CRISPR Technology Development: Gene Drives and Genome Editing

THE UNIVERSITY of ADELAIDE

Chandran Pfitzner

Department of Molecular & Cellular Biology

School of Biological Sciences

University of Adelaide

October 2020

1 CONTENTS

2 THESIS DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Chandran Pfitzner

16/10/2020

3 ACKNOWLEDGEMENTS

Firstly, I would like to thank Paul Thomas, I feel incredibly lucky to have had such an excellent supervisor. From the first time talking with you in undergrad and your excited discussion of the potential for gene drives, your passion and enthusiasm for science was obvious and I knew you were the kind of scientist I aspired to be. As my supervisor, you exceeded my expectations with your endless support, advice, and encouragement, I could always count on you to make sure I had everything I needed to get the science done and to grow into a scientist myself.

To my dearest friends Ella, Louise, Mel, and Ruby, I have a million things to thank you for, both on a personal level and professionally. Your intelligence, capability, experience and friendly willingness to help or provide advice on any manner of problem made my PhD a fantastic experience. Even if I hadn't come out of this as a scientist, my life would have been enriched and improved from having had you in my life. You were the best group of lab mates I could have asked for, thank you.

Thank you to all the other Thomas lab members, past and present who I have shared a great camaraderie with and have been a great source of insight into technical problems and overall have just been a good group of friends to hang out with. To James, thank you for kickstarting my lab life, your knowledge and experience with all manner of techniques and theory was invaluable. Sandie, thanks for the many, many mice you generated and cryopreserved for me, but more importantly thanks for being the number one source for wholesome podcast recommendations. Josh, thanks for all the games nights, bouldering, and various other shenanigans we've been doing since well before you joined the dark side. Luke, thanks for all the late, late Friday knockoffs. Adi, thanks for being an inspiration and diving head first after your ideas and then giving me the chance to join in. Stef, thanks for sharing your TC knowledge and your keen sense of comedic timing. Mark, thanks for providing so many essential resources to keep myself and the lab going. Michaela, thanks for providing so many actual resources to keep myself and the lab going. Jayshen, thanks for always looking over everyone in the lab, your laser-sharp knowledge of music and movies will be missed. Caleb, thanks for your ever-present school spirit, keep being inspirational. Ash, thanks for bringing that fiery, competitive edge into the lab. Dale, thanks for leaving behind your legacy, I always felt like I had to live up to it, hopefully I got at least part way there. Alvaro, donde esta la biblioteca? Thank you to the other lab members Gelshan, Athena, Connor, Dan, and Eli who I unfortunately didn't get to know so well or who have long since moved on. And finally, thanks to my co-supervisor Phill Cassey for bringing another perspective into my research and always offering to help.

4 ABSTRACT

The broad theme of this thesis is the development of CRISPR-based genetic engineering technology, primarily focusing on an exploration of mammalian gene drives.

The CRISPR/Cas9 system utilises a complex composed of the Cas9 nuclease for DNA cleavage and a guide RNA (gRNA) for targeting it to a specific genomic locus. Since its revolutionary discovery and utilisation as a genome editing tool, one pioneering application is the CRISPR-based gene drive: The insertion of the genes for both Cas9 and the gRNA into a specific chromosome in an animal such that the gRNA targets the homologous locus of the wild type (WT) chromosome. In the offspring of a cross with a WT animal, the gene drive is initially hemizygous. Subsequently the nuclease and gRNA complex together and cleave the WT chromosome, resulting in copying of the nuclease and gRNA genes into that WT chromosome via homology-directed repair (HDR), termed "homing" in the context of gene drives. When viewed at a population level, this results in the rapid spread of the gene drive throughout a wild population.

Due to this "Super-Mendelian" inheritance, a gene drive offers the potential to modify entire wild populations. This opens numerous possibilities such as the eradication or suppression of populations of invasive pests or immunising natural populations against human pathogens such as malaria in mosquitoes. These are extremely powerful outcomes that could reduce human disease burden, reverse the devastating impact of invasive pests on ecosystems, or greatly reduce the agricultural cost of dealing with pests.

Gene drives have been experimentally tested in a small number of species including the fly *Drosophila melanogaster*, the yeast *Saccharomyces cerevisiae*, and the three mosquito species *Anopheles stephensi*, *Anopheles gambiae*, and *Aedes aegypti*. All of these have had a very high homing rate. A low rate of homing has also been observed in *Mus musculus* in the female germline but otherwise no vertebrates have experimentally developed gene drives.

This thesis describes the generation of four experimental gene drive approaches in mice, two of which used Cas9 as the nuclease under the control of either zygotic (CAG) or germline (Vasa) promoters, and another two that used Cas12a with either zygotic (CMV) or germline (Vasa) promoters. Gene drives were constructed with the key safety features of a "split drive" and a "synthetic target" to avoid any ecological impact in case of accidental release.

Homing did not occur at any detectable rate in any of the gene drives. Both the Cas9 zygotic-homing gene drive and the germline-homing gene drive in males showed a high percentage of indels at the synthetic target, indicating a high rate of Cas9-induced cleavage. It was concluded that zygotic-homing likely failed to occur due to lack of proximity between the gene drive chromosome and the synthetic target chromosome, as they remain separated a full 18-20 hours post-fertilisation until after the first G_2 phase.

Germline-homing likely didn't occur in the males as Vasa-induced expression begins during a period of mitotic proliferation of the primordial germ cells, a cellular state that likely doesn't promote the HDR required for homing. Contrasting this, the female oocytes are undergoing meiosis at this time point, where the homologous chromosomes are aligned and in an ideal position to promote HDR. However, Cas9 expression levels in the female germline were very low and likely reduced the chances of any homing occurring.

The Cas12a gene drives all failed to generate an appreciable level of Cas12a cleavage (0-4.3% across all Cas12a lines), as evidenced by expression levels and percentage of indels observed. As such, the

mouse models used to test the Cas12a gene drive here were not sufficient to accurately assess its functionality.

This thesis also discusses the design and testing of a suite of all-in-one CRISPR gene editing plasmids that allowed one-step generation of said plasmids containing two unique, customisable gRNAs. These were all successfully made and showed consistent, simultaneous cleavage of multiple target sites within cell culture, allowing for multiple knockdowns, large deletions, or reduction of off-target cleavage via the use of the Nickase variant of Cas9.

5 PUBLICATIONS

This thesis is a combination of a conventional thesis and the following publications:

Adikusuma F*, **Pfitzner C***, Thomas PQ. Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. PloS one 2017;12(12), e0187236.

Pfitzner C, White MA, Piltz SG, Scherer M, Adikusuma F, Hughes JN, Thomas PQ. Progress toward zygotic and germline gene drives in mice. The CRISPR Journal 2020;3(5), (accepted for publication).

*Co-first authors

The following publication was also produced but does not contribute to this thesis:

Prowse TAA, Cassey P, Ross JV, **Pfitzner C**, Wittmann TA, Thomas P. Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. Proc Biol Sci 2017;284(1860).

6 INTRODUCTION

6.1 CRISPR TECHNOLOGY

6.1.1 Terminology

Generically, the term "CRISPR" (clustered regularly interspaced short palindromic repeats) refers to both the use of and the suite of genome editing and related molecular tools that are composed of CRISPR associated (Cas) proteins and the short RNA molecules that bind to them. Technically and historically, the term was only used to refer to a class of DNA repeats in prokaryotes which the above-mentioned RNA is derived from and function as part of an acquired immune system for prokaryotes.[1-3] In the context of this thesis, "CRISPR", "CRISPR-based", etc. will refer to the genome editing tools.

Homology-directed repair (HDR) is a term that encompasses several different DNA repair pathways, most of which utilise donor DNA as a template, and are discussed in section 6.1.7 in further detail.[4] Homologous recombination (HR) is a specific HDR pathway[4] which is often used in the literature interchangeably with HDR to mean all of the HDR pathways.

6.1.2 Origins

The first CRISPR system that was developed for genome editing applications utilised SpCas9, derived from *Streptococcus pyogenes*.[3] This provided a programmable and easy-to-use system for inducing double-stranded breaks (DSBs) in DNA for the purposes of creating targeted small insertions or deletions (indels) or inserting custom DNA sequences at specific genomic loci.[3, 5]

6.1.3 Components

The CRISPR/Cas9 system is composed of the nuclease Cas9 and an associated gRNA – a short, single-stranded RNA molecule (Figure 1).[3] When both are present in a cell's nucleus, Cas9 initially binds to the gRNA which induces conformational changes in Cas9 to allow it to bind to genomic DNA.[3, 6] This ribonucleoprotein (RNP) complex is able to target a specific genomic position by two different elements: Firstly, Cas9 recognises a protospacer adjacent motif (PAM) on the non-complementary DNA strand (composed of the sequence NGG for SpCas9).[3, 7] Secondly, the ~20 bp "guide" sequence at the 5' end of the gRNA binds to the complementary gRNA binding site on the DNA.[3, 6] Double-stranded DNA cleavage then occurs 3 bp upstream of the PAM by the Cas9 HNH nuclease and RuvC-like domains.[3]

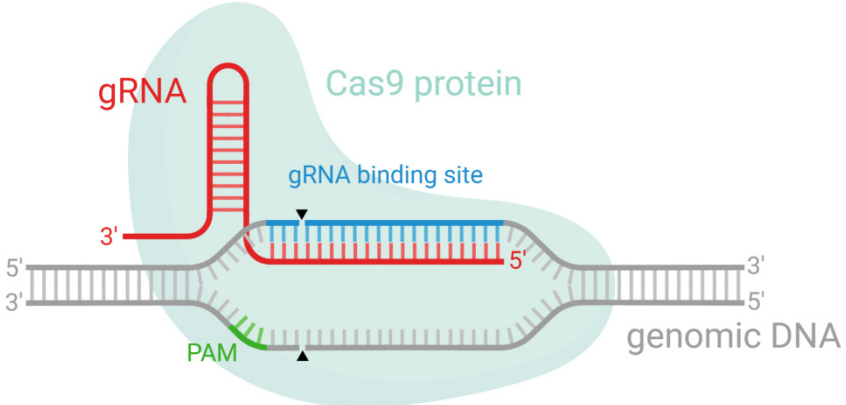


Figure 1. Schematic of CRISPR/Cas9 system. Cas9 protein (teal) is bound to a gRNA (red) and genomic DNA (grey).[3, 6] The PAM, with sequence NGG (green), is bound to the protein, and the gRNA binding site (blue) is bound to the ~20 bp guide portion of the gRNA. Black triangles show where DNA cleavage occurs.[3, 7] Image created with BioRender.[8]

6.1.4 Targeting Efficiency and Specificity

The efficiency of a particular guide to induce cleavage is influenced by a number of different factors including the epigenetic landscape of the target site and the composition of the guide itself.[9, 10] Guide composition can cause reduced or early termination of gRNA transcription, secondary structure formation which interferes with Cas9/DNA binding, or have inefficiencies due to unknown mechanisms.[9, 10] Ultimately the efficiency of a guide can only be experimentally determined but there are bioinformatic processes that have been developed to predict the efficiency.[9, 10]

Another consideration when designing guides is whether or not they will generate DSBs at unintended genomic sites (off-target cleavage). A basic local alignment search tool (BLAST) analysis of the sequence will ensure there aren't any identical off-target gRNA binding sites adjacent to PAMs.[11] It is more complex than this however, as there is a complicated relationship between the guide sequence and potential off-target sites that have one or more base-pair mismatches (MMs). It has been shown that the more MMs there are in an off-target site the less likely it will be cut. It has also been shown that off-target sites with MMs near the PAM are least likely to be cut, whereas sites with MMs further away from the PAM are most likely to be cut.[12-14] It is further complicated by whether there are multiple MMs next to each other and what the specific bases are but these details are yet to be completely understood.[12-14] In addition to this, Cas9 has been shown to have a very low level of activity with an NAG PAM site instead of NGG.[14]

In the case where there are potential off-target sites that may be cut, it is prudent to look at each specific genomic locus and determine if indels at that location would be deleterious or not. Oftentimes the sites will be intergenic or intronic and thus unlikely to affect the cell or organism's viability. There are now multiple tools available online that can perform most of the analyses described above including the initial design of the guide and assessing its potential cutting efficiency. The tools used throughout this thesis were CCTop (https://crispr.cos.uni-heidelberg.de/)[15], GT-Scan (https://gt-scan.csiro.au/)[16], Benchling (https://benchling.com/)[17], and the now-decommissioned Zhang lab tool (formerly at http://crispr.mit.edu/)[14].

6.1.5 Methodology

There are two primary uses for the generation of a DSB with CRISPR. The first is to induce the formation of a random, small indel via the cell's error-prone repair pathways (see section 6.1.7). This can disrupt a critical amino acid in a gene or cause frameshift mutations, both potentially leading to loss-of-function. The second is to provide a site for the insertion of custom donor DNA (see section 6.1.6) to generate gene knock-in models, tags, conditional alleles, and many other modifications.

For the generation of a mouse model, gRNA and Cas9 (protein or mRNA) can be delivered to cultured zygotes via microinjection, which are then transferred into pseudopregnant females and allowed to come to term (Figure 2).[18] Typically multiple different indels will be generated which can be easily separated by outcrossing to WT mice.[18] An added benefit of outcrossing to WT mice is that off-target mutations will be segregated away from the desired mutants at the same time, assuming the loci are not tightly linked.

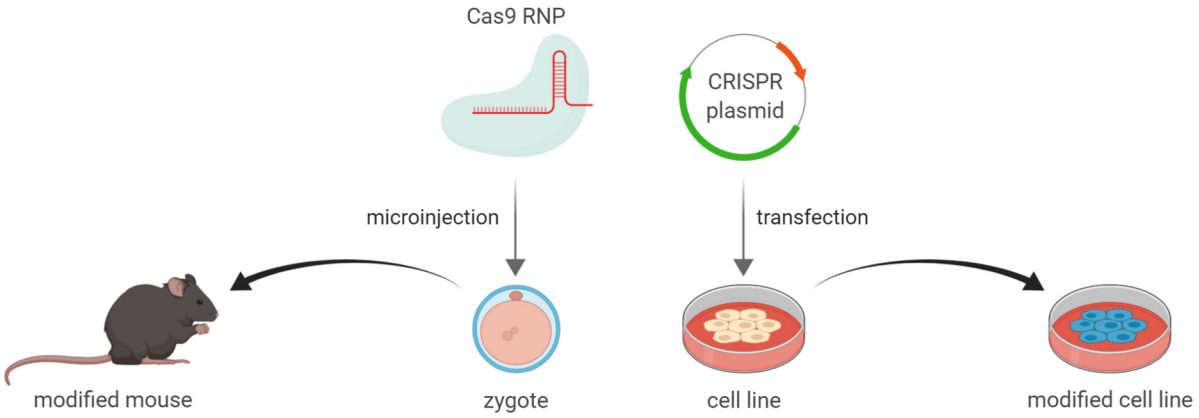


Figure 2. Examples of CRISPR/Cas9 delivery techniques and applications. Left shows delivery of a Cas9 protein pre-loaded with a gRNA into a mouse zygote via microinjection.[18] Right shows delivery of a CRISPR plasmid to cells via transfection, the plasmid(s) encode genes for both the Cas9 protein and the gRNA so the cells themselves will express them.[19, 20] Image created with BioRender.[8]

For cell work, CRISPR components can be delivered in a variety of ways. Most notably are all-in-one plasmids which contain a Cas9 gene under a suitable promoter (e.g. the cytomegalovirus (CMV) promoter for ubiquitous expression), a gRNA scaffold that allows easy cloning in of a guide (also under a suitable promoter such as U6), and a variety of different selectable markers.[19-21] These can be delivered to cells through typical methods such as transfection or infection.[22, 23] Once in the cells, Cas9 and the gRNA are expressed, allowing formation of the RNP complex and subsequent DNA cleavage, repair, and indel formation.[12]

6.1.6 Transgenics

As mentioned above, the other basic application of CRISPR is the ability to generate transgenic mice with the addition of donor DNA once a DSB has been generated. Technically, this is a separate technique from CRISPR, but the ease of which targeted DNA cleavage can now be accomplished with CRISPR has strongly tied the two together. Older methods to generate transgenic animal models such as gene targeting by homologous recombination could not be performed in zygotes.[24] For rodents it relied on very rare integration events in embryonic stem cells (1 in 10^6-10^9 cells) which then had to be transferred to blastocyst embryos to generate chimeras. Breeding of chimeras was then used to transmit the mutation to the next generation (a process that did not always occur). Typically, generation of mutant a mouse using ES cells took around a year but can now be accomplished with CRISPR in a few weeks.[24]

The basic concept is illustrated in Figure 3A where donor DNA in various forms can be delivered to living cells along with the CRISPR components. Whether it is linear double-stranded DNA (dsDNA), single-stranded DNA (ssDNA)[25] or circular plasmid DNA, all 3 contain the transgene of interest flanked by homology arms that are identical in sequence to either side of the CRISPR cut site.[5, 26] Whether these are delivered to cells in culture, zygotes, or specific animal tissues, the outcome is the same as per Figure 3B, in that the cell's HDR pathways can integrate the transgene into the same location as the CRISPR cut site.[5, 19, 20]

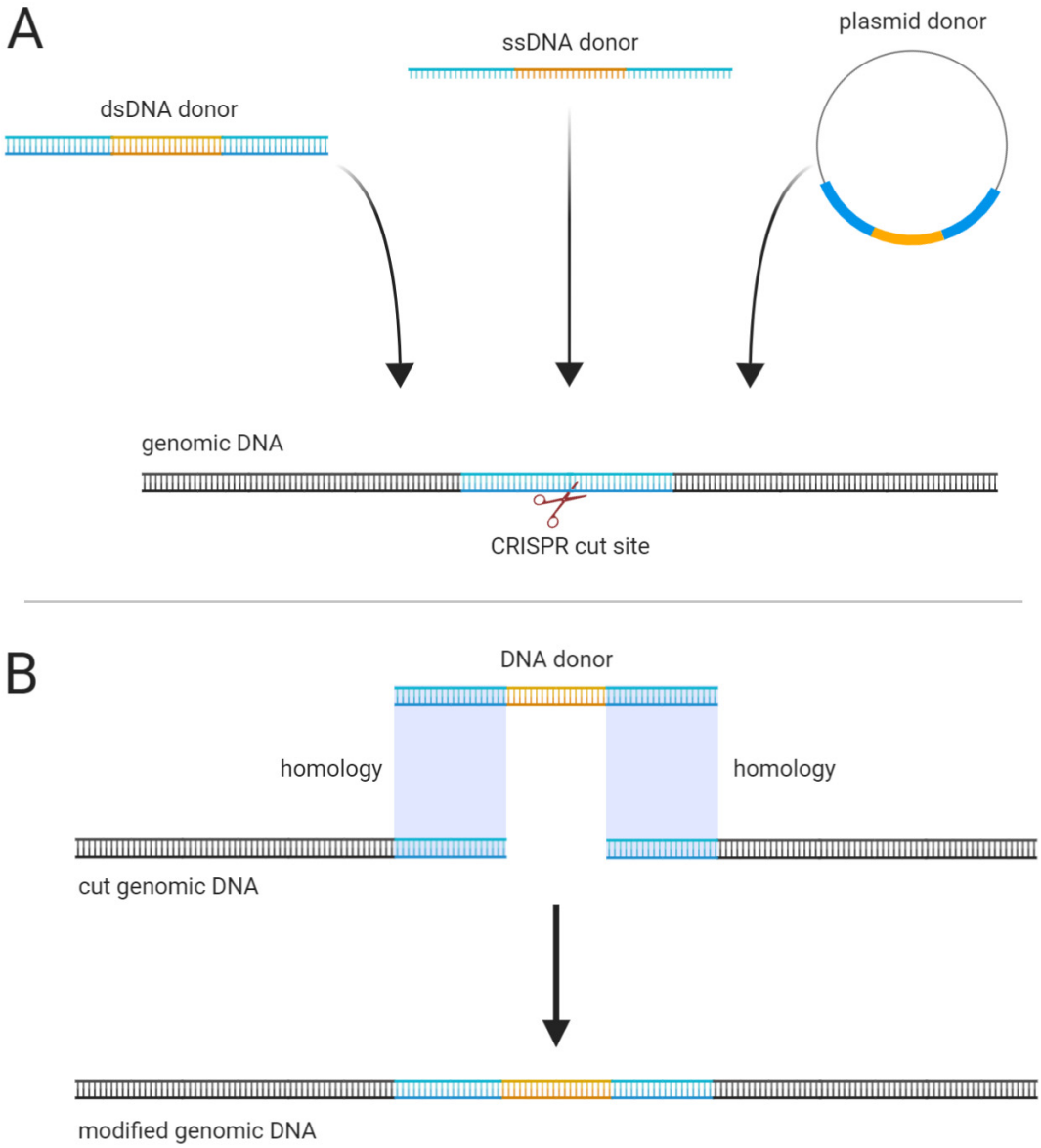


Figure 3. (A) Schematic showing different methods of custom DNA insertion. A custom DNA sequence (orange), flanked by homology arms (blue) which have identical sequence to the genomic regions flanking the CRISPR cut site.[5, 26] (B) After the genomic DNA is cut, the homology arms are utilised by the cell's HDR pathways and the custom DNA is inserted into the cut site.[5] Image created with BioRender.[8]

6.1.7 DNA Repair Outcomes

Transgene integration as depicted in Figure 3B occurs via HDR, which includes several different cellular pathways.[27] Although it is the desired outcome, it typically occurs at very low efficiency.[5] There are a handful of other pathways that predominate and they all lead to the generation of indels at the target site.[28] This includes non-homologous end-joining (NHEJ),[29] microhomology-mediated end-joining (MMEJ), also called alternative end-joining,[30] and single-strand annealing (SSA).[31] Although the latter is technically a type of HDR, it also leads to the generation of indels.[31]

The factors that determine the repair pathways and how exactly they function are not fully understood.[28] An overview of these pathways is shown in Figure 4, the major decision point between NHEJ and all other pathways occurs after the initial processing of the free DNA ends when a DSB is generated.[28] If the activity of specific 5' to 3' exonucleases proceeds on the DNA ends, termed

"resectioning" which leaves 3' overhangs on the DNA, then the non-NHEJ pathways take control. And if this does not occur, then NHEJ proceeds.[28]

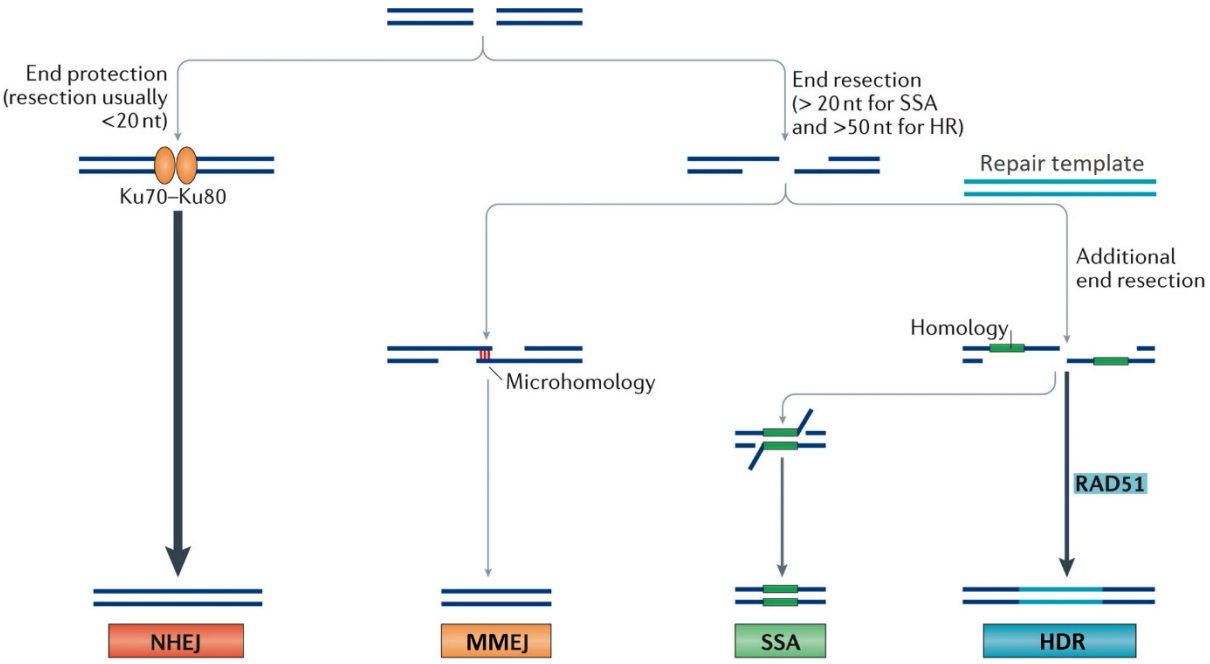


Figure 4. Schematic of major DNA repair pathways. DNA with a DSB shown at top, leading to the repaired DNA at the bottom. Image modified from Chang *et al.* (2017).

A couple of the reasons why NHEJ is favoured are that it is active in all phases of the cell cycle,[4] and it occurs more quickly than HDR,[4] completely repairing the DSB in as little as 30 minutes.[32] Initiation of NHEJ occurs as any free DNA ends are quickly bound by the Ku70–Ku80 complex (Ku)[33] which recruits a suite of nucleases, polymerases, ligases, and other proteins to catalyse NHEJ.[28] Although short resectioning can occur (<20 bp) with Ku bound,[28] it otherwise inhibits longer resectioning and thus inhibits HDR and MMEJ.[4] Enzymatic action at each end of the different DNA ends happens sequentially and can consist of a nuclease removing nucleotides, a polymerase adding nucleotides, or a ligase joining the ends together.[28] For an individual DSB, multiple rounds of each of these activities can occur, a different number of times for each enzyme and in different orders, resulting in multiple different indels generated from an identical DSB.[28] Although mechanistic details of why specific indels are generated from specific DSB events are not known, they can be somewhat reliably predicted via machine-learning algorithms employed by software such as inDelphi (https://indelphi.giffordlab.mit.edu/).[34]

MMEJ is similar to NHEJ in that it is an error-prone repair pathway that doesn't make use of a DNA repair template.[30] However, it also shares similarity with HDR in that it requires longer 5' to 3' resectioning (≥20 bp) to occur.[28, 35] Comparatively, the resectioning is a lot shorter than HDR,[35] although the recent demonstration of large deletions (several kb in length) after CRISPR cleavage may challenge that assumption.[36] With the free 3' overhangs present on both DNA ends, sections of microhomology between those two strands are now exposed, allowing them to anneal together.[30] At this point, if those microhomologies are present and have annealed, this prevents the HDR pathways from going forward.[4] After annealing, the MMEJ pathway proceeds, with the end result being the loss of all intervening sequence between the microhomologies.[30]

Likely a major reason why HDR is much less prolific than other repair pathways is that it is restricted to the S and G_2 phases during the normal cell cycle.[37] HDR also takes 7-8 hours to complete, a much

longer time in comparison to NHEJ/MMEJ.[4] After the initial short-range resectioning with the same exonuclease that is active in MMEJ, a separate exonuclease takes over to perform long-range resectioning of 1 kb or more.[38] At this point there are large sections of homology present between the template and the exposed ssDNA that facilitate the HDR process.[4] HDR then proceeds through a number of different pathways as shown in Figure 5.[4] Break-induced replication (BIR) proceeds when only one end of the DSB is found.[4] For synthesis-dependent strand annealing (SDSA), once a single strand is synthesised using the template DNA it dissociates and anneals to the other end of the DSB where synthesis of the second strand proceeds without involvement of the template.[4] In contrast, homologous recombination (HR) causes synthesis of both strands from the template DNA via a double Holliday junction, the same process that occurs in crossover during meiosis.[4] A key protein here is RAD51 which is responsible for the pairing of a repair template to the resectioned DNA and leads to BIR, SDSA, and HR.[39]

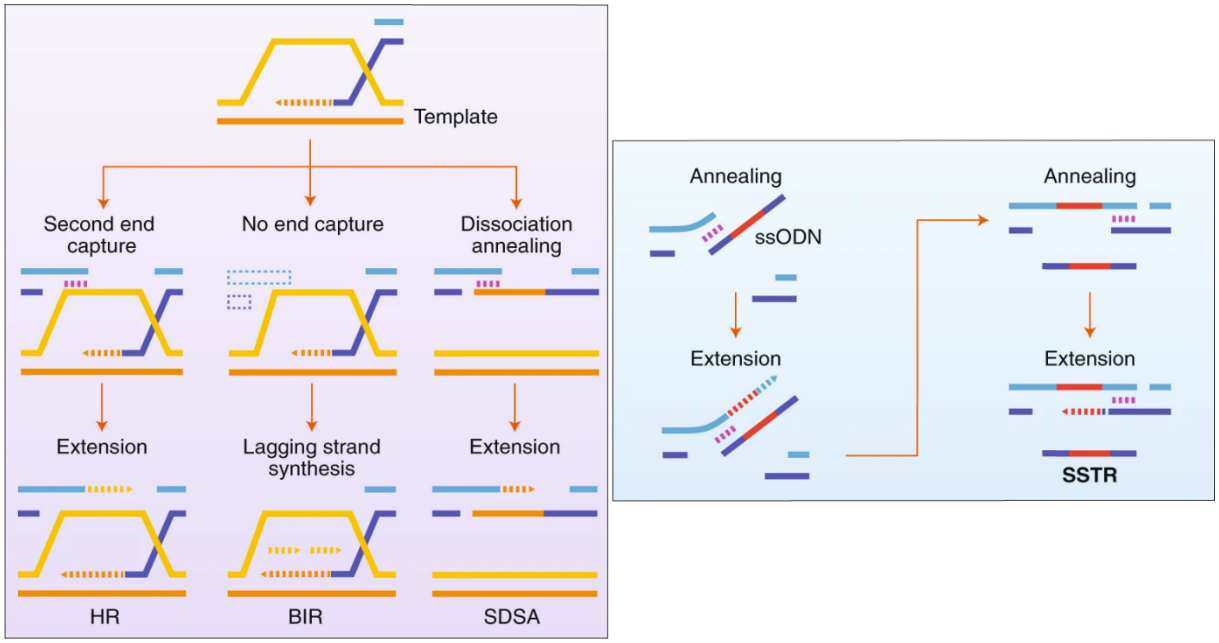


Figure 5. Overview of HDR pathways. Left: dsDNA-donor-templated HDR. Right: ssDNA-donor-templated HDR. Image modified from Yeh *et al.* (2019)

There are also two RAD51-independent HDR pathways that have been characterised,[4] SSA (Figure 4) is one of these that causes deletions and functions conceptually similarly to MMEJ[31] except with larger homologies (≥20-200 bp)[4]. The second RAD51-independent pathway is single-stranded templated repair (SSTR) which takes place when a synthetic ssDNA template is supplied.[40] A proposed repair pathway is shown in Figure 5 but the mechanisms have yet to be elucidated.[4]

6.1.8 Expanded CRISPR Toolkit

Although the CRISPR genome editing applications so far discussed are the most ubiquitously used, there is a constantly expanding toolkit making use of CRISPR for a myriad of different functions. SpCas9 homologs have been identified from many other species, these have different PAMs (e.g. `NNGRRT` PAM from SaCas9)[41] and different cleavage mechanisms (e.g. Cas12a which leaves a staggered DSB with 5' overhangs)[42].

Many modified CRISPR nucleases have also been developed. Two "Nickase" variants of SpCas9 are the D10A mutant which inactivates the RuvC nuclease domain[3] and the H840A mutant which inactivates the HNH nuclease domain[19] so that they only induce single-stranded cleavage ("nicking") instead of double-stranded cleavage. Nickases are a valuable tool as it is possible to use two

different guides with closely located targets to induce a DSB.[43] This allows the reduction of off-target cutting as any off-target sites for either guide are likely not going to be located anywhere near each other and therefore only cause nicking of the DNA and thus no DSB and consequential indel formation.[43]

To expand the targeting range of Cas9, many modified Cas9 variants have been developed to use different PAM sites. Initial variants included SpCas9 VQR (NGAN PAM), SpCas9 VRER (NGCG PAM), and SaCas9 KKH (NNNRRT PAM).[44] Further modifications led to the development of two different nucleases that recognised a simple NG PAM: xCas9[45] and Cas9-NG[46]. Further to this, reducing the prevalence of off-target cleavage via the generation of "high fidelity" Cas9 variants has been one of the key goals of Cas9 modification. A number of these have been generated now including SpCas9-HF1[47], eSpCas9[48], HypaCas9[49], evoCas9[50], and Sniper-Cas9[51].

To address the need to perform a precise, single-nucleotide conversion, techniques of "base editing" have been developed which utilise the CRISPR platform. Initial base editors were able to convert cytosine (C) to thymine (T) or guanine (G) to adenine (A).[52] A "dead" Cas9 (dCas9) which incorporates both Nickase mutations mentioned above was conjugated to a cytidine deaminase.[52] A gRNA then directs the dCas9-cytidine deaminase fusion protein to the target location and the deaminase operates on the small stretch of unpaired ssDNA that has been made available by the activity of dCas9.[52] Among other improvements, switching from dCas9 to Nickase to induce ssDNA cleavage on the non-target strand pushed the cellular repair processes in favour of the edited base when resolving the mismatch.[52] Similarly, A to G and T to C base editors were generated by directed evolution on a transfer RNA adenosine deaminase that was conjugated to the dCas9 in the same manner as the previous base editors.[53]

To complement base editors the prime editing system was developed that can perform all base editing transition and transversion mutations in addition to performing small insertions and deletions.[54] The base editing transition mutations already possible with the base editors could also be done with prime editing but with variable efficiency such that base editing is not obsolete.[54] Prime editing functions by replacing the gRNA with a prime editing guide RNA (pegRNA) that is identical to a standard gRNA except the 3' end is extended to include a primer binding site (PBS) homologous to the gDNA adjacent to the gRNA target site, then following the PBS is a template containing the intended edited or deleted base(s) surrounded by short regions of further homology to the gDNA.[54] Prime editing utilises a Nickase conjugated to a reverse transcriptase (RT).[54]

When delivered to cells (Figure 6), the Nickase binds the target site and nicks the gDNA, the PBS then anneals to the homologous region adjacent to the nick and the RT.[54] Using the free, 3' end of the DNA as a primer, RT incorporates the template into that strand.[54] As per Figure 6, the DNA is left in a state of equilibrium before cellular processes remove the flap DNA and ligate the nicked DNA.[54] This can be pushed in the favour of incorporating the edit by including a silent "PAM-killing" mutation such that if the template is incorporated, the target site is lost and can't be cut again, but if the template is not incorporated the Nickase + pegRNA is likely to bind and cut again.[54] Resolving the heteroduplex that is left behind when the template is incorporated can then also be pushed towards favouring maintaining the templated edit by inducing a second Nick on the other strand at a nearby location, allowing the cell to repair that strand with the edited side of the heteroduplex as a template.[54]

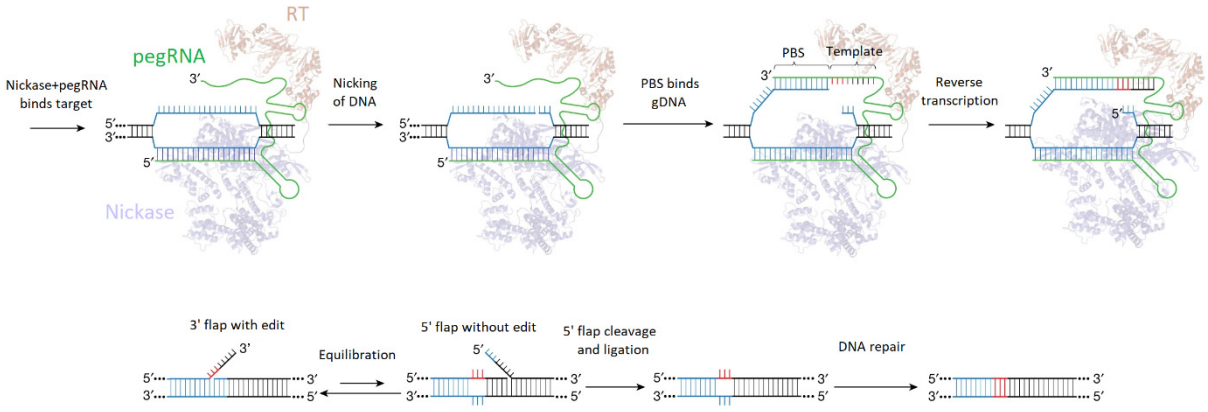


Figure 6. Schematic of DNA modification via prime editing. Top: pegRNA (green) and Nickase (purple) binds target site (blue). After the gDNA is nicked, the PBS with adjacent template containing edit (red) binds the ssDNA, and reverse transcriptase (RT, brown) extends the ssDNA to incorporate the edit. Bottom: After RT activity, DNA is in a state of equilibrium with a 3' flap containing the edit or a 5' flap without the edit. After 5' flap cleavage, DNA is repaired to incorporate the edit. Image modified from Anzalone *et al.* (2019).

CRISPR applications are not only limited to inducing DNA cleavage however. CRISPR interference (CRISPRi) is a technique that uses dCas9.[55] When dCas9 is coupled with a gRNA that targets specific promoter regions of a gene it will subsequently induce transcriptional repression by blocking the activity of RNA polymerase and other transcription factors.[55]

dCas9's intact targeting ability has also been employed to enhance gene expression. To achieve this, a virus protein-64 acidic transactivation domain (VP64) was conjugated to the C-terminus of dCas9.[56, 57] When delivered to cells in combination with gRNAs that target gene-promoter regions, they strongly induce expression of that gene.[56, 57] Further enhancements have been made to this system including conjugation to dCas9 of a tripartite activator composed of a fusion between VP64, p65, and the rightward transactivator (Rta) to increase expression by several orders of magnitude over a simple dCas9-VP64 fusion protein.[58]

6.2 Gene Drives

6.2.1 Population Level Effects and Natural Gene Drives

A gene drive is simply a way to "drive" or spread a gene through a population of wild animals at a rate greater than Mendelian inheritance would allow (Super-Mendelian inheritance).[59] This can occur even if the gene imparts a neutral or negative fitness cost for the animal.[59] As shown in Figure 7, a mouse initially heterozygous for a typical gene will only transmit that gene to 50% of progeny.[60] In contrast, a gene drive can force the inheritance of that initially heterozygous gene in up to 100% of the offspring.[59]

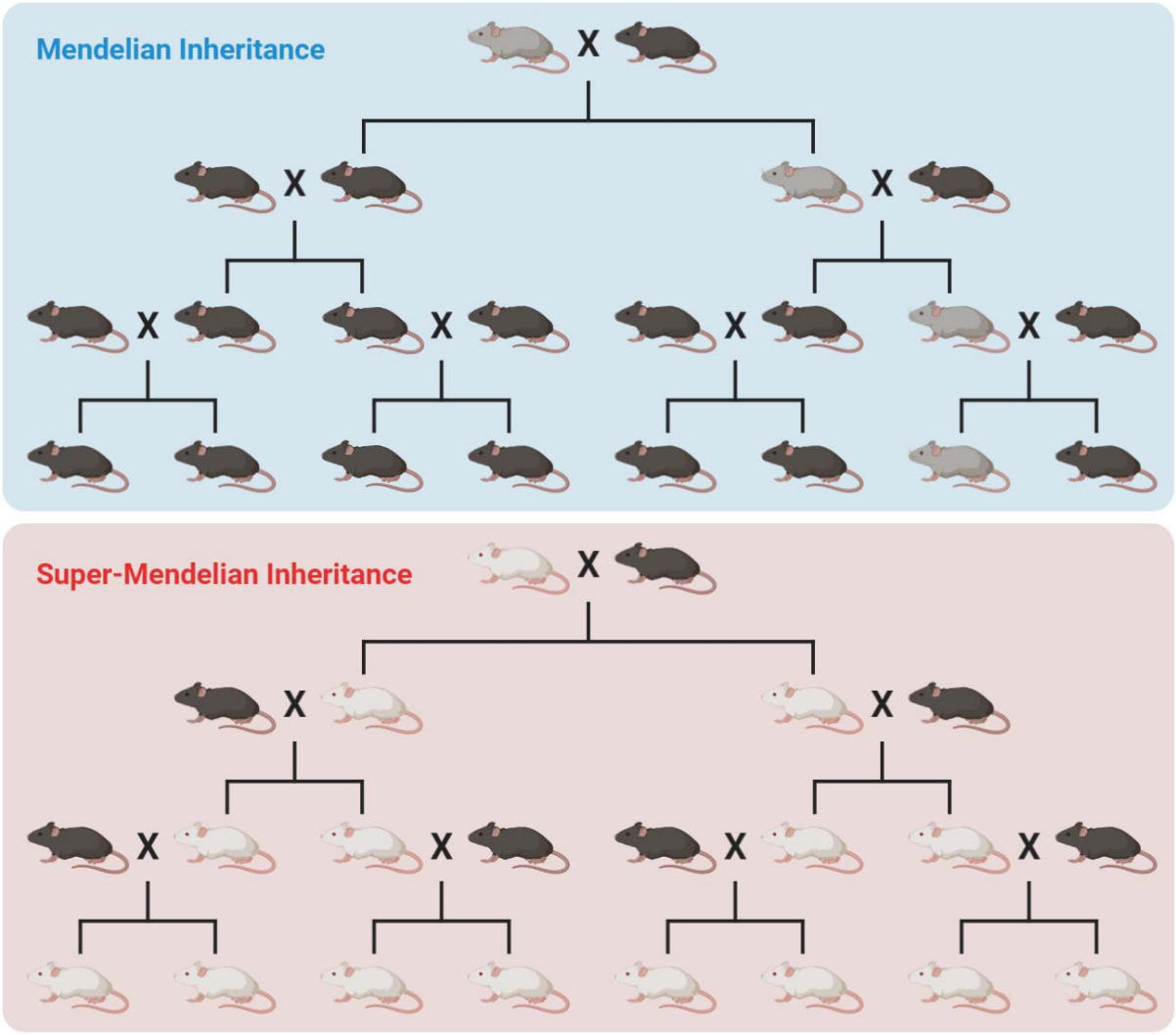


Figure 7. Schematic showing comparison of different inheritance modes. Top shows a mouse heterozygous for a simple modified gene (grey), released into the wild it will breed with wild mice (black) and following the principles of Mendelian inheritance, it will not spread through the population at any appreciable rate.[60] Bottom shows a mouse with a gene drive (white) released instead, it will rapidly spread through the population showing the properties of Super-Mendelian inheritance.[59] Image created with BioRender.[8]

Although typical genes or genetic elements cannot act as a gene drive by themselves, there are some naturally occurring gene drives that do exist. The three major types are transposable elements (TEs), homing endonuclease genes (HEGs), and meiotic drives.

TEs are mobile genetic elements capable of replicating and relocating within their host genome.[61] TEs can be broken down into Class I (RNA-mediated) or Class II (DNA-mediated).[62] Long terminal repeat (LTR) retrotransposons are typical Class I TEs, these primarily encode a polymerase-like protein that includes reverse transcriptase (RT), ribonuclease, protease, and integrase domains.[61] After transcription of the element, complementary DNA (cDNA) is generated using the RT domain which is then inserted into a target site using the integrase domain.[61] Conversely, Class II TEs bypass the RNA intermediate and either cut or copy the TE via a transposase protein.[61]

From the perspective of a molecular biologist however, TEs are not ideal tools to use. Rates of transposition greatly vary, they can have deleterious effects on the organism when a new location they're copied into disrupts a gene, and they might not adapt to having extra genetic elements

inserted into them.[63] They can also end up being silenced due to the genomic loci that they move into.[63]

HEGs are selfish genetic elements that bias their own inheritance by cutting chromosomes homologous to the ones they are present in and copying themselves over.[64] So far they have only been found in fungi, plants, algae, and bacteria so are not necessarily a good candidate for manipulation in animals.[64] They also rely on site-specific nucleases that cut homologous loci and are difficult to modify to target a custom genomic locus.[64, 65]

Meiotic drives cause a specific allele to be inherited at a greater ratio than 0.5 (Mendelian inheritance).[66] These are usually found in males and typically function by killing or inactivating gametes that don't contain the meiotic drive.[66] The prime example is the *t*-haplotype, which is found throughout *M. musculus* species[67] and has up to 99% transmission in males[68]. Functionally it is not yet fully understood, but the key cellular outcomes have been characterised.[68] During late spermatogenesis when haploid sperm are still joined by the syncytium, inhibitors of sperm motility are expressed from the *t*-haplotype locus that are freely transferred between sperm.[68] The inhibitor of this pathway is also expressed from the *t*-haplotype locus but is not transferred to other sperm, conferring a competitive advantage to the *t*-haplotype-containing sperm.[68] The *t*-haplotype spans a region of 40 Mb, which is not an ideal size for easy manipulation and modification.[67] Furthermore, there is no drive in females (i.e. it has a Mendelian inheritance ratio of 0.5).[66, 67]

6.2.2 CRISPR Gene Drives

A CRISPR gene drive is a synthetic gene drive that makes use of the CRISPR system. It can carry cargo genes, be placed in potentially any location in the genome, and is relatively easy to design and create. A schematic for a basic gene drive is shown in Figure 8. A gene drive is composed of a single cassette that is inserted into a genomic locus of interest. The cassette contains three elements: A CRISPR nuclease gene (typically SpCas9), a gRNA expression cassette, and an optional cargo element containing, for example, a gene you want to spread through the population. The gRNA in the cassette contains a guide that targets the homologous WT chromosome at the same locus as the gene drive construct.

Depending on what promoters are chosen for the nuclease there are two main types of gene drive. A zygotic-homing gene drive when expression is ubiquitous or a germline-homing gene drive when expression is limited to cells within the germline. As both the nuclease and gRNA need to be transcribed for the gene drive to be active, the gRNA can be left under the control of a ubiquitously expressed promoter such as U6.[69]

6.2.2.1 Zygotic-homing Gene Drive

The structure and functional outcomes of a hypothetical zygotic-homing gene drive using the Cas9 nuclease in mice are shown in Figure 8. When a gene drive mouse homozygous for the gene drive cassette mates with a WT mouse, the resultant zygotes would be hemizygous for the gene drive cassette. The gene drive chromosome expresses both Cas9 protein and the gRNA, which subsequently form the Cas9/gRNA RNP. In these new zygotes, this is the first time a WT chromosome is "seen" by the RNP and so they subsequently bind to their target site on the WT chromosome and cut it. Because there is now homology between both sides of the cut site and the sequence flanking the gene drive cassette, HDR could occur using the gene drive cassette as the repair template, potentially resulting in it being copied to the WT chromosome (homing).[59] If this homing occurs, the new zygote will have recapitulated the same genotype as the parental homozygous gene drive mouse.

