1 What are Gene Drive Organisms and how do Gene Drives work?

Synthetic gene drives are a new form of genetic engineering that provide the tools for permanently modifying or potentially even eradicating species in the wild. Unlike the previous GMOs, Gene Drive Organisms (GDOs) are not meant to stay where they are released, but instead are designed and purpose-built to spread and to drive their modified genes far and wide. One idea for example is to push wild populations back and replace them with specially designed populations that additionally will cause offspring to die if neighbouring populations interbreed with each other. The intention for synthetic gene drives is to rapidly alter the genetic make-up of wild populations, with the aim of either changing certain of their characteristics or eliminating them. An example would be using gene drive technology to genetically prevent mice from having any daughters, then releasing such gene drive mice into an island ecosystem so that they breed with the wild population. Producing only male offspring the whole population eventually collapses and eradicate all the mice on that island. The list of targets is manifold - fruit flies, mosquitoes, snails, rats, mice, plants, feral cats, possums – and new proposals appear frequently.

Whilst primarily aimed at organisms that are perceived as a problem by some parts of human society, whether or not they have been classified as agri-cultural pests, disease-spreading insects or invasive alien species, the fact is that the technology could be applied much more widely, and indeed could be weaponised or used for industrial sabotage.

A gene drive target may be any organism that sexually reproduces and that does so with reasonable frequency. Thus, gene drives are specific to organisms that reproduce through a process called meiosis, and only work as intended in such organisms. Meiosis is a dedicated and particular form of cell division that ultimately generates non-identical sex cells. It is common in eukaryotes (higher organisms) but absent from prokaryotes, which thus discounts all bacteria and archaea. Moreover, the genes that are subject to alteration by engineered drives are those situated in the nucleus, and not genes elsewhere, such as in mitochondria. As noted above, they will only work when the organism reproduces via meiosis and not if for example it does so instead via vegetative reproduction, of which many kinds of organisms are capable, including some plants and probably all fungi.

Looking specifically at organisms that are diploid, i.e. whose nucleic DNA is made up of two sets of genetic material with one set from each parent, this information will be mixed and halved before it is passed on through sex cells e.g. sperm or egg cells.

---

1 The nucleus is only found in eukaryotic organisms. It is a compartment within a eukaryotic cell that holds the chromosomal DNA and is surrounded by a porous membrane. Other compartments holding DNA are for example mitochondria or chloroplasts.
In that process, the DNA of the parents gets redistributed randomly with the only proviso being that it will have to result in a complete set again.

Whilst normal genes have this 50% chance of inheritance, gene drive elements have changed the odds in their own favour. Some call this “super-Mendelian” inheritance (Chevalier and Stoddard 2001; Hammond et al. 2016; Grunwald et al. 2019). When genetically engineering an organism with special or specifically designed gene drives, these may force their own inheritance to a level of 80% or nearly 100% (see Figure 1). This is particularly the case with CRISPR/Cas9-based gene drives, which have been reported to resemble a “mutagenic chain reaction” (Gantz and Bier 2015). Depending on whether they have been designed as a population suppression drive (to reduce or eliminate a target species), or as some kind of modification drive (to spread a specifically designed or desired trait), the release of GDOs with such gene drives may - and is generally intended to - either lead to the collapse of a population or to a change of traits and characteristics throughout the entire population.

For example, a population suppression drive could be a gene drive that will spread female sterility. In theory, when passed on to each and every offspring, and also carried and thus spread by each of the males, the wild populations would be quickly reduced and eventually collapse. However, in reality - and depending on gene drive type and species - there might be significant practical difficulties, as well as significant unintended and unpredicted consequences.

There are various genetic mechanisms that have evolved in nature that will result in an increased inheritance rate of specific genetic elements, genes or even whole chromosomes, irrespective of whether their presence negatively impacts the fitness of the organism. Whilst initially not referred to as ‘gene drives’ or even ‘drives’, the term has now come to mean a whole broad spectrum of mechanisms, many of which are being investigated, proposed or developed for application as synthetic (or engineered) gene drives.

Gene drives are currently defined as systems where genetic elements have a biased inheritance trait, irrespective of a benefit or lack of benefit for the organism. They can be used to carry additional genes often referred to as “payload genes” and spread these and their traits throughout a population.

There is a further aspect to synthetic gene drives that is of crucial importance. Deploying GDOs is also a form of ecosystem engineering, either as an intentional aim, or an accidental and unintentional consequence of suppressing or eradicating one or more species, or of the intentional modification of biological functions and characteristics of one or more species. The risks and potential serious negative consequences of such ecosystem engineering cannot be addressed in this chapter, yet this aspect needs to be born in mind when discussing the science and technologies.

2 Short historical background

The idea behind using gene drives to suppress or modify whole populations, especially those regarded as pests, by genetic control methods and strategies is not new; it is only the technical capabilities scientists have very recently developed that are. As early as the 1940s, scientists such as Serebrovskii (1940) and Vanderplank (1944) proposed redirecting an insect’s own genetic system (against itself), in order to either destroy insect populations or to make them less destructive to human endeavours.

2 Different names are being used for such genes that are being linked to and transported by the engineered gene drives: e.g. payload genes (Champer, Buchman, and Akbari 2016), effector genes (Sinkins and Gould 2006; Marshall and Akbari 2018) or simply cargo (Hay et al. 2010).
Figure 1: Synthetic gene drives & super-Mendelian inheritance: in a population, a mutation that has no fitness benefit will quickly disappear with the Mendelian inheritance rate of 50%. On the other hand, a synthetic gene drive system with a near 100% inheritance rate ensures the spread of a trait, even if that trait has a clear fitness cost.
such as growing crops (for review see Gould and Schliekelman 2004).

Serebrovskii’s theory involved reducing the fitness of insect populations or causing sterile offspring by releasing large numbers of mutated strains, in particular strains with chromosomal translocations (the exchange of whole segments between different chromosomes). This concept was revived by Curtis in the late 1960s (Curtis 1968). Vanderplank’s work was different in that he suggested and later demonstrated for tsetse flies in Tanzania, that releasing a closely related species or subspecies that will mate with the target species would lead to reduced viability or sterility in the resulting hybrids (Vanderplank 1947; 1948). This approach is now referred to as ‘hybrid sterility’.

In a separate approach, Knipling spearheaded the sterile insect technique (SIT), which works by releasing vast quantities of sterile males, with sterility caused by chromosomal abnormalities induced by radiation (Knipling 1955). A massive and successful screwworm fly SIT-eradication program was carried out in the US, starting in Florida in 1957, succeeding in 1966, and begun in Mexico in 1972 and succeeding in 1991 (Gould and Schliekelman 2004).

Other research projects also got under way, seeking to affect at least 31 insect species in specific regions, with particular focus on agricultural pests (e.g. fruit flies, bollworms, boll weevils) and vectors of diseases (e.g. different species of mosquitoes and tsetse flies) (LaChance 1979 in Gould and Schliekelman 2004).

By the early 1980s, the ‘golden era’ of research on ‘autocidal control and strain replacement’ had come to an end (Gould and Schliekelman 2004), and funding was drying up, partly due to the lack of further such ‘loud’ and ‘easy’ successes. Theoretical population genetics intended to design genetic control programs were equally put back. There were a variety of reasons for this, including lack of function outside laboratories3 (see also Box 1), funders losing interest, and the need for large governmental infrastructure.

In short, they couldn’t quite make it work, although as so often happens with new genetic and technology approaches, there was a great deal of talk and publicity, optimistic claims that brought a great deal of funding and prestige, but then calmed down when results were less than had been expected.

Box 1: Density-dependence: Problems not just for SIT (sterile insect techniques)

Applying SIT to mosquitoes is complicated by what scientists call “density-dependent” effects on mosquito populations. The size of a population of mosquitoes does not depend only on how well the mosquitoes reproduce, but also on other factors, such as competition for food between larvae and for breeding sites. Reducing reproductive fitness may have little effect if the size of the mosquito population is limited mainly by these factors, rather than by its ability to reproduce. Density-dependent effects mean that reducing the numbers of mosquitoes that breed successfully can sometimes have little effect on total numbers of adult mosquitoes, and paradoxically might sometimes even increase populations: for example, because reducing breeding success also reduces competition between larvae for resources, resulting in increased survival rates or a rebound in numbers. Density-dependent effects can influence the current generation of mosquitoes or only affect future generations (delayed density-dependent effects) (Gould and Schliekelman 2004; Juliano 2007; Walsh et al. 2011; 2012).

Excerpt taken from p.2 (GeneWatch_UK 2012)

These early efforts, often referred to as ‘classical genetic pest manipulation,’ (Gould and Schliekelman 2004), basically showcase the history of the broad concept of using genetics to exert control over or destroy undesired insect populations. Whilst linked to the same aim of suppression or replacement of wild populations, there had been no ‘drive’ element, no notion of altering the evolution of entire species in any of the examples or strategies mentioned above. These relatively recent strategies required repeated releases of such altered organisms

---

3 These include issues of mass rearing, sterilisation methods, release methods (e.g. the unintended release of females alongside, who are often not as easily sterilised as males; or greatly reduced male fitness, or loss of fitness over time); but also issues of density dependent populations in the context of SIT.
Chapter 1: What are Gene Drives?

on a large scale; but none of them were capable of being ‘run-away’ or true ‘gene-drive’ technologies, nor were they intended or designed to actively and aggressively spread into all future generations and neighbouring populations.

However, by the early 1990s, with the advent of genetic engineering and the ability to construct artificial genes, recombining different DNA sequences and inserting novel gene sequences into an organism became more commonplace. The possibility of using various ‘drive elements’ and ‘drive mechanisms’ to actively spread genetic traits in a population seemed to open up. One idea was, for example, to use parasite resistance genes in mosquitoes in order to stop the spread of pathogens, like the malaria-causing pathogen *Plasmodium*. Drive elements, like the ‘piggy-back transposon,’ a mobile and highly active genetic element with the ability to carry extra DNA and insert it into host DNA, were also utilised and developed for producing transgenic insects, as well as for modelling the spread of infertility (Ribeiro and Kidwell 1994).

Research intensified around drive elements, that is to say genetic sequences that had an increased inheritance ratio above the standard 50%. Sometimes referred to as ‘selfish genes’, these drive elements include transposable elements (TEs), first discovered in maize, as well as homing endonuclease genes (HEGs). As transposable elements had also been found in *Drosophila* and other insects, they were the first to be proposed as drive elements in the early 1990s (Curtis 1992; Kidwell and Ribeiro 1992). Burt, however, advocated the use of site-specific selfish genes as drive elements (such as the homing endonuclease genes), to be able to target and knock out essential host genes, and in this way to eradicate entire populations (Burt 2003).

There are a number of genetic elements, as well as specific genetic mechanisms, that can give rise to ‘drive’, that is, to increase the rate of inheritance of a specific trait. Some of these elements and mechanisms have been investigated or are being developed as gene drive systems, such as for example MEDEA (Maternal Effect Dominant Embryonic Arrest), underdominance, meiotic drive, t-complex in mice, and X-shredder. They will be introduced and addressed in Section 4 of this chapter.

3 The breakthrough: the CRISPR/Cas9 or RNA guided Gene Drive system

Whilst development of Gene Drives and GDOs was moving relatively slowly, the advent of CRISPR/Cas9 in 2012 (Jinek et al. 2012) radically changed the pace of developments and advances, igniting a fevered push for application.

CRISPR/Cas9 has its origin in bacteria, where it was found to act as a natural defence system against viruses. Utilising this bacterial ability to recognise and cut up the DNA of its invaders, it was eventually developed as an easy-to-use ‘genome editing’ tool, designed to cleave (or break) a double-strand of DNA at a specific recognition site (Jinek et al. 2012). As such it is made up of two components: the CRISPR part is a single strand of RNA termed ‘single guide’ RNA (sgRNA), able to recognise a specific DNA sequence and to ‘guide’ the Cas9 protein to this location. The ‘CRISPR-associated’ protein 9 (Cas9) is an endonuclease, and is thus the part of CRISPR/Cas that will cause the DNA ‘double strand break’ (DSB) at the target site. In order for the sgRNA to recognise and bind to a DNA target site, the nucleotide sequence of both the RNA and the target DNA need to be near mirror images of each other.

---

4 CRISPR/Cas9 stands for: ‘clustered regularly interspaced short palindromic repeats’ / ‘CRISPR-associated’ protein 9
5 It is not tested or known whether this is the cause of CRISPR evolution in bacteria or indeed if there is another cause.
It is easy to make Cas9 cleave the DNA at a different site by altering the nucleotide sequence of the guide RNA. This is sometimes referred to as ‘programming’ a site-specific nuclease and requires detailed knowledge of the DNA sequence of an organism, since a target sequence should be unique so as to avoid cleaving the DNA unintentionally at multiple places. Different endonucleases have been identified and developed, like Cas12a (formerly Cpf1) which creates staggered ends of DNA strands rather than the blunt ends of Cas9. Others are also tested, to work under different conditions, though to do the same task, namely the cleavage (or ‘breakage’) of a DNA double-strand.

3.0.1 DNA repair mechanisms after double-strand breaks (DSBs)

In genome editing the major role, if not the only role, for site-directed nucleases (SDNs) is to create a DNA double strand break (DSB) at a specific location within the genome. A breakage will in turn induce the cell’s own native DNA repair, which has two main repair mechanisms available: the error-prone, non-homologous end joining (NHEJ), and the more specific homology-directed repair (HDR\(^6\)), which requires a template (see Figure 2).

DSBs are preferentially repaired through non-homologous end joining (NHEJ), whereby the ends of the broken DNA molecule will often be further processed, and sequence information can be lost or altered upon rejoining, making this an error-prone repair process (Wyman and Kanaar 2006). NHEJ thus often results in small insertions or deletions (indels), or in the substitution of a few nucleotides at or near the cutting site. Because the mechanisms behind this type of DNA repair are not yet fully understood, it cannot be controlled or predicted what exact type of DNA modification will occur when no external DNA template is supplied. The nature and detail of these small mutations are thus regarded as random.

