Tampere University of Applied Sciences

CRISPR-Cas9 based mutagenesis and gene silencing in zebrafish

Pieter Maene

BACHELOR'S THESIS
June 2020

Biomedical Laboratory Science

Odisee
DE CO-HOGESCHOOL

Bachelorproef
Studiegebied Gezondheidszorg

Pieter Maene
Professionele Bachelor Biomedische laboratoriumtechnologie
Afstudeerrichting Farmaceutische en biologische laboratoriumtechnologie
Academiejaar 2019/2020

CRISPR-Cas9 based mutagenesis and gene silencing in zebrafish

Tampere University
Experimental Immunology
Arvo Ylpön katu 34
33520 Tampere
Finland

CRISPR-Cas9 based mutagenesis and gene silencing in zebrafish

INTERNSHIP DETAILS

Internship details

Period	:	3 February 2020 – 15 June 2020
Internship	:	Tampere University Experimental Immunology Arvo Ylpön katu 34, 33520 Tampere Finland Tel.: +358 40 190 9834 e-mail: info@biomeditech.fi
Mentor	:	Dr. Anni Saralahti
Tutor	:	Delphine Goethals

PREFACE

This thesis marks the completion of not only my internship, but also my time at Odisee University college. Therefore, I would like to start by thanking Odisee and its lecturers in helping me succeed and giving me the opportunity for an internship abroad. I’d also like to thank my parents for the opportunity to study and the support they gave me during the whole period. Mr. Gryson also deserves gratitude as he aided all the students abroad to get everything in order and was always ready to help when there were questions.

It was planned that this thesis would be a reflection of practical work, coupled with data acquired from experiments. This practical part of the thesis and internship was cut short because of the Covid-19 pandemic and is therefore not mentioned in this thesis. However, I still got the opportunity to gain experience by working in the research facility alongside some great, inspiring researchers. So, I would like to express my gratitude to professor Mika Rämet for giving me the opportunity to be part of the research team. The lab work was an educational experience, mostly thanks to my mentor, Dr. Anni Saralahti. I would like to take the opportunity to sincerely thank Dr. Saralahti for mentoring me. Not only did she take the time to aid me in the lab work to help me become a better laboratory technician, but she kept mentoring me when I had already returned home. Thanks to Dr Saralahti, I was able to write and finish this thesis. Apart from Dr Saralahti, I would also like to thank the other researchers and laboratory technicians of the experimental immunology group in helping develop better laboratory skills and answering my questions when needed.

I found the subject of the bachelor’s thesis really interesting and biotechnology is a field that has always interested me. I was really pleased when this subject was assigned. I learned a lot by writing this thesis and I’m sure the gained knowledge will prove useful in my further career.

This Bachelor’s thesis took a lot of hours to write, that’s why I would like to give thanks to my parents and friends for motivating and supporting me. The last people I’d love to thank are the ones that proofread my thesis: Vilters Katrien, Van de Steene Tessa, Röther Valentina, and Fiers Mathias.

ABSTRACT

Keywords: Gene editing, CRISPR, CRISPRi, guided nucleases, zebrafish, CRP

Genetic modification has come a long way since its inception. Biotechnological laboratories are impossible to imagine without it. Guided nucleases from zinc finger nucleases to TALENs and, most notably, the CRISPR-Cas systems are readily available. They have revolutionized the use of model organisms such as *Danio rerio*, the zebrafish. The zebrafish has proven very advantageous as a model organism and has been used in a plethora of studies based around genetic modification. The guided nucleases, as well as other tools for genetic modification (such as Morpholino oligonucleotides), have been thoroughly used in zebrafish studies.

With the CRISPR system being used extensively in laboratories, a lot of progress has been made since its first use in genetic modification. In a field that is ever evolving, it is important to keep an overview of the possibilities and the advantages as well as the disadvantages of the different techniques. Different techniques have been developed based on the CRISPR system initially found in bacteria. One system worth mentioning is CRISPRi. It is based on a modified Cas enzyme. Cas enzymes have been modified as Cas9 nickase or dCas9 and have broadened the applications of the CRISPR system. This review focusses on the use of CRISPR-Cas based gene mutagenesis and CRISPR based gene silencing in zebrafish and the delivery methods thereof.

Student: Pieter Maene

pieter.maene@hotmail.com

Mentor: dr. Anni Saralahti

anni.saralahti@tuni.fi

TABLE OF CONTENTS

INTERNSHIP DETAILS **5**

PREFACE **6**

ABSTRACT **7**

TABLE OF CONTENTS **5**

LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS **7**

1 PRESENTATION PLACE OF INTERNSHIP **9**

2 AIMS OF THE STUDY **10**

3 REVIEW OF THE LITERATURE **11**

3.1 INTRODUCTION 11

3.1.1 What is gene modification? **11**

3.1.2 Classification of gene modifications **12**

3.1.3 Challenges and limitations **14**

3.2 ZEBRAFISH 15

3.2.1 Zebrafish: general **15**

3.2.2 Zebrafish as a model organism **16**

3.2.3 Immunology of the zebrafish **18**

3.2.4 Streptococcal infection in zebrafish **19**

3.3 APPLICATIONS OF MUTANT ZEBRAFISH LINES 20

3.4 NON-CRISPR BASED MUTAGENESIS METHODS 21

3.4.1 ZFN **21**

3.4.2 TALENs **23**

3.4.3 Morpholinos **24**

3.5 CRISPR-CAS SYSTEM 25

3.5.1 CRISPR-Cas system in nature **25**

3.5.2 CRISPR-Cas system in laboratories **28**

3.5.3 CRISPR mediated mutagenesis in zebrafish **29**

3.6 CRISPR-I 30

3.6.1 Gene silencing **30**

3.6.2 CRISPRi **30**

3.7 METHODS OF CRISPR-CAS9 DELIVERY 31

3.7.1 General methods of delivery **31**

3.7.2 Delivery in zebrafish 32
3.8 POST-MUTATIONAL CONTROL 34
3.8.1 Genotyping 34
3.8.2 Sample collection and preparation 34
3.8.3 Genomic analysis 35
3.8.4 Phenotypic analysis 35
4 SUMMARY AND CONCLUSIONS 36
5 REFERENCES 38
6 ANNEXES 52
6.1 FIGURES 52

LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS

(A)GE	(agarose) gel electrophoresis
AAV	adeno-associated virus
AD	activation domain
ap.	approximately
ATP	adenosine triphosphate
bp	base pair(s)
Cas	CRISPR-associated
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRi	CRISPR interference
CRP	C-reactive protein
crRNA	CRISPR-RNA
dCas9	nuclease deficient Cas9 protein/catalytically inactive dead Cas9 protein
DNA	deoxyribonucleic acid
dpf	days post fertilization
DSB	double stranded breaks
dsDNA	double stranded DNA
E. coli	*Escherichia coli*
e.g.	exempli gratia, for example
etc	et cetera
GFP	green fluorescent protein
GOI	gene of interest
gRNA	guide RNA
HDR	homology-directed repair
hpf	hours post fertilization
hpi	hours post injection
HR	homologous recombination
HRMA	high-resolution melting analysis
Indels	insertions/deletions
KI	knock-in
KO	knock-out
KRAB	Krüppel associated box
MET	Faculty of medicine and health technologies
MGE	mobile genetic elements
MO	morpholino oligonucleotide
mRNA	messenger RNA
ngs	next-generation sequencing
NHEJ	non-homologous end joining
NLS	nuclear localization signals
NOD	nucleotide-binding oligomerization domain
nt	nucleotide
OPEN	Oligomerized Pool Engineering
PAGE	polyacrylamide gel electrophoresis
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PCV(13)	(13-valent) pneumococcal conjugate vaccine
PDK	polycystic kidney disease
PPV(23)	(23-valent) pneumococcal polysaccharide vaccine
pre-crRNA	precursor CRISPR RNA

Pre-mRNA	precursor mRNA
PTGS	posttranscriptional gene silencing
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNAi	RNA interference
RNP	ribonucleoprotein
RVD	repeat-variable diresidue
saCas9	*Staphylococcus aureus* Cas9
sgRNA	single stranded guide RNA
siRNA	small interfering RNA
SMR	SWISS-MODEL repository
SNP	single nucleotide polymorphism
spCas9	*Streptococcus pyogenes* Cas9
ssDNA	single stranded DNA
T3SS	type III secretion system
TAGE	T7-digestion AGE
TAL	transcription activator-like
TALE(N)	TAL-effector (nucleases)
TGS	transcriptional gene silencing
TLR	toll-like receptor
tracrRNA	transactivating CRISPR RNA
wpf	weeks post fertilization
WT	wild type
ZAL	zwitterionic amino lipids
ZF	zinc finger
ZFIN	the Zebrafish Information Network
ZFN	zinc finger nucleases
ZFP	zinc finger protein
ZNP	ZAL nanoparticle

1 PRESENTATION PLACE OF INTERNSHIP

The internship took place in the Arvo building, associated with Tampere University of applied sciences in Tampere, Finland.

Finland is a Nordic country in Europe, neighboring Finland are Sweden and Russia. It is, however, not part of the Scandinavian countries because Finnish is not related to the Scandinavian languages. It is rather a Finnic (part of the Uralic language family) language. Tampere is a city in Southern Finland. It is the second most populated city in Finland after Helsinki. Tampere is known for its two lakes surrounding the city, Näsijärvi and Pyhäjärvi.

Tampere University of applied sciences, or TAMK (Tampereen ammattikorkeakoulu), is part of the Tampere University community. The different universities (TAMK, Tampere University, and Tampere University of Technology) merged at the beginning of 2019 to form the community as to have closer communication with each other. Through this, they want to ensure access to higher education and research by working together. TAMK is one of the largest and most popular universities of applied sciences in Finland.

The Arvo building, located at Arvo Ylpön katu 34 in Tampere, Finland, has offered premises for both Tampere University and Tampere University of applied sciences since 2011. The location is close to TAMK main campus and TAYS medical center. The Arvo building is part of the Faculty of Medicine and Health Technologies (MET) at Tampere University. There are currently 35 different research groups working at the faculty, all working in different fields such as cancer genomics, immunoregulation, toxicology and pharmacology, and more. One of the research groups is the experimental immunology research group, under the leadership of professor Mika Rämet. The experimental Immunology research group focuses on the mechanisms of the immune response. Its goal is to find the genes involved in the immune system regulation or activation. Besides the immune response to certain pathogens, the virulence or pathogenicity of the pathogens is studied.

The research team uses primarily non-mammalian animal models such as the fruit fly (*Drosophila melanogaster)* and the zebrafish (*Danio rerio)*. Professor Rämet's research group has multiple ongoing projects. The project I got involved in focused on the role of C-reactive protein (CRP) in the defense against streptococcal infections, using CRISPR-based mutagenesis in Zebrafish (*Danio rerio*). This project is led by postdoctoral researcher Anni Saralahti.

2 AIMS OF THE STUDY

In a previous study (Saralahti et al., unpublished results), the zebrafish were genetically modified and subjected to survival assays after pneumococcal (*Streptococcus pneumoniae)* infection. The genes responsible for the affected phenotypes were determined. The *crp* gene was one of the genes that, when altered, responded to a lowered survival. The lack of expression of *crp* was shown to result in a higher mortality in the zebrafish.

CRP, or C- reactive protein, is a protein of which the level in plasma increases greatly during acute phase response to tissue injury, infection, or other inflammatory stimuli. It is also an important factor in the innate defense against bacterial and viral infections.

The ultimate aim of the project is to assess the role of CRP in the defense against streptococcal infections. This will be done by mutating the *crp* gene in zebrafish using CRISPR-based mutagenesis and silencing the *crp* expression with CRISPRi-based knockdown.

This review explores the different types of gene modification methods with a strong emphasis on both CRISPR-Cas9 and CRISPRi. The goal is to find out how CRISPR and CRISPRi can be administered in zebrafish and how these methods differ from each other or influence the experiments. This comparison will be achieved through comparison of the available literature on the matter.

3 REVIEW OF THE LITERATURE

3.1 INTRODUCTION

3.1.1 What is gene modification?

Genetic modification or editing, sometimes also referred to as gene engineering, is the practice of changing the genetic makeup of organisms. It is a practice that has been happening for a long time in the form of selective breeding of crops or animals. Ever since the discovery of the DNA double helix in 1953 (Franklin & Gosling, 1995; Watson & Crick, 1953) scientist have been trying to introduce site-specific modifications in the genomes of organisms because the ability to change the nucleic acid bases at precise locations holds a lot of value. It's a technique with applications in pharmaceuticals, healthcare, biotechnology, farming, etc.

