variation at the target site (largely through multiplex gRNA targeting).

In addition to this pairing, DA promoters, nucleases and target sequences should be tailored to the target wild population; for example, at present, the *zpg* promoter is most suitable for DAs designed to propagate in *A. gambiae* (Hammond et al. 2018). In addition, by extensive genomic sequencing of wild populations, the most appropriate, highly conserved target sequences can be found (Ming et al. 2019). Furthermore, specifying the best promoter:target site combinations can be facilitated by assays such as that developed by Hammond et al. (2018). As DRA rates have been evidenced as partially locus dependent (Champer et al. 2018), identifying the best genomic locations for a specific promoter to drive expression by assaying to find sites of least resistance formation would

improve resistance rates further (Hammond et al. 2017). Moreover, Cas9 pairing with multiple gRNAs can limit target choice through PAM requirements – optimal nuclease selection dependent on desired target sequences is also crucial to maximising DA success.

The two mechanisms of resistance generation that a combinatorial DA does not address, however, are incomplete HDR events and inbreeding behaviour of the target population. Both present obstacles to a drive release free of DRA out-competition. As incomplete HDR-generated DRAs become more common as DA size increases (Unckless et al. 2017), and combinatorial DAs require many more components, incomplete HDR events may become frequent enough to prevent drive propagation. A viable solution to both inbreeding and incomplete HDR events has yet to be demonstrated, and necessitates future focus.

CONCLUSIONS AND CONTEMPLATION

C-GDT is currently in a most exciting period of development, between initial proof-of-concept studies and real-world application. This review has explored the field's recent shift from basic theoretical models and DA constructs to second-generation design solutions in order to address resistance emergence limitations in drive-subjected populations. Overcoming the impeding effects of DRAs will become a reality through the future combining of multiple novel DA design aspects: (i) species-specific promoter selection for optimal HDR:NHEJ, (ii) multiplexing gRNAs targeting adjacent sequences within an essential gene in a ClvR + targeted effector drive to increase the fitness cost of any DRAs generated, and (iii) GWAS methodology utilisation to study direct and background natural genetic variation within a wild population, in order to choose optimal combinations of target sequences, promoters and Cas9 variants that are unique to the

population of interest. On top of this, DRA generation could be utilised as a mechanism of controlling drive propagation by tailoring the DA purpose to be attainable before DRA out-competition.

However, uncertainties still remain: (i) inbreeding is likely to reduce drive propagation, hence prior study of target population mating behaviours is necessary before application, (ii) incidences of incomplete HDR will increase with larger, second-generation DAs, which remains unaddressed, (iii) the mechanisms by which background genetic variation in wild populations influence drive efficiency rates and DRA generation are complex, population-specific and understudied. These are therefore areas of necessary focus before future drive projects are approved. Moreover, CRISPR editing utilisation is currently limited to certain taxa; the development of new Cas9 platforms for use in novel target species, (for instance, the CopyCat cassette as a mammalian drive platform (Grunwald et al. 2019, *Figure 1*)), must continue in order for GDT to reach its full potential.

It is important to note that creating spatiotemporally controllable C-GDs that can be tailored to specific goals is also a requirement for real-world use, to prevent uncontrolled drive transmission throughout populations. C-GD regulation mechanisms explored so far include Cas9 protein regulators (for example, with Cas9 fusion proteins, or control of its nucleocytoplasmic localisation), temporal confinement mechanisms (for example, with split gene drive systems) and replacement strategies (by secondary drive release) (Zentner & Wade, 2017, Noble et al. 2019).

Addressing these knowledge gaps will define the final step to successfully utilising this new technology, providing cost-effective solutions to many extensive health, ecological and environmental issues.

REFERENCES

Bull, J. J., Remien, C. H., & Krone, S. M. (2019). Gene-drive-mediated extinction is thwarted by population structure and evolution of sib mating. *Evolution, medicine, and public health*, *2019*(1), 66-81.

Burt, A. (2003). Site-specific selfish genes as tools for the control and genetic engineering of natural populations. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270(1518), 921-928.

Champer, J., Buchman, A., & Akbari, O. S. (2016). Cheating evolution: engineering gene drives to manipulate the fate of wild populations. Nature Reviews Genetics, 17(3), 146.

Champer, J., Liu, J., Oh, S. Y., Reeves, R., Luthra, A., Oakes, N., ... & Messer, P. W. (2018). Reducing resistance allele formation in CRISPR gene drive. *Proceedings of the National Academy of Sciences*, *115*(21), 5522-5527.

Champer, J., Wen, Z., Luthra, A., Reeves, R., Chung, J., Liu, C., ... & Clark, A. G. (2019). CRISPR gene drive efficiency and resistance rate is highly heritable with no common genetic loci of large effect. *Genetics*, *212*(1), 333-341.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., ... & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science, 339(6121), 819-823.

DiCarlo, J. E., Chavez, A., Dietz, S. L., Esvelt, K. M., & Church, G. M. (2015). Safeguarding CRISPR-Cas9 gene drives in yeast. Nature biotechnology, 33(12), 1250.

Drury, D. W., Dapper, A. L., Siniard, D. J., Zentner, G. E., & Wade, M. J. (2017). CRISPR/Cas9 gene drives in genetically variable and nonrandomly mating wild populations. Science advances, 3(5), e1601910.