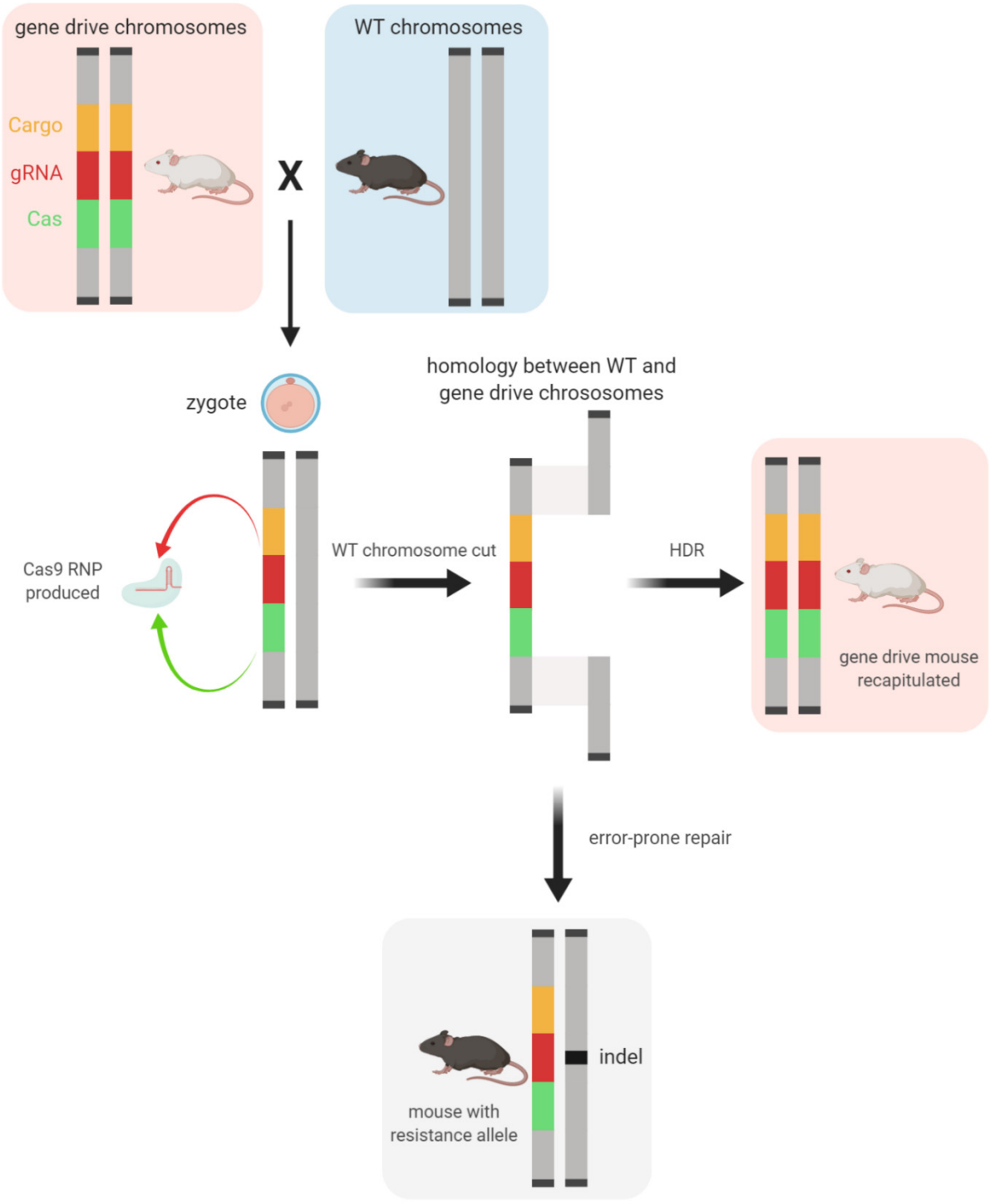


Figure 8. A zygotic-homing gene drive mouse (red box, top left), two homologous chromosomes are shown each with the Cas9 gene (green), a gRNA expression element (red), and a cargo element (yellow) making up the gene drive cassette. The zygote formed from a mating between a gene drive mouse and a WT mouse is shown and the subsequent actions and outcomes of the gene drive activity. Image created with BioRender.[8]

However, an alternative outcome to homing is the generation of an indel on the WT target chromosome via the error-prone repair pathways discussed in section 6.1.7 (e.g. NHEJ).[30, 59] This results in the generation of a mouse with only one copy of the gene drive cassette. This can be especially detrimental in a gene drive situation as any indels present in this new allele will likely no longer have a complete gRNA binding site and are thus resistant to future cutting by the same Cas9/gRNA RNP (hence the term “resistance allele”). This is not always strictly detrimental though

because if a gene drive's purpose is actually to knock out the gene at the target site, then an indel may still accomplish this.

6.2.2.2 Germline-homing Gene Drive

Genetically, a germline-homing gene drive differs from a zygotic-homing gene drive only in the promoters used to express the CRISPR nuclease and/or the gRNA, although in practice, modifying only the Cas9 promoter is simpler and no published experimental gene drives have done otherwise. The functional outcomes of an example germline-homing gene drive using Cas9 in mice are shown in Figure 9.

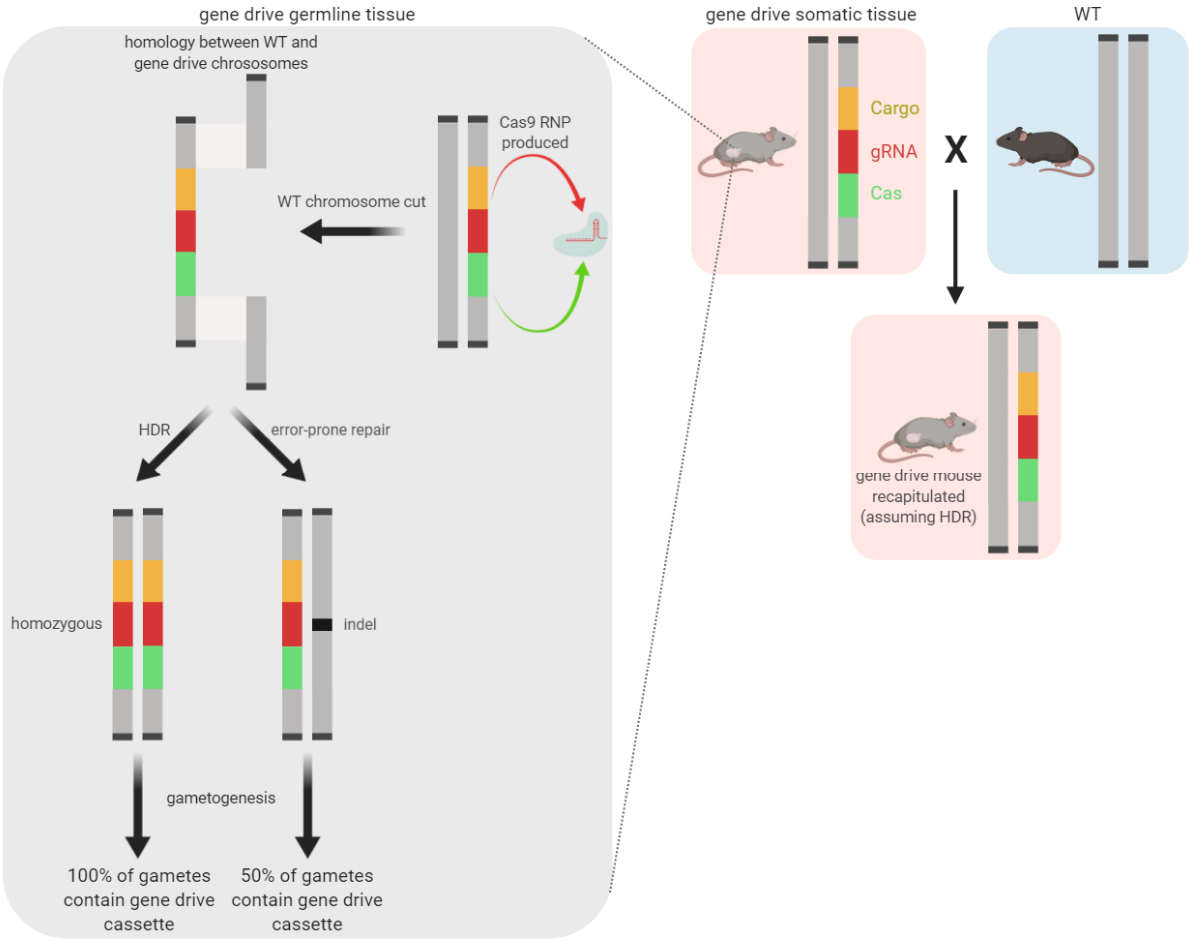


Figure 9. A germline-homing gene drive mouse (red box, top middle), two chromosomes are shown for the somatic tissue, one is WT and the other contains the Cas9 gene (green), a gRNA expression element (red), and a cargo element (yellow) making up the gene drive cassette. No gene drive activity occurs in the somatic tissue. The germline tissue (grey box) is shown and the subsequent actions and outcomes of the gene drive activity in that tissue. If homing occurs, when mated to a WT mouse (blue box) the gene drive mouse is recapitulated in the offspring (red box, middle right). Image created with BioRender.[8]

In this case, the gene drive mouse's somatic tissue remains hemizygous for the gene drive because the gene drive is only active in the germline. In early embryogenesis in that tissue, before any gametes are produced, the Cas9 and gRNA are first expressed, coming together to form the RNP complex. The RNP then binds to the target site on the WT chromosome and cuts it. Just as described above in zygotic-homing, HDR then facilitates copying of the gene drive cassette to the WT chromosome due to the homology between the two chromosomes around the gene drive and the cut site.[27, 59] Assuming successful homing, subsequent gametogenesis then leads to all gametes containing a copy of the gene drive cassette.[27, 59] Again there is the alternative error-prone pathways

which can result in indel formation instead, if this were to occur all of the time, only half of the resulting gametes would have the gene drive cassette whilst the other half would have an indel on the WT chromosome.[30, 59] In the case of a perfectly efficient gene drive, subsequent crossing of this mouse to a WT mouse means all offspring from this mating will be somatically hemizygous for the gene drive cassette, exactly recapitulating the original hemizygous parent.[59]

6.3 Gene Drive Applications

6.3.1 Spreading Transgenes vs. Knocking Out Endogenous Genes

At the molecular level, there are two key types of gene drive. As discussed above, one type carries along with it a cargo gene that will have some sort of effect on the population of wild organisms.[59] This can include beneficial genes that might do such things as confer resistance to particular diseases or it can include genes which negatively impact the target species such as removing resistance to pesticides, which will positively benefit agricultural industries or the environment.[59]

The other type of gene drive contains no cargo and is instead placed in a location in order to disrupt a gene of interest.[59] If the target site is chosen carefully, this can help reduce any deleterious effects seen by an indel instead of homing because the indel may disrupt the gene just as copying of the gene drive into that location does. A gRNA targeting an early coding region in the protein often leads to a frameshift mutation, knocking out the gene.[70] Statistically speaking however, if an indel is an insertion or deletion that is a multiple of 3, no frameshifting will occur and potentially the gene won't be knocked out.[70] Some strategies to help increase the odds of generating a knockout are choosing an important location in a gene such as an active site of an enzyme or a splice donor site[71], or use of multiple gRNAs that cut multiple important sites, or multiple sites close together to cause a large deletion between the two sites.[21] To allow homozygous lethal or otherwise deleterious gene disruptions it is necessary to use a germline homing strategy which restricts the gene disruption to the germline tissue. The drive can thus spread through a population with minimal ill effects until it saturates the population and gene drive animals breed with one another. It is only then that any new offspring are homozygous for the gene disruption in somatic tissue. One caveat for this kind of drive is that the gene targeted for disruption must not be required in the germline or the animals won't be fertile and thus won't pass on the gene.

6.3.2 Pest Control

Invasive species incur a huge monetary cost on society, including the damage they do to the environment, agriculture and the cost of controlling or removing them, all adding up to an estimated $120 billion USD per year.[72] The economic burden to the agriculture industry is also enormous, costing tens of millions of dollars to individual countries each year [73]. In Australia alone, it costs an estimated $1 billion per year through agricultural and environmental damage.[74]

Invasive rodents, including mice, pose a significant threat to biodiversity, particularly on islands and are the likely cause of hundreds of species extinctions.[75-78] Previous attempts at invasive vertebrate pest control have had some success most commonly using the methods of poisoning, trapping, and hunting.[79] Despite this, there are still many challenges including cost and ethical considerations.[79] Non-selective toxins are often used which can result in detrimental effects to non-target, native species.[79] Concerns over the undue suffering of the pest species is also an issue with regards to how they are killed, how long it takes, and what pain they are put through.[79] The cost of failed eradication attempts is especially important, as the reduction in pest species may only be temporary and lead to no further attempts at controlling them due to the perceived uselessness of previous eradication attempts.[79]

CRISPR gene drives offer an enticing alternative either for the suppression or eradication of pest species. These can function a number of ways. A “suppression drive” can potentially disrupt a gene that causes lethality or infertility when both copies are lost yet are haplosufficient and have no effect when one copy is lost.[80, 81] If developed as a germline-homing gene drive, mice that inherit the gene drive from only one parent will be unaffected as the somatic tissue remains hemizygous, and thus non-deleterious.[81] This allows spread through the population without a harmful effect before it reaches saturation (Figure 10). At this point, gene drive carrier mice will mostly be mating with other gene drive carriers resulting in the lethality or infertility of their offspring, causing a population crash.[81]

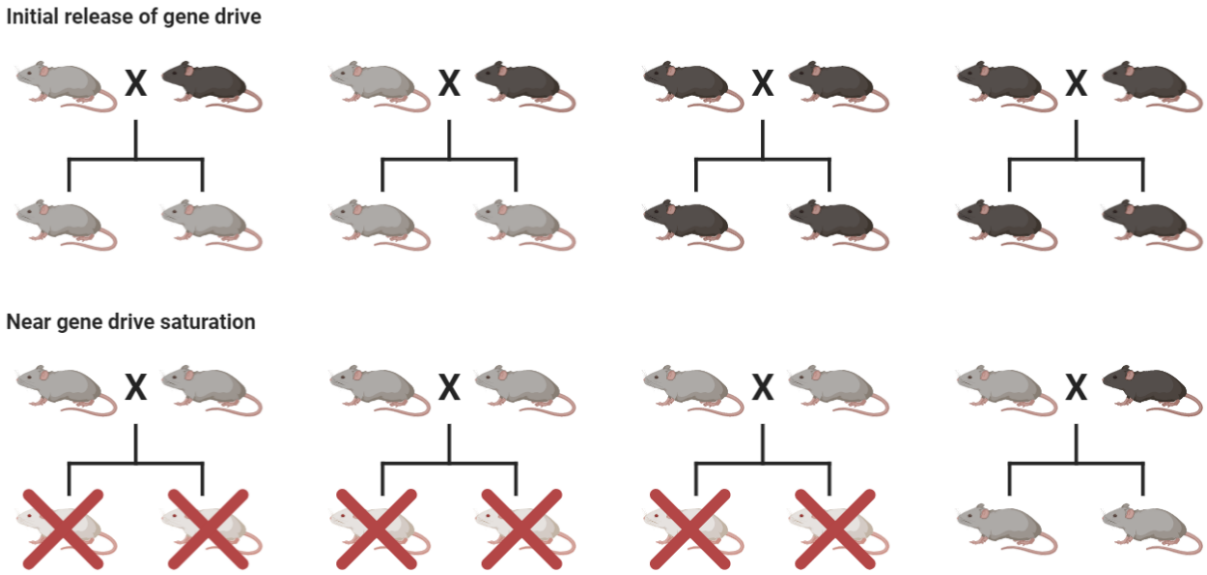


Figure 10. Suppression drive mating outcomes. Top: When a germline-homing suppression drive is initially released into the population. Drive-carrying mice (grey) are hemizygous for the drive in their somatic tissue and homozygous for the drive in their germline tissue. Matings with WT mice (black) result in more drive-carrying offspring. Matings between two WT mice continue as normal too. Bottom: When the suppression drive has spread to most of the population, most matings are between two drive-carrying mice and result in offspring that are homozygous for the suppression drive (white), resulting in lethality/infertility. Image created with BioRender.[8]

6.3.3 Other Uses

There are many other potential applications for gene drives that are not the focus of this thesis. A very promising example is eradication of diseases carried by wild animal hosts, such as malaria.[59] This could be achieved by providing immunity to the disease for the animal host, or even eradicating the animal host entirely.[59] The latter may be possible and environmentally feasible for certain species of mosquitoes, although the benefits of eradication need to be weighed against the importance of their role in the ecosystem.[59] If a species is eradicated, it could also potentially be later re-introduced once the disease is no longer present.[59] Applications also exist in the realm of agriculture in the form of a gene drive removing pesticide resistance in crop pests, or introducing new alleles making them sensitive to different pesticides.[59]

6.4 Key Experimental Gene Drives

An important concept when discussing gene drives is the efficiency at which they operate. This is simply a measure of how often they convert a WT allele into a gene drive allele. If no alleles are converted, then it has an efficiency of 0% whereas if all alleles are converted it has an efficiency of 100%. There is an important distinction here between alleles inherited and alleles converted, because even if a gene drive has an efficiency of 0%, it will still pass on the gene drive allele either

100% of the time or 50% of the time depending on whether it is homozygous or hemizygous (respectively) for the gene drive allele. In a cross between a WT animal and an animal initially hemizygous for a germline-homing gene drive, efficiency as a percentage is calculated as shown below, where n is the total number of offspring and x is the number of offspring carrying the gene drive.

$$\frac{2(x - 0.5n)}{n} \times 100\%$$

In recent years, a small number of studies have been published that experimentally tested gene drives. Below I summarise several key publications that advanced the field by implementing and investigating various features and types of gene drives.

6.4.1 "The mutagenic chain reaction: A method for converting heterozygous to homozygous mutations"

After initially being proposed by Esvelt *et al.* (2014), the first experimental CRISPR gene drive was published by Gantz and Bier (2015). A gene drive cassette was inserted into the *D. melanogaster* genome containing Cas9 driven by a germline promoter (Vasa) and a gRNA driven by a ubiquitously expressed promoter (U6) targeting the same locus on the homologous WT chromosome.

As initially calculated in this paper, when female *D. melanogaster* were crossed to WT, out of 6 females and 436 offspring, there was an allele conversion efficiency of 94.6-100%. Similarly, in the 2 males and 91 offspring resulting from a cross to WT, there was a conversion efficiency of 96.2-100% There was also an overall 4% rate of mosaicism seen in the offspring and several cases where indels were created instead of homing events. It should be noted however, that these results were later called into question as a misinterpretation of the data, this is discussed further in section 6.4.5 below where a follow-up paper analysed a similar gene drive system.

Notwithstanding the controversy over the conversion rate, this was a very impressive result, especially as the first experimental CRISPR gene drive to be constructed. The efficiencies were incredibly promising and caused a widespread interest in generating more gene drives in various species. It also increased discussion on concerns over the safety of both gene drives and simply working on gene drives from an experimental standpoint.

6.4.2 "Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*"

The Gantz and Bier paper was quickly followed up by another paper, this time in *An. stephensi* by Gantz *et al.* (2015). This paper took a more applied approach to the problem in its attempt to use cargo genes in a gene drive that have been shown to confer resistance to the malaria parasite *Plasmodium falciparum*.[84] Apart from the addition of the cargo element, the gene drive was constructed in the same way as Gantz and Bier (2015) with Vasa-Cas9 and U6-gRNA elements, although the gene drive was now located on another chromosome with a corresponding change to the gRNA guide to target the new location in the homologous WT chromosome.

Much larger numbers of animals were crossed in this paper; 3894 offspring were screened and a very impressive gene drive efficiency of 98.8% was observed. However, a critical issue was identified in the form of maternal carryover of Cas9 mRNA and gRNA. They found that in nearly every case (1781 of 1784 mosquitoes) where a female gene drive mosquito was a parent to the above progeny, the WT gene drive locus had been mutagenised by error-prone repair pathways. Contrasting this they found none of the progeny with a male gene drive parent showed this genotype. This is consistent with the concept of maternal carryover where the Cas9 mRNA and gRNA remain in the

egg and thus are still active when a zygote is formed via fertilisation, and can target the WT chromosome inherited from the male parent inducing error-prone repair. The offspring of the gene drive mosquitoes were also crossed another time where a high gene drive efficiency was again seen at 96.9%, although this time only in the offspring of the males from the previous generation due to the maternal carryover issue.

This was an important paper as it showed that a gene drive was still effective even when the entire construct including cargo was very large (~17 kb). It also highlighted an important issue with the promoters being used with regards to maternal carryover of RNA being highly problematic.

6.4.3 "Safeguarding CRISPR-Cas9 gene drives in yeast"

As a follow-up to the initial paper by Esvelt *et al.* (2014), an *S. cerevisiae* gene drive was tested by DiCarlo *et al.* (2015) out of the same lab. Key experiments within this paper showed that the drive worked in a vastly different species to the mosquitoes demonstrated by Gantz and Bier (2015) and Gantz *et al.* (2015). Several different gene drives were tested, including ones with and without different types of cargo and all of them showed homing efficiencies of ≥98%.

This paper importantly introduced the use of a safety feature for gene drives in the form of a "split drive" as discussed by Akbari *et al.* (2015), and discussed further in section 6.5.1 below. Another important concept tested out in this paper was the idea of a reversal drive, first introduced in the previously mentioned paper by Esvelt *et al.* (2014). This is another important feature concerning the safety of drives and whether it would be possible to reverse whatever changes have been caused by an already released gene drive in a wild population. It was shown to work here very effectively with an efficiency of >99%. This success is tempered somewhat considering the use of *S. cerevisiae* which is known to have relatively higher rates of HDR.[87, 88]

6.4.4 "A CRISPR-Cas9 Gene Drive System Targeting Female Reproduction in the Malaria Mosquito vector *Anopheles gambiae*"

Moving closer again to real-world applications was this paper by Hammond *et al.* (2016). *An. gambiae* was used here as it is the main vector for malaria. The goal was to construct three different gene drives that disrupt genes conferring recessive female sterility. The same U6 promoter was again used for the gRNAs, allowing the gene drives to be controlled by Cas9 expression with another, similar germline promoter *vasa2*.

Multiple generations of the gene drives were crossed which gave homing efficiencies between 69-98% and varied depending on sex and the specific gene drive used. Interestingly, among the individual mosquitoes (32) where homing did not occur, 15 showed an indel, and 17 had WT sequence, indicating the gRNA target site was not cut at all or was repaired perfectly after cleavage. This point was not discussed within the paper, however.

Unfortunately, the recessive sterility aspect of the gene drives failed. Females hemizygous for the gene drives showed a highly penetrant, dominant sterility based on number of larvae produced (0%, 4.6%, and 9.3% of WT for the three drives). The *vasa2* promoter was shown to be leaky, expressing significantly in somatic tissue and was proposed as the direct cause of this dominant sterility.

6.4.5 "Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations"

Further developments in gene drives came in this paper by Champer *et al.* (2017) which focussed on resistance alleles, one of the key stumbling blocks for an efficient gene drive. Modelling has shown that when these resistance alleles arise in a population they can rapidly confer complete resistance

to the gene drive at a population level, essentially immunising that population from any further modification by the gene drive.[81, 91]

In order to investigate these issues, both Vasa and nanos promoter-driven gene drives in *D. melanogaster* were created, the latter promoter also being germline-active, but with a lower expression level and reduced leaky somatic expression compared to Vasa.[92] The nanos drive produced a 62% homing rate, 35% rate of resistance allele formation, and 1% were left with an intact WT allele. The Vasa drive generated a lower homing rate of 52% and a higher resistance allele formation rate of 48%. Their study design allowed them to determine when these resistance alleles were forming, and they showed that they could happen in the germline before fertilisation all the way through to post-zygotic embryos where maternally deposited Cas9 was still present.

These drives were also tested across multiple lines of *D. melanogaster* where homing efficiency greatly varied. Among 7 different lines the nanos gene drive for instance had efficiencies between 40-62%. These data suggest that genetic background may be an important determinant of gene drive activity.

This paper also called in to question the original results from Gantz and Bier (2015) considering they both used the same promoter, at the same site, and in the same animal but got vastly different homing efficiencies. Their interpretation was a significant number of post-fertilisation indels were formed that disrupted the target site in a similar manner to the gene drive, but as those flies were only phenotyped and not genotyped by sequencing, the homing rate could not be accurately determined.

6.4.6 "A CRISPR–Cas9 gene drive targeting *doublesex* causes complete population suppression in caged *Anopheles gambiae* mosquitoes"

An *An. gambiae* gene drive paper by Kyrou *et al.* (2018) provided key experimental data as a demonstration of a suppression gene drive in a context which much more closely resembles how a functional gene drive release would occur.

The gene drive was positioned to block formation of the female splice variant of the *doublesex* (*dsx*) gene. When homozygous in females, they produced an infertile, intersex animal. When homozygous in males or hemizygous in either sex, the animals developed normally and were fertile. The intended outcome would be gene drive spread until the entire population is composed of only males and infertile, intersex females.

Real world conditions were emulated by use of a cage trial where mosquitoes could breed and interact freely with one another. Populations were started with a gene drive allele frequency of 12.5% (150 gene drives males, 150 WT males, and 300 WT females). Each generation, 650 randomly selected eggs were used to seed subsequent generations and all hatched larvae were phenotyped for presence of the gene drive allele.

The gene drive allele reached 100% frequency by generations 7 and 11 in two replicate cage trials, with no eggs produced in the subsequent generation. To investigate the effect that resistance alleles had, generations 2-5 were genotyped for indels at the gRNA target site. Up to 1.16% of non-gene drive alleles were found to contain indels but none were found to encode functional female *dsx* transcript.

By performing a cage trial, which had yet to be done before this paper, Kyrou *et al.* (2018) showed very strong evidence for the efficacy of a field release of a gene drive. This was also an excellent representation of the concept discussed in section 6.2.2.1 whereby the effect of resistance alleles

can be abated by relying on any indels present to be deleterious in the same manner as the presence of the gene drive allele.

6.4.7 "Super-Mendelian inheritance mediated by CRISPR–Cas9 in the female mouse germline"

Several more mosquito and fly experimental gene drive papers have been published for *An. gambiae*[94, 95], *An. stephensi*[96, 97], *Ae. aegypti*[98], and *D. melanogaster*[99-107]. However, the next big development came when a *Mus musculus* paper was publish by Grunwald *et al.* (2019). Although technically not a gene drive, it is the first (and only) paper to demonstrate the ability for homing to occur in vertebrates, or indeed any species outside of flies, mosquitoes, and the single yeast paper.

The homing was tested by use of a "split drive" approach. The homing element was a ~2 kb construct containing U6-gRNA inserted into the Tyrosinase (*Tyr*) gene. Cas9 was under the control of a ubiquitous CAG promoter which was transcriptionally inactivated via a *loxP*-Stop-*loxP* (LSL) site. Homing was assessed by crossing mice with the split drive transgenes to either a *Vasa-cre* or *Stra8-cre* mouse. Expression of Cre recombinase in the germline by either promoter excised the stop motif between the loxP sites, thus irreversibly turning on expression of Cas9 under the CAG promoter. As such, homing could be tested on a single generation of mice.

When activated by *Vasa-cre*, in females only, there was a homing efficiency of between 0-72% for individual animals (31% overall). This was from a relatively low number of potential homing events (132) from 10 different females. There was also a high rate of both indels (35%) and uncut WT alleles present (34%). No homing was seen in the males at all, with a 100% rate of error-prone repair. The *Stra8-cre* lines saw no homing in females or males. It was hypothesised that likely differential expression patterns between Cre and subsequently Cas9 were responsible for the lack of homing in males vs. females and that more precise timing of Cas9 expression would be needed.

These were important results, as although the homing efficiencies were significantly lower than those seen in insects, it showed it was at least possible to get homing in a vastly different species than shown in all previous papers. It also revealed that a similar system of Vasa-driven gene drives cassettes that were so successful in insects was likely not going to translate directly to the mouse.

6.4.8 "A male-biased sex-distorter gene drive for the human malaria vector *Anopheles gambiae*"

This subsequent paper by Simoni *et al.* (2020) was a follow up to the Kyrou *et al.* (2018) paper discussed above in section 6.4.6, further refining a suppression drive in *An. gambiae*.

The drive was located in the same position in the *dsx* gene and added a cargo gene. That gene, I-PpoI is a site-specific nuclease which cleaves repeat sequences in the X-chromosome and was controlled with a promoter expressed during spermatozoa development (a *beta2-tubulin* variant). This led to the selective cleavage of X-bearing sperm and thus causing a sex ratio bias towards male offspring. Thus, this drive used both the position in *dsx* and its cargo gene as separate means of biasing the population towards males. This essentially made the gene drive a more robust system, where loss of the I-PpoI nuclease meant that a functional gene drive targeting female sterility in the *dsx* gene continued to function, and loss of the Cas9 or gRNA would still produce non-functional *dsx* alleles contributing to the same sterility phenotype.

This paper was also significant in that it used the *zero population growth* (*zpg*) promoter instead of *vasa* which successfully dealt with both the leaky expression and the maternal carryover previously seen when under the control of *vasa*.

This was again tested in cage trials where all aspects of the drive were successful: A male bias of 93.1% of offspring was seen, with homing efficiencies of 92% for males and 99.8% for females. Two

replicate cage trials starting at 2.5% and 25% gene drive allele frequencies were all successful at eliminating the population (within 9 or 13 generations and 5 or 6 generations respectively).

6.5 Gene Drive Safety

Because of the enormous potential of CRISPR gene drives to rapidly alter wild populations, consideration of the potential ecological and environmental risks they pose has been intensive. There have been numerous papers discussing many different aspects of their safety, including safety during the experimental phase and safety concerns regarding potential release of gene drives into the wild. Nearly all experimental gene drive papers discussed above also address the safety issues and it was first brought up and discussed extensively in the original paper proposing CRISPR gene drives by Esvelt *et al.* (2014).

6.5.1 Experimental Safety

Safety when conducting experiments was of prime concern during the work in this thesis. Strategies were employed to confine the experimental gene drives so that they either cannot escape the laboratory, or if they do, their spread and effect on wild populations would be as minimal as possible. There are a number of key confinement strategies that have been laid out and are generally agreed upon in multiple different papers, initially outlined by Akbari *et al.* (2015). These include molecular, ecological, reproductive, and barrier confinement strategies.[86]

The two key molecular strategies are a “split drive” and a synthetic target.[86] These are both demonstrated in Figure 11. A split drive functions by separating the Cas9 nuclease and the gRNA on to separate chromosomes and allowing only one of these elements to be the homing cassette. This means that if a split drive were ever to be accidentally released into the wild, the Cas9 and gRNA can segregate away from each other as shown in Figure 12. This severely limits any continued homing except when the Cas9 and gRNA are inherited together by chance.

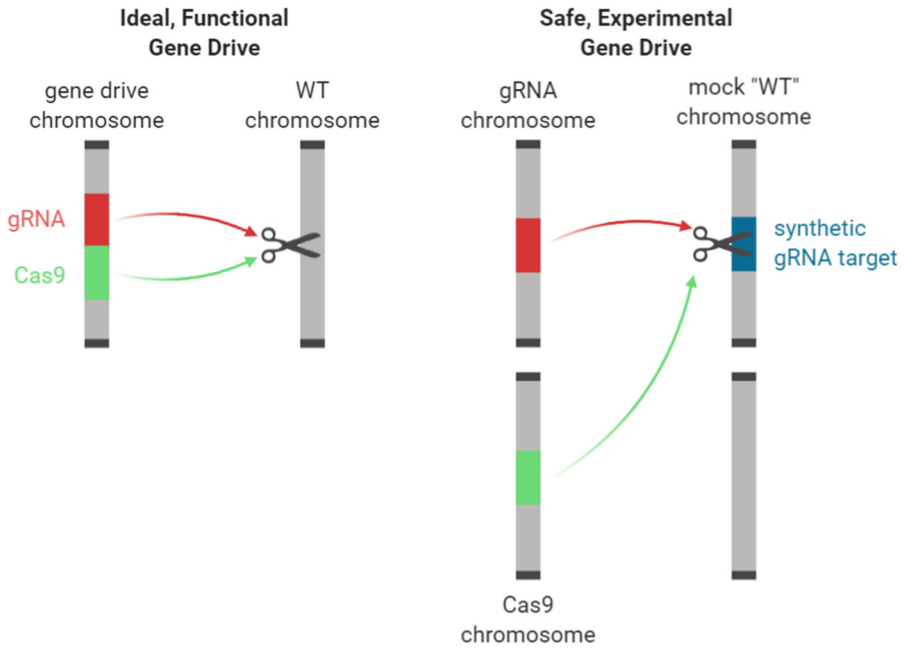


Figure 11. Gene drive experimental safety. Left shows an ideal, functional gene drive that would be a likely design in a released drive. The Cas9 gene and gRNA are situated next to each other and they will target the homologous locus in a WT chromosome. Right shows a safer alternative which is more suited for experimentally testing in a lab. The mock “WT” chromosome contains a synthetic gRNA target not present in WT chromosomes. The Cas9 gene and the gRNA are also not on the same chromosome to allow segregation and halting of gene drive activity when bred with WT mice. Image created with BioRender.com.[8]

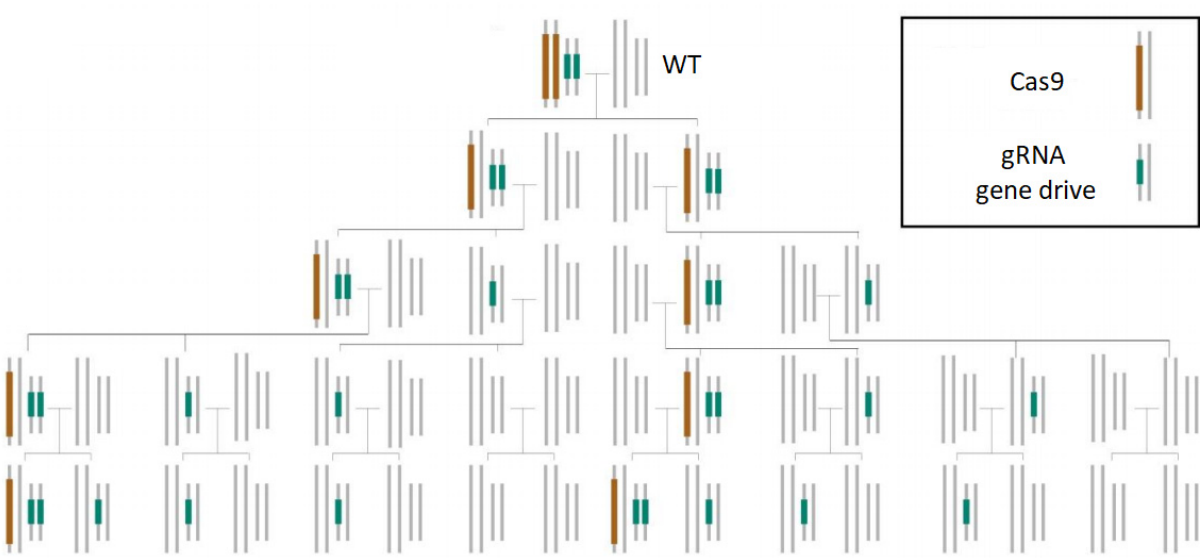


Figure 12. Inheritance of a split drive system. Pattern of inheritance shown if a split gene drive with 100% homing efficiency were to be released into the wild. The gRNA gene drive is the homing construct that contains only a gRNA. The Cas9 is separated onto a different chromosome. Figured modified from Akbari *et al.* (2015).

A synthetic target simply means that an extra animal model needs to be made with a synthetic sequence inserted into the chromosomal region that is homologous to the gene drive homing cassette. The gRNA in the homing cassette is designed to cleave this synthetic target sequence. Thus, the gene drive is incapable of cutting WT chromosomes, so if it were accidentally released into the wild, it could not function at all. The drawback to this system is the extra animal model that needs to be made and bred at relatively high numbers which then essentially functions as the experimental "WT" animal that the gene drive animal is crossed to.

Ecological confinement strategies are basically directives to perform experiments outside of the habitable range of the wild organism so any escaped gene drive animals have no wild mates to breed with.[86] This is of course, not always easy or even possible depending on what organism is being studied, especially given that pest species are often widespread.

Reproductive confinement means that the laboratory strain is incapable of breeding with the wild strains.[86] This is often not possible, but one example is using strains of *D. melanogaster* which have compound autosomes which can breed with one another but are incapable of producing offspring with wild *D. melanogaster* without those same compound autosomes.[86]

The final, and likely most important confinement strategy is the physical barriers.[86] This is simply the barriers and procedures that are in place to effectively prevent any gene drive animals from escaping into the wild. All the aforementioned confinement strategies are essentially in place as a backup in case the physical barrier strategies fail. Physical barriers include features such as the number of doors that are between the organisms and the outside world, whether the organisms should be handled when they are not anesthetised, and the use of low-temperature rooms or air-blast fans to prevent any flying organisms from escaping. Physical containment strategies will differ considerably between different organisms, especially when considering small, flying insects in contrast to comparably much larger animals like mice or rats which can be more easily tracked and accounted for.

6.5.2 Release Safety

Release safety has also been extensively discussed. This covers concepts such as trying to limit the spread of a gene drive and attempting to reverse or prevent changes induced by a gene drive.

A “reversal drive”, as the name implies, is a secondary gene drive that attempts to reverse the changes induced by a previously released gene drive.[59] It functions in the exact same way as a normal gene drive, except it is targeting the genomic locus of the previously released gene drive to change it back into essentially a WT animal.[59] The only caveat being that the Cas9 and gRNA components can never be removed by the gene drive because they are still required to spread the reversal drive.[59] As mentioned in section 6.4.3, a reversal drive has been constructed and exhibited similar homing rates as standard gene drives.[85]

An “immunisation drive” is still a conceptual idea but is a potentially valid strategy that can be deployed in response to another gene drive that has been released (either deliberately or accidentally). An immunisation drive is again another form of gene drive, but it would function by targeting the same target site of another gene drive with the intention of changing that target sequence to prevent the other gene drive from spreading.[59]

Trying to limit the spread of gene drives is a key concern among researchers, and there are no easy solutions for when a successful gene drive is released. It is likely to spread rapidly, crossing national borders, and potentially affecting very closely related species or different sub-populations that are unintended targets. A couple of different solutions do present themselves though - in certain circumstances a sub-population may be entirely isolated from others by means of being geographically isolated on an island, indeed invasive rodents on islands are an excellent target for an eradication drive.[59] A genetic solution to this problem is also apparent in the form of polymorphisms that are fixed in the target population so long as the same polymorphism is not fixed in non-target populations.[59] However, this requires a great deal of sequencing data about the genomes of the different populations.[59]

6.6 CRISPR Gene Drive Modelling

Experimentally, in small populations of tightly controlled insects, gene drives are looking promising. This is not necessarily going to translate well into large populations where many new and challenging variables can come into play. To this end, it is essential to employ computational modelling of these systems to get some idea of how they will behave.

Many factors influence how a model is constructed and simulated. Primary considerations include the specific species and how they mate and breed. Depending on the model, many other aspects can be taken into consideration for the simulation too, such as social behaviour, how spread out the mice are, how likely they are to mingle with one another, etc. Often modeling is panmitic such that mate selection is random and does not depend on the animals’ likelihood to be near one another, whereas spatial modelling can account for this but is more computationally intensive. For modelling gene drives, some key pieces of data specific to them are the rate of homing, rate of resistance allele generation, and rate of DNA cleavage. It’s also important to consider the type of gene drive - different types of suppression drives will affect the population in different manners to one another and to a simple gene drive carrying a neutral cargo gene.

A key outcome from the modelling has shown that a CRISPR gene drive will indeed spread rapidly through a large population[91, 110] similar to the small, cage trials.[93, 109] Modelling has also shown

resistance allele formation will severely limit their spread.[91, 110, 111] Even naturally occurring resistance alleles might already be present in wild populations due to natural genetic variation.[91, 110]

When considering the spread of a cargo element as opposed to any sort of gene knockdown or population suppression gene drive, it has become clear that unless the error-prone repair pathways are suppressed, then resistance to the drive will inevitably evolve.[90]

When looking at modelling CRISPR gene drives in vertebrate pests for population suppression on islands, Prowse *et al.* (2017) published a paper where we showed that simple gene drives containing only a single gRNA were incapable of eradicating the population. More complex modelling was done however, whereby multiple gRNAs are present in a gene drive that target separate, but nearby locations as a potential method to remove resistance alleles due to subsequent cleavage events giving the gene another chance to "overwrite" the resistance allele with the gene drive.[81] Depending on the type of gene drive, we showed that multiple gRNAs effectively eradicated the island populations of invasive vertebrates.[81]

A simpler approach to modelling this in Marshall *et al.* (2017) and Noble *et al.* (2017) showed that multiple gRNAs were sufficient to overcome the problem of resistance. This does not adequately account for the molecular mechanisms that are likely to take place with multiple gRNAs however, as multiple nearby cut sites generally result in a large deletion between the two sites which would be a single resistance allele in this case that cannot be recovered from by the gene drive. In Prowse *et al.* (2017), we modelled with the assumption that the multiple gRNAs were expressed at different time periods however, preventing the occurrence of large deletions. It should be noted, that if the purpose of the gene drive is to knock out a gene, as discussed in section 6.3.1 and demonstrated in the mosquito cage trials,[93] a deletion of a critical gene may also be beneficial to the drive and prevent the resistance allele from being passed on.[81]

7 AIMS

The broad aim of this thesis was to research and develop advances in CRISPR technology. This can be broken down into a major and minor aim.

The major aim was to design, implement, and test CRISPR-based gene drives in *M. musculus*. This was done separately with Cas9 and Cas12a nucleases. Both versions of these were also implemented as zygotic and germline-homing gene drives controlled using appropriate promoters for the nucleases. After validation of the individual components of the gene drives, specific outcomes to be measured for each gene drive were the homing efficiency, number of resistance alleles generated, and number of uncut WT alleles. All gene drive research was to be carried out with the key safety features of a split drive system and a synthetic target as discussed in section 6.5.1.

The minor aim was to develop a suite of CRISPR plasmids allowing for the streamlining and simplification of genome editing where two CRISPR target sites were required instead of only one.

8 CAS9 DUAL GRNA PAPER

8.1 INTRODUCTION

As discussed in section 6.1.5, one of the commonly used CRISPR genome editing tools is an all-in-one plasmid that contains both a CRISPR nuclease and the associated gRNA under appropriate promoters that allow the expression of the nuclease and gRNA. This greatly simplifies the use of CRISPR as the plasmid can be introduced into cultured cells, tissue, or zygotes to induce editing without a requirement to generate Cas9 protein or gRNA beforehand.

What was not previously mentioned is that often it is beneficial to simultaneously induce two cuts, using two different gRNAs. This can be used to create large deletions between the two sites, to knockout multiple genes simultaneously, to induce chromosomal translocations, to insert multiple sequences into the genome at once, or, in the case of Cas9 Nickase it can be used to induce two nearby nicks to induce a single DSB whilst greatly reducing off-target cutting.[113, 114]

Previous methods to accomplish double-gRNA activity either involved the use of multiple plasmids or had polymerase chain reaction (PCR)-dependent or complicated and time-consuming cloning strategies to create the required plasmid expressing multiple gRNAs.[115-118] As such, the publication included in this thesis chapter provides a significant advance in CRISPR tool development, and describes a suite of plasmids that allow a PCR-free, one-step generation of Cas9/Nickase genome editing plasmids containing two distinct gRNAs.

8.2 "VERSATILE SINGLE-STEP-ASSEMBLY CRISPR/CAS9 VECTORS FOR DUAL GRNA EXPRESSION" PAPER

The 11 pages following the statement of authorship below contain the full, published paper. The following 4 pages after that contain the supporting information published with the paper.

Statement of Authorship

Title of Paper	Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Adikusuma F*, Pfitzner C*, & Thomas PQ (2017). Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. PloS one,12(12), e0187236 *Co-first authors

Co-Principal Author

Name of Principal Author (Candidate)	Chandran Pfitzner		
Contribution to the Paper	Performed significant amount of investigation, data analysis, and data interpretation. Handled significant amount of data curation. Contributed significantly to manuscript and creation of visualisations.		
Overall percentage (%)	50%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	14/9/2020

Co-Principal Author

Name of Principal Author	Fatwa Adikusuma		
Contribution to the Paper	Performed majority of conceptualisation. Designed the methodology. Performed significant amount of investigation, data analysis, and data interpretation. Handled significant amount of data curation. Wrote first draft and contributed significantly to manuscript and creation of visualisations.		
Signature		Date	14/9/2020

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Paul Q Thomas		
Contribution to the Paper	Contributed to conceptualisation. Acquired funding, performed project administration and supervision. Contributed significantly to manuscript.		
Signature		Date	16/9/2020

Please cut and paste additional co-author panels here as required.

PLOS ONE

RESEARCH ARTICLE

Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression

Fatwa Adikusuma[1,2☯], **Chandran Pfitzner**[1,3☯], **Paul Quinton Thomas**[1,3,4*]

1 School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, Australia, **2** Centre for Biomedical Research, Faculty of Medicine, Diponegoro University, Semarang, Central Java, Indonesia, **3** Robinson Research Institute, The University of Adelaide, Adelaide, South Australia, Australia, **4** South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia

☯ These authors contributed equally to this work.
* paul.thomas@adelaide.edu.au

Citation: Adikusuma F, Pfitzner C, Thomas PQ (2017) Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. PLoS ONE 12(12): e0187236. https://doi.org/10.1371/journal.pone.0187236

Editor: Tony T. Wang, SRI International, UNITED STATES

Received: August 31, 2017

Accepted: October 10, 2017

Published: December 6, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

CRISPR/Cas9 technology enables efficient, rapid and cost-effective targeted genomic modification in a wide variety of cellular contexts including cultured cells. Some applications such as generation of double knock-outs, large deletions and paired-nickase cleavage require simultaneous expression of two gRNAs. Although single plasmids that enable multiplex expression of gRNAs have been developed, these require multiple rounds of cloning and/or PCR for generation of the desired construct. Here, we describe a series of vectors that enable generation of customized dual-gRNA expression constructs via an easy one-step golden gate cloning reaction using two annealed oligonucleotide inserts with different overhangs. Through nucleofection of mouse embryonic stem cells, we demonstrate highly efficient cleavage of the target loci using the dual-guide plasmids, which are available as Cas9-nuclease or Cas9-nickase expression constructs, with or without selection markers. These vectors are a valuable addition to the CRISPR/Cas9 toolbox and will be made available to all interested researchers via the Addgene plasmid repository.

Introduction

CRISPR/Cas9 technology is a powerful genome editing tool that has become widely used by researchers to generate targeted genetic modifications in many contexts including cultured cell lines and zygotes. CRISPR/Cas9 offers several advantages over preexisting genome editing technologies including ease of use, relatively low cost and high activity [1–5]. The CRIPSR/Cas9 platform comprises two components; Cas9, which functions as a programmable endonuclease that generates a blunt-ended double-stranded break (DSB) and a ~100 nt guide RNA (gRNA), in which the ~20 nt at the 5' end directs Cas9 to the target site via RNA:DNA complementary base pairing [6–8]. Generation of a targeted DSB can be achieved by delivery of Cas9 and gRNA components in plasmid, RNA or ribonucleoprotein (RNP) forms. For some applications, such as cultured cells, plasmids are generally preferred due to their ease of generation and stability. Commonly used plasmids for expression of Cas9 or Cas9-nickase (D10A) and single gRNA are available from the Zhang laboratory and can be obtained through the

Addgene plasmid repository. These plasmids contain both gRNA and Cas9 expression cassettes in a single plasmid with optional selection markers such as puromycin or GFP to facilitate screening. Importantly, generation of a unique customized gRNA of interest can be performed easily as the gRNA cloning site contains BbsI restriction sites, allowing a one-step golden gate cloning approach for insertion of a pair of annealed oligonucleotides containing the specific ~20 bp guide sequence [6, 9].