One of the earliest and most important discoveries in the search for precise gene editing was the discovery of restriction endonucleases, found in bacteria. The first restriction enzyme was found in *Escherichia coli* (Meselson & Yuan, 1968). The endonuclease was shown to degrade foreign DNA, but not the host-DNA. The isolated restriction enzyme was site-specific and required Mg^{2+} ions, S-adenyl methionine and ATP. After the isolated enzyme from *E. coli*, a restriction enzyme was isolated from *Haemophilus influenzae* (Smith & Welcox, 1970), this enzyme only needed magnesium ions as a co-factor. Restriction enzymes inflict a limited number of double stranded breaks (DSBs) in target DNA that contains certain recognition sites (Kelly & Smith, 1970). The inflicted DSBs can then be repaired using the cell's own repair mechanisms (Rudin et al., 1989). The most important repair mechanisms are non-homologous end joining (NHEJ) and homology-directed repair (HDR) (see chapter 3.1.2). The restriction enzymes were the first step to modifying DNA.

The next step in the journey to modern gene editing tools was when homologous recombination (HR) was discovered in yeast cells (Scherer & Davis, 1979) and mammalian cells (Smithies et al., 1985). This was not the first method for stable incorporation of gene inserts into a genome (Wigler et al., 1977) but it was, however, the first method where the insertion was site-specific. The homologous recombination technique had a low frequency of success ($\sim 10^{-3}$). It embodies the mechanism where exogenous DNA can be directed into the host genome by homologous sequences within that genome. Later, it was found that the efficiency of HR increased greatly when applying it to DNA after a DSB was inflicted (Rudin et al., 1989). The combination of isolated restriction endonucleases and HR resulted in a technique for targeted gene integration. The downside was that the restriction nucleases were not easy to use. The recognition sites occur multiple times in the target genome and there are confined available recognition sites since these are limited to the number of isolated restriction enzymes (some enzymes even share recognition sites). These problems combined meant that only a small fraction of genes could be specifically targeted. To solve this problem, researchers developed targeted nucleases to inflict specific DSB in target DNA (Gupta & Shukla, 2017).

First, meganucleases, also called homing nucleases, were used. Meganuclease technology consists of re-engineering the DNA-binding domain in naturally occurring restriction enzymes. This technique faces a lot of challenges. The DNA-binding domain and cleavage domain are hard to separate and re-engineering proteins is relatively difficult. These two properties made meganucleases less useful in practice (Maeder & Gersbach, 2016).

Later, scientists stumbled upon Zinc Finger-mediated DNA binding. By combining these zinc fingers with nuclease domains of restriction enzymes (more specifically FokI), zinc finger nucleases (ZFN) were created. ZFN were among the first site-specific nucleases (see chapter

3.4.1). Following the use of ZFN, reports surfaced of transcription activator-like (TAL) effectors. These were also coupled with the same FokI-endonuclease domain to create TAL-effector nucleases (TALENs) (see chapter 3.4.2).

The latest big step for gene editing was the implementation of RNA-guided nucleases, based on the CRISPR-Cas systems, found in prokaryotes (see chapter 3.5). CRISPR-Cas was already discovered earlier and studied in the field of bacterial immunology but was not used for gene integration. It was later integrated into the field of genome editing (Doudna & Charpentier, 2014).

3.1.2 Classification of gene modifications

There are different kinds of gene modifications. They can be classified under gene knock-down, where the gene transcription is lowered, gene knock-in (KI) where a new gene is expressed, overexpression where the rate of transcription will be higher or gene knock-out (KO) where the gene is silenced. The modifications can be either stable or transient. Transient modifications are modifications that are only temporarily incorporated. Each modification has a different working mechanism. The most common modifications will be outlined briefly in this chapter. The type of modification is strongly dependent on the repair mechanism used to repair the inflicted DSB. Therefore, to fully grasp the concept, an understanding of the repair mechanisms is needed. DSBs are common in eukaryotic cells and occur in nature. They can, for instance, happen during cell division. The occurring DSBs can be repaired by two different pathways: NHEJ or HDR. NHEJ is the most frequent repair pathway for DSB in mammalian cells (H. H. Y. Chang et al., 2017). When DSB are present, the cells engage NHEJ (Hustedt & Durocher, 2017). In this repair mechanism, the DSB is repaired through direct re-ligation of the cleaved DNA ends. Repair by NHEJ is error prone and often causes small indels (insertion/deletion).

HDR can be further divided into different forms. The most common form is homologous recombination, but it also includes single-strand annealing and breakage-induced replication. The different forms of HDR have different requirements concerning the length of the homologous sequences between donor and acceptor (Lieber, 2010). For cells to use HDR, a template must be present and the organism has to be diploid (even transient diploidy, as in dividing bacteria). The template is copied during the repair, making it possible to introduce an insert in the target sequence (Bannikov & Lavrov, 2017). To utilize HDR, the plasmid DNA template needs big homology regions (ap. 1000bp) flanking the DSB. If using a smaller ssDNA as a template, it is required for it to be a maximum of 120 nt (nucleotides).

HDR is less effective than NHEJ, but there are methods to get a better efficiency from HDR when incorporating templates (Richardson et al., 2016). NHEJ repairing causes small indels at the place of the DSB, this makes it an efficient way of gene KO. HDR repairing, on the other hand, incorporates an insert in the desired site and is thus a good way for gene KI (Figure 1).

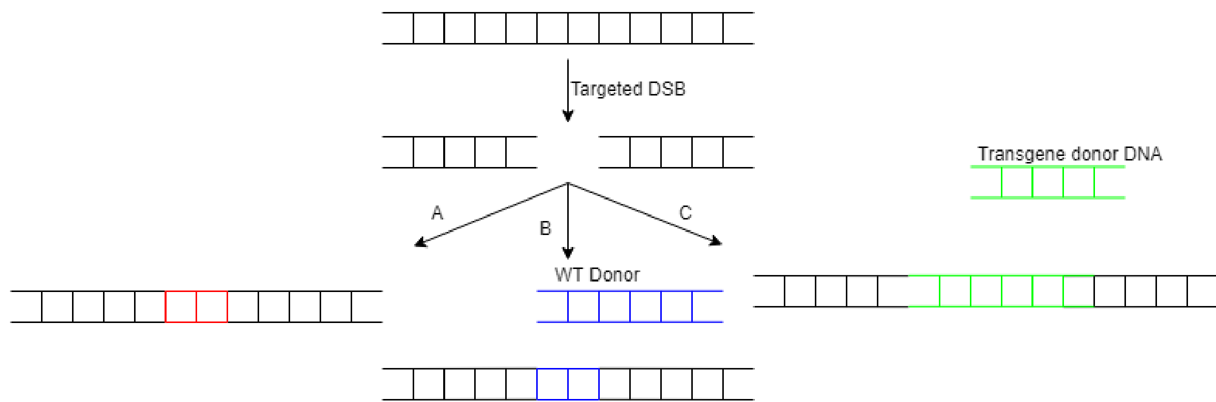


Figure 1: Genome engineering with DSB. This figure gives an overview of the different kinds of genome engineering. At first, a (targeted) DSB is inflicted in the target DNA, this can be done through targeted nucleases such as ZFN, TALENs, and CRISPR/Cas. In reaction A, the DSB is repaired with the NHEJ mechanism resulting in indels (shown in red) and subsequent gene disruption or knockout. In reaction B and reaction C, the DSB is repaired by HDR. In mechanism B, the donor DNA consists of the wild type (WT), but with the correct base sequence. This is done for genetic diseases where a mutation distorts the gene activity. In reaction C the donor DNA is a transgenic DNA, this can be used for gene knock-ins. A: NHEJ, knockout; B: HDR, Gene correction; C: HDR, Knock-in

Gene knockouts (Figure 1A) can be caused by choosing the nuclease target site in the coding region of a gene. Because of the mutagen character of NHEJ repair mechanisms, the repair of the cleavage will result in an indel. This indel may cause a frameshift mutation. The frameshift mutation causes the transcript most likely to be wrong and non-functional. This is called a knockout. The knockouts can be used to target gain-of-function mutational diseases. One example is Huntington's disease. Patients with Huntington's disease have a repeat expansion on a protein coding gene (Ramaswamy & Kordower, 2012). Using knockdown technology, this expansion can be silenced, resulting in a non-mutant protein. The same idea, although less common, can be adopted to repair certain genetic diseases. Some diseases are caused by frameshift mutations. By targeting these indels with the guided nuclease systems, the correct reading frame can be restored (Ousterout et al., 2013). This is known as reading frame correction.

Knock-in therapy, sometimes referred to as gene therapy (although this review steers clear of that term because it has multiple possible definitions), is based on the HDR repair mechanism. By using the HDR repair mechanism, the DSB can be repaired by addition of a wanted exogenous DNA sequence. Gene correction is similar, but the used template has an incorporated correction for a point mutation or other disease-causing mutation. The opposite of Knock-in by insertions, would be gene inactivation/knockout by deletion. This is possible by introducing two DSBs, each flaking the gene of interest.

The gene modifications can undergo stable integration or transient expression (T. K. Kim & Eberwine, 2010). When an inserted gene is integrated in the host genome, this is known as stable integration. Some edits are not integrated and are only expressed for a short period of time (~days). Transiently expressed genes are not heritable (Condreay et al., 1999).

3.1.3 Challenges and limitations

The challenges and limitations for each introduced method will be discussed in chapters 3.4 and 3.5. In this chapter, the general challenges and limitations for gene modification in eukaryotic organisms will be discussed.

When editing genomes, one of the most common challenges is the off-target effects. This happens when the nuclease (whether it is ZFN, TALEN, or CRISPR based) cuts in a place other than the target DNA. Off-target effects can occur with unintended binding of the enzymes (Knott & Doudna, 2018). Bigger genomes usually contain multiple regions that are (highly) identical. This increases the chances of unspecific binding and thus increases the chances of off-target effects. One of the goals in the development and optimization of gene editing technologies lies in the specificity of the targeted nucleases. An improvement or solution for the unspecificity can lie in modifying the FokI nuclease domain, for example (Szczepek et al., 2007). This greatly decreased the cytotoxicity when working with ZFN, presumably because of lowered off-target effects.

Another problem for gene knockout or perturbations is a problem known as genetic robustness. Genetic or biological robustness is an inherent property of biological systems. Genetic robustness is a combination of naturally occurring mechanisms aimed to conserve biological or genetic functions even when natural perturbations occur (Kitano, 2004). Perturbation or gene knockout can activate compensatory mechanisms in the organism. These compensatory mechanisms can alter the expression of different genes and can happen through transcriptional or signaling pathways (El-Brolosy & Stainier, 2017). Genetic robustness can occur in several forms, one of which, dosage compensation, was already found in 1932 (Muller, 1932) and later shown in the *Drosophila melanogaster* in 1965 (Mukherjee & Beermann, 1965). In this experiment, it was shown that male fruit flies have twice as much transcription from their single X-chromosome compared to female fruit flies that have two X-chromosomes. Eukaryotes also possess genetic redundancies (Tautz, 1992). With genetic redundancies, one protein is encoded by multiple genes; This can also be referred to as overlapping gene function. Knockdown assays in one of these genes will not have full effect because the other genes will mask the effect. Insertional mutagenesis can also cause problems in genome-editing. With insertional mutagenesis, an otherwise inactive gene is activated and transcribed. This can be manifested in the form of proto-oncogenes (Baum et al., 2004).

Furthermore, the efficiency is also a challenging factor. Not every organism will carry the wanted mutation or perturbation. This conjures two extra challenges in the experimental design. First, it is a challenge to improve efficiencies. Whether it is by using a different method or by altering existing methods. Secondly, because not all organisms will carry the mutation, a screening method is needed. It is not always easy to screen for the genetic modification if the phenotype is not clearly distinguishable. The available methods for the control of mutation will be discussed in chapter 3.8.

3.2 ZEBRAFISH

3.2.1 Zebrafish: general

The zebrafish, or *Danio rerio,* are members of the Cyprinidae (minnow) family of the order Cypriniformes. (Vishwanath, 2010). The taxonomy of the zebrafish can be seen in Table 1: (McCluskey & Postlethwait, 2015)

Table 1: Taxonomy of Danio rerio

Kingdom	*Animalia*
Phylum	*Chordata*
Class	*Actinopterygii*
Order	*Cypriniformes*
Family	*Cyprinidae*
Genus	*Danio*

The zebrafish's natural habitat is inland (freshwater) waters. They are found in slow moving, small streams or rivers, nearby streams and rice paddies (Engeszer et al., 2007). Zebrafish are native to southern Asia and can be found in many different countries. They breed all year round and feed on insects, worms, and small crustaceans.

Zebrafish are named after the blue horizontal stripes on the side of their bodies (Figure 2). The adults are 3-5 cm long and up to 1 cm wide. To a trained eye, the male and female zebrafish can be relatively clearly distinguished. The zebrafish live in shoals.