Homology-directed repair (HDR) requires a repair template with extensive regions of homology to the DNA sequences neighbouring the breakage site. If such a template is present and if the HDR pathway is triggered, the DSB will be ‘repaired’ according to the sequence provided in the template, which might be small alterations or the insertion of longer DNA sequences, including whole genes. CRISPR/Cas-based gene drives rely on the actions of this pathway.

Which DNA repair mechanism will be triggered depends largely on the species and taxonomic groups, developmental stage of the organism, cell type, and presence of environmental factors. The predominant repair mechanism is the NHEJ pathway.

3.0.2 CRISPR/Cas variants

Other variants have been developed that will perform other tasks than the initial double strand break of Cas9 and Cas12a. Modifying the Cas9 endonuclease to cut only a single strand of DNA, such nickases (nCas9) are also used to initiate mutations or to create double strand breaks when used as a pair. A highly versatile variant is the deactivated dCas9, which can be fused with other enzymes such as deaminases, able to alter individual nucleotides, such as cytidine, and are now termed ‘base editors’. It can also be linked with gene activators or deactivators, which are in fact used in specific synthetic gene drive systems described in Section 4.1.2, underdominance, referred to as CRISPRa.

---

\(^6\) This repair system is given different names by different authors, e.g. homologous recombination (HR), homologous repair (HRI), homologous recombination repair (HRR), and homology dependent repair (HDR).
Chapter 1: What are Gene Drives?

Figure 2: Genome editing and the DNA repair pathways: Top row: After having constructed a site-specific nuclease (SDN) that will recognise a chosen DNA sequence, this SDN will create a DNA double-strand break at its target site. Second row: degradation of both DNA ends by endogenous enzymes may occur, enlarging the area of damage (pink letters above) before the cell’s own repair mechanisms gets triggered. Bottom half: Depending on circumstances, one of the following three actions may occur. (1) The error-prone ‘non-homologous end joining’ (NHEJ) repair pathway is triggered, resulting in smaller random mutations near or at the cleavage site. This is predominantly used to create gene knock-outs. (2) & (3): If repair templates are added, the homology directed repair (HDR) pathway may be triggered. For (2) this is a short template with a few mismatches intended to achieve small specific sequence alterations, either to ‘correct’ a gene or to ‘set’ specific mutations. (3) resembles the insertion of a longer DNA sequence at a pre-determined site, by supplying a template framed by DNA sections of high sequence homology with the site for insertion. Such inserts could be regulatory sequences or a gene coding for a protein.

Whilst NHEJ (1) is regarded as a routine application for many plant species, HDR (2 & 3) remains challenging, in particular for plants.

In the EU, for regulatory purposes, a classification has been suggested according to the intended outcomes of the actions of site-directed nucleases (SDN), i.e. gene disruption (1) gene correction with template (2) and gene addition (3). In the EU classification these are termed SDN1, SDN2 and SDN3 respectively (Lusser et al. 2012).

---

7 Site directed endonucleases used in genome editing include: Zinc Finger Nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs), meganucleases and CRISPR/Cas9. They have different modes of action and of sequence recognition.
3.1 Limitations and uncertainties of CRISPR/Cas

It is well established that CRISPR/Cas9 works by inducing a double strand breakage at a (genomic) DNA target site, with sufficient sequence homology to its own guide RNA8. After the repair there will commonly be the intended on-target effects, but additionally there is also evidence of unintended on-target effects, as well as unintended off-target effects.

**Unintended on-target effects:** Working with both mouse and human cell lines, researchers from the Sanger Institute, UK, for example, reported evidence of significant on-target mutations, such as large deletions -- of up to 9.5 kb -- and complex rearrangements around the DNA breakage site (Kosicki, Tomberg, and Bradley 2018). Additionally, they found mutations (deletions, rearrangements and even insertions) away from the target site, i.e. not physically linked to or running on from it. Whilst the implications of these specific and complex rearrangements have not been investigated, such rearrangements constitute a clear risk, as they can alter gene expression, give rise to further mutations during reproduction, as well as disable or alter the sequence of genes at the site of rearrangement. However, the same situation may also arise for unintended off-target sites, as the action of CRISPR/Cas9 would work under the same rules for both. Off-target sites though have not yet been investigated for complex rearrangements, a fact that needs urgent attention, given that the risks are likely to be the same.

Furthermore, a recent pre-publication is indicating that intended on-target indel mutations -- set by the error-prone NHEJ repair mechanism -- may have very unexpected and problematic consequences. Tuladhar investigated the consequences of intended knock-out mutations, in particular of frameshift mutations induced by indel mutations. The researchers looked at the processing of the resulting RNAs, their translation into proteins as well as the impact on gene regulation (Tuladhar et al. 2019). Pre-publishing in a not yet peer-reviewed form they reported: “By tracking DNA-mRNA-protein relationships in a collection of CRISPR/Cas9-edited cell lines that harbor frameshift-inducing INDELS in various targeted genes, we detected the production of foreign mRNAs or proteins in ~50% of the cell lines.” (Tuladhar et al. 2019, 1). The news here is the generation of new internal ribosomal entry sites (IRES) leading to the production of truncated proteins and the alteration of pseudo-mRNAs resulting in protein coding RNAs. Were these findings found to be common in CRISPR/Cas induced indel mutations, this would have serious implications for safety as well as predictability. These findings are however a reminder that CRISPR/Cas9 is a new technology that due to its ease of use, has found wide-spread application without the necessary time to establish all the consequences and risks of that use.

**Unintended off-target effects:** A problem that has already been long recognised is that of off-target effects. At the DNA level, off-target effects are those where the RNA-guided nuclease cuts at a site that is not the intended target site. This is thought to primarily happen at sites that are not identical but that have high sequence similarities to the guide-RNA. Experiments have shown CRISPR/Cas9 may cut DNA even with 2-3 nucleotide mismatches between the DNA sequence and the guide-RNA, albeit with lowered efficiency. There does not seem to be a hard rule as to how many nucleotide mismatches are tolerated by the HD repair mechanism, as this also depends on the species, cell type, the actual nuclease variant and the experimental conditions.

Whilst there is an increasing reliance on the use of algorithms to calculate and predict the potential off-target sites, according to the degree of homology (based on the number and position of mismatches), there is also increasing concern about this. In fact, the sole reliance on algorithms to accurately predict the potential off-target sites or regions for off- or on-target effects has come into question repeatedly, as only whole genome sequencing, an increasingly affordable technology, would be able to pick up some of the mutational effects observed. This does not only refer to extensive mutations delinked from the actual cutting site (Kosicki, Tomb-

---

8 And in the presence of a PAM site.
erg, and Bradley 2018), but also to the integration of vector backbone DNA derived from the plasmid used in the original transgene construct, and for example observed in genome editing experiments with oilseed rape (Braatz et al. 2017).

(Akcakaya et al. 2018) find that many studies reporting no or few off-target effects (mutations) will have failed to identify actual off-target effects due to the limitations of the “in silico” (i.e. computer modelling) predictions of potential off-target sites.

Since CRISPR/Cas has been found to cut sites with even seven mismatches, or to bind to sequences with as many as 9 consecutive mismatches, Chakraborty argues that restricting searches for potential off-target sites to 3 or 4 mismatches is failing to investigate properly (Chakraborty 2018, 226). He states: “In conclusion, the off-target problems associated with CRISPR-Cas have not been addressed conclusively, which does not bode well, since non-specificity is an intrinsic feature of CRISPR-Cas, evolved over billions of years—otherwise hyper-variable viruses would evade this microbial immune-system with ease and render it ineffective.”

It is important to bear in mind that any predictions of potential off-target sites require extremely good and accurate knowledge of the DNA sequence of an organism. This will be a real difficulty when dealing with wild and diverse populations, and the degree of variation present within a whole species (see Section 3.3, Limitations). Thus far, only laboratory data has been generated, and it can be anticipated from the findings, that unintended on-target effects as well as off-target effects will take place. This is a serious concern, as it adds additional risks to the release of GDOs into wild populations.

3.2 How does a CRISPR/Cas Gene Drive work?

When CRISPR/Cas is used as a homing endonuclease, it becomes a ‘drive element’. In this scenario, the whole CRISPR/Cas construct will need to be copied across into its own target site (see Figure 3). This can only work properly when the construct contains sequences at its outer borders that are homologous to those present next to the specific CRISPR/Cas target site. Once the construct is present on one chromosome it will produce the CRISPR/Cas molecule, i.e. the RNA-guided endonuclease, which will cleave the DNA at its target site in the parallel (homologous) chromosome. Once the target site is cut, the repair mechanism kicks in and uses the homologous chromosome as a repair template, in this case containing the CRISPR gene drive construct. With the homology-directed repair (HDR) mechanism activated, the construct gets copied into the target site, and thus ensures the ‘inheritance’ and spread of the construct.

The construct may or may not contain a ‘payload gene’ (Champer, Buchman, and Akbari 2016), opening up the possibility of introducing new genes and traits into the target organism (see Figure 3). One category of payload genes aimed for are so-called refractory genes, that will stop a disease from spreading or being transmitted, such as genes for malaria resistance (Gantz et al. 2015).

The adaptation of CRISPR/Cas9 as a drive element was suggested in 2014 (Esvelt et al. 2014), and is based on Burt’s proposal of using site-specific homing endonuclease genes (Burt 2003). CRISPR/Cas9 gene drive systems were quickly put to test by different research teams, offering proof of principle in four different species. Gantz and Bier were the first with drosophila, entitling their finding aptly as a ‘Mutagenic Chain Reaction’ (Gantz and Bier 2015), followed by yeast (DiCarlo et al. 2015), Anopheles stephensi mosquitoes (Gantz et al. 2015), and Anopheles gambiae mosquitoes (Hammond et al. 2016). More recently mice were added to this list of proof of principle, though conversion rates were low (Grunwald et al. 2019). See also Table 1.

---

9 “To our knowledge, our report provides the first demonstration that CRISPR-Cas nucleases can robustly induce off-target mutations in vivo. Previous in vivo studies have reported no or very few off-target mutations, but used the cell-based ‘genome-wide unbiased identification of double-strand breaks enabled by sequencing’ (GUIDE-seq) method12–14 or other in silico approaches that have not been validated to effectively identify these sites in vivo (see Supplementary Discussion).” (Akcakaya et al. 2018, 419)
Chapter 1: What are Gene Drives?

Figure 3: The homing of a CRISPR/Cas gene drive: The CRISPR/Cas construct is composed of the gene sequence for the Cas protein and the sequence for the guide RNA (gRNA). Additionally there are flanking sequences that are identical to the sequences found on either side of the DNA cleavage site, and which are required for the homology-directed repair. (1) Once activated (depending on the promoter used) CRISPR/Cas molecules will be assembled, find the recognition site on the homologous chromosome and sever the DNA. (2) The areas of sequence homology align; i.e. the areas bordering the cleavage site align with the flanking sequences of the CRISPR/Cas construct. (3) In the process of homology-directed repair the CRISPR/Cas construct is copied across to the homologous chromosome. Left panel: Here the drive is meant to disrupt a gene, e.g. for female fertility or pesticide tolerance. Right panel: Here the construct carries an additional payload gene, e.g. for pathogen resistance.

3.3 Limitations, shortcomings and uncertainties of CRISPR/Cas Gene Drives

Here we summarise the limitations encountered, such as emergence and build-up of resistance, inefficiency, off-target effects, lack of specificity and inability to be recalled once released. These shortcomings are being clearly recognised not just by researchers, but also by funders like DARPA, which has developed a ‘Safe Genes Project’, not simply to get gene drives to work, but to find ways to counter or undo them. This will be detailed below.

3.3.1 Resistance

Resistance to CRISPR/Cas almost inevitably happens. Endonucleases recognise specific DNA sequences as their target sites. If the sequence of a target site changes for whatever reason, the nuclease will not recognise the target and thus will not or cannot cut. In the case of a CRISPR/Cas-based gene drive, this means the drive will be stopped. The organism with such an altered target site has become resistant to the gene drive. There are two sources for such altered target sites to occur: First, there is
natural variation of the target site sequence within a population; and second, new mutations arise due to the activity of the CRISPR/Cas-based gene drive itself.

Natural variation is a common phenomenon, yet there are genes with more highly conserved sequences. Early advocates of endonuclease-based gene drives thus proposed using such conserved genes as the target of choice (Burt 2003). ‘Highly conserved’ means that the DNA sequence of a gene (or the corresponding amino acid sequence) has remained the same over time on an evolutionary scale, and that it has not been changed by random mutations. Such genes are commonly essential genes. This strategy has recently been picked up by Kyrou et al. for mosquitoes (Kyrou et al. 2018) (see also Section 4.1.2 on ZFNs, TALENs and CRISPR/Cas based homing systems).

The problem of new mutations arising is a consequence of the actions of the cell’s own repair mechanisms. In fact, what makes CRISPR such a popular mechanism for genome editing for breeding or research purposes is also its biggest weakness for gene drives. As detailed in Figure 2, a cell has two main pathways to deal with a double-strand break of the DNA: to either stick the ends roughly back together again with the non-homologous end joining pathway, (NHEJ) or to find and use a DNA template for the homology-directed repair (HDR). The NHEJ pathway commonly results in random mutations. The frequency by which the NHEJ is triggered will depend on species, cell type and developmental stage, but also on other factors not yet fully understood.

Looking at mutation rates of two different CRISPR/Cas9 gene drives in the fruitfly Drosophila melanogaster, Jackson Champer and colleagues reported: “We observed resistance allele formation at high rates both prior to fertilization in the germline and post-fertilization in the embryo due to maternally deposited Cas9. Assessment of drive activity in genetically diverse backgrounds further revealed substantial differences in conversion efficiency and resistance rates. Our results demonstrate that the evolution of resistance will likely impose a severe limitation to the effectiveness of current CRISPR gene drive approaches, especially when applied to diverse natural populations.” (Champer et al. 2017, 1, emphasis added)

One suggested approach to dealing with this problem is “multiplexing”, where a gene drive is equipped with multiple guide-RNAs capable of targeting different sequences. So far, no one has been able to experimentally advance the idea far enough to show a convincing avoidance of resistance. This will be discussed in Section 6.