To simultaneously target a pair of genomic regions, expression of two gRNAs is required. While this can be achieved by co-transfection of two plasmids, this process can be inefficient. To achieve efficient dual cuts, all CRISPR/Cas9 components with dual-gRNAs should be expressed from a single plasmid. Single plasmids expressing multiple gRNAs have been developed, however generation of the desired constructs using those available plasmids require multiple cloning and/or PCR steps. Here we modify commonly-used vectors from the Zhang laboratory so that each plasmid can express two gRNAs and can be generated via a simple one-step cloning method. We show that these plasmids, termed dual-gRNA plasmids, provide an efficient tool for experiments requiring simultaneous expression of two gRNAs such as multiplexed knock-out of two genes, generation of large deletions and generation of indels using Cas9-nickase. These vectors are a valuable addition to the CRISPR/Cas toolbox and will be made available through the Addgene plasmid repository.

Results

Generation of vectors

To generate plasmids that permit simultaneous expression of two gRNAs, we inserted an additional hU6-gRNA expression cassette into the available CRISPR plasmids from the Zhang laboratory. The second cassette was positioned in the opposite orientation to the original hU6-gRNA expression cassette to reduce the possibility of recombination (Fig 1A). The additional cassette also contains a BbsI golden gate site at the guide insertion site as per the original cassette. However, unlike the original BbsI site which generates GTTT and GGTG overhangs, the new site generates CGGT and TTTA overhangs (Fig 1B) allowing simultaneous targeted insertion of two annealed oligonucleotides with different complementary overhangs in a one-step digestion-ligation reaction (Fig 1C; see below). We added the extra gRNA cassette to the following Cas9 nuclease vectors: pX330 (no selection marker), pX458 (GFP selection marker) and pX459.V2.0 (puromycin selection marker), and to the following Cas9-nickase vectors: pX335 (no selection marker), pX461 (GFP selection marker) and pX462.V2.0 (puromycin selection marker). Those vectors were named pDG330, pDG458, pDG459, pDG335, pDG461 and pDG462, respectively.

Efficient generation of custom dual-gRNA vector using a one-step cloning protocol

Having generated the dual-gRNA vectors, we next tested whether we could simultaneously insert two annealed oligonucleotide duplexes in a one-step cloning process. We designed two gRNA oligonucleotide inserts targeting the mouse *Sox1* and *Sox3* genes. These inserts carried BspMI and SacI restriction sites at the original and second hU6-gRNA sites, respectively. Annealed oligonucleotide duplex pairs and pDG459 vector were subjected to a one-step digestion-ligation cycling protocol followed by bacterial transformation (Fig 1C). All 12 colonies analyzed contained vectors with correct assembly based on their RFLP pattern (Fig 1D). Similar results were obtained with other dual-gRNA plasmids (pDG330, pDG335, pDG461 and pDG462) with correct assembly in 21/23 colonies based on RFLP, confirmed by sequencing in

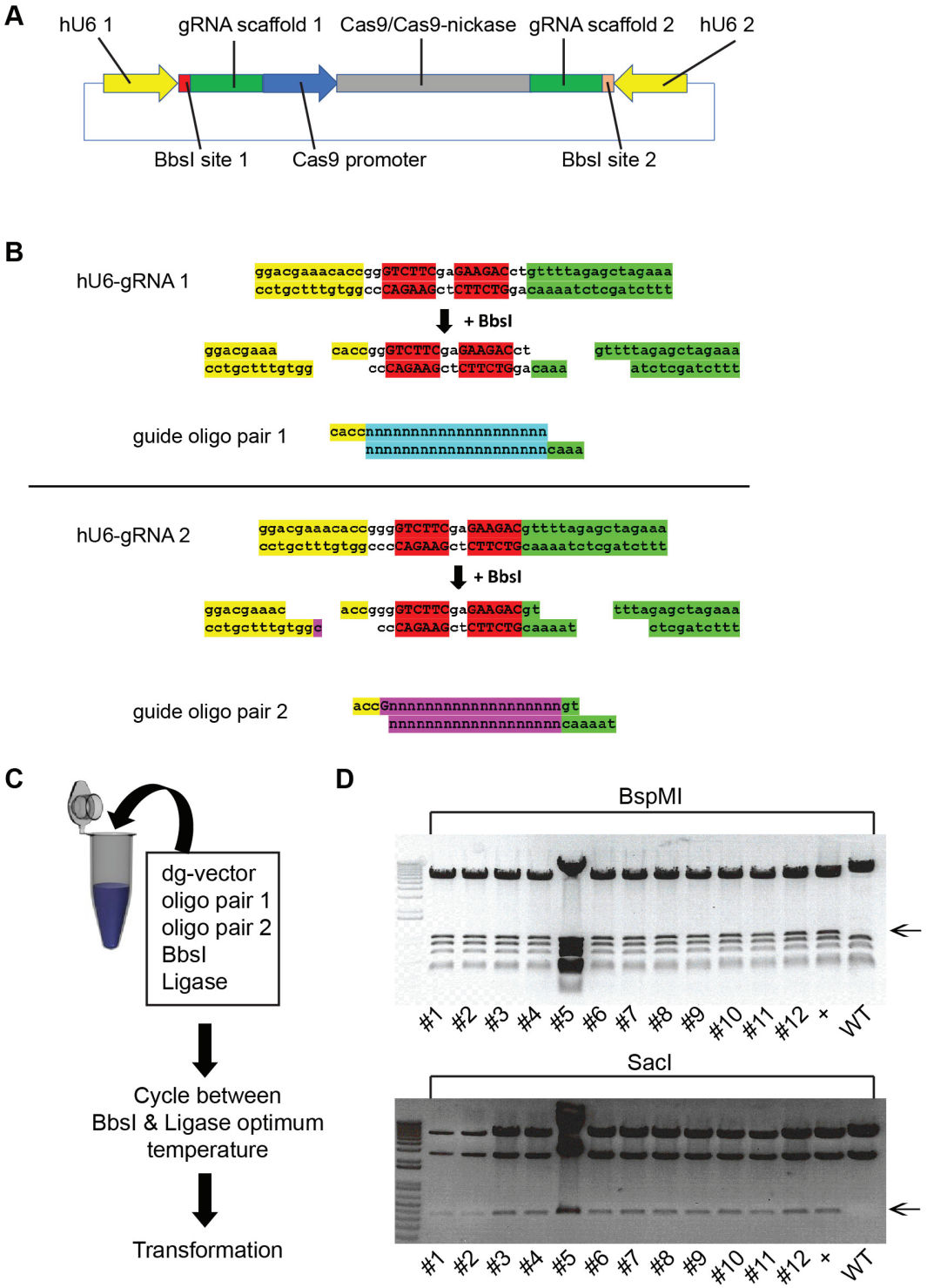


Fig 1. Generation of dual-gRNA expressing vectors. (A) Schematic of dual-gRNA vectors. (B) Golden gate cloning strategy for insertion of specific guide sequences into each cassette. Note that the BbsI sites generate different overhangs after restriction digest. Red highlights indicate the BbsI sites, yellow and green highlights are part of hU6 promoter and gRNA, respectively, that are necessarily present in the plasmid. Blue and purple highlights indicate the unique customized guide sequences (C) One-step cloning protocol for the generation of customized dual-gRNA vectors. (D) Insertion of Sox1A and Sox3A oligonucleotide duplexes into pDG459 resulted in correct insertions in all 12

colonies as indicated by BspMI and SacI restriction digest. The black arrow indicates the diagnostic band for correct insertion.

https://doi.org/10.1371/journal.pone.0187236.g001

9 samples (data not shown). This demonstrates that our dual-gRNA vector design combined with the one-step cloning protocol can allow easy and efficient generation of CRISPR/Cas9 vectors with dual-gRNA expression cassettes.

Efficient generation of DSB at two sites using vectors expressing Cas9 nuclease and dual-gRNAs

We next tested whether the dual-gRNA Cas9-nuclease vectors could efficiently induce indels or deletions through simultaneous digestion at two target sites. Four different pDG459 derivatives were initially generated; the first targeted *Sox1* site A and *Sox3* site A (pDG459 Sox1A/Sox3A), the second targeted *Sox1* site B and *Sox3* site B (pDG459 Sox1B/Sox3B), the third targeted *Sox1* site A and *Sox1* site B (pDG459 Sox1A/Sox1B) which are separated by 51 bp and the last targeted *Sox3* site A and *Sox3* site B (pDG459 Sox3A/Sox3B) which are separated by 47 bp (Fig 2A). All target sequences contained restriction sites and hence indel generation at each site could be assayed by RFLP analyses. In addition, efficient digestion by pDG459 Sox1A/Sox1B or pDG459 Sox3A/Sox3B gRNAs should cause a deletion of ~50 bp which can be readily detected by PCR. Each of the four constructs were separately transfected to the mouse ES cells followed by puromycin selection to ensure only transfectants were harvested. *Sox1* and *Sox3* PCRs were performed on Sox1A/Sox3A-treated samples followed by a BfuAI (isoschizomer of BspMI) and SacI RFLP assay to assess indel generation at Sox1A and Sox3A sites, respectively. Both RFLP analyses indicated that pDG459 Sox1A/Sox3A plasmid induced mutations with ~100% efficiency at both Sox1A and Sox3A sites (Fig 2B and S1 Fig). Highly efficient mutagenesis of the Sox1B and Sox3B sites was also detected by ApaI and SfoI RFLP assays in pDG459 Sox1B/Sox3B-trasfected cells (Fig 2B and S1 Fig). We next examined whether deletion of the sequences between the cut sites could be induced by pDG459 Sox1A/Sox1B or Sox3A/Sox3B transfection. PCR products corresponding to deletion alleles were readily generated in pDG459 Sox1A/Sox1B- or Sox3A/Sox3B-treated samples but not in the WT and the unpaired controls upon *Sox1* or *Sox3* PCR (Fig 2B, S1 Fig). Efficient dual nuclease activity was also demonstrated using pDG330- and pDG458-derived constructs (S2 Fig). Together, these data indicate that all-in-one dual-gRNA Cas9 nuclease vectors can facilitate efficient simultaneous cutting at two gRNA target sites.

Efficient DSBs induced by plasmids expressing Cas9-nickase and dual paired-gRNAs

Expression of Cas9-nickase with a single gRNA results in a ssDNA break that is typically repaired without causing a mutation. In contrast, expression of Cas9-nickase and two gRNAs targeting closely spaced sites on opposite DNA strands will generate a staggered DSB, repair of which results in indel mutations [10, 11]. We next tested the dual-gRNA Cas9-nickase vectors to assess whether they could efficiently induce DSBs via expression of gRNA pairs. We generated pDG462 derivatives targeting Sox1A/Sox1B and Sox3A/Sox3B which have the requisite orientation and spacing to permit mutagenesis by paired-nickase activity (Fig 2A). As negative controls, we also generated pDG462 targeting Sox1A/Sox3A and Sox1B/Sox3B which are not paired therefore should not generate indel mutations. Vectors were transfected to mouse ES cells followed by puromycin selection. T7E1 heteroduplex assays revealed that pDG462 Sox1A/Sox1B and Sox3A/Sox3B efficiently generated mutations at *Sox1* and *Sox3*,

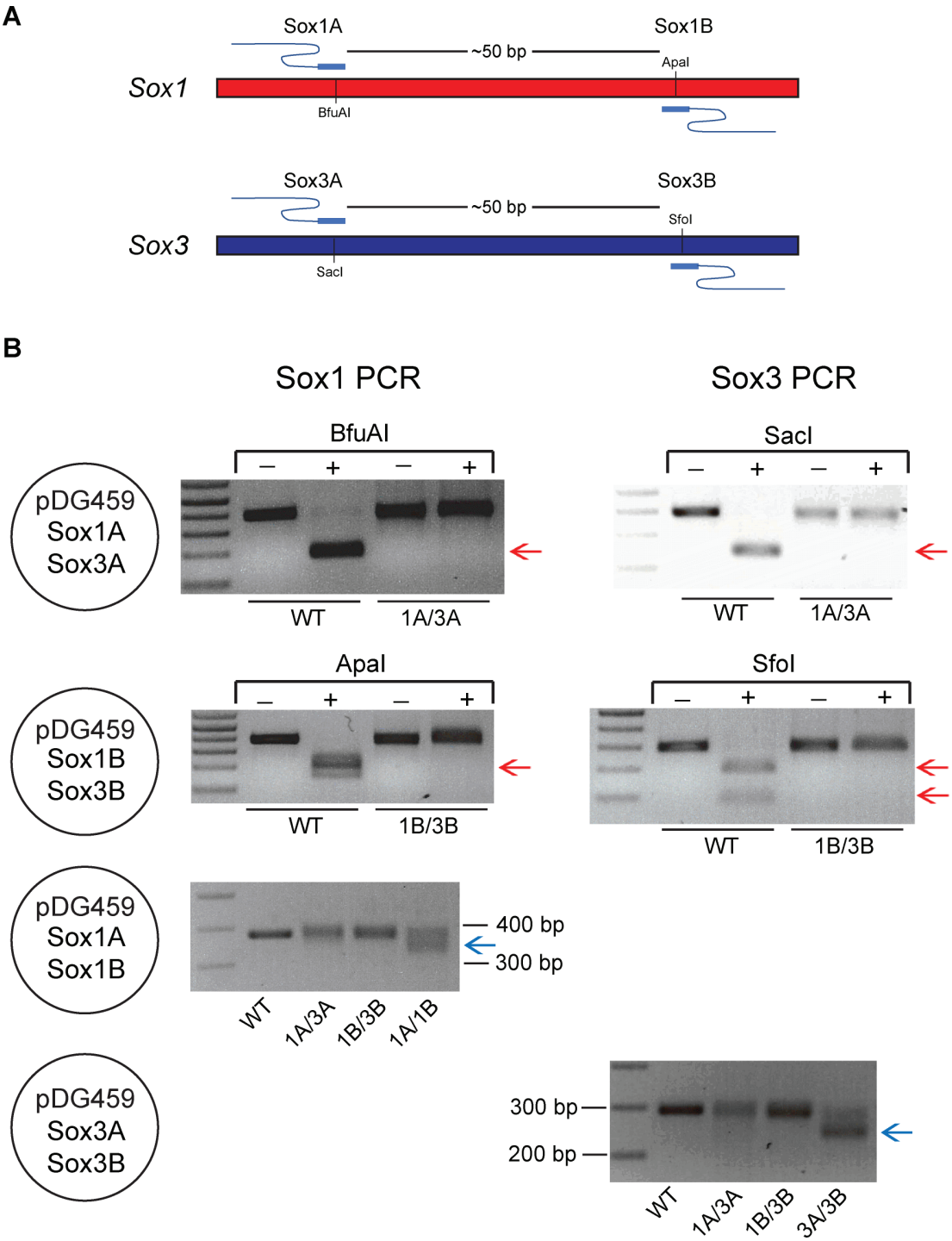


Fig 2. Efficient dual cutting mediated by pDG459 vector. (A) Schematic of gRNA target sites in the *Sox1* and *Sox3* genes. (B) Highly efficient dual cuts induced by vectors derived from pDG459 as indicated by PCR and RFLP analyses. WT products were cut by restriction enzymes resulting in bands indicated by the red arrows. Absence of these bands in dual-gRNA vector-treated samples indicated that the Cas9 nuclease and the gRNAs efficiently induced mutations thus destroying the restrictions sites. Efficient cuts from pDG459 Sox1A/Sox1B and pDG459 Sox3A/Sox3B were indicated by deletion of ~50 bp regions between cuts (blue arrows). Complete figures with more independent samples can be found in S1 Fig.

https://doi.org/10.1371/journal.pone.0187236.g002

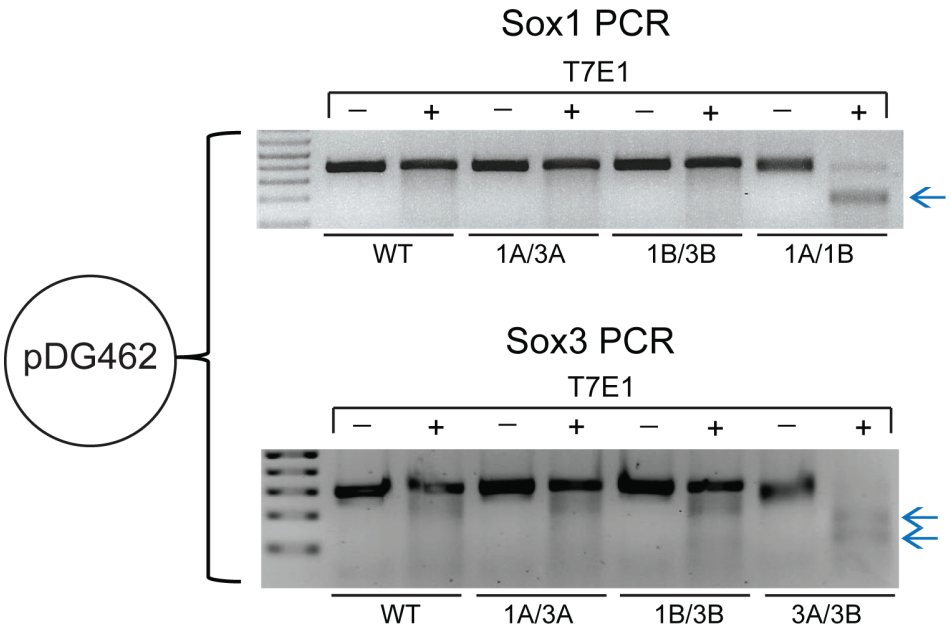


Fig 3. Paired-nickase DSB induction by pDG462. *Sox1* or *Sox3* PCR followed by T7E1 assay was performed on pDG462-transfected samples. Mutations in *Sox1* and *Sox3* were induced by pDG462 Sox1A/Sox1B or pDG462 Sox3A/Sox3B, respectively, as indicated by the digested products after T7E1 treatment (blue arrows). Mutations were not induced by non-paired-nickase control plasmids (pDG462 Sox1A/Sox3A or pDG462 Sox1B/Sox3B). Complete figures with more independent samples can be found in S3 Fig.

https://doi.org/10.1371/journal.pone.0187236.g003

respectively (Fig 3 and S3 Fig). In contrast, there was no evidence of mutations after transfection of the non-paired control plasmids (Fig 3 and S3 Fig). Efficient mutation of *Sox3* was also achieved using dual-gRNA nickase vectors pDG335 and pDG461 expressing Sox3A/Sox3B (S4 Fig). Together, these data demonstrate efficient targeted mutagenesis using dual-gRNA paired-nickase vectors.

Discussion

Plasmids from the Zhang laboratory have greatly simplified generation of customized gRNA-Cas9/Cas9-nickase expression constructs through utilization of the golden gate cloning strategy. Users only need to anneal a pair of oligonucleotides and ligate them into the vectors via a one-step cloning process, circumventing multiple rounds of PCR and cloning [6, 9]. We modified available plasmids to allow simultaneous insertion of two oligonucleotide duplex inserts using the simple one-step cloning method. These modified vectors provide a user-friendly and cost-effective system to perform experiments that require simultaneous expression of two gRNAs. Additionally, we have shown that both gRNA cassettes are active and induce mutations with high efficiency at both target sites when combined with reliable transfection and selection methods.

Other recent studies have also generated CRISPR/Cas9 vectors that are able to express dual-gRNAs simultaneously, most of which also take advantage of golden gate cloning. However, unlike the dual-gRNA vectors described herein, these require multiple rounds of cloning and/or PCR [12–15]. Additionally, the strategy to express dual-gRNA as a polycistronic transcript that is split by Csy4 RNA polymerase [16] has been shown to have low efficiency [17]. Furthermore, our dual-gRNA vectors are available with Cas9 nuclease or nickase, and with or without selection markers, and can therefore be utilized in a broad range of experimental contexts.

Vectors from other studies, although more complicated, are useful when conducting experiments requiring more than 2 gRNAs since those vectors can bear up to 7 gRNAs in a single vector [13, 14].

Our one-step cloning strategy could be applied to generate multiple gRNAs by adding more hU6-gRNA cassettes. To do so, the BbsI sites of the new cassettes would need to be modified to produce different unique overhangs upon digestion. This cloning approach could also be combined with other commonly used CRISPR platform variants such as Cpf1, dCas9-Fok1, Cas9-HF, eSpCas9, and other Cas9 orthologs or mutants that recognize different PAM sequences.

Off-target mutagenesis is one of the most significant issues of CRISPR/Cas9 genome editing [18, 19], particularly for therapeutic applications. The paired-nickase strategy has previously been shown to minimize the off-target effects that are a feature of Cas9 nuclease [20, 21]. We therefore anticipate that the dual-gRNA nickase vectors will be an attractive option for users who require efficient mutagenesis and with maximum specificity.

Efficient dual nuclease cuts are useful for generating targeted large deletions for many purposes such as studying the function of enhancers or long non-coding RNA. In some situations, targeted large deletions are required to delete an exon such as for DMD therapeutics via exon skipping [22–24] or to delete a centromere for chromosome removal [25]. Dual-gRNA Cas9 vectors could also be used for simultaneous KO of two different genes. We also offer our dual-gRNA nuclease vectors for efficient generation of chromosome translocations to model diseases such as Burkitt's lymphoma or acute myeloid leukemia [12]. Dual DSBs may also aid insertion of flanking loxP sequences for conditional deletion and for insertion of gene swap constructs [26, 27]. Furthermore, these vectors can also be used for injection into mouse zygotes for the generation of mutant mice [28]. Taken together our vectors are a valuable addition to the CRISPR/Cas9 toolbox and should be useful for many CRISPR/Cas9-based applications.

Materials and methods

Plasmid and gRNA design

Plasmids pX330, pX335, pX458, pX459.V2.0, pX461 and pX462.V2.0 were gifts from Feng Zhang (Addgene plasmid 42230, 42335, 48138, 62988, 48140 and 62987, respectively) [6, 9]. The Cas9 or Cas9-nickase of those plasmids are derived from *Streptococcus pyogenes* Cas9 which recognizes NGG PAM sequences. The BbsI sequences from pX330 were replaced with the second version of BbsI sequences (see Fig 1B). The hU6-gRNA region was then amplified using primers containing NotI sites. PCR products were then ligated to original plasmids at the NotI site. Guide sequences targeting Sox1A, Sox1B, Sox3A and Sox3B were `5′-GCCGCC GGGCGAGTGCAGGT-3′`, `5′-GCCCACGAACCTCTCGGGCC-3′`, `5′-GCTGACCCACATCTG AGCTC-3′` and `5′-GACCGCAGTCCCGGCGCCC-3′`, respectively, which were designed using online CRISPR design tool http://crispr.mit.edu/. The modified plasmids have been submitted to Addgene with plasmid reference number #100898–100993.

One step cloning for the generation of customized dual-gRNA plasmid

Forward and reverse oligonucleotides containing the guide sequences for Sox1A, Sox1B, Sox3A and Sox3B with appropriate overhangs (Table 1) were phosphorylated and annealed by mixing 100 pmol of each pair and 0.5 μL T4 PNK (NEB) then incubated at 37˚C for 30 minutes, 95˚C for 5 minutes and slowly ramped to RT. Annealed oligonucleotides were diluted 1 in 125. Pairs of oligonucleotide duplexes were ligated into the empty vectors in a one-step digestion ligation reaction by mixing the diluted duplex oligonucleotide pairs (1 μL each) with

Table 1. List of oligos used to generate the dual-gRNA targeting plasmids.

Target	Oligo pair 1 (5'-3')	Oligo pair 2 (5'-3')
Sox1A/Sox3A	CACC GCCGCCGGGCGAGTGCAGGT	ACC GCTGACCCACATCTGAGCTC GT
	AAACACCTGCACTCGCCCGGCGGC	TAAAACGAGCTCAGATGTGGGTCAG
Sox1B/Sox3B	CACC GCCCACGAACCTCTCGGGCC	ACC GACCGCAGTCCCGGCGCCC GT
	AAACGGCCCGAGAGGTTCGTGGGC	TAAAACGGGCGCCGGGACTGCGGT
Sox1A/Sox1B	CACC GCCGCCGGGCGAGTGCAGGT	ACC GCCCACGAACCTCTCGGGCC GT
	AAACACCTGCACTCGCCCGGCGGC	TAAAACGGCCCGAGAGGTTCGTGGG
Sox3A/Sox3B	CACC GCTGACCCACATCTGAGCTC	ACC GACCGCAGTCCCGGCGCCC GT
	AAACGAGCTCAGATGTGGGTCAGC	TAAAACGGGCGCCGGGACTGCGGT

Grey highlights indicate the sequence of the guides

https://doi.org/10.1371/journal.pone.0187236.t001

100 ng empty vector, 100 μmol of DTT, 10 μmol of ATP, 1 μL of BbsI (NEB), 0.5 μL of T4 ligase (NEB) and NEB-2 buffer in 20 μL of reaction. The mixture was placed in a thermocycler and cycled 6 times at 37˚C for 5 minutes and 16˚C for 5 minutes before bacterial transformation. Plasmids were prepared using miniprep kit (Qiagen) or PureLink® HiPure Plasmid Midiprep Kit (Life Technologies). Correct insertion of oligonucleotide duplexes into the vectors was confirmed by Sanger sequencing using the following primers: GGTTTCGCCACCT CTGACTTG (first insert) and TGCATCGCATTGTCTGAGTAGG (second insert). It is recommended to digest the vectors using BbsI before sequencing as correct insertion should remove the BbsI sites.

Cell culture and transfection

R1 mouse embryonic stem cells from Andras Nagy's laboratory (Established from a male blastocyst hybrid of two 129 substrains (129X1/SvJ and 129S1/SV-+^{P}+$^{Tyr\text{-}c}$ Kitl $^{Sl\text{-}J}$/+)) were used for all experiments. Cells were cultured in 15% FCS/DMEM supplemented with LIF, 3 μM CHIR99021 (Sigma), 1 μM PD0325901 (Sigma), 2 mM Glutamax (Gibco), 100 μM non-essential amino acids (Gibco) and 100 μM 2-mercaptoethanol (Sigma). One million ES cells were nucleofected with 3 μg of plasmid DNA using the Neon™ Transfection System 100 μL Kit (Life technologies) at 1400 V, 10 ms and 3 pulses according to the manufacturer's protocol. For transfection of pDG459 and pDG462, puromycin selection (2 μg/mL was initiated 24 hours post transfection for 48 hours. GFP FACS was performed on cells transfected with pDG458 and pDG461 48 hours post transfection. Surviving cells were cultured for 4–7 days without selection before harvesting. Cells transfected with plasmid pDG330 and pDG335 did not undergo any selection.

DNA extraction, PCR, RFLP and T7E1 assay

Genomic DNA was extracted from 1–2 million cells using High Pure PCR Template Preparation Kit (Roche) according to the manufacturer's instructions. *Sox1* PCR was performed using primers F: 5′-CCCTTCTCTCCGCTAGGC-3′ and R: 5′-GTTGTGCATCTTGGGGTTTT-3′. *Sox3* PCR used primers F: 5′-CAGCATGTACCTGCCACCT-3′ and R: 5′-ACAAAACCCCG ACAGTTACG-3′. RFLP or T7E1 assay was performed by mixing 5 μL of PCR products (without purification) with the restriction enzymes or T7E1 enzyme (NEB) in a total volume of 20 μL and incubated for 1 hour at the suggested optimal temperatures. Prior to T7E1 assay, PCR products were slowly re-annealed to form heteroduplex products by heating the PCR products at 95˚C for 5 minutes and slowly ramped down to room temperature.

Supporting information

S1 Fig. Efficient dual cutting mediated by pDG459 vector. Extended figures of Fig 2B with more independent samples. (A) BfuAI and SacI RFLP analyses indicated efficient dual cuts from pDG459 Sox1A/Sox3A. (B) ApaI and SfoI RFLP analyses indicated efficient dual cuts from pDG459 Sox1B/Sox3B. WT products after digestions (red arrows) were absent in pDG459-treated samples. (C) Large deletions were induced in the *Sox1* region in pDG459 Sox1A/Sox1B-treated samples. (D) Large deletions were induced in the *Sox3* region in pDG459 Sox3A/Sox3B-treated samples. Large deletion fragments are indicated with blue arrows. Each sample came from independent transfection ($n \geq 3$).
(TIF)

S2 Fig. Mutation inductions mediated by vectors pDG330 and pDG458. (A) Transfection of pDG330 Sox1A/Sox3A into mouse ES cells induced mutations at both targets which were indicated by smaller fragments after T7E1 assay (blue arrows). (B) BfuAI and SacI RFLP were used to assess the mutation induction in Sox1A and Sox3A sites, respectively, after treatment of pDG458 Sox1A/Sox3A followed by GFP FACS enrichment. Presence of WT products produced smaller bands after restriction digestions (red arrows) which were absent in pDG458 Sox1A/Sox3A-treated samples. Each sample came from independent transfection.
(TIF)

S3 Fig. Paired-nickase DSB induction by pDG462. Extended figures of Fig 3 with more independent samples. Smaller bands produced after T7E1 digestion (blue arrows) indicated presence of mutation in samples treated with paired-nickase pDG462 Sox1A/Sox1B (A) or Sox3A/Sox3B (B). Each sample came from independent transfections.
(TIF)

S4 Fig. Paired-nickase-mediated mutation inductions by pDG335 and pDG461 vectors. T7E1 assay showed that expression of paired-nickase gRNAs Sox3A/Sox3B from pDG335 (A) or pDG461 (B) induced mutations in the *Sox3* locus as indicated by the presence of cut products (blue arrows). Each sample came from independent transfections.
(TIF)

Acknowledgments

The plasmids used in this study were provided by the Feng Zhang laboratory through Addgene. R1 ES cell line was a kind gift from Andras Nagy. F.A is supported by Beasiswa Unggulan DIKTI (Directorate General of Higher Education, Indonesian Government)

Author Contributions

Conceptualization: Fatwa Adikusuma, Paul Quinton Thomas.

Data curation: Fatwa Adikusuma, Chandran Pfitzner.

Formal analysis: Fatwa Adikusuma, Chandran Pfitzner.

Funding acquisition: Paul Quinton Thomas.

Investigation: Fatwa Adikusuma, Chandran Pfitzner.

Methodology: Fatwa Adikusuma.

Supervision: Paul Quinton Thomas.

Validation: Fatwa Adikusuma, Chandran Pfitzner.

Visualization: Fatwa Adikusuma, Chandran Pfitzner.

Writing – original draft: Fatwa Adikusuma, Paul Quinton Thomas.

Writing – review & editing: Fatwa Adikusuma, Chandran Pfitzner, Paul Quinton Thomas.

References

1. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014; 157(6):1262–78. https://doi.org/10.1016/j.cell.2014.05.010 PMID: 24906146; PubMed Central PMCID: PMC4343198.
2. Komor AC, Badran AH, Liu DR. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. Cell. 2017; 168(1–2):20–36. https://doi.org/10.1016/j.cell.2016.10.044 PMID: 27866654; PubMed Central PMCID: PMC5235943.
3. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014; 32(4):347–55. https://doi.org/10.1038/nbt.2842 PMID: 24584096; PubMed Central PMCID: PMC4022601.
4. Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. Nat Methods. 2013; 10 (10):957–63. https://doi.org/10.1038/nmeth.2649 PMID: 24076990; PubMed Central PMCID: PMC4051438.
5. Barrangou R, Doudna JA. Applications of CRISPR technologies in research and beyond. Nat Biotechnol. 2016; 34(9):933–41. https://doi.org/10.1038/nbt.3659 PMID: 27606440.
6. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013; 339(6121):819–23. https://doi.org/10.1126/science.1231143 PMID: 23287718; PubMed Central PMCID: PMC3795411.
7. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337(6096):816–21. https://doi.org/10.1126/science.1225829 PMID: 22745249.
8. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013; 339(6121):823–6. https://doi.org/10.1126/science.1232033 PMID: 23287722; PubMed Central PMCID: PMC3712628.
9. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013; 8(11):2281–308. https://doi.org/10.1038/nprot.2013.143 PMID: 24157548; PubMed Central PMCID: PMC3969860.
10. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013; 154(6):1380–9. https://doi.org/10.1016/j.cell.2013.08.021 PMID: 23992846; PubMed Central PMCID: PMC3856256.
11. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. 2013; 31(9):833–8. https://doi.org/10.1038/nbt.2675 PMID: 23907171; PubMed Central PMCID: PMC3818127.
12. Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han YC, et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature. 2014; 516 (7531):423–7. https://doi.org/10.1038/nature13902 PMID: 25337876; PubMed Central PMCID: PMC4270925.
13. Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci Rep. 2014; 4:5400. https://doi.org/10.1038/srep05400 PMID: 24954249; PubMed Central PMCID: PMC4066266.
14. Kabadi AM, Ousterout DG, Hilton IB, Gersbach CA. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. Nucleic Acids Res. 2014; 42(19):e147. https://doi.org/10.1093/nar/gku749 PMID: 25122746; PubMed Central PMCID: PMC4231726.
15. Vidigal JA, Ventura A. Rapid and efficient one-step generation of paired gRNA CRISPR-Cas9 libraries. Nat Commun. 2015; 6:8083. https://doi.org/10.1038/ncomms9083 PMID: 26278926; PubMed Central PMCID: PMC4544769.
16. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol. 2014; 32(6):569–76. https://doi.org/10.1038/nbt.2908 PMID: 24770325; PubMed Central PMCID: PMC4090141.
17. Han K, Jeng EE, Hess GT, Morgens DW, Li A, Bassik MC. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nat Biotechnol. 2017; 35(5):463–74. https://doi.org/10.1038/nbt.3834 PMID: 28319085.

18. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013; 31(9):827–32. https://doi.org/10.1038/nbt.2647 PMID: 23873081; PubMed Central PMCID: PMC3969858.

19. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol. 2013; 31(9):822–6. https://doi.org/10.1038/nbt.2623 PMID: 23792628; PubMed Central PMCID: PMC3773023.

20. Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. Nat Biotechnol. 2015; 33(2):179–86. https://doi.org/10.1038/nbt.3101 PMID: 25503383; PubMed Central PMCID: PMC4320661.

21. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. 2014; 24(1):132–41. https://doi.org/10.1101/gr.162339.113 PMID: 24253446; PubMed Central PMCID: PMC3875854.

22. Tabebordbar M, Zhu K, Cheng JK, Chew WL, Widrick JJ, Yan WX, et al. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science. 2016; 351(6271):407–11. https://doi.org/10.1126/science.aad5177 PMID: 26721686; PubMed Central PMCID: PMC4924477.

23. Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Castellanos Rivera RM, et al. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science. 2016; 351(6271):403–7. https://doi.org/10.1126/science.aad5143 PMID: 26721684; PubMed Central PMCID: PMC4883596.

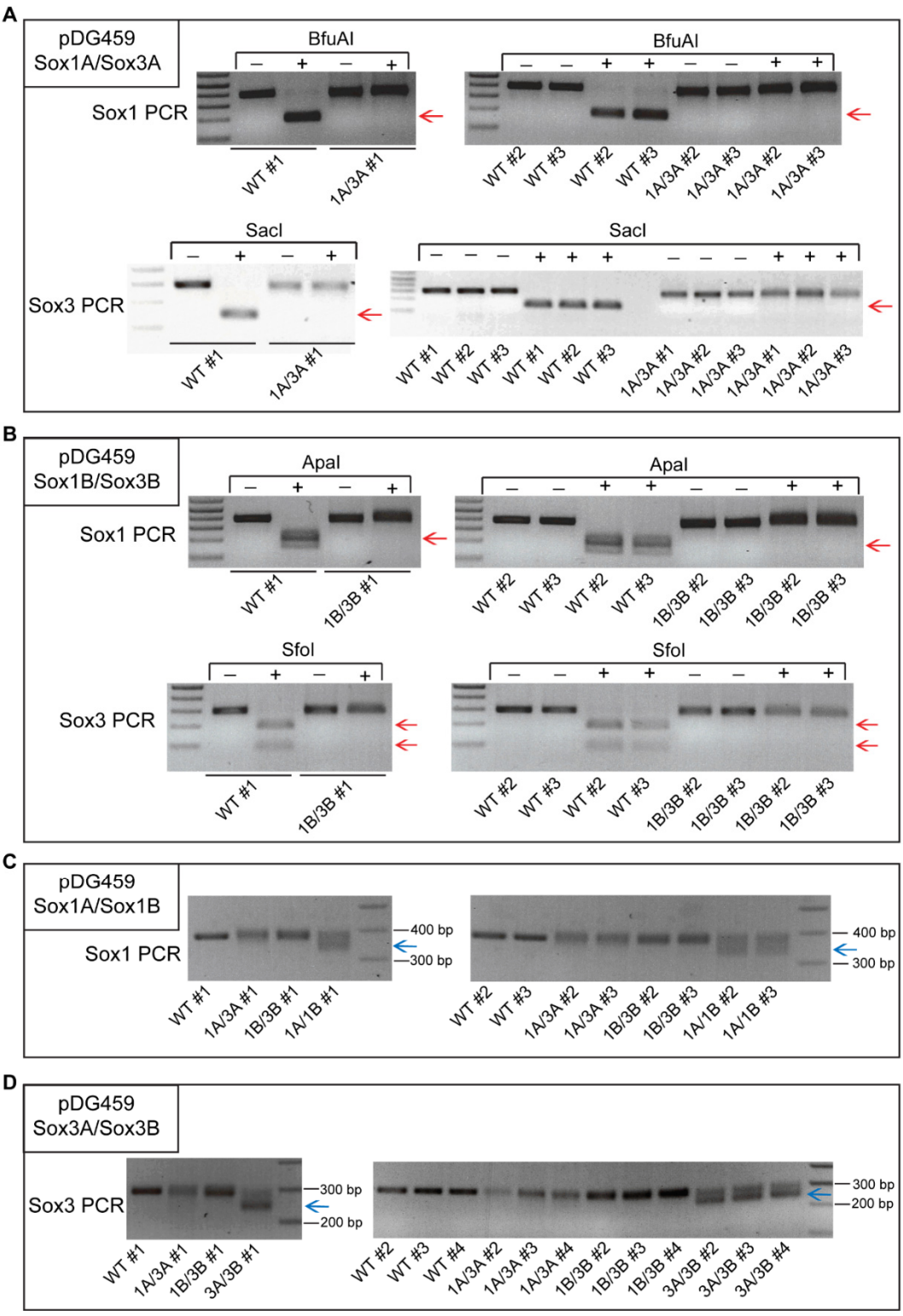
24. Long C, Amoasii L, Mireault AA, McAnally JR, Li H, Sanchez-Ortiz E, et al. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science. 2016; 351(6271):400–3. https://doi.org/10.1126/science.aad5725 PMID: 26721683; PubMed Central PMCID: PMC4760628.

25. Adikusuma F, Williams N, Grutzner F, Hughes J, Thomas P. Targeted Deletion of an Entire Chromosome Using CRISPR/Cas9. Mol Ther. 2017. https://doi.org/10.1016/j.ymthe.2017.05.021 PMID: 28633863.

26. Adikusuma F, Pederick D, McAninch D, Hughes J, Thomas P. Functional Equivalence of the SOX2 and SOX3 Transcription Factors in the Developing Mouse Brain and Testes. Genetics. 2017; 206(3):1495–503. https://doi.org/10.1534/genetics.117.202549 PMID: 28515211; PubMed Central PMCID: PMC5500146.

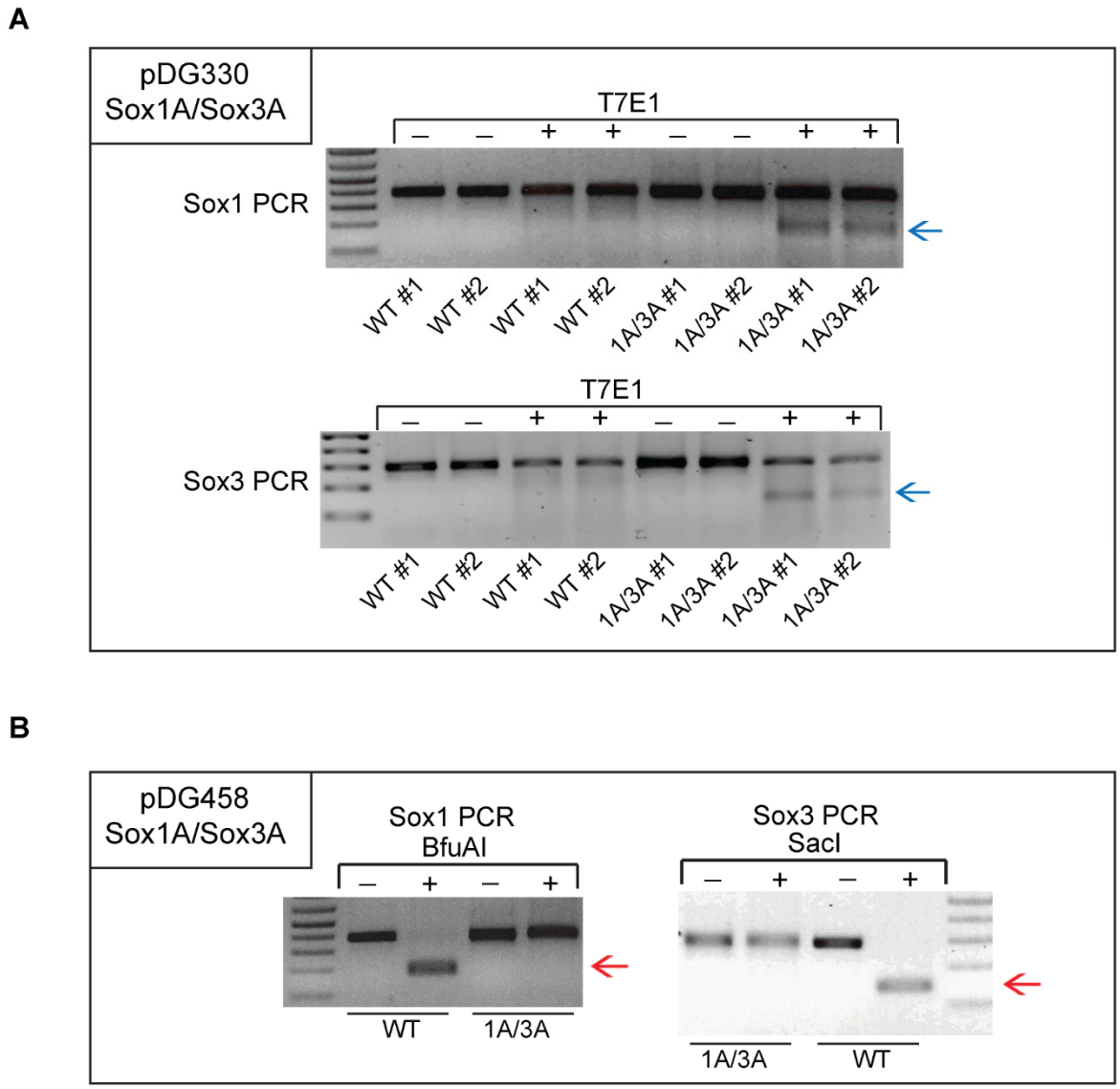
27. Quadros RM, Miura H, Harms DW, Akatsuka H, Sato T, Aida T, et al. Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. Genome Biol. 2017; 18(1):92. https://doi.org/10.1186/s13059-017-1220-4 PMID: 28511701; PubMed Central PMCID: PMC5434640.

28. Mashiko D, Fujihara Y, Satouh Y, Miyata H, Isotani A, Ikawa M. Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. Sci Rep. 2013; 3:3355. https://doi.org/10.1038/srep03355 PMID: 24284873; PubMed Central PMCID: PMC3842082.

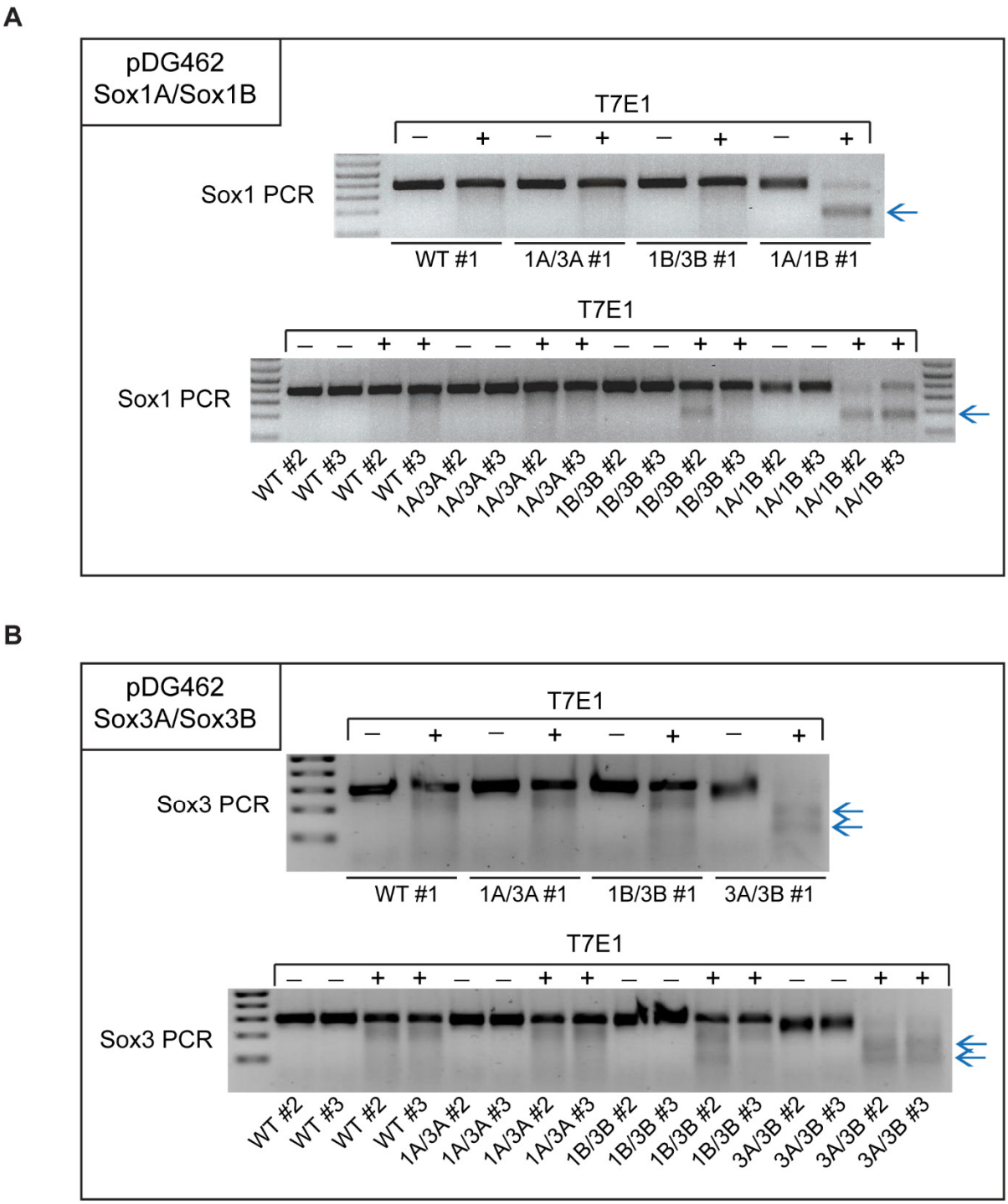


S1 Fig. Efficient dual cutting mediated by pDG459 vector.