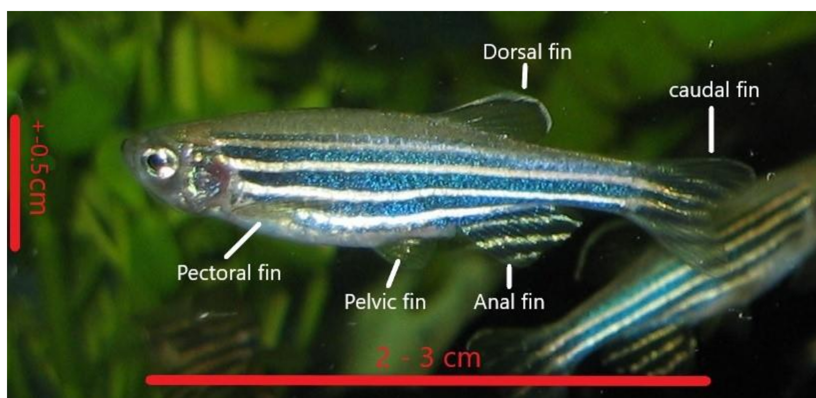


Figure 2: Adult female zebrafish. The blue stripes running across the sides of the zebrafish clearly show how the fish got its name. The fish in this figure is a female fish. This can be determined by looking at the abdomen, but one usually needs some practice before being able to clearly distinguish between sexes. The abdomen (located between the pectoral fin and pelvic fin) of the female fish is bigger than a males' abdomen.

3.2.2 Zebrafish as a model organism

The use of animal models has helped researchers tremendously in biomedical research. Animal models are beneficial because they give a better understanding of the biological responses in comparison to *in vitro* analyses. Animal models have proved useful in the past and are at the base of many important discoveries. For example, RNA interference (RNAi) has first been proven (and studied) in *Caenorhabditis elegans*, a type of roundworm (Fire et al., 1998; Guo & Kemphues, 1995). The Toll-like receptors (TLRs), also found in humans, were first discovered in *Drosophila melanogaster* (Hansson & Edfeldt, 2005). Mammalian models would be the most representable, but given their ethical questions and financial issues, other model organisms are often easier to use. Opposed to the first-mentioned non-mammalian organisms, zebrafish are vertebrate organisms. Zebrafish have been used extensively as a model organism for the study of developmental biology and toxicology (Bambino & Chu, 2017) but have seen an uprising in the use for immunological studies as well as cancer studies (Amatruda et al., 2002). These vertebrates show great usability in animal studies compared to mammalian models (Vargas et al., 2015). Because of their small size, zebrafish are easy to keep and maintain with high population density (5-7 fish/liter). They also give a lot of offspring (high fecundity) and are relatively cheap (Allen & Neely, 2010). The embryos develop rather quickly (Figure 4). During development, the zebrafish are transparent through the first two weeks, they also develop *ex utero*, which makes monitoring development convenient. They have a great genetic homology to humans; At least 70% of the coding human genome has direct orthologues in the zebrafish. On top of that, their genome is completely sequenced (Howe et al., 2013). The high level of homology suggests that numerous genes connected to human diseases are also found in the zebrafish genome. This explains why the zebrafish has seen a big surge in biomedical research in general (Figure 3).

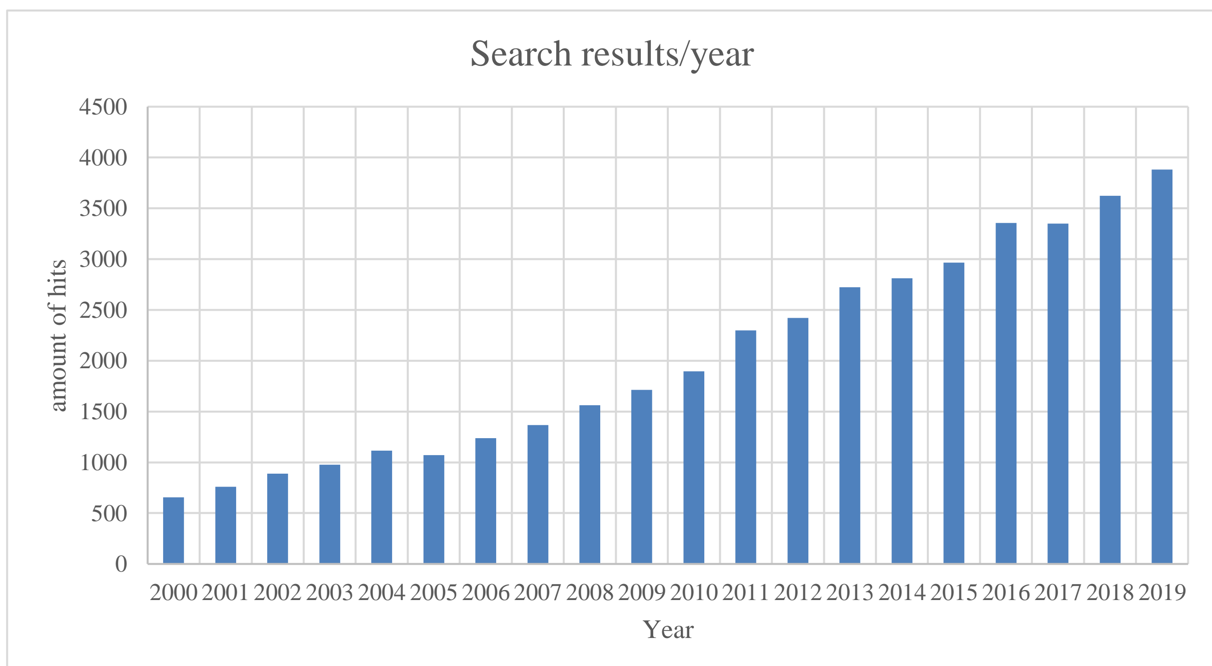


Figure 3: Graph of search results per year. This graph shows the amount of hits on PubMed search when using the search query 'zebrafish'. While there are a modest 657 hits in the year 2000, this number rose to a staggering 3882 hits in 2019. Data retrieved from pubmed.ncbi.nlm.nih.gov (02-05-2020)

Zebrafish have been used extensively in research and an online database has been opened where researchers can access information regarding genes, genomes and other information about working with zebrafish (*ZFIN The Zebrafish Information Network*, n.d.). The ZFIN is an important factor in why the zebrafish is a useful model organism. It manages the information coming from all the research conducted using zebrafish. It includes mutants, genes, publications, and similar (Sprague, 2006). The development of zebrafish was already documented earlier (Figure 4).

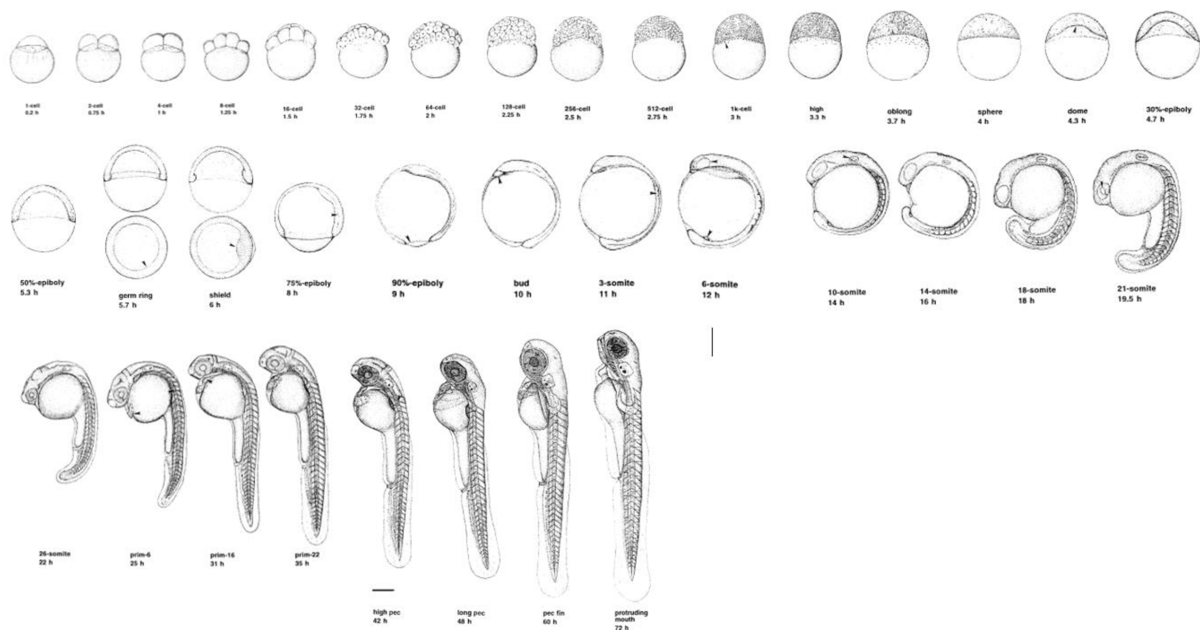


Figure 4: Developmental stages of Danio rerio (Kimmel et al., 1995). *In this figure, the developmental stages of the zebrafish are pictured. Starting from the one-cell stage (0,2 hours post fertilization (hpf)) up until the formation of the protruding mouth (72 hpf). It is clear that in 72 hours, the zebrafish embryo develops. The rate at which the embryo develops is pictured at 27,5°C. These rates can be slowed down by using cold water of 24°C* (Villamizar et al., 2012).

The zebrafish model organism also has some limitations. One of which is that there is no one-to-one link between the homologues of humans and zebrafish. The zebrafish have multiple duplicate genes which could make genetic screening more complex (Woods et al., 2000). This can be especially troublesome with knockout strategies. Silencing one gene could have no or little effect on the amount of gene transcript produced. Having duplicate genes means that both (or more) of the genes must be silenced. Even though the genome is fully sequenced, the annotation of the zebrafish genome is still not complete and is still limited. Apart from the genetic elements, zebrafish have different environmental parameters in comparison to humans or mammals, such as the adapted body temperature (28°C in zebrafish, 37°C in humans). There are also big differences in anatomical and physiological structure between zebrafish and mammals (Tobin et al., 2012). Zebrafish do not possess lymph nodes, the zebrafish has a different reproductive system and they use gills instead of lungs (Menke et al., 2011). The difference in anatomy can have a big impact when studying host-pathogen interactions because they limit the range of diseases that can be studied in zebrafish. On top of these limitations, the zebrafish is still a relatively new model organism and while there has been a lot of progress, there is still a scarcity of zebrafish specific antibodies. This issue can be partially resolved by employing the cross-reactivity exhibited by anti-mammalian antibodies in zebrafish (Villarreal et al., 2017).

3.2.3 Immunology of the zebrafish

The human immune system is quite complex and these complex host-pathogen interactions explain why the reactions might happen differently *in vivo* and *in vitro*. That is why there is a need for a model organism that resembles the human immune system to study the complex host-pathogen interactions *in vivo* (H. Meijer & P. Spaink, 2011). The transparency of zebrafish in the early life stages combined with the *in vivo* imaging possibilities, make the zebrafish an attractive model organism for immunological studies (Hall et al., 2008). The zebrafish model has helped to study the different areas of human pathology, which include, but are not limited to, microbe-host interaction and heritable diseases and they are also readily used in drug screens (Goldsmith & Jobin, 2012).

In zebrafish, the innate and adaptive immune systems develop at different stages (Lam et al., 2004). This makes it possible to study the influence of certain immune cell types in the host-pathogen interactions. During the first 4 days post fertilization (dpf), the zebrafish does not exhibit adaptive immunity markers. The adaptive immune system is only functionally mature after 4-6 weeks post fertilization (wpf); This means the larvae survive solely on the innate immune system up until at least 4 wpf (Novoa & Figueras, 2012). Infectious experiments at different developmental stages are easily done. Therefore, it is possible to study either innate immune response only or in combination with the adaptive immune response.

The zebrafish has an immune system that shows many similarities to mammalian systems and thus to the human immune system. For one, the zebrafish immune system has different waves of hematopoiesis that match with the hematopoietic waves found in mammalian embryos (Bertrand & Traver, 2009). The hematopoietic waves take place in anatomically distinct tissues (Davidson & Zon, 2004). The zebrafish immune system utilizes macrophage-like cells (25 hpf), granulocytes (48 hpf), cytokines and chemokines, Toll-like receptors (Trede et al., 2004). It has become clear that, when comparing the zebrafish's immune system to the mammalian immune system (both innate and adaptive), that these are highly conserved (Sullivan & Kim, 2008). All main components of the innate immune system found in humans can also be found in zebrafish (Kanther & Rawls, 2010; Y. Li et al., 2017; Meijer et al., 2004) (Table 2). The zebrafish can also be used as a model for the adaptive immune response, but the focus in this paper will be on the innate immune response.

Inflammation is an important part of the innate immune defense. It has been shown that this response in zebrafish can be induced easily by transecting the tail of the larvae and that inflammation resolves over a similar time to mammalian models (Renshaw et al., 2006). The zebrafish thus shows promising results as a model organism to study infectious diseases (O'Callaghan & Vergunst, 2010; Tobin et al., 2010; Torraca & Mostowy, 2018).

