A beginner's guide to understanding and implementing the genetic modification of zebrafish

Sara A. Rafferty, T. Alexander Quinn

Article history:
Received 18 May 2018
Received in revised form 7 July 2018
Accepted 10 July 2018
Available online 19 July 2018

Keywords:
Gene editing
Transgenics
Tol2
TALENs
CRISPR/Cas9
Cardiovascular

Abstract

Zebrafish are a relevant and useful vertebrate model species to study normal- and patho-physiology, including that of the heart, due to conservation of protein-coding genes, organ system organisation and function, and efficient breeding and housing. Their amenability to genetic modification, particularly compared to other vertebrate species, is another great advantage, and is the focus of this review. A vast number of genetically engineered zebrafish lines and methods for their creation exist, but their incorporation into research programs is hindered by the overwhelming amount of technical details. The purpose of this paper is to provide a simplified guide to the fundamental information required by the uninhibited researcher for the thorough understanding, critical evaluation, and effective implementation of genetic approaches in the zebrafish. First, an overview of existing zebrafish lines generated through large scale chemical mutagenesis, retroviral insertional mutagenesis, and gene and enhancer trap screens is presented. Second, descriptions of commonly-used genetic modification methods are provided including Tol2 transposon, TALENs (transcription activator-like effector nucleases), and CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9). Lastly, design features of genetic modification strategies such as promoters, fluorescent reporters, and conditional transgenesis, are summarised. As a comprehensive resource containing both background information and technical notes of how to obtain or generate zebrafish, this review compliments existing resources to facilitate the use of genetically-modified zebrafish by researchers who are new to the field.

© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

When choosing a model organism to study human pathophysiology, careful consideration must be given to the similarity of relevant processes in each species. The utility of experimental animals to further our mechanistic understanding of normal and diseased function is greatly expanded by their genetic modification, as particular genes can be manipulated to determine the phenotypic outcome. For cardiac research, genetic modification has been performed in fruit flies (Wolf and Rockman, 2011), zebrafish (Gut et al., 2016; Gut et al., 2017; Poon and Brand, 2013), mice (Davis et al., 2012), rats (Pinto et al., 1998), rabbits (Peng, 2012), and pigs (Perleberg et al., 2018), which has served to increase our knowledge of heart development, function, and pathology.

Each of the above species has specific advantages and limitations. The zebrafish in particular is becoming an increasingly powerful vertebrate model for cardiac research, and an alternative to the mouse as an ‘easily’ genetically modifiable organism with current tools. In the zebrafish, ~70% of genes correspond to a human gene encoding a protein with the same function (Howe et al., 2013). Further benefits of the zebrafish as a model organism include external and translucent embryonic development (allowing in vivo imaging), large number of offspring, and relatively rapid maturation to adult. Despite having a two-chambered heart, zebrafish are relevant for cardiac research given their similarity to humans in terms of fundamental characteristics, such as heart rate, excitation patterns (reflecting the electrocardiogram), myocyte action potential morphology, and underlying ion channels (Genge et al., 2016; Gut et al., 2017; Poon and Brand, 2013).

A large array of approaches have been developed for editing the zebrafish genome, yet for the uninitiated, the associated terminology and particular advantages of each technique can be overwhelming and difficult to comprehend. This review is meant as a guide for the non-expert, aiming to explain the techniques available for genetically modifying zebrafish, to improve access and facilitate their use by labs performing cardiac research. We first provide a brief overview of genetic resources available specifically

---

**Abbreviations**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-HT</td>
<td>4-hydroxytamoxifen</td>
</tr>
<tr>
<td>amhc</td>
<td>atrial myosin heavy chain</td>
</tr>
<tr>
<td>Cas</td>
<td>CRISPR-associated protein</td>
</tr>
<tr>
<td>cmic2</td>
<td>cardiac myosin light chain 2</td>
</tr>
<tr>
<td>cmr</td>
<td>chloramphenicol resistance</td>
</tr>
<tr>
<td>CRISPR</td>
<td>clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>crRNA</td>
<td>CRISPR RNA</td>
</tr>
<tr>
<td>CZRC</td>
<td>China Zebrafish Resource Centre</td>
</tr>
<tr>
<td>dCas9</td>
<td>catalytically inactive Cas9</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dpf</td>
<td>days post fertilisation</td>
</tr>
<tr>
<td>DSB</td>
<td>double-strand break</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double stranded DNA</td>
</tr>
<tr>
<td>ENU</td>
<td>N-ethyl-N-nitrosourea</td>
</tr>
<tr>
<td>ER</td>
<td>estrogen receptor</td>
</tr>
<tr>
<td>EZRC</td>
<td>European Zebrafish Resource Centre</td>
</tr>
<tr>
<td>Flp</td>
<td>flipase</td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>HDR</td>
<td>homology-directed repair</td>
</tr>
<tr>
<td>HR</td>
<td>homologous recombination</td>
</tr>
<tr>
<td>hsp70</td>
<td>heat shock protein 70</td>
</tr>
<tr>
<td>ilk</td>
<td>integrin-linked kinase</td>
</tr>
<tr>
<td>IRES</td>
<td>internal ribosome entry site</td>
</tr>
<tr>
<td>ITR</td>
<td>inverted terminal repeats</td>
</tr>
<tr>
<td>KCNH2/HERG</td>
<td>potassium voltage-gated channel subfamily H member 2/human ether-a-go-go-related gene</td>
</tr>
<tr>
<td>kcnh6a</td>
<td>potassium voltage-gated channel subfamily H member 6a</td>
</tr>
<tr>
<td>MMEJ</td>
<td>microhomology-mediated end-joining</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>myh6</td>
<td>myosin heavy polypeptide 6</td>
</tr>
<tr>
<td>myl7</td>
<td>myosin light chain 7</td>
</tr>
<tr>
<td>NHEJ</td>
<td>non-homologous end-joining</td>
</tr>
<tr>
<td>PAM</td>
<td>protospeacer adjacent motif</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>REAL</td>
<td>restriction enzyme and ligation</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RVD</td>
<td>repeat variable diresidue</td>
</tr>
<tr>
<td>sgRNA</td>
<td>single guide RNA</td>
</tr>
<tr>
<td>ssODN</td>
<td>single stranded oligodeoxynucleotide</td>
</tr>
<tr>
<td>TALEN</td>
<td>transcription activator-like effector nuclease</td>
</tr>
<tr>
<td>TILLING</td>
<td>targeting induced local lesions in genomes</td>
</tr>
<tr>
<td>tracrRNA</td>
<td>transactivating CRISPR RNA</td>
</tr>
<tr>
<td>UAS</td>
<td>upstream activating sequence</td>
</tr>
<tr>
<td>vmhc</td>
<td>ventricular myosin heavy chain</td>
</tr>
<tr>
<td>ZFIN</td>
<td>Zebrafish Information Network</td>
</tr>
<tr>
<td>ZFRC</td>
<td>Zebra fish International Resource Centre</td>
</tr>
</tbody>
</table>
for zebrafish, including lines created through large scale genetic screens, followed by descriptions of the most commonly used methods for genetic modifications to create stable transgenic lines (including Tol2 transposon, TALENs [transcription activator-like effector nucleases], and CRISPR/Cas9 [clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9]). These approaches are summarised in Fig. 1. Design considerations including promoters, fluorescent reporters, and conditional transgenesis are also discussed, highlighted by recent examples applied to the heart.

2. Genetically-modified zebrafish resources

While it may ultimately be beneficial for a research laboratory to be capable of generating their own transgenic zebrafish to meet their particular needs, the zebrafish research community is highly collaborative, allowing access to most existing zebrafish lines. This is facilitated by numerous databases (a few of which are mentioned below), with comprehensive lists available in recent reviews (Gut et al., 2017; Varshney et al., 2015). The most prominent of these resources is the Zebrafish Information Network (ZFIN), the main hub for zebrafish genetic information, providing an easily searchable database of genetically-modified strains, plasmids containing engineered sequences of DNA, antibodies, and links to relevant publications. This is complemented by large zebrafish stock centres, including the Zebrafish International Resource Centre (ZIRC), the European Zebrafish Resource Centre (EZRC), and the China Zebrafish Resource Centre (CZRC), from which many genetically-modified zebrafish strains can be ordered.

2.1. Large scale chemical mutagenesis screens

Many zebrafish lines containing mutations of interest have been generated using forward genetic strategies, in which mutations are induced, phenotypes observed, and causative genes later identified. The first such strategy was accomplished in the 1990s through examination of developmental abnormalities in families of offspring from male zebrafish exposed to mutagenic N-ethyl-N-nitrosourea (ENU). This process requires screening the phenotype of thousands of individuals (Driever et al., 1996; Haffter et al., 1997), with the site containing the responsible mutation identified by post hoc positional cloning (involving linking DNA from mutants to known markers in increasingly smaller regions of the genome) (Amsterdam and Hopkins, 2006; Gut et al., 2017). This has led to the identification of many cardiac phenotypes, the most informative being those resulting from mutations that are embryonic lethal in
mice (zebrafish can survive without heart function for ~7 days post fertilisation [dpf]) (Dahme et al., 2009). For instance, positional cloning identified *integrin-linked kinase* (ilk) as the gene responsible for the heart failure zebrafish mutants *main squeeze* (Bendig et al., 2006) and *lost-contact* (Knöll et al., 2007), which improved our understanding of the heart’s ability to sense mechanical stretch and led to the identification of an ILK mutation in human patients with dilated cardiomyopathy. Yet, as this method is extremely laborious, ten years after the initial screens were published only 154 affected genes had been identified (Amsterdam and Hopkins, 2006).