Extended figures of Fig 2B with more independent samples. (A) BfuAI and SacI RFLP analyses indicated efficient dual cuts from pDG459 Sox1A/Sox3A. (B) ApaI and SfoI RFLP analyses indicated efficient dual cuts from pDG459 Sox1B/Sox3B. WT products after digestions (red arrows) were absent in pDG459-treated samples. (C) Large deletions were induced in the *Sox1* region in pDG459 Sox1A/Sox1B-treated samples. (D) Large deletions were induced in the *Sox3* region in pDG459 Sox3A/Sox3B-treated samples. Large deletion fragments are indicated with blue arrows. Each sample came from independent transfection (n ≥ 3).

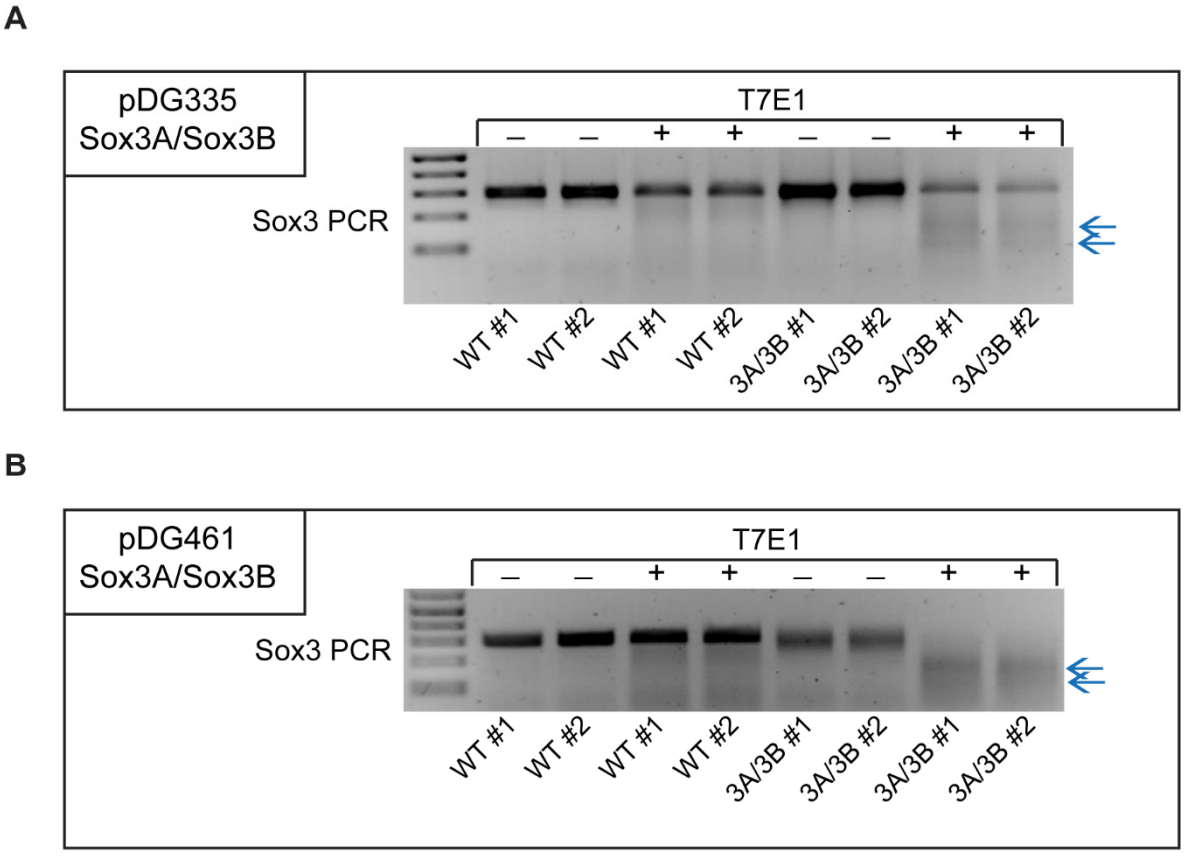


S2 Fig. Mutation inductions mediated by vectors pDG330 and pDG458.
(A) Transfection of pDG330 Sox1A/Sox3A into mouse ES cells induced mutations at both targets which were indicated by smaller fragments after T7E1 assay (blue arrows). (B) BfuAI and SacI RFLP were used to assess the mutation induction in Sox1A and Sox3A sites, respectively, after treatment of pDG458 Sox1A/Sox3A followed by GFP FACS enrichment. Presence of WT products produced smaller bands after restriction digestions (red arrows) which were absent in pDG458 Sox1A/Sox3A-treated samples. Each sample came from independent transfection.



S3 Fig. Paired-nickase DSB induction by pDG462.

Extended figures of Fig 3 with more independent samples. Smaller bands produced after T7E1 digestion (blue arrows) indicated presence of mutation in samples treated with paired-nickase pDG462 Sox1A/Sox1B (A) or Sox3A/Sox3B (B). Each sample came from independent transfections.



S4 Fig. Paired-nickase-mediated mutation inductions by pDG335 and pDG461 vectors.
T7E1 assay showed that expression of paired-nickase gRNAs Sox3A/Sox3B from pDG335 (A) or pDG461 (B) induced mutations in the *Sox3* locus as indicated by the presence of cut products (blue arrows). Each sample came from independent transfections.

9 CAS9 ZYGOTIC AND GERMLINE GENE DRIVE PAPER

9.1 INTRODUCTION

This paper addresses the major aim of this thesis by the design, implementation, and testing of both zygotic and germline CRISPR/Cas9-based gene drives in *M. musculus*.

As discussed in sections 6.2-6.4, CRISPR gene drive research, although offering enormous potential, is still in its infancy. When the research presented in this thesis started, there had only been 3 published papers that contained experimental CRISPR gene drives.[82, 83, 85] Since then, that number has increased to 21 papers with all experiments performed in flies and mosquitoes except for a single yeast paper.[82, 83, 85, 89, 90, 93-107, 109] The idea to generate a CRISPR gene drive in a vertebrate, a radically different species to that reported in the other papers, presented a great opportunity for novel and important research. The "single-generation" homing demonstrated in mice by Grunwald *et al.* (2019) further bolstered the promise of vertebrate gene drives for invasive pests.

The approach taken here was with safety as paramount, utilising both a split drive and a synthetic target to reduce the risk of an accidental release and spread of the gene drive to basically nil. The research outcomes, methods and rationale for experimental designs are fully discussed within the paper below. Further details regarding additional experiments and methods not relevant to the final publication are presented in section 9.3.

9.2 "PROGRESS TOWARD ZYGOTIC AND GERMLINE GENE DRIVES IN MICE" PAPER

The 10 pages following the statement of authorship below contain the full paper proof, currently accepted for publication in The CRISPR Journal. The following 15 pages after that contain the supplemental information to be published with the paper.

Statement of Authorship

Title of Paper	Progress toward zygotic and germline gene drives in mice
Publication Status	☐ Published ☑ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Pfitzner C, White MA, Piltz SG, Scherer M, Adikusuma F, Hughes JN, & Thomas PQ (2020). Progress toward zygotic and germline gene drives in mice, The CRISPR Journal.

Principal Author

Name of Principal Author (Candidate)	Chandran Pfitzner
Contribution to the Paper	Contributed to conceptualisation. Contributed significantly to methodology. Performed majority of investigation, data analysis, and data interpretation. Handled all data curation, wrote first draft of manuscript, and created all visualisations. Reviewed all other author manuscript suggestions and edited final draft.
Overall percentage (%)	90%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 1/10/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Melissa A White
Contribution to the Paper	Performed investigation on some aspects. Reviewed manuscript during drafting.
Signature	Date 1/10/20

Name of Co-Author	Sandra G Piltz
Contribution to the Paper	Acquired some material for investigation and performed investigation on some aspects. Reviewed manuscript during drafting.
Signature	Date 1/10/20

Name of Co-Author	Michaela Scherer

Contribution to the Paper	Acquired some material for investigation and performed investigation on some aspects. Reviewed manuscript during drafting.
Signature	Date: 21/09/2020

Name of Co-Author	Fatwa Adikusuma
Contribution to the Paper	Performed investigation on some aspects. Reviewed manuscript during drafting.
Signature	Date: 21/09/2020

Name of Co-Author	James N Hughes
Contribution to the Paper	Contributed to conceptualisation and methodology. Acquired some material for investigation and performed investigation on some aspects. Reviewed manuscript during drafting.
Signature	Date: 11/10/2020

Name of Co-Author	Paul Q Thomas
Contribution to the Paper	Performed majority of conceptualisation and methodology. Acquired funding, performed project administration and supervision. Acquired some material for investigation. Contributed to data interpretation. Performed significant revision and editing to manuscript during drafting.
Signature	Date: 2/9/2020.

Please cut and paste additional co-author panels here as required.

The CRISPR Journal
Volume 3, Number 5, 2020

DOI: 10.1089/crispr.2020.0050

The CRISPR Journal

ORIGINAL ARTICLE

Progress Toward Zygotic and Germline Gene Drives in Mice

Chandran Pfitzner,[1,2] Melissa A. White,[2,3,4] Sandra G. Piltz,[2,3,4] Michaela Scherer,[2,3,4] Fatwa Adikusuma,[2,3,5] James N. Hughes,[1] and Paul Q. Thomas[2,3,4,*]

Abstract

CRISPR-based synthetic gene drives have the potential to deliver a more effective and humane method of invasive vertebrate pest control than current strategies. Relatively efficient CRISPR gene drive systems have been developed in insects and yeast but not in mammals. Here, we investigated the efficiency of CRISPR-Cas9-based gene drives in *Mus musculus* by constructing "split drive" systems where gRNA expression occurs on a separate chromosome to Cas9, which is under the control of either a zygotic (CAG) or germline (Vasa) promoter. While both systems generated double-strand breaks at their intended target site *in vivo*, no homology-directed repair between chromosomes ("homing") was detectable. Our data indicate that robust and specific Cas9 expression during meiosis is a critical requirement for the generation of efficient CRISPR-based synthetic gene drives in rodents.

Introduction

Invasive rodents, including mice, pose a significant threat to biodiversity, particularly on islands, and are the likely cause of many species extinctions.[1–3] The economic burden to the agriculture industry is also considerable, costing tens of millions of dollars to many countries each year.[4] Previous attempts at invasive vertebrate pest control have had some success, but there are still many challenges, including cost and ethical considerations.[5]

Manipulation of natural gene drives has been proposed as a tool for population suppression of invasive rodent pests by rapidly spreading a gene through a wild population that would ultimately have a deleterious effect on reproductive fitness. These include transposable elements (TEs),[6] homing endonuclease genes (HEGs),[7] and meiotic drives.[8] However, all these have significant drawbacks. TEs copy themselves to unpredictable locations, causing their own repression or disruption of endogenous genes.[6] HEGs are not found in animals, and rely on site-specific nucleases that are difficult to tailor to a specific location.[7,9] Meiotic drives are rare, poorly understood, and generally very large, making them difficult to manipulate.[10,11]

Synthetic clustered regularly interspaced short palindromic repeats (CRISPR)-based gene drives provide an alternative that has a small genomic footprint, can be inserted almost anywhere in the genome, and can be created relatively quickly and easily.[12,13] CRISPR gene drives are composed of a cassette integrated into a specific genomic site that expresses CRISPR-associated protein 9 (Cas9) endonuclease and a customizable guide RNA (gRNA) designed to cut the homologous wild-type (WT) locus.[12,13] Repair of the double-strand break (DSB) by homology-directed repair (HDR; using the gene drive allele as a repair template) results in conversion of the WT allele to a gene drive allele, in a process termed ''homing,'' which renders the cell homozygous for the gene drive allele. Homing can be restricted to the gamete precursors, resulting in selective homozygosity in the germline, while the somatic cells remain heterozygous.[12,13] The homing event will ensure that the gene drive allele will be present in all of the gametes and will be transmitted to all of that organism's progeny.[12,13] Thus, over several generations, gene drives will rapidly spread through a given population. Alternatively, repair of the DSB by error-prone pathways such as

[1]School of Biological Sciences and [3]Adelaide Medical School, The University of Adelaide, Adelaide, Australia; [2]Precision Medicine, South Australian Health and Medical Research Institute, Adelaide, Australia; [4]Robinson Research Institute, The University of Adelaide, Adelaide, Australia; and [5]CSIRO Synthetic Biology Future Science Platform, Canberra, Australia.

*Address correspondence to: Paul Q. Thomas, Room 1.208, SAHMRI, North Tce, Adelaide, SA 5000, Australia, Email: paul.thomas@adelaide.edu.au

non-homologous end joining (NHEJ), microhomology-mediated end joining, and single-strand annealing can generate an indel that interrupts the gRNA binding sequence, thereby creating a mutant allele that is immune to homing.[12,13] These so-called resistant alleles pose a significant barrier to gene drive spread.[14]

Highly efficient CRISPR gene drive systems have already been generated in the following species, where germline homing rates (in parentheses) are defined as the percentage of WT alleles converted to gene drive alleles: *Drosophila melanogaster* (62–94%),[15–17] *Anopheles stephensi* (>96.9–99%),[18] *Saccharomyces cerevisiae* (99%),[19] and *Anopheles gambiae* (87.3–99.3%).[20] A gene drive system in mice was recently generated by Grunwald *et al.* whereby a CAG-driven Cas9 was activated in the germline by a Vasa-driven Cre recombinase through deletion of a floxed-stop cassette.[21] Highly variable and relatively inefficient (0–72%) homing was observed in females, but no homing was detected in males. While this study suggests that it may be possible to develop deployable gene drives in mice, key components, including the choice of promoter for Cas9 expression, are yet to be assessed.

Here, we describe the first attempt to develop mice with functional gene drive systems in which Cas9 is directly driven by either zygotic or germline promoters. To ensure that the mice we generated did not pose any threat to the environment if unintentionally released,[22] we employed a ''synthetic target'' strategy as a molecular safeguard. In addition, we used a ''split drive'' system where Cas9 is located on a separate chromosome from the gRNA-containing homing cassette, preventing high rates of homing unless the two lines are deliberately crossed.[22]

Methods

Mouse model generation

Cas9 mRNA for zygotic injections was generated from the *Xho*I-digested pCMV/T7-hCas9 plasmid (Toolgen) using the mMESSAGE mMACHINE® T7 ULTRA Transcription Kit (Ambion) and purified using a RNeasy Mini Kit (Qiagen).

gRNA was generated with T7 promoter-containing oligos (Sigma–Aldrich/IDT) from PX459 V2.0 (Addgene; 62988)[23] using a HiScribe™ T7 Quick High Yield RNA Synthesis Kit (NEB) and purified using a RNeasy Mini Kit (Qiagen).

Gibson assembly was used to generate *Vasa-Cas9* dsDNA. The Vasa promoter fragment and associated *β*-globin-II intron was amplified from pVasa-Cre (Addgene; 15885), previously characterized by Gallardo *et al.*[24] Cas9-BGH was amplified from PX459 V2.0 (Addgene; 62988).[23] Vasa-*β*-globin-II and Cas9-BGH were assembled into pStart-K (Addgene; 20346). See Supplementary Table S3 for primers.[25] This plasmid was digested with *Bam*HI, purified using Gel DNA Recovery Kit (Zymoclean), and further purified on a floating dialysis membrane for 2.5 h. Construct had no homology arms and was injected into zygotes as per Table 1 for random integration. Rosa26-gRNA and Cas9 were included in the injection mix to stimulate cellular repair pathways. Lines *Vasa-Cas9-2* and *Vasa-Cas9-4* were generated from the same construct. The transgene integration site was not characterized for either line.

An unanticipated side effect of the presence of the $Tyr^{gRNA\text{-}Tomato}$ allele resulted in the sudden and largely unexplained deaths of mice carrying that allele. Lethality was observed for homozygotes between P23 and 36, and the earliest heterozygote death was P67. As genotyping for homing experiments was performed on blastocysts, embryos, and weanlings, it is very unlikely this phenotype masked *bona fide* homing events. To preclude this possibility completely, and due to the ethical concerns surrounding this unexpected toxicity, the $Tyr^{gRNA\text{-}Lite}$ mouse was designed and generated as shown in Figure 1B. Mouse models were generated as detailed in Table 1.

Mouse crosses

Sequencing of Tyr^{Target} loci in *Vasa-Cas9-2/+* ; $Tyr^{gRNA\text{-}Tomato/Target}$ and *Vasa-Cas9-4/+* ; $Tyr^{gRNA\text{-}Lite/Target}$ mice confirmed no carryover of maternal *Cas9* mRNA into eggs, as no indels were seen.

Table 1. Zygotic Injection Mixtures for Mouse Model Creation

Mouse	*Location*	*Donor DNA (ng/μL)*	*Cas9 (ng/μL)*	*gRNA (ng/μL)*
$Tyr^{gRNA\text{-}Tomato}$	Pronuclei	Plasmid (10)	mRNA (25)	Tyr-gRNA (10)
$Tyr^{gRNA\text{-}Lite}$	Pronuclei	ssDNA (10)	PNA Bio Cas9 protein (50)	Tyr-gRNA (25)
Tyr^{Target}	Cytoplasm	dsDNA (100)	mRNA (12.5)	Tyr-gRNA (5)
Vasa-Cas9	Pronuclei	dsDNA (3)	PNA Bio Cas9 protein (50)	Rosa26-gRNA (25)
$Rosa26^{Cas9}$	N/A	N/A	N/A	N/A

Reagents were injected into C57BL/6J zygotes, as described in the table, transferred to pseudo-pregnant recipients, and allowed to develop to term. $Tyr^{gRNA\text{-}Tomato}$ plasmid was purchased from GenScript. $Tyr^{gRNA\text{-}Lite}$ ssDNA was purchased from IDT as a Megamer®. Tyr^{Target} dsDNA was purchased from IDT. Tyr-gRNA targeted Tyrosinase (*Tyr*) intron 1 and was designed using the Zhang lab tool.[26] The Rosa26-gRNA design has been published previously.[27] $Gt(ROSA)26Sor^{tm1.1(CAG\text{-}cas9*,\text{-}EGFP)Fezh}$/J ($Rosa26^{Cas9}$) mice were supplied by JAX.[27] See Supplemental Table S4 for all sequences.

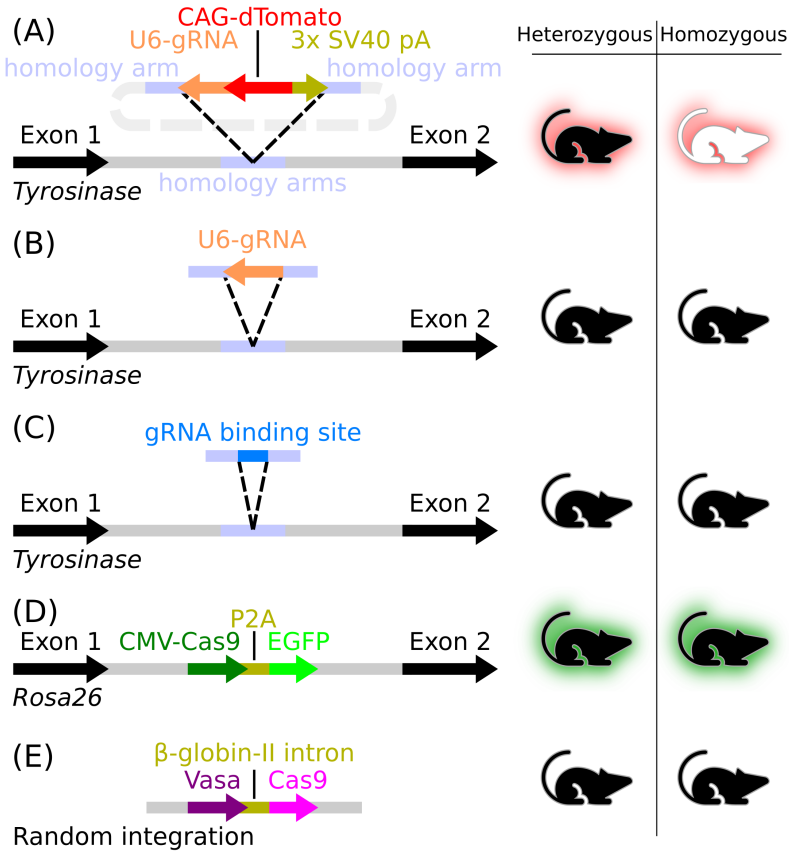


FIG. 1. Mouse models. **(A)** *Tyr*$^{gRNA\text{-}Tomato}$ mice contained a U6-driven Neo-gRNA, a CAG-driven dTomato gene, and a 3×SV40 polyA signal inserted into *Tyr* intron 1. **(B)** *Tyr*$^{gRNA\text{-}Lite}$ mice contained a U6-driven Neo-gRNA inserted into the same location as **(A)** in *Tyr* intron 1. **(C)** *Tyr*^{Target} mice contained the Neo-gRNA target sequence and NGG protospacer adjacent motif inserted into the same location as **(A)** in *Tyr* intron 1. **(D)** The *Rosa26*^{Cas9} mouse contained a CAG-driven *cas9* linked via P2A to *eGFP* in the Rosa26 locus.[27] **(E)** The *Vasa-Cas9* mouse lines contained a Vasa-driven *hSpCas9* linked via the *β*-globin-II intron (non-targeted integration). Coat color and fluorescence phenotypes are shown on the right.

DNA extractions

gDNA was extracted from embryo, tail tip, or ear notch biopsies using the High Pure PCR Template Preparation Kit (Roche), KAPA Express Extract kit (Roche), or MyTaq™ Extract-PCR Kit (Bioline). gDNA was prepared from blastocysts in a 20 μL solution of 185.5 mM pH 8.3 Tris-HCl, 185.5 mM KCl, 7.4×10^{-6}% gelatin, 8.3×10^{-4}% Polysorbate 20, 1.48% tRNA from baker's yeast (Sigma–Aldrich), and 1.15 mg/mL Proteinase K (Thermo Fisher Scientific), which was incubated at 56°C for 10 min and 95°C for 10 min.

gDNA was prepared from sperm by washing epididymides in phosphate-buffered saline (PBS) at 37°C followed by transfer to the center well of a 37°C Center-Well Organ Culture Dish (Falcon), containing 500 μL M2 medium (Sigma–Aldrich) in the center well and 3 mL PBS in the outer well. Several incisions were made across epididymides before incubating at 37°C/5% CO_2 for 5 min. The M2 medium was then centrifuged at 400 *g* for 10 min. The supernatant was discarded, 500 μL PBS was added, and then the mixture was centrifuged at 8,000 *g* for 1 min. The supernatant was discarded, and the pellet was re-suspended in 100 μL PBS. Tissue Lysis Buffer (400 μL; Roche) and Proteinase K (50 μL; Roche) was added and vortexed. The solution was then incubated at 55°C for 1 h, 50 μL of 1 M DTT was added, vortexed rigorously, and then incubated O/N at 55°C. A High Pure PCR Template Preparation Kit

(Roche) was then used for DNA extraction, eluting in 80 μL EB.

RNA extraction

Acid guanidinium thiocyanate-phenol-chloroform RNA extraction was performed on testes, ovaries, and spleens. RNA was purified using the RNeasy Mini/Micro kits (Qiagen) in conjunction with RNase-Free DNase Set (Qiagen). cDNA was generated using the High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems).

Genotyping analysis

Polymerase chain reaction (PCRs) and quantitative PCRs (qPCRs) were performed using the primers shown in Supplementary Table S3.

To assay copy number of the $Tyr^{gRNA\text{-}Lite}$ homing construct in sperm, droplet digital PCR (ddPCR) was performed using a Bio-Rad-designed assay targeting $Tyr^{gRNA\text{-}Lite}$ (dCNS586703446). A second assay targeting Rpp30 (dMmuCNS822293939) was used in every sample as a reference locus to normalize for quantity and quality differences between samples. Copy number was determined by comparison to genomic DNA with known copy numbers as shown in Figure 5 and by comparison to control DNA taken from somatic (ear) tissue of the same mice from which the sperm DNA was extracted. DNA (12.5 ng ear/sperm) was digested with *Mse*I, and droplet generation was performed on a QX200 (Bio-Rad). PCR was run, and droplets were read on a QX200 (Bio-Rad).

Sanger sequencing was performed by the Australian Genome Research Facility (Adelaide, Australia). Indels at Tyr^{Target} were detected by restriction fragment length polymorphism (RFLP) in combination with T7 endonuclease digestion to cleave heteroduplexes.

Ethics

All animal work was conducted in accordance with Australian guidelines for the care and use of laboratory animals following approval by the University of Adelaide Animal Ethics Committee (approval number S-2016-024) and SAHMRI Animal Ethics Committee (approval numbers SAM253 and SAM271).

Results

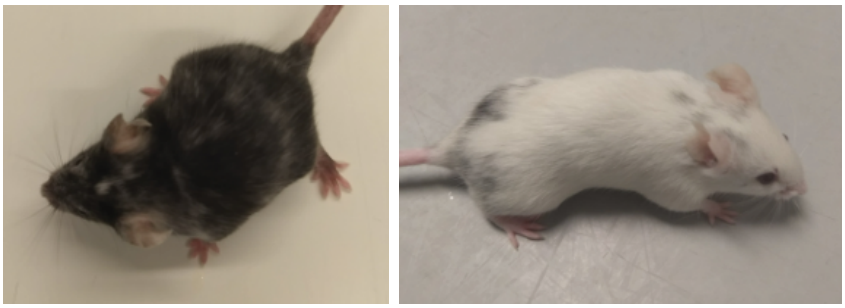
Design and generation of the homing system

To generate a ''synthetic target'' site, we identified a candidate homing gRNA (Neo-gRNA) targeting the bacterial kanamycin kinase gene and confirmed its on-target cleavage activity using embryonic stem cells harboring that gene (Supplementary Fig. S1A). Next, we generated a ''donor'' mouse line ($Tyr^{gRNA\text{-}Tomato}$) by inserting a Neo-gRNA expression cassette[28] and a ubiquitous dTomato reporter gene[29] into the first intron of the *Tyr* gene (Fig. 1A and Supplementary Fig. S2A and B). To generate a null allele, we included a 3×SV40 polyadenylation signal in the same orientation as *Tyr* so that the *Tyr* transcript is likely to be subject to premature termination. Inactivation of *Tyr* was confirmed by the white coat of $Tyr^{gRNA\text{-}Tomato}$ homozygotes (Supplementary Fig. S2C). We also generated a complementary ''receiver'' mouse line (Tyr^{Target}) with the Neo-gRNA target sequence at the same locus as $Tyr^{gRNA\text{-}Tomato}$ (Fig. 1C). We reasoned that by targeting an intronic location, indel mutations generated by error-prone repair pathways would not significantly reduce tyrosinase (TYR) function. Therefore, conversion of the Tyr^{Target} allele to a $Tyr^{gRNA\text{-}Tomato}$ allele via zygotic homing would generate a white mouse, whereas error-prone repair pathways would generate a black mouse. As expected, the Tyr^{Target} insertion did not detectably alter TYR activity, as demonstrated by the black coat of Tyr^{Target} homozygous mice (Supplementary Fig. S3).

To assess homing, Cas9 was expressed using a separate transgene. For zygotic homing, we used $Rosa26^{Cas9}$ mice,[27] which express Cas9 (and enhanced green fluorescent protein (eGFP)) ubiquitously (Fig. 1D). For germline homing experiments, we generated *Vasa-Cas9* lines via random integration (Fig. 1E).

Zygotic homing

To assess homing in the zygote, we generated 96 $Rosa26^{Cas9}/+$; $Tyr^{gRNA\text{-}Tomato/Target}$ mice. Two groups were identified based on coat color: black mice ($n=88$) and dappled mice ($n=8$). Of the 88 black mice, 83 (94%) carried indels in the Tyr^{Target} allele, indicating efficient production and high cleavage activity of the Neo-gRNA/Cas9 complex (Fig. 2). The presence of dappled mice suggested that somatic homing may be occurring in multicell embryos. To assess whether somatic homing had occurred in some cells of the dappled mice, we quantitated the $Tyr^{gRNA\text{-}Tomato}$ allele using ddPCR (Supplementary Fig. S4). Surprisingly, only a single copy of the $Tyr^{gRNA\text{-}Tomato}$ allele was present, indicating a lack of homing. Based on our previous observation that DSB repair in zygotes often generates large (>100 bp) deletions,[30] we amplified the target locus using primers distant from the cleavage site. Large deletions were identified in all dappled mice via PCR (Supplementary Fig. S5A). Sanger sequencing of those samples (Supplementary Fig. S5B) demonstrated the deletions extended into exon 1 to generate null alleles consistent with the partial (mosaic) albino phenotype. Notably, multiple Tyr^{Target} indel alleles were also



Genotype	Coat Color	
	Black	Dappled
Large Δ mosaic	0	2
Indel	17	0
Indel mosaic	39	0
No activity/indel mosaic	27	6
No activity	5	0

	Totals
Any indel/large Δ	83
No activity	5

FIG. 2. Zygotic-homing gene drive results. *Top*: Representative *Rosa26*^{Cas9}/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice with dappled coats, showing wide variability in mosaicism. *Bottom*: Table of genotypes of *Tyr*^{Target} allele(s) in *Rosa26*^{Cas9}/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ gene drive mice. "Large Δ mosaic" contain a large deletion and one or more indels. "Indel" contain a single indel. "Indel mosaic" contain one or more indels. "No activity/indel mosaic" contain an uncut *Tyr*^{Target} allele and one or more indels/large deletions. "No activity" contain a single uncut *Tyr*^{Target} allele. Dappled is a mix of black and white fur.

detected in many black mice (Fig. 2), indicating mosaicism. Taken together, these data indicate that despite efficient cleavage of the target sequence, homing did not occur in zygotes or cleavage-stage embryos.

Germline homing

For germline homing experiments, we initially generated a *Vasa-Cas9* transgenic mouse using the previously characterized 5.6 kb Vasa promoter fragment (*Vasa-Cas9-2*; Fig. 1E), which has been shown to drive robust Cre expression in the male and female germline.[24] Reverse transcription qPCR analysis showed that Cas9 was expressed in both testes and ovaries, although, unexpectedly, expression in the latter was extremely low (Supplementary Fig. S6A). To assess germline homing, *Vasa-Cas9-2*/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice were generated and mated with WT partners. If homing occurred in the germline, >50% of their progeny would carry the *Tyr*$^{gRNA\text{-}Tomato}$ red fluorescence marker, while no homing would result in ~50% transmission (Fig. 3). Altogether, we screened 355 offspring from six *Vasa-Cas9-2*/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ males (119 offspring) and 12 *Vasa-Cas9-2*/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ females (236 offspring; Fig. 4 and Supplementary Table S1). No significant increase (χ^2 test) in *Tyr*$^{gRNA\text{-}Tomato}$ transmission was observed in males (53.8%; $p=0.46$) or females (48.7%; $p=0.75$). To assess gRNA/Cas9 cleavage activity, non-fluorescent offspring were screened for indels in the *Tyr*^{Target} allele by RFLP (Supplementary Fig. S7 and Supplementary Table S1). Only a low percentage of indels were present in progeny from both males (14.5%) and females (9.3%), with the majority of *Tyr*^{Target} alleles remaining uncut.

To investigate if higher Cas9 expression levels would promote germline homing, we generated an additional *Vasa-Cas9* mouse line (*Vasa-Cas9-4*). Expression analysis showed a similar Cas9 mRNA level to *Vasa-Cas9-2* in the testes (Supplementary Fig. S6A). Cas9 mRNA levels in the ovaries were higher than *Vasa-Cas9-2* but still much lower than in the testes. Transgene copy number was higher in *Vasa-Cas9-4* (Supplementary Fig. S6B), consistent with the higher expression in ovaries, although it remains unclear why elevated expression was not

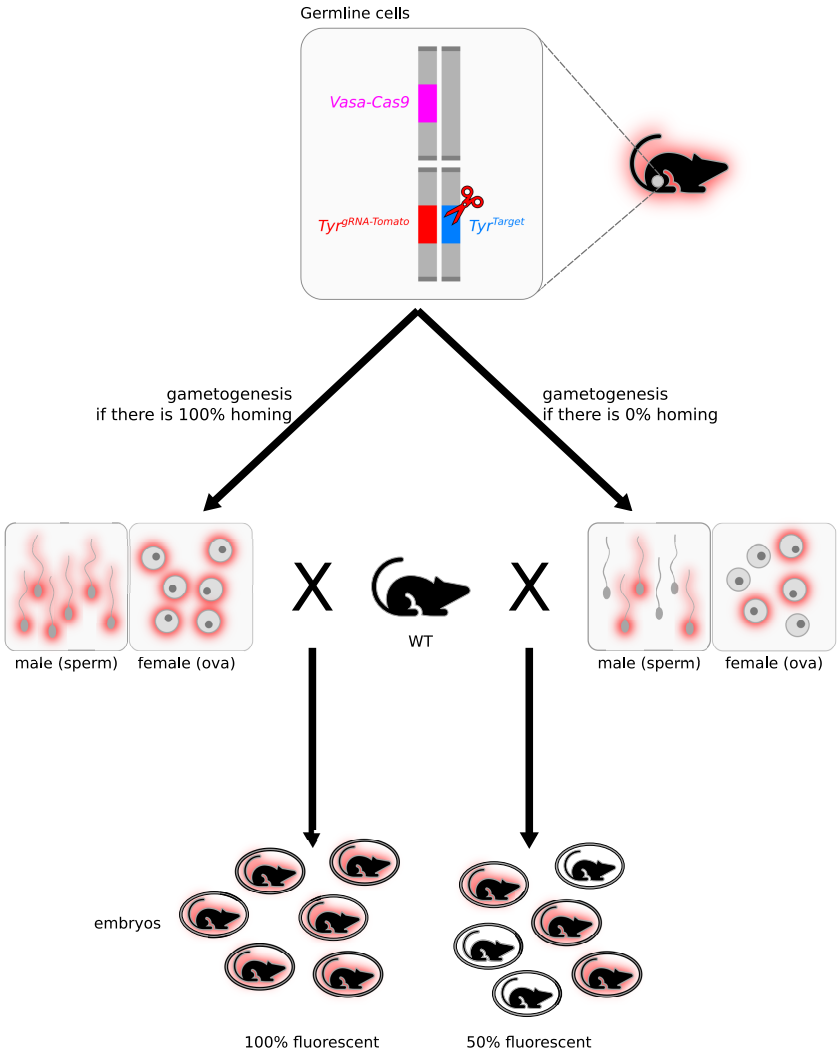


FIG. 3. Germline-homing gene drive activity. Experimental mice (top right) contain *Vasa-Cas9*, $Tyr^{gRNA\text{-}Tomato}$ and Tyr^{Target} alleles. Somatic tissue will not express Cas9 and thus maintain an intact Tyr^{Target}. Cas9 produced in germline tissue (top) will complex with Neo-gRNA and generate a double-strand break. DNA repair mechanisms would then either copy the $Tyr^{gRNA\text{-}Tomato}$ allele over from the donor chromosome (homing) or create an indel ($Tyr^{Target\Delta}$). The sperm and ova produced (middle) are shown in the two most extreme alternative possibilities, 100% or 0% homing. The actual homing percent is calculated after crossing the original mouse with a wild-type mouse and counting the number of fluorescent offspring.

detected in *Vasa-Cas9-4* testes. Germline homing experiments were performed as described for *Vasa-Cas9-2*, except with a shorter donor construct ($Tyr^{gRNA\text{-}Lite}$) lacking the dTomato expression cassette to exclude any possibility of influence from an observed lethality phenotype (see Methods). Offspring from *Vasa-Cas9-4*/+ ; $Tyr^{gRNA\text{-}Lite/Target}$ × WT matings were genotyped by PCR (Fig. 4 and Supplementary Table S2). Altogether, we screened 620 offspring from 10 transgenic males (255 offspring) and 10 transgenic females (365 offspring). Similar to *Vasa-Cas9-2*, no significant increase (χ^2 test) in $Tyr^{gRNA\text{-}Lite}$ transmission >50% was detected from transgenic males (49.0%; $p = 0.80$) or females (51.8%; $p = 0.53$). A total of 87.1% of Tyr^{Target}/+

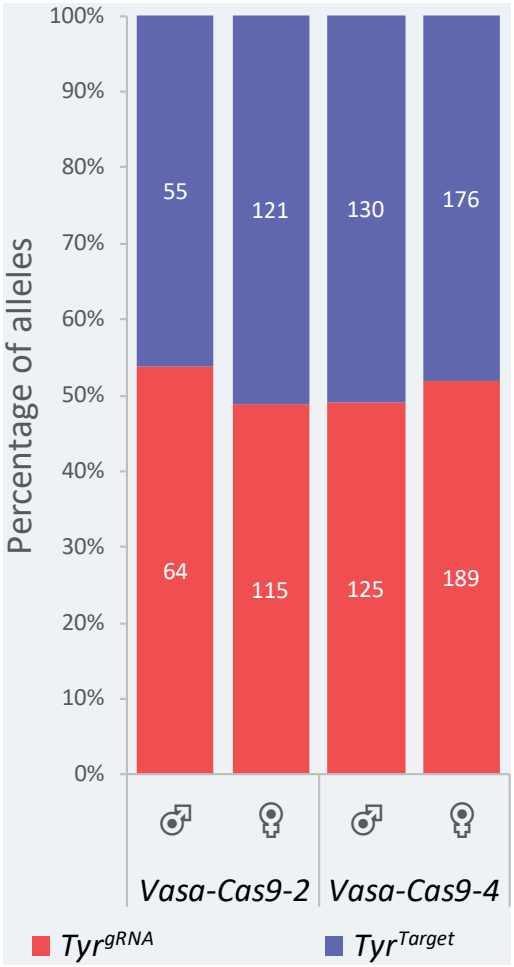


FIG. 4. Germline-homing genotyping data. Bar graph showing the percentage of different alleles in the offspring of gene drive mice *Vasa-Cas9-2*/+ ; $Tyr^{gRNA\text{-}Tomato/Target}$ and *Vasa-Cas9-4*/+ ; $Tyr^{gRNA\text{-}Lite/Target}$. Total number of offspring shown in columns.

offspring of males carried indels at the target locus compared to 11.0% of Tyr^{Target}/+ offspring of females (Supplementary Fig. S7A and Supplementary Table S2). Finally, to investigate whether a low level of homing was occurring in *Vasa-Cas9-4*/+ ; $Tyr^{gRNA\text{-}Lite/Target}$ male mice, we investigated whether >50% of sperm carried the $Tyr^{gRNA\text{-}Lite}$ allele using ddPCR (Fig. 5). No difference in $Tyr^{gRNA\text{-}Lite}$ level between sperm and the somatic control (ear) was detected, indicating an absence of homing.

Discussion

Zygotic homing

While the vast majority of gene drive experiments have employed germline promoters to drive Cas9, it is also notionally possible for homing to occur in the zygote. Indeed, targeted integration of dsDNA templates can occur (albeit inefficiently) in zygotes, confirming that HDR is active in the early embryo.[31] Further, it has recently been reported that ''interhomolog repair'' (effectively the same process as gene drive homing) can occur in human embryos,[32] although concerns have been raised about the interpretation of these results.[30,33] In our experiments using the constitutive CAG promoter, we found no evidence of homing in zygotes, despite the high efficiency of DSB generation (>95% of $Rosa26^{Cas9}$/+ ; $Tyr^{gRNA\text{-}Tomato/Target}$ mice have indels). Thus, we conclude that error-prone repair pathways predominate over HDR in the early embryo, consistent with the recent observations of Grunwald *et al.*[21] and earlier observations due to maternal carryover in flies[16] and mosquitos.[18] While this DSB repair bias may reflect the availability of endogenous DNA repair proteins, the separation of the donor and receiver alleles into distinct pronuclei until after G_2 phase (18–20 h after fertilization)[34] would also limit the opportunity for zygotic homing to occur.

Approximately 77% of $Rosa26^{Cas9}$/+ ; $Tyr^{gRNA\text{-}Tomato/Target}$ mice were mosaic, having multiple Tyr^{Target} alleles, including many with more than three alleles. Zygotic chromosomes are transcriptionally repressed until the G_2 phase,[35] meaning that the formation of the gRNA/Cas9 complex is delayed until after S phase or cell division, resulting in mosaicism. The strength of the CAG and U6 promoters may also be limiting, resulting in insufficient gRNA/Cas9 complex formation and/or translocation to the Tyr^{Target} locus in the zygote. It is also possible that flawless NHEJ-mediated DNA repair occurs in a proportion of zygotes, thereby delaying the generation of indels until the two-cell stage or later. Surprisingly, given that the Cas9 and gRNA promoters are constitutive, 5% of $Rosa26^{Cas9}$/+ ; $Tyr^{gRNA\text{-}Tomato/Target}$ mice had an intact Tyr^{Target} allele, and a further 34% had a mix of both intact alleles and indels. This high level of post-zygotic cleavage leading to mosaicism indicates that it is unlikely the intact alleles remain uncut, but instead suggests that flawless NHEJ-mediated DNA repair occurs at a relatively high level in somatic cells, as has been seen in other gene drive systems.[20,21] In the five mice that were free of indels, continued production of eGFP (linked to Cas9 via P2A) indicates that Cas9 protein is still being produced, and continued production of

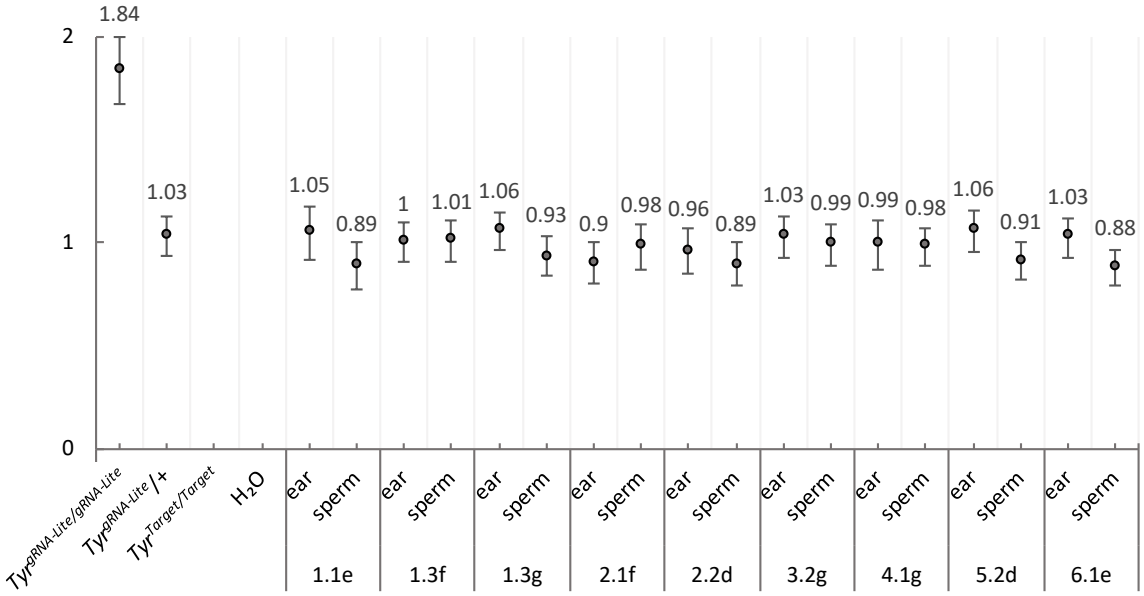


FIG. 5. *Tyr*$^{gRNA\text{-}Lite}$ droplet digital polymerase chain reaction assay. Comparison of *Tyr*$^{gRNA\text{-}Lite}$ levels (with 95% confidence intervals) in genomic DNA from somatic tissue (ear) and sperm from *Vasa-Cas9-4/+* ; *Tyr*$^{gRNA\text{-}Lite/Target}$ mice. Genomic DNA with copy number of 2 (*Tyr*$^{gRNA\text{-}Lite/gRNA\text{-}Lite}$), 1 (*Tyr*$^{gRNA\text{-}Lite/+}$) and 0 (*Tyr*$^{Target/Target}$) shown for reference.

dTomato tells us there is no gene silencing happening in the region of Neo-gRNA transcription (Supplementary Fig. S8), although it is possible that specific downregulation of the U6 promoter (for Neo-gRNA) may be occurring, or increased Cas9 mRNA/protein degradation may occur through unknown mechanisms.

It is possible that carryover of maternal mRNA or protein could have influenced the probability of homing and the developmental stage at which indels are generated. However, we found no evidence for a difference in indel frequency or timing in *Rosa26*^{Cas9}/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice based on the parental origin of the Cas9 or gRNA expression alleles (Supplementary Fig. S9).

Germline homing

Studies in *D. melanogaster*, *A. stephensi*, and *A. gambiae* have shown that homing can occur with high efficiency when Cas9 expression is driven in the germline using promoter fragments.[15,18,20] Here, we used a similar strategy to express Cas9, employing the 5.6 kb murine Vasa proximal promoter, which is one of the few characterized mammalian promoter fragments that is expressed in the male and female germline.[24] In contrast to insects, transmission of the donor allele was not statistically >50% in either of the *Vasa-Cas9* lines. While this does not completely rule out homing, especially in light of the level of homing seen by Grunwald *et al.*,[21] it is clear that it is not occurring at a substantial or useful level.

Although significant homing did not occur, sufficient Cas9 was produced to generate indels in the *Tyr*^{Target} allele in some germ cells. Analysis of Cas9 levels revealed that transgene expression was much higher in testes than in ovaries for both *Vasa-Cas9* lines, in contrast to the published *Vasa-Cre* line.[24] Consistent with the sexually dimorphic expression level of each line, indel generation was higher in males than in females. However, it remains unclear why indels were generated much more frequently in the *Vasa-Cas9-4* line versus *Vasa-Cas9-2*, given that transgene expression in the testes was similar.

Recent studies that investigated the timing of Cas9 expression in the germline, either by the comparison of different promoters[21,36] or by a small-molecule controlled system,[37] suggest that it is critical for homing efficiency. It is thought that homing is most likely to occur during meiosis when homologous chromosomes are aligned for meiotic recombination.[21,36] Generation of DSBs before or after meiosis instead promotes generation of indels via error-prone repair pathways. As the timing of meiosis is sex specific, activity of the Vasa transgene must be considered in both males and females. It has previously been shown that the Vasa promoter fragment is strongly induced at E15–18 in testes and before

P3 in ovaries.[24] In the E15–18 testes, primordial germ cells (PGCs) are undergoing mitotic proliferation and do not enter meiosis until around P9.[38] As a consequence, Vasa expression is likely too early for homing, consistent with the high frequency of indels that we observed in both *Vasa-Cas9* lines. In contrast, nascent oocytes enter meiosis at E13.5, moving through zygotene to pachytene before they arrest at around E19 in dictyotene, which is maintained until around P21.[38] As alignment of homologous chromosomes is maintained throughout this period, homing should be promoted. Why, then, did we not observe super-Mendelian transmission of the Tyr^{gRNA} alleles in females? The answer probably relates to the very low level of transgene expression in the ovary, which, despite the substantial transgene copy number, was barely above background levels. Given only 9% and 11% of female *Vasa-Cas9* offspring carried indels, the frequency of DSB generation in the female germline was likely very low, and therefore homing, if it occurred at all, was below the limit of our ability to detect. It is also possible that the Vasa promoter is activated too early in the ovary, given that endogenous Vasa expression starts at E10.5–11.5 in PGCs of both sexes.[39] Notably, in contrast to our results, a similar study by Grunwald *et al*.[21] did observe homing in females (although not in males). An important difference between our study and that of Grunwald *et al*. is their use of the *Vasa-Cre* line to express Cas9 from the CAG promoter via removal of a stop-flox cassette. Thus, Cas9 levels are likely much higher in the Grunwald *et al*. experiment due to higher levels of CAG-Cas9 compared to Vasa-Cas9 (see Supplementary Fig. S10 for a direct comparison). Although not possible in our study due to the uniform genetic background, genotyping of polymorphic markers on the ''receiver'' chromosome may be a useful strategy to identify homing events definitively.