Table 2: Comparison of the selected components of humans and zebrafish innate immune system

Immune cell in humans	**Homologs in zebrafish**
Monocytes	+
Macrophages	+
Neutrophiles	+
Eosinophiles	+
Basophils	-
Mast cells	+
NK cells	+
Patters recognition receptors in humans	
10 different TLR (TLR1 – TLR10)	Several TLR homologs identified (not for TLR6 and TLR10)
23 NOD-like receptors	>400 NOD-like receptors

3.2.4 Streptococcal infection in zebrafish

The human Gram-positive pathogen *Streptococcus pneumoniae* (pneumococcus) has a polysaccharide capsule that contains essential virulence factors. It is usually present as a harmless commensal of the nasopharynx but is capable of pathogenesis (Nobbs et al., 2015) when spreading to deeper areas. It is one of the most common causes of bacterial pneumonia, meningitis, otitis media, and septicemia. The mortality of pneumococcal infections is highest in extreme ages, such as elderly or young children. According to estimates from the WHO, 476 000 children younger than five (HIV negative) lost their lives in 2008 due to pneumococcal infection (WHO, 2013). Around 90 distinct serotypes of *Streptococcus pneumoniae* have been reported. Multiple reports about antibiotic resistance of the pneumococcus have surfaced (Men et al., 2019). This antibiotic resistance urges for the use of vaccinations. There are two types of pneumococcal vaccines: pneumococcal conjugate vaccine (PCV) based on bacterial polysaccharides bound to a carrier protein and pneumococcal polysaccharide vaccine (PPV) based on the polysaccharides from the bacterial capsule. Both the PCV and PPV provide serotype-specific protection. This means that the vaccines do not protect against every serotype. The PPV protects against 23 serotypes and is often referred to as 23-valent pneumococcal polysaccharide vaccine (PPV23). The newest version of the PCVs offers immunization against 13 serotypes and is referred to as 13-valent pneumococcal conjugate vaccine (PCV13). These vaccinations still do not offer full protection. Consequently, the pneumococcus is an important pathogen to study.

There have been multiple studies where bacterial infections have been carried out in zebrafish to check if these pathogens can, in fact, infect the model organism. For example, the disease caused by *Streptococcus pyogenes* in zebrafish is similar to the disease caused by the same bacterium in humans (Neely et al., 2002). The infection with *S. iniae* in zebrafish reproduces a lot of the features also caused in humans by a variety of streptococcal infection. The interests of this review go out to the *Streptococcus pneumoniae* infection. The *S. pneumoniae* is not a natural fish pathogen and its optimal growth temperature is the human body temperature (37°C), while the zebrafish are kept in 28°C water.

Previous studies have shown that injection (into the bloodstream) of zebrafish embryos with *S. pneumoniae* causes fatal infections (Rounioja et al., 2012). As mentioned in the aim of the study (chapter 2), this is the relevant bacterial infection to study. The same study also states that zebrafish embryos are capable to combat the pneumococcal infection. The zebrafish was shown to be a valuable vertebrate model to study *S. pneumoniae* infections. Adult zebrafish are susceptible to bacterial meningitis, caused by *S. pneumoniae*. It is reported that there is a limited adaptive immune response against these infections in zebrafish, meaning this model organism is not suited for the study of adaptive immune responses to pneumococcus or the research for human vaccines against streptococcal infections (Saralahti et al., 2014).

3.3 APPLICATIONS OF MUTANT ZEBRAFISH LINES

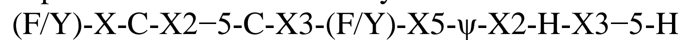
Because of the suitability of the zebrafish model for immunological studies, they are often used in experimental immunology studies. Experimentally manipulated genes and the possibility for high-throughput-screens help to investigate the role of genes or proteins in disease or immunity (Zon & Peterson, 2005). Assigning function of genes can happen either through forward or reverse genetics. Forward genetic approaches randomly mutate genes with a mutagen (e.g. chemical mutagens, UV light etc.). The mutated organisms are then screened for different phenotypes and their genome is analyzed. In reverse genetics however, the goal is to find the function of a particular gene of interest (GOI) by specifically mutating that GOI and searching for the phenotypical alteration. Because of the rising use of zebrafish in research, reverse and forward genetic methods are well described, and several studies have reported about the efficiency of these methods. With forward and reverse genetic approaches, it is possible to identify genes that are (potentially) relevant for human disease (Shah et al., 2016) because, as mentioned above, the zebrafish genes have a high percentage of homologues in humans. These homologues can be used in genetic screens to identify genes responsible for genetic diseases. This, for example, is how genes responsible for polycystic kidney disease were identified (Sun, 2004). Chapter 3.2.2 mentioned that the zebrafish is also used for developmental studies. This is the case for the study where genome-wide screens were used to identify genes responsible for abnormal embryogenesis (Amsterdam et al., 1999; Driever et al., 1996). Zebrafish have also been used in pharmaceutical studies (Langheinrich et al., 2003) to study the effect of certain drugs on QT-prolonging and even in neuronal circuit studies (Albadri et al., 2017).

3.4 NON-CRISPR BASED MUTAGENESIS METHODS

3.4.1 ZFN

Zinc-finger nucleases are a way of targeted (sequence-specific) genome cleavage. Designed ZFN can inflict DSB in target DNA and the DSB can trigger cellular DNA repair mechanisms: HDR or NHEJ (as described in chapter 3.1.1). Zinc finger nucleases were created by fusing two zinc finger proteins (ZFP) (Left ZFP & right ZFP) and the cleavage domain of FokI (from *Flavobacterium okeanokoites*) endonuclease together (Y. G. Kim et al., 1996). Zinc fingers (ZF), together with helix-turn-helix motifs and leucine zippers, are the most common DNA binding proteins found in eukaryotes. ZFPs are eukaryotic transcription factors found in multiple possible configurations. Cys_2His_2 zinc fingers (Figure 5) are the most commonly found DNA-binding motifs in eukaryotes. ZFPs were shown to contain 7-11 zinc atoms and are built from nine similar tandems with two invariant cysteines and histidines pairs (J. Miller et al., 1985).

The Cys_2His_2 domain is not only used for protein-DNA interaction but is also used *in vivo* for protein-RNA and protein-protein interaction. They contain the consensus sequence



(F/Y)-X-C-X2−5-C-X3-(F/Y)-X5-ψ-X2-H-X3−5-H

where X= any amino acid ψ= hydrophobic residue

(Wolfe et al., 2009).

Figure 5: Human Zinc Finger protein structure. (Q9NZL3, ZN224_HUMAN). On the left the structure from Uniprot. The blue dot in the middle (and the purple dot of the right structure) represents the zinc ion. The alpha helix can be seen on the left-hand side and the double beta sheet on the right. The right structure is from SMR. The residues in yellow depict the cysteine residues. The alpha Helix can be found at the C-terminus while the beta sheets are connected to the N-terminus.

When cleaving the complete FokI enzyme with trypsin, a 41 kDa amino-terminal end and a 25 kDa carboxyl-terminal fragment are found. The amino-terminal fragment was found to be the FokI recognition domain and the carboxyl-end fragment was found to be the cleavage domain. (L. Li et al., 1992). By combining these two domains, the ZFN combine the qualities of both domains. They have the DNA binding quality of ZFP and the nuclease activity of a restriction enzyme.

When constructing the ZFN, it is possible to use three to six fingers per ZFN. With each zinc finger added, the recognition site gets expanded with 3 bp because each ZF has a 3bp recognition site (Carroll, 2011). The most used technique for making ZFN is called "modular assembly" where the DNA-binding domains of the ZFs (each with a 3bp recognition site) are combined to the desired recognition site (Segal et al., 2003) (Figure 6). The ZFP assembly can be carried out by inserting the necessary genes in bacterial expression vectors.

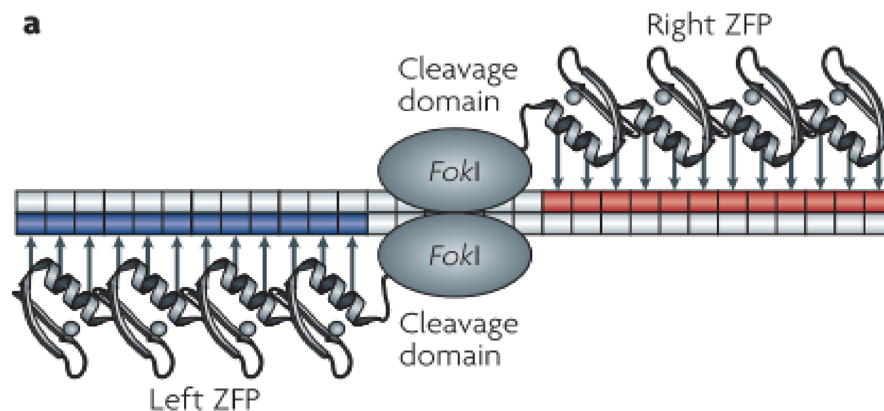


Figure 6: Structure of ZFN dimer bound to target DNA. This figure clearly shows the modular mechanisms of ZFN. The FokI nuclease cleavage domain is shown in the middle, bound to ZFP. There are four ZFP on each side. Every single zinc finger has a recognition site of 3 bp, making this particular example a ZFN with a recognition site of 24 bp. (Urnov et al., 2010)

A big advantage for the use of ZFN lies in the fact that the FokI endonuclease domain requires dimerization to cleave the target (Bitinaite et al., 1998). This might sound counterintuitive, but the need for dimerization means that the FokI domain will only cleave the DNA when both parts of the ZFN dimer are bound to their respective recognition sites. In other words, cleavage will not happen when solely one monomer is bound to the DNA, this lowers the chance of non-specific cuts (and off-target effects).

High gene targeting frequencies were first reported in *Drosophila*. This was done by direct injection in embryos with ZFN-coding mRNA (Beumer et al., 2008). Later, the injections were also shown to yield a high effectivity in zebrafish. When injected during the 1-cell stage, a high frequency of the animals carried the expected mutation (Doyon et al., 2008). The mutations were also transmitted quite frequently through the germline (Meng et al., 2008), resulting in viable adults.

There are problems and challenges with ZFN, though. One of which is that the modular assembly is keen to failure (Ramirez et al., 2008). ZF motifs usually bind to their 3 bp recognition site, however, when an aspartic acid residue is present at the +2 position of the alpha-helix, it can change the recognition site to a 4 bp (an extra adenine or cytosine) via cross-strand contact (Chandrasegaran & Carroll, 2016). This 4th bp contact can influence the specificity of the flanking ZFs. This further complicates the design of ZFP and ZFN because it means that not each ZF acts independently from its neighbor. These problems with the modular assembly of ZFP forced scientists to resolve to other methods for ZFP selection: the sequential selection and bipartite selection. The protocols for these selection methods are outside of the scope of this review and will not be discussed. Both the approaches yield ZFP with a high specificity and affinity but take a lot of time and do not lend themselves for routine performance.

Customized zinc finger nucleases created with Oligomerized Pool Engineering (OPEN) are proven to work in zebrafish (Foley et al., 2009), but this will also not be discussed any further.

3.4.2 TALENs

TALENs, TAL-effector nucleases, or transcription activator-like effector nucleases were the follow-up to ZFN. Due to the relative difficulty in design with ZFN and the fact that there are context-dependent effects with ZFN (see chapter 3.4.1), TALENs were introduced as an alternative method. In comparison to ZFN, TALENs are easier to design and engineer and the specificity is similar (Hockemeyer et al., 2011).

Like ZFN, TALENs are engineered protein nucleases that can induce site-specific DSB. These DSB can then be repaired using NHEJ or HDR (see chapter 3.1.1) to induce site-specific gene modifications. TALENs are based on the DNA binding domain of the TAL-effectors and the FokI nuclease domain (Bi & Yang, 2017). Similar to the ZFN, the TALENs are composed of an engineered array of repeats fused to the cleavage domain. The TAL-effector was discovered in a bacterial genus, the *Xanthomonas spp.*

Xanthomonas is a genus of 27 Gram-negative bacterial species known for infecting plants (Ryan et al., 2011). The *Xanthomonas spp.* genome contains four gene clusters associated with its pathogenesis: *xps, rpf, hrp, gum. T*he *hrp* (hypersensitive reaction and pathogenicity) codes for a type III secretion system (Tampakaki et al., 2004). The type III secretion system (T3SS) was the base for the discovery of TALEs. When infecting a plant in nature, the T3SS injects > more than 25 effector proteins in the plant cell. These effector proteins play an important role in pathogenicity (Bonas, 1990). Some of these effector proteins induce gene transcription in the host plant (Kay et al., 2007; Marois et al., 2002). The proteins are effectively called transcription activator-like effectors (TALEs). These TALEs contain a DNA binding domain to be able to induce transcription by binding to the DNA.