Advancements in genetic techniques have enabled the generation of mutant zebrafish lines using reverse genetic approaches, in which DNA regions of interest can be studied without the prior requirement of an observable phenotype. The first such method successfully applied to zebrafish is known as TILLING (targeting induced local lesions in genomes) and it identifies genes mutated by exposure to chemical mutagens. In its early form, DNA was amplified using polymerase chain reaction (PCR) and then digested with CEL1 endonuclease, an enzyme that specifically cuts DNA at mismatched single nucleotides. Genomic regions found to contain sequence abnormalities were sequenced and individual fish were confirmed as a mutation carrier if their outcross to wild-type zebrafish produced mutant offspring (Amsterdam and Hopkins, 2006; Pan et al., 2015). TILLING is performed at large centres (Moens et al., 2008) and has become relatively high-throughput, using next-generation sequencing to examine multiple DNA regions of interest simultaneously (Pan et al., 2015). Mutant zebrafish lines generated and characterised through chemical mutagenesis and TILLING now number in the thousands. A searchable database is available through the Zebrafish Mutation Project (https://www.sanger.ac.uk/resources/zebrafish/zmp/), whose ultimate goal is to generate a phenotyped knockout zebrafish for each protein coding gene (Kettleborough et al., 2013; Varshney et al., 2015).

### 2.2. Insertional mutagenesis screens using retroviral vectors

Retroviral insertional mutagenesis is an alternative forward genetic strategy used to cause random mutations responsible for observable phenotypes and subsequently identify the affected genomic location. Retroviral vectors are replication-defective viruses that retain the ability to reverse transcribe their RNA into DNA and integrate into a host genome. Specifically for fish, the Moloney murine leukemia virus with an envelope protein of vesicular stomatitis virus has been used for infection (Varshney et al., 2015). Injecting this retroviral vector among the approximately 1000 cells of a zebrafish blastula results in very high rates of retroviral DNA insertion into the genome, with offspring inheriting up to 10 independent mutations. Because the sequence of the introduced DNA is known, identification of the disrupted genomic region is simpler than with positional cloning. After breeding to obtain homozygotes, a mutant phenotype can be linked to the affected gene by identifying DNA on either side of the insertion using PCR-based methods and sequencing (Amsterdam and Hopkins, 2006). Next-generation sequencing has been applied to increase throughput of mutant gene identification, with a few labs aiming to accomplish a similar goal as for chemical mutagenesis: to mutate every protein-coding gene in the zebrafish genome (Varshney et al., 2013). Current progress is compiled in the Zebrafish Insertion Collection database (https://research.nih.gov/zinc).

### 2.3. Gene trap and enhancer trap screens

Gene and enhancer traps are specifically constructed DNA sequences that randomly insert into the zebrafish genome and allow identification of genomic regions responsible for observable phenotypes. Advantages over chemical or retroviral insertional mutagenesis include efficient genome incorporation accomplished by a transposable element and the inclusion of a reporter sequence (typically a fluorscent protein), which can allow visualisation of gene expression changes and therefore recognition of more subtle phenotypes (Auer and Del Bene, 2016; Trinh and Fraser, 2013).

A transposable element (commonly called a ‘transposon’) is a DNA sequence that can change its position within the genome, allowing genomic sequences to be removed and replaced. Specifically, a transposase enzyme recognises and cuts inverted terminal repeats (ITR) on either side of a sequence of DNA. These ITRs are said to flank the DNA sequence and their cleavage by a transposase removes the sequence from its current context and allows it to be inserted elsewhere (outlined in Fig. 2A). Random integration of foreign DNA in zebrafish can therefore be accomplished by injecting one-cell stage embryos with transposase messenger RNA (mRNA) along with a plasmid containing a foreign DNA sequence flanked by the ITRs. A plasmid is a circular double stranded DNA molecule derived from bacteria that holds a DNA sequence to allow its introduction into other organisms. Upon injection, transposase mRNA is translated to protein, cuts the exogenous DNA sequence from its plasmid at the ITRs, and randomly pastes it into the host genome (Clark et al., 2011). Although this is performed at the one-cell stage, not all cells of the resulting zebrafish may incorporate the injected DNA. Integrations that occur in germ line cells, however, can be transmitted to offspring (known as germline transmission) and stable lines of zebrafish containing the new gene can then be maintained (Felker and Mosimann, 2016). Specific procedures for introducing a transgene into zebrafish using transposons are covered later in this review, but their ease of design and construction have allowed the creation of hundreds of stable transgenic zebrafish lines containing a gene or enhancer trap (gene and enhancer trap DNA sequences have also been delivered using retroviruses (Trinh and Fraser, 2013)). Though random insertion and potential post-translational modification necessitate large scale screens to identify and confirm expression of any particular ‘trapped’ gene (Auer and Del Bene, 2016), a number of research groups have compiled characterisation and ordering information in publically accessible databases (Clark et al., 2012; Kawakami et al., 2010) or through stock centres (Varshney et al., 2015).

A gene trap is defined as a constructed DNA sequence lacking a promoter (a region of DNA that initiates transcription of a particular gene) and start codon (the first codon of an mRNA transcript translated by a ribosome), so the reporter is expressed only if its DNA inserts within an expressed gene. Such an insertion disrupts the function of the ‘trapped’ gene, as the gene trap sequence is designed to prevent downstream transcription, but the reporter provides a visual indication of the endogenous gene’s expression (Auer and Del Bene, 2016; Trinh and Fraser, 2013; Weber and Köster, 2013). In a variation of the gene trap known as the protein trap, the reporter is expressed only if the introduced DNA inserts within a gene without causing disruption such that a functional protein fused to the reporter protein is produced, enabling visualisation of normal expression (Trinh and Fraser, 2013). The foreign DNA of an enhancer trap, on the other hand, contains a low-activity promoter followed by a reporter gene, allowing expression to only occur if its DNA inserts upstream of an expressed gene within the region controlled by enhancer elements responsible for that gene’s expression. As such, the translated reporter protein can represent expression of an endogenous gene as controlled by the ‘trapped’ enhancer (Auer and Del Bene, 2016; Trinh and Fraser, 2013).

### 3. Methods for genetically modifying zebrafish

Often zebrafish with a desired genetic modification may not
already exist or existing zebrafish lines may not be generally accessible. However, knowledge of the zebrafish genome and advances in reverse genetic strategies now allow a gene of interest to be modified in nearly any desired way. Individual labs may therefore wish to collaborate with others who possess the necessary technical expertise for generation of mutant zebrafish lines or develop the capabilities themselves. A fundamental requirement for any genetic modification strategy is to have standard microbiology procedures for handling plasmids in place, including their transformation into E. coli, growth of colonies on agar plates containing appropriate antibiotic, propagation of individual colonies in antibiotic-containing liquid media, DNA extraction using mini- or maxiprep protocols, DNA purification using purchased kits or phenol/chloroform and ethanol precipitation, and analysis by agarose gel electrophoresis. Next, the technical ability for micro-injection of one-cell stage zebrafish embryos, the identification of founder and F1 (offspring of founders) positive fish, and the raising of zebrafish to establish stable lines is needed. A wealth of resources exist that describe relevant methods (Auer and Del Bene, 2016; Clark et al., 2011; Felker and Mosimann, 2016; Ma et al., 2016a) and expertise can often be found from other labs or zebrafish breeding facilities within an institution, so details will not be provided here. Publicly accessible resources are also available to assist in designing mutagenesis strategies and obtaining DNA plasmids, which are highlighted in the following sections.

Initial efforts for the generation of transgenic zebrafish involved the injection of embryos solely with plasmids containing the desired foreign DNA (Amsterdam et al., 1995; Culp et al., 1991; Stuart et al., 1990, 1988; Udvarda and Linney, 2003). Transgenic efficiency by this approach was improved by linearisation of plasmid DNA (Thermes et al., 2002), yet it remained hindered by cumbersome methods for plasmid construction and relatively low efficiency of transgene insertion and germline transmission. Improvement in the ease of plasmid design and the efficacy of gene insertion have been accomplished by transposon-mediated and targeted genome editing approaches, described below.

### 3.1. Introduction of exogenous DNA using Tol2 transposons

Soon after the discovery of Tol2 as a transposon active in fish (Koga et al., 1996), it was shown to enable integration of foreign DNA excised from an injected plasmid into the zebrafish genome, creating offspring with inherited insertions (Kawakami et al., 2000). Later, it was discovered that a promoter sequence known to drive gene expression in a particular tissue could be placed upstream of a reporter gene in a Tol2-compatible plasmid to restrict reporter expression to a tissue of interest, despite unpredictable integration location within the zebrafish genome (Kawakami et al., 2004). While other transposons have been used for gene insertion in zebrafish, including the synthetic Sleeping Beauty system (Davidson, 2003) and plant-derived Ac/Ds (Emelyanov et al., 2006), Tol2 is by far the most commonly used method of transposon-mediated transgenesis and has the most resources developed for its implementation (Clark et al., 2011; Felker and Mosimann, 2016).

The general method of generating transgenic zebrafish using the Tol2 transposon system will be described here to aid researchers in understanding the principal steps involved (more detailed protocols can be found in Chen et al., 2016, Clark et al., 2011, and Kawakami et al., 2016). Briefly, the first step in Tol2-mediated transgenesis is to design and generate an expression plasmid containing the relevant DNA sequences in the correct orientation and flanked by ITRs specific to the Tol2 transposase. The components required for successful protein expression of an introduced gene have been well summarised by others (Felker and Mosimann, 2016). To accompany this, Tol2 transposase must be prepared. Transposase mRNA (generated in vitro by linearisation of a plasmid via treatment with a restriction enzyme followed by reverse transcription) is most often recommended because once in vivo both translation into functional enzyme and degradation of mRNA are rapid, which favours integration at an early stage while avoiding subsequent rearrangements (Clark et al., 2011). Using transposase in the form of a DNA fragment within a plasmid directly is another common strategy and Tol2 transposase protein has also recently been shown to be effective for the generation of transgenic zebrafish (Ni et al., 2016). Once prepared, expression plasmid DNA and transposase are injected into one-cell stage embryos. Expression plasmids have traditionally been generated using molecular cloning and restriction enzymes to insert desired DNA sequences into Tol2 transposon-compatible plasmids. Others have chosen to use the Gateway system for plasmid construction (Chen et al., 2016; Clark et al., 2011; Kawakami et al., 2016), due to its modular nature and amenability to users lacking extensive molecular cloning expertise (Villefranc et al., 2007).