3.3.2 Inefficiency in plants

In order for CRISPR/Cas gene drives to work, a major prerequisite is the triggering of the homology-directed repair (HDR) mechanism after the DNA double strand breakage (induced by CRISPR/Cas). Without this, the gene-drive element (or construct) cannot align next to the breakage point and be copied across into the target site (see Figure 3). However, the predominant repair mechanism in plants is the ‘non-homologous end joining’ (NHEJ) pathway, which simply sticks the loose ends of the broken/severed DNA strand back together in a haphazard way. Usually such a repair site will contain small mutations when compared to the original DNA sequence, and will in future be immune to being cut again by the same CRISPR/Cas. Researchers found that homology directed repair will rarely happen in plants, even in the presence of templates with high homologies. There are only a few examples in plants where CRISPR/Cas9 as a genome editing tool using a sequence insertion has been successful. The efficiency rate has been very low. Hahn for example reported, that even with specific frequency enhancing methods they only achieved a frequency of 0.12% HD-repair in the model plant Arabidopsis (Thale cress) (Hahn et al. 2018).

---

10 See also Grunwald et al.: “The alternative DSB repair pathway, non-homologous end joining (NHEJ), frequently generates small insertions and deletions (indels) that make CRISPR-Cas9 an effective means of mutating specific sites in the genome.” (Grunwald et al. 2019, 105)
3.3.3 Inefficiency in mice

Experiments did show that whilst a CRISPR/Cas9 gene drive can work in mice, it does so with only very limited efficiency (Grunwald et al. 2019). In this case the gene drive was designed to spread a mutation, which, instead of trying to cause infertility, attempts to change coat colour from grey to white. When inherited through the female germline, the gene was transmitted to 73% of offspring, exceeding the 50% expected from Mendelian inheritance. However super-Mendelian inheritance was not observed when the CRISPR/Cas9 construct was passed through the male germline, for reasons that are not yet understood. The researchers state that levels of transmission efficiency fall short of what is needed to rapidly drive a gene through a wild population without resistance arising, and comment that, “...both the optimism and concern that gene drives may soon be used to reduce invasive rodent populations in the wild is likely premature.” (Grunwald et al. 2019, 108)

3.3.4 Issues with p53

A further complexity has emerged for CRISPR genome editing from experiments on cultured human (and mouse) cells. This showed that genome editing is often counteracted by the cell’s natural defences against DNA damage, and that such editing is most likely to be successful in cells in which such a protective defence is somehow not active. These protective mechanisms, which are mediated through a tumour-suppressor protein known as p53, may represent an unanticipated hurdle in designing gene-drives in some organisms, in particular mammalians, as explored in Box 2.

3.3.5 CRISPR/Cas off-target effects

There are a number of issues regarding off-target effects, in particular in relation to the behaviour of CRISPR/Cas9 in wild populations. Firstly, and as detailed above, there have been cases where substantial mutations have occurred either at the target site or at a distance from the target site. Whilst the mechanisms and reasons behind this are not understood, there is even less knowledge concerning how CRISPR/Cas may behave under ‘natural’ conditions, outside laboratory settings and in wild and diverse populations. Secondly, the same is true for off-target effects, where DNA breakage occurs at non-target sites and where repair occurs via the error-prone NHEJ pathway. Will the rate of such off-target breakages change once released into the wild, where the conditions may be significantly different from any settings previously tested? Furthermore, the genomic DNA sequence of wild populations will entail substantial variations as compared to any laboratory reared strain. CRISPR/Cas may thus find accidental target sites that were not intended as target sites and which once cut may be repaired with mistakes.

In fact, releasing CRISPR-Cas9 gene drives into the wild is placing the laboratory and genetic modification procedures into wild populations, with no means of any control at hand. Unintentional mutations arising may be harmless or may be highly problematic, such as disrupting genes, altering gene regulation or producing new proteins or RNAs, clearly adding significant risk to any release.

3.3.6 Invasiveness and potential global reach

The majority of gene drives designed or modelled to date, including the CRISPR-Cas9 based homing drives, have the potential to be highly invasive (see Section 5). Because populations of species breed with other neighbouring groups, over the course of many generations genetic material can spread throughout a whole species. Therefore, a drive released in one country or region could spread to other neighbouring areas, and eventually

11 first published online in 2018 without peer review on https://www.biorxiv.org/content/early/2018/07/07/362558
12 The same authors state: “Although HDR of CRISPR-Cas9-induced DSBs does occur in vitro and in vivo in mammalian cells and embryos, usually from a plasmid or single-stranded DNA template, NHEJ is the predominant mechanism of DSB repair in somatic cells.” (Grunwald et al. 2019, 105)
could reach all reproductively linked populations of a species around the globe. This means that there are considerable technical difficulties in designing a gene drive that can be confined to a particular geographic area. This has been recognised widely as a serious problem. Invasiveness, combined with a lack of recallability and reversibility (see Table 2), and also combined with the potentially increased rate of mutations and modifications within a wild gene drive population, is likely to make the risks incalculable and potentially very high.

### 3.3.7 Irreversibility

Almost all of the gene-drive designs constructed so far, and especially the homing CRISPR-Cas9 technology, make effectively irreversible changes to the genome. One possible exception may be very low release rates of underdominance based drives, which however are not CRISPR-based homing drives. This leads to difficult questions about what steps could be taken if a CRISPR gene drive behaves in an unpredicted or harmful way, and what can be done in order to prevent this. Some suggest that a way to reverse the effects of a drive would be to release another gene drive to ‘overwrite’ the changes induced by the first drive. However, alongside the obvious potential for further unpredicted effects, this approach could not completely restore the genomes of affected species to the baseline states, because now both the sequence, as well as the CRISPR/Cas activity of the second gene drive, would be present in the natural population as well, gradually spreading throughout. This is such a thorny technical question, the irreversibility of CRISPR/Cas based homing drives has now become a major funding focus of DARPA’s ‘Safe Gene’ program, a circumstance which underlines the degree and urgency of this problem.

**Box 2: p53 and CRISPR**

Introducing p53, the ‘Guardian of the Genome’

Since its discovery in cultured mouse cells in 1979 (Lane and Crawford 1979), the tumour suppressor p53 has become one of the most intensively studied proteins in the mammalian cell.

Biologists’ fascination with p53 arises from its role in protecting organisms from cancer. In response to signals triggered by events such as DNA damage (Kastan et al. 1991) or uncontrolled cellular replication, p53 activates protective responses which include halting cellular replication (cell-cycle arrest), DNA repair or programmed cell death (apoptosis) (reviewed by Kastenhuber and Lowe 2017). This role has earned p53 the title ‘guardian of the genome’ (Lane 1992). Inactivation of p53 through mutation of its gene, TP53, allows a potentially cancerous cell to avoid these defence mechanisms, and hence this is the most commonly mutated gene in human cancers (Kandoth et al. 2013).

**p53 versus CRISPR**

Following its invention only a few years ago, in 2012, CRISPR-Cas9 genome editing of mammalian cells has become a routine laboratory experimental procedure, generating great interest in its potential to treat human disease (Adli 2018). However, it was only in 2018 that evidence emerged from studies of cultured human cells showing that the cell’s natural defence mechanisms, centred on p53, counteract the genome editing process by inducing a DNA damage response and cell cycle arrest (Ihry et al. 2018, Haapaniemi et al. 2018). Equally, when p53 is inactivated, CRISPR/Cas9 becomes more effective. In human stem cells gene editing was observed to be 17 times more efficient in the absence of p53 (Ihry et al. 2018). This means that cells which have been successfully edited are likely to lack this vital protective mechanism, posing a significant cancer risk if they are re-implanted in a patient without appropriate screening. This discovery has led to calls for caution in applying the technology in the clinic.

**How widespread are p53 type protective mechanisms?**

Given the intention for widespread application of CRISPR/Cas9, the question arises: in which organisms and in which cell types might we expect similar interference in genome editing? Answering this question requires a little more background. In humans and most mammals, p53 has two additional and similar ‘sister’ proteins, p63 and p73, which perform a variety of functions, some of which are related to controlling development, and some of which resemble or even overlap with those of p53 (Belyi and Levine 2009). All three have some role in protecting genomic integrity by responding to

---

13 More than 90,000 publications on this protein can be found on Pubmed
DNA damage and inducing apoptosis (Lin et al. 2009; Zaika et al. 2011). The protective activity of each form appears to vary in different cell types: p63, for example, has been shown to protect female germline cells (i.e. reproductive cells) (Suh et al. 2006), whereas the core function of p53 is protecting somatic (non-reproductive) cells (El Husseini and Hales 2018), with an apparently less prominent role in protecting the germline (Muller, Teresky, and Levine 2000).

The p53/p63/p73 family can be found in all bony fish, mammals and birds, whilst related proteins are found in other vertebrates, mollusks, insects and nematodes, but not yeast (Belyi et al. 2010). The common ancestor of the p53 family is believed to have been present in very early animals: a descendant of this ancestral p53 is found in modern day sea anemones, which diverged from other animals around a billion years ago, and has been shown to protect the genome by inducing apoptosis in response to DNA damage (Pankow and Bamberger 2007) in germline cells, but not somatic ones. This implies that the ancestral p53 evolved to protect germline genomic integrity and that this function has been retained within the p53 family throughout the animal kingdom.

What does this mean for gene drive research?

The complexity of the many processes involved make it difficult to fully predict the consequences of the interplay between a CRISPR/Cas9 homing gene drive and the protective mechanisms mediated by the p53 family. It might be expected that if CRISPR/Cas9 genome editing is activated in germline cells to propagate a gene drive, members of the p53 family could promote repair of the DNA breaks (without integration of the gene drive) or activation of programmed cell death, both of which could interfere with propagation of the drive. Whilst gene drives have been shown to work in insects (Kyrou et al. 2018) which possess p53 family proteins, this does not rule out that the p53 type responses could be activated in other cases, either due to a higher sensitivity to DNA damage, or to differences in gene drive design. This may account for the observed low efficiency of CRISPR/Cas9 gene drives in mice (Grunwald et al. 2019), which possess a p53/p63/p73 family similar to that found in humans with obvious potential to counter-act CRISPR/Cas9. Indeed, the presence of this powerful set of defence mechanisms may present a significant obstacle to efforts to apply gene drives in many animal taxa beyond insects.

3.4 CRISPR/Cas as enabler for many Gene Drive systems

CRISPR/Cas9 has become a key element in the development and feasibility of gene drives. This is not exclusively but is particularly true for homing drive systems, as just mentioned above, where CRISPR/Cas9 has become the prime agent in the role of an RNA-guided homing endonuclease (CRISPR/Cas-based homing drive).

However, there are other gene drive systems that have received little attention, but because of the utilisation of CRISPR/Cas (as a site-directed endonuclease and genome editing tool), are now experiencing accelerated development (Marshall and Akbari 2018), for example the X-shredder and toxin-antidote systems, discussed in the following section.

4 Mechanisms and techniques used

Engineered gene drives have two main goals for practical and/or commercial use in the fields of human health, industry or agriculture: to either alter (modify) and replace a population; or to suppress and eliminate a population or a species. In this, some applications depend on the ability to transport or carry a ‘payload’ or ‘cargo’ gene, together with or linked to, the drive element. Payload or cargo genes envisaged for altered characteristics may include, for example: toxin genes, disease resistance genes or disease-refractory genes. The latter are genes coding for compounds that will stop vectors (e.g. mosquitoes or flies) from transmitting diseases (e.g. Zika or Malaria), by blocking the pathogen from developing or spreading in the host-vector. Whether a particular drive mechanism has the ability to relia-
bly transport such extra cargo is a criterion indicated in the sections below and summarised in Table 2.

Most drive mechanisms are based on so-called ‘selfish genetic elements’, also referred to as ‘active genetic elements’ by some (e.g. Gantz and Bier 2016, Grunwald et al. 2019). As detailed below, these drive mechanisms can be categorised as two main types, which attempt to achieve inheritance bias through: 1.) over-replication of the genetic element, thus also referred to as ‘replication distorter’; 2.) preferential segregation or transmission of the genetic element, also referred to as ‘transmission distorter’.

The utilisation of CRISPR-endonucleases features strongly in both types, enabling the development of synthetic gene drives for different mechanisms and modes of action. In fact, there is an overlap between the use of different genetic elements, the mechanisms, and the modes of action, that is detailed below.

4.1 Selfish genetic elements:

Genes within a genome are commonly seen as working together collaboratively to produce a viable organism (Runge and Lindholm 2018). As part of this collaboration, all genes get an equal chance of transmission during sexual reproduction, which, according to Mendel’s Law of Inheritance, gives each gene from each of the two parents a 50:50 chance of being passed on to the next generation (i.e. from ‘child’ to ‘grandchild’). Most multicellular organisms are (at least) diploid, meaning they have two complete sets of genetic material in the form of chromosomes, one set from each parent. This means that each gene is present in two copies, occupying the same position or ‘locus’ on the parallel (or homologous) chromosome, often with slight variation. The different variations of a gene are termed ‘alleles’, coding for example for yellow or green seed colour. If the two copies or alleles of a gene are identical within an organism, this organism is termed to be ‘homozygous’ for that gene or allele (or sometimes trait) (see Figure 4).

If there are two different versions or alleles within the organism, the organism is termed ‘heterozygous’ for that allele (or trait). These alleles will be passed on (transmitted) in a 50% ratio to the offspring.