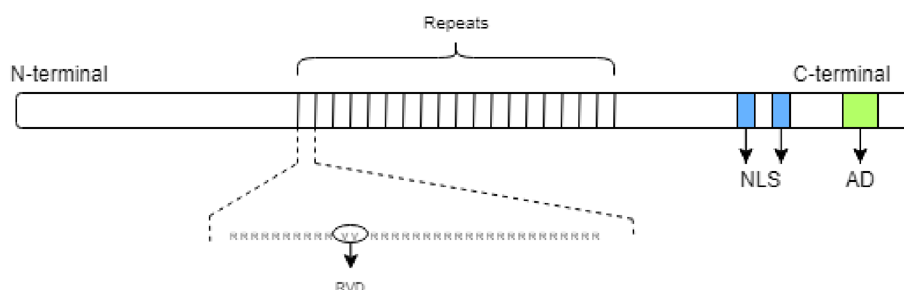


Figure 7: TAL-effector domain. representation of the TAL protein. In the middle section, the tandem repeat array can be found. The first repeat has been enlarged. 'R' represents an amino acid residue. The 34 amino acids residues can be seen; The residues on position 12 and 13 are the hypervariable residues (represented by 'V') that make up the RVD and define the DNA binding specificity. Closer to the C-terminus, Nuclear localization signals (NLS) can be found as well as an acidic activator domain (AD)

The TAL effector protein consists of 13 to 28 repeats, nuclear localization signals, and an activation domain. One TALE repeat typically consists of 34 amino acid repeats (Figure 7). The last repeat is usually truncated to 20 amino acids. At positions 12 and 13 there is a polymorphism, the repeat-variable diresidue (RVD). The RVD directly correlates with the target nucleotides (Moscou & Bogdanove, 2009). The two variable amino acid residues of the RVD recognize one base pair in the DNA (Boch et al., 2009). The central repeats bind with the target DNA by forming a supersolenoid around the DNA double helix. Each repeat connects with one nucleotide through the RVD hypervariable residue (Bogdanove & Voytas, 2011).

It is possible to construct an artificial protein that recognizes any given DNA sequence of interest by generating TALEs with new repeat sequences. In the same fashion as how ZFN were build, TALENs can be built i.e combining it with the FokI nuclease domain. The TALEs can also be used for gene perturbations by fusing the DNA binding domain to other effector domains, e.g. the VP64 domain (see chapter 3.6.2) (J. C. Miller et al., 2011). As described with ZFN, the TALENs also work in pairs in order to dimerize the FokI domain. To create a TALEN with a specific target, the four types of repeats (each targeting one of the four DNA bases) are combined in the wanted fashion (Bi & Yang, 2017). Dual TALENs have been proven to work as effectively in homodimeric as in heterodimeric configurations (Christian et al., 2010). By injecting RNA encoding TALE nuclease pairs into the 1-cell stage of zebrafish embryos, the activity of TALENs in zebrafish was shown to be successfully tested (Sander et al., 2011). TALENs show lower off-target effects compared to ZFN.

3.4.3 Morpholinos

Morpholinos oligonucleotides (MOs), sometimes just called Morpholinos, are synthetic antisense modified DNA oligonucleotides. They carry a six-ring heterocycle backbone (called the morpholine ring) and nonionic phosphorodiamidate linkages (Figure 8) (Summerton & Weller, 1997). The morpholine ring and phosphorodiamidate linkages provide the MOs with a high *in vivo* stability compared to regular oligonucleotides (Kok et al., 2015). They are designed to bind to mRNA and effectively block translation by steric hindrance making MOs a method for (transient) gene knockdown (Stainier et al., 2017). Morpholinos do not introduce mutations like ZFNs, TALENs, and CRISPR-based methods. Instead, they are a gene perturbation technique. The outcome is comparable to small interfering RNAs (siRNA) and CRISPRi (see chapter 3.6.2).

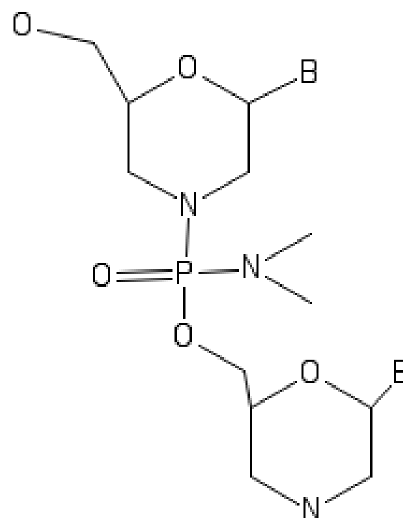


Figure 8: Morpholino oligomer backbone. One repeating unit of the Morpholino oligomer backbone. Where B represents one of the nucleobases (A,C,G,T). Each subunit contains one base bound to the morpholine ring structure. The phosphorodiamidate linkages connect the subunits.

When using MOs as a method for translation blockage, the oligomers are designed to match the region neighboring the first ATG. To interfere with mRNA splicing, the oligomers are designed to match a splice donor or acceptor site (Schulte-Merker & Stainier, 2014). MOs have been used in different model organisms, for example in zebrafish and *Xenopus* (Heasman et al., 2000). In zebrafish, MOs have been used extensively. They can be injected during the 1-cell stage. By inhibiting either splicing or translation of mRNA, gene function knockdown can be observed for up to 10 days (Nasevicius & Ekker, 2000). Thus, MO effects last long enough to inactivate genes that regulate the immune system. MOs targeting different genes can also be co-injected, thus providing a unique opportunity to test multiple gene functions simultaneously within a living organism.

Because MOs only affect RNA transcripts rather than genome structure, like CRISPR, it is harder to determine the phenotype/genotype correlation. Morpholinos are also prone to misleading results and off-target effects (Eisen & Smith, 2008). The off-target effects are less common than the previously used siRNAs. This happens due to the fact that the non-ionic MOs interact less with extracellular and cellular structures that cause the off-target effects (Summerton, 2007). The off-target effects, however, are still common. Therefore, the CRISPRi system can be a better option for transient gene knockdown studies. For permanent loss-of-function studies, the CRISPR system can be used.

3.5 CRISPR-CAS SYSTEM

3.5.1 CRISPR-Cas system in nature

The CRISPR/Cas immunity or Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated proteins immunity is a mechanism found in nature, first observed in *Escherichia coli* (Ishino et al., 1987) and later found in numerous archaea and bacteria (Mojica et al., 2000). It is used by bacteria and archaea as an adaptive sequence-specific defense mechanism against invading foreign DNA, or mobile genetic elements (MGEs), such as plasmids and phages (Hwang et al., 2013). CRISPR-Cas works by introducing double stranded breaks (DSBs) in the invading target DNA. This defense mechanism is based on a DNA region referred to as Clustered Regularly Interspaced Palindromic Repeats (CRISPR) and is capable of functioning in a single cell (Rath et al., 2015). The CRISPR array consists of repeats and spacers. The spacers have unique sequences and are derived from MGEs (Terns & Terns, 2012). They are used as the recognition patterns of the invading nucleic acids.

The CRISPR locus on itself is not active. It needs CRISPR-associated proteins, coded for by CRISPR-associated genes. Usually these CRISPR-associated (Cas) genes are situated next to the CRISPR array (Figure 9). There are multiple Cas proteins. Some of these Cas proteins are associated with almost all CRISPR systems (e.g. cas1, cas2) while other Cas proteins are only associated with certain CRISPR-cas systems. Preceding the CRISPR array is another conserved, AT-rich area, the leader. The leader contains a promotor and a recognition site for spacer insertion.

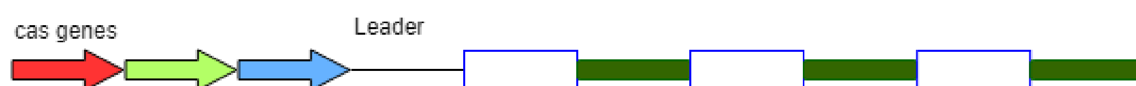


Figure 9: CRISPR Locus organization. (simplified representation). Depicted here is a type II CRISPR locus organization. The cas operon genes can be found on the left, followed by the leader and on the right the CRISPR array. The blue rectangles represent the repeats, while the filled green rectangles represent the unique spacers, derived from protospacers or invading MGEs.

The CRISPR-Cas systems were first classified into three different types (type I, type II, type III) depending on the *cas* genes, organization of cas-operons, and the structure of the repeats (Chylinski et al., 2014). More recently, also a type IV, type V, and type VI systems were identified, but their mechanisms are less understood (Makarova et al., 2020). The simplest (and most widely used) is the CRISPR-Cas type II, also called "HNH-systems" (more specifically the type II system from *Streptococcus pyogenes*). In the type II system, the endonuclease activities required for the interference with foreign DNA are concentrated in a single protein, namely the Cas9 protein. The type II system also codes for the Cas1 and Cas2 proteins and sometimes a fourth protein. The type II systems can be divided into three different subtypes, namely type II-A, type II-B, and type II-C. Type II-A and type II-B both code for a fourth cas protein: the csn2 protein and the cas4 protein, respectively. Type II-C does not have a fourth cas-protein coding gene. When talking about CRISPR-Cas systems in this thesis, the type II system is meant, unless stated otherwise.

The Cas9 protein is a large, multidomain protein consisting of about 800-1400 amino acids. It contains two different nuclease domains: the HNH domain and the RuvC-like domain. Both domains each cut a different strand of the target DNA and are thus both needed to cleave a double stranded target (Nishimasu et al., 2014). In the Cas9 protein, two clearly distinct lobes can be found. One of the lobes has all the nuclease domains, except a part of the HNH domain, while the other lobe contains the binding sites.

The immunity works by incorporating short fragments of the foreign invading DNA into the CRISPR loci in the bacteria (Amitai & Sorek, 2016). In short, the foreign DNA is incorporated in the genome as a spacer. The sequence from which the spacer is derived is called the protospacer. Later, this spacer is transcribed to crRNA (CRISPR-RNA). A single crRNA contains a spacer and a part of the repeat sequences. The formed crRNA interacts with Cas9 to form a crRNA-CasRNP complex. In a type II CRISPR-Cas system, the conversion uses tracrRNA (transactivating CRISPR RNA), RNase III, and an unknown factor. The tracrRNA is a non-coding RNA and is required for associating with the cas9 protein for the activity of the cas9 protein.

This complex can basepair (Watson-Crick base-pairing) with the invading DNA complementary to the spacer (thus 'recognizing' the repeated invasion) and cleaving the alien DNA. The HNH nuclease domain in the Cas9 protein cleaves the complementary DNA strand, while the RuvC-like domain cleaves the non-complementary DNA strand (Jinek et al., 2012). Combining these two actions results in a double stranded break in the target DNA. The foreign DNA contains a PAM (protospacer adjacent motif) (Amitai & Sorek, 2016). The PAM is a short sequence whose presence is needed for the cleavage to take place. PAM recognition in type II CRISPR requires the cas1-cas2 complex as well as the cas9 protein. Because the PAM is needed for the correct binding of the Cas protein to the DNA, and because this PAM sequence is not present in the bacterial genome, but only in the invading MGE's, the bacterial genome is protected from cleavage by the Cas nuclease activity. The PAM sequence is a sequence of 3-5 bp. *Streptococcus pyogenes* Cas9 (spCas9), for example, recognizes an NGG-motif or sometimes a less effective NAG motif (Pallarès Masmitjà et al., 2019). While the recognized PAM motif for *Staphylococcus aureus* Cas9 (saCas9) is an NNGRRT (Mekler et al., 2020). PAM-motifs are used in type I and type II systems. The type III systems use special repeats that flank the spacer DNA; These spacers prevent the cleavage.

This CRISPR-Cas immunity response can be divided into three phases: adaptation, expression, and interference.

The adaptation stage is the stage where the invading MGE gets detected and is incorporated in the CRISPR as a spacer. It is the stage where the bacterium gets the genetic memory. The spacers are usually incorporated at the same one side (proximal end) of the CRISPR array. This means that CRISPR has a chronological order of incorporated spacers. It becomes clear that this stage grants the host immunity at the genomic level. This means that the immunity (gain or loss) against invading nucleic acids will be hereditary and passed on to the offspring (van der Oost et al., 2009). There are two different ways of spacer integration. There is naïve acquisition, when the MGE has not yet been encountered, and primed acquisition, where the MGE has been encountered before. The Cas1-Cas2 complex is an important factor in the acquisition and integration of new spacers (Fineran & Charpentier, 2012). Both Cas1 and Cas2 are needed for this integration (He et al., 2018).

During the expression stage, the transcription starts (Figure 10). The CRISPR DNA is transcribed to form pre-mRNA that, when processed enzymatically, gives crRNA that acts as gRNA (guide RNA). At the same time, the *cas* genes are transcribed into Cas proteins (Sternberg et al., 2016). crRNA has the spacer at the 5'-end and a fragment of the CRISPR repeat at the 3'-end. The processing of pre-crRNA to crRNA requires the actions of tracrRNA (trans activating crRNA), the Cas9 protein, and RNaseIII (endoribonuclease III).

Stage three, the interference, is the stage in which the invading nucleic acid is recognized by the ribonucleoprotein complex (crRNA-Cas RNP) and is silenced. In this process, the tracrRNA hybridizes with the repeat sequence and forms a dual RNA hybrid structure (Jiang & Doudna, 2017). This hybrid can guide the Cas9 protein to the DNA sequence that contains the compatible 20 nt sequence and adjacent PAM-sequence from the invading nucleic acid. Once the Cas9 protein is guided to the invading nucleic acids, it can introduce a double stranded break using its nuclease activity. The DSB is then repaired by error-prone NHEJ or HDR.