#### 3.1.1. Gateway cloning

While conventional cloning uses restriction enzyme sites within plasmids to allow insertion of foreign DNA sequences, Gateway...
cloning is based instead on DNA recombination. Specifically, it is based on reactions between E. coli and an infecting virus, in which a short DNA sequence in the virus (known as the attP site) recombinates with the E. coli DNA (at the attB site) resulting in insertion of the virus DNA into the E. coli genome (flanked by attL and attR sites) (Wallhout et al., 2000). These naturally occurring attP, attB, attL, and attR sequences have been modified to expand their functionality, allowing them to be used to cut and paste DNA fragments among various plasmids, while maintaining the correct DNA orientation and reading frame (Cheo, 2004; Hartley et al., 2000). As such, combining different plasmids in the presence of particular enzymes in vitro allows recombination between specific sites to occur, resulting in flanked DNA fragments to be removed from their initial plasmid and inserted into others (Hartley et al., 2000).

Overall, the Gateway cloning system (outlined in Fig. 3) consists of four types of plasmids: donor, entry, destination, and expression. These plasmids are recombined by two reactions: BP (occurring between the attB and attP sites), in which a DNA sequence of interest is inserted into a donor plasmid to create an entry plasmid, and LR (occurring between the attL and attR sites), in which the entry plasmid and a destination plasmid are combined to create an expression plasmid for embryo injection (Hartley et al., 2000).

For the BP reaction, first the DNA sequence of interest is amplified by PCR from genomic DNA using primers that add attB1 and attB2 sequences to the 5' and 3' end of the amplified fragment. This product is then mixed with a donor plasmid containing the ccdB gene (which is toxic to most strains of E. coli, hence must be grown in the DB3.1 strain) and the cmr (chloramphenicol resistance) gene, flanked by attP1 and attP2 sequences. The addition of BP clonase to this mixture causes attB1 and attB2 to recombine with attP1 and attP2, resulting in replacement of the ccdB/cmr gene in the donor plasmid with the desired DNA sequence (in the correct orientation and flanked by attL1 and attL2 sites). The resulting entry plasmid is then transformed into a regular E. coli strain, followed by standard procedures for propagation and plasmid DNA extraction (Hartley et al., 2000; Wallhout et al., 2000).

In the following LR reaction, a destination plasmid containing the ccdB/cmr gene flanked by attR1 and attR2 sequences is linearised by enzymatic digestion and then combined with the entry plasmid DNA and LR clonase. Similar to the BP reaction, attL1 and attL2 then recombine with attR1 and attR2 resulting in the ccdB/cmr sequence being replaced in the destination plasmid by the DNA fragment of interest from the entry plasmid (Hartley et al., 2000; Wallhout et al., 2000). The resulting expression plasmid (containing the DNA sequence of interest flanked by attB1 and attB2 sites) is propagated in standard E. coli, with its expected sequence confirmed by restriction digest analysis and sequencing (as can also be done for the entry plasmid) (Kwan et al., 2007).

While the above approach can be used in general for transgenesis in multiple species, generation of an injectable expression plasmid for zebrafish by Tol2 Gateway cloning is often performed using an expanded system (the MultiSite Gateway® Three Fragment Vector Construction Kit from ThermoFisher Scientific) in which additional versions of the recombination sites allow three different, specifically designed entry plasmids to be inserted into a destination plasmid in a predefined order in the correct orientation using a single LR reaction (Kwan et al., 2007; Villefranc et al., 2007). The first (5') entry plasmid is typically used to provide enhancer-promoter elements that drive the expression of the gene of interest, and can be ubiquitous, tissue-specific, or conditional. The second (middle) entry plasmid contains the sequence of the gene to be introduced into the genome followed by a stop codon, which may be preceded by the sequence of a fluorescent or Myc tag (antibody.recognising) reporter protein (creating a fusion protein that will allow visualisation of expression). The third (3') entry plasmid contains a 3' polyA signal that stabilises mRNA and promotes translation. If visualisation is desired but a reporter gene was not included in the middle entry plasmid, it can instead be included in the 3' entry plasmid, resulting in a similar fusion protein (if effects on function resulting from protein fusion are a concern, the 3' entry plasmid can include an internal ribosome entry site (IRES) sequence upstream of the reporter sequence, which results in the gene of interest and reporter being translated as two separate proteins) (Weber and Köster, 2013).

**Fig. 3.** Gateway cloning. See Section 3.1.1 for details.
Fortunately, the zebrafish community benefits from a large collection of Multisite Gateway®-compatible entry and destination plasmids (available from ZIRC. Addgene, or directly from the labs who developed them), which contain the necessary recombination sites and are designed to be combined with Tol2 transgenesis, allowing the BP reaction step to be skipped. A single LR reaction combines the ready-made entry plasmids into a destination plasmid to generate a Tol2 ITR-containing expression plasmid to be injected with transposase into embryos (Kwan et al., 2007; Villefranc et al., 2007). Overall, Tol2-mediated transgenesis combined with the Gateway cloning system has maximised accessibility, ease of use, and efficiency of generating transposon-mediated genetically-modified zebrafish and is commonly used alongside or in combination with newer transgenic technologies.

### 3.2. Targeted gene editing

Although Tol2-mediated transgenesis is a powerful method for the generation of mutant zebrafish, it is limited by the fact that it results in the random insertion of genetic sequences into the genome. Alternative methods that allow genetic modification to be targeted to genomic locations are rapidly progressing and have already greatly expanded the potential for gene knock-out and tagging, and the introduction of disease-relevant single nucleotide mutations. These newer technologies, generally called engineered endonucleases, are based on the precise cutting of DNA, and the two most common are amenable to implementation by non-expert labs (TALENs and CRISPR/Cas9).

As with transposon-mediated transgenesis, targeted gene editing approaches involve injection into one-cell stage zebrafish embryos, however rather than random insertion, they utilise a protein or RNA molecule designed to recognise a specific DNA sequence and an associated nuclease to cut the DNA at a specific site (Varshney et al., 2015). The resulting double-strand break (DSB) in DNA is repaired by endogenous mechanisms, and these natural processes can be harnessed to dictate the functional consequence of repair, allowing the creation of either gene knock-out or knock-in animals. In general, gene knock-out is accomplished by injecting a DNA recognition molecule and the associated nuclease, resulting in DNA repair that introduces random mutations and renders the targeted gene incapable of producing a functional protein. Gene knock-in, on the other hand, is created by also including exogenous DNA in the injection mix, allowing it to be incorporated into the targeted site during repair (Hoshijima et al., 2016a; Morita et al., 2017). DNA repair mechanisms responsible for sequence modifications are described below, followed by specific principles of TALENs, CRISPR/Cas9, and their common considerations.

The first method developed for targeted genome editing utilised zinc-finger nucleases (ZFNs), which bind to specific 3-nucleotide sequences. Briefly, in this approach two strings of ZFNs, each designed to recognise one of the strands of the DNA at a site of interest (Gut et al., 2017), are each fused to the catalytic domain of Fold nuclease (an enzyme which itself cannot discriminate particular DNA sequences, but can induce a DSB when dimerised). When the two ZFN strings come together at the targeted site, the DNA is precisely cut (Joung and Sander, 2013; Morita et al., 2017). While there has been progress in making ZFN design and construction more user-friendly, as well as the development of new versions of Fold that increase the specificity of DSB induction (Campbell et al., 2013; Hoshijima et al., 2016b), this method has been surpassed in applicability by both TALENs and CRISPR/Cas9.

#### 3.2.1. Role of DNA double strand breaks and repair mechanisms in gene editing

DSBs occur naturally and must be repaired to prevent genmic instability and subsequent cellular dysfunction (Thoms et al., 2007). There are three well-characterised mechanisms for repairing DSBs: non-homologous end-joining (NHEJ), microhomology-mediated end joining (MMEJ), and homology-directed repair (HDR), all of which are functional in embryonic zebrafish (Liu et al., 2012; Morita et al., 2017).

NHEJ is the least precise DSB repair strategy and is used in gene knock-out approaches by inducing a DSB without providing a template for repair. As the broken strands are reattached, the original sequence is lost as nucleotide insertions and deletions occur. These mutations frequently result in a frameshift wherein the reading frame of codons is shifted, so the translated gene product does not encode the correct sequence of amino acids (Albadri et al., 2017; Joung and Sander, 2013; Morita et al., 2017; Sander and Joung, 2014). It is possible for NHEJ to also be used for the generation of knock-in animals if DNA is included in the injection mix, but as the exogenous and targeted DNA will in most cases not contain complementary sequences, the introduced DNA will contain mutations at each end where it attaches to endogenous DNA (Albadri et al., 2017; Auer and Del Bene, 2016; Morita et al., 2017).

MMEJ occurs in cases where the DSB exposes sequences of 3–30 nucleotides on both strands of DNA that are complementary to each other, known as areas of microhomology. These short homologous sequences bind together and fill-in the remaining sequence to mend the DSB, resulting in a predictable loss of nucleotides (Hoshijima et al., 2016a; Prykhozhij et al., 2016). MMEJ can be mobilised in gene editing without the provision of exogenous DNA if the induced DSB is targeted to a genomic location containing microhomology (Hoshijima et al., 2016a). If the goal of gene editing is knock-out, regions conducive to MMEJ should be avoided to increase the likelihood of inducing a mutation that results in a non-functional protein (Prykhozhij et al., 2016). MMEJ can be harnessed for gene knock-in by introducing exogenous DNA that shares short sequences (20–40 nucleotides) of homology with the targeted site of the genome (Morita et al., 2017).