Conclusion

Based on our observations and those of Grunwald *et al*.,[21] we suggest that zygotic homing is not a feasible strategy in mice. Efficient homing in the germline will require identification of additional promoters, potentially with robust and specific expression during meiosis in oocytes and spermatocytes. Thus, considerable experimental development is required before rodent gene drives can be considered for deployment to address conservation, agricultural, or health objectives.

Acknowledgments

The authors thank John Godwin and members of the Genetic Biocontrol for Invasive Rodents consortium for useful discussions during the course of this research project. The authors acknowledge the facilities and the scientific and technical assistance of the South Australian Genome Editing (SAGE) Facility, the University of Adelaide, and the South Australian Health and Medical Research Institute. SAGE is supported by the Australian Phenomics Network (APN). The APN is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS) program.

Author Disclosure Statement

No competing financial interests exist.

Funding Statement

This study was funded by a US Defense Advanced Research Projects Agency (DARPA) ''Safe genes'' grant to Paul Thomas (HR00111720046).

Supplementary Material

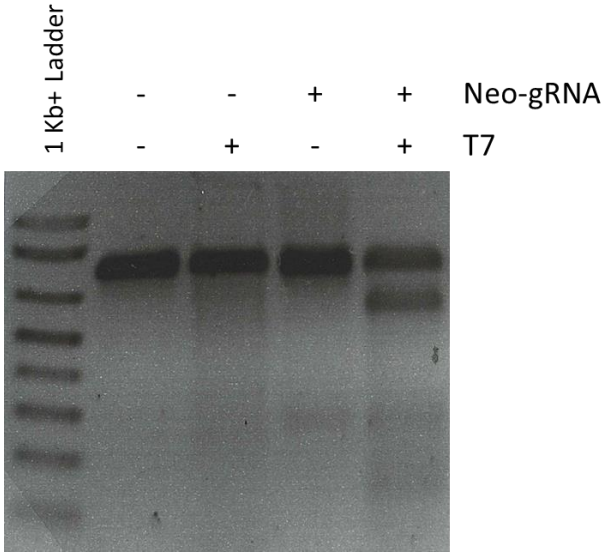
Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3
Supplementary Figure S4
Supplementary Figure S5
Supplementary Figure S6
Supplementary Figure S7
Supplementary Figure S8
Supplementary Figure S9
Supplementary Figure S10
Supplementary Table S1
Supplementary Table S2
Supplementary Table S3

References

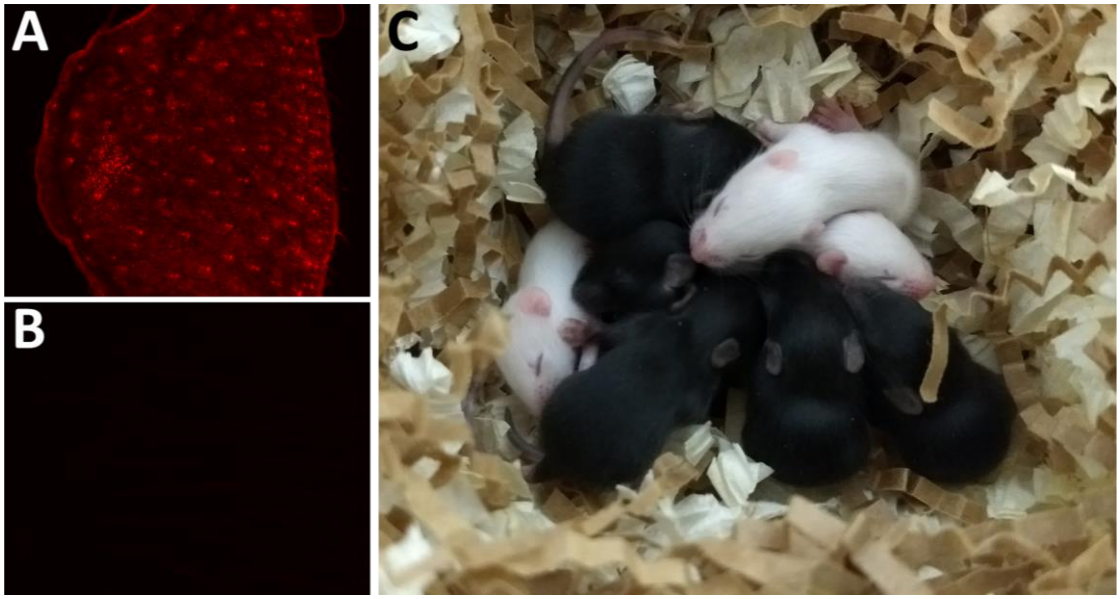
1. Blackburn TM, Cassey P, Duncan RP, et al. Avian extinction and mammalian introductions on oceanic islands. *Science* 2004;305:1955–1958. DOI: 10.1126/science.1101617.
2. Harris DB. Review of negative effects of introduced rodents on small mammals on islands. *Biol Invasions* 2008;11:1611–1630. DOI: 10.1007/s10530-008-9393-0.
3. Doherty TS, Glen AS, Nimmo DG, et al. Invasive predators and global biodiversity loss. *Proc Natl Acad Sci U S A* 2016;113:11261–11265. DOI: 10.1073/pnas.1602480113.
4. Stenseth NC, Leirs H, Skonhoft A, et al. Mice, rats, and people: the bio-economics of agricultural rodent pests. *Front Ecol Environ* 2003;1:367–375. DOI: 10.2307/3868189.
5. Gregory S, Henderson W, Smee E, et al. Eradications of Vertebrate Pests in Australia: A Review and Guidelines for Future Best Practice. Canberra, Australia: Invasive Animals Cooperative Research Centre, 2014, 1–90.
6. Sinkins SP, Gould F. Gene drive systems for insect disease vectors. *Nat Rev Genet* 2006;7:427–435. DOI: 10.1038/nrg1870.
7. Burt A, Koufopanou V. Homing endonuclease genes: the rise and fall and rise again of a selfish element. *Curr Opin Genet Dev* 2004;14:609–615. DOI: 10.1016/j.gde.2004.09.010.
8. Lyttle TW. Segregation distorters. *Annu Rev Genet* 1991;1991:511–577. DOI: 10.1146/annurev.ge.25.120191.002455.
9. Chan YS, Takeuchi R, Jarjour J, et al. The design and *in vivo* evaluation of engineered I-Onul-based enzymes for HEG gene drive. *PLoS One* 2013;8:e74254. DOI: 10.1371/journal.pone.0074254.
10. Kelemen RK, Vicoso B. Complex history and differentiation patterns of the *t*-haplotype, a mouse meiotic driver. *Genetics* 2018;208:365–375. DOI: 10.1534/genetics.117.300513.
11. Charron Y, Willert J, Lipkowitz B, et al. Two isoforms of the RAC-specific guanine nucleotide exchange factor TIAM2 act oppositely on transmission ratio distortion by the mouse *t*-haplotype. *PLoS Genet* 2019;15:e1007964. DOI: 10.1371/journal.pgen.1007964.

12. Esvelt KM, Smidler AL, Catteruccia F, et al. Concerning RNA-guided gene drives for the alteration of wild populations. *Elife* 2014;3:e03401. DOI: 10.7554/eLife.03401.
13. Gantz VM, Bier E. The dawn of active genetics. *Bioessays* 2016;38:50–63. DOI: 10.1002/bies.201500102.
14. Prowse TAA, Cassey P, Ross JV, et al. Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. *Proc Biol Sci* 2017;284:20170799. DOI: 10.1098/rspb.2017.0799.
15. Gantz VM, Bier E. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science* 2015;348:442–444. DOI: 10.1126/science.aaa5945.
16. Champer J, Reeves R, Oh SY, et al. Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations. *PLoS Genet* 2017;13:e1006796. DOI: 10.1371/journal.pgen .1006796.
17. Champer J, Liu J, Oh SY, et al. Reducing resistance allele formation in CRISPR gene drive. *Proc Natl Acad Sci U S A* 2018;115:5522–5527. DOI: 10.1073/pnas.1720354115.
18. Gantz VM, Jasinskiene N, Tatarenkova O, et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proc Natl Acad Sci U S A* 2015;112:E6736–6743. DOI: 10.1073/pnas.1521077112.
19. DiCarlo JE, Chavez A, Dietz SL, et al. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat Biotechnol* 2015;33:1250–1255. DOI: 10.1038/nbt.3412.
20. Hammond A, Galizi R, Kyrou K, et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nat Biotechnol* 2016;34:78–83. DOI: 10.1038/nbt.3439.
21. Grunwald HA, Gantz VM, Poplawski G, et al. Super-Mendelian inheritance mediated by CRISPR-Cas9 in the female mouse germline. *Nature* 2019;566:105–109. DOI: 10.1038/s41586-019-0875-2.
22. Akbari OS, Bellen HJ, Bier E, et al. Safeguarding gene drive experiments in the laboratory. *Science* 2015;349:927–929. DOI: 10.1126/science.aac7932.
23. Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 2013;8:2281–2308. DOI: 10.1038/nprot.2013.143.
24. Gallardo T, Shirley L, John GB, et al. Generation of a germ cell-specific mouse transgenic Cre line, Vasa-Cre. *Genesis* 2007;45:413–417. DOI: 10.1002/dvg.20310.
25. Wu S, Ying G, Wu Q, et al. A protocol for constructing gene targeting vectors: generating knockout mice for the cadherin family and beyond. *Nat Protoc* 2008;3:1056–1076. DOI: 10.1038/nprot.2008.70.
26. Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 2013;31:827–832. DOI: 10.1038/nbt.2647.
27. Platt RJ, Chen S, Zhou Y, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 2014;159:440–455. DOI: 10.1016/j.cell.2014.09.014.
28. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;339:819–823. DOI: 10.1126/science.1231143.
29. Shaner NC, Campbell RE, Steinbach PA, et al. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol* 2004;22:1567–1572. DOI: 10.1038/nbt1037.
30. Adikusuma F, Piltz S, Corbett MA, et al. Large deletions induced by Cas9 cleavage. *Nature* 2018;560:E8–E9. DOI: 10.1038/nature23305.
31. Yang H, Wang H, Shivalila CS, et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 2013;154:1370–1379. DOI: 10.1016/j.cell.2013.08.022.
32. Ma H, Marti-Gutierrez N, Park SW, et al. Correction of a pathogenic gene mutation in human embryos. *Nature* 2017;548:413–419. DOI: 10.1038/nature23305.
33. Egli D, Zuccaro MV, Kosicki M, et al. Inter-homologue repair in fertilized human eggs? *Nature* 2018;560:E5–E7. DOI: 10.1038/s41586-018-0379-5.
34. Ciemerych MA, Sicinski P. Cell cycle in mouse development. *Oncogene* 2005;24:2877–2898. DOI: 10.1038/sj.onc.1208608.
35. Jukam D, Shariati SAM, Skotheim JM. Zygotic genome activation in vertebrates. *Dev Cell* 2017;42:316–332. DOI: 10.1016/j.devcel.2017.07.026.
36. Li M, Yang T, Kandul NP, et al. Development of a confinable gene drive system in the human disease vector *Aedes aegypti*. *Elife* 2020;9:e51701. DOI: 10.7554/eLife.51701.
37. Lopez Del Amo V, Bishop AL, Sanchez CH, et al. A transcomplementing gene drive provides a flexible platform for laboratory investigation and potential field deployment. *Nat Commun* 2020;11:352. DOI: 10.1038/s41467-019-13977-7.
38. Hilz S, Modzelewski AJ, Cohen PE, et al. The roles of microRNAs and siRNAs in mammalian spermatogenesis. *Development* 2016;143:3061–3073. DOI: 10.1242/dev.136721.
39. Tanaka SS, Toyooka Y, Akasu R, et al. The mouse homolog of *Drosophila Vasa* is required for the development of male germ cells. *Genes Dev* 2000;14:841–853. DOI: 10.1101/gad.14.7.841.

Supplemental Figures

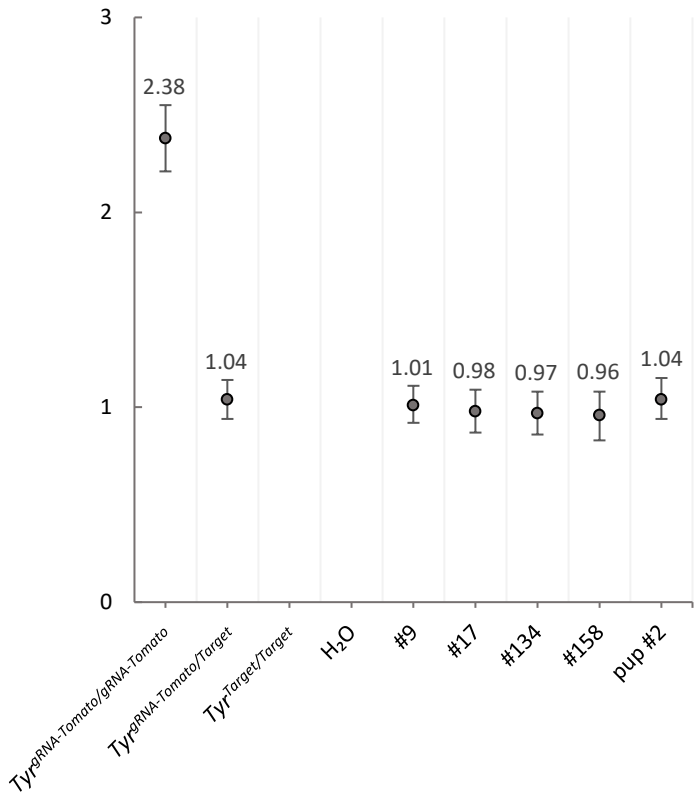


Supplemental Figure 1. Demonstration of Neo-gRNA cleavage activity at its target binding site. Mouse ES cells containing the target binding site for Neo-gRNA were transfected with a plasmid containing U6-driven Neo-gRNA and a plasmid (pX459) containing CMV-driven *hSpCas9*. RFLP analysis was performed around the cut site. Digestion of Neo-gRNA/T7 band demonstrates successful cutting by Neo-gRNA.

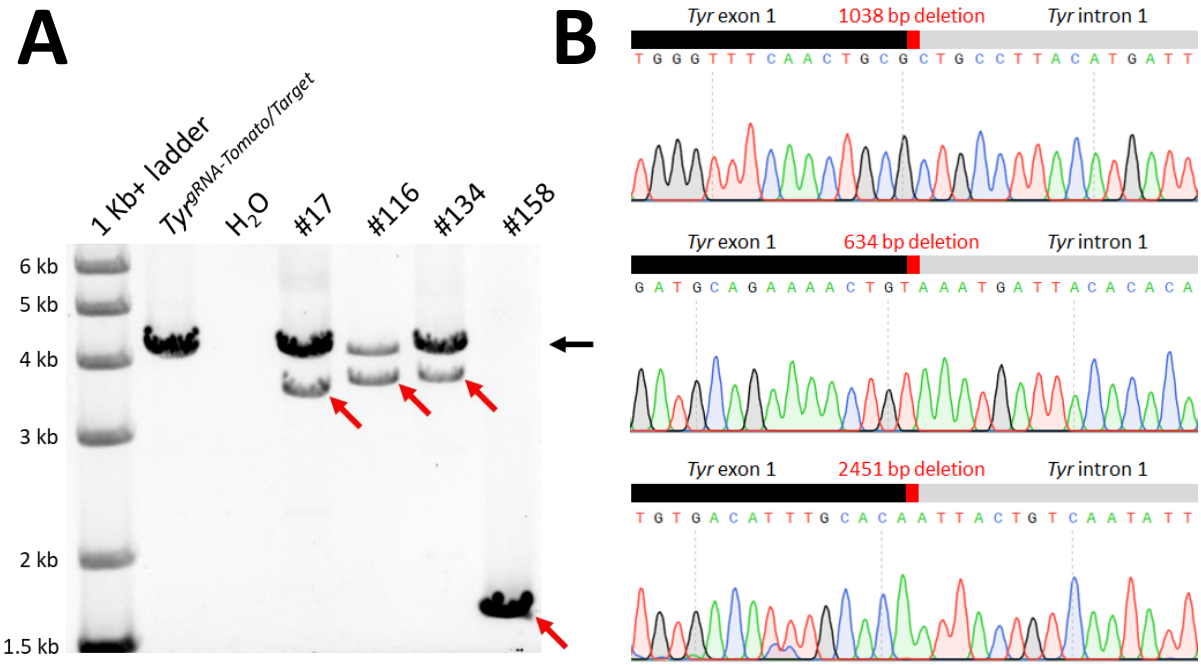


Supplemental Figure 2. Phenotyping *Tyr*$^{gRNA\text{-}Tomato}$ mice. (A, B) Representative immunofluorescence images of ear notches of **(A)** *Tyr*$^{gRNA\text{-}Tomato}$ mice and **(B)** WT mice, showing expression of dTomato. **(C)** A representative F2 litter of *Tyr*$^{gRNA\text{-}Tomato}$ mice. Black mice were heterozygous for *Tyr*$^{gRNA\text{-}Tomato}$. White mice were homozygous for *Tyr*$^{gRNA\text{-}Tomato}$.

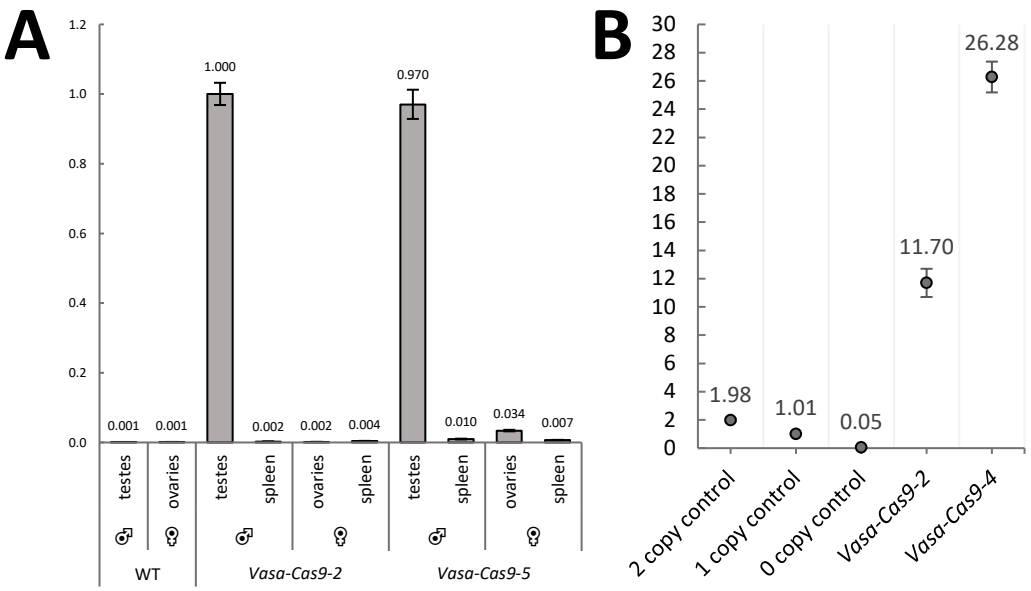
Supplemental Figure 3. *Tyr*^{Target} homozygote phenotyping, showing black coats.



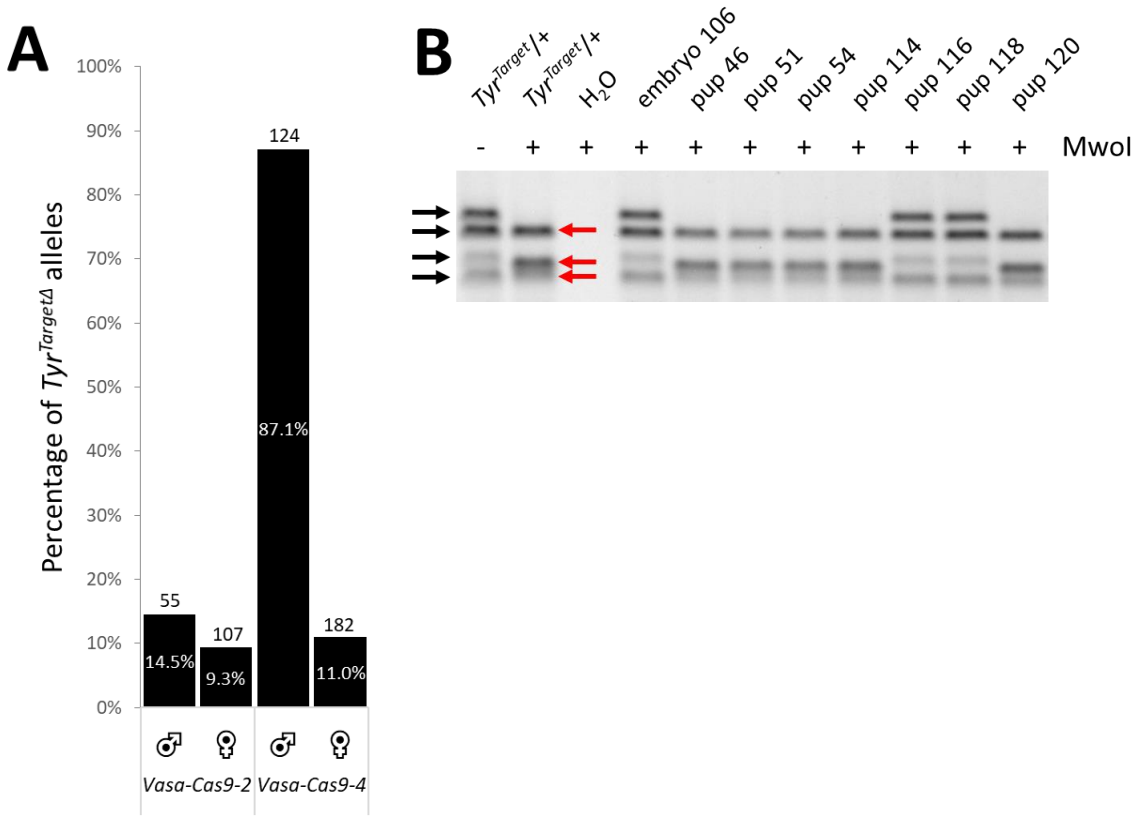
Supplemental Figure 4. *Tyr*$^{gRNA\text{-}Tomato}$ copy number determination for zygotic-homing mice. **Representative** ddPCR showing copy number of *Tyr*$^{gRNA\text{-}Tomato}$ (with 95% CI) in five *Rosa26*^{Cas9}/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice with dappled coats.



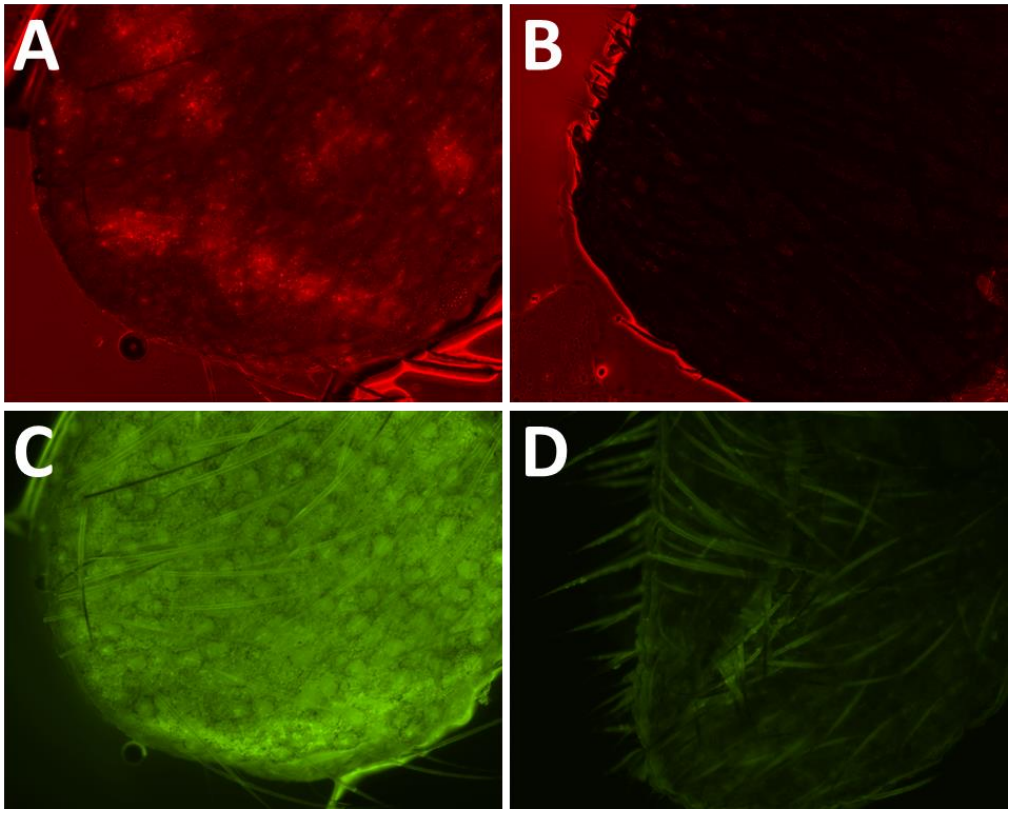
Supplemental Figure 5. Large deletion genotyping. (A) PCR showing large deletions around *Tyr*^{Target} in four *Rosa26*^{Cas9}/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice with dappled coats. The black arrow shows expected band (~4 kb) for no deletion/small indels. The red arrows show large deletions of varying size. (B) Sanger sequencing traces from three *Rosa26*^{Cas9}/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice with dappled coats showing large deletions extending into *Tyr* exon 1.



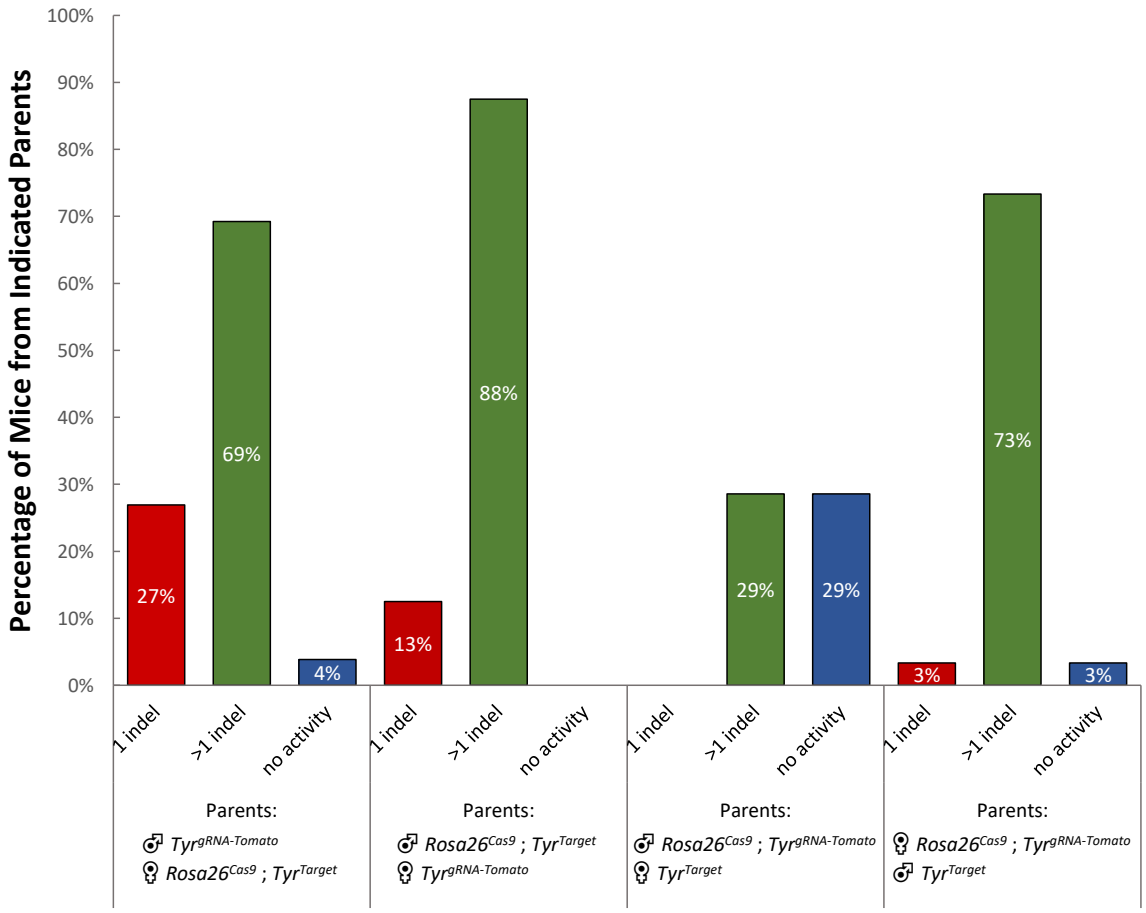
Supplemental Figure 6. *Vasa-Cas9* expression. (A) RT-qPCR showing expression levels (with 95% CI) of the *Vasa-Cas9* lines in various tissue types. **(B)** ddPCR copy number assay showing the genomic transgene copy number (with 95% CI) of the construct in *Vasa-Cas9-2* and *Vasa-Cas9-4* hemizygotes.



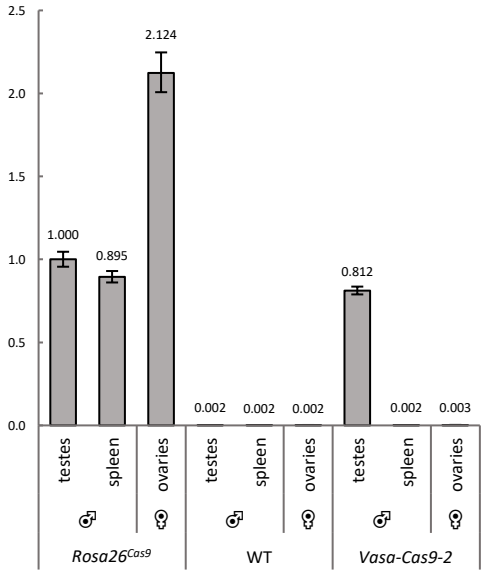
Supplemental Figure 7. Extended germline-homing genotyping data. (A) Data shows number of *Tyr*^{Target} alleles with an indel (*Tyr*^{TargetΔ}) as a percentage of total *Tyr*^{Target} alleles with and without indels in the offspring of gene drive mice (as per Figure 3), broken down into *Vasa-Cas9* line and sex. Exact percentage of indels is shown in each bar and total number of *Tyr*^{Target} alleles is shown at the top of each column. **(B)** Representative *Tyr*^{Target} cut site digestion with MwoI (as indicated) and T7 Endonuclease (all samples), black arrows show expected uncut bands due to destruction of MwoI site indicating presence of an indel, red arrows show cut bands due to intact MwoI site and thus no indel.



Supplemental Figure 8. Phenotyping of *Rosa26*^{Cas9}/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice containing no indels. Representative immunofluorescence images of ear notches of aforementioned mice **(A, C)** and WT mice **(B, D)**, showing expression of dTomato **(A, B)** and expression of EGFP **(C, D)**.



Supplemental Figure 9. Breakdown of zygotic-homing indel formation based on parent of origin. Sex symbols at bottom show which alleles were inherited from which parents.



Supplemental Figure 10. Cas9 expression comparison. RT-qPCR showing expression levels (with 95% CI) of the CAG-Cas9 construct in *Rosa26*^{Cas9} mice in comparison to *Vasa-Cas9-2* in various tissue types.

Supplemental Tables

Mouse ID	Sex	*Tyr*$^{gRNA\text{-}Tomato}$ alleles	Total *Tyr*^{Target} alleles	*Tyr*$^{Target\Delta}$ alleles
3.2b	male	23	18	4
27.1d	male	16	17	3
27.1e	male	3	6	0
33.2e	male	2	5	0
33.2f	male	9	3	0
33.2g	male	11	6	1
24.1a	female	8	5	0
33.2a	female	19	14	0
33.2b	female	3	5	0
35.1b	female	3	1	1
35.1d	female	4	4	0
35.1e	female	3	3	0
35.2a	female	24	27	3
35.2c	female	0	8	4
36.2d	female	21	26	1
37.2a	female	8	6	1
37.2b	female	10	3	0
37.2c	female	12	19	0

Supplemental Table 1. Germline-homing individual mouse genotyping data for *Vasa-Cas9-2*. Breakdown of the ratio of inherited alleles for the offspring of the listed *Vasa-Cas9-2*/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice crossed to WT. Data collated to produce *Vasa-Cas9-2* data in Figure 4 and Supplemental Figure 7.

Mouse ID	Sex	*Tyr*$^{gRNA\text{-}Lite}$ alleles	Total *Tyr*^{Target} alleles	*Tyr*^{TargetΔ} alleles
1.2e	male	15	17	13
1.3f	male	7	10	8
1.3g	male	9	12	11
2.1f	male	20	16	16
2.2d	male	13	17	17
3.2g	male	15	13	10
4.3f	male	12	7	6
4.1g	male	8	6	5
5.2d	male	14	18	16
6.1e	male	12	14	13
1.1b	female	21	21	4
1.2b	female	28	32	3
3.1b	female	24	23	0
3.2e	female	12	6	0
4.1b	female	27	24	1
4.2b	female	27	15	1
4.3d	female	17	20	1
5.2c	female	11	9	1
6.1b	female	8	14	2
6.1c	female	14	12	1

Supplemental Table 2. Germline-homing individual mouse genotyping data for *Vasa-Cas9-4*. Breakdown of the ratio of inherited alleles for the offspring of the listed *Vasa-Cas9-4*/+ ; *Tyr*$^{gRNA\text{-}Tomato/Target}$ mice crossed to WT. Data collated to produce *Vasa-Cas-4* data in Figure 4 and Supplemental Figure 7.

Locus	Primer 1	Primer 2	Primer 3	Type
Tyr$^{gRNA\text{-}Tomato}$	CCAGACAGCCCTTGTAATCATTAGC	GGCTATCGTGGCGTTTTAGA		PCR
Tyr$^{gRNA\text{-}Lite}$	CCAGACAGCCCTTGTAATCATTAGC	AACTTGAAAAAGTGGCACCGAG	GCACCTCCTATGGTATCTGGAA	PCR
Tyr^{Target}	ACTGTTTGAGAGTCAGCAACGT	TCTCTGGCCAAAACCAAGACTT		PCR/RFLP
Tyr^{Target}	GGGTTCTGTCCTCAACTGGT	TTTGATGTAAGAAGGGGAGTGGT		Large Δ PCR
Rosa26^{Cas9}	AAGGGAGCTGCAGTGGAGTA	CCGAAAATCTGTGGGAAGTC	CCATAAGGTCATGTACTGGGC	PCR
Vasa-Cas9	ATTGTACTTCAGCACAGTTTTAGAG	AGTCTCCGTCGTGGTCCTTA		PCR
Vasa-Cas9	ACCTGAACCCCGACAACA	CTGGCGTTGATGGGGTTTTC		RT-qPCR
Vasa-β-globin-II	GCTTTAAAGGAACCAATTCAGTCG-ACTGGATCCGGTACCGTGTGCCAC-CATGCCTGGCCC	TCCTTATAGTCCATCTGTAGGA-AAAAGAAGAAGGCATGAACATG-GTTAGCAGAGG		Gibson assembly
Cas9-BGH	TTCATGCCTTCTTCTTTTTCCTAC-AGATGGACTATAAGGACCACGACGG	TACAAGAAAGCTGGGTCTAGAT-ATCTCGAGTGCGGCCGCGGGAT-CCTCCCCAGCATGCCTGCTATT		Gibson assembly

Supplemental Table 3. List of PCR primers. Shown in 5’ to 3’ orientation.

Tyr-gRNA guide	ATCAGGCAATCATGTAATAA
Rosa26-gRNA guide	ACTCCAGTCTTTCTAGAAGA
Tyr$^{gRNA\text{-}Tomato}$ donor DNA	TTAATGGTGTGTTCATTACGGAAACTCTATTTTCATAAAGGAAAGTAAAGAGATAACTAAACAATCTCCATATTAATTAAATCTATG CATCTCAGTGCTGACTGCACATTAGAACCGCATTAAGAGTCTTAAAAGTATAAACACCCTTGGGAGGCTGAAGGCAGGAGGATTGCT GAGATTTGGAGGCCAGCCTAGGTTGTAGAGTAAGTTCTAAGTTAGTCTAGGGTTACAGTGTTAAACCTCTTCCTCAAATAGACAAAG AACACAGACACTGGAGATTTACAATAGACTCATTGATTCTTTGTTTCCAGTATGTAGTTGGATGATTGATTTTCAAGCAACTACAGA ACAATAGATTGTAGGTGATCATAATATTGACAGTAATTGAATTAAAGCACTGTAAACATGCATTATTTATTAGTTCATTTATTAATT TATTATTTACTGTCTAAGGGTGTTTTGTTGACATGTATGTCTGTATACCATATGTGTGCCTGGCTCTCAGAGTGGTCAAAAGTGGCT ATCAGACCCCCAAGAACTGGAGTTACAAGTGGCTGTAAAATACCATCTGAGTTTTGGGAACTGAACCAAGTTTGTCTGAAAGAATTA ACCAGTGCTCTTAAATTCTGAGCCATCTTTCCAGGCCAGAACAGTCAAATTTGCTTTGTTTCATTTCTGTAAGTCATTATTCACTTT TCTTTTGAAATGGTACATATTCTTTTTCTTCAGCACATATTCACTGAACAGGGTTTCATAAAGCCACATTCATCTACACATATCATC TGATTTGACAATATCCCTTTATTCATGCTCTCTAAACTTCCTCTAATGGGAAATGAGGAGGAGAAAAAAAGTCTTCATTTTAGAAGG ATTTTAACAAGACAGGTTGTGTGAATAAAGAAAACAGACAAGCAGAATTAGCTTATTTGTTATGTCTGCAACTGTTTATCCCTGGAC AGTGCAAAACTTAAGTATATAGAGCAAACTGTATTCATTTGACCTCATTTTAAAATAATGTTGTTTGTTTGCAAAGTTATTGTTGTT TTGTTAATTATCTTTTTTCAAATTTTGCTTGCTTTTCTTAGTGATTTTTTCCCCCTCTACTTTCTAATTTTACATAAAAAGTCCTTT TCCTCCATAGCCCTATAAATGCAATACAAGGTTGTAGGCAATATTTTAGCAGTTATACATCAGGACTTACTACAATAGGCAAGTCTG CTGTGGTTTTATTGCAACTTTTGATAATTTCAATAATCCTTCAGACTTTCTGCATCAAGTAATGCTATATAACAAATCGAATATTGT AACTGAGAGAATTCCACACCTGAGCCTGATAGAACTTAAAGCAATAGCTAGAGAAGTAAAAAGCTAGATGAGAGTTTTGCAGTTCAT GAATTAAGGTAGAATGACTGAAAATATACAGCAATGTTACTTCGCAGCAGAGCCCCATGCTATAAAGTACTCACTAGGTTTAGTAAA CACTTATCACCATTGGCACCAGGAATTAGCAGCTGTAATTTACAGTGGCCCTTTAAAGGTTTCCATTTTGGAAGTGTTAGACCAGCA GAGACAGAGGACATTCATATTTCAAAATTTCAAGTGTGGGGTAGCCTATGTTCTTTTTTCCCCATTTGAACTGGTAATCAGAATAGT TGCAATAGAAAATTTTACAAGATAGGCAGAAAAACATATTTCACATTTGTGTTCATTTAAAGGAAAACAATTTAACATATATTTCTT TCTTCATCTTTTGCTTCCATTCCAGAATGAGGAAAAATAATAGGAGAAAAGAAAATTTAATCTCAAAATAACAATCCCTTAATTCTA TGTCTAACATTTTTAAAGTAGGAAGAAGGCTAATTTTTCTCCATCCTAGAGAAAAAAAGTTAGAAAGTGCCCATTCCTTAAAAAGGA CTGTCTTCTCCCTGAACATAACGTAAGTTAGCCAGACAGCCCTTGTAATCATTAGCAAATATCACCTTTGCTATTTACCAACAATGC TATCAAAAAATAATTTTTAACAAATTAAGTAGTATCTTATGACTGTAAAATAGAAGTCTGTTTATATTAATTGTACAATAAATTCTT TCAATAGGTATTCACATATTCTAAATACACAGCAGGCTTTAACTCTTTTATTACTTAACTGTTTGAGAGTCAGCAACGTTTAAAATT TAAGCAACTGAAGATTTTTTTGTGTGTGTGTAATCATTTTTTCTATCAGGCAATCATGTAAGGCAGCGCGGCTATCGTAGTCTTCGA GAAGACCTGTTTCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTA TTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTT TGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCTGCGACTCTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTT ACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGC TTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAA ACTCATCAATGTATCTTATCATGTCTGGATCTGCGACTCTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCT TTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAAT GGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATC AATGTATCTTATCATGTCTGGATCCCCATCAAGCTGATCCGGAACCCTTAATGGACTCTAGAACGCGTCTCGACATTGATTATTGAC TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCC GCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTG ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGT CAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT CATCGCTATTACCATGGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCTCCCCACCCCCAATTTTGTATTT ATTTATTTTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCG GGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCG GCGGCCCTATAAAAAGCGAAGCGCGCGGCGGGCGGGAGTCGCTGCGCGCTGCCTTCGCCCCGTGCCCCGCTCCGCCGCCGCCTCGCG CCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTT GGTTTAATGACGGCTTGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTGAGGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTC GGGGGGTGCGTGCGTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCTCCGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGC GGGGCTTTGTGCGCTCCGCAGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGGCTGCGAGGGGAACAA AGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGTCGGTCGGGCTGCAACCCCCCCTGCACCCCCCTCC CCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTACGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCG GCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCCGCCCCCGGAGCGCCGGCGG CTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGC GGAGCCGAAATCTGGGAGGCGCCGCCGCACCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGG GGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTCCCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGGGG GGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTCATGCCTTCTTCTTTTT CCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTGGCCACCATGGTGAGCAAGGGCGAGGA GGTCATCAAAGAGTTCATGCGCTTCAAGGTGCGCATGGAGGGCTCCATGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGG CCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCCCA GTTCATGTACGGCTCCAAGGCGTACGTGAAGCACCCCGCCGACATCCCCGATTACAAGAAGCTGTCCTTCCCCGAGGGCTTCAAGTG GGAGCGCGTGATGAACTTCGAGGACGGCGGTCTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCACGCTGATCTACAAGGT GAAGATGCGCGGCACCAACTTCCCCCCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCACCGAGCGCCTGTA CCCCCGCGACGGCGTGCTGAAGGGCGAGATCCACCAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGACCAT CTACATGGCCAAGAAGCCCGTGCAACTGCCCGGCTACTACTACGTGGACACCAAGCTGGACATCACCTCCCACAACGAGGACTACAC CATCGTGGAACAGTACGAGCGCTCCGAGGGCCGCCACCACCTGTTCCTGTACGGCATGGACGAGCTGTACAAGGCCGTGGATCCAAA AAAGAAGAGAAAGGTAGATCCAAAAAAGAAGAGAAAGGTAGATCCAAAAAAGAAGAGAAAGGTACACGGCTAATTGTTTATTGCAGC TTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAA ACTCATCAATGTATCTTATGCGGCCGCGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAG ATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCA GTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGG AAAGGACGAAACACCGGCAGCGCGGCTATCGTGGCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTT GAAAAAGTGGCACCGAGTCGGTGCTTTTTGCTAGCGGCTGGTAAAGGTCATCTCTGATACTTAGTGAATTATCTCAGCACTTTCTCT TGAAGATCCTGAATAATAAGTCTTGGTTTTGGCCAGAGATTTATATTTTAAAGTAGCATACTTATGGATTTTACTGTTAATATTAAT TTTAATATTACTCAGAAACATGAACTAAGATAATCTGGTTTCTCATCCATTCTTAAATAACTGAGTTACTGTTTATTCCAGATACCA TAGGAGGTGCTAACAATACAGTGGTGAACCAATCAGTCCTTGTTTATAGCAGCTTAGAAATCTAACTAAGCTGTTAGATATTAAGAC AAATAAGCACATAGATATTTGTTTAATAACAACTTTGAAGAGTTCTCAAAAGAGAAAAAGAAATCAAGAAGCATTAACATCTGTTAG TAAGGCACATTTTCTGCCCTGAGATATTATCAGAGATGCTCACAGAGCATGAAAATGTGGCTGCTGAAGTACCAGTCTTTGCAACTC TCTCTATATAGTGCATCTTACCTGCCAGGAGGAGAAGAAGGATGCTGGGCTGAGTAAGTTAGGATTTTCAGGGTGACGACCTCCCAA GTACTCATCTGTGCAAATGTCACAGTTTTCTGCATCTCTCCAATCCCAGTATGGAACAGTGAAGTTCTCATCCCCAGTTAGTTCTCG AATTTCTTGTTCCCACAATAACAAGAAAAGTCTGTGCCAAGGCAGAAACCCTGGTGCTTCATGGGCAAAATCAATGTCCCTCCATAT TTCAGAGCCCCCAAGCAGTGTGTCCCTTGACACATAGTAATGCATCCATACAAAGAGGTCGTAGATGTTGATATCATTAAACATGGG TGTTGACCCATTGTTCATTTGGCCATAGGTGCCTGTGGGGATGACATAGACTGAGCTGATAGTATGTTTTGCTAAAGTGAGGTAAGA AAAGAACTTATTCTTTTCGGAGACACTCAAATCAAAAATGTTTCTTCTAATCAAGACTCGCTTCTCTGTACAATTTGGGCCCCCAAA TCCAAACTTACAGTTTCCGCAGTTGAAACCCATGAAGTTGCCTGAGCACTGGCAGGTCCTATTATAAAACACAGAGGGCCAGGACTC ACGGTCATCCACCCCTTTGAAGGGGAACTGAGGTCCAGATGGTGCACTGGACAGAAGGATATCCTGGCAGGAACCTCTGCCTGAAAG CTGGCCGCAGGGACTCCCATCACCCATCCATGGTGGGCAGCATTCTTTTGCCAACAAGTTCTTAGAGGAGGCACAGGCTCGAGGAAA

	ATGGCCATCAGAGATCTGGAAACTCCACAGAAGGCAATACAAAACAGCCAAGAACATTTTCTCCTTTAGATCATACAAAATCTGCAC CAATAGGTTAATGAGTGTCACAGACTTCTTTTCCAGCAACCCCTGGAGTGACTATCACATGTTTTGGCTAAGACCTATATAACCACT CCCCTTCTTACATCAAATACTCTCAGCCTGTTTTACACTAAGCTTTTATCTTCTGCAAAGCACATGACTAACTTTTTCTTGGAGTTT GTACATAGCCCATAGTGAGGTAACTAAATTGAAGGAAGATATATTATTCTAATTGATATGAAATTAATAATAATTGGAATTTGCAAA TGAGAAGTAGTATCAAGAATACATCATCTCTTATGAATAAAGGTCATGAAACATCTAGATTCTGAATACAAGCCTGTTGTATTTAGA GATCATAAGTAGACAATAACATTTATAGAACTCATTTTGTCAAAAATTTTTCTTCAGAACTCAGAATATTGCTTATATCTGGCTAAG CTGAGCTTATTTATTTTATAAAAGTAATCAGTCTAAAATAGTTGCCTCTCATTTTTCCTTGATTATCTTGAATTGATTATTCCAAAG ATATCTTACTTTGTTACTTTGTTATTTTGGTGAACTGCTTCAATGCATAAAGTTTCTAGTTAAGAAAACACC
Tyr$^{gRNA\text{-}Lite}$ donor DNA	AGCAGGCTTTAACTCTTTTATTACTTAACTGTTTGAGAGTCAGCAACGTTTAAAATTTAAGCAACTGAAGATTTTTTTGTGTGTGTG TAATCATTTTTTCTATCAGGCAATCATGTAAGGCAGCGCGGCTATCGTGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTG TCTTCCCAATCCTCCCCCTTGCTGTCCTGCCCCACCCCACCCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCC TCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAGATGGCTGGCAA CTAGAAGGCACAGTCGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATT AATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTA TGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAAC ACCGGCAGCGCGGCTATCGTGGCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCA CCGAGTCGGTGCTTTTTGGCTGGTAAAGGTCATCTCTGATACTTAGTGAATTATCTCAGCACTTTCTCTTGAAGATCCTGAATAATA AGTCTTGGTTTTGGCCAGAGATTTATATTTTAAAGTAGCATACTTATGGATTTTA
Tyr^{Target} donor DNA	GTTTGAGAGTCAGCAACGTTTAAAATTTAAGCAACTGAAGATTTTTTTGTGTGTGTGTAATCATTTTTTCTATCAGGCAATCATGTA AGGCAGCGCGGCTATCGTGGCTGGTAAAGGTCATCTCTGATACTTAGTGAATTATCTCAGCACTTTCTCTTGAAGATCCTGAATAAT AAGTCTTGGTTTTGGCCAGAGATTTA

Supplemental Table 4. DNA sequences for generation of mouse models. Homology arms highlighted in red.