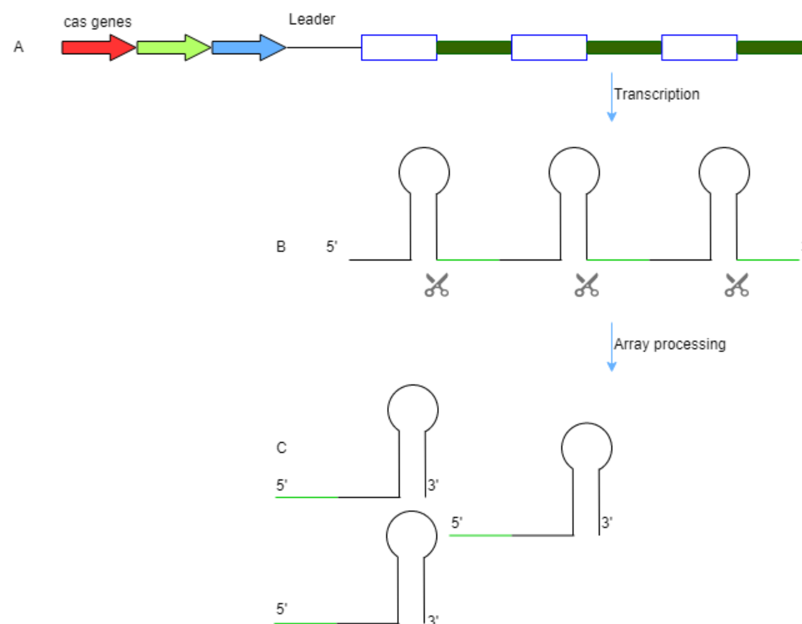


Figure 10: Processing of CRISPR array to crRNA. Expression stage of CRISPR immunity response. Figure A shows the CRISPR locus. The CRISPR array with its spacers and repeats is transcribed. Figure B shows the pre-crRNA, with hairpin loops. The acquired spacers are shown in green. The CRISPR transcript is then processed by Cas proteins and cut, resulting in crRNA. crRNA can be seen in figure C. The crRNA from figure C is later combined with the Cas9 protein to form the RNP.

3.5.2 CRISPR-Cas system in laboratories

The CRISPR/Cas9 system can be used for precise gene modification in plant cells and animal cells. The use of CRISPR/Cas9 type II has greatly facilitated the use of animal models and genetic modifications of the animal models. Unlike ZFNs and TALENs, that use protein motifs for target recognition, CRISPR-based methods use RNA-DNA recognition. This is a big advantage because RNA design (for one single guide RNA) makes it easier to implement in laboratories. Additionally, Watson-Crick base pairing is more specific than protein-DNA interactions. One important discovery showed that the CRISPR system was still active when transferred between bacteria (Sapranauskas et al., n.d.). This sparked a lot of interest among scientists wanting to use the CRISPR system to induce specific DSB in DNA. The CRISPR/Cas type II has been engineered for genome editing as RNA guided nuclease (Ma et al., 2014). In nature the CRISPR/Cas9 system consists of a dual RNA structure (mature crRNA and tracrRNA). In an experimental context, it is a lot more beneficial to use if this is reduced to a single RNA structure. The gRNA is thus engineered as a single transcript (figure 11). As described above, the crRNA and tracrRNA are artificially combined and form a single chimeric RNA strand, the sgRNA (single guide RNA). The sgRNA consists of a nucleotide sequence of $\pm$ 80 nt, of which a 20 nt spacer. This 20 nt spacer in the sgRNA can be engineered to match with a sequence of interest (Pallarès Masmitjà et al., 2019), causing the Cas9 protein to target that specific sequence and inflict a DSB. This is useful for lab work because it is a way to easily create RNA guided nucleases. Apart from the 20 bp target recognition site, the sgRNA also contains a hairpin scaffold structure that mimics the formed structure when tracrRNA and crRNA basepair.

Any DNA sequence that contains the needed PAM motif and 20 nucleotides can be used as a target for the gRNA (guide RNA) (Gasiunas et al., 2012). As the guide RNA's can be relatively easily designed in the lab, this opens a lot of possibilities to target the genes of interest and to study e.g. loss of function. By introducing multiple gRNA's, this system can be used to target multiple genes at once.

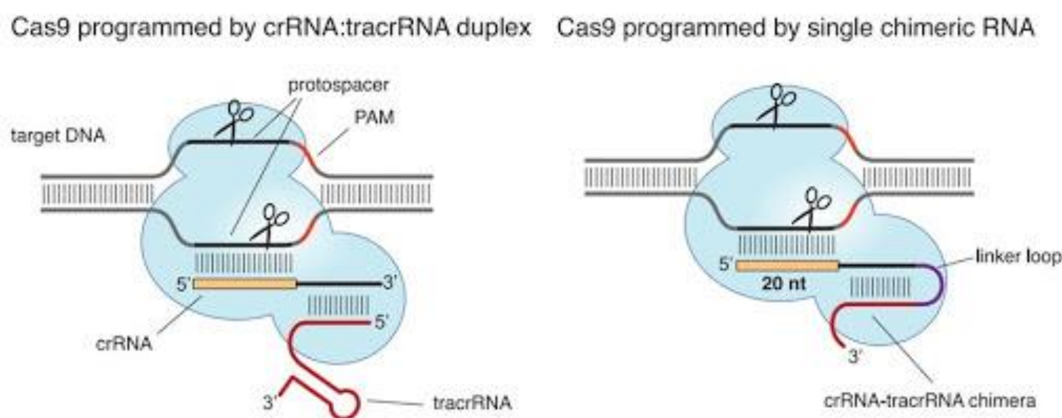


Figure 11: Comparison between the interference stage of the CRISPR-Cas9 systems found in nature (left) and the CRISPR-Cas9 system used in the laboratories, using a single chimeric RNA (sgRNA) (right). Figure adapted from Jinek et al., 2012.

Only one year after Jinek et al. published their findings about producing chimeric RNA molecules (Jinek et al., 2012), it has been shown that the Cas9 endonuclease can be 'programmed' for site-specific modification of eukaryotic, and therefore human, genomes (Jinek et al., 2013). Micro-injections with Cas9 and sgRNA in embryonic 1-cell stages have

been proven to be effective in zebrafish and mice (Sung et al., 2014). Mutant lines of zebrafish can be made with an efficiency up to 86,0% (Hruscha et al., 2013) and the mutation is heritable.

Scientists are still searching for new applications and to increase specificity (Kleinstiver et al., 2016) of the CRISPR systems (Adli, 2018). Furthermore, they aim to engineer smaller Cas9 mutants, as this would have more therapeutic applications. To reduce the off-target effects, there is the possibility to produce mutant Cas9 proteins with a higher fidelity (Bannikov & Lavrov, 2017). "Dead Cas9" or Cas9 nickase (nCas9) proteins are also being used, the dead Cas9 proteins are nuclease deficient because of certain mutations. nCas9 proteins are deficient in one of the two nuclease domains, giving single stranded breaks. Different Cas9 proteins gave rise to the CRISPR/Cas9 system being employed in more applications than gene editing alone. It has also found applications in gene regulation (e.g. CRISPRi), epigenome and chromatin editing (Hilton et al., 2015; Thakore et al., 2015), and has also been used in imaging applications. CRISPR-based methods for nucleic acid detection have also been developed (Broughton et al., 2019; Kellner et al., 2019). These tests are being employed now as a detection test for the COVID-19 disease, caused by the SARS-CoV-2 virus (Zhang et al., 2020).

Even with the improved specificity there are still some off-target effects and unwanted integration effects to account for when using CRISPR-Cas systems, but new methods or variations are being developed. One example of an alteration is the usage of a dCas9 fused to the same FokI domain found in ZFN and TALENs. This FokI domain requires dimerization (see chapter 3.4.1, 3.4.2) and this decreases off-target effects (Tsai et al., 2014). The delivery method also has an influence on the efficiency of the CRISPR system. When injecting a purified Cas9 ribonucleoprotein in contrast to injecting plasmids coding for the Cas9 and sgRNA, the off-target effects can be reduced and there is less unwanted integration (S. Kim et al., 2014).

Table 3: Comparison of the discussed programmable nucleases

	ZFN	**TALEN**	**Cas9**
Target sequence	9-18 bp	14-20 bp	±23 bp
Ease of engineering	Difficult	Moderate	Easy
Target recognition	Protein-DNA	Protein-DNA	RNA-DNA

3.5.3 CRISPR mediated mutagenesis in zebrafish

It has been established that the zebrafish is a great vertebrate model organism to study immunology and development and is also readily used for genetic screens (see chapters 3.2.2 & 3.2.3). CRISPR-Cas mediated methods are great for these applications. This chapter will explore the CRISPR applications in zebrafish. For the different methods of delivery in zebrafish, please refer to chapter 3.7.

The zebrafish was the first vertebrate organism to demonstrate the efficiency of CRISPR-Cas genome editing *in vivo* (Hwang et al., 2013). They discovered mutational frequencies of more

than 50%. The use of CRISPR-Cas in zebrafish allowed for unprecedented possibilities in gene editing. One such new possibility is the knockout of conditional genes. One method of achieving this is by tissue-specific expression of Cas9 (Ablain et al., 2015). For example, in Ablain et al. the erythrocyte-specific *gata1* promotor (or muscle-specific *mylz2* promotor) was used for the muscle specific expression of *cas9* which enabled tissue-specific translation of the mRNA to Cas9 protein and thus enabled tissue specific gene modification in zebrafish.

Using codon optimized *cas9*, biallelic mutations can be achieved when injecting both the *cas9* mRNA and the sgRNA in the one-cell stage of the embryo (Jao et al., 2013). CRISPR-Cas9 mediated mutagenesis is being used in zebrafish not only for knockout but also for knock-in, epigenetic editing, and live imaging (K. Liu et al., 2019). Knock-in experiments were effectively carried out using 5.7 kb donors (Auer et al., 2014) as well as with reporter genes (Ota et al., 2016). There is, in fact, a germline transmission of simultaneously targeted genes (using multiple gRNAs) (Ota et al., 2014). Large-scale deletions are also possible, with efficiencies ranging from 3%-70%. These are carried out using two gRNAs (Xiao et al., 2013). The reported mutation rates are of varying degrees. Furthermore, there are a lot of factors that can have an influence on the efficiency of mutagenesis. These factors include sgRNA design, delivery methods, and the nature of the target gene (Gagnon et al., 2014).

3.6 CRISPR-I

3.6.1 Gene silencing

RNA silencing is a gene regulatory mechanism first found in plants but later described in almost all eukaryotic organisms. The basic idea of gene silencing lies in the limiting of transcript levels. This can be achieved by either suppressing transcription (= transcriptional gene silencing) (TGS) or by RNA degradation that happens in a sequence-specific manner. The latter is referred to as RNAi (RNA interference) in animals or posttranscriptional gene silencing (PTGS) in plants (Agrawal et al., 2003). CRISPRi makes use of transcriptional gene silencing.

3.6.2 CRISPRi

Researchers have been searching for new ways to use the CRISPR system and have been repurposing the Cas9 protein for other functions than to inflict DSB's. As mentioned above, the Cas9 protein has been engineered and re-engineered for other functions.

CRISPRi, or CRISPR interference, is a knock-down technique based on the mechanisms of CRISPR-Cas9. A nuclease-deficient version of the Cas9 protein is used. This protein, called dCas9 (or dead Cas9), still possesses the DNA-binding activity found in regular CRISPR-Cas9 systems, but will not introduce DSB's in the target DNA. Instead, the protein will bind the target DNA, which sterically hinders gene transcription by RNA polymerase. (Kampmann, 2018). This means that only the protein and a sgRNA is needed to silence the target. Target silencing with CRISPRi can have efficiencies of 99.9% in *E. coli* (H Larson et al., 2012). CRISPRi was shown to be effective in mammalian cells as well. However, the efficiency of this CRISPRi system was limited (Qi et al., 2013). Combining the dCas9 protein with effector domains can be used to further regulate the transcription. Fusion of the dCas9 with a repressor domain, such as the KRAB (Krüppel associated box) domain of Kox1 (a member of the KRAB C2H2 zinc finger family) results in a more efficient repression of the transcription in eukaryotic cells (Gilbert et al., 2013). In the same manner, dCas9 has been fused to the VP64 activation domain (virus protein 64, tetrameric virus protein 16 transcription activator domain) in order to

activate gene transcription (Mali et al., 2013). This is often referred to as CRISPR activation (CRISPRa).

CRISPRi has been applied in *D. rerio* with dCas9 fusion proteins. Both mRNAs for the fusion proteins with KRAB (repressor) and with the VP160 domain (activator) were constructed *in vitro* and injected into the one-cell stages of the zebrafish embryos. Both fusion proteins are proven to have the ability to alter the endogenous gene expression in zebrafish (Long et al., 2015). However, the efficiency of using a single *in vitro* constructed gRNA was low. Using multiple gRNAs gave a better knockdown efficiency. CRISPRi has since been used effectively in gene knock-down experiments (Chowdhury et al., 2018, 2019). In Chowdhury et al., CRISPRi was used with sgRNA targeting the transcription initiation site of *tie1AS* to reduce the expression of this gene by injecting the sgRNAs together with a dCAS9-KRAB fusion protein.