HDR is the most precise DSB repair strategy as it uses long sequences of complementary DNA as a template to replicate the lost sequence when a DSB occurs (Albadri et al., 2017; Hoshijima et al., 2016a; Morita et al., 2017; Sander and Joung, 2014). The exogenous DNA can be supplied as a single stranded oligodeoxynucleotide (ssODN) of less than 100 nucleotides or as a double stranded DNA (dsDNA) molecule typically contained within a plasmid (Campbell et al., 2013; Hoshijima et al., 2016a). Because plasmids can contain DNA sequences of up to several thousand nucleotides, they are generally used to introduce sequences of full (or even multiple) genes. While ssODNs induce a form of HDR that is not well characterised, plasmids designed to include a DNA sequence of 400–2000 nucleotides in common with the targeted region (known as a homology arm) trigger repair of DSBs via homologous recombination (HR), a well understood process that naturally occurs between homologous sister chromatids during cell division, thus facilitating the most precise possible method of genome editing (Campbell et al., 2013; Hoshijima et al., 2016a; Morita et al., 2017). A plethora of genetic modifications are possible by engineered endonucleases harnessing these DNA repair mechanisms with or without introduction of foreign DNA and examples of their application will be highlighted in following sections.

#### 3.2.2. Targeted gene editing using TALENs

TALENs are protein-based and are comprised of TALE domains and the catalytic domain of the FokI nuclease. TALE proteins were originally discovered as transcription activators in a bacterium pathogenic to plants and consist of multiple 33–35 amino acid-containing TALE domains. Positions 12 and 13 are called a repeat
variable diresidue (RVD) and particular pairs of amino acids making up the RVD bind specifically to one of the four DNA nucleotides. Therefore, sequences of TALE domains can be assembled to recognise longer DNA sequences, with their practical use typically requiring a sequence of 15–21 TALE domains. Because FokI is used as the nuclease to generate DSBs, similar to ZFNs, two TALEN monomers must be targeted to the same genomic region, complementary to the forward and reverse strands respectively, with a spacer of 13–20 base pairs between the two recognition sequences such that the DSB is usually induced in the centre of the spacer when the two domains of FokI dimerise (Campbell et al., 2013; Doyle et al., 2012; Gut et al., 2017; Ma et al., 2016a; Morita et al., 2017). It is often desirable for the spacer sequence to contain a unique restriction site whose sequence will be lost during DSB and subsequent repair, so that extracted DNA can quickly be analysed using restriction digest to determine whether genomic modification was successful (Doyle et al., 2012; Ma et al., 2016a). When pursuing a TALEN-mediated gene editing strategy (outlined in Fig. 2B), a gene or genomic region must be chosen for modification followed by TALEN design. Various open-access web-based programs are available to help determine the specific region to target and the sequence of TALE domains required for recognition and subsequent DSB induction. After inspecting a sequence of interest, the software uses customisable constraints and recommended guidelines for the sequence and length of TALEN recognition sites and the spacer sequence to suggest candidates (Auer and Del Bene, 2016; Cermak et al., 2011). Commonly referenced design tools include ZIFIT Targeter (https://zifit.partners.org/ZIFIT/), TAL Effector Nucleotide Targeter 2.0 (https://tale-nt.cac.cornell.edu/), Mojo Hand 2.0 (https://www.talendesign.org/), and CHOPCHOP (https://chopchop.cbu.uibk.no/). Most offer interpretation of outputs consisting of candidate target sequences, their location within the genomic region, and the corresponding RVD sequence for each TALEN monomer. Other outputs may include recognition sites for restriction enzymes included within the spacer sequence, primer sites on either side of the spacer sequence, an evaluation of potential recognition of unintended sequences, and an overall score to rank options (Doyle et al., 2012; Montaguel et al., 2014; Neff et al., 2013).

Once the required TALENs are determined, the sequence of TALE domains must be assembled. Although commercial synthesis is available, methods have been developed to simplify the process and reduce the time required for in-house assembly (Gut et al., 2017), which have been nicely summarised in recent reviews (Campbell et al., 2013; Joung and Sander, 2013). Considerations in choosing an assembly method include scale of production desired, the skills and equipment available to the researcher, and the characteristics of the TALEN scaffold to be used (which is contained within a plasmid and includes all regions of the TALEN protein other than the TALE domains). Scaffolds used in TALEN assembly systems for zebrafish all have a nuclear localisation sequence and have been shown to allow germline transmission. Variations occur in the promoter, epitope tag, and N- and C-termini truncations, as well as whether the FokI catalytic domain is homodimeric or heterodimeric (multiple options may exist for a given application, however the optimal method for most circumstances is still unknown). Various kits, plasmids, and protocols can be obtained for TALE assembly from Addgene or by directly contacting individual labs. One method of TALE assembly involves the use of standard restriction enzyme and ligation (often referred to as REAL). This process comprises individually cutting TALE domains from existing plasmids using several types of restriction enzymes that generate characteristic overhangs and then pasting them into progressively longer ordered sequences in other plasmids. It typically takes at least 9 days to build a TALEN containing 13–17 TALE domains by this procedure (Campbell et al., 2013; Joung and Sander, 2013). An alternative method with more reasonable time, reagent, and molecular cloning skill requirements uses the Golden Gate TALEN and TAL Effector Kit (Cermak et al., 2011), which allows TALENs of 12–21 TALE domains to be generated in 3–5 days and allows the use of a variety of scaffolds (Campbell et al., 2013; Joung and Sander, 2013). Though still relying on restriction digest and ligation and requiring storage of a set of 50 module plasmids, the Golden Gate TALEN system is simplified by the use of only two restriction enzymes and a multi-fragment ligation that reduces the number of required reactions (Cermak et al., 2011). A recent update to the Golden Gate TALEN system, the FusX TALEN Assembly System, has been shown to have a germline transmission rate of 45% in zebrafish. This new method accomplishes TALEN monomer construction within 3 days through a single reaction by utilising module plasmids containing TALE domain triplets in all possible combinations, but requires obtaining and storing hundreds of plasmids for full flexibility (Ma et al., 2016a, 2016b).

Once a plasmid is confirmed to contain the required TALEN sequence (by restriction digest and often sequencing), mRNA is transcribed in vitro, purified, and then injected to induce a DSB at the desired genomic site (Ma et al., 2016a). The decision of whether and how to introduce exogenous DNA for TALEN-mediated gene insertion will be discussed after introducing CRISPR/Cas9, as the same considerations apply to both gene editing strategies.

3.2.2. Targeted gene editing using CRISPR/Cas9

CRISPR/Cas systems are found in many species of archaea and bacteria serving as an adaptive immune strategy to degrade foreign DNA, with Streptococcus pyogenes providing the spCas9 protein most commonly used for gene editing. In the naturally-occurring form, CRISPR RNA (crRNA) and transactivating CRISPR RNA (tracrRNA) are transcribed from the bacteria’s genome and come together in an RNA structure with an exposed 20-nucleotide spacer sequence. Cas9 is recruited and the resulting RNA-protein complex can interact with a DNA sequence complementary to the spacer sequence if it is immediately upstream of a particular 3-nucleotide sequence known as a protospacer adjacent motif (PAM). The PAM, which for spCas9 is 5‘-NGG-3’ (where N can be any of the four nucleotides and G is guanine), is opposite the strand complementary to the spacer sequence and is strictly required for CRISPR/Cas9 to execute its function. Once an RNA-Cas9-DNA complex is formed, the two nuclease domains of Cas9 each cleave one strand of DNA, resulting in a DSB within the 20-nucleotide region recognised by the spacer (Gut et al., 2017; Morita et al., 2017; Prykhozhij et al., 2017; Sander and Joung, 2014).

The CRISPR/Cas9 system has been modified for gene editing (outlined in Fig. 2C) by combining aspects of the crRNA and tracrRNA into a single guide RNA (sgRNA). The sgRNA is designed to contain a sequence derived from crRNA and tracrRNA that is the same for all CRISPR/Cas9 applications in addition to about 20 nucleotides complementary to a targeted genomic region adjacent to a PAM (equivalent to the spacer sequence). Injecting the sgRNA along with Cas9 into a one-cell stage zebrafish embryo allows induction of a DSB 3 nucleotides upstream of the PAM in the targeted DNA sequence (Auer and Del Bene, 2016; Gut et al., 2017; Sander and Joung, 2014). The option to include exogenous DNA in the injection mix will be discussed in a later section.

Many resources exist to simplify the construction of sgRNA and Cas9 to facilitate CRISPR/Cas9 gene editing. A recently published tool guide (Prykhozhij et al., 2016) is an excellent starting point for identifying optimal software and appropriate protocols. In general, one must first choose the genomic region to target for DSB (such as a specific gene) along with the desired modification, followed by identification of a candidate sequence of DNA within the targeted
region that can be recognised by an sgRNA (i.e., a unique 20-nucleotide sequence immediately upstream of a PAM). Though software can be used to accomplish this, it is helpful to have an understanding of sgRNA design since it is a key feature of any CRISPR/Cas9 gene editing strategy. Important considerations can include a G nucleotide preceding the target sequence for optimal sgRNA binding (Auer and Del Bene, 2016; Gagnon et al., 2014; Hwang et al., 2013a; Sander and Joung, 2014) and a recognition site for a restriction enzyme within the target sequence to allow for post hoc characterisation of the successfully induced genetic modification (Prykhozhij et al., 2016). The form of Cas9 to be used, whether multiple sgRNAs will be employed, and potential effects on unintended genomic sites will be addressed later in this section.