There are, however, specific genetic elements that do not play according to the same rules. This is why they are termed ‘selfish genes’ or ‘selfish genetic elements’ (Werren, Nur, and Wu 1988), they seem to solely look after their own interests rather than the new offspring’s. They are not part of the collaborative effort of upholding or enhancing the viability and fitness of an organism, and do not follow the 50:50 rule of inheritance. Instead, they have gained control over their own transmission. They are also capable of altering the odds of inheritance in their own favour, and thus are able to rapidly propagate through populations (Manser et al. 2017). They can even do so at a high fitness cost to the organism. This in turn leads to “counter-adaptations” by the rest of the organism’s genome “that generate unique selection pressures on the selfish genetic element. This arms race is similar to host–parasite co-evolution …” (Runge and Lindholm 2018, 1). In this sense, genes can be seen as a type of society, in which most members behave in a certain way, but there are occasionally outliers or other aberrant members. As in society, the outliers sometimes confer an advantage or disadvantage to the group, but unpredictably. Yet the story does not stop there.

Rather than casting the selfish genetic elements (SGEs) in a negative light, and focusing on the aspect of ‘selfish’, it is precisely this co-evolution and co-adaptation that is becoming a focus of research. Instead, some of the elements or mechanisms, for example the over-replication ability of transposable elements, are increasingly regarded as vital components for genome evolution and even speciation (Biemont 2010). John H. Werren importantly noted: “The story that is emerging increasingly supports a central role of SGEs [selfish genetic elements] in shaping structure and function of genomes and in playing an important role in such fundamental biological processes as gene regulation, development,
Chapter 1: What are Gene Drives?

The evolution of genetic novelty, and evolution of new species.” (Werren 2011, 10863). In fact, the study of these elements and the processes involved are now contributing to an emerging re-think of what a genome is and how it interacts with its environment. (Werren 2011; Lindholm et al. 2016).

Consequently, when talking about engineered gene drives, which are all based on and are exploiting the mechanism of these SGEs, there is a level that we cannot comprehend at this point in time. SGEs are vital evolutionary players, deeply embedded in a long evolved and complex regulatory structure. What therefore does it mean to take SGEs out of their own context, reshape and alter them, and place them back into this interactive system? What might the consequences be at that specific level?

This is an important discussion that needs to take place now. It would be wrong to inadvertently assume that SGEs are only a “tool” that can be readily adapted and utilised for the purpose of modifying organisms and whole populations in the wild. They are much more than that and it could be most unwise to disregard this.

Setting aside their important role in evolutionary dynamics and focusing on the aspect of gene drive, there are many different ways by which ‘selfish’ genetic elements enhance their own presence in a population or species. Such genetic elements may for example be genes, sections of chromosomes or even whole chromosomes. There are the ‘over-replicators’ (McLaughlin and Malik 2017) that flourish by copying or moving themselves to other parts of the genome, which are termed ‘transposable elements’, or into their own (allelic) locus in the parallel chromosome, termed ‘homing endonuclease genes’. The other group are the ‘transmission-distorters’, which ensure they are the genes transmitted to the next generation, not “the other” ones. This is often done by actively destroying “the other”, whether that is at the DNA level (e.g. X-shredders) or at the level of cells or embryos that will die (e.g. Medusa).

Figure 4: Diploid chromosomes, alleles and their terminology. Depicted are two pairs of chromosomes, with one chromosome of each pair (here a short and a long one) derived from each parent. The chromosomes within a pair are termed homologous chromosomes and are basically the same in that the position (locus) of the genes are the same, though they may be different alleles. If the alleles are identical then the organism is homozygous for that gene or trait (e.g. purple flower colour). Otherwise the organism is heterozygous for a gene or a trait (e.g. seed colour).
A number of these selfish genetic elements, their multiplication or transmission mechanisms, and their mode of action, are being or have been considered for the construction and use of engineered gene drives. The following briefly describes these selfish genetic elements and/or drive mechanisms by which these elements multiply or change the odds of their inheritance and transmission. Also indicated is the extent to which these may potentially lend themselves as gene drive systems for population eradication or modification, the latter also depending on the ability and reliability of carrying and spreading payload genes. This section also lists limitations and risk factors, such as inactivation of or build-up of resistance to gene drive mechanisms, the lack of reversibility, and vertical gene transfer to neighbouring populations or closely related species. A summary is given in Section 5 (Table 2).

The elements and drive mechanisms described in this section work at very different levels and cannot easily be compared. Some focus on the intentional outcome (e.g. sex-ratio distorters) or the mode of action (e.g. toxin-antidote based drives), whilst others refer to the mobile element itself (e.g. ‘transposable’ element). An overview is given in Figure 5, which is based on discussions in (McLaughlin and Malik 2017; Lindholm et al. 2016; Simoni et al. 2014).15

4.2 Over-replicators / replication-distorters

As outlined above, over-replicators achieve an inheritance bias in their favour by creating extra copies of themselves in the genome. There are two members in this group, the transposable elements

---

15 Simoni for example states: “Naturally occurring selfish elements include transposable elements, meiotic drive chromosomes, sex ratio distorting elements and homing endonuclease genes (HEGs). HEGs are highly specific endonucleases that generate double-strand breaks (DSB) at specific loci in the host genome.” (Simoni et al. 2014)
Transposable elements (TEs) and homing endonuclease genes (HEGs). TEs do not have a mechanism to guide the insertion of a new copy to a particular site in the genome, although the insertions do not occur randomly either; whilst the HEGs have the means to guide their insertion to a precise location on the genome, namely at the exact same position where they are, but on the parallel chromosome.

4.2.1 Transposable elements (TEs)

Transposable elements (TEs) are a type of ‘mobile genetic elements’ that are found in almost all species. Discovered by Barbara McClintock in maize in the late 1940s (McClintock 1950) and later described as ‘jumping genes’, a TE is a segment of DNA that can change its position within the genome of an organism on its own accord. This process is commonly referred to as transposition. During transposition, TEs will often increase their copy number within the host genome, thus leading to a higher inheritance rate in subsequent generations.

Whilst there are many different families of TEs, they can all be grouped into two classes: (I) those that move by copy & paste mechanisms (via an RNA intermediate); and (II) those that move by cut & paste mechanisms, often referred to as ‘DNA transposons’. Transposition of class-I TEs will automatically result in multiplication, as the original TE remains in its place whilst a copy inserts itself at a different location in the genome. This is not the case with class II TEs, although replication may occur through a number of mechanisms (summarised in Marshall and Akbari 2016), thus leading to a higher enhanced rate of inheritance.

If, for example, a DNA transposon moves position during the stage of DNA replication of the cell cycle, it could jump from a location that has already been replicated and land in a location that has not yet been replicated, resulting in a net gain of one TE. Additionally, once the TE is excised for transposition, there will be a gap in the DNA where it had been. The cell’s repair mechanism may simply rejoin the loose ends by non-homologous end joining (NHEJ) or it may fill the gap via homology-directed repair (HDR), using the duplicated DNA strand containing the TE as a template.

Whilst TEs are commonly referred to as selfish genetic elements, (Munoz-Lopez and Garcia-Perez 2010), this view is not shared by all. Biemont for example states: “TEs are no longer seen as ‘junk’ and ‘selfish’ pieces of DNA—the predominant view from the 1960s through the 1990s—but as major components of genomes that have played a significant role in evolution, an idea also first proposed by McClintock (1984: her Nobel Prize lecture).” (Biemont 2010, 1085). Whilst high TE activity outbursts can at times be associated with speciation, such outbursts are usually very time limited, as the host organism will soon generate counter-measures to shut down the activity of the TEs and have them back in order. Measures like gene-silencing will for example disable the production of compounds that TEs need in order to multiply or jump.

TEs of class II are known to be able to spread widely throughout populations. The ‘P element’ is a good example of this. Now commonly found throughout populations of the fruit fly Drosophila melanogaster, the P element seems to have only arrived in this species in the 1930s, rapidly spreading throughout all its populations within 50 years (Anxolabehere, Kidwell, and Periquet 1988). They are thought to have horizontally transferred from Drosophila willistoni, possibly via the semiparasitic mite Proctolaelaps regalis (Houck et al. 1991).

Due to their ability to spread widely and to enhance their presence in a genome, TEs can make up a substantial portion of the genome of a species. However, most of these TEs will have been inactivated over time through acquired mutations and various host defense mechanisms, including gene-silencing mechanisms such as DNA methylation and RNA interference (Munoz-Lopez and Garcia-Perez 2010). For example, the genome of silk-

---

16 Though a selfish element, TEs are not considered by everyone to be a gene drive.

17 “On the other hand, host organisms have developed different mechanisms of defense against high rates of transposon activity, including DNA-methylation to reduce TE expression [...] several RNA interference mediated mechanisms [...] mainly in the germ line [...], or through the inactivation of
worms (*Bombyx mori*) is comprised of around 45% TEs, that of honeybees (*Apis mellifera*) only of 1%, in (Biemont 2010) and that of the main malaria-carrying mosquito (*Anopheles gambiae*) about 15% (Holt et al. 2002). Whilst some plants, especially maize, have a genome with more than 70-80% TE sequences, humans have around 45%, mice about 37% and some fish 10%, see (Munoz-Lopez and Garcia-Perez 2010 and Biemont 2010).

**TEs as gene drive mechanisms:**

The rapid spread of the *P* element initially raised hopes that class II TEs could be used as gene drive systems, transporting engineered ‘payload’ genes throughout populations of intentionally modified/engineered organisms. It is these TEs of class II that are regarded as having the potential for gene drive applications. They basically consist of a transposase gene framed by terminal inverted repeats (TIRs). A payload gene placed adjacent to the transposase gene would thus – at least in theory - move together with the TE construct and spread above the 50:50 odds of inheritance.

Genetically engineered TEs have been used to transform and genetically modify insects, first achieved in *Drosophila melanogaster*, utilising the *P*-element (Spradling and Rubin 1982). The *P*-element, however, only works in drosophilid insects. The *Hermes*, *mos1/mariner*, Minos and *piggyBac* elements were identified to work in some mosquito species (listed in Macia et al. 2017B and reviewed in O’Brochta et al. 2003).

Attempts have been made to harness TEs as gene drives in mosquitoes, yet when engineered into an organism the integrated engineered TEs have very low remobilisation rates, meaning they stay where they are and do not jump. Macias et al. commented recently: “It was imagined that transposons would also be useful as a gene drive system, but transposons that could mediate insertion into a mosquito’s genome were not so easily remobilized [...]. Only recently has a synthetic construct based on the piggyBac transposon been demonstrated to mobilize itself once inserted into a mosquito genome, but rarely […]” (Macias, Ohm, and Rasgon 2017, 3). Even special attempts to improve the post-integration mobility of artificial *Hermes* and *piggyBac* elements have only resulted in mobilisation rates less than 1% (Smith and Atkinson 2011; Macias, Ohm, and Rasgon 2017) and 6% (O’Brochta et al. 2011). This is much too low for gene drive requirements.

**Draw backs & limitations:**

- TEs do not integrate at specific recognition sites – and therefore cannot be used to disrupt or knock out a specific target gene, such as a gene crucial for development, fertility or gender. TE-based gene drives could thus only be used in replacement strategies for – in theory at least – spreading particular trait or effector genes.
- Low efficiency: As pointed out, experiments have shown that the post-integration mobilisation of engineered TEs has so far been extremely low and insufficient for gene drive purposes.
- Insert size (payload gene or cargo) can be a problem – size matters. Frequently, TEs with large cargo sizes don’t jump or spread easily, e.g. *Sleeping Beauty* element (Izsvák, Ivics, and Plasterk 2000). Lampe et al., for example, found for the *Himar1 mariner* element (in horn fly) that “transposition frequency decreased exponentially with increasing transposon size” (Lampe, Grant, and Robertson 1998).
- Stability and integrity of the construct and insert: It has been observed that engineered TEs can lose (or throw out) the added cargo DNA sequences, for example those found with the *P* element (Carareto et al. 1997), or that the sequence of the insert is being mutated. The consequence would be the spread of the element, but not the additional effector gene sequences (Marshall 2008), which would not be desirable.
- Furthermore, cells or organisms are able to develop defence mechanisms against the mobilisa-
tion, jumping and spread of TEs. Proof lies in the large quantities of stationary and mostly deactivated, mutated or silenced TEs that often make up substantial portions of the genome.

- The specificity of particular and well-adapted TEs for particular species would make the use of TEs a new challenge for each new species.

- A serious drawback is that of horizontal gene transfer, where the TE (and a linked payload gene) is transferred to another species by mechanisms that are not fully understood. Transfer via sexual reproduction will keep a TE within the same species, yet TEs specific to one species have been found to appear in other species. Bourque et al. have summarised this recently: “There is now a large body of evidence supporting the idea that horizontal transposon transfer is a common phenomenon that affects virtually every major type of TE and all branches of the tree of life [...]. While the cellular mechanisms underlying horizontal transposon transfer remain murky, it is increasingly apparent that the intrinsic mobility of TEs and ecological interactions between their host species, including those with pathogens and parasites, facilitate the transmission of elements between widely diverged taxa [...].” (Bourque et al. 2018, 4).

A number of papers have drawn the conclusion that if TEs were to be used as gene drive systems, they would require a lot more experimentation, knowledge and research (Sinkins and Gould 2006) – and are in that sense regarded as either too costly (Marshall and Akbari 2016) or superseded by other gene drive methods.

A key question is also if genetically engineered gene-drive TEs will keep to the same pattern of semi-random integration, e.g. avoid inserting itself into the actual coding sequence of a gene. If it were to insert into coding sequences the outcomes would be highly unpredictable and could be problematic, even if the TE were blocked by the organism from further movements.

### 4.2.2 Homing endonuclease genes (HEGs)

Homing endonuclease genes (HEGs) are another type of ‘mobile genetic element’, originally discovered in budding yeast in the 1970s and early 1980s by researchers in the Pasteur Institute (Dujon 1980, Jacquier and Dujon 1985). HEGs have since been found in many bacteria, bacteriophages, fungi, and plant chloroplasts.

The principles underlying how HEGs are constructed and how they achieve drive are at the centre of current gene drive development, with the CRISPR/Cas-based homing gene drive being the best known category of engineered HEGs.