3.7 METHODS OF CRISPR-CAS9 DELIVERY

3.7.1 General methods of delivery

As mentioned in chapter 3.5.2, the delivery method has an influence on the efficiency of mutagenesis. A successful delivery method should deliver the CRISPR-Cas9 components into the cell nucleus (Knight et al., 2015). The CRISPR-Cas9 based mutagenesis methods contain two main components: The Cas9 protein (native, nCas9, dCas9) and the sgRNA. These can be delivered into the host in different ways. This chapter aims to give an overview of the different available methods and their characteristics.

Because the protein and sgRNA are macromolecules, they have difficulty entering the cell nucleus. Different forms have been developed, including plasmids, *cas9* mRNA and sgRNA, and direct protein delivery (complexed Cas9-RNP). These can be delivered in the cell using different techniques. Some common techniques include viral vectors, non-viral vectors such as nano-particles, and physical methods including micro-injection and electroporation (Chandrasekaran et al., 2018).

Plasmids encoding the Cas9 protein and the sgRNA have been used in mice with relatively high efficiencies. The plasmids were injected into pronuclei of fertilized eggs (Mashiko et al., 2014). The plasmids usually contain a promotor, NLS, and sometimes even green fluorescent protein genes to follow expression *in vivo* (Cong et al., 2013).

Four major viral vectors have been utilized in CRISPR delivery: lentiviral, retroviral, adenoviral, and adeno-associated viral (AAV) vectors. Because of the relatively small size of the *cas9* (4.2 kb for spcas9) and gRNA, the CRISPR system is compatible with all four of these vectors (Schmidt & Grimm, 2015). In viral vectors, the packaging size is possibly a limiting factor, and this certainly needs attention. The *spcas9* is 4.2 kb and the used gRNA cassette is usually not longer than 700 bp. AAV vectors have the smallest packaging size of these four, with a maximum packaging size of 4.9 kb (Dong et al., 1996). Exceeding this size, will diminish the transfection efficiency (Wu et al., 2010). This is, however, still big enough for the 4.2 kb *spcas9* and gRNA but leaves only 0.7kb for the signals needed for (eukaryotic) expression i.e. the eukaryotic promotor and poly-A signal. A packaging size of 0.7kb can be enough to create an AAV vector carrying the needed signals. The other viral vectors have higher packaging size upper limits. This is around 7 kb for retroviral vectors and around 30 kb for adenoviral vectors (Bouard et al., 2009; Kumar et al., 2001). It has even been found that AAV vectors that express the gRNA were similarly potent as plasmid co-transfection (Senís et al., 2014). Apart from the limiting packaging sizes, the viral vectors can also provoke immune responses that can hinder

(repeated) gene delivery and are known to promote oncogenesis (Baum et al., 2006). The limiting factor of packaging size can partly be resolved by using smaller engineered variants of the *Cas9* such as the saCas9 instead of the spCas9 (Ran et al., 2015) or a Cas9 orthologue derived from *Campylobacter jejuni* (E. Kim et al., 2017).

Plasmid delivery is highly inefficient unless a vesicle or physical delivery method is used. Plasmid vectors are considerably easier to produce than viral vectors and show a low immunogenicity. This makes gene delivery through plasmids an attractive method, granted that a high delivery efficiency can be obtained. Direct injection of plasmid DNA shows low efficiencies and is dependent on a lot of factors, such as the solute, the plasmid design, but also the injection technique. Expression when using electroporation as a physical method, improved 100- to 1000-fold compared to injecting naked plasmid DNA (Wells, 2004). Electroporation, or electropermeabilization, is a technique where the plasmid delivery is followed by electrical pulses (Zou et al., 2016). Using electroporation can damage the cells, which in turn lowers the efficiency (McMahon et al., 2001). This should be considered when deciding on the delivery method.
Direct protein delivery uses pre-formed RNP complexes to induce the genetic modification in the organism. Electroporation of pre-formed RNP in diploid *Medaka* (fish) cells achieved mutation efficiencies of 61.5% (Q. Liu et al., 2018). This can also be delivered using gold particles or other nanoparticles (Mout et al., 2017).

Nanoparticles package the components in a bio-compatible vehicle. One kind of nanoparticle being used is based on zwitterionic amino lipids (ZALs) and is called the ZAL nanoparticle (ZNP). ZNP delivery enables high protein expression *in vitro* and *in vivo* even when administered in low doses (J. B. Miller et al., s2017). Magnetic nanoparticles showed rapid transfection with fairly high efficiency both *in vitro* and *in vivo*. With magnetic nanoparticles, the genes are attached to magnetic nanoparticles and are focused to the target cells with high-gradient magnets (Dobson, 2006). Other non-viral vectors include PEGylated nanoparticles, lipoplexes, polyplexes, and more (S. D. Li & Huang, 2006). These are mostly used for tumor therapy and will therefore, not be considered relevant in this review.

3.7.2 Delivery in zebrafish

The following section will primarily focus on the delivery in the one-cell stage of zebrafish, as this is the scope of the thesis. Genetic modification can be done, not only during the single cell stage, but also during other larval stages and even in mature zebrafish (Zou et al., 2016). Delivery in zebrafish embryos is established using microinjection, into the cell during the single cell stage (Figure 12). One possibility is the co-injection of *cas9* mRNA and GOI-targeting sgRNA (Albadri et al., 2017; Hwang et al., 2013; Ota et al., 2016; Prykhozhij et al., 2018; Sung et al., 2014) and when using HDR, the donor DNA is also co-injected (Auer et al., 2014). In this approach, the (codon-optimized) *cas9* is translated, followed by *in vivo* folding of the protein and formation of the RNP complexes. The embryonic cells that develop sufficiently high levels of the Cas9-gRNA RNP become mutated and these cells further develop to become mosaic zebrafish. The mutations manifest in a dose-dependent manner. The sgRNA and the *cas9* mRNA can be constructed *in vitro* prior to injecting; The sgRNA can be synthetized directly using PCR or oligonucleotides when there is no need for labor-intensive cloning procedures (Hruscha et al., 2013). The *cas9* mRNA can be constructed by cloning a plasmid containing *cas9* in bacteria, linearizing these plasmids, and transcribing them *in vitro* (e.g. mMESSAGE mMACHINE SP6 kit)(Invitrogen, n.d.).

Instead of injecting cas9 mRNA, Cas9 protein can be used. In these studies, the *in vitro* assembled Cas9-gRNA RNP complexes are injected. These complexes are active immediately

after injection and seem to have higher efficiency compared to *cas9* mRNA (Kotani et al., 2015a). In one study, when using Cas9 protein, the first mutations were discovered as early as 4 hours post-injection (hpi) while injecting *cas9* mRNA and the specific sgRNA showed mutations at around 6 hpi (Sung et al., 2014). It is also argued that this method has a lower chance of sgRNA degradation because it is already complexed in the RNP (Burger et al., 2016). Plasmid injections (with the possibility to add a tissue specific promotor) are also being used (Ablain et al., 2015). The plasmids contain the sgRNA, *cas9* and sometimes encoded marker such as Green Fluorescent Protein (GFP). It is possible to inject multiple gRNAs (Jao et al., 2013) and thus induce multiple gene disruptions in F0 zebrafish embryos.

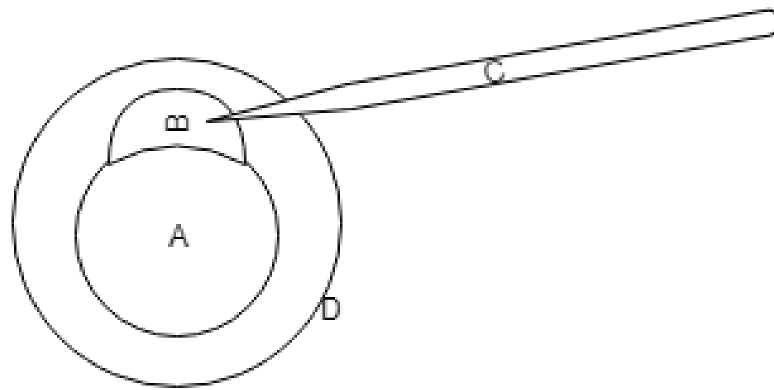


Figure 12: Representation of microinjection in single cell stage of zebrafish embryos. The single cell stage is shown at 0.2 hpf, the 2-cell stage is shown at 0.75 hpf. A is the yolk; B is the single cell; C is the needle used for injection (usually borosilicate), and D is the chorion.

3.8 POST-MUTATIONAL CONTROL

Following the CRISPR injection, it is important to find and select the wanted mutant zebrafish before continuing the assay. There are multiple ways of achieving this. The most commonly used techniques on zebrafish embryos will be discussed in this chapter.

3.8.1 Genotyping

Genotyping determines the genetic changed caused by injection of CRISPR-Cas9 components. Genotyping is usually done in combination with phenotyping or with the usage of markers(N. Chang et al., 2013; Kotani et al., 2015a), such as GFP. Phenotyping will only be applicable if a clear connection between phenotype and genotype has been shown. All genotyping methods are based on the same few steps: the first step is the sample collection where cells or tissue is harvested from the zebrafish. In the second step, the goal is to isolate the DNA from these samples. This should be done with regard to the quality as it is important that the isolated DNA is of sufficient quality for the downstream applications. The finals steps are the actual genotyping steps. Usually these consist of a target-specific polymerase chain reaction (PCR) followed by an analysis technique. Some techniques aim to automate this process in order to make it less labor-intensive and thus more suited for upscaling or high-scale screens. One example of an automated system was coined the “ZEG” or Zebrafish Embryo Genotyper, where the genetic material of 24 embryos or larvae can be obtained simultaneously. (Lambert et al., 2018)

3.8.2 Sample collection and preparation

Tissue collection and DNA extraction can be done following numerous different protocols depending on the downstream applications. For preliminary research where only the efficiency of modification needs to be assessed, the DNA can be extracted from lysed embryonic cells (Foley et al., 2009; Simpson et al., 1999), eggs (Aranishi, 2006) or tissue, such as larval eyes (Hyde et al., 2005) or from fin tissue (Zilberman et al., 2006). When the zebrafish is needed for later experiments, non-lethal (and preferably non-invasive) methods should be utilized to ensure the survivability of the zebrafish and the possibility of further applications. A very feasible method is the tail biopsy of a live zebrafish embryo. This can be done as early as 72 hpf (larval stage) and the tail will regenerate as the zebrafish grows (Wilkinson et al., 2013). This allows for early selection of mutant fish lines (sometimes referred to as ‘crispants’ (Burger et al., 2016)). After the tail biopsy (or other non-lethal tissue collecting) and subsequent genotyping, the crispants with desired mutations can be separated and/or raised to adult age for further experiments/assays and possibly to out-cross the chimer founders. Fin or tail clipping from embryos, however, is labor-intensive and requires technical skills. It is still one of the most commonly used methods for non-lethal DNA extraction. The DNA can also be extracted in a non-invasive way from environmental DNA (Turner et al., 2015) from larvae (Janelle Espinoza et al., 2017). There is also the “shake-and-stew” method where DNA is released from individual fish larvae without the use of biopsies or destructive interactions with the subject (Alvarado Bremer et al., 2014). In this method, epidermal cells are released from the larvae by vortexing. After which the epidermal cells are lysed with heated alkaline solution.

Adult tissues can also be harvested by tail or fin clippings. Using hot sodium and Tris, DNA of sufficient quality can be obtained in as little as 10-20 minutes (Meeker et al., 2007). Less invasive techniques are sometimes preferred. This can be done for post-larval zebrafish by swabbing the animal with sterile paper to collect cells and extracting DNA from these cells using chelating-resin methods (Mirimin et al., 2011). Harvesting of fish scales and DNA

extraction from these scales is possible and is a method with little experimental interference in subsequent experiments (Yue & Orban, 2001).

3.8.3 Genomic analysis

When the DNA is extracted from the expected crispant, the DNA should be genotyped. The most common methods rely on PCR followed by an analysis technique. In its most basic form, the PCR is followed by (agarose) gel electrophoresis ((A)GE) or western blot (Kotani et al., 2015) , sometimes in combination with a restriction digest. PCR combined with GE is widely used but is only applicable when the mutated locus has big indels that would be clearly visible on gel. A small alteration of this method is the restriction fragment length polymorphism (RFLP) analysis (Schmid et al., 2013). RFLP-PCR relies on a point mutations or single nucleotide polymorphisms (SNPs) inserting or deleting a restriction enzyme recognition site making this method relatively limited. Comparing the WT DNA with the expected mutant DNA will show distinct size differences when analyzing using AGE. The AGE can be substituted for a polyacrylamide GE (PAGE, polyacrylamide gel electrophoresis). PAGE has a higher sensitivity and smaller insertions or deletions can be visualized. For a more precise and sensitive genotyping method, the PCR product can be analyzed using Sanger sequencing or next-generation sequencing (ngs) (N. Chang et al., 2013). Sequencing and ngs can take up to several days to get the result and do not allow for large-scale screens. Other frequently used methods are heteroduplex mobility assays. In this method, cycles of denaturing and annealing are used on a mixture of expected mutant DNA and WT DNA. This can, for example, be done using a PCR thermocycler (Nagamine et al., 1989). During these cycles, heteroduplexes and homoduplexes will be formed when there is a mutation present. Heteroduplexes show a reduced mobility on gels and point mutations can effectively be visualized. There are also protocols available where heteroduplex assays are coupled with capillary electrophoresis systems (Kozlowski, 2001). Another heteroduplex mismatch assay is called the TAGE (T7-digestion AGE) assay and uses a T7 endonuclease (Chen et al., 2012). The T7 endonuclease I recognizes and cuts heteroduplex DNA at the site of mismatches, giving shorter fragments compared to WT homoduplexes (Babon et al., 2003). The size difference can be visualized using AGE.