An alternative for determining an appropriate target site in the zebrafish genome is to consult relevant publications or the searchable CRISPR2 database [https://research.nhgri.nih.gov/CRISPR2/] to find sites previously targeted by CRISPR/Cas9. CRISPR2 can be particularly helpful as, if a validated target site exists (the database currently comprises 1417 target sites), the target and PAM sequence, form of Cas9, genotyping primers used, success rate of inducing somatic and germline mutations, and links to published references are provided (Varshney et al., 2016). Once a target sequence is determined, an sgRNA unique to that region must be generated. Perhaps the greatest advantage of CRISPR/Cas9 over TALENs is that the desired sgRNA sequence can be ordered as a DNA oligonucleotide (Gut et al., 2017; Hwang et al., 2013b) from many commercial suppliers for a relatively low price. Additionally, a cloning-free method of sgRNA generation exists (Gagnon et al., 2014), so the only requirements for assembling a functional sgRNA are access to a thermocycler and the ordering of basic molecular biology reagents and two oligonucleotides. One of these oligonucleotides is the same for all CRISPR/Cas9 applications, consisting of an 80-nucleotide sequence that forms the modified crRNA:tracrRNA structure optimal for in vivo binding of Cas9. The other oligonucleotide is unique to each application, containing the 17-20-nucleotide gene-specific recognition sequence flanked by a sequence required for in vitro transcription and a sequence that overlaps with the first oligonucleotide. Briefly, the method involves annealing the two oligonucleotides together, treating them to form dsDNA, reverse transcribing the DNA into RNA, and purifying the RNA so it is injection ready. Alternatively, synthetic sgRNA can be purchased for a reasonable cost and has the theoretical advantage of avoiding the addition of extra nucleotides during in vitro transcription (although little improvement in performance of synthetic sgRNA has been reported) (Prykhozhij et al., 2017). Of note, depending on the software used for design, the target site may be given as the recognition site with or without inclusion of the necessary upstream G nucleotide(s) and downstream PAM sequence. When ordering the oligonucleotide specific to the sgRNA recognition sequence, the G nucleotide(s) must be included (Auer and Del Bene, 2016), while the three PAM nucleotides (NGG if using spCas9), must be excluded (Gagnon et al., 2014). Also, while design tools are extremely helpful, they do not reliably predict in vivo sgRNA performance (Prykhozhij et al., 2016). It is therefore recommended that multiple sgRNAs be designed and tested for each CRISPR/Cas9 gene editing application, which is reasonable considering the relative ease and low cost of ordering and construction (Albadri et al., 2017; Auer and Del Bene, 2016).

Regarding the form of Cas9, it can be included in the injection mix as either mRNA or protein. Although some report decreased success of exogenous mRNA insertion with Cas9 protein compared to mRNA (albeit with increased embryo survival and DSB induction) (Albadri et al., 2017), others have demonstrated impressive insertion rates with Cas9 protein (Hoshijima et al., 2016b). Further studies are needed to determine the optimal form of Cas9, but the ease of in vitro mRNA transcription versus protein synthesis favours the use of mRNA for most labs (Prykhozhij et al., 2017). If Cas9 mRNA is to be used, a plasmid containing Cas9 is needed followed by plasmid linearisation, in vitro transcription, and purification (Gagnon et al., 2014). A wide variety of Cas9 plasmids exist (many of which are available from Addgene) and the choice is typically determined by researcher preference and specific experimental needs, as all options should include a nuclear localisation signal and be compatible with any sgRNA (Auer and Del Bene, 2016). Although spCas9 remains the most commonly used, smaller or better performing versions of Cas9 are available, along with different PAM requirements to increase possible target sites (Prykhozhij et al., 2016). If instead Cas9 protein is to be used, it can be purchased from a commercial vendor or produced in-house, if appropriate technical expertise are available (Gagnon et al., 2014).

A major limitation of CRISPR/Cas9 versus TALENs is the increased likelihood of inducing a DSB in an unintended region of the genome due to sgRNA recognition of non-exact DNA sequence complements or an alternative PAM sequence (Gut et al., 2017; Sander and Joung, 2014). Many of the freely available software design tools consider potential off-target sites of sgRNAs as they search for target sites in the zebrafish genome, offering a score of possible sgRNAs to indicate their exclusivity. Other tools can be used to predict potential off-target sites for a given sgRNA and such sites can be assessed for mutation after sgRNA and Cas9 injection by comparing to non-injected controls (Auer and Del Bene, 2016; Prykhozhij et al., 2016). Off-target DSBs are less likely with the use of TALENs because of the requirement for TALEN pairs to create a functional FokI nuclease. A similar strategy has been applied to CRISPR/Cas9 by fusing catalytically inactive Cas9 (dCas9) to FokI, creating the need for two sgRNAs targeting the same region of DNA (complementary to the forward and reverse strands respectively) to bring two FokI domains together to form a functional dimer to cleave the dsDNA. Alternatively, the use of Cas9 nickases, which are mutated forms of Cas9 that cleave only a single strand of DNA and thus necessitate two sgRNAs targeting the same region of DNA for a DSB to occur, have been demonstrated, but there is concern of single strand DNA damage with this method (Sander and Joung, 2014) and its utility has not been proven in zebrafish (Prykhozhij et al., 2016). Yet, while DSBs occurring in unintended regions of the genome is a legitimate concern, their effects can be mitigated by confirming successful gene editing with various molecular techniques, including sequencing, breeding out off-target effects in distant genomic regions through multiple rounds of outcrossing (Gut et al., 2017; Prykhozhij et al., 2017), and performing experiments in more than one line established from separate founders (Hwang et al., 2013a).

An additional advantage of CRISPR/Cas9 over TALENs is the potential for targeting multiple genomic regions in one injection (Gut et al., 2017), which can be especially useful when trying to inactivate the often multiple versions of a zebrafish gene analogous to a single human gene (true for ~24% of human genes (Howe et al., 2013)) or multiple genes thought to function in a similar process or pathway to study cumulative effects without performing multiple injections or additional breeding (Shah et al., 2015). CRISPRseek [www.bioconductor.org/packages/release/bioc/html/CRISPRseek.html] and CRISPR MultiTargeter [https://multicrispr.net/] are design tools that allow simultaneous design of sgRNAs for more than one target site (Prykhozhij et al., 2016).

3.2.4. Considerations common to the use of TALENs and CRISPR/Cas9

Many of the considerations for the use of TALENs and CRISPR/Cas9 for targeted gene editing are similar, as both are used to recognise precise genomic locations for DSB induction. The first
step is to choose the genomic site to target, which is crucial for obtaining the desired result and for its interpretation. If attempting gene knock-out, causing a DSB in an early exon often causes the greatest damage by shifting the entire downstream sequence out of frame, thus increasing the likelihood of producing a non-functional protein (Campbell et al., 2013; Morita et al., 2017). However, it is generally best to target a region after the start codon or in a domain key to the protein’s function (Gut et al., 2017; Ma et al., 2016a; Prykhozhij et al., 2016). Two pairs of TALENs or sgRNAs can be employed to induce large deletions by targeting two sites on the same chromosome, or translocations by targeting sites on different chromosomes (Joung and Sander, 2013; Ma et al., 2016a). If instead attempting gene knock-in (such as a fluorescent reporter), targeting insertion near the stop codon of the endogenous gene of interest will link their expression (Kimura et al., 2014; Li et al., 2015; Shin et al., 2014). Alternatively, exogenous DNA can be targeted upstream of the endogenous gene’s start codon (Kimura et al., 2014; Ota et al., 2016) or within an intron (Hoshijima et al., 2016b; Li et al., 2015). Introducing foreign DNA can also be used as a more controlled method of gene knock-out than DSB alone, by including a stop codon in every frame of the introduced sequence, which will result in a truncated, non-functional protein regardless of reading frame (Sagnon et al., 2014).

Once the genomic region to target is identified, the surrounding genetic sequence must be determined and entered into a TALEN or sgRNA design tool (Hoshijima et al., 2016a). If a lab can implement both TALEN and CRISPR/Cas9 approaches, the decision of which method to use will be based on which will provide the best target site recognition. CRISPR/Cas9 is often favoured due to the ease of sgRNA construction, but TALENs can target sites not possible with sgRNA, as their only requirement is that there is a T nucleotide immediately preceding the site of TALEN binding in the targeted genomic sequence (Campbell et al., 2013). sgRNAs, on the other hand, work best when there is a G nucleotide upstream of the recognition sequence (Auer and Del Bene, 2016; Gannong et al., 2014; Sander and Joung, 2014) and require a downstream PAM sequence (Gut et al., 2017; Morita et al., 2017; Prykhozhij et al., 2017; Sander and Joung, 2014). The effectiveness of TALENs, however, is reduced in methylated regions and it is recommended to avoid targeting sites that contain more than one cytosine nucleotide directly followed by a guanine nucleotide (Chen et al., 2013).

For any targeted gene editing method used in zebrafish, one must also consider the potential for polymorphisms in the particular strain being used, as individual differences in nucleotides in the targeted region can affect sgRNA or TALEN sequence recognition and subsequent DSB and repair. This is particularly important if targeting an intron, as they generally harbour more polymorphisms than exons. To check for polymorphisms, it is recommended to PCR amplify and sequence 300–500 nucleotides of DNA surrounding the intended site of modification, so that the precise target sequence can be incorporated into the TALE domain or sgRNA design (Albadri et al., 2017; Auer and Del Bene, 2016; Hoshijima et al., 2016a; Ma et al., 2016b).