In general, HEGs are genes that code for an enzyme (endonuclease) that is able to recognise and cut a specific DNA sequence of 14-40 base pairs (Stoddard 2005), and to then have themselves copied into the middle of that sequence via homology repair. This overall process is called ‘homing’ (Stoddard 2005). The homing gene thus resides within the recognition sequence cut by the endonuclease.

If one chromosome contains a HEG and the equivalent homologous chromosome does not, the endonuclease will detect the recognition site on that chromosome and induce a site-specific double strand break. The HEG will be copied across via homology-directed repair (see Section 3 and Figure 2). The severed recognition sequence now becomes the flanking sequences (see Figure 3). If, however, the cleaved ends get rejoined by the NHEJ repair, mutations will occur in the recognition/cleavage site, making this chromosomal site unrecognisable to, and ‘resistant’ to, the specific endonuclease. Equally, natural sequence variation at the recognition site could protect the site from HEG insertion.

**Homing gene drives**

Austin Burt was the first to suggest utilising HEGs and their ability to spread by inserting themselves into the parallel chromosome, with the aim of either altering or eradicating natural populations of target}

---

18 “Homing is the transfer of an intervening sequence (either an intron or intein) to a homologous allele that lacks the sequence (Dujon 1989; Dujon et al. 1989; Belfort & Perlman 1995), leading to gene conversion and dominant transmission and inheritance of the mobile element.” (Stoddard 2005, 50)
species. As naturally occurring HEGs are limited in the DNA sequences they may recognise, Burt envisaged that through genetic engineering it would be possible to alter the site-specificity of the endonuclease in order to make it target essential genes (Burt 2003). With an appropriate promoter (e.g. for meiosis), and ensuring that the resulting knock-out was recessive, Burt suggested a population could – in theory – be eradicated in 20 generations. Another possible action suggested was to engineer the HEG construct to contain an extra gene, a so-called payload or cargo gene (Champer, Buchman, and Akbari 2016), which would get copied across together with the HEG and thus also spread in a super-Mendelian fashion.

The potential use of homing endonucleases as gene drive systems in animals was first tested in the fruit fly *Drosophila melanogaster* (Chan et al. 2011, Chan, Huen, et al. 2013) and the mosquito *Anopheles gambiae* (Windbichler et al. 2011). Using HEGs derived from yeast and inserting artificial target sequences into the respective genomes, the experiments provided a proof of principle for homing processes to work in these species, although drive conversion levels were low.

These experiments also indicated the importance of the proper timing for when an HE gene is activated and for the resulting homing endonuclease to find and sever its target sequence. Depending on which point during gametogenesis and meiosis a DNA double-strand break occurs, different repair mechanisms dominate the process (see Section 3 and Figure 2 for repair pathways). Increased frequency of NHEJ-induced repair will result in increased frequency of target site mutations, thus enhancing the rise of resistance. This means that choosing the right promoter element in an HEG construct is crucial to the outcome (see Chan et al. 2011, Table 1).

If HEGs were to be used as gene drive systems, their target specificity would need to be adaptable for different sequences. Researchers from Cambridge, UK and Seattle in the US redesigned and engineered HEGs based on yeast HEGs with altered specificity, and also used artificial target sites. Working with *Drosophila*, they found that the redesigned homing endonucleases had a reduced conversion frequency, as compared to the original yeast HEG (Chan, Takeuchi, et al. 2013). The authors suggested that site specificity alone is not sufficient for successful homing.

**ZFNs, TALENs and CRISPR/Cas based homing systems**

To test if site-specific nucleases could be used as homing endonucleases, researchers turned to zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). Using the same experimental system and design as Chan in 2011 and 2013, researchers at Imperial College, London, tested ZFN-based and TALEN-based HEGs. They found these site-specific nucleases triggered a much higher rate of NHEJ repair than the actual/original HEGs (Simoni et al. 2014). As a consequence, there was an increased level of target-site mutations, which in turn creates resistance to the gene drive, because the enzyme is no longer able to recognise or cut the target site.

However, with the arrival of RNA-guided site-specific nucleases such as CRISPR/Cas9, the picture drastically changed. As already detailed in Section 3 of this chapter, the recent uptake of CRISPR/Cas as a homing gene-drive system has resulted, most of the time, in high conversion rates (rates of homing) in laboratory studies, although these also demonstrated the occurrence and build-up of resistance (see Table 1). Whilst demonstrating proof of principle for the fruit fly *Drosophila melanogaster*, the yeast *Saccharomyces cerevisiae* and the mosquitoes *Anopheles stephensi* and *A. gambiae*, the issue of gene drive resistance remained an insurmountable hurdle.

This changed suddenly in 2018, when researchers from Imperial College London, UK, succeeded...
in completely crashing a laboratory population of caged *Anopheles gambiae* mosquitoes (Kyrou et al. 2018). Strictly following Burt’s original strategy (Burt 2003), they did so after 7–11 cage generations without any emergence of resistance. This was a first, and has brought this technology to a further step of proof of principle, at least in enclosed, caged, artificial systems.

The strategy was to choose a target gene that was both highly conserved and essential in gender determination, the *doublesex* gene. Disrupting this gene at a particular site with a CRISPR-Cas9 gene drive results in sterility in females carrying the drive. ‘Highly conserved’ means that the DNA sequence of a gene has remained the same over time on an evolutionary scale, and that it has not been changed by random mutations. A ‘highly conserved sequence’ implies a conserved and highly protected gene, where any alteration to that gene sequence would result in a non-viable life form. Choosing a highly conserved gene sequence, in particular the sex determination ‘*doublesex*’ gene, as the gene drive target site, means that no viable resistance alleles (gene variants) arise and spread to save the caged (or potentially, the wild) population. This is a new strategy on the path towards overcoming this type of gene drive resistance and so far has resulted in the above-mentioned crash of a population of caged mosquitoes.

Because of its vital role, the doublesex gene has very little scope for mutation and therefore the minor mutations which normally allow resistance to evolve do not appear; this is likely to be the mechanism allowing this drive to completely eradicate laboratory populations. Significantly, the gene sequence is completely conserved across the *Anopheles gambiae* species complex, meaning the drive would function just as effectively in these sibling species. Given the capacity of members of this complex to hybridise, if this drive were released in the wild it could potentially affect the entire species complex alongside *gambiae* – along with the ecosystems linked to them.

The strategy of targeting highly conserved genes to avoid the build-up of target site resistance thus adds an extra layer of risks and concerns to what is already perceived to be a very high-risk technology.

Experiments have also been carried out in mammals, in this case mice. When both gene copies (alleles) of a targeted gene controlling coat colour are disrupted, in order to change grey coat to white, the researchers found that the gene drive did not work readily. Drive activity in early embryo or male germlines resulted in mutations, rather than drive conversion, as the predominant repair mechanism was NHEJ. Of the various strategies, limiting the gene drive activity to the female germline gave an efficiency or conversion rate of 73%, but also showed NHEJ-induced mutations. Whilst providing a proof of concept, the authors noted that the “precise timing of the Cas9 expression may present a greater challenge in rodents than in insects” in terms of efforts to prevent resistance to the gene drive (Grunwald et al. 2019). (see also Chapter 2, case study on mice).

### 4.3 Segregation & transmission distorters

One mechanism for genetic elements to achieve drive or super-Mendelian inheritance is by their own duplication (over-replication) and insertion into other chromosomal loci. This is the mechanism used for TEs and HEGs.

The drive mechanisms of transmission distortions described in this section result in an inheritance bias by means of eliminating, outmanoeuvering or outracing the competition. The term ‘distortion of transmission’ was first coined by L.C. Dunn in the late 1930s\(^2\) (Dunn and Bennett 1971).
Table 1: CRISPR/Cas9 based drives: Relevant proof of concept work published on RNA-guided gene drives. This table compares the proof of concept drives both for the conversion rate as well as for the degree of resistance to the drive observed. Whilst not directly comparable due to differing experimental procedures, including numbers of generations observed, the development of resistance is common to all except for Kyrou et al. (2018).

<table>
<thead>
<tr>
<th>Species</th>
<th>Trait</th>
<th>Kind</th>
<th>Conversion rate (homing rate)</th>
<th>Resistance</th>
<th>Institute</th>
<th>REF</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Drosophila melanogaster</em> (fruitfly)</td>
<td>Yellow body colour</td>
<td>Loss of function (X-linked gene)</td>
<td>97%</td>
<td>Not tested [3% ?]</td>
<td>University of California, San Diego, US</td>
<td>(Gantz and Bier 2015)</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em> (yeast)</td>
<td>Colour change</td>
<td>Gene addition, Gene correction, Gene disruption</td>
<td>&gt;99%</td>
<td>Not found</td>
<td>Harvard Medical School, Boston, US (+)</td>
<td>(DiCarlo et al. 2015)</td>
</tr>
<tr>
<td><em>Anopheles stephensi</em> (mosquito)</td>
<td>anti-Plasmodium falciparum effector genes (refractory)</td>
<td>Introgression (population modification)</td>
<td>99.5% (in germline) ~50% in egg</td>
<td>Yes. In particular, if homing action leaked to egg from females with drive. (&gt;70%)</td>
<td>University of California, San Diego &amp; Irvine, US</td>
<td>(Gantz et al. 2015)</td>
</tr>
<tr>
<td><em>Anopheles gambiae</em> (mosquito)</td>
<td>Female sterility</td>
<td>Loss of function, suppression &amp; payload gene</td>
<td>91.4 to 99.6% initially. 69–98% at later generations</td>
<td>Yes, including in-frame mutations (6 bp deletion)</td>
<td>Imperial College London, UK, University of Cambridge, UK, University of Perugia, Italy (+)</td>
<td>(Hammond et al. 2016)</td>
</tr>
<tr>
<td><em>Anopheles gambiae</em> (mosquito)</td>
<td>Female sterility</td>
<td>Loss of function, suppression</td>
<td>100%</td>
<td>unlikely</td>
<td>Imperial College London, UK</td>
<td>(Kyrou et al. 2018)</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>Sex-conversion to males</td>
<td>Loss of function, suppression</td>
<td>30% (third of which in-frame mutations)</td>
<td></td>
<td>University of Göttingen, Germany</td>
<td>(KaramiNejad-Ranjbar et al. 2018)</td>
</tr>
</tbody>
</table>
Chapter 1: What are Gene Drives?

Meiotic drive

Meiosis is the key phase for sexual reproduction in a higher organism where the sexual reproductive cells (gametes - e.g. egg cells or sperm) are being produced and the genetic material is divided up in a random fashion in line with Mendel’s Law. During meiosis, diploid cells are divided into haploid cells, in which only one copy of a gene or a chromosome will be present, originating from either of the parents.

‘Meiotic drive’ is an overarching term referring to any selfish genetic element or drive mechanism that manipulates the processes of meiosis\(^{21}\) and the “production of gametes\(^{22}\) to increase their own rate of transmission, often to the detriment of the rest of the genome and the individual that carries them.” (Lindholm et al. 2016, 315).

In the meiotic drive system, the determining tasks are ‘who’ will succeed in getting a ride in the gametes, eliminating or outmanoeuvering the competitors. And which embryo will survive, which may depend on the presence of an antidote to counter the toxin produced at an earlier stage by part of a selfish genetic element team. This for example would be the tactic pursued by the ‘toxin-antidote’ based drives.

The forces present in this dynamic between genetic elements, mode of action and responses and protective efforts by the individual organism and species, are tremendous. The review “The Ecology and Evolutionary Dynamics of Meiotic Drive” by Anna Lindholm et al. (2016) offers an insight that reminds us that synthetic gene drives based on the mechanisms of naturally-occurring meiotic drives will not only be exposed to (and have to withstand) the same counter-forces, but synthetic drives will also shape responses in their turn and thus influence ecology and evolution.

There are many types of meiotic drive systems, the main ones of which are explained below.

4.3.1 Sex-ratio distorters

Sex ratio distorters are drive systems that skew the gender ratio, resulting in either predominantly male or female offspring. Also referred to as ‘sex-linked meiotic drive’ (Champer, Buchman, and Akbari 2016), they are the main drive systems under development for synthetic gene drives. If the key factor for determining population size is the number and productivity of females, then eradicating females becomes the action of choice for gene drive systems. There are two options: to place the modified selfish element either on the male sex chromosome (the Y-chromosome), or on an autosome (a chromosome other than a sex chromosome). The highest and swiftest suppression rate can be achieved if the drive is linked to the Y-chromosome (Champer, Buchman, and Akbari 2016; Marshall and Akbari 2018).

Sex ratio distorters achieve an inheritance advantage by destroying ‘the other’. If, for example, the X-chromosomes gets destroyed during spermatogenesis, there can be no female offspring of that organism, if the determinant for female is XX. A mechanism ensuring that only the Y chromosome gets through spermatogenesis would change the sex ratio drastically towards male. Such a strategy would – in theory - cause a population to collapse over time.

Male bias (sex-ratio distortion) is found in nature, for example in *Aedes* and *Culex* mosquitoes (Craig, Hickey, and Vandehey 1960; Newton, Wood, and Southern 1976; Sweeny and Barr 1978). Although the actual molecular mechanism behind this is not understood, there is a specific type of Y chromosome that will result in 90% male offspring. Somehow the presence of this driving Y chromosome during spermatogenesis leads to breakages in the X chromosome, disabling or preventing female progeny (Burt and Crisanti 2018).

However, it is obvious that those natural occurring sex-distortion mechanisms have not resulted in the elimination of those populations or species, as

---

21 This action and timing require meiosis-specific regulatory elements (promoters).
22 Gametes are mature sexual reproductive cells, with female gametes being egg cells (ovules) and male gametes sperm (pollen).
they can still be found. This was already being discussed in the 1960s when the use of organisms with naturally occurring drives were suggested as a form of biocontrol. Hamilton (1967) argued that there would be a response, a counter mutation or counter elements, to contain and ‘mask’ any sex distortion factor, especially a strong Y-linked male bias factor. In this way, a co-evolutionary process would bring the sex-ratio back to equilibrium\(^23\) (Hamilton 1967).