More efficient methods to supplement the ones mentioned above have been developed but are not yet as frequently used. One such method is based on high-resolution melting analysis (HRMA). HRMA is a sensitive method capable of detecting SNPs and other mutations (Xing et al., 2014). The PCR-HRMA method is based on the differences in thermal dissociation characteristics of the dsDNAs between WT and mutagen DNA.

3.8.4 Phenotypic analysis

Supplementing the genomic analysis is usually a phenotypic analysis. Phenotyping (i.e. assessing the change of the phenotype) is an easy, convenient, and accessible method. The phenotype of an organism consists of the set of observable characteristics and is dependent on the genome of the organism. There is a genotype-phenotype association and in order to successfully use phenotyping, this association should be known. Phenotyping can only be done if the wanted mutation causes a distinct phenotypic change in the zebrafish (Ablain et al., 2015; Sun, 2004). This method is quick and easy and does not need extra equipment as the phenotype of the larvae or adult zebrafish can be assessed using (fluorescence) microscopy. There are different possible phenotypes, these can range from behavioral changes (C. X. Liu et al., 2018) to a lowered immune response, loss-of-function phenotypes (Burger et al., 2016) and null-phenotypes for knock-down experiments (Jao et al., 2013).

4 SUMMARY AND CONCLUSIONS

The zebrafish has shown to be very advantageous in research and was at the foundation of many discoveries. The many qualities of zebrafish quickly attracted the attention of many scientists and with good reason. The extensive use of zebrafish and the research and discoveries coupled with it, has made the zebrafish a great model organism for multiple fields. The similarities of the human and zebrafish immune systems provide the zebrafish model organism with strong advantages to be used in immunological studies.

Guided nucleases revolutionized the scientific world and have made an enormous impact on research. The revolution started with ZFN but when researchers stumbled on the design difficulties, they soon adopted TALENs as the go-to method. While TALENs are still being used to this day, the focus shifted to the CRISPR-Cas system and the CRISPR-based methods. The adaptation of a single sgRNA to guide the Cas9 endonuclease to the target site reformed the use of guided nucleases.

Combining the flexibility and ease of use of the CRISPR system with the advantages of the zebrafish model organism, results in great potential for genetics, immunology, and pharmacology. New improvements to the CRISPR system continue to be made, making it not only a great current system, but also a system with tremendous prospects. CRISPRi is one alteration of the CRISPR system and is also being improved continuously. Although the usage of the CRISPRi system in zebrafish is rather limited at this time, it seems like this system will be used far more in the future and has potential for genetic knock-down studies in the zebrafish. When working with CRISPR-based methods in zebrafish, there are multiple techniques at your disposal, each technique offering their own advantages and disadvantages. Experimental design is of uttermost importance to assure a successful outcome and clear data, therefore it’s important to be aware of the different possibilities and take these into account before starting research.

“The power to control our species’ genetic future is awesome and terrifying. Deciding how to handle it may be the biggest challenge we have ever faced.”

-Jennifer A Doudna, A Crack in Creation: gene editing and the unthinkable power to control evolution.

5 REFERENCES

Ablain, J., Durand, E. M., Yang, S., Zhou, Y., & Zon, L. I. (2015). A CRISPR/Cas9 Vector System for Tissue-Specific Gene Disruption in Zebrafish. *Developmental Cell*, *32*(6), 756–764. https://doi.org/10.1016/j.devcel.2015.01.032

Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-04252-2

Agrawal, N., Dasaradhi, P., Mohammed, A., Malhotra, P., Bhatnagar, R. K., & Mukherjee, S. K. (2003). RNA Interference: Biology, Mechanism, and Applications. *Microbiology and Molecular Biology Reviews*, *67*(4), 657–685. https://doi.org/10.1128/MMBR.67.4.657

Albadri, S., De Santis, F., Di Donato, V., & Del Bene, F. (2017). CRISPR/Cas9-Mediated Knockin and Knockout in Zebrafish. In *Research and Perspectives in Neurosciences* (Issue 9783319601915, pp. 41–49). Springer Verlag. https://doi.org/10.1007/978-3-319-60192-2_4

Allen, J. P., & Neely, M. N. (2010). Trolling for the ideal model host: zebrafish take the bait. *Future Microbiology*, *5*(4), 563–569. https://doi.org/10.2217/fmb.10.24

Alvarado Bremer, J. R., Smith, B. L., Moulton, D. L., Lu, C. P., & Cornic, M. (2014). Shake and stew: A non-destructive PCR-ready DNA isolation method from a single preserved fish larva. *Journal of Fish Biology*, *84*(1), 267–272. https://doi.org/10.1111/jfb.12280

Amatruda, J. F., Shepard, J. L., Stern, H. M., & Zon, L. I. (2002). Zebrafish as a cancer model system. *Cancer Cell*, *1*(3), 229–231. https://doi.org/10.1016/S1535-6108(02)00052-1

Amitai, G., & Sorek, R. (2016). CRISPR-Cas adaptation: Insights into the mechanism of action. *Nature Reviews Microbiology*, *14*(2), 67–76. https://doi.org/10.1038/nrmicro.2015.14

Amsterdam, A., Burgess, S., Golling, G., Chen, W., Sun, Z., Townsend, K., Farrington, S., Haldi, M., & Hopkins, N. (1999). A large-scale insertional mutagenesis screen in zebrafish. *Genes and Development*, *13*(20), 2713–2724. https://doi.org/10.1101/gad.13.20.2713

Aranishi, F. (2006). Single fish egg DNA extraction for PCR amplification. *Conservation Genetics*, *7*(1), 153–156. https://doi.org/10.1007/s10592-005-5387-y

Auer, T. O., Duroure, K., De Cian, A., Concordet, J.-P., & Del Bene, F. (2014). Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Research*, *24*(1), 142–153. https://doi.org/10.1101/gr.161638.113

Babon, J. J., McKenzie, M., & Cotton, R. G. H. (2003). The Use of Resolvases T4 Endonuclease VII and T7 Endonuclease I in Mutation Detection. *Molecular Biotechnology*, *23*(1), 73–82. https://doi.org/10.1385/MB:23:1:73

Bambino, K., & Chu, J. (2017). Zebrafish in Toxicology and Environmental Health. In *Current Topics in Developmental Biology* (Vol. 124, pp. 331–367). Academic Press Inc. https://doi.org/10.1016/bs.ctdb.2016.10.007

Bannikov, A. V., & Lavrov, A. V. (2017). CRISPR/CAS9, the king of genome editing tools. *Molecular Biology*, *51*(4), 514–525. https://doi.org/10.1134/S0026893317040033

Baum, C., Kustikova, O., Modlich, U., Li, Z., & Fehse, B. (2006). Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors. *Human Gene Therapy*, *17*(3), 253–263. https://doi.org/10.1089/hum.2006.17.253

Baum, C., von Kalle, C., Staal, F. J. T., Li, Z., Fehse, B., Schmidt, M., Weerkamp, F., Karlsson, S., Wagemaker, G., & Williams, D. A. (2004). Chance or necessity? Insertional Mutagenesis in Gene Therapy and Its Consequences. *Molecular Therapy*, *9*(1), 5–13. https://doi.org/10.1016/j.ymthe.2003.10.013

Bertrand, J. Y., & Traver, D. (2009). Hematopoietic cell development in the zebrafish embryo. *Current Opinion in Hematology*, *16*(4), 243–248. https://doi.org/10.1097/MOH.0b013e32832c05e4

Beumer, K. J., Trautman, J. K., Bozas, A., Liu, J. L., Rutter, J., Gall, J. G., & Carroll, D. (2008). Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(50), 19821–19826. https://doi.org/10.1073/pnas.0810475105

Bi, H., & Yang, B. (2017). Gene Editing With TALEN and CRISPR/Cas in Rice. In *Progress in Molecular Biology and Translational Science* (1st ed., Vol. 149, pp. 81–98). Elsevier Inc. https://doi.org/10.1016/bs.pmbts.2017.04.006

Bitinaite, J., Wah, D. A., Aggarwal, A. K., & Schildkraut, I. (1998). Foki dimerization is required for dna cleavage. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(18), 10570–10575. https://doi.org/10.1073/pnas.95.18.10570

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., & Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, *326*(5959), 1509–1512. https://doi.org/10.1126/science.1178811

Bogdanove, A. J., & Voytas, D. F. (2011). TAL effectors: Customizable proteins for DNA targeting. *Science*, *333*(6051), 1843–1846. https://doi.org/10.1126/science.1204094

Bonas, U. (1990). Isolation of a gene cluster from Xanthomonas campestris pv. vesicatoria that determines pathogenicity and the hypersensitive response on pepper and tomato.pdf. *Mol Plant Microbe Interact*, *4*, 81–88. https://doi.org/10005382285

Bouard, D., Alazard-Dany, N., & Cosset, F.-L. (2009). Viral vectors: from virology to transgene expression. *British Journal of Pharmacology*, *157*(2), 153–165. https://doi.org/10.1038/bjp.2008.349

Broughton, J. P., Deng, X., Yu, G., Fasching, C. L., Servellita, V., Singh, J., Miao, X., Streithorst, J. A., Granados, A., Sotomayor-gonzalez, A., Zorn, K., Gopez, A., Hsu, E., Gu, W., Miller, S., Pan, C., Guevara, H., Wadford, D. A., Chen, J. S., & Chiu, C. Y. (2019). *CRISPR – Cas12-based detection of SARS-CoV-2*.

Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., Zaugg, J., Weber, L. M., Catena, R., Jinek, M., Robinson, M. D., & Mosimann, C. (2016). Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development*, *143*(11), 2025–2037. https://doi.org/10.1242/dev.134809

Carroll, D. (2011). *Genome Engineering With Zinc-Finger Nucleases*. https://doi.org/10.1534/genetics.111.131433

Chandrasegaran, S., & Carroll, D. (2016). Origins of Programmable Nucleases for Genome Engineering. *Journal of Molecular Biology*, *428*(5), 963–989. https://doi.org/10.1016/j.jmb.2015.10.014

Chandrasekaran, A. P., Song, M., Kim, K. S., & Ramakrishna, S. (2018). Different Methods of Delivering CRISPR/Cas9 Into Cells. In *Progress in Molecular Biology and Translational Science* (1st ed., Vol. 159). Elsevier Inc. https://doi.org/10.1016/bs.pmbts.2018.05.001

Chang, H. H. Y., Pannunzio, N. R., Adachi, N., & Lieber, M. R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair HHS Public Access. *Nat Rev Mol Cell Biol*, *18*(8), 495–506. https://doi.org/10.1038/nrm.2017.48

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J. W., & Xi, J. J. (2013). Genome editing with RNA-guided Cas9 nuclease in Zebrafish embryos. *Cell Research*, *23*(4), 465–472. https://doi.org/10.1038/cr.2013.45

Chen, J., Zhang, X., Wang, T., Li, Z., Guan, G., & Hong, Y. (2012). Efficient detection, quantification and enrichment of subtle allelic alterations. *DNA Research*, *19*(5), 423–433. https://doi.org/10.1093/dnares/dss023

Chowdhury, T. A., Koceja, C., Eisa-Beygi, S., Kleinstiver, B. P., Kumar, S. N., Lin, C. W., Li, K., Prabhudesai, S., Joung, J. K., & Ramchandran, R. (2018). Temporal and spatial post-transcriptional regulation of zebrafish tie1 mRNA by Long Noncoding RNA during brain vascular assembly. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *38*(7), 1562–1575. https://doi.org/10.1161/ATVBAHA.118.310848

Chowdhury, T. A., Li, K., Ramchandran, R., Daniel, J. F. N., David L, K. S., David H, W., Holmes, Chowdhury, T. A., Li, K., Ramchandran, R., Daniel, J. F. N., David L, K. S., David H, W., & Holmes. (2019). Lessons learned from a lncRNA odyssey for two genes with vascular functions, DLL4 and TIE1. *Vascular Pharmacology*, *114*(1), 103–109. https://doi.org/10.1016/j.vph.2018.06.010

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A. J., & Voytas, D. F. (2010). Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics*, *186*(2), 757–761. https://doi.org/10.1534/genetics.110.120717

Chylinski, K., Makarova, K. S., Charpentier, E., & Koonin, E. V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Research*, *42*(10), 6091–6105. https://doi.org/10.1093/nar/gku241