After designing and assembling the required TALEN or sgRNA, if the goal is to knock-in an exogenous DNA sequence, an ssODN or plasmid containing the DNA sequence to be inserted is also required. When designing the ssODN or plasmid, it is ideal to include a mutation that will prevent recognition and cleavage by the TALENs or sgRNA/Cas9 and form a recognition site for a restriction digest enzyme. This design feature helps ensure an additional DSB is not generated by the targeted nuclease once the new sequence has been inserted into the host genome (Doyle et al., 2012; Hoshijima et al., 2016a; Prykhozhij et al., 2016) and allows restriction digest analysis of the region of interest to confirm insertion (Albadri et al., 2017; Ma et al., 2016a). If insertion is targeted to an exon of an endogenous gene and a functional protein is required, the introduced mutation must result in a different codon that specifies the same amino acid to ensure no disruption of expression (Hoshijima et al., 2016a; Prykhozhij et al., 2016).

ssODNs are designed to have a sequence complementary to the targeted genomic region, other than the included mutation(s) (Ma et al., 2016a; Prykhozhij et al., 2016). They can be ordered like any oligonucleotide (Gagnon et al., 2014), but it may be necessary to test multiple versions to find the best configuration for high efficiency editing (Hwang et al., 2013a). Although there is some debate, evidence from CRISPR/Cas9 experiments in other species suggests that any mutations should not be in the centre of the ssODN (Prykhozhij et al., 2017).

A plasmid must be used if the desired insertion is sufficiently large (i.e., one or more complete genes). Though plasmids containing genes of interest are often available from other labs or databases (such as Addgene), it may be necessary to modify existing plasmids using conventional cloning techniques to obtain the particular sequence one wants to introduce. Plasmids consist of a backbone, which is a sequence that will not be incorporated into the host genome, as well as multiple restriction enzyme sites that allow integration of foreign DNA that can later be cut from the plasmid and inserted into the host genome. Thus, a gene of interest from another plasmid or from extracted DNA can be PCR amplified using primers that add unique recognition sites to the ends of the fragment and then inserted into the desired plasmid (with successful plasmid construction confirmed by sequencing) (Hoshijima et al., 2016a). If the goal is HR-mediated gene editing, the plasmid additionally requires inclusion of long homology arms flanking the exogenous DNA. One kilobase of homologous sequence downstream and upstream of the insertion site (obtained by PCR amplifying the region being targeted in the strain of zebrafish to be injected) has been shown to allow efficient precise insertion (Prykhozhij et al., 2016) and is a manageable length for later PCR and sequencing to confirm incorporation (Albadri et al., 2017), although longer sequences may improve success rate (Shin et al., 2014; Zu et al., 2013). Success of such an approach can be improved by targeting an intron as the insertion site, as exogenous gene expression will occur even if imprecise incorporation at the junctions between foreign and host DNA occurs (i.e., mutations) (Albadri et al., 2017; Hoshijima et al., 2016a).

A consensus does not yet exist on whether plasmid DNA injected into zebrafish should be linearised or left circularised. While increased efficiency has been demonstrated using plasmids linearised in vitro prior to injection (particularly when cut within the shorter of the two homology arms) (Shin et al., 2014), other groups have had success injecting circularised plasmids that include sites to allow linearisation to occur in vivo via an sgRNA or restriction digest enzyme (Auer et al., 2014a; Hoshijima et al., 2016b; Iron et al., 2014; Kimura et al., 2014). There has been some suggestion that circularised plasmids are less toxic than linearised plasmids when injected into zebrafish embryos, allowing them to be introduced at a higher concentration and therefore increasing the likelihood of exogenous gene incorporation (Auer et al., 2014a; Iron et al., 2014).

When all components of an injection mix are ready, an initial round of injections followed by a cleavage assay on 1 dpf embryo DNA should be performed to assess the rate of mutation induction. Specifically, DNA amplified from the targeted region can be checked using restriction digest, high resolution melt analysis, endonuclease mismatch detection assays (Surveyor or T7), or sequencing. Engineered endonucleases that cause mutations in greater than 70% of injected embryos are ideal for generating stable lines, though a minimum mutation rate of 30–40% is recommended (Albadri et al., 2017; Hoshijima et al., 2016a; Prykhozhij et al., 2016;
Promoters have been found that restrict expression of an inserted gene in unexpected cells and conditions (Prykhozhij et al., 2017). Reliable expression of the downstream gene is necessary. It is important to note that injected reagents take time to induce DSBs and repair, while cell division of a zebrafish embryo occurs rapidly (Shah et al., 2015), therefore not all injected fish will contain mutations in every cell (and only those with the mutation in germ cells will transmit the genetic modification to offspring for establishment of a stable line) (Varshney et al., 2015). By 7 dpf, molecular analysis of the targeted region of the genome in a subset of injected fish should be performed by extracting DNA and performing PCR (of either the entire modified region or the junctions where exogenous and endogenous DNA meet) and agarose gel electrophoresis with or without restriction enzyme digest, ideally confirmed by sequencing and Southern blots. Molecular confirmation must also be obtained in the outcrossed offspring of the injected fish (Albadri et al., 2017; Auer and Del Bene, 2016; Prykhozhij et al., 2016; Shin et al., 2014; Zu et al., 2013). Screening for fluorescence, if applicable, should be performed in all injected fish within 7 dpf, remembering that germline transmission is more likely in fish exhibiting strong reporter expression, but is possible in fish that do not (if the strength was incorporated in the germ but not somatic cells) (Albadri et al., 2017; Hisano et al., 2015; Kimura et al., 2014; Shin et al., 2014). If the gene editing strategy does not involve expression of a protein useful for visual screening, some suggest inclusion of mRNA of a fluorescent protein in the injection mix to transiently indicate successfully injected embryos and therefore assist in prioritising those to raise.

4. Specific considerations for design of genetic modification approaches

Inducing a global change in gene expression using genetic modification techniques has proven to be an incredibly powerful tool for determining a gene’s function. Alternative strategies, however, can allow for visualisation and manipulation of gene expression at specific time points and in distinct cell types. These refined approaches avoid unwanted secondary effects and are often more relevant for the modelling of human disease (Hoshijima et al., 2016a; Prykhozhij et al., 2017).

4.1. Spatial control

A common goal in the refinement of gene editing strategies is the introduction of spatial control of modified gene expression by use of an endogenous or inserted promoter that drives expression exclusively in particular cell types. Upstream of every gene’s start codon is a regulatory region that controls where and under what conditions that gene is transcribed. Although some genes are ubiquitously expressed in all cell types, many genes are transcribed exclusively in particular cells, as their upstream promoter sequence is responsive to a unique combination of transcription factors. Importantly, promoter regions can be separated from the endogenous gene whose expression they normally promote, and another gene can be inserted downstream, so that the promoter drives the new gene’s expression. The choice of promoter is crucial when aiming for cell-specific expression and protocols exist to identify the sequence upstream of a coding gene that controls its particular expression (Weber and Köster, 2013). Important considerations also include promoter strength, to ensure that the level of expression of the downstream gene is sufficient, and the potential for promoter ‘leakiness’, manifested as low-level off-target expression in unexpected cells and conditions (Prykhozhij et al., 2017). Reliable promoters have been found that restrict expression of an inserted gene to specific regions of the zebrafish heart including the promoter of cardiac myosin light chain 2 (cmlc2, also known as myosin light chain 7 (myl7)) (Huang et al., 2003), ventricular myosin heavy chain (vmhc) (Zhang and Xu, 2009), and atrial myosin heavy chain (amhc, also known as myosin heavy polypeptide 6 (myh6)) (Berdougu et al., 2003).

Early work using Tol2 transgenesis included the introduction of promoter sequences combined with a reporter gene. Although insertion location within the zebrafish genome is random and stable lines contain the insertion in every cell, inclusion of a promoter sequence was able to restrict transgene expression to specific cells and allowed them to be visualised (Kawakami et al., 2004). Targeted gene editing approaches continue to use established promoters, but also allow for cell-specific expression without knowledge of a gene’s promoter. As discussed above, TALENs or CRISPR/Cas9 can be used to incorporate a foreign gene at a specific location within the host genome, causing it to be expressed whenever and wherever the endogenous target gene is expressed. This is accomplished by specifically targeting insertion immediately upstream of an endogenous gene’s start codon or within the transcribed portion of a gene (Hisano et al., 2015; Hoshijima et al., 2016b; Li et al., 2015; Ota et al, 2016; Shin et al., 2014). Therefore, by targeting an endogenous gene that is known to be cell-type specific, expression of an inserted gene can match the restricted expression pattern without requiring full characterisation of the relevant promoter region (Kimura et al., 2014).

4.2. Temporal control

Although controlling the cell-type in which an introduced genetic change is expressed is an incredibly useful tool, controlling timing of expression or function adds even greater utility, particularly for genetic manipulations that are embryonically lethal or that are relevant for study when occurring only at a later life stage. Expression of introduced genetic modifications can be turned on or off by specific interventions, such as heat shock or exposure to chemicals or light. Heat shock promoters, including the commonly used hsp70, allow a downstream gene to remain silent until exposure to heat, which in zebrafish is accomplished by 1 hour of incubation at a temperature about ten degrees higher than normal (Kwan et al., 2007; Mayrhofer and Mione, 2016; Weber and Köster, 2013). Chemically inducible systems link expression of a gene to the presence of a specific drug. Typically, this involves fusing a transcription factor to a hormone receptor that prevents the transcription factor from accessing DNA until pharmacological treatment (Fekler and Mosimann, 2016; Mayrhofer and Mione, 2016). The use of light-activated proteins in the burgeoning field of optogenetics allows for control by the application of light, which results in a more rapid response than chemically-inducible systems that rely on the processing of a drug. Optogenetics is particularly useful for studying excitable cells in the heart as light-sensitive proteins that function as ion channels (Ambrosi et al., 2014; Boyle et al., 2018; Crocini et al., 2017) or fluorescent indicators of intracellular calcium or membrane voltage (Koopman et al., 2017; Quinn et al., 2016, Schneider-Warme, 2018) have recently been discovered or developed.