The question remains open as to what extent such ‘counter-measures’ would arise in response to engineered sex-ratio distortion drives. And if so, would they be quick enough to save the population or species? Or can gene drives be developed in such a way that such counter-measures can be blocked? There are currently three types of engineered sex-ratio distorters under consideration and being investigated as synthetic gene drives.

As already detailed above: A CRISPR/Cas9-based homing drive developed at Imperial College London was engineered to target the highly conserved doublesex gene in the mosquito Anopheles gambiae and resulted in male only offspring (Kyrou et al. 2018). Already presented in Section 4.1.2 under HEGs, this gene drive has been highly effective in the artificial environment of caged trials.

The second sex-ratio distortion drive is a mouse specific \(t\)-haplotype-based gene drive, being developed at Texas A&M University intended to produce ‘daugtherless’ mice, to eradicate mouse populations. There is no proof of concept so far. See below under Section 4.1.2 (a), \(t\)-complex or \(t\)-haplotype.

The third drive is the synthetic X-shredder gene drive, developed at Imperial College London.

\subsection{t-complex or \(t\)-haplotype}

The \(t\)-haplotype or \(t\)-complex is a meiotic drive and sex-ratio distorer located on chromosome 17 that naturally occurs in mice. Its discovery goes back to 1927, when Nadine Dobrovolskaia-Zavadskaya, evaluating X-ray experiments in mice, first thought this to be the gene for short tails or taillessness (gene symbol \(T\)), hence the name (Herrmann and Bauer 2012). However, further evaluations with crosses showed that ‘tailless mice produced only tailless litters upon intercrossing, but neither short-tailed nor normal-tailed pups. Inspection of the embryos from such crosses showed that about half of the embryos died in utero.” (Herrmann and Bauer 2012)

It was much later that it became evident that this region of chromosome 17 was what would later be called a selfish genetic element, containing not only genes for transmission distortion, but also for male infertility and embryonic lethality. Mice that are homozygous for the \(t\)-complex (i.e. where both parallel chromosomes contain the gene for embryonic lethality), will die before birth. And males with a copy of the \(t\)-complex will pass this on to 90% of offspring (Lyon 2003; Lindholm et al. 2013). This lethality is based on a toxin-antidote system, where a toxin will be released into the cells during spermatogenesis (sperm development) and only those sperm will survive or be able to fertilise an egg cell that carry the gene for the antidote, which is located on the \(t\)-complex.

This meiotic drive system is specific to mice. Where mice are perceived as a problem, e.g. on islands, plans are underway to try to alter and convert this system into a synthetic gene drive to turn against the mice. The idea here is to create ‘daughterless’ mice by modifying the \(t\)-complex with a mouse gene called \(Sry\).\(^24\) This gene will act during embryo development and trigger the development of male characteristics irrespective of the actual gender of the mouse. Released into the wild, any offspring should have male characteristics. With no females left to breed, the idea is that the population would collapse.

---

\(^{23}\) Such a sex-ratio equilibrium is also known as the ‘Fisher’s principle’.

\(^{24}\) The \(Sry\) gene is a mammalian sex-determining gene located on the the Y-chromosome. It determines maleness and its name is short for ‘sex-determining region of the Y’. Here, as payload gene of the drive system, it would be on an ‘autosome’, i.e. not on a sex chromosome, and thus present and active also in genetically female offspring. These would thus develop male characteristics but be infertile.
It is highly uncertain, however, whether this genetically engineered gene drive would or could perform as envisaged, and also what actual outcomes would be. This is conveyed by the fact that the underlying drive system (the \( t \)-complex) has been part of the mouse population and its evolution dynamics, but is still not present at a higher level than it is. This brings up the question of how mice manage to handle this species-specific selfish genetic element. It is a co-evolved system, so wouldn’t the mouse genome evolve a response? There are also reports that mating or fertilisation rates are lower with \( t \)-complex males than with wild type males (Manser et al. 2017). Altogether this may well mean that the system cannot deliver what researchers hoped for. However, the second question of course is whether this engineered gene drive is more aggressive and invasive than the natural mechanism it is based on, and what would happen to all and related species of mice if it found its way to other locations? As no experimental data are available, the performance of this gene drive remains speculative.

b. X-shredder

As indicated by its name, this type of transmission distorter will shred the X-chromosome during spermatogenesis (male meiosis), i.e. cleaving the chromosome at multiple sites by using a site-directed nuclease. Homology directed repair would not be possible at this stage, as there would be no second X chromosome that could serve as a template for such repair. First suggested by Austin Burt in 2003 (Burt 2003), it was in 2014 that Galizi reported on experiments with the malaria mosquito \( Anopheles gambiae \), providing a proof of principle for this approach. Using a homing endonuclease (\( I-PpoI \) nuclease from a slime mold) to cleave the X-chromosome at multiple sites led to a sex ratio of up to 95\% males (Galizi et al. 2014). He repeated the same experiments with a CRISPR/Cas-based site directed nuclease, giving rise to the same degree of transmission distortion and resulting in 86-95 \% male offspring (Galizi et al. 2016). The authors suggest that higher gender distortion could be achieved by placing the gene drive on the Y-chromosome rather than on a normal chromosome, an ‘autosome’. This would mean that all offspring (which in this case are always male) would automatically carry the X-shredder mechanism, rather than leaving its distribution to Mendelian inheritance rates.

4.3.2 Underdominance / Heterozygous disadvantage

This gene drive approach is envisaged to function as a tool for population replacement, that is, to spread a payload gene with its trait throughout a population and bring it to fixation, meaning that every individual will carry it.

Underdominance - also called ‘heterozygous disadvantage’ or ‘heterozygous inferiority’ - is a phenomenon where the heterozygous offspring is less fit than either of its homozygous parents. Where a gene has two variants (alleles - see Figure 4) and where it is an advantage to be homozygous for either one of these alleles (i.e., to have a set of either one variant or the other), a cross of the two differently homozygous parents will result in a heterozygous offspring with one of each allele. Under these circumstances, such a heterozygous offspring will have a lower fitness level than either parent and will over time be selected against. Ultimately, either one of the alleles will become firmly established (fixed) in a population, usually the one with the higher initial frequency (Davis, Bax, and Grewe 2001; Champer, Buchman, and Akbari 2016). There is also the chance that two distinct populations may arise, each occupying different or neighbouring territories. See footnote\(^{25}\) for different definition. Examples of underdominance in nature are mentioned in (Champer, Buchman, and Akbari 2016).

Although resulting in transmission distortion, naturally occurring underdominance does not lend itself as a gene drive system for spreading payload genes (Sinkins and Gould 2006). However, an idea arose to create and use genetically engineered un-

\(^{25}\) “UD is a genetic property classically defined as the condition where, at a single locus, the fitness of heterozygotes is lower than that of either corresponding homozygotes; generating this effect has been proposed as achievable through either chromosomal translocations or mutually suppressing transgenic toxin-antidote elements (though none have as yet been developed in exactly the latter format).” (Leftwich et al. 2018, 1204)
derdominance systems instead, utilising either reciprocal translocations or toxin-antidote systems, and increasingly also employing CRISPR/Cas, albeit not as a gene drive system, but rather to act as a ‘toxin’. However, there are multiple hurdles remaining.

Davis et al. (2001) were the first to suggest using a double toxin-antidote system for this purpose (including modelling for it). This system would employ two separate constructs (see Figure 6), each consisting of a copy of the desired payload gene, a suppressor gene and a lethal gene, with a promoter that can be suppressed by the suppressor present on the other construct (Davis, Bax, and Grewe 2001). As long as both constructs are present within an organism or an embryo, nothing will happen, as the lethal gene is suppressed. But if the offspring ends up with one construct only, no matter which one, it will die – an approach the authors called ‘extreme underdominance’. If sufficient numbers of individuals harbouring both constructs were repeatedly released, the expectation is that the engineered trait (and construct) would become established in the wider population.

Other teams developed these ideas and systems further, either keeping Davis’s idea of a ‘one locus’ drive (where both constructs will be at the same chromosomal location but on the opposite homologous chromosome), or suggesting a ‘two-locus’ drive system, where the second construct would be on a separate chromosome all together. For example, Akbari & Matzen, then at the California Institute of Technology, designed and built a synthetic ‘two-locus’ gene drive system they termed ‘maternal-effect lethal underdominance’ (UDMEL). Here the toxins are expressed maternally during egg-production, and it is the embryos that will die as a consequence unless they have the genes for the corresponding antidotes (Akbari et al. 2013). Males carrying the toxin genes will not express them. Tested in the model organism Drosophila melanogaster, UD^{MEL} was the first engineered threshold dependent gene drive system, and modelling suggesting a required release frequency of above 24%. Most papers published on underdominance drives however are solely theoretical papers, modelling both the thresholds and the multiple releases required for the payload gene to become fixed in a population, or at least to be present temporarily. These models also attempt to predict whether reversals might be possible, what variables need to be taken into account, etc. These models look at different theoretical gene drive constructs with different genes (toxins, suppressors, antidotes or replacement genes) as well as combinations with different promoters that will make drive components active either in females, males, adults or during embryogenesis or other stages of development or in different cell types.

Whilst the theoretical behaviour of underdominance has been run through various models and simulations, data from actual laboratory experiments is limited. Reeves et al. genetically engineered the fruit fly Drosophila melanogaster with a ‘one-locus’ system, using an RNAi transgene for blocking a vital ribosomal protein gene, but also adding an RNAi-resistant version of the ribosomal protein gene as the
‘antidote’ (rescue). In this system, offspring will not die but be weakened if they have only one underdominance construct. This experiment is viewed as a laboratory proof of concept for this particular type of drive, with potential for application in other species. However, it would “require releases to exceed an allele frequency of 61% in a given wild population”, making this approach impractical for large populations (Reeves et al. 2014, 6).

A recent review on threshold-dependent gene drives divides underdominance into 4 subcategories, which includes CRISPRa (Leftwich et al. 2018). This mechanism does not use CRISPR/Cas9 as an RNA-guided homing-endonuclease gene drive mechanism, but rather uses a deactivated form of Cas9 (dCas9) that has been modified into a ‘transactivator’. The transactivator is capable of triggering the overexpression or the untimely (ectopic) expression of chosen genes, which results in the death of the organism (Waters et al. 2018). It is devised as a toxin-antidote gene drive system, currently under development. First experiments in D. melanogaster however evidenced that much more research and understanding is required (Waters et al. 2018).

4.3.3 Toxin-antidote based drives

These systems use combinations of toxins and antidotes to achieve an inheritance bias. If we use these terms ‘toxins’ and ‘antidotes’ in the broadest sense, such combinations could, for example, consist of mechanisms to silence a vital gene (e.g. via RNAi), and then provide a replacement gene as an antidote that will not be silenced by the mechanism; or it could use a toxic protein neutralised by an enzyme, or by an RNA-based silencing mechanism, that will stop the production of the toxin.

In order for such gene drives to work, the presence of the ‘toxin’ needs to create a serious disadvantage (e.g. death) that can only be remedied if the ‘antidote’ is available. If the antidote is not present, the cell or organism will die. There are two main ways to achieve this: either by physically separating the genes for the toxin and the antidote so they will not automatically be inherited together; or by a time separation of the activities of the respective genes, combined with using a long durability toxin product. A payload gene could be tightly linked to the antidote gene and thus achieve drive.

In the first case, if the toxin and antidote genes were placed on different chromosomes, it would be easy for an organism to just de-select the toxin gene and make it disappear from a population, thus halting the drive mechanism, something that could happen rather quickly. This theoretical model was put forward by Gould et al. in 2008 as the ‘killer-rescue’ system, regarding its weakness as a benefit (Gould et al. 2008). Suggested as a ‘self-limiting’ system, it would place a time limit to the lifespan of the drive, although the organisms genetically modified with both the rescue antidote and the payload gene would not necessarily vanish but possibly remain in the population.

In a different form, the toxin-antidote system could also be utilised to achieve Underdominance (see above), in situations where the toxin is tightly linked to an advantage, for example, carrying an antidote for a different toxin. In that case, a toxin could be produced in females during oogenesis, i.e. during egg production. If this substance is toxic
only at a later stage, such as during embryonic development, and if it is able to remain present into that stage of development, then only those offspring with the antidote gene will survive. In this scenario, the antidote gene would become active only in early embryogenesis, counteracting the toxin and thus rescuing the embryo. Examples of this are the UD^MEL underdominance design mentioned above and especially the Medea system detailed below, together with other variations, such as inverse Medea, Merea, Semele and Medusa.

Toxin-antidote components are found as part of other gene drive designs and systems, in particular underdominance and sex-ratio distortion. Increasingly, CRISPR-based nucleases are incorporated as toxins in these theoretical designs, with modified resistant genes added as the antidote.

a. Medea, Merea, inverse Medea and Semele

The following are all single-construct designs, meaning all genetic elements involved are tightly linked and transfer as a unit. All, with the exception of Medea, are theoretical designs.

Medea stands for ‘maternal effect dominant embryonic arrest’. It is a selfish genetic element, in which the female will make a toxin during egg-production (oogenesis) that will lead to the death of the embryos--unless any of them has inherited a copy of the Medea element from its mother or father--as this also holds the antidote within the same element.

This phenomenon was first discovered in the flour beetle Tribolium castaneum (Beeman, Friesen, and Denell 1992) and takes its name from Greek mythology, where Medea is said to have killed her own children (with ancient sources differing as to whether by intent or accident).

In 2007, researchers from the California Institute of Technology, Pasadena, genetically engineered the first gene drive system, based on the principles of Medea, which is basically a toxin-antidote system (Chen et al. 2007). They did so in the model fruit fly Drosophila melanogaster. Using microRNA as the toxin to silence an essential embryonic gene (here Myd88), the antidote was the same embryonic gene, but modified with an altered sequence so it could not be silenced by the microRNA.

At least according to the models, Medea is regarded as a strong drive system that could spread payload genes rapidly, so long as it is released at high frequencies and the fitness cost is kept low (Sinkins and Gould 2006; Akbari et al. 2014). Recent laboratory experiments carried out with the spotted wing drosophila (D. suzukii), an agricultural pest in soft fruit production in California, confirmed the need for high release frequencies, and also showed in long term cage trials that selection for resistance to the microRNA-based toxin being used is a concern (Buchman, Marshall, et al. 2018).