9.3 Additional Methods, Results, and Discussion

9.3.1 Sperm Sequencing

A more in-depth analysis of Cas9 cleavage via NextGen sequencing in germline-homing gene drive mice was thought to be an informative experiment to perform as this approach could definitively show that the gene drive activity is occurring in the germline cells and is not due to a leaky promoter or mRNA carryover. It could also provide a more accurate measure of the rate of error-prone repair in comparison to WT sequence as each individual sequence read provides the same amount of information on a repair event as the analysis on one offspring. The exact indels and their different ratios could further be analysed too.

DNA was extracted from sperm from *Vasa-Cas9-2/+ ; Tyr*$^{gRNA\text{-}Tomato/Target}$ mice as per the DNA extraction methods section 9.2. The *Tyr*^{Target} locus was amplified using primers `5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTTAGCCAGACAGCCCTTGT-3'` and `5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCTCTGGCCAAAACCAAGACT-3'`, and paired end Illumina MiSeq sequencing was performed by AGRF. Resulting data was analysed using CRISPResso2[119] and is shown in Figure 13.

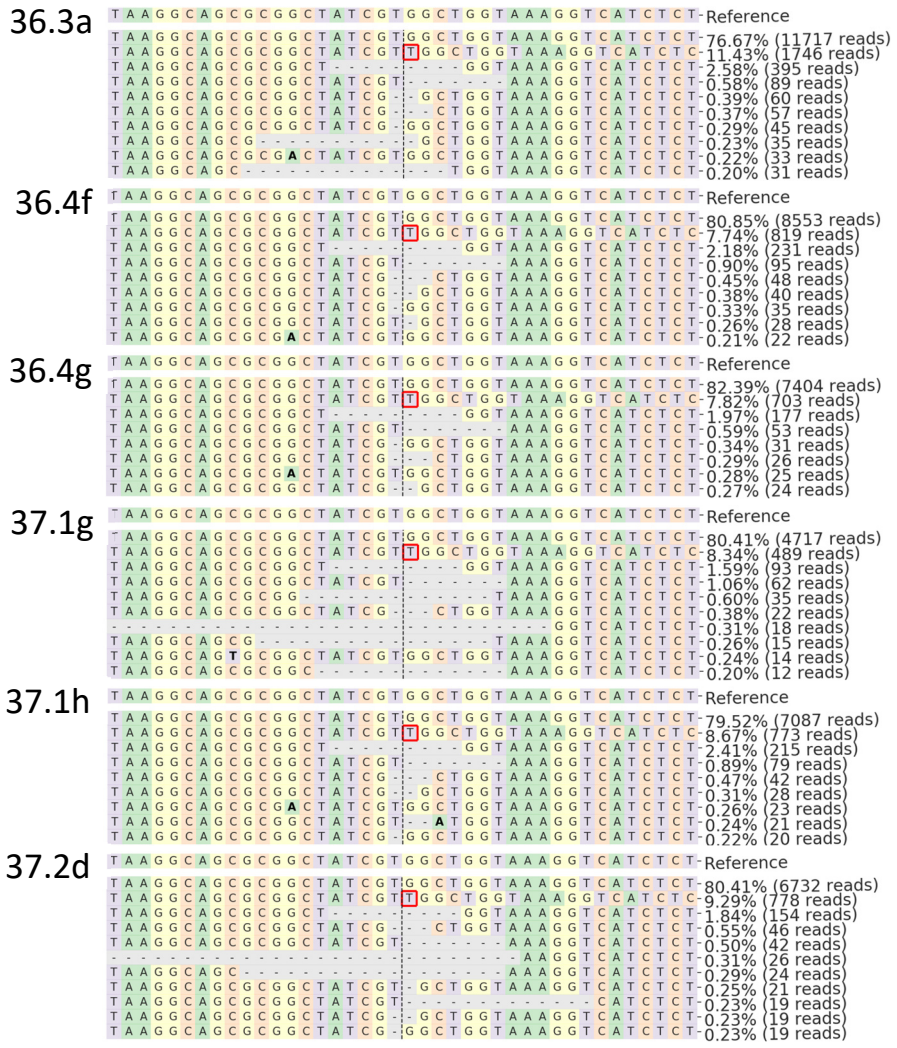


Figure 13. NextGen sequencing data of *Tyr*^{Target} locus in sperm from *Vasa-Cas9-2/+ ; Tyr*$^{gRNA\text{-}Tomato/Target}$ mice. The dotted line indicates the Cas9 cleavage site. Dashes represent deletions, red boxes indicate an insertion, and a bold letter indicates a base change. Data shows reference sequence and number of reads for each allele resulting from error-prone repair. Individual mouse IDs are shown on the left. Images generated with CRISPResso2.[119]

The sequencing data shows an indel formation rate of ~20%, reflecting a similar cut rate (14.5%) to what was determined by genotyping offspring (Supplemental Figure 7A in section 9.2). Considering the rate of indels in the sperm, we can infer that the indels as discussed in the manuscript were from germline activity of the gene drive and not a result of any mRNA carryover. However, as female germline cells were not assayed, we can't definitively conclude there was no maternal mRNA carryover.

This sequencing demonstrates specific indels formed and the corresponding rate of formation are very predictable. A single T insertion is the most common indel in every sample at a rate of 7.74-11.43% and a 9 bp deletion is the second most common in every sample at 1.59-2.58%. A 7 bp deletion is the third most common in 5 of 6 samples at 0.58-1.06%. The rest of the indel alleles occur at very low rates (0.55% and under), which includes 11 indels unique to particular sperm samples but also a further 5 indels that are present in multiple samples.

These data present an interesting opportunity for gene drives in general. Since we know that resistance allele formation is a major hurdle to overcome for the effective spread in wild populations, so any strategies to help reduce the rate of resistance allele formation are going to be extremely beneficial. These data suggest a strategy to reduce that rate, if a gene drive were to be made with multiple, secondary gRNAs within it, each one could target one of the major resistance alleles based on sequencing data, therefore giving the cell another opportunity for HDR instead of following error-prone repair pathways.

9.3.2 *Tyr*$^{gRNA\text{-}Tomato}$ High Penetrance Lethality and *Tyr*$^{gRNA\text{-}Lite}$ Mouse Model

As briefly discussed in the mouse model generation methods in section 9.2 above, during routine breeding of the *Tyr*$^{gRNA\text{-}Tomato}$ mice highly penetrant lethality was observed. Initially this was only seen in homozygotes which all died before weaning so it was assumed to be a homozygous lethal trait. Unfortunately, as breeding continued, spontaneous lethality occurred in hemizygotes too. The earliest death was at P67 and of all mice that reached this age 34% died prematurely. Stressors such as mating or pregnancy tended to increase the death rate. Affected mice were mostly found deceased showing no earlier signs of distress or illness.

The specific cause of death is unknown, considering the subsequently generated *Tyr*$^{gRNA\text{-}Lite}$ mouse lacking the dTomato fluorophore showed no lethality phenotype, we posited that perhaps a high level of dTomato was being produced under the CAG promoter, aggregating, causing complications, and leading to death. Extremely high levels of fluorescence were seen and the levels were so high such that upon autopsy internal organs were visibly redder compared to controls. Another potential cause of this phenotype was that upon insertion of the Tyr$^{gRNA\text{-}Tomato}$ cassette, there was off-target cutting that disrupted a critical gene. However, given that the colony was established by out-crossing resulting in segregation of the founder's chromosomes over several generations, this seems very unlikely.

The *Tyr*$^{gRNA\text{-}Lite}$ mice that were subsequently generated, although not specifically mentioned in the manuscript, contained a bovine growth hormone (BGH)-polyadenylation (polyA) signal in the same orientation as *Tyr*. This was intended to mimic the same effect as the 3x simian virus 40 (SV40) polyA signal in *Tyr*$^{gRNA\text{-}Tomato}$ mice where hemizygotes would have a black coat and homozygotes would have a white coat. However, upon testing this system, the BGH-polyA signal was insufficient to terminate transcription of *Tyr* in homozygotes, having no effect on the coat colour. The white coat phenotype was not a critical aspect to the experiments however, so the experiments continued without this aspect.

10 Cas12a Zygotic and Germline Gene Drives

10.1 Introduction

10.1.1 Cas12a

A gene drive is highly dependent upon what cellular repair mechanisms take priority (HDR vs. error-prone repair),[59] which is influenced by the state of the DNA after cleavage.[4, 28] Cas9 DNA cleavage generates a DSB with two blunt DNA ends,[3] we hypothesised that a gene drive employing a CRISPR nuclease that induced a different kind of DSB may favour the HDR pathway over error-prone repair.

Cas12a (formerly Cpf1)[120] is one such system. In contrast to Cas9, Cas12a produces a staggered cut that is not consistently positioned (Figure 14).[120] On the non-complementary strand it has been shown to be 14 or 16 bp 3' of the PAM and 23-25 bp in the same direction on the complementary strand.[120] Further to this, there are trimming events post-cleavage which can cause the loss of up to 4 bp on the non-complementary strand.[120] All this results in a DSB with 5' overhangs on the end of each strand. This may be useful in promoting HDR or at least reducing NHEJ as the DSB with overhangs will be in a state that potentially favours HDR and MMEJ over NHEJ, the latter of which proceeds with blunt ends.[28] Indeed, subsequent to the design of this project, when testing targeted DNA integration in zebrafish there was shown to be a higher rate of HDR in comparison to Cas9.[121]

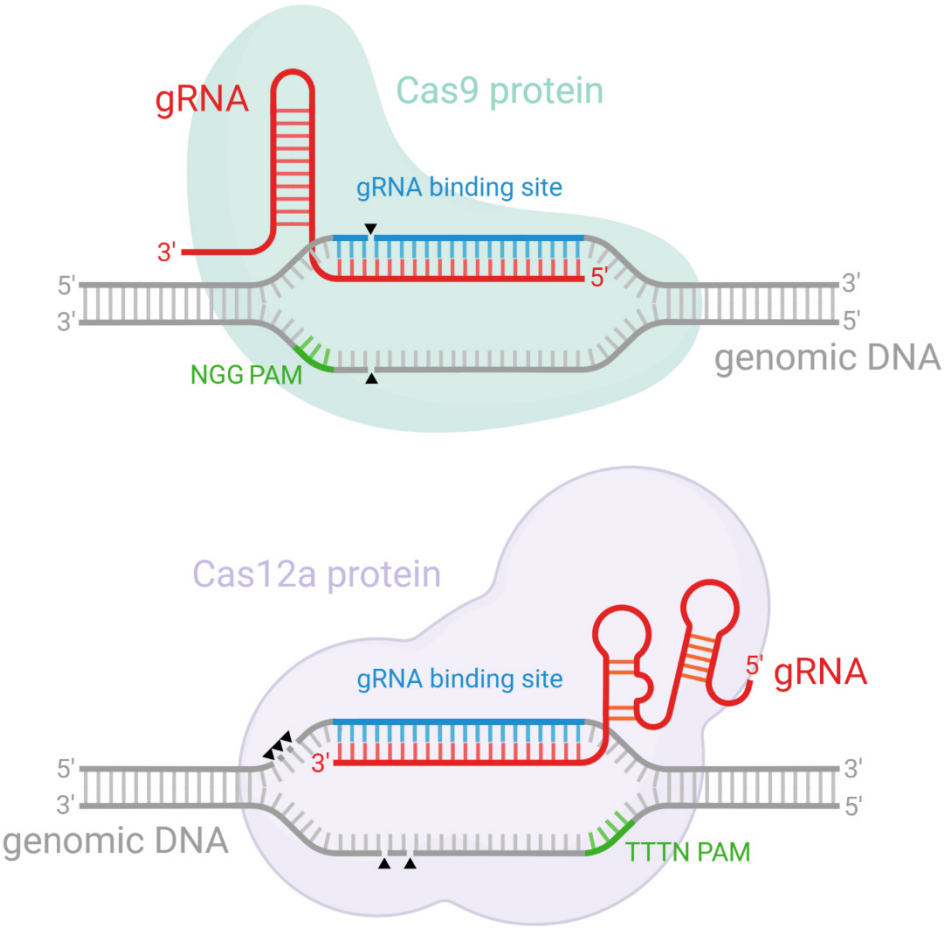


Figure 14. Schematic outlining the differences between Cas9 and Cas12a. Cas12a requires a different gRNA (red),[42] the PAM (green) is at the opposite end and has a different sequence (TTTN),[42] and whilst the Cas9 produces a blunt cut (black triangles),[3] Cas12a produces a staggered cut at inconsistent positions.[120] Image created with BioRender.[8]

Aside from the cleavage method, Cas12a has several significant differences from Cas9. Most importantly is that target sites in the gDNA require a TTTN PAM on the 5' end of the non-complementary strand as opposed to the NGG at the 3' end for Cas9.[42]

Cas12a also uses a different gRNA, the guide region (18-26 nt) is on the 3' end compared to Cas9 where it is on the 5' end.[42] Considering that Cas12a cuts at the PAM-distal region,[42] any indels are far from the PAM where, similar to Cas9 it is less likely to be important for binding.[12-14] This suggests Cas12a would be more likely to rebind and cut that same site again even if there are already indels present from a previous cleavage event. However, the relationship Cas12a has with off-target binding and what MMs do to the binding affinity are less understood than Cas9. The data are contradictory and indicate both that the bases closer to the PAM are more important and also that the bases along the entire gRNA binding site are equally important.[120, 121] If indeed Cas12a was more likely to re-cut an already cut allele with indels present at the PAM-distal region, this would be greatly beneficial to a gene drive. Considering it is desirable to get as few indels as possible, this would allow the gene drive multiple chances at homing via the HDR pathway instead of a single error-prone repair event generating a resistance allele.

10.2 Experimental Design

10.2.1 Aims

The major aim was to develop both zygotic- and germline-homing gene drives utilising Cas12a. These were designed to incorporate both a split drive system and a synthetic target, keeping safety as paramount. The major aim is broken down into sub-aims 1, 2 and 3.

Sub-aim 1 was to design and generate the following mouse models: A mouse expressing Cas12a under a germline promoter, a mouse expressing Cas12a under a zygotic promoter, a mouse expressing a Cas12a gRNA, and a mouse containing a synthetic target.

Sub-aim 2 was to cross the Cas12a gRNA line, synthetic target line, and zygotic Cas12a expression line, then to analyse its genotype for evidence of homing.

Sub-aim 3 was to cross the Cas12a gRNA line, synthetic target line, and germline Cas12a expression line, then to cross that line with WT mice and analyse the inheritance pattern for evidence of homing.

10.2.2 Mouse Models

Figure 15 shows all the required mouse models.

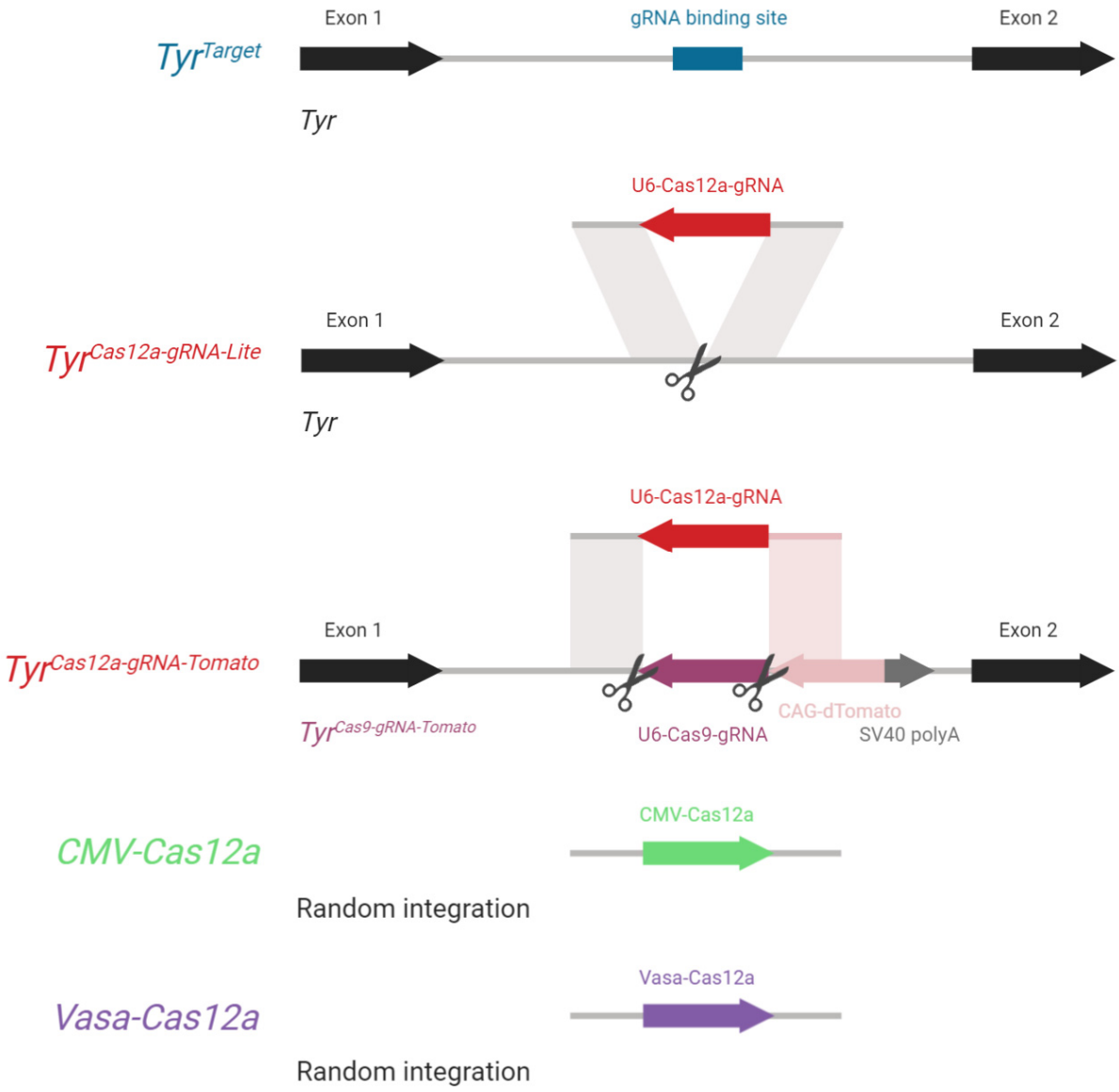


Figure 15. Schematic showing the mouse models needed for the Cas12a gene drives. Scissors show CRISPR cut sites for insertion of custom DNA. Image created with BioRender.[8]

The *Tyr*^{Target} mouse with a synthetic gRNA binding site in intron 1 of the *Tyr* gene which was previously made for the Cas9 gene drives was re-used here. This was possible as a Cas12a `TTTN` PAM site is directly adjacent the Cas9 `NGG` PAM site as shown in Figure 16.

Figure 16. *Tyr* intron 1 with *Tyr*^{Target} insert. Shown are the Cas12a PAM (green) and Cas12a gRNA binding sequence (red) which encompasses the Cas9 PAM (blue). The Cas9 binding sequence (purple) is also shown on the complementary strand.

Tyr$^{Cas12a\text{-}gRNA\text{-}Tomato}$, the Cas12a version of *Tyr*$^{Cas9\text{-}gRNA\text{-}Tomato}$, is a mouse that contained a U6-promoter driven[69] Cas12a gRNA ("Tyr12a-gRNA") that targets *Tyr*^{Target} as shown in Figure 16 and is located in the same genomic locus as *Tyr*^{Target}. It also contained a ubiquitously expressed CAG-promoter driven[122] dTomato gene[123] used as a fluorescent reporter and an SV40 late polyA signal to interrupt

the *Tyr* gene.[124] This was made by cutting out the U6-Cas9-gRNA cassette present in the *Tyr*$^{Cas9\text{-}gRNA\text{-}Tomato}$ mouse and inserting in a U6-Cas12a-gRNA cassette. U6 promotes transcription by Pol III and a poly(T) signal terminates it, in this case TTTTT was used to ensure efficient termination.[125]

Tyr$^{Cas12a\text{-}gRNA\text{-}Lite}$ was in the same locus and contained only the U6-Tyr12a-gRNA cassette from *Tyr*$^{Cas12a\text{-}gRNA\text{-}Tomato}$ without any other elements. This was created as an alternative to *Tyr*$^{Cas12a\text{-}gRNA\text{-}Tomato}$ due to the low penetrance lethality seen in *Tyr*$^{Cas9\text{-}gRNA\text{-}Tomato}$ mice as discussed in section 9.3.2 which emerged likely due to the dTomato protein.

For generating Cas12a lines, a choice needed to be made between AsCas12a and LbCas12a, two variants with similar function and activity. At the time of making this decision, the published literature comparing the two was inconclusive in regards to what would be a better choice: Zhang *et al.* (2017) showed greater percentage of indel formation with LbCas12a in mouse fibroblasts and Toth *et al.* (2016) showed a slightly higher rate of HDR for LbCas12a (24% vs. 15%) in N2a mouse neuroblastoma cells. However, in mouse embryos, *in vivo* data from Kim *et al.* (2016) showed a similar number of indels were generated for each type of Cas12a, varying from 2-80%. As gene drives operate *in vivo*, this data was considered particularly relevant. Combined with data from our own lab, published by Robertson *et al.* (2018) showing that AsCas12a gave a 33% rate of indel formation *in vivo*, AsCas12a was chosen for the gene drives.

For zygotic homing, CMV-Cas12a, a ubiquitously expressed, CMV promoter driven[122] *AsCas12a* gene, randomly integrated into the mouse genome was generated. For germline homing, *Vasa-Cas12a*, the *AsCas12a* gene driven by the germline promoter Vasa,[129] also randomly integrated into the mouse genome was generated.

10.2.3 Experiments

10.2.3.1 Zygotic-homing Gene Drive

Figure 17 shows the design for the zygotic-homing gene drive experiment. We generated *CMV-Cas12a*/*CMV-Cas12a* ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Cas12a\text{-}gRNA\text{-}Lite}$ mice, *Tyr*$^{Target/Target}$ mice, and crossed them together to produce *CMV-Cas12a*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ zygotes. If homing occurred, we would expect a *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite}$ homozygote, if no homing occurred we would expect error-prone repair to produce a resistance allele with a *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target\Delta}$ genotype.

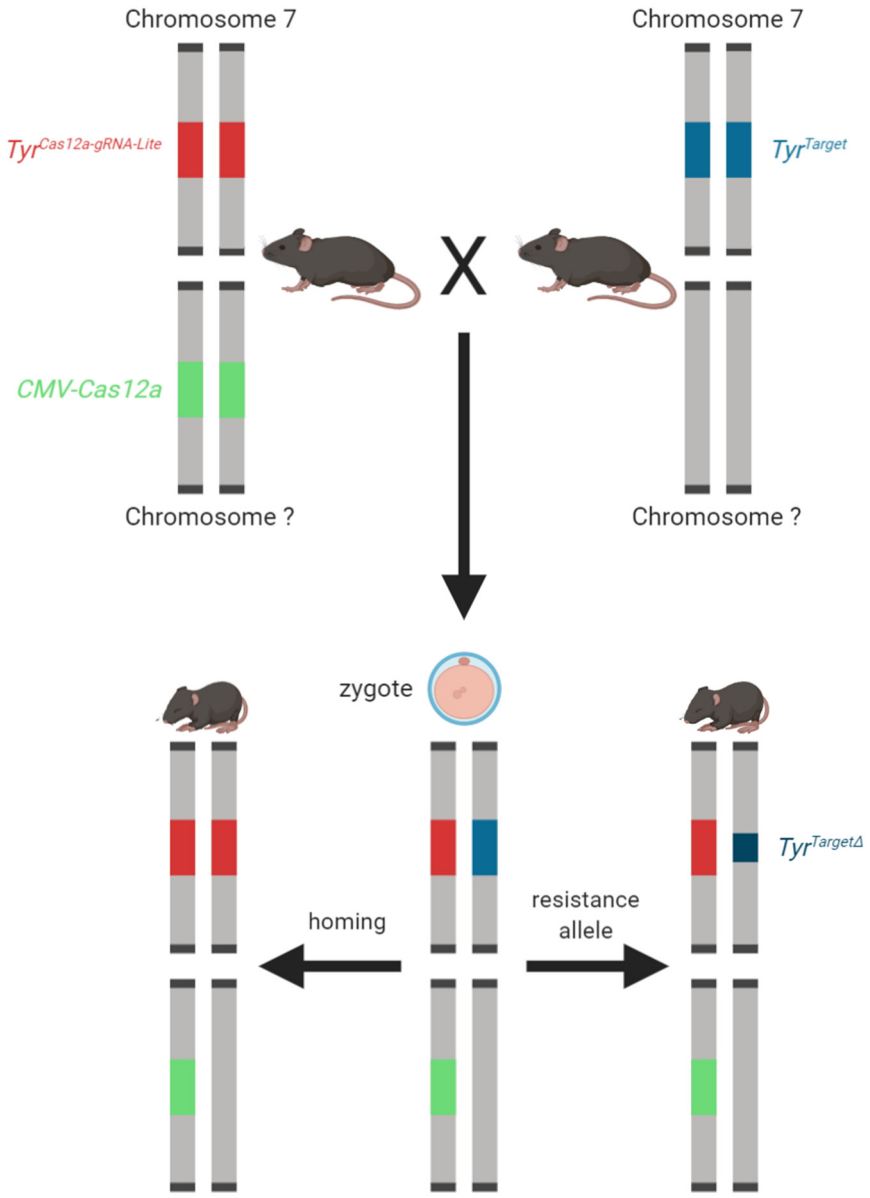


Figure 17. Zygotic-homing Cas12a gene drive experimental design, showing relevant homologous chromosomes. Expected genetic outcomes shown as either homing or resistance allele formation. Image created with BioRender.[8]

10.2.3.2 Germline-homing Gene Drive

Figure 18 shows the experimental setup for the germline-homing gene drive. *Vasa-Cas12a*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice were generated as a first step. As homing rates could potentially vary anywhere from 0 to 100%, the schematic shown illustrates the expected genetic outcomes in those two extreme cases. If homing occurred with 100% efficiency, all sperm or ova in the experimental mice would contain the $Tyr^{Cas12a\text{-}gRNA\text{-}Lite}$ gene. Then, after crossing to a WT mouse, 100% of the produced embryos would be $Tyr^{Cas12\text{-}gRNA\text{-}Lite}$/+. In contrast, if homing did not occur, 50% of the haploid sperm or ova in the experimental mouse would contain the $Tyr^{Cas12a\text{-}gRNA\text{-}Lite}$ gene with the other 50% containing Tyr^{Target}. Then, after crossing to a WT mouse, the embryos would be on average 50% $Tyr^{Cas12\text{-}gRNA\text{-}Lite}$/+ and 50% Tyr^{Target}/+.

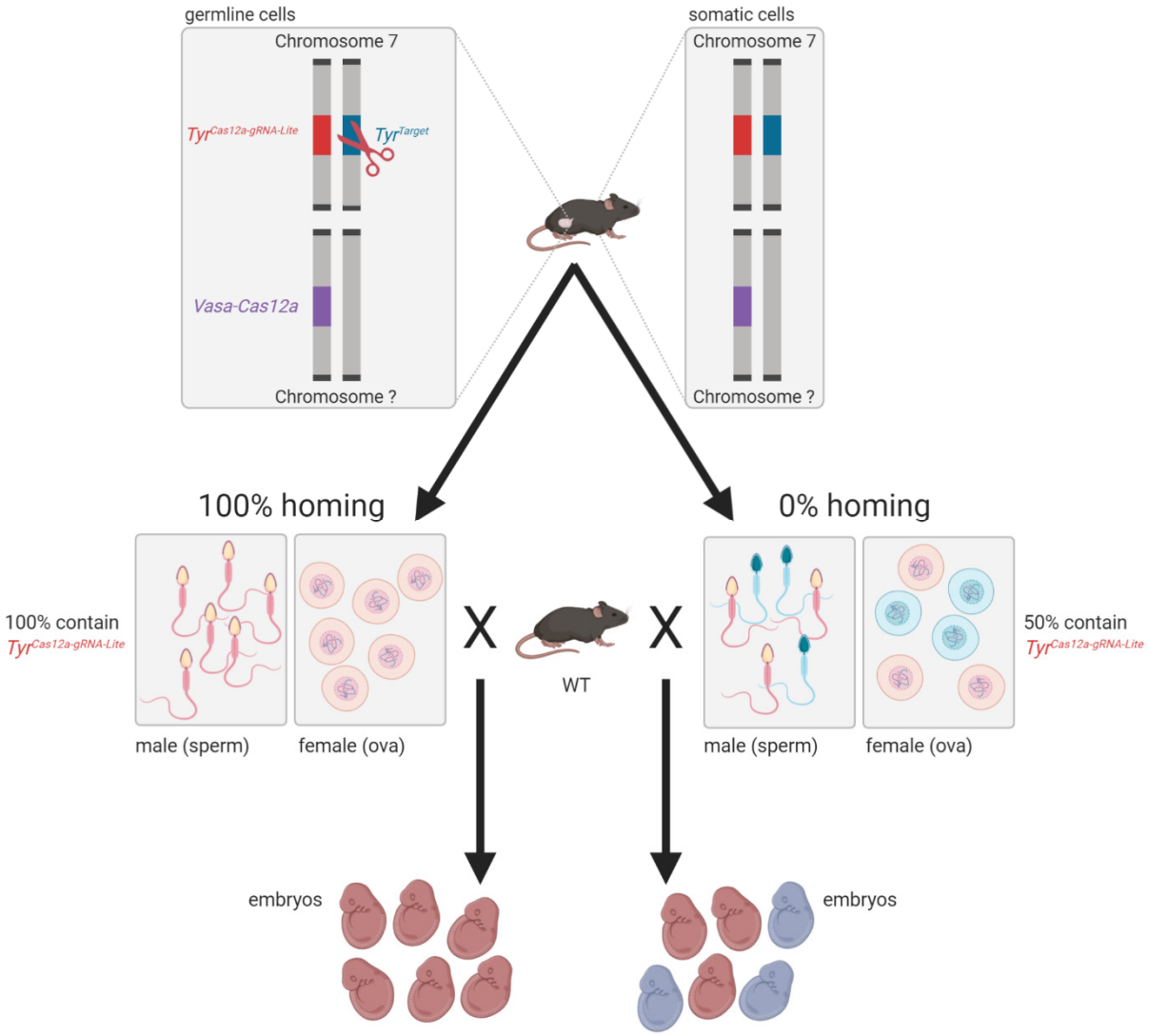


Figure 18. Germline-homing Cas12a gene drive experimental design, showing relevant homologous chromosomes. The schematic shown here illustrates the expected genetic outcomes in the two extreme cases where homing rates are either 100% or 0%. Red sperm and ova (middle boxes) contain the *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite}$ gene. Red embryos (bottom) are *Tyr*$^{Cas12\text{-}gRNA\text{-}Lite}$/+. Blue sperm and ova (middle boxes) contain the *Tyr*^{Target} gene. Blue embryos (bottom) are *Tyr*^{Target}/+. Image created with BioRender.[8]

10.3 Methods

10.3.1 Mouse Model Generation

C57BL/6JSah females were superovulated by injecting 5 IU Folligon® PMSG (Intervet India) followed by 5 IU Chorulon® hCG (Intervet India) 47.5 h later. Superovulated females were mated to male C57BL/6JSah mice overnight.

The following morning presumptive zygotes were collected from oviducts in EmbryoMax® FHM Mouse Embryo Media (Sigma-Aldrich) and hyaluronidase (15 ng/mL) for denudation of cumulus cells. Presumptive zygotes were washed in EmbryoMax® FHM Mouse Embryo Media (Sigma-Aldrich), transferred to EmbryoMax® KSOM Mouse Embryo Media with 1/2 Amino Acids (Sigma-Aldrich) at 37°C in 5% CO_2, and screened for the presence of two pronuclei (indicating fertilisation).

Zygotes were transferred to EmbryoMax® FHM Mouse Embryo Media under Paraffin (#107160, Merck Millipore) and microinjected into a pronucleus with Cas9 protein, donor DNA, and gRNA (as

detailed in subsequent sections) buffered in pH 8 EDTA (0.05 M) and pH 7.5 Tris (0.1 M) filtered at 0.45 µm. Zygotes were then transferred to EmbryoMax® KSOM Mouse Embryo Media with 1/2 Amino Acids under Paraffin at 37°C in 5% CO_2. Either the same day or the following, zygotes were washed in EmbryoMax® FHM Mouse Embryo Media (Sigma-Aldrich) before being transferred unilaterally into the oviducts of pseudopregnant CD1 female mice and allowed to come to term.

10.3.1.1 *Tyr^{Target}*

This mouse had already been generated for the Cas9 gene drive paper (see the design and generation of homing system results in section 9.2).

10.3.1.2 *Tyr$^{Cas12a\text{-}gRNA\text{-}Lite}$*

The Cas9 "Tyr-gRNA" guide targeting *Tyr* intron 1 (5′-ATCAGGCAATCATGTAATAA-3′) was designed using the Zhang lab tool at http://crispr.mit.edu/[14] (now decommissioned). Plasmid containing complete Tyr-gRNA sequence was generated in pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene; 62988)[130] with oligos purchased from IDT (5′-GGCTATCGTGGCGTTTTAGA-3' and 5′-AAACTTATTACATGATTGCCTGATC-3'). Tyr-gRNA dsDNA was amplified with the addition of a T7 promoter using oligos purchased from IDT (5′-TTAATACGACTCACTATAGATCAGGCAATCA-TGTAATAA-3' and 5′-AAAAGCACCGACTCGGTGCC-3'). Tyr-gRNA was generated using HiScribe™ T7 Quick High Yield RNA Synthesis Kit (NEB) and purified using RNeasy Mini Kit (Qiagen).

U6-Cas12a-gRNA ssDNA described in Figure 15 for the *Tyr$^{Cas12a\text{-}gRNA\text{-}Lite}$* line purchased from IDT as a Megamer®:

```
  1 TAAATCTCTG GCCAAAACCA AGACTTATTA TTCAGGATCT TCAAGAGAAA GTGCTGAGAT
 61 AATTCACTAA GTATCAGAGA TGACCTTTAA AAAAAGGCAG CGCGGCTATC GTGGCTGGAT
121 CTACAAGAGT AGAAATTACG GTGTTTCGTC CTTTCCACAA GATATATAAA GCCAAGAAAT
181 CGAAATACTT TCAAGTTACG GTAAGCATAT GATAGTCCAT TTTAAAACAT AATTTTAAAA
241 CTGCAAACTA CCCAAGAAAT TATTACTTTC TACGTCACGT ATTTTGTACT AATATCTTTG
301 TGTTTACAGT CAAATTAATT CCAATTATCT CTCTAACAGC CTTGTATCGT ATATGCAAAT
361 ATGAAGGAAT CATGGGAAAT AGGCCCTCTT ACATGATTGC CTGATAGAAA AAATGATTAC
421 ACACACACAA AAAAATCTTC AGTTGCTTAA ATTTTAAACG TTGCTGACTC TCAAAC
```

0.75 µL PNA Bio SpCas9 protein (1 µg/µL) and 0.75 µL Tyr-gRNA (500 ng/µL) were incubated 10 min on ice, then mixed with U6-Cas12a-gRNA ssDNA (final concentration 10 ng/µL) to a total volume of 15 µL with buffer as detailed in section 10.3.1. This mixture was injected into zygotes as detailed in section 10.3.1.

10.3.1.3 *Tyr$^{Cas12a\text{-}gRNA\text{-}Tomato}$*

The Cas9 "U6-gRNA1" and "U6-gRNA2" guides targeting the U6-Cas9-gRNA cassette in the *Tyr$^{Cas9\text{-}gRNA\text{-}Tomato}$* mouse (5′-AGGACGAAACACCGGCAGCG-3′ and 5′-CGGTGCTTTTTGCTAGCG-GC-3′) were designed using the GT-Scan tool at https://gt-scan.csiro.au/gt-scan.[16] Two separate plasmids containing complete U6-gRNA sequences were generated in pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene; 62988)[130] with oligos purchased from Sigma-Aldrich (5′-ACCGAGGACGAAACACC-GGCAGCG-3' paired with 5′-AAACCGCTGCCGGTGTTTCGTCCTC-3' and 5′-ACCGCGGTGCT-TTTTGCTAGCGGC-3' paired with 5′-AAACGCCGCTAGCAAAAAGCACCGC-3' for the two respective gRNAs/plasmids). U6-gRNA1 and U6-gRNA2 dsDNA were amplified with the addition of a T7 promoter using oligos purchased from Sigma-Aldrich/IDT (5′-TTAATACGACTCACTATAGAGG-ACGAAACACCGGCAGCG-3' paired with 5′-AAAAGCACCGACTCGGTGCC-3' and 5′-TTAAT-

ACGACTCACTATAGCGGTGCTTTTTGCTAGCGGC-3' paired with 5'-AAAAGCACCGACTCGGTG-CC-3' for the two respective gRNAs). U6-gRNAs were generated using HiScribe™ T7 Quick High Yield RNA Synthesis Kit (NEB) and purified using RNeasy Mini Kit (Qiagen).

U6-Cas12a-gRNA ssDNA described in Figure 15 for the *Tyr*$^{Cas12a\text{-}gRNA\text{-}Tomato}$ line purchased from IDT as a Megamer®:

```
  1 CCAAAACCAA GACTTATTAT TCAGGATCTT CAAGAGAAAG TGCTGAGATA ATTCACTAAG
 61 TATCAGAGAT GACCTTTACC AGCCAAAAAA GGCAGCGCGG CTATCGTGGC TGGATCTACA
121 AGAGTAGAAA TTACGGTGTT TCGTCCTTTC CACAAGATAT ATAAAGCCAA GAAATCGAAA
181 TACTTTCAAG TTACGGTAAG CATATGATAG TCCATTTTAA AACA
```

0.75 µL PNA Bio SpCas9 protein (1 µg/µL) and 0.75 µL U6-gRNA1 (500 ng/µL) were incubated 10 min on ice. 0.75 µL PNA Bio SpCas9 protein (1 µg/µL) and 0.75 µL U6-gRNA2 (500 ng/µL) were incubated 10 min on ice. These were mixed with U6-Cas12a-gRNA ssDNA (final concentration 10 ng/µL) to a total volume of 15 µL with buffer as detailed in section 10.3.1. This mixture was injected into a mix of *Tyr*$^{Cas9\text{-}gRNA\text{-}Tomato}$/+ and WT C57BL/6J zygotes as detailed in section 10.3.1. The mix of zygote genotypes was due to *Tyr*$^{Cas9\text{-}gRNA\text{-}Tomato}$/+ being used as one parent because of the high rate of spontaneous deaths of *Tyr*$^{Cas9\text{-}gRNA\text{-}Tomato}$ homozygotes due to unknown reasons as detailed in section 9.3.2.

10.3.1.4 *CMV-Cas12a*

Although this was primarily an attempt at random integration, a gRNA targeting Rosa26 ("Rosa26-gRNA") was also used to promote integration at that locus and potentially encourage activation of DNA repair machinery to facilitate more integrations. Rosa26-gRNA design was previously published (5'-ACTCCAGTCTTTCTAGAAGA-3').[131] Plasmid containing complete Rosa26-gRNA sequence was generated in pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene; 62988)[130] with oligos purchased from Sigma-Aldrich (5'-CACCGACTCCAGTCTTTCTAGAAGA-3' and 5'-AAACTCTTCTAGAAAGACT-GGAGTC-3'). Rosa26-gRNA dsDNA was amplified with the addition of a T7 promoter using oligos purchased from Sigma-Aldrich/IDT (5'-TTAATACGACTCACTATAGACTCCAGTCTTTCTAGAAG-A-3' and 5'-AAAAGCACCGACTCGGTGCC-3'). Rosa26-gRNA was generated using HiScribe™ T7 Quick High Yield RNA Synthesis Kit (NEB) and purified using RNeasy Mini Kit (Qiagen).

The CMV-Cas12a construct in its entirety was already present in pY010 (pcDNA3.1-hAsCpf1) (Addgene; 69982)[42], this was digested with MfeI and DraIII to isolate CMV-AsCas12a-BGH(polyA) as per Figure 19. Construct was purified with Gel DNA Recovery Kit (Zymoclean).

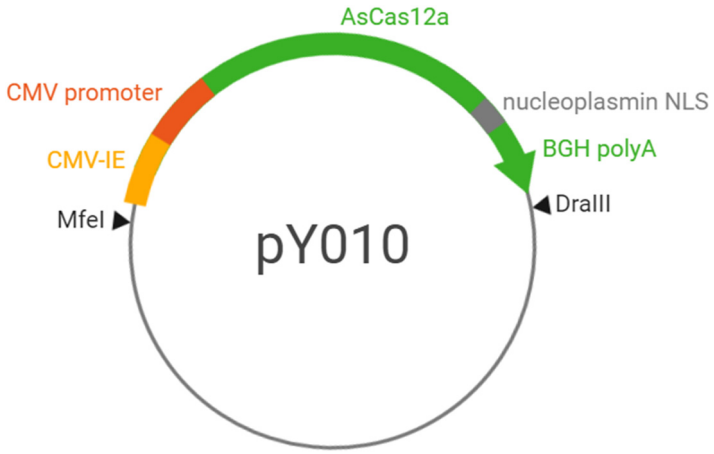


Figure 19. Generation of the *CMV-Cas12a* construct. pY010 plasmid[42], black triangles indicate restriction sites. Image created with BioRender.[8]

0.75 µL PNA Bio SpCas9 protein (1 µg/µL) and 0.75 µL Rosa26-gRNA (500 ng/µL) were incubated 10 min on ice, then mixed with CMV-Cas12a dsDNA (final concentration 3 ng/µL) to a total volume of 15 µL with buffer as detailed in section 10.3.1. This mixture was injected into zygotes as detailed in section 10.3.1.

10.3.1.5 Vasa-Cas12a

Gibson assembly fragments were constructed as follows, shown in Figure 20. pStart-K (Addgene; 20346)[132], a small, low-copy number plasmid functioning as the backbone was digested with EcoRI. The Vasa-β-globin-II fragment was amplified from pVasa-Cre (Addgene; 15885)[129] using primers `5'-GCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCGTGTGCCACCATGCCTGG-3'` and `5'-AAGCCCTCGAACTGTGTCATGGAGCTGTAGGAAAAAGAAGAAGG-3'`. The AsCas12a-BGH(polyA) fragment was amplified from pY010 (pcDNA3.1-hAsCpf1) (Addgene; 69982)[42] using primers `5'-CTTCTTTTTCCTACAGCTCCATGACACAGTTCGAGGGCTTTAC-3'` and `5'-TACAAG-AAAGCTGGGTCTAGATATCTCGAGTGCGGCCGCGGTCGACCGCCTCAGAAGCCATAGAGC-3'`, the latter included the addition of a SalI restriction site. After Gibson assembly and transformation in MAX Efficiency™ DH5α™ Competent Cells (Invitrogen), pSK-VCas12a was expanded with PureLink™ HiPure Plasmid Midiprep Kit (Invitrogen), digested with SalI, and purified using Gel DNA Recovery Kit (Zymoclean).