4.3. Fluorescent reporters

Zebrafish are particularly conducive to the use of fluorescent reporters due to the relatively easy visualisation of internal structures aided by nearly translucent embryos and existing transgenic models that prevent pigment through older stages (Karlsson et al., 2001; White et al., 2008). A multitude of fluorescent proteins exist, with each excited by a specific spectrum of light that causes them to emit fluorescence. The first and most commonly used fluorescent
reporter is green fluorescent protein (GFP), but options now exist across the colour spectrum and with increased functionality (Halpern et al., 2008; Weber and Köster, 2013; Zhang et al., 2002), including photo-induced colour conversion (Felker and Mosimann, 2016; Halpern et al., 2008; Zhang et al., 2002) or functional reporting, such as indicators of membrane potential or intracellular calcium dynamics (Zhang et al., 2002). Versions of fluorescent proteins have also been generated that are specifically localised to the sarcolemma, nucleus, mitochondria (Halpern et al., 2008; Weber and Köster, 2013; Zhang et al., 2002), and actin cytoskeleton (Lin et al., 2012; Reischauer et al., 2014), allowing detailed subcellular visualisation. Further utility can be gained by expressing multiple fluorescent reporters in the same animal or cell type, accomplished by either interbreeding lines of zebrafish or generating plasmids containing multiple promoters driving expression of different fluorescent proteins (Weber and Köster, 2013).

It is especially beneficial to incorporate a fluorescent reporter within an inserted DNA sequence as it serves as an easy screening tool for indicating successful genetic modification and for visualising affected cell types. Multiple sequence configurations within a plasmid can accomplish this goal. For example, the fluorescent protein can be driven by the same promoter as the gene of interest (inserted or endogenous) and translated either into a single fusion protein linked to the gene of interest or as a protein separate from the gene of interest (Weber and Köster, 2013). Alternatively, a different promoter can be used to drive fluorescence expression in an easily observable tissue distinct from that being studied (Clark et al., 2011). Commonly used options are lens-specific crystallin (Davidson, 2003) or heart-specific cmlc2 (Huang et al., 2003) promoters. Thus, although successful gene editing must always be confirmed by molecular techniques, fluorescence expression in zebrafish larvae can save time and resources by facilitating prioritisation of fish to be raised and tested (Hoshijima et al., 2016a).

4.4. Gal4/UAS

Greater control of modified gene expression can be achieved by separating cell-specific activators of expression from the genes whose expression they control. The principal example of a conditional activator/effect approach used in zebrafish is the Gal4/UAS system (outlined in Fig. 4A), which was initially discovered in yeast. Gal4 is a protein that specifically activates transcription of genes located downstream of an upstream activating sequence (UAS). Introducing the gal4 sequence downstream of a cell-specific promoter or linked to a cell-specific gene results in Gal4 protein expression restricted to that particular cell type, which on its own will have no effect as zebrafish do not naturally contain UAS regions in their genome (Weber and Köster, 2013). Including UAS upstream of an inserted gene, on the other hand, will leave the gene silent, as zebrafish do not naturally contain the Gal4 transcription factor required to activate the UAS sequence. When Gal4-expressing and UAS-containing zebrafish are bred, however, the presence of Gal4 will cause activation of UAS and expression of the gene of interest. In this way, the Gal4/UAS system is a method of restricting expression of a gene of interest to a specific cell type. An advantage of the Gal4/UAS system is that one can maintain separate activator and effector lines (each of which has no phenotype on its own) and interbreed combinations as needed. This allows the study of a particular cell type by expressing various genes within it or the study a gene of interest by expressing it in multiple cell types without the need to generate multiple transgenic zebrafish lines (Mayrhofer and Mione, 2016; Weber and Köster, 2013).

The first stable Gal4 and UAS zebrafish lines were generated by direct injection of plasmid DNA into embryos (Scheer and Campos-Ortega, 1999), but soon after Tol2-mediated transgene insertion was employed (Halpern et al., 2008; Mayrhofer and Mione, 2016) (including some Gal4 lines generated using enhancer traps (Kawakami et al., 2010)) and more recently targeted gene editing has expanded the available options (including conversion of existing lines expressing GFP in a particular cell type to expressing Gal4 from the same genomic location) (Auer et al., 2014a, 2014b). Labs can generate their own Gal4 and UAS zebrafish lines following the general processes described for gene editing techniques, with the Gal4 sequence placed downstream of a previously characterised cell-specific promoter or within a gene of interest and the UAS sequence placed upstream of a gene one wants to express in a particular cell type. Various Gal4 and UAS zebrafish lines already exist, with many found on ZFIN, the Zebrafish Enhancer Trap Database (https://burgesslab.nichd.nih.gov/), or the zTrap database (https://kawakami.lab.nig.ac.jp/zttrap/).

Extensions of the Gal4/UAS system have also recently been developed. Attempts to increase activity of Gal4 led to versions including Gal4VP16, which is more effective at driving UAS transcription but highly toxic to embryos, and Gal4FF, which contains two modified versions of VP16 and is more active than the original Gal4 but with fewer toxic side effects. UAS responsiveness has been improved by introducing sequential repeats of its sequence (Mayrhofer and Mione, 2016; Weber and Köster, 2013). One concern of the Gal4/UAS system is UAS silencing, wherein even with Gal4 expression and UAS present in the genome of the same individual, the UAS-related gene is not expressed (Prykozhij et al., 2017). This silencing is thought to be caused by methylation, as the UAS sequence contains side-by-side cytosine-guanine nucleotides, which are the common targets of methylation. Silencing can therefore be avoided by including fewer UAS repeats to offer less sites for methylation and 5 UAS repeats have been shown to be maximally effective (Halpern et al., 2008; Mayrhofer and Mione, 2016; Prykozhij et al., 2017; Weber and Köster, 2013). Silencing may also depend on the location of transgene insertion, however this is less well understood (Halpern et al., 2008; Prykozhij et al., 2017).

---

![Fig. 4. Methods for conditional transgene expression. See Sections 4.4 and 4.5 for details.](https://kawakami.lab.nig.ac.jp/zttrap/)
4.5. Site-specific recombinases

Another powerful tool for conditional genetic manipulation is the use of site-specific recombinases, which involves the use of a recombinase enzyme to cleave a pair of specific short DNA sequences. When this occurs, the DNA sequence flanked by the two recognition sites is excised. Alternatively, a different orientation of the recognition sites results in inversion of the flanked DNA sequence rather than its removal (Felker and Mosimann, 2016; Mayrhofer and Mione, 2016). The most commonly used site-specific recombinase system is Cre recombinase and its 34 base pair recognition site called loxP (although various variations of the lox recognition sites exist, namely Lox2272 and LoxN) (Felker and Mosimann, 2016; Mayrhofer and Mione, 2016). In a common use of the system (outlined in Fig. 4B), loxP sequences are included on either end of a gene of interest within a plasmid. Upon insertion into the host genome, the gene is expressed normally, as the loxP sequences have no effect on expression. Separately, zebrasid lines are created that express Cre recombinase driven by a specific promoter, which will have no effect as loxP sites do not naturally occur in the zebrasid genome. If these loxP-containing and Cre-expressing zebrasid lines are then bred, Cre recombinase is expressed in the desired cell type and will bind and cleave at the loxP sites, resulting in removal of the intervening gene. This results in knock-out of the gene in cells containing the specific promoter (with the remaining loxP having no effect). Alternatively, a sequence that prevents transcription of a downstream gene (a ‘stop’ sequence) can be flanked by loxP sites and placed upstream of a gene of interest, so that it will not be expressed. However, breeding with a Cre recombinase-expressing zebrasid will result in offspring exhibiting Cre recombinase-mediated removal of the ‘stop’ sequence and therefore expression of the gene (Felker and Mosimann, 2016; Mayrhofer and Mione, 2016; Weber and Köster, 2013).

An alternative site-specific recombinase system is Flp/frt, wherein Flipsase (Flp) recognises and cleaves two frt sequences in a similar fashion as Cre-loxP. Importantly, the LoxP and frt recognition sites are cut exclusively by Cre and Flp, so both systems can be incorporated into one gene editing strategy to maximise capabilities (Mayrhofer and Mione, 2016; Weber and Köster, 2013). For example, a DNA sequence called FT1 can be integrated into zebrasid in a forward or reverse orientation using Tol2 transgenesis. Insertion of FT1 in the forward orientation within an intron produces a truncated non-functional protein of the affected gene and expression of the fluorescent protein mCherry; insertion in the opposite orientation allows for seamless integration with no impact on endogenous gene function and no fluorescence. Subsequent exposure to Cre or Flp inverts the foreign DNA sequence, allowing rescue or creation of a gene knock-out depending on the initial orientation. Though the affected gene must be identified by PCR due to random integration of FT1, once zebrasid lines are established they serve as powerful models of reversible gene knock-out and rescue (Ni et al., 2012).

4.6. Novel applications of spatio-temporal control of genetic modification for cardiac research in zebrasid

The combination of spatial and temporal control of genetic modification in zebrasid is enabling novel approaches for understanding structure and function of the heart. Recent examples are given below, which serve to illustrate the power of applying such techniques for cardiac research.

An impressive use of genetic techniques was used by Hoshijima et al. (2016b) to study the potassium voltage-gated channel subfamily H member 6a (kcnh6a) gene, which encodes a cardiac-specific potassium channel equivalent to KCNH2/HERG in human and when mutated causes an autosomal dominant form of Long-QT syndrome. They designed a TALEN pair to induce a DSB within intron 6 of the kcnh6a gene, slightly downstream of exon 6. Concurrently they introduced a plasmid containing: (i) a left homologous arm extending upstream, identical to the endogenous kcnh6a sequence but with an loxP site in intron 5; (ii) exon 6 of the kcnh6a gene; (iii) a frt site; (iv) the alpha-crystallin::Venus reporter gene (which expresses a form of GFP exclusively in the lens of the eye); (v) a second frt site; (vi) a second loxP site; and (vii) a right homologous arm extending downstream. When insertion was successful, kcnh6a continued to produce its functional cardiac-specific potassium channel, as exon 6 was replaced but remained intact. The alpha-crystallin promoter produced fluorescence exclusively in the lens and aided in identification of individuals capable of germline transmission. Offspring embryos were then injected with Cre mRNA to cause excision of the entire sequence contained within the two loxP sites, resulting in both loss of the alpha-crystallin::Venus reporter gene (and therefore lens fluorescence) and exon 6 of kcnh6a, resulting in non-functional kcnh6a channels and the associated cardiac disease phenotype. The presence of the frt sites offered the option of injecting embryos with Flp mRNA to remove the alpha-crystallin::Venus reporter gene without affecting lens function. This use of a lens-specific promoter to indicate genetic modification and of site-specific recombinases to enable removal of inserted gene sequences may prove useful in many experimental contexts.