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

Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E (2016). "The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA". *Nature*. 532 (7600): 517–521.

Friess, J. L., von Gleich, A., & Giese, B. (2019). Gene drives as a new quality in GMO releases—a comparative technology characterization. PeerJ, 7, e6793.

Gantz, V. M., & Bier, E. (2015). The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. Science, 348(6233), 442-444.

Gantz, V. M., Jasinskiene, N., Tatarenkova, O., Fazekas, A., Macias, V. M., Bier, E., & James, A. A. (2015). Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. Proceedings of the National Academy of Sciences, 112(49), E6736-E6743.

Grunwald, H. A., Gantz, V. M., Poplawski, G., Xu, X. R. S., Bier, E., & Cooper, K. L. (2019). Super-Mendelian inheritance mediated by CRISPR–Cas9 in the female mouse germline. Nature, 566(7742), 105.

Hammond, A. M., & Galizi, R. (2017). Gene drives to fight malaria: current state and future directions. Pathogens and global health, 111(8), 412-423.

Hammond, A.M, Galizi, R., Kyrou, K., Simoni, A., Siniscalchi, C., Katsanos, D., ... & Burt, A. (2016). A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. Nature biotechnology, 34(1), 78.

Hammond, A. M., Kyrou, K., Gribble, M., Karlsson, X., Morianou, I., Galizi, R., ... & Nolan, T. (2018). Improved CRISPR-based suppression gene drives mitigate resistance and impose a large reproductive load on laboratory-contained mosquito populations. bioRxiv, 360339.

Hirano H, Gootenberg JS, Horii T, Abudayyeh OO, Kimura M, Hsu PD, Nakane T, Ishitani R, Hatada I, Zhang F, Nishimasu H, Nureki O (2016). "Structure and Engineering of Francisella novicida Cas9". *Cell*. 164 (5): 950–961.

Kandul, N. P., Liu, J., Buchman, A., Gantz, V. M., Bier, E., & Akbari, O. S. (2019). Assessment of a split homing based gene drive for efficient knockout of multiple genes. *bioRxiv*, 706929.

KaramiNejadRanjbar, M., Eckermann, K. N., Ahmed, H. M., Dippel, S., Marshall, J. M., & Wimmer, E. A. (2018). Consequences of resistance evolution in a Cas9-based sex conversion-suppression gene drive for insect pest management. Proceedings of the National Academy of Sciences, 115(24), 6189-6194.

Kyrou, K., Hammond, A. M., Galizi, R., Kranjc, N., Burt, A., Beaghton, A. K., ... & Crisanti, A. (2018). A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nature biotechnology*, *36*(11), 1062.

Li, M., Yang, T., Kandul, N. P., Bui, M., Gamez, S., Raban, R., & Bennett, J. (2020). Development of a confinable gene drive system in the human disease vector Aedes aegypti. eLife, 9.

Marshall, J. M., Buchman, A., & Akbari, O. S. (2017). Overcoming evolved resistance to population-suppressing homing-based gene drives. Scientific reports, 7(1), 3776.

McVey, M., & Lee, S. E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends in Genetics, 24(11), 529-538.

Noble, C., Min, J., Olejarz, J., Buchthal, J., Chavez, A., Smidler, A. L., ... & Esvelt, K. M. (2019). Daisy-chain gene drives for the alteration of local populations. Proceedings of the National Academy of Sciences, 116(17), 8275-8282.

Noble, C., Olejarz, J., Esvelt, K. M., Church, G. M., & Nowak, M. A. (2017). Evolutionary dynamics of CRISPR gene drives. Science advances, 3(4), e1601964.

Oberhofer, G., Ivy, T., & Hay, B. A. (2019). Cleave and Rescue, a novel selfish genetic element and general strategy for gene drive. *Proceedings of the National Academy of Sciences*, *116*(13), 6250-6259.

Raban, R. R., Marshall, J. M., & Akbari, O. S. (2020). Progress towards engineering gene drives for population control. Journal of Experimental Biology, 223(Suppl 1).

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nature protocols, 8(11), 2281.

Shapiro, R. S., Chavez, A., Porter, C. B., Hamblin, M., Kaas, C. S., DiCarlo, J. E., ... & Wang, Y. (2018). A CRISPR–Cas9-based gene drive platform for genetic interaction analysis in *Candida albicans*. Nature microbiology, 3(1), 73.

Unckless, R. L., Clark, A. G., & Messer, P. W. (2017). Evolution of resistance against CRISPR/Cas9 gene drive. Genetics, 205(2), 827-841.

Walter, M., & Verdin, E. (2019). Viral gene drive in herpesviruses. bioRxiv, 717017.

Zaidi, S. S. E. A., Mahfouz, M. M., & Mansoor, S. (2017). CRISPR-Cpf1: a new tool for plant genome editing. *Trends in plant science*, *22*(7), 550-553.

Zentner, G. E., & Wade, M. J. (2017). The promise and peril of CRISPR gene drives: genetic variation and inbreeding may impede the propagation of gene drives based on the CRISPR genome editing technology. BioEssays, 39(10), 1700109.

Zhang, F., Wen, Y., & Guo, X. (2014). CRISPR/Cas9 for genome editing: progress, implications and challenges. Human molecular genetics, 23(R1), R40-R46.