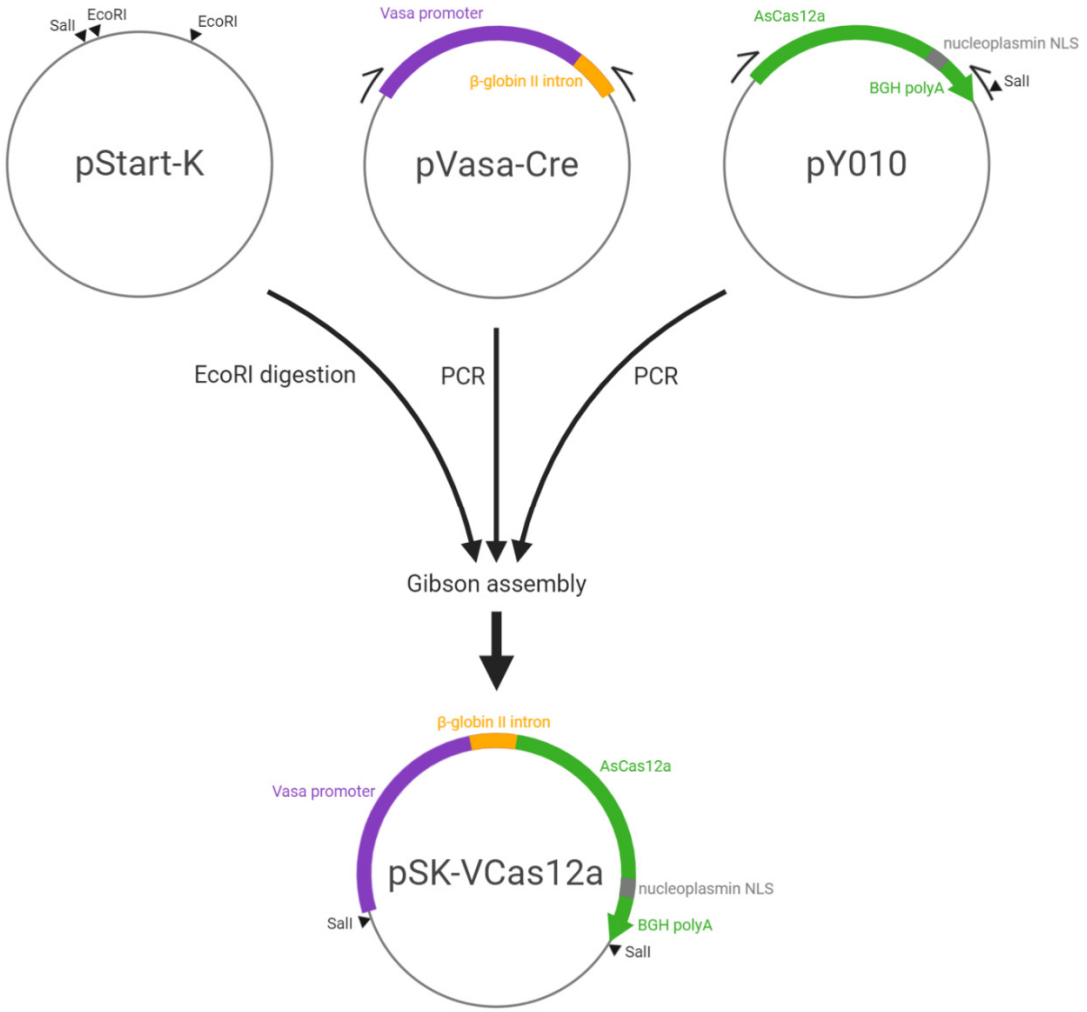


Figure 20. Generation of the *Vasa-Cas12a* construct. Black triangles indicate restriction sites. Half arrows indicate PCR primers. Image created with BioRender.[8]

0.75 µL PNA Bio SpCas9 protein (1 µg/µL) and 0.75 µL Rosa26-gRNA (500 ng/µL) were incubated 10 min on ice, then mixed with Vasa-Cas12a dsDNA (final concentration 3 ng/µL) to a total volume of 15 µL with buffer as detailed in section 10.3.1. This mixture was injected into zygotes as detailed in section 10.3.1.

10.3.2 *In Vivo* Cleavage Assessment

The Cas12a guide Tyr12a-gRNA targeting Tyr^{Target} (`5'-CCAGCCACGATAGCCGCGCTGCCT-3'`) was analysed using the GT-Scan tool at https://gt-scan.csiro.au/gt-scan.[16] Plasmid containing CMV-AsCas12a and complete Tyr12a-gRNA sequence was generated in pY026 (Addgene; 84741)[133] with oligos purchased from Sigma-Aldrich (`5'-AGATCCAGCCACGATAGCCGCGCTGCCTT-3'` and `5'-AAAAAAGGCAGCGCGGCTATCGTGGCTGG-3'`), and expanded with PureLink™ HiPure Plasmid Midiprep Kit (Invitrogen).

The same process to prepare for microinjection of zygotes as described in 10.3.1 was performed except using $Tyr^{Target/Target}$ mice instead of C57BL/6JSah and the injection mix contained only $pY026^{Tyr12a\text{-}gRNA}$ plasmid (100 ng/µL) in the same buffer.

10.3.3 Sample Collection

For colony genotyping, tissue was collected via ear notching live mice. For zygotic-homing gene drive mouse genotyping and genotyping the offspring of female germline-homing gene drive mice, pups

were culled at ~P0 and tail tissue was collected for DNA extraction. For genotyping the fetal offspring of male germline-homing gene drive mice and fetuses for the *in vivo* cleavage assay (section 10.3.2), pregnant females were culled at ~E12.5 and tail tissue was collected from fetuses.

For genotyping the blastocyst offspring of female germline-homing gene drive mice, females were superovulated by injecting 5 IU Folligon® PMSG (Intervet India) followed by 5 IU Chorulon® hCG (Intervet India) 47.5 h later at ~5 PM. The following morning, oviducts were collected from females into M2 medium (#M7167, Sigma-Aldrich) and kept at 4 °C. Cauda epididymides were collected from a C57BL/6J male into M2 medium and kept at 4 °C. IVF was performed as per the Online Manual for CARD Mouse Reproductive Technology sections "*In Vitro* Fertilization using Epididymal Sperm Transported at Cold Temperature"[134] and "*In Vitro* Fertilization (IVF)"[135]. Embryos were cultured in mHTF (#KYD-008-02-EX-X5, United Bioresearch) under Paraffin (#107160, Merck Millipore) at 37°C in 5% CO_2/5% O_2 overnight. All 2-cell embryos were transferred to fresh culture media in the same conditions and incubated for 72-84 h. Embryos that developed a blastocoel were washed in M2 medium and transferred with 1 µL M2 medium into 9 µL MQ H_2O.

For mRNA transcription data, whole testes, ovaries, and spleens were collected from adult mice.

10.3.4 Sample Extraction

For ear notch/tail tissue DNA extraction, gDNA was extracted from samples using High Pure PCR Template Preparation Kit (Roche), KAPA Express Extract kit (Roche), or MyTaq™ Extract-PCR Kit (Bioline).

For blastocyst DNA extraction, a blastocyst lysis buffer base was made with 2 mL pH 8.3 Tris-HCl (1 M), 2 mL KCl (1 M), 40 µL 2% gelatin, 90 µL Polysorbate 20, and 5.87 mL MQ H_2O, stored at RT. For each extraction, blastocyst lysis buffer was made with 125 µL blastocyst lysis buffer base, 2 µL tRNA from baker's yeast (Sigma-Aldrich), and 7.75 µL Proteinase K (20 mg/mL) (Thermo Scientific). 10 µL of blastocyst lysis buffer was then added to blastocysts before heating to 56 °C for 10 min, then inactivated at 95 °C for 10 min.

For RNA extraction, acid guanidinium thiocyanate-phenol-chloroform RNA extraction was performed on testes, ovaries, and spleens with clean-up performed using RNeasy Mini/Micro kit (Qiagen) in conjunction with RNase-Free DNase Set (Qiagen).

10.3.5 Genotyping

Tyr^{Target} and *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite}$ mice were genotyped using primers `5'-CCAGACAGCCCTTGTAATCAT-TAGC-3'` and `5'-TCTCTGGCCAAAACCAAGACTT-3'`, giving a 376 bp band for WT, 399 bp for *Tyr*^{Target}, and 675 bp *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite}$.

Tyr^{Target}/+ mice were screened for *Tyr*^{Target} cleavage using a modified restriction fragment length polymorphism (RFLP) protocol as follows. A 20 µL PCR around the region was performed with primers `5'-ATACACAGCAGGCTTTAACTCTTTT-3'` and `5'-GGATCTTCAAGAGAAAGTGCTGAG-A-3'`, giving a 202 bp band for *Tyr*^{Target} and a 179 bp band for WT. 10 µL was added to a 20 µL T7 Endonuclease I (NEB) digestion for 15 min, 10 µL of that reaction was added to a 20 µL TseI (NEB) digestion for 15 min, loaded onto a 2.5% agarose gel, and ran at 100 V for 90 min. Due to the very small difference in size between *Tyr*^{Target} PCR products and WT PCR products, a complicated mix of heteroduplexes formed during amplification. The added step of including an endonuclease digested all the heteroduplexes to give a cleaner band pattern before restriction digest. The TseI binding site covered the Cas12a cleavage position so that only uncut *Tyr*^{Target} alleles would be digested by it.

Some *Tyr*^{Target}/+ mice were screened for *Tyr*^{Target} cleavage using a combination of Sanger sequencing (with primers 5′-CCAGACAGCCCTTGTAATCATTAGC-3′ and 5′-TCTCTGGCCAAAACCAAGAC-TT-3′) and Tracking of Indels by DEcomposition (TIDE)[136]/Inference of CRISPR Edits (ICE)[137] analysis. These tools require a WT Sanger sequencing trace to use as a background for comparison. As the *Tyr*^{Target} and WT alleles only differ by 23 bp, they are hard to separate on a gel for extraction and independent sequencing. Instead, they were sequenced together, and *Tyr*$^{Target/Target}$ DNA was used as background which results in a "false" 23 bp deletion showing up for all TIDE/ICE analyses.

Tyr$^{Cas12a\text{-}gRNA\text{-}Tomato}$ mice were sequenced using primers 5′-ATGCAGAAGAAGACCATGGGCT-3′ and 5′-GCACCTCCTATGGTATCTGGAA-3'.

Routine genotyping for *CMV-Cas12a* mice was performed using primers 5′-AAGAATCACGAGAGC-CGCAA-3′, 5′- GATTGGAGATGCCGTTCTGC-3′, 5′-AAGGGAGCTGCAGTGGAGTA-3′, and 5′-CCGAAAATCTGTGGGAAGTC-3′, giving a 297 bp band for WT and *CMV-Cas12a* and 622 bp for *CMV-Cas12a*. F_0 genotyping was performed using primers 5′-TGTACGGGCCAGATATACGC-3′ and 5′-GCCCTCGAACTGTGTCATGG-3′, giving a 735 bp band for *CMV-Cas12a* and no band for WT. qPCR was performed to determine zygosity using primers 5′-AAGAATCACGAGAGCCGCAA-3′/ 5′-AGGATGAAGTCGCCGGTTTT-3′ for *CMV-Cas12a*, and 5′-AGGATGAAGTCGCCGGTTTT-3′/ 5′-CCTCTCAGACGGTGGAGTTATATT-3′ for *sox1* as a reference gene.

CMV-Cas12a mice were sequenced using the primer pairs in Table 1.

5′-TGTACGGGCCAGATATACGC-3′	5′-GCCCTCGAACTGTGTCATGG-3′
5′-TGGCACCAAAATCAACGGGA-3′	5′-TGTGCTGATATCCTCGGCG-3′
5′-GAGACACGCCGAGATCTACAA-3′	5′-CAGTGTATCCCAGTGGTCGC-3′
5′-GGAGTTTAAGAGCGACGAGG-3′	5′-CGAAGCTCAGGGCCTTATACC-3′
5′-ACAAGGCCAGAAATTATGCCACC-3′	5′-CGGTCCAATACAGTGTGTGCAG-3′
5′-ACCAAGACAACCTCTATCGATCT-3′	5′-TGTTGTCCAGCTTCTTCTGGTA-3′
5′-ACTATCAGGCCGCCAATTCC-3′	5′-AAGTGCAGGATGAAGTCGCC-3′
5′-CCAGCTGACAGACCAGTTCA-3′	5′-GATTGGAGATGCCGTTCTGC-3′
5′-TGAATGGCGTGTGCTTCGAC-3′	5′-GTGGGGATACCCCCTAGAGC-3′

Table 1. Paired primers for sequencing *CMV-Cas12a*.

Vasa-Cas12a mice were genotyped using primers 5′-GCACGTGCAGCCGTTTAAG-3′, 5′- CTGG-CGTTGATGGGGTTTTC-3′, 5′-GGCTGATCCGTGTGGAGTAT-3′, and 5′-AGGGCCACAACAGT-AAATGG-3′, giving a 647 bp band for WT and *Vasa-Cas12a* and 719 bp for *Vasa-Cas12a*. F_0 genotyping was performed using the first 2 primers which only gave the 719 bp band. qPCR was performed to determine zygosity using the same primer set as for *CMV-Cas12a* above.

Vasa-Cas12a mice were sequenced using the primer pairs in Table 2.

5′-GCACGTGCAGCCGTTTAAG-3′	5′-GCCCTCGAACTGTGTCATGG-3′
5′-CTGCATATAAATTCTGGCTGCGT-3′	5′-TGTGCTGATATCCTCGGCG-3′
5′-GAGACACGCCGAGATCTACAA-3′	5′-CAGTGTATCCCAGTGGTCGC-3′
5′-GGAGTTTAAGAGCGACGAGG-3′	5′-CGAAGCTCAGGGCCTTATACC-3′
5′-ACAAGGCCAGAAATTATGCCACC-3′	5′-CGGTCCAATACAGTGTGTGCAG-3′
5′-ACCAAGACAACCTCTATCGATCT-3′	5′-TGTTGTCCAGCTTCTTCTGGTA-3′
5′-ACTATCAGGCCGCCAATTCC-3′	5′-AAGTGCAGGATGAAGTCGCC-3′
5′-CCAGCTGACAGACCAGTTCA-3′	5′-GATTGGAGATGCCGTTCTGC-3′
5′-TGAATGGCGTGTGCTTCGAC-3′	5′-CCGCCTCAGAAGCCATAGAG-3′

Table 2. Paired primers for sequencing *Vasa-Cas12a*.

The Rosa26-gRNA cut site was sequenced using primers 5′-AAGGGAGCTGCAGTGGAGTA-3′ and 5′-CCGAAAATCTGTGGGAAGTC-3′.

10.3.6 RT-qPCR

cDNA was generated using High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems).

mRNA levels were quantitated using primers `5'-AAGAATCACGAGAGCCGCAA-3'`/`5'-AGGATGA-AGTCGCCGGTTTT-3'` targeting *AsCas12a*, and primers `5'-TGATGGCACTGGCCCCAACAT-3'`/ `5'-GCGCCCTCCTTAGTAGCCCAC-3'` targeting the reference gene eukaryotic elongation factor 2 (*eEF2*).

10.4 RESULTS

10.4.1 Mouse Model Generation

10.4.1.1 *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite}$

Of 39 pups born from CRISPR microinjection, 2 contained the complete insert in *Tyr* intron 1 (Figure 21). Sanger sequencing of the entire insert revealed one mouse had a MM and the second had a dual base call. Subsequent breeding of the latter showed only the correct base was passed on.

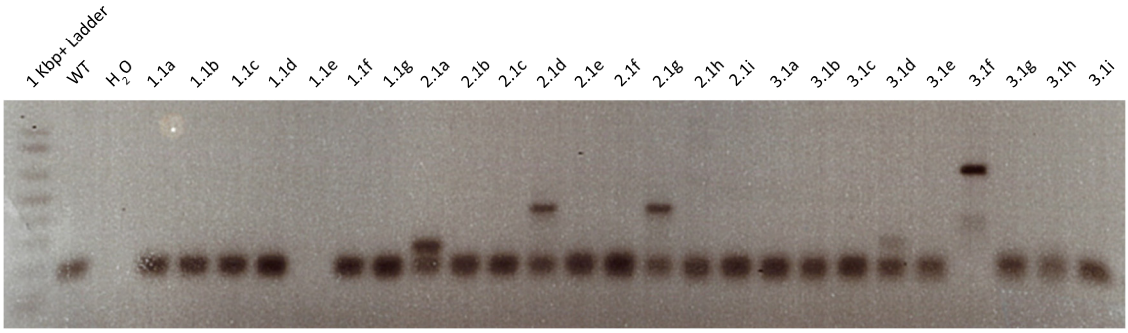


Figure 21. Representative gel showing genotyping of *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite}$ F_0 mice. Samples 2.1d and 2.1g show a band matching the expected size.

10.4.1.2 *Tyr*$^{Cas12a\text{-}gRNA\text{-}Tomato}$

3 of 4 pups born from microinjection were positive for TxRed fluorescence in ear notches, confirming the *Tyr*$^{Cas9\text{-}gRNA\text{-}Tomato}$ allele had been inherited from the heterozygous parent. However, Sanger sequencing revealed that no insertions were detected and no CRISPR cleavage had taken place. Only one microinjection session was performed to attempt generation of this line before it was decided not to proceed. The number of spontaneous deaths of both homozygotes and hemizygotes of the *Tyr*$^{Cas9\text{-}gRNA\text{-}Tomato}$ line was increasing (as per section 9.3.2) and the reason was not known. One possible explanation was a high rate of production of dTomato protein with associated toxicity. Due to this risk it was decided to complete the Cas12a experiments without the aid of a fluorescent marker and to use the *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite}$ line exclusively.

10.4.1.3 *CMV-Cas12a*

Of 15 pups born from CRISPR microinjection, genotyping revealed that 3 contained the insert (Figure 22). Although PCR and Sanger sequencing of the Rosa26 target site showed the formation of indels, the CMV-Cas12a construct was not inserted there and all were in random locations. The complete sequence for all 3 founders was confirmed by Sanger sequencing.

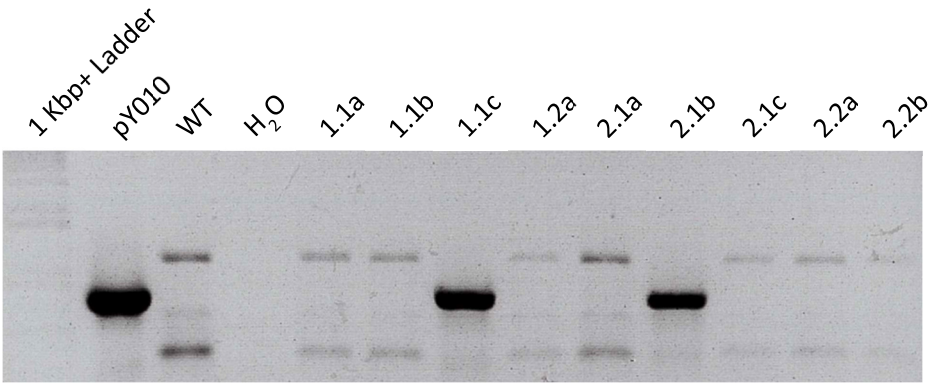


Figure 22. Representative gel showing genotyping of *CMV-Cas12a* F_0 mice. The plasmid pY010 contained the full CMV-Cas12a construct and was used as a positive control. Samples 1.1c and 2.1b show a clear band matching the positive control.

One of the F_0 *CMV-Cas12a* mice failed to transmit the gene after screening 39 pups, indicating it was a low-level mosaic and likely no germ cells contained the CMV-Cas12a construct. Another F_0 female was unable to be fully characterised as it died due to birthing issues with its first litter.

The third F_0 *CMV-Cas12a* mouse successfully transmitted and was determined to have integrated on the X-chromosome. This was confirmed by 2 pieces of data: No males that could potentially be homozygous were homozygous (n=72) and the hemizygous offspring of male *CMV-Cas12a* crossed to WT females were always female (n=30).

Cas12a mRNA levels were characterised in the testes, ovaries, and spleen by RT-qPCR (Figure 23). Significant background amplification could be seen, likely primer-dimer formation as can be inferred from the WT melt curves from all tissues which exhibit a lower peak in comparison to the *CMV-Cas12a* testis tissue melt curve where the highest level of mRNA can be seen. Very low levels of *Cas12a* mRNA can be seen in the ovary and both male and female spleens as evidenced by their melt curves which all show peaks for both primer-dimers and the true RT-qPCR product.

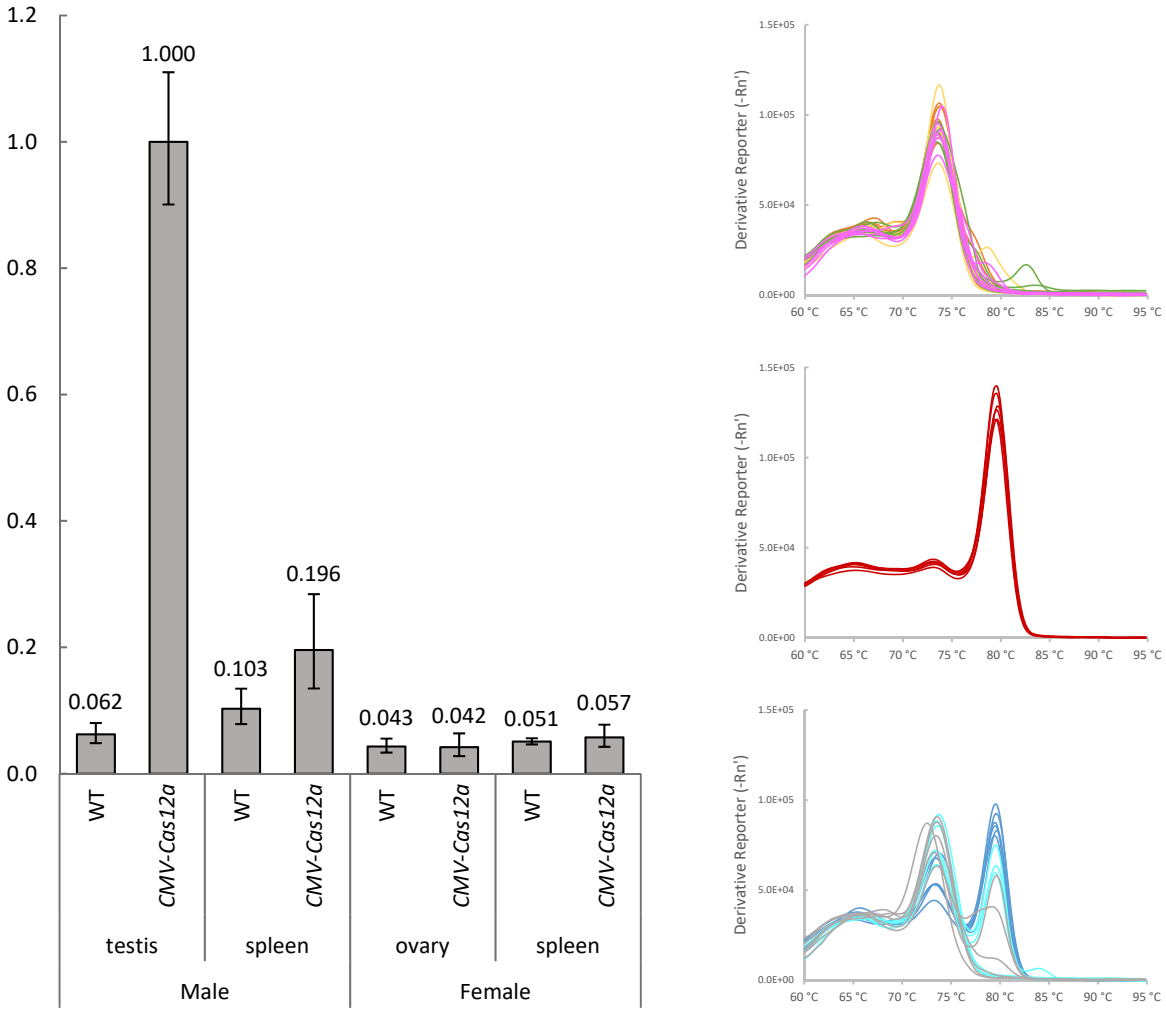


Figure 23. Characterisation of Cas12a mRNA levels in *CMV-Cas12a* mice by RT-qPCR. Left: Relative quantification of mRNA (with 95% CI). Top right: WT melt curves from all tissues. Middle right: *CMV-Cas12a* testis tissue melt curves. Bottom right: *CMV-Cas12a* ovary and both male and female spleen melt curves.

10.4.1.4 Vasa-Cas12a

Of 23 pups born from CRISPR microinjection, 7 contained the insert (4 male/3 female, Figure 24). The Rosa26 target site again only showed indels, so all were in random locations. The β-globin-II, Cas12a, and BGH polyA sequences were verified by Sanger sequencing in the males and colonies were established for each (*Vasa-Cas12a-A*, *Vasa-Cas12a-B*, *Vasa-Cas12a-C*, and *Vasa-Cas12a-D*). The Vasa sequence was not verified as the original Vasa-Cre plasmid had not been sequenced since it was created.

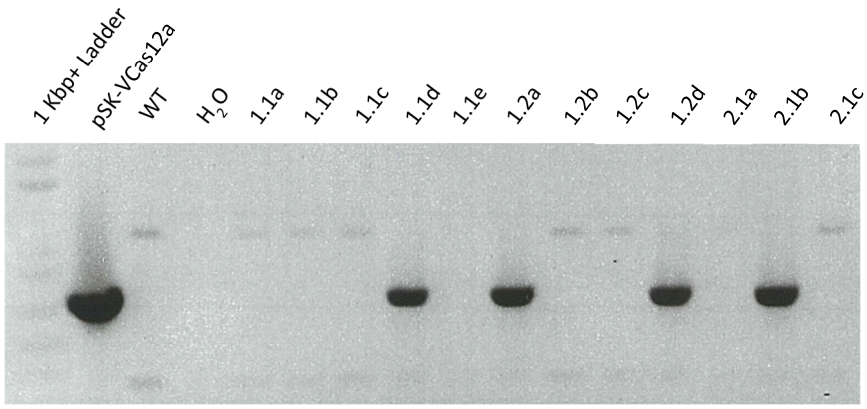


Figure 24. Representative gel showing genotyping of *Vasa-Cas12a* F_0 mice. The plasmid pSK-VCas12a (Figure 20) contained the full Vasa-Cas12a construct and was used as a positive control. Samples 1.1d, 1.2a, 1.2d and 2.1b show clear bands matching the positive control.

Cas12a mRNA levels were characterised in testes, ovaries, and spleens by RT-qPCR in all 4 lines (Figure 25). *Vasa-Cas12a-D* was the first to be characterised, using the *CMV-Cas12a* line as a positive control, as *CMV-Cas12a* mRNA levels were very low in comparison, *Vasa-Cas12a-D* was used as a positive control for testing lines *A*-*C*.

Vasa-Cas12a-D was rejected for experimental use as there was a low level of background expression in all tissue types, this was not seen in any of the other 3 lines. *Vasa-Cas12a-B* was also rejected as it had a much lower level of expression in testes. None of the lines had any substantial expression in ovaries, an unexpected outcome given the expression of Vasa-Cre from the same promoter in Gallardo *et al.* (2007), although similar to what was seen in *Vasa-Cas9* lines as discussed in the germline homing results in section 9.2. Experimental crosses were performed with lines derived from *Vasa-Cas12a-A* and *Vasa-Cas12a-C* as they both had high testes expression.

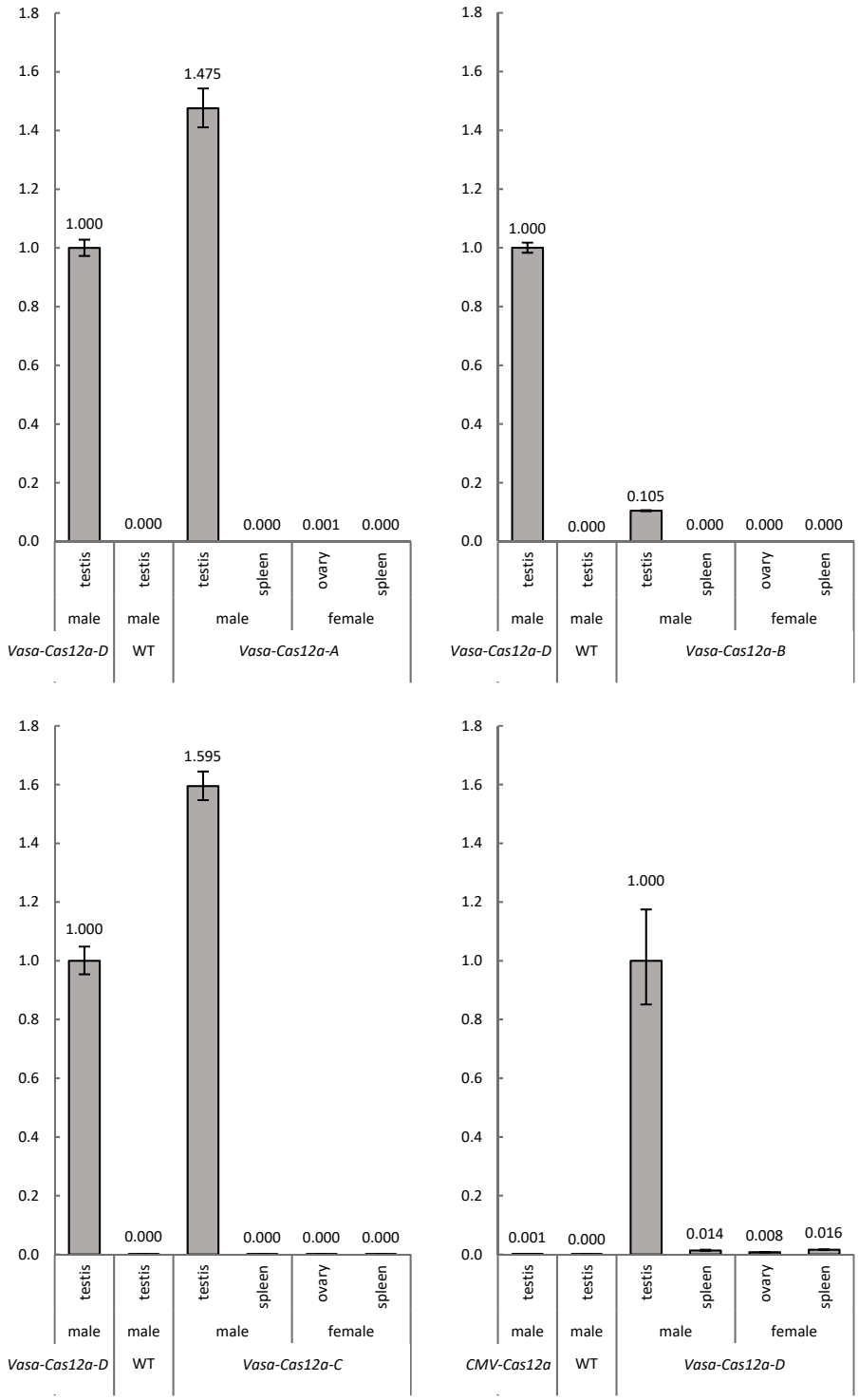


Figure 25. Relative quantification of *cas12a* mRNA levels in 4 different transgenic lines of *Vasa-Cas12a* mice by RT-qPCR (with 95% CI).

10.4.2 Experimental Results

10.4.2.1 *In Vivo* Cleavage Assessment

The *in vivo* effectiveness of Cas12a with Tyr12a-gRNA to cleave Tyr^{Target} needed to be tested in a context that would be as similar to the gene drive mice as possible (see Figure 17 and Figure 18). To this end, microinjection of a plasmid containing *CMV-Cas12a* and U6-Tyr12a-gRNA was performed on $Tyr^{Target/Target}$ zygotes. 27 fetuses were collected and Tyr^{Target} was analysed via Sanger sequencing,

no indels were detected. Experimental crosses were continued despite this data as it could not be definitively concluded that indels were being created and repaired. It was also possible that delivery of the CRISPR components in plasmid form meant Cas12a and the gRNA were not produced quick enough or in sufficient quantity before being degraded or lost.

10.4.2.2 Zygotic Homing

To assess homing in the zygote, we generated 35 *CMV-Cas12a*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice. The Tyr^{Target} locus was analysed via Sanger sequencing for all 35 mice and no cleavage activity was detected. This was confirmed further by Sanger sequencing trace analysis with TIDE[136] and ICE[137] to identify potential low-level cleavage and mosaicism. As DSB repair in zygotes often generates large (>100 bp) deletions,[36] we amplified Tyr^{Target} using primers distant from the cleavage site but did not detect large deletions (Figure 26).

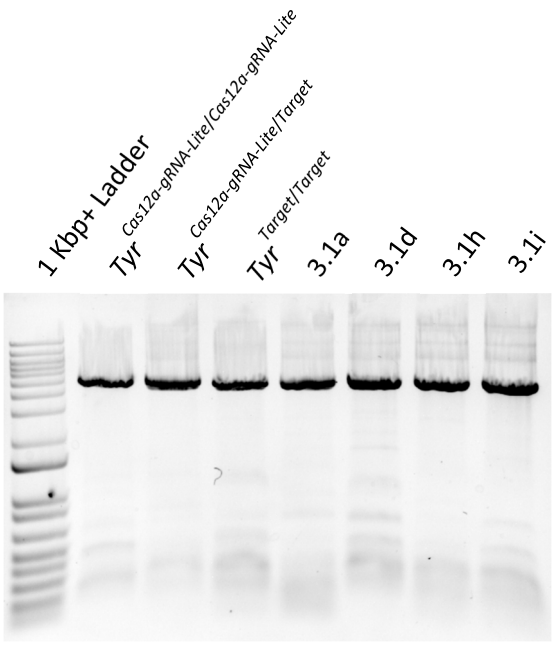


Figure 26. Large deletion genotyping. Representative PCR showing no detectable large deletions around Tyr^{Target} in 4 *CMV-Cas12a*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice. The only band is the expected ~4 kb for no deletion, $Tyr^{Cas12a\text{-}gRNA\text{-}Lite}$ and Tyr^{Target} bands are not different enough in size to be separated on this gel.

10.4.2.3 Germline Homing

To assess germline homing, *Vasa-Cas12a-A*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice were generated and crossed to WT partners. For female *Vasa-Cas12a-A*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice, blastocysts were collected and genotyped instead of embryos. This was done in an attempt to allow a greater number of offspring to be genotyped at a quicker rate.

Altogether, 93 fetuses from 4 *Vasa-Cas12a-A*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ males and 53 blastocysts from 3 *Vasa-Cas12a-A*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ females were screened. If homing was occurring, we would expect greater than 50% of these offspring to contain the $Tyr^{Cas12a\text{-}gRNA\text{-}Lite}$ allele and if no homing was occurring it would remain at 50% (Figure 18). No significant change (χ^2 goodness of fit) from the Mendelian inheritance ratio of 50% $Tyr^{Cas12a\text{-}gRNA\text{-}Lite}$ alleles occurred in either males (49.5%, p=1.0) or females (47.2%, p=0.78) as shown in Figure 27.

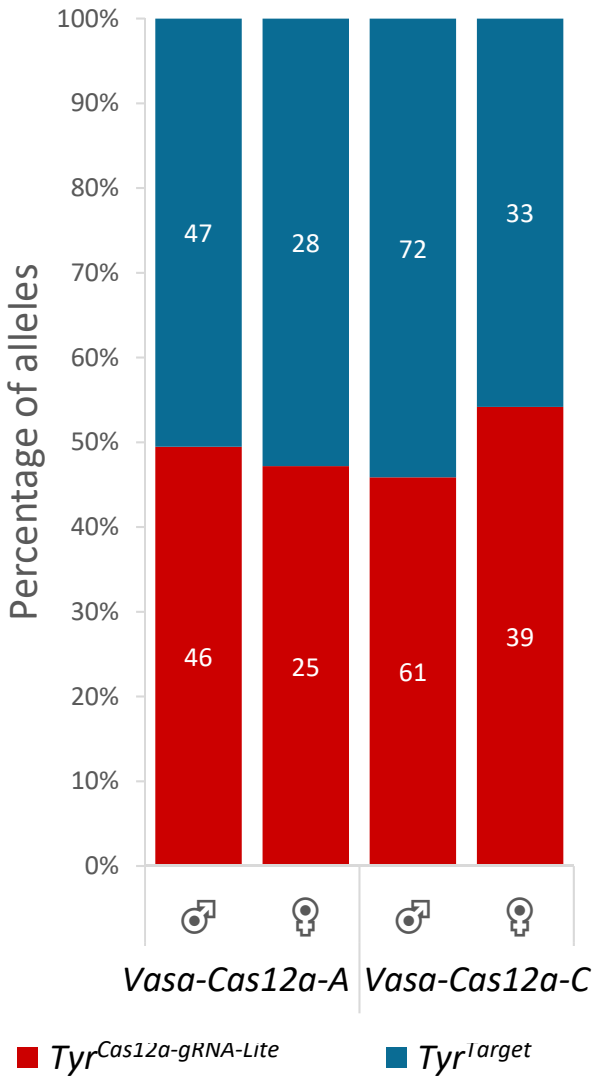


Figure 27. Germline-homing genotyping data. Bar graph showing the percentage of different alleles in the offspring of the gene drive mice *Vasa-Cas12a-A*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ and *Vasa-Cas12a-C*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$. Total number of offspring shown in columns.

To assess the second *Vasa-Cas12a* line used, *Vasa-Cas12a-C*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice were also generated and mated with WT partners. For female *Vasa-Cas12a-C*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice, pups were collected for genotyping instead of blastocysts or embryos. This was done so they would not have to be culled and could continue producing offspring to allow a higher number of offspring from each female.

133 fetuses from 4 *Vasa-Cas12a-C*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ males and 72 pups from 2 *Vasa-Cas12a-C*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ females were screened. No significant change (χ^2 goodness of fit) from the Mendelian inheritance ratio of 50% *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite}$ alleles occurred in either males (45.9%, p=0.38) or females (54.2%, p=0.56) as shown in Figure 27.

To assess the cleavage activity of both lines, a combination of Sanger sequencing, TIDE[136]/ICE[137] analysis (Figure 28) and RFLP (Figure 29) was performed around the *Tyr*^{Target} site in the offspring of all *Vasa-Cas12a-A*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ and *Vasa-Cas12a-C*/+ ; *Tyr*$^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice. No cleavage was seen in female offspring of either line and only very low rates in the males (4.3% and 4.2%, respectively) as shown in Figure 30.

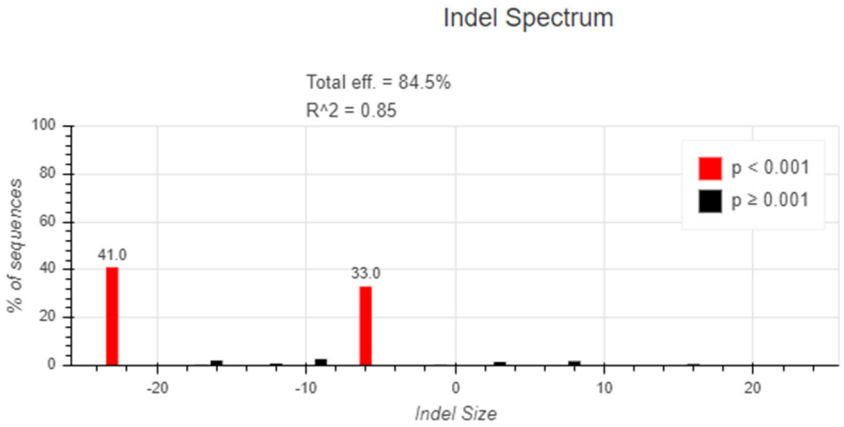


Figure 28. Example TIDE analysis. Shown are results of a single offspring of a WT mouse crossed to *Vasa-Cas12a*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$. $Tyr^{Target/Target}$ DNA is used as the background, which shows the inherited WT allele as a 23 bp deletion. A 6 bp deletion is also seen. Image generated with TIDE.[136]

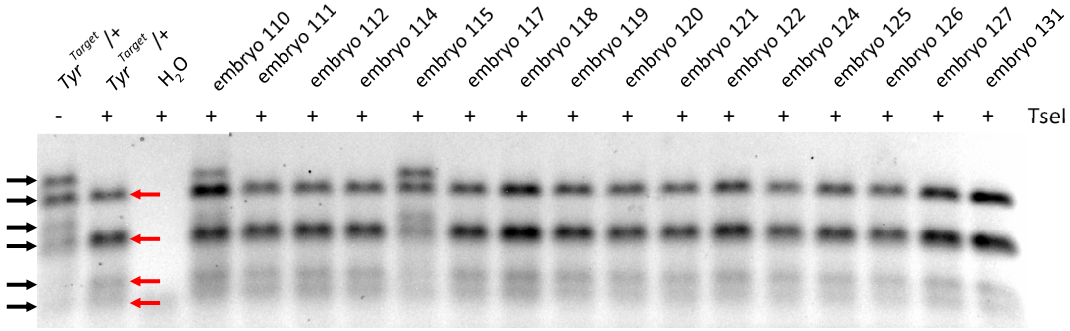


Figure 29. Example RFLP analysis of Tyr^{Target} for detection of indels. Digestion with TseI (as indicated) and T7 Endonuclease I (all samples), black arrows show expected uncut bands due to destruction of TseI site indicating presence of an indel, red arrows show cut bands due to intact TseI site and thus no indel. Samples from embryo 110 and 115 show clear extra bands matching the $Tyr^{Target/+}$ control without TseI, indicating the presence of indels in both samples.

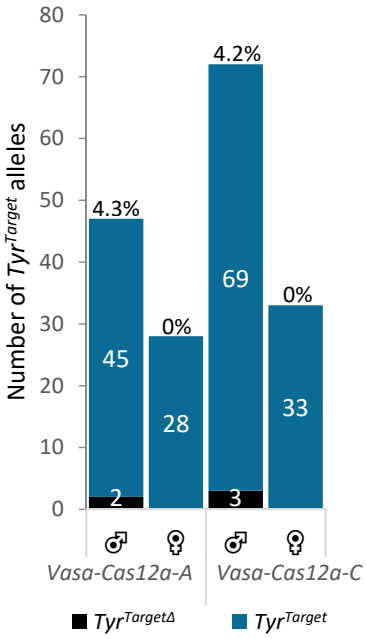


Figure 30. Tally of Tyr^{Target} alleles in offspring of the two *Vasa-Cas12a* gene drive lines. The columns show the number of indels ($Tyr^{Target\Delta}$) and the number of uncut alleles (Tyr^{Target}). The cut rate shown as percentage of Tyr^{Target} alleles with indels is shown at the top.

10.5 Discussion

We found no evidence of either zygotic homing or germline homing in any of the Cas12a lines tested. This is likely due to the very low level of cleavage seen throughout all the lines tested: 0% for all *CMV-Cas12a*, 0% for females from both *Vasa-Cas12a* lines, and only 4.2% and 4.3% in the males of the *Vasa-Cas12a* lines.

What then can we ascribe this low level of cleavage? Two possibilities can be ruled out immediately: The use of an endogenously expressed CRISPR system in general and possible position effects of the genomic locus that both the gRNA ($Tyr^{Cas12a\text{-}gRNA\text{-}Lite}$) and its target ($Tyr^{Target}$) are located in. These can be ruled out as the endogenously expressed Cas9 system which had its corresponding U6-gRNA ($Tyr^{Cas9\text{-}gRNA\text{-}Lite}$ or $Tyr^{Cas9\text{-}gRNA\text{-}Tomato}$) and target ($Tyr^{Target}$) at the exact same genomic locus had high levels of cleavage for multiple configurations. This includes a cleavage rate of 94.8% for the zygotic-homing $Rosa26^{Cas9}$/+ ; $Tyr^{Cas9\text{-}gRNA\text{-}Tomato/Target}$ mice compared to 0% cleavage for the equivalent Cas12a system with *CMV-Cas12a*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ mice. The germline-homing *Vasa-Cas9-4*/+ ; $Tyr^{gRNA\text{-}Lite/Target}$ males also had 87.1% cleavage compared to 4.2% and 4.3% for the equivalent germline-homing *Vasa-Cas12a*/+ ; $Tyr^{Cas12a\text{-}gRNA\text{-}Lite/Target}$ males.

The U6 promoter used for Cas12a is the same construct used in the Cas9 gene drive and previously published Cas9 gene drives in insects. Therefore, it is unlikely that this promoter is causing any issues such as low transcription levels or producing truncated gRNA transcripts that might lead to the low cleavage rates.

Looking specifically at the zygotic-homing gene drives we see the most striking difference between Cas12a and Cas9, 0% and 94.8% cleavage, respectively, and both of these systems should be expressing Cas12a/Cas9 ubiquitously. An obvious explanation presents itself here in that there is a very large difference in protein expression levels, which was indirectly measured by RT-qPCR. Whilst Cas12a and Cas9 were not directly compared to each other we can see that the tissue with the highest level of *CMV-Cas12a* (testes, Figure 25) is negligible in comparison to the testes' expression in *Vasa-Cas12a-A* and *Vasa-Cas12a-C* where there is more than 1000 times as much mRNA. The ovary and spleen expression levels for *CMV-Cas12a* are also barely distinguishable from the WT controls (Figure 23).

The likely cause for the significant differences in mRNA levels between the CMV-Cas12a and $Rosa26^{Cas9}$ transgenes is that they use slightly different versions of the CMV promoter. The CMV in *CMV-Cas12a* is the CMV immediate-early (CMV-IE) enhancer with the CMV promoter downstream of this, whereas the CAG promoter used in $Rosa26^{Cas9}$ is the CMV-IE coupled to the modified chicken β-actin promoter.[122] Original expression data showed CAG produced anywhere from 15% to 1600% more protein than CMV depending on tissue type.[122] Position effects also need to be considered for the Cas12a gene here, as Cas9 was expressed in the Rosa26 locus, well known to provide a safe harbour for highly-expressing transgenes, whereas the Cas12a was randomly integrated in an unknown location which could have compromised its expression.

An initial assumption may be that the complete lack of cleavage in both *Vasa-Cas12a* female lines might be explained by what appears to be no Cas12a expression at all (Figure 25). However, the data for the *Vasa-Cas9-2* line showed a similar level of Cas9 expression that was not distinguishable from background using RT-qPCR and yet there must have been expression as cleavage of Tyr^{Target} was detected. The reason for seeing cleavage in the *Vasa-Cas9-2* line with no detectable Cas9 expression is likely due to the low percentage of cells in the ovary actively expressing under a Vasa promoter compared to a much higher percentage of cells in testes.[129] Therefore the assumption that there is no Cas12a expression and that explains the lack of cleavage is not something that we can decisively conclude.

One further point of difference between the Cas9 and Cas12a constructs is the nuclear localisation signals (NLS's) utilised. The Cas9 constructs contained 5' SV40 NLS's and 3' nucleoplasmin NLS's, whereas the Cas12a constructs contained only 3' nucleoplasmin NLS's. It's possible that this was leading to a lower level of Cas12a in the nucleus compared to Cas9 but this is unlikely considering previously reported comparisons between these different NLS configurations with Cas9 didn't lead to any significant differences in cleavage rates.[138]

Moving away from the endonucleases themselves, another important difference between the Cas12a and Cas9 systems is the gRNA and associated transcription levels of that gRNA. A recent paper by Gao *et al.* (2018) investigated transcription of gRNAs under the U6 promoter and what effect that had on the gRNA product. The first thing to note is that traditionally a thymine (T)-stretch of T4 (4 Ts) or more has been used to terminate U6 transcription[125] and in both the Cas12a and Cas9 systems in this thesis a T5 was used. Gao *et al.* (2018) found that a T6 was required to completely terminate transcription, any less and there would be increasing amounts of read-through transcripts that did not terminate until they found another poly-T signal further downstream.

The above has the potential to reduce the average binding affinity of the gRNA for Cas12a and the subsequent cleavage efficiency. Secondary structure formation in the gRNA when it is longer or a decreased binding efficiency from the longer gRNAs are potential mechanisms here. However, a more important discovery was made in that even when a T6 is present, a variable number of uracil bases are left on the 3' end of the gRNA.[139] In the case of T5 it is anywhere from 1 to 5 uracils.[139] Even this small addition of 1-5 uracils to the end of the guide portion of the 3' end of the Cas12a gRNA was enough to decrease the cleavage efficiency.[139, 140] This decrease in cleavage efficiency was not seen for Cas9,[139] likely because the guide region of the Cas9 gRNA is present at the 5' end instead of the 3' end like Cas12a gRNAs where the poly-U tail is present.[139, 140]

The above features of U6 transcription are likely a significant contributor to the differences seen between Cas9 and Cas12a cleavage. This can potentially be accounted for in future experiments by not relying on the poly-T termination of U6 transcription and instead using a more robust method such as the insertion of a hepatitis delta virus (HDV) ribozyme sequence to cleanly cleave the Cas12a gRNA product directly after the guide sequence.[139, 140]

Setting aside any problematic differences with gRNA expression between the two endonucleases, although the Cas9 and Cas12a guides target the same 20-24 bp sequence, they do target different strands. That, along with inherent differences between the endonucleases mean they would need to be tested in more controlled and isolated conditions to assess whether Cas9 simply cuts more efficiently at that site regardless of promoters and expression levels.