Another striking example of genetic modification in zebrasid comes from a study by Gupta and Poss (2012), in which they traced the lineages of cells to reveal how clonal populations form the zebrasid heart. In this case, they exploited the fact that Cre recombinase can be made chemically inducible by fusion with a mutant estrogen receptor (ER). The resulting CreER protein is expressed exclusively in the cytoplasm, preventing it from acting on loxP sites in nuclear DNA. Upon treatment with tamoxifen or 4-hydroxytamoxifen (4-HT) (which can be introduced into the tank water), CreER can enter the nucleus and exert its effect on loxP (Felker and Mosimann, 2016; Mayrhofer and Mione, 2016). Gupta and Poss used a plasmid containing the β-actin 2 promoter followed by the sequences for red, cyan, and yellow fluorescent protein, with two pairs of different lox sites placed between each gene. After injection, zebrasid containing multiple copies of the DNA insert were used to generate a stable line called prizm, which was then crossed to a cmk2:CreER line expressing an inducible Cre exclusively in cardiomyocytes. At 2 dpf, a timepoint at which the heart consists of a single layer of cells, embryos were exposed to 4-HT. Cre recombinase was then able to enter the myocyte nuclei and randomly act on a pair of lox sites within each copy of the introduced DNA sequence. Each interaction caused excision of one or more fluorescence genes and allowed expression of those remaining. Because of the multiple copies of the insert, greater than twenty variations in the colour of fluorescence colour were possible. All cells present at the time of treatment that later proliferated produced daughter cells expressing the same unique fluorescent colour (Fig. 5). Of note, CreER is available under the control of a ubiquitous promoter and a range of cell-specific promoters (information can be found in the CreZoo database (https://cre zoo.crt-dresden.de/crezoo)) (Felker and Mosimann, 2016; Mayrhofer and Mione, 2016; Varshney et al., 2015).

A novel method for spatio-temporal control of gene editing in zebrasid developed by Ablian et al. (2015) used CRISPR/Cas9 for cell-specific knock-out to prevent heme biosynthesis in erythrocytes. The regular CRISPR/Cas9 approach for targeted mutation results in global gene knockout and thus does not allow interro-
site-specific sgRNA were generated as usual, however a Tol2-compatible CRISPR plasmid (available from Addgene) and Gateway cloning were used to construct a plasmid containing a ubiquitous promoter driving sgRNA expression and a cell-specific promoter driving Cas9 expression. Upon injection of the plasmid and Tol2 transposase mRNA, the designed DNA fragment randomly inserted into the genome, creating a stable line in which sgRNA is expressed in every cell but only results in gene knock-out in the specific cell types expressing Cas9.

A further development has been reported by Reade et al. (2017) in which they activated *lefty1*, an antagonist of the Nodal signalling pathway, to interfere with left-right determination in early development. The technique used a blue light-activated transcription activator (an optimised version of the protein TALEN, TA4-EL222), which binds to C120 sequences to stimulate expression of a downstream gene, essentially creating a system (TAEL/C120) comparable to Gal4/UAS, but with the added control of light activation. Specifically, a stable transgenic line expressing *lefty1* downstream of C120 was generated followed by injection of embryos with TAEL mRNA. Embryos exposed to blue light from 12 to 24 hours post fertilisation (a period crucial to left-right patterning) were found to have hearts with both chambers along the midline, rather than with left-right orientation. Furthermore, they combined the light-activated spatial control of TAEL/C120 with the precise gene targeting of CRISPR/Cas9. Embryos were injected with TAEL mRNA and a plasmid containing both Cas9 downstream of C120 and a ubiquitous promoter driving expression of an sgRNA targeted to tyrosinase, a protein required for pigment development. Blue light exposure resulted in CRISPR/Cas9-mediated mutations in the tyrosine gene and prevented normal pigment formation.

5. Conclusion

Zebrafish serve as an important genetically modifiable experimental model, as gene editing is relatively easy, inexpensive, and high-throughput compared to other vertebrates (especially mammals), a plethora of resources for implementation now exist, and they offer a viable model for many human diseases. The current genetic toolbox makes nearly any desired combination of genetic edits possible, yet access to existing fish and ability to generate custom lines remains a challenge for many researchers. While the zebrafish community has worked hard to make methods including Tol2, TALEN, and CRISPR/Cas9-mediated gene editing available to researchers, existing resources suffer from a lack of central organisation. In this review we have attempted to gather the essential knowledge and form a comprehensive resource for the inexperienced researcher exploring the potential of including genetically-modified zebrafish as part of their research program. The aim was

![Image](image-url)
to summarise fundamental knowledge and explain technical details for those wishing to obtain or generate genetically-modified zebrafish lines.

Although a researcher may wish to develop the ability to use gene editing techniques to modify genes of interest themselves, large scale projects are underway to generate a genetically-modified zebrafish for each of the 26,000 protein-coding genes (Howe et al., 2013) in the zebrafish genome (Kettleborough et al., 2013; Varshney et al., 2015). Thus, when considering using zebrafish to study a specific gene, cell-type, or disease it is worthwhile to first search publications and resources for an existing relevant model. An understanding of random and targeted mutagenesis approaches enables judgement of whether existing fish can serve experimental needs. If transgenic zebrafish lines have been generated, but no longer exist or cannot be obtained, the plasmids used in initial generation are often accessible and can be used to save time and resources.

Tol2, TALEN, and CRISPR/Cas9 can in theory be implemented by any lab to generate genetically-modified zebrafish, however each technique has advantages and limitations related to the experimental goal and available technical capabilities. As one considers gene editing in zebrafish, the first step is to carefully consider the intended genetic change. An understanding of available tools can enable thoughtful up-front design to maximise utility of the resulting transgenic line. If random insertion into the genome is sufficient or desired, Tol2-mediated transgene insertion is most likely the best choice, and the many available resources and existing plasmids for use in Gateway cloning should ease the process of engineering the foreign DNA required. If a specific gene or region of the genome must be targeted, TALEN or CRISPR/Cas9 is required. Given the ease of generating an sgRNA compared to a series of TALE domains, most labs will opt for a CRISPR/Cas9 strategy unless a specific genomic region can only be targeted with TALENs. Though potential off-target effects of CRISPR/Cas9 must be investigated, the increasing flexibility and capability of this modality make it likely to continue to be the most commonly used targeted gene editing option.

If targeted editing is employed, one must also decide how to best harness the repair of the induced DSB to generate the genetic modification desired. Most fundamentally, the difference between gene knock-out or knock-in is accomplished by relying on NHEJ alone or providing an external DNA template to be inserted during repair. Targeted DSB followed by random mutations during NHEJ repair that lead to a non-functional protein product is efficient and relatively easily accomplished, but design of the site of the DSB is important. Introducing plasmid DNA allows precise changes not possible with any other strategy but creating the DNA sequence to be inserted can be challenging for a lab lacking molecular cloning experience. ssODNs offer an easy alternative, though the breadth of edits that can be accomplished is limited.

With the rapidly evolving field of gene editing, it can be difficult for the average researcher to keep up. Labs understandably have different access to and preferences for methods and the outcome of any given study will be impacted by the specifics of the genomic site targeted and the gene editing strategy employed. These factors complicate the comparison of studies and decisions about which approach to take. High variability currently exists, for instance, in length of homology arms, form of injected plasmids, the method used to investigate mutation induction efficiency, and reporting germline transmission. The importance of educating oneself to employ best practices, work efficiently, and increase rates of success would be simplified if access to updated resources and protocols was improved. Increased understanding of molecular mechanisms triggered in a zebrafish embryo following injection and through subsequent rapid cell division will also help inform improved methods.

Finally, it is worth noting that although some innovative genetic techniques have been developed specifically for cardiac research in zebrafish, many transformative developments have come from experts in other fields and it is important to consider them for application to the heart. Genetically-modified zebrafish offer experimental capabilities not possible using other models and are worthy of consideration in many research programs. Increased development of user-friendly tools and accessible methods are key to expanding their use, as researchers are better equipped to understand available options, obtain existing lines, or implement techniques to design and generate custom ones. It is hoped that improved awareness of the zebrafish as a genetically modifiable organism will increase its consideration as a powerful tool to be used in combination with others for a comprehensive understanding of cardiac physiology and disease.

Editors’ note

Please see also related communications in this issue by Segert et al. (2018) and Ojehomon et al. (2018).

Acknowledgements

Disclosure

The authors have no conflicts of interest to declare relating to this work.

Author contributions

S.A.R. drafting of manuscript. All authors approved of the final version for publication.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2016-04879 to T.A.Q.) and the Canadian Institutes of Health Research (MOP 342562 to T.A.Q.), TAQ is a National New Investigator of the Heart and Stroke Foundation of Canada.

References

Bendig, G., Grimminger, M., Huttnier, I.G., Wessels, G., Dahme, T., Just, S., Trano, N., Katus, H.A., Fishman, M.C., Rothbauer, W., 2006. Integrin-linked kinase, a novel component of the cardiac mechanical stretch sensor, controls contractility in...