Different possible Medea variations have been suggested and modelled, e.g. (Akbari et al. 2014). In fact, there is a multitude of different systems inspired by these Medea principles, the closest of which are Merea, where the gene for the antidote is recessive, and inverse Medea, where the toxin is produced during early embryonic development, that is, unless the antidote was produced maternally during egg production (Marshall and Hay 2012b, 2011). These are all theoretical designs used for modelling of engineered gene drives in order to see if, for example, payload genes would easily find fixation in a population, gene drives would be less invasive, or suppression could lead to population collapse.

Semele is yet another design variant of Medea, except that in its case the toxin is produced during sperm development, so it is the father killing the offspring (Marshall et al. 2011).

b. Medusa

Medusa is a two-construct design within the toxin-antidote system. It has not gone past the model stage and again is simply a theoretical design intended for population suppression, in which a population crash might be kept to geographical limits (Marshall and Hay 2014).
Medusa is made up of four components, two toxins and the two respective antidotes. One toxin and the antidote will be located on the X-chromosome, the second toxin with the antidote to the first will be located on the Y-chromosome. One without the presence of the other could therefore not survive. This system will thus select for individuals with both the transgenic X and Y chromosome, thus selecting against females (XX); and, if initially released at a sufficiently high frequency, could bring the population to collapse.

5 Gene Drive categories and attributes, their limitations and risks

There are many ways to categorise and compare these various potential or theoretical gene drive systems. Depending on the purpose of such an evaluation or comparison, different attributes and parameters will be of importance. Some of these will largely be of interest to the developers, such as the ability (or inability) of each system to stably carry payload genes, or its susceptibility to resistance and inactivation.

Naturally, a quite different selection or combination of attributes and parameters are considered when the main goal is to understand the risks to the environment, health or biodiversity, that are, or can be associated with, a particular (potential) gene drive system. Parameters of relevance here are such qualities as invasiveness, the potential for global spread, the speed of spread, the lack of reversibility or removability, horizontal gene transfer, and the potential for eradication (suppression) or alteration (replacement) of a population/species.

A fundamental difficulty presents itself when different gene drive systems are assessed within the various parameters. The main problem is that there is very little reliable data. Only a few gene drive systems currently have a proof of concept, and these are restricted to laboratory conditions and largely to laboratory strains (with the exception of Drosophila suzukii). Many gene drive systems are merely at the design stage or in an early state of development and assessments of them are thus also largely theoretical. Whilst a gene drive system may have a specific design, current evaluations are based on the assumption that the gene drive system, once engineered into a gene drive organism (GDO), will behave and perform as designed. Some will draw attention to this, like Champer et al. who state in their table 1: “The characteristics listed here are variable and depend on a range of factors (for example, ecology of the target species, population distribution, movement patterns, fitness costs, payload characteristics, and so on); therefore, only ideal-case scenarios are compared to emphasize intrinsic differences of the various types of drives.” (Champer, Buchman, and Akbari 2016, emphasis added). This means that, overall, we are actually talking about “potential gene drive systems” (a term used by numerous authors when presenting their assessments).

A major question relevant to biosafety is whether a particular gene drive system can be confined, once it has been released or has escaped into the wild--or if its design will favour uncontrollable spread, with potential global eradication or permanent genetic modification of the entire species. In the following we will briefly introduce gene drive categories that are relevant to this question: a.) threshold-dependent drives; threshold-independent drives and temporally self-limiting drives; b.) suppression (eradication) vs. replacement or modification; and c.) recallability and reversibility.

As far as these drives rely on CRISPR/Cas for its ability to cause a DNA breakage at the site of a specific target sequence, the systems are vulnerable to the development of resistance.
5.1 Threshold-dependent, threshold-independent and temporally self-limiting drive systems

These categories have been given different names by different groups and authors. Min et al., for example, refer to these same categories as threshold, standard and self-exhausting drive systems (Min et al. 2018). We are using here the terminology chosen by Marshall & Akbari in their 2018 review entitled “Can CRISPR-based gene drive be confined in the wild? A question for molecular and population biology” (Marshall and Akbari 2018).

Predictions and statements made to date rely heavily on modelling, which again relies heavily on population biology, meaning that answers differ from species to species or even from population to population, as well as from ecosystem to ecosystem. In fact, many more factors will come into play. Looking at threshold-dependent drive systems, Gould’s group found “that to determine the best method of spatial release, and the total number of engineered insects that must be released, it is important to take into account the age and sex of the released insects and spatial structure of the population.” (Huang et al. 2011, 415).

Another major point is that of mating behaviour. Many models assume (implicitly) a deterministic representation of a randomly mating (panmictic) population (Edgington and Alphey 2018), which may well not reflect reality. Modelling outcomes will be different, if, for example, assortative mating (non-random) and polyandry strategies (female mating with multiple males) are taken into consideration eg. (Bull 2017). Leitschuh points out that wild rodents will “exhibit mating strategies such as polyandry and assortative mating, ...and have seasonal population fluctuations...”, while laboratory rodents have very controlled reproductive environments.” (Leitschuh et al. 2018, S132). There is real concern that sexual selection might develop against drive-carrying individuals: “The costs associated with drive create a benefit to avoiding mating with individuals carrying a driver, and thus preferences against driver carriers are expected to evolve...” (Lindholm et al. 2016, 322) – as found for example in stalk-eyed flies (Johns, Wolfenbarger, and Wilkinson 2005, Cotton et al. 2014) or in mice carrying the t-complex (Manser, Konig, and Lindholm 2015). The mechanisms behind this are not understood.

It is important to remember that any predictions and assumptions made about gene drives and GDOs are most likely not realistic.

Threshold-dependent drive systems

Depending on its frequency level when being released, this category of drive will spread into a population and achieve fixation, unless it is below the threshold frequency, in which case it will quickly vanish from the population (Davis, Bax, and Grewe 2001, Min et al. 2018). The determining factor here is quantity. Examples are engineered underdominance, Medea, or autosomal X-shredder (see Table 2).

According to some models, this drive category offers local confinement with local fixation. It is being argued that the dynamic assumed in simple population models may not hold true in the wild. Marshall and Akbari state, “…whether this holds true or not depends crucially on the dispersal patterns and population structure of the species being considered.” (Marshall and Akbari 2018, 426). There may well be numerous other hurdles to this category functioning as intended, for instance selective mating behaviour.

‘Dilution’ has been suggested as a ‘remedy’ to counteract the drive and its effects, but that would necessitate a large-scale release of wild specimens (see ‘reversibility’ below).

Threshold-independent drive systems

This category of drive does not require a specific minimal frequency for its proliferation. Instead, it can spread from an initially very low occurrence. It is characterised by high invasiveness and high risk of spreading throughout populations, affecting a targeted species and its linked ecosystems globally. As mentioned above, this category is also referred

---

26 Frequency here means the proportion that the released GDOs constitute as compared to the whole of the existing population.
to as ‘standard’, with other researchers also calling it a ‘self-sustaining drive, or a ‘global drive’ (Delborne et al. 2018) as well as ‘global drive system’ or ‘global gene drive’ (Noble et al. 2016).

For example, the engineered homing endonuclease drives, especially the CRISPR/Cas based homing drives, and the Y-linked X-shredder are all threshold independent gene drives (see Table 2).

Whilst there exist proof of concept from laboratory experiments for some of these, it is not clear at all whether or how this will perform in the wild. Success is debatable, but the risks of negative impacts are not; they are indisputably serious. It is in this context that a number of researchers are clearly indicating that this category of drive should not be attempted or released into the wild unless it can be stopped from spreading and reversed. One statement reads: “Before robust and efficient homing-based gene drive systems can be implemented in the wild, tools are required to remove the effector gene and possibly the entire drive system from the environment in the event of unwanted consequences.” (Marshall and Akbari 2018, 427).

There are no such tools or countermeasures currently available, and none of the current conceptual models are capable of even hypothetically restoring the populations to a non-GM (and non-GD) population.

**Temporarily self-limiting drives:**

This category is highly theoretical, and conceptual models such as the so-called ‘daisy chain drive’ will be described in Section 6. To summarise, the idea behind this category is that a synthetic drive can be designed that will stop functioning after a given number of generations, for example by including elements with Mendelian inheritance. Whilst some suggest such drives will therefore be transient, others counter this optimism, stating, among other points, that this will largely depend on the fitness cost. For example:

“For payloads that incur relatively low fitness costs (up to 30%), a simple daisy-chain drive is practically incapable of remaining localized, even with migration rates as low as 0.5% per generation.” (Dhole et al. 2018, 794)

Due to its theoretical time limitations and Mendelian non-drive component, there is a mistaken assumption that this means it will not be able to spread outside the target population. This result again is, entirely dependent on fitness costs (e.g. Dhole et al. 2018), as well as on biological and behavioural factors in a target population, such as dispersal rate.

**5.2 Suppression (elimination/eradication) vs modification**

**Suppression drives** are intended to reduce or eliminate a population. If combined with a threshold-independent drive, they may spread to a global scale and result in the eradication of an entire species. This may particularly become the case if the drive system itself has a low intrinsic fitness cost and no resistance develops to the drive (Champer, Buchman, and Akbari 2016).

**Modification drives** are intended to spread specific traits through a population, for example with the aid of payload genes. Again, if linked to threshold-independent global drives, and if it overcomes resistance problems, this may genetically modify an entire species. Modification drives are sometimes referred to as ‘alteration drives’, and occasionally as ‘replacement drives’. The latter is problematic though, as it is not quite accurate here and it confuses the issue with population replacement seen for the underdominance system. Such replacement would strictly mean to replace one population with the other, without those populations mixing, or without relying on the spread of the modification via homing-CRISPR/Cas9 drives. Replacement would require a large scale release of a modified population reared in laboratories and cages.

**5.3 Removability and reversibility**

The inability to predict the behaviour and consequences of a gene drive once it is released has
Table 2: Comparison of potential gene drive systems.
Please note, all table entries are not based on actual studies in the field, but are based on modelling in combination with laboratory findings or deductions. The entries thus largely reflect the potential if the specific gene drive system were to succeed in working as envisaged.

The table shows that hardly any gene drive system is confinable, with the potential exception of high-threshold lethal underdominance systems, which though whilst they may not mix with and spread into wild populations, would push them back and replace them.

(*1) large transposons will commonly not jump easily, and transposons are regarded as having a tendency to lose components and be mutated by host organism. PLG: payload gene.

<table>
<thead>
<tr>
<th>Transposable Elements</th>
<th>Threshold-dependent</th>
<th>Threshold-independent</th>
<th>Temporally self-limiting</th>
<th>Intended as Suppression drive (eradication)</th>
<th>Modification / replacement with payload gene (PLG)</th>
<th>Confineability</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>yes</td>
<td>-</td>
<td>-</td>
<td>Probably not (*1)</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>HEGs</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>Maybe (in theory - with tightly linked PLG)</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>CRISPR-HEG</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>Maybe (in theory - with added PLG)</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>autosomal X-shredder</td>
<td>yes</td>
<td>-</td>
<td>yes</td>
<td>-</td>
<td>medium-high (Mendalian inheritance)</td>
<td></td>
</tr>
<tr>
<td>Y-linked X-shredder</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>-</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>T-haplotype Medusa</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Medusa</td>
<td>yes</td>
<td>-</td>
<td>yes</td>
<td>yes, theoretically (esp. if PLG is placed between toxin &amp; antidote genes)</td>
<td>yes (though can replace population)</td>
<td></td>
</tr>
<tr>
<td>Engineered Translocations</td>
<td>yes</td>
<td>-</td>
<td>-</td>
<td>yes</td>
<td>yes (though can replace population)</td>
<td></td>
</tr>
<tr>
<td>Engineered Underdominance</td>
<td>yes</td>
<td>-</td>
<td>-</td>
<td>yes</td>
<td>depends on design &amp; threshold</td>
<td></td>
</tr>
<tr>
<td>(toxin-antidote) killer-rescue</td>
<td>yes</td>
<td>-</td>
<td>-</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical: (Daisy) Chain-Drive</td>
<td>(can in theory be engineered for that)</td>
<td>in theory</td>
<td>hypothetically</td>
<td>not necessarily</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chapter 1: What are Gene Drives?
led many researchers and scientists to call for the requirement that gene drives be stoppable and reversible, e.g. (Marshall and Akbari 2016) – at the very least in initial testing phases.

Given their potential to modify entire ecosystems\footnote{In Champer et al. 2016 its reference 40 & 41.}, the ability to stop and remove a gene drive system is seen by many as a necessary or at least highly desirable prerequisite for any release of a gene drive organism. This is particularly the case for any releases in the trial phase, where it is being highlighted as a ‘must have’.

It is important to understand here that frequently a crucial distinction is being made between ‘removability’ and ‘reversibility’, which differentiates between sequence reversibility and trait reversibility (Min et al. 2018). ‘Removal’ here means the ability to restore the population to its original wild-type state. In the case of high-threshold gene drives the suggested means to attain this is the large-scale release of wild-type organisms, which in itself is a serious challenge and may not be possible. ‘Reversal,’ however, is to genetically counteract and to block, disable or neutralise a gene drive system once released or escaped into the wild, by releasing an additional second gene drive as a “reversal drive”. This measure, however, does not mean that the original population will be reinstated. Crucially, Champer et al. point out: “Of note, despite their name, reversal gene drives do not restore the original modification to the wild type; rather, they induce further changes that may undo a phenotypic alteration caused by the initial gene drive.” (Champer, Buchman, and Akbari 2016, 148). In short, this means there are additional biosafety concerns to be addressed for “reversibility”.

It is important to note that some authors are not making this separation and may refer to both as ‘removal’ (Marshall and Akbari 2016).

\section*{6 Real problems and the search for safety}

In July 2017, the U.S. Defense Advanced Research Projects Agency (DARPA) announced a significant programme of research, named ‘Safe Genes’, that gives insights into concerns in the wider research community about the potential for irreversible and global effects from the use of synthetic/engineered gene drives. The programme itself directs considerable funding towards proving methodologies that might address these risks, to achieve “spatial, temporal, and reversible control of genome editors [including gene drives] in living systems” and to “eliminate unwanted engineered genes from systems and restore them to genetic baseline states” (DARPA 2017).

Reading this not through the perspective of a goal, but as a clear reflection of the real safety problems involved, and in particular (although not only) for the CRISPR/Cas9-based homing drive, it makes clear (to paraphrase) that: We do not know, nor do we have the technical tools, to reliably stop gene drives from spreading and being active once in the wild. There is no spatial or temporal control system or mechanism that has any demonstration of proof of concept. It further states that once the (genetic) changes have occurred in the wild there are no tools or mechanisms for undoing this, i.e. for reversing it and restoring a living system back to its genetic baselines.

Synthetic gene drives, and not uniquely but in particular CRISPR/Cas9-based homing drives, are very potent instruments, with potentially serious and significantly negative, even devastating consequences for biodiversity and ecosystems. These issues of lack of control over an instrument and of its technical limitations have already come up in previous sections of this chapter and are of deep concern to many.

In the following we will give some examples of CRISPR gene drive concepts and designs that have
been proposed as possibly providing some answers to the above concerns – should they themselves be made to work and be shown to work predictably. These examples are either DARPA-funded projects or related research.

6.1 Restrictive Gene Drives: (daisy) chain drives, split drives and global vs. local

The ‘daisy drive’ is a theoretical gene drive technology proposed by Kevin Esvelt, that theoretically would only ‘drive’ a gene into a population for a certain number of generations, and which is thus described as ‘temporally self-limiting’.

It is a variation on a CRISPR/Cas homing drive described in detail by Noble et al., where the different elements have been split up and spread across different chromosomes (Noble et al. 2016). In fact, the principle is quite simple.

There are three separate elements: element A: a payload gene that is inserted into location (L-a); element B: a gene coding for a CRISPR-based endonuclease designed to cut the DNA at location (L-a) and itself inserted at location (L-a); finally element C: a gene coding for another CRISPR-based endonuclease, here designed to cut the DNA at location (L-b) and itself inserted at location (L-c). All these elements can function in a chain-like manner: element A (the payload gene) needs a helper to achieve drive and to get itself copied to its ‘home’ location on the homologous chromosome; element B will be that helper, in that it cuts the DNA at the spot element A needs to be copied into; yet element B also requires a helper to achieve drive, which will be element C. The only element that cannot obtain drive in this design is element C, which will be passed on

Figure 7

Figure 7: Design of Daisy Chain Drive. In a Daisy chain no drive element works on its own. It is a series of split drives; the guide-RNA for A is placed in construct B, and the guide RNA for B is placed in construct C. Thus, in a chain, A needs B, and B needs C. The last element in the chain, here C, does not have a drive and will thus not be copied across to the homologue chromosome. Instead it will be passed on in a 50:50 ratio. Thus gamete a (e.g. sperm) will have a full daisy-chain, whilst gamete b lacks construct C, and thus the ability to drive the next element.
according to Mendelian Law. So, if a certain number of such gene drive organisms are released that are homozygous for all three elements, i.e. where all elements are present twice, then there will be drive for A and B but not for C. Once C gets thinned out in the following generations, then there will be reduced drive for element B, which will ultimately result in a loss of drive for element A.

If however recombination takes place between the gRNA of C with A, then this “could create a “daisy necklace” capable of self-sustaining global drive.” (Noble et al. 2016)

‘Daisy-chain drive’ is one of the potential methodologies being proposed which can supposedly limit the geographical spread of a gene drive and its trait(s), and so is portrayed as a potential means of alleviating concerns about gene drives uncontrollably spreading through a whole global population. No proof of concept of this technology has yet been published, though Esvelt’s team has secured significant funding to attempt to achieve their goal of such a local and reversible daisy-chain drive (DARPA 2017).

Whilst it is referred to as a ‘local’ drive by its designers, this may actually be somewhat misleading, as it is designed as a drive with temporal limited spread, a so-called ‘self-exhausting’ drive. Modelling studies carried out by researchers at North Carolina State University, Raleigh, suggest that there are flaws in this chain-drive concept once it is exposed to interbreeding with neighbouring populations. As mentioned previously, they concluded that daisy-chain will not be locally contained unless fitness costs are above 30% and migration rates are below 0.5% per generation (Dhole et al. 2018).

Furthermore, modelling studies also show that, should the drive perform as planned, a payload gene which has a 10% fitness cost may readily reach fixation in a population with an introduction frequency of as little as 3% (Dhole et al. 2018). There is also the potential for the genetic modifications, including payload gene, to spread more widely via normal Mendelian inheritance through the interbreeding of linked populations, depending on fitness costs.

A serious technical hurdle is the potential for rapid emergence of resistance, as two different CRISPR/Cas homing drives must be active for this drive design to work, and not just one, though the authors hope that the use of multiple sgRNAs may help the situation.

**Split drives & synthetic target site drives**

Concerns about accidental escape of GDOs from laboratories, transport or cage trials into the wild are shared by many, especially in terms of CRISPR-based homing drives. Given the possibility of human error or unforeseen natural events, Champer and his colleagues point out that relying on physical containment solely is insufficient and note: “Since very few escapees can establish an effective drive in a population (Unckless et al. 2015; Noble et al. 2018; Marshall and Hay 2012[a]; Marshall 2009), additional safety measures should be employed in any experiments with drives potentially capable of spreading indefinitely.” (Champer et al. 2019, 3). Both split drives and synthetic target site drives have thus been designed for this particular purpose, namely to add an additional layer of safety to experimentation with GDOs in the laboratory (Akbari et al. 2015; Champer et al. 2019).

In a **split drive** the endonuclease (Cas9) is physically separated from the drive construct (the guide RNA and potentially a payload gene), and both need to be present for drive to occur. In this system, the endonuclease is inherited according to Mendelian Law. A **synthetic target site drive** is designed to recognise a DNA sequence that has previously been added to the laboratory strain through genetic engineering but that will be absent in wild populations. Both these designs have recently been tested in the model insect *Drosophila melanogaster* (Champer et al. 2019).

Whilst this is a good step for increasing the safety of laboratory experimentation, it should be a binding requirement, and not a voluntary approach taken by some.
6.2 Gene Drives targeting geographic sequence variants

An alternative proposed means of limiting the spatial range of a gene drive is to identify geographic variations in genome sequences, and target a gene drive to such particular variants. Also funded by DARPA (DARPA 2017), this strategy is being explored by teams developing gene drives in mosquitoes (Wood 2017), mice and feral cats (AWC 2018), but again, no proof of concept in a laboratory has yet been published. Furthermore, unless the target gene is a highly conserved and essential gene (see Kyrou et al. 2018), resistance is bound to arise. Still, a highly conserved and essential gene would commonly have the same sequence across the whole species or even species group, thus being able to leak the gene drive quickly into the wider population. This leaves this approach clearly at the hypothetical stage.

6.3 Gene Drive ‘catchers’ – ideas and approaches for ‘anti-gene drive’-drives

The question of how to counteract a gene drive, whether it is a drive which behaves in unforeseen ways, has unpredicted negative impacts, is an unintended release, or is used maliciously, is evidently a key concern in the research and development community. The theoretically most plausible method if it can’t be countered with CRISPR inhibitors\(^{28}\), which seems somewhat of an unlikely strategy at eco-system scale, then perhaps the only imaginable strategy for counteracting a gene drive is to release a second gene drive to ‘catch’ it, an ‘anti-gene drive’. DARPA is clearly aware of this and is directing funding to developing “drives that can overwrite every copy of a ‘rogue’ gene drive” (Esvelt 2017) through their $65 million ‘Safe Genes’ research programme. Again, no proof of concept in a laboratory has yet been published.

6.4 ‘Immunising’ drives

A yet different idea is that of target prevention. In the case of an unwanted or problematic CRISPR/Cas-based homing drive being ‘on the move’, the idea would be to release a separate synthetic gene drive that could over-write the target sequence of the first gene drive, thus ‘immunising’ the population (Esvelt et al. 2014). This would be an approach that relied on genetically modifying a wild population with, once again, unforeseeable consequences. It is for example unclear what the implications are of leaving active Cas9 constructs/endonucleases in a population, and whether that might produce a background toxicity or give rise to off-target effects. Equally it is not understood what the chances are of accidentally arriving at an equilibrium between two counteractive gene drives and its possible consequences. (Vella et al. 2017). Again, no proof of concept in a laboratory has yet been published.

7 Summary and conclusions

As discussed in this chapter, gene drives are defined as a genetic element or mechanism that imposes a greater than 50% inheritance rate of itself or an associated trait, even if this inflicts a high fitness cost on the organism. Such elements and mechanisms have been found or observed in nature, and their roles are as yet not really understood. Termed ‘selfish genetic elements’ from early on, some of these, for example the over-replication ability of transposable elements, are increasingly regarded as vital components for genome evolution and even speciation (Biemont 2010). John H. Werren for example noted: “The story that is emerging increasingly supports a central role of SGEs [selfish genetic elements] in shaping structure and function of genomes and in playing an important role in such
fundamental biological processes as gene regulation, development, evolution of genetic novelty, and evolution of new species.” (Werren 2011, 10863). In fact, the study of these elements and the processes involved are now contributing to an emerging re-think of what a genome is and how it interacts with its environment. (Werren 2011; Lindholm et al. 2016).

The origin and design of synthetic - or engineered - gene drives is based on these genetic elements and mechanisms found in nature. There are basically two main categories of gene drives. The “over-replicators” actively pursue their own multiplication within a species by copying themselves directly across to the genomes of the next generation without passively awaiting their distribution during sexual reproduction according to Mendelian Law. These are the transposable elements and the homing endonuclease genes (homing drives). Then there are the ‘transmission distorters’, that will seriously disadvantage or eliminate ‘competitors’, such as sibling gametes or embryos that did not inherit the specific genetic element (e.g. some sex ratio-distorters), or that will weaken or kill the developing organism if an element is not inherited from both parents (underdominance drive). A major instrument with the genetic ability to kill rival cells is the toxin-antidote mechanism, which can be used in almost any transmission distorser system.

There are real complexities involved in the different mechanisms and systems described in this chapter. To imagine and build functional synthetic gene drives has long eluded the abilities of researchers. This changed completely with the arrival of the genome editing tool CRISPR/Cas9. This versatile endonuclease is currently being employed in almost all categories of gene drives. Whilst it can be used as a ‘toxin’ in the toxin-antidote mechanism, it is most prominent in the CRISPR/Cas-based homing drive (also known as the RNA-guided homing drive), and the CRISPR/Cas-based ‘X-shredder’, both of which quickly attained proof of concept in laboratory settings with model organisms.

The arrival of CRISPR/Cas thus seemed like a breakthrough, making gene drive construction appear simple and easy, and with this came many claims, hopes, dreams, promises and projections - and funding. At the same time, the discussion began to focus on the risks, complexities and limitations that were rapidly emerging from the research.

What follows is just a summary of concerns that voiced in the literature as well as at the CBD and other bodies. For example it was recognised from early on that the concept of the CRISPR/Cas-based homing drive, one of the most advanced and potent systems so far developed, poses major risks precisely because it is a global gene drive, because it is self-sustaining and thus threshold-independent. Its potential power to eliminate or modify means that there is no room for errors in the technology or for unintended effects on the target species or the ecosystem into which it is released. There is also no room for accidental escape from labs or cages, unintended spread, or crosses into closely related species, nor for a spreading payload gene to turn out to be problematic in the wild, or change a species such that it becomes invasive, a stronger or different pest or vector. There must also be no negative impacts on the resilience of ecosystems, or on biological diversity including agricultural biodiversity, food systems or human health, livelihoods, and cultural practices of indigenous peoples and local communities.

As pointed out, research is also revealing serious limitations and malfunctions to this technology, such as its inefficacy in many organisms, the rapid emergence of resistance, off-target effects, irreversibility and the impossibility of containment or recall once released. In view of both risks and limitations, this technology as it stands is not fit for application.

So are these issues being addressed? Major efforts are being undertaken to circumvent or overcome resistance. The other issues, so far, are stuck at the stage of theoretical models and designs, such as the various daisy drive designs, or the ‘anti-gene drive’-drives, e.g. immunisation drive, reversal drive, drive catchers.

All these efforts are still lacking proof of concept and often merely exist in computer modelling,
which carries its own limitations. It is, however, important to recognise that, as the technology develops, so new problems and challenges emerge, requiring new layers of ‘solutions’, which in turn carry their own risks and limitations, and add to the overall complexities involved in assessing both performance and impacts of proposed applications.

It also applies to other gene drive systems with proof of principle that are being followed up, such as the CRISPR/Cas-based X-shredder (a potentially self-sustaining drive system), underdominance (Buchman, Ivy, et al. 2018) and Medea (Chen et al. 2007; Buchman, Marshall, et al. 2018), which will also have their own limitations and risks.

Whilst this, our report’s first chapter, has looked at the technical and technological aspects of gene drives and related elements, it is still impossible to say very much about either the actual performance or the potential impacts of release under real life conditions, e.g.: on high genetic variation in wild populations, or on interactions with other species and response to the complexities of ecosystems. The behaviour of gene drives and gene drive organisms in such settings may be very different from laboratory experiments and modelled predictions, thus adding an extra layer of risk whose nature and gravity may be impossible to accurately predict in advance.


Curtis, C. F. 1968. “Possible use of translocations to fix desirable genes in insect pest populations.” Nature 218 (5139):368-368a0. doi: 10.1038/218368a0.


human malaria mosquito.” *Nature Communications* 5:8. doi: 10.1038/ncomms4977.


Chapter 1: What are Gene Drives?


Chapter 1: What are Gene Drives?


Chapter 1: What are Gene Drives?