The final factor to account for is what repair pathways predominate when cleavage occurs within the zygote or germline tissue. The Cas9 data definitively informed us that even when there is a high rate of cleavage and under a ubiquitous promoter, there were still some cases where no indels were found. This cannot be unequivocally equated with the term cleavage efficiency as it's been used in this thesis, as the endonuclease may indeed be cleaving the DNA but the cellular repair mechanisms may also be perfectly repairing these DSBs so that when the cut site is analysed there is no evidence of cleavage. The differences in the type of DSB that is produced by Cas9 (blunt) vs. Cas12a (staggered) may also greatly influence this. It is possible that because Cas12a creates a staggered DSB with single-stranded overhangs the DNA is in a state that favours non-NHEJ repair processes,[28] affording a greater probability of the DSB being perfectly repaired unlike Cas9. This is something that needs to be investigated further independent of a gene drive system however, as not enough information on the post-Cas12a cleavage repair pathways are currently known.

11 THESIS DISCUSSION

11.1 ZYGOTIC-HOMING GENE DRIVES

Zygotic homing has been seen as a possible method for controlling the activity of a gene drive, especially as it's known that HDR can occur within zygotes as shown originally by Yang *et al.* (2013) and by the generation of the mouse lines throughout this thesis including Tyr^{Target}, $Tyr^{Cas9\text{-}gRNA\text{-}Tomato}$, $Tyr^{Cas9\text{-}gRNA\text{-}Lite}$, and $Tyr^{Cas12a\text{-}gRNA\text{-}Lite}$. It has also been an attractive option considering there are more options in promoter choice for controlling the gene drive as ubiquitous expression is all that's needed and that is much easier to implement than the more specific control required for germline homing. Although this aspect of less-restrictive promoter choice here needs to be tempered significantly considering the extremely low expression levels seen by the CMV-Cas12a construct.

The data presented in this thesis however, especially that seen for the Cas9 zygotic-homing gene drive is strong evidence against a zygotic-homing gene drive ever being feasible. This negative data is primarily the high rate of error-prone repair (94%) which competes with the desired outcome of homing. The high rate of mosaicism (81%) is also not beneficial to a gene drive. Although that is not as deleterious as the high rate of error-prone repair because even if the embryo maintains its WT alleles into the 2-cell or later stages, there is still a chance that one or more homing events can occur in those cells. In fact, even if a mixture of homing and error-prone repair events occur, if the cell(s) that underwent homing end up being the precursor(s) to the germline tissue, that homing event will be just as effective as a homing event in the zygote. Statistically however, fewer homing events (one) are required if it happens in the zygote in comparison to a 2-cell or later stage. It's possible that the mosaicism could be reduced or even eliminated with the use of a different promoter with a higher activity level, but this is likely to just generate more indels instead of homing, so it probably isn't worth investigating. If we also consider the 6% of mice that had no indels in them at all, it is not unreasonable to assume that the Cas9 activity can't be increased to a level where it guarantees cutting in the zygote, so it might not be possible to completely remove the possibility of mosaicism in zygotic-homing gene drives.

It should be acknowledged that this thesis has only tested a single zygotic promoter, with a single Cas nuclease (successfully), in a single experimental configuration, and in a single organism. Even taking into consideration the zygotic-homing data from Grunwald *et al.* (2019), they tested the same combination of promoter, nuclease, genomic locus of homing cassette, and model organism, perhaps therefore, sweeping conclusions should not be made about zygotic-homing gene drives. However, this argument can really only be made by ignoring the most important factor: That of the state of the chromosomes in the zygote. As discussed in the germline homing discussion in section 9.2, there are two major factors at play that are impediments to a functional zygotic-homing gene drive. First is the repression of transcription from zygotic chromosomes until the G_2 phase[141] and second is the complete separation of maternal and paternal chromosomes into distinct pronuclei until after the G_2 phase, whereupon mitosis starts and the zygote rapidly divides[142]. This leaves a period of approximately 4 hours during G_2[142] where HDR isn't physically possible and rapid generation of indels can take place. The zygote then quickly enters mitosis, dividing within approximately 1 hour,[143] after which it is a 2-cell embryo where if DNA cleavage is still occurring, the likely outcome is mosaicism, regardless of whether resistance alleles are being generated or homing is taking place.

It should be noted that it's possible that homing could occur in both cells of the 2-cell embryo, or as discussed above, homing could occur in only one cell and that cell could be the precursor to the

germline tissue. However, having to rely on more than one homing event in a single animal or relying on it occurring in a specific cell in the developing embryo inherently reduces the likelihood of getting that homing event in comparison to a single event needed in the zygote. We also need to consider how likely it is that homing can occur in rapidly dividing cells that are repeatedly moving through the cell cycle. As discussed in section 6.1.7, HDR is mostly restricted to the S and G_2 phases which actually gives an advantage when it comes to timing for 2-cell embryos. In comparison to the zygote or 4-cell and later stages of development where G_2 takes approximately 4 hours, 2-cell embryos take an extended 10-12 hours in G_2.[142, 143] This is obviously not a guarantee of increased HDR however, but it is worth taking into consideration.

There was no evidence of maternal carryover in the zygotic-homing system presented in this thesis. However, as discussed in section 6.4, this has been seen to occur at very high rates in germline-homing systems in mosquitoes and flies. Considering this, it cannot be discounted that maternal carryover may occur if zygotic-homing gene drives were developed further. Maternal carryover would likely be more detrimental with zygotic-homing as the separated maternal and paternal chromosomes (until G_2) would mean a full 18-20 hours[142] where maternal Cas9 can be actively generating indels without the possibility of homing.

An additional disadvantage of zygotic-homing drives is that they are less amenable to be implemented as a suppression drive. A typical suppression drive spreads a mutation that is homozygous lethal or causes infertility when homozygous, this necessitates the need to keep the gene heterozygotic in the somatic tissue. An efficiently homing zygotic-homing drive however, would cause all cells in the developing embryo to be homozygous and thus they would either not survive or be infertile and as such would not be able to propagate the gene drive further.

A notable, potential exception to the above is a “Y-shredder” gene drive. The fundamental concept behind this kind of drive is that it contains a gRNA with many targets (up to ~300) along the entire Y chromosome (or around the centromere) and the Cas-induced DNA cleavage results in the loss of the entire Y chromosome.[144] Multiple, Y-shredder gRNAs have been assayed and found to effectively convert XY males into XO females containing a single X-chromosome,[144, 145] which are known to be both viable and fertile.[146] Hypothetically this type of system could function as a zygotic-homing gene drive, but given the drawbacks discussed above the prospects are limited at best.

Given all this, it is likely that zygotic-homing gene drives do not have any kind of a future and research should be more focused on germline-homing gene drives and other systems.

11.2 Germline-homing Gene Dives

11.2.1 Cas9 and the Vasa Promoter

A number of papers have now shown that germline homing can occur with high efficiency in insects and yeast[82, 83, 85, 89, 90, 93-107, 109], many of which used various versions of the Vasa promoter, very similar to the one utilised within this thesis. Additionally, Grunwald *et al.* (2019) showed that homing can occur at a relatively low rate in female mice if activity is induced in the germline.

Considering the above, as discussed in the germline homing discussion in section 9.2, this suggests the reason that both the Vasa-Cas9 and Vasa-Cas12a gene drives show no evidence of homing is due to low expression or activity levels in the female germline. The Vasa-Cas9 line with the highest rate of DNA cleavage, *Vasa-Cas9-4*, had ovary expression levels close to background compared to the testes which had ~29-fold higher expression. The *Vasa-Cas12a* lines were indistinguishable from background and didn’t generate a single indel at the target site.

The difference in expression and activity levels of Cas9 and Cas12a between males and females was unexpected. This difference was observed across all four *Vasa-Cas12a* lines (Figure 25), both *Vasa-Cas9* lines (Supplemental Figure 6A in section 9.2), and an additional *Vasa-Cas9* line that was not used (Figure 31). The Vasa promoter fragment used for all of these lines was previously characterised by Gallardo *et al.* (2007) where it was used to drive expression of Cre, also randomly integrated into the mouse genome. Expression levels of Vasa-Cas9 or Vasa-Cas12a cannot be compared directly to Vasa-Cre in Gallardo *et al.* (2007) as northern blots and β-galactosidase assays were used to check and localise expression. A northern blot is semi-quantitative however and the analysis of Cre expression through a northern blot did indicate very similar levels of Cre mRNA in both the testes and ovaries. Therefore, it is unlikely that the difference in expression levels between males and females seen in *Vasa-Cas12a* and *Vasa-Cas9* lines is due to the Vasa promoter fragment. The Vasa-Cre construct does differ somewhat however, it contains an SV40 polyA signal whereas the Vasa-Cas9/Cas12a constructs utilise BGH polyA signals instead. This leaves open the possibility that there were sex-specific differences in rates of mRNA degradation leading to higher mRNA levels in testes. Considering previous studies have shown that the BGH polyA signal leads to higher levels of mRNA than the SV40 polyA signal[147] this is an unlikely explanation.

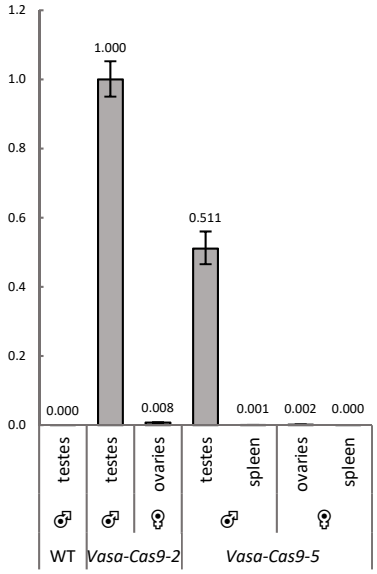


Figure 31. Characterisation of *cas9* mRNA levels in P21 *Vasa-Cas9-5* mice by RT-qPCR (with 95% CI).

Another potential explanation for the sex bias seen in expression levels is position effect. Considering all the Vasa mice were integrated into the mouse genome at random locations, it is possible that the localised chromatin state or nearby regulatory elements are downregulating expression in females. However, considering that 7 different Vasa mice were tested in this thesis, all of them would likely be randomly integrated at different loci, it seems unlikely that the same kind of female-specific downregulation would be occurring in all the lines. It is also possible the Vasa-Cre line used by Gallardo *et al.* (2007) was the exception in that it showed similar expression between sexes as there were an additional 7 Vasa-Cre lines generated for the paper that were not discussed and they possibly could have had similar expression patterns to the Vasa lines generated in this thesis.

Perhaps more important than the lack of activity in the female germline is the complete lack of homing seen in the males. Considering that Vasa-Cas9 expression levels were high and showed a high rate of cleavage, based on the insect gene drives, homing could have reasonably been

expected. The same effect (lack of male homing) was reported by Grunwald *et al.* (2019) but not by any of the insect papers. Considering how critical it is to have a low rate of error-prone repair to avoid generation of resistance alleles (as discussed in section 6.6), this is likely the most important finding to be taken from the work performed in this thesis. It tells us the Vasa promoter fragment is not a suitable promoter and likely does not activate the gene drive at a time where HDR is likely to occur in males, instead promoting error-prone repair pathways.

The timing of expression under the Vasa promoter coupled with sex differences in embryo development provides some insight into why Vasa-driven gene drives in mice may only be viable for homing in females. Although endogenous Vasa is first detected between E10.5-11.5 in both sexes,[148] the Vasa promoter fragment used within this thesis was initially demonstrated to induce expression between E15-18 in testes and before P3 in ovaries.[129] A later study indirectly showed that expression in ovaries was induced between E14.5-16.5, the same approximate window as in the testes.[149]

Stages of spermatogenesis and oogenesis relative to embryonic days (days post-fertilisation) are shown in Figure 32. In the testes, the gonocytes (spermatogonial stem cell progenitors) proliferate until P8.5 when they enter meiosis (leptotene).[150] This means there is approximately 14.5 days of activity of any Vasa-controlled gene drive that occurs during which the gonocytes are undergoing mitosis. Considering that during the normal cell cycle HDR can generally only occur in S and G_2 phases and even then, isn't common (as discussed in 6.1.7), these are not ideal conditions for homing.

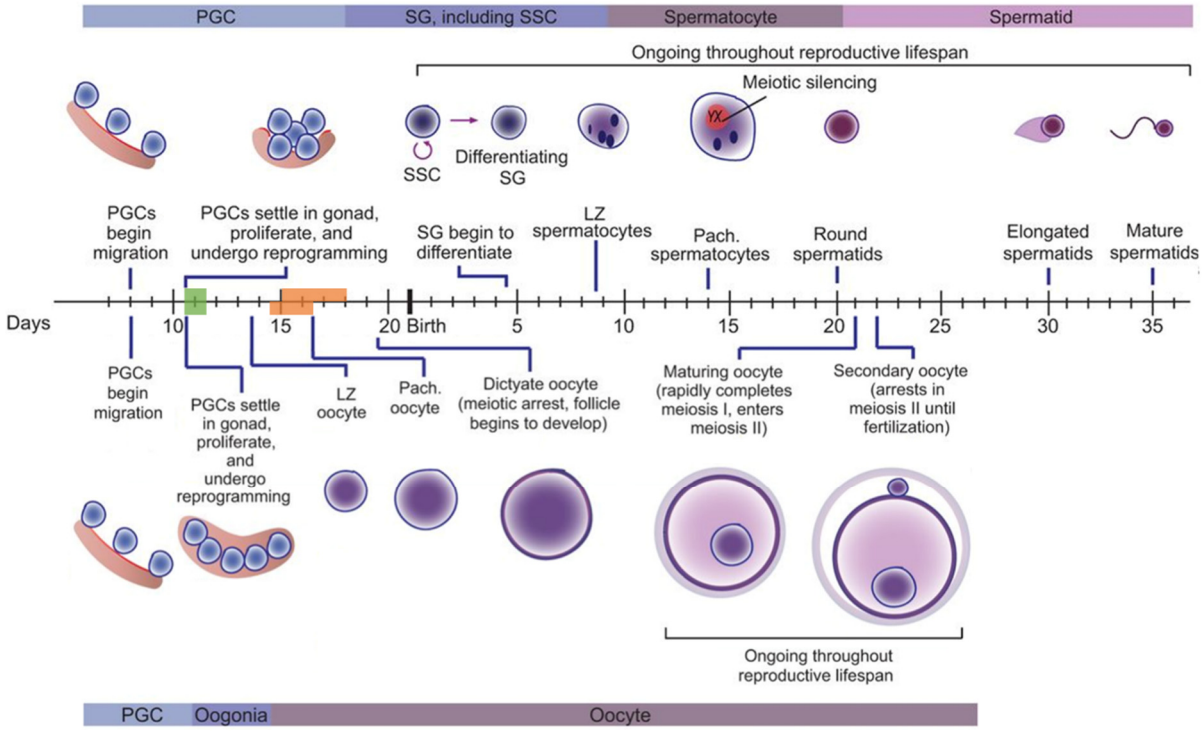


Figure 32. Spermatogenesis (top) and oogenesis (bottom) in mice. Start of expression of endogenous Vasa (green box) and the Vasa promoter fragment (orange box). PGCs, primordial germ cells. SG, spermatogonia. SSC, spermatogonial stem cells. LZ, leptotene/zygotene. Pach, pachytene. Image modified from Hilz *et al.* (2016).

Contrastingly, in the ovaries, the PGCs (oocytes) have already entered meiosis at E13.5 and homologous chromosomes have paired up by E16.5 when the oocytes enter pachytene.[150] It should be noted that the E16.5 timepoint coincides with the first timepoint where evidence of expression from the Vasa promoter fragment has been detected in females.[149] The developing oocyte then has

a further ~25.5 days where the homologous chromosomes are bound together until it reaches P21 and meiosis I rapidly completes.[150] With this data, we can hypothesise that during this lengthy period there could be an increase in HDR simply due to the proximity of the homologous chromosomes considering that they need to interact with one another in order for HDR to take place. Although specific information on the repair pathways in oocytes is not known, spermatocytes undergoing meiosis have been extensively investigated by Enguita-Marruedo *et al.* (2019). They showed that from leptotene up to mid-pachytene, HDR is the preferred method of DNA repair, likely to facilitate crossover between homologous chromosomes.[151] Then from mid-pachytene to late-diplotene NHEJ and HDR were both acting with NHEJ outcompeting HDR early on in late-pachytene.[151] HDR pathways again became dominant closer towards the late-diplotene stage where NHEJ was downregulated.[151] So, although we can't definitively claim these processes occur in an identical fashion in oocytes, it is a reasonable assumption and would mean that for the majority of those first ~25.5 days of Vasa expression, the oocytes are in a cellular state more favourable to homing.

Another important point to consider when comparing *Vasa-Cas9* and *Vasa-Cas12a* mice to those used in Grunwald *et al.* (2019) is that for the latter, a Vasa-Cre gene was used to irreversibly activate ubiquitous expression of Cas9 by the CAG promoter. This likely lead to slightly delayed expression compared to Vasa-Cas9 and probably much higher expression levels. Even considering this, they did not see an impressive homing rate in females compared to the insect gene drives. Lacking further data, it is isn't possible to know what stages of meiosis I that the majority of Cas9-induced cleavage took place and whether that was a time point more conducive to NHEJ or HDR as discussed above.

All this together points to the conclusion that Vasa is simply not a good promoter for a mouse gene drive. It is either not active at the optimum time and/or not enough Cas9 is produced to substantially promote HDR. In addition, the germ cell development state is vastly different between males and females when it is turned on. So ultimately there needs to be better control of Cas9 expression and it needs to be more tightly linked to the cell cycle or the HDR pathways instead of tissue type, allowing for expression only when the homologous chromosomes are lined up and in an ideal state for HDR.

11.2.2 Cas12a

With the low rate of indel formation in both sexes induced by Vasa-Cas12a-A and Vasa-Cas12a-C, should future gene drives be developed with Cas12a? Cas12a generates a staggered break that puts the DNA in a potentially favourable state for non-NHEJ repair pathways[28] and it is critical to avoid all types of error-prone repair, with NHEJ being the major error-prone repair pathway. Considering this, Cas12a has potential to function as an optimal gene drive. However, at this stage the data to support this is indirect and Cas12a has not yet been used for any gene drive system so a lot more preliminary work needs to be done before anything more conclusive can be stated.

The Cas12a system itself needs improvement from what was presented in this thesis. The key outcome needs to be an increase in the cleavage efficiency of the Cas12a/gRNA RNP which needs to be addressed by looking at both components of the complex. Firstly, the expression level of Cas12a needs to be increased, perhaps initially by a similar strategy as demonstrated by Grunwald *et al.* (2019) where a germline promoter drives the expression of Cre to activate a CAG-Cas12a cassette. This would provide a simple test of whether Cas12a is more likely to induce homing compared to Cas9 without having to consider more critical aspects such as the particular germline promoter used. The gRNA is the other component of the RNP that needs to be investigated. As discussed in section 10.4.2.3, the U6 terminator for the gRNA does not cleanly terminate transcription which leads to a reduction in the cleavage efficiency of the RNP. This can be addressed relatively simply by moving

away from the reliance on the poly-T terminator sequence and changing to something else such as a self-cleaving HDV which can cleanly remove the trailing nucleotides on the gRNA.

11.3 FUTURE DIRECTIONS

There are many future avenues of research in the gene drive field and a lot more work needs to be done, especially in mice and other vertebrates where high levels of homing have yet to be seen. Improvements or changes to most aspects of the system need to be investigated, although as discussed in section 11.1, zygotic homing doesn't appear to offer much promise.

Considering that we already know homing can occur with Cas9 in mice,[108] future experiments perhaps should focus on improving this via selection of different promoters until a reasonably high level of homing can be seen and then consider further modifications. Due to the highly divergent timing of gametocyte development between males and females, germline promoters similar to Vasa which don't have any sexual dimorphism are not going to be good candidates. Instead, a better strategy might be to investigate promoters that are active during meiosis, or perhaps even more specifically promoters that are active during the early and late stages of prophase I where we have evidence of HDR being the dominant repair pathway as discussed in section 11.2.1. Promoter candidates can be identified by looking at already characterised proteins that are expressed at these times and fusing their promoter regions with a Cas9 transgene. Many such candidates exist such as the meiotic recombinase-inducing endonuclease Spo11,[152] SYCP3, a structural component of the synaptonemal complex,[153] or the meiotic recombinase DMC1,[154] all of which are active in early leptotene.

It may be feasible to investigate promoters that are sex-specific instead, which might induce a higher level of expression at the ideal time but only in one sex. Whilst this would not be the ideal outcome, a gene drive that functions in only one sex is likely still better than one that doesn't function at all. In addition to this, non-gene drive applications for such a system still exist, such as being able to generate homozygous mice with large transgenes at a greater efficiency.

Once a relatively efficient gene drive can be constructed in mice, there are further modifications that could be investigated to improve the homing efficiency. A more efficient gene drive is not necessarily a precursor to investigating other elements however, and it may be more beneficial to investigate these other options before, or in tandem with investigations on alternative promoters.

Building upon the work performed in this thesis, a functional Cas9-based gene drive may be improved upon by switching to Cas12a to see if that will improve homing further by reducing the rate of indel formation. As discussed in section 10.1.1 and 6.1.7, the staggered DSBs generated by Cas12a may leave the DNA in a state that does not favour NHEJ, increasing the rate of HDR. Building upon this idea, although Cas12a generates staggered DSBs, it does leave a 5' overhang whereas the DNA resectioning that occurs leading to HDR begins by generating 3' overhangs as discussed in 6.1.7. Considering this, a more promising strategy might be to use a Cas9-Nickase with multiple gRNAs targeting closely adjacent loci. If designed appropriately, the ssDNA cleavage by Cas9-Nickase will leave 3' overhangs, perhaps better promoting the non-NHEJ repair pathways.

Moving away from Cas9-Nickase, as discussed in section 6.6, modelling of multiplexed gRNAs targeting adjacent loci is a promising strategy to allow a "second chance" at homing if the second target site is cut after a resistance allele has already formed at the first target site. This has been shown to be a successful strategy when implemented in *D. melanogaster*, reducing resistance allele formation and increasing the homing efficiency.[99]

Once a system for multiplexing gRNAs is implemented, this could be leveraged further by utilising NextGen sequencing data on the target site in germline cells. As discussed in section 9.3.1 the most common resistance alleles can be discovered and further gRNAs can be added to the gene drive homing cassette so that those resistance alleles can be cut a second time, giving the system another chance at homing, effectively increasing the homing efficiency and reducing the generation of resistance alleles.

Because there is a growing body of knowledge on the proteins involved in the different repair pathways, this can potentially be utilised by co-expressing proteins with the gene drive nuclease that either promotes HDR or inhibits error-prone repair pathways such as NHEJ. A couple of examples that have shown some success at enhancing HDR outside of the context of a gene drive include: Increased expression of Rad51,[155] a key protein in the HDR pathway, and separately, expression of an shRNA to knockdown Ku70,[156] part of the Ku complex which binds DSBs in a first step towards NHEJ as discussed in 6.1.7.

11.4 Conclusion

Numerous configurations of gene drive systems in mice were experimentally constructed and extensively tested throughout this thesis. Although no detectable level of homing was seen, a great deal of information has been learnt. Zygotic-homing gene drives are likely not a viable option going forward and a germline-homing gene drive will require a good deal more investigation before something functional is generated. Nevertheless, the knowledge gained here pushes the direction of future research, clearly indicating the need to focus on the activation and timing of the gene drive to favour HDR. Exploration of promoters expressing during meiosis in combination with other HDR-enhancing strategies are likely to yield promising results.

12 REFERENCES

1. van der Zanden AG, Kremer K, Schouls LM, Caimi K, Cataldi A, Hulleman A, et al. Improvement of differentiation and interpretability of spoligotyping for Mycobacterium tuberculosis complex isolates by introduction of new spacer oligonucleotides. J Clin Microbiol 2002;40(12):4628-39.
2. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science 2007;315(5819):1709-12.
3. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA—Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 2012;337(6096):816–21.
4. Yeh CD, Richardson CD, Corn JE. Advances in genome editing through control of DNA repair pathways. Nature cell biology 2019;21(12):1468-78.
5. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell 2013;154(6):1370-9.
6. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. Science 2014;343(6176):1247997-.
7. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 2014;156(5):935-49.
8. BioRender. BioRender.com. A tool to help scientists create and share scientific figures. World Wide Web: https://biorender.com/
9. Chari R, Mali P, Moosburner M, Church GM. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nature methods 2015;12(9):823-6.
10. Wong N, Liu W, Wang X. WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. Genome Biol 2015;16:218.
11. F.Altschul S, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology 1990;215(3):403-10.
12. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 2013;31(9):822-6.
13. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol 2013;31(9):839-43.
14. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 2013;31(9):827-32.
15. Stemmer M, Thumberger T, Del Sol Keyer M, Wittbrodt J, Mateo JL. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. PloS one 2015;10(4):e0124633.
16. O'Brien A, Bailey TL. GT-Scan: identifying unique genomic targets. Bioinformatics 2014;30(18):2673-5.
17. Benchling [Biology Software]. 2020. World Wide Web: https://benchling.com
18. Shen B, Zhang J, Wu H, Wang J, Ma K, Li Z, et al. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. Cell research 2013;23(5):720-3.
19. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science 2013;339(6121):819-23.
20. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science 2013;339(6121):823-6.
21. Adikusuma F, Pfitzner C, Thomas PQ. Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. PloS one 2017;12(12):e0187236.
22. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science 2014;343(6166):80-4.
23. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 2014;343(6166):84-7.

24. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell 2014;157(6):1262-78.
25. Quadros RM, Miura H, Harms DW, Akatsuka H, Sato T, Aida T, et al. Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. Genome Biol 2017;18(1):92.
26. Meyer M, de Angelis MH, Wurst W, Kuhn R. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proceedings of the National Academy of Sciences of the United States of America 2010;107(34):15022-6.
27. Maher RL, Branagan AM, Morrical SW. Coordination of DNA replication and recombination activities in the maintenance of genome stability. J Cell Biochem 2011;112(10):2672-82.
28. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat Rev Mol Cell Biol 2017;18(8):495-506.
29. Roth DB, Wilson JH. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. Mol Cell Biol 1986;6(12):4295-304.
30. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends in genetics : TIG 2008;24(11):529-38.
31. Sugawara N, Ira G, Haber JE. DNA Length Dependence of the Single-Strand Annealing Pathway and the Role of Saccharomyces cerevisiae RAD59 in Double-Strand Break Repair. Mol Cell Biol 2000;20(14):5300–9.
32. Mao Z, Bozzella M, Seluanov A, Gorbunova V. Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair (Amst) 2008;7(10):1765-71.
33. Griffith AJ, Blier PR, Mimori T, Hardin JA. Ku Polypeptides Synthesized in Vitro Assemble Into Complexes Which Recognize Ends of Double-Stranded DNA. J Biol Chem 1992;267(1):331-8.
34. Shen MW, Arbab M, Hsu JY, Worstell D, Culbertson SJ, Krabbe O, et al. Predictable and precise template-free CRISPR editing of pathogenic variants. Nature 2018;563(7733):646-51.
35. Truong LN, Li Y, Shi LZ, Hwang PY, He J, Wang H, et al. Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America 2013;110(19):7720-5.
36. Adikusuma F, Piltz S, Corbett MA, Turvey M, McColl SR, Helbig KJ, et al. Large deletions induced by Cas9 cleavage. Nature 2018;560(7717):E8-E9.
37. Orthwein A, Noordermeer SM, Wilson MD, Landry S, Enchev RI, Sherker A, et al. A mechanism for the suppression of homologous recombination in G1 cells. Nature 2015;528(7582):422-6.
38. Shibata A, Moiani D, Arvai AS, Perry J, Harding SM, Genois MM, et al. DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. Molecular cell 2014;53(1):7-18.
39. Renkawitz J, Lademann CA, Kalocsay M, Jentsch S. Monitoring homology search during DNA double-strand break repair in vivo. Molecular cell 2013;50(2):261-72.
40. Storici F, Snipe JR, Chan GK, Gordenin DA, Resnick MA. Conservative repair of a chromosomal double-strand break by single-strand DNA through two steps of annealing. Mol Cell Biol 2006;26(20):7645-57.
41. Feng Y, Chen C, Han Y, Chen Z, Lu X, Liang F, et al. Expanding CRISPR/Cas9 Genome Editing Capacity in Zebrafish Using SaCas9. G3 (Bethesda) 2016;6(8):2517-21.
42. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 2015;163(3):759-71.
43. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 2013;31(9):833-8.
44. Robertson L, Pederick D, Piltz S, White M, Nieto A, Ahladas M, et al. Expanding the RNA-Guided Endonuclease Toolkit for Mouse Genome Editing. CRISPR J 2018;1:431-9.

45. Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 2018;556(7699):57-63.
46. Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, Okazaki S, et al. Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science 2018;361(6408):1259-62.
47. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature 2016;529(7587):490-5.
48. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. Science 2016;351(6268):84-8.
49. Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB, et al. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. Nature 2017;550(7676):407-10.
50. Casini A, Olivieri M, Petris G, Montagna C, Reginato G, Maule G, et al. A highly specific SpCas9 variant is identified by in vivo screening in yeast. Nat Biotechnol 2018;36(3):265-71.
51. Lee JK, Jeong E, Lee J, Jung M, Shin E, Kim YH, et al. Directed evolution of CRISPR-Cas9 to increase its specificity. Nat Commun 2018;9(1):3048.
52. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 2016;533(7603):420-4.
53. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 2017;551(7681):464-71.
54. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 2019;576(7785):149-57.
55. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 2013;152(5):1173-83.
56. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nature methods 2013;10(10):973-6.
57. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. Nature methods 2013;10(10):977-9.
58. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, E PRI, et al. Highly efficient Cas9-mediated transcriptional programming. Nature methods 2015;12(4):326-8.
59. Esvelt KM, Smidler AL, Catteruccia F, Church GM. Concerning RNA-guided gene drives for the alteration of wild populations. Elife 2014:e03401.
60. Mendel JG. Versuche über Pflanzenhybriden. Verhandlungen des Naturforschenden Vereins Brünn 1865;(Band 4):3-47.
61. Tu Z, Coates C. Mosquito transposable elements. Insect Biochem Mol Biol 2004;34(7):631-44.
62. Finnegan D. Transposable elements. Current opinion in genetics & development 1992;2(6):861-7.
63. Sinkins SP, Gould F. Gene drive systems for insect disease vectors. Nature reviews Genetics 2006;7(6):427-35.
64. Burt A, Koufopanou V. Homing endonuclease genes: the rise and fall and rise again of a selfish element. Current opinion in genetics & development 2004;14(6):609-15.
65. Chan YS, Takeuchi R, Jarjour J, Huen DS, Stoddard BL, Russell S. The design and in vivo evaluation of engineered I-OnuI-based enzymes for HEG gene drive. PloS one 2013;8(9):e74254.
66. Lyttle TW. Segregation Distorters. Annu Rev Genet 1991;1991:511-77.
67. Kelemen RK, Vicoso B. Complex History and Differentiation Patterns of the t-Haplotype, a Mouse Meiotic Driver. Genetics 2018;208(1):365-75.
68. Charron Y, Willert J, Lipkowitz B, Kusecek B, Herrmann BG, Bauer H. Two isoforms of the RAC-specific guanine nucleotide exchange factor TIAM2 act oppositely on transmission ratio distortion by the mouse t-haplotype. PLoS genetics 2019;15(2):e1007964.
69. Ohkawa J, Taira K. Control of the Functional Activity of an Antisense RNA by a Tetracycline-Responsive Derivative of the Human U6 snRNA Promoter. Human Gene Therapy 2004;11:577-85.

70. Gaj T, Gersbach CA, Barbas CF, 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 2013;31(7):397-405.
71. Garcia-Tunon I, Alonso-Perez V, Vuelta E, Perez-Ramos S, Herrero M, Mendez L, et al. Splice donor site sgRNAs enhance CRISPR/Cas9-mediated knockout efficiency. PloS one 2019;14(5):e0216674.
72. Pimentel D, Zuniga R, Morrison D. Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics 2005;52(3):273-88.
73. Stenseth NC, Leirs H, Skonhoft A, Davis SA, Pech RP, Andreassen HP, et al. Mice, rats, and people: the bio-economics of agricultural rodent pests. Frontiers in Ecology and the Environment 2003;1(7):367-75.
74. McLeod R, Counting the cost: impact of invasive animals in Australia. Canberra: Cooperative Research Centre for Pest Animal Control, 2004.
75. Blackburn TM, Cassey P, Duncan RP, Evans KL, Gaston KJ. Avian Extinction and Mammalian Introductions on Oceanic Islands. Science 2004;305(5692):1955-8.
76. Harris DB. Review of negative effects of introduced rodents on small mammals on islands. Biological Invasions 2008;11(7):1611-30.
77. Doherty TS, Glen AS, Nimmo DG, Ritchie EG, Dickman CR. Invasive predators and global biodiversity loss. Proceedings of the National Academy of Sciences of the United States of America 2016;113(40):11261-5.
78. Piaggio AJ, Segelbacher G, Seddon PJ, Alphey L, Bennett EL, Carlson RH, et al. Is It Time for Synthetic Biodiversity Conservation? Trends in ecology & evolution 2016.
79. Gregory S, Henderson W, Smee E, Cassey P, Eradications of vertebrate pests in Australia: A review and guidelines for future best practice. Canberra, Australia: Invasive Animals Cooperative Research Centre, 2014;1-90.
80. Burt A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. Proc Biol Sci 2003;270(1518):921-8.
81. Prowse TAA, Cassey P, Ross JV, Pfitzner C, Wittmann TA, Thomas P. Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. Proc Biol Sci 2017;284(1860).
82. Gantz VM, Bier E. The mutagenic chain reaction A method for converting heterozygous to homozygous mutations. Science 2015;348(6233):442-4.
83. Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. Proceedings of the National Academy of Sciences of the United States of America 2015;112(49):E6736-43.
84. Isaacs AT, Jasinskiene N, Tretiakov M, Thiery I, Zettor A, Bourgouin C, et al. Transgenic Anopheles stephensi coexpressing single-chain antibodies resist Plasmodium falciparum development. Proceedings of the National Academy of Sciences of the United States of America 2012;109(28):E1922-30.
85. DiCarlo JE, Chavez A, Dietz SL, Esvelt KM, Church GM. Safeguarding CRISPR-Cas9 gene drives in yeast. Nat Biotechnol 2015;33(12):1250-5.
86. Akbari OS, Bellen HJ, Bier E, Bullock SL, Burt A, Church GM, et al. Safeguarding gene drive experiments in the laboratory. Science 2015;349(6251):927-9.
87. DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res 2013;41(7):4336-43.
88. Akhmetov A, Laurent JM, Gollihar J, Gardner EC, Garge RK, Ellington AD, et al. Single-step Precision Genome Editing in Yeast Using CRISPR-Cas9. Bio Protoc 2018;8(6).
89. Hammond A, Galizi R, Kyrou K, Simoni A, Siniscalchi C, Katsanos D, et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. Nat Biotechnol 2016;34(1):78-83.

90. Champer J, Reeves R, Oh SY, Liu C, Liu J, Clark AG, et al. Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations. PLoS genetics 2017;13(7):e1006796.
91. Unckless RL, Clark AG, Messer PW. Evolution of Resistance Against CRISPR/Cas9 Gene Drive. Genetics 2017;205(2):827-41.
92. Port F, Chen HM, Lee T, Bullock SL. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proceedings of the National Academy of Sciences of the United States of America 2014;111(29):E2967-76.
93. Kyrou K, Hammond AM, Galizi R, Kranjc N, Burt A, Beaghton AK, et al. A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. Nat Biotechnol 2018;36(11):1062-6.
94. Hammond AM, Kyrou K, Bruttini M, North A, Galizi R, Karlsson X, et al. The creation and selection of mutations resistant to a gene drive over multiple generations in the malaria mosquito. PLoS genetics 2017;13(10):e1007039.
95. Carballar-Lejarazu R, Ogaugwu C, Tushar T, Kelsey A, Pham TB, Murphy J, et al. Next-generation gene drive for population modification of the malaria vector mosquito, Anopheles gambiae. Proceedings of the National Academy of Sciences of the United States of America 2020;117(37):22805-14.
96. Pham TB, Phong CH, Bennett JB, Hwang K, Jasinskiene N, Parker K, et al. Experimental population modification of the malaria vector mosquito, Anopheles stephensi. PLoS genetics 2019;15(12):e1008440.
97. Adolfi A, Gantz VM, Jasinskiene N, Lee H-F, Hwang K, Bulger EA, et al. Efficient population modification gene-drive rescue system in the malaria mosquito Anopheles stephensi. bioRxiv 2020.
98. Li M, Yang T, Kandul NP, Bui M, Gamez S, Raban R, et al. Development of a confinable gene drive system in the human disease vector Aedes aegypti. Elife 2020;9.
99. Champer J, Liu J, Oh SY, Reeves R, Luthra A, Oakes N, et al. Reducing resistance allele formation in CRISPR gene drive. Proceedings of the National Academy of Sciences of the United States of America 2018;115(21):5522-7.
100. KaramiNejadRanjbar M, Eckermann KN, Ahmed HMM, Sanchez CH, Dippel S, Marshall JM, et al. Consequences of resistance evolution in a Cas9-based sex conversion-suppression gene drive for insect pest management. Proceedings of the National Academy of Sciences of the United States of America 2018;115(24):6189-94.
101. Oberhofer G, Ivy T, Hay BA. Behavior of homing endonuclease gene drives targeting genes required for viability or female fertility with multiplexed guide RNAs. Proceedings of the National Academy of Sciences of the United States of America 2018;115(40):E9343-E52.
102. Guichard A, Haque T, Bobik M, Xu XS, Klanseck C, Kushwah RBS, et al. Efficient allelic-drive in Drosophila. Nat Commun 2019;10(1):1640.
103. Lopez Del Amo V, Bishop AL, Sanchez CH, Bennett JB, Feng X, Marshall JM, et al. A transcomplementing gene drive provides a flexible platform for laboratory investigation and potential field deployment. Nat Commun 2020;11(1):352.
104. Lopez Del Amo V, Leger BS, Cox KJ, Gill S, Bishop AL, Scanlon GD, et al. Small-Molecule Control of Super-Mendelian Inheritance in Gene Drives. Cell reports 2020;31(13):107841.
105. Champer J, Yang E, Lee E, Liu J, Clark AG, Messer PW. A CRISPR homing gene drive targeting a haplolethal gene removes resistance alleles and successfully spreads through a cage population. Proceedings of the National Academy of Sciences of the United States of America 2020.
106. Xu XS, Bulger EA, Gantz VM, Klanseck C, Heimler SR, Auradkar A, et al. Active Genetic Neutralizing Elements for Halting or Deleting Gene Drives. Molecular cell 2020.
107. Kandul NP, Liu J, Buchman A, Gantz VM, Bier E, Akbari OS. Assessment of a Split Homing Based Gene Drive for Efficient Knockout of Multiple Genes. G3 (Bethesda) 2020;10(2):827-37.
108. Grunwald HA, Gantz VM, Poplawski G, Xu XS, Bier E, Cooper KL. Super-Mendelian inheritance mediated by CRISPR-Cas9 in the female mouse germline. Nature 2019;566(7742):105-9.

109. Simoni A, Hammond AM, Beaghton AK, Galizi R, Taxiarchi C, Kyrou K, et al. A male-biased sex-distorter gene drive for the human malaria vector Anopheles gambiae. Nat Biotechnol 2020.
110. Unckless RL, Messer PW, Connallon T, Clark AG. Modeling the Manipulation of Natural Populations by the Mutagenic Chain Reaction. Genetics 2015;201(2):425-31.
111. Marshall JM, Buchman A, Sanchez CH, Akbari OS. Overcoming evolved resistance to population-suppressing homing-based gene drives. Sci Rep 2017;7(1):3776.
112. Noble C, Olejarz J, Esvelt KM, Church GM, Nowak MA. Evolutionary dynamics of CRISPR gene drives. Sci Adv 2017;3(4).
113. Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. Nat Biotechnol 2015;33(2):179-86.
114. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res 2014;24(1):132-41.
115. Kabadi AM, Ousterout DG, Hilton IB, Gersbach CA. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. Nucleic Acids Res 2014;42(19):e147.
116. Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han YC, et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature 2014;516(7531):423-7.
117. Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci Rep 2014;4:5400.
118. Vidigal JA, Ventura A. Rapid and efficient one-step generation of paired gRNA CRISPR-Cas9 libraries. Nat Commun 2015;6:8083.
119. Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol 2019;37(3):224-6.
120. Strohkendl I, Saifuddin FA, Rybarski JR, Finkelstein IJ, Russell R. Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a. Molecular cell 2018;71(5):816-24 e3.
121. Moreno-Mateos MA, Fernandez JP, Rouet R, Vejnar CE, Lane MA, Mis E, et al. CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. Nat Commun 2017;8(1):2024.
122. Hitoshi N, Ken-ichi Y, Jun-ichi M. Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene 1991;108(2):193-9.
123. Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol 2004;22(12):1567-72.
124. Schek N, Cooke C, Alwine JC. Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analyses. American Society for Microbiology Journals 1992;12(12):5386-93.
125. Arimbasseri AG, Rijal K, Maraia RJ. Transcription termination by the eukaryotic RNA polymerase III. Biochimica et biophysica acta 2013;1829(3-4):318-30.
126. Zhang Y, Long C, Li H, McAnally JR, Baskin KK, Shelton JM, et al. CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice. Sci Adv 2017;3(4).
127. Toth E, Weinhardt N, Bencsura P, Huszar K, Kulcsar PI, Talas A, et al. Cpf1 nucleases demonstrate robust activity to induce DNA modification by exploiting homology directed repair pathways in mammalian cells. Biol Direct 2016;11:46.
128. Kim Y, Cheong SA, Lee JG, Lee SW, Lee MS, Baek IJ, et al. Generation of knockout mice by Cpf1-mediated gene targeting. Nat Biotechnol 2016;34(8):808-10.
129. Gallardo T, Shirley L, John GB, Castrillon DH. Generation of a germ cell-specific mouse transgenic Cre line, Vasa-Cre. Genesis 2007;45(6):413-7.
130. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nature protocols 2013;8(11):2281-308.
131. Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell 2014;159(2):440-55.

132. Wu S, Ying G, Wu Q, Capecchi MR. A protocol for constructing gene targeting vectors: generating knockout mice for the cadherin family and beyond. Nature protocols 2008;3(6):1056-76.
133. Zetsche B, Heidenreich M, Mohanraju P, Fedorova I, Kneppers J, DeGennaro EM, et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat Biotechnol 2017;35(1):31-4.
134. Online Manual for CARD Mouse Reproductive Technology, In Vitro Fertilization using Epididymal Sperm Transported at Cold Temperature. Center for Animal Resources and Development, Kumamoto University. World Wide Web: http://card.medic.kumamoto-u.ac.jp/card/english/sigen/manual/lowtempce.html
135. Online Manual for CARD Mouse Reproductive Technology, In Vitro Fertilization (IVF). Center for Animal Resources and Development, Kumamoto University. World Wide Web: http://card.medic.kumamoto-u.ac.jp/card/english/sigen/manual/mouseivf.html
136. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res 2014;42(22):e168.
137. Synthego Performance Analysis, ICE Analysis. 2019. v2.0. Synthego. World Wide Web: https://ice.synthego.com/
138. Hu P, Zhao X, Zhang Q, Li W, Zu Y. Comparison of Various Nuclear Localization Signal-Fused Cas9 Proteins and Cas9 mRNA for Genome Editing in Zebrafish. G3: Genes, Genomes, Genetics 2018;8(3):823-31.
139. Gao Z, Herrera-Carrillo E, Berkhout B. Delineation of the Exact Transcription Termination Signal for Type 3 Polymerase III. Mol Ther Nucleic Acids 2018;10:36-44.
140. Gao Z, Herrera-Carrillo E, Berkhout B. Improvement of the CRISPR-Cpf1 system with ribozyme-processed crRNA. RNA Biol 2018;15(12):1458-67.
141. Jukam D, Shariati SAM, Skotheim JM. Zygotic Genome Activation in Vertebrates. Dev Cell 2017;42(4):316-32.
142. Ciemerych MA, Sicinski P. Cell cycle in mouse development. Oncogene 2005;24(17):2877-98.
143. Gu B, Posfai E, Rossant J. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. Nat Biotechnol 2018;36(7):632-7.
144. Prowse TA, Adikusuma F, Cassey P, Thomas P, Ross JV. A Y-chromosome shredding gene drive for controlling pest vertebrate populations. Elife 2019;8.
145. Adikusuma F, Williams N, Grutzner F, Hughes J, Thomas P. Targeted Deletion of an Entire Chromosome Using CRISPR/Cas9. Mol Ther 2017;25(8):1736-8.
146. Probst FJ, Cooper ML, Cheung SW, Justice MJ. Genotype, phenotype, and karyotype correlation in the XO mouse model of Turner Syndrome. The Journal of heredity 2008;99(5):512-7.
147. Pfarr DS, Rieser LA, Woychik RP, Rottman FM, Rosenberg M, Reff ME. Differential Effects of Polyadenylation Regions on Gene Expression in Mammalian Cells. DNA 1986;5(2):115-22.
148. Tanaka SS, Toyooka Y, Akasu R, Katoh-Fukui Y, Nakahara Y, Suzuki R, et al. The mouse homolog of Drosophila Vasa is required for the development of male germ cells. Genes & development 2000;14(7):841-53.
149. Holt JE, Pye V, Boon E, Stewart JL, Garcia-Higuera I, Moreno S, et al. The APC/C activator FZR1 is essential for meiotic prophase I in mice. Development 2014;141(6):1354-65.
150. Hilz S, Modzelewski AJ, Cohen PE, Grimson A. The roles of microRNAs and siRNAs in mammalian spermatogenesis. Development 2016;143(17):3061-73.
151. Enguita-Marruedo A, Martin-Ruiz M, Garcia E, Gil-Fernandez A, Parra MT, Viera A, et al. Transition from a meiotic to a somatic-like DNA damage response during the pachytene stage in mouse meiosis. PLoS genetics 2019;15(1):e1007439.
152. Chicheportiche A, Bernardino-Sgherri J, de Massy B, Dutrillaux B. Characterization of Spo11-dependent and independent phospho-H2AX foci during meiotic prophase I in the male mouse. J Cell Sci 2007;120(Pt 10):1733-42.
153. Syrjanen JL, Pellegrini L, Davies OR. A molecular model for the role of SYCP3 in meiotic chromosome organisation. Elife 2014;3.

154. Tarsounas M, Morita T, Pearlman RE, Moens PB. RAD51 and DMC1 form mixed complexes associated with mouse meiotic chromosome cores and synaptonemal complexes. J Cell Biol 1999;147(2):207-20.
155. Song J, Yang D, Xu J, Zhu T, Chen YE, Zhang J. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. Nat Commun 2016;7:10548.
156. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 2015;33(5):543-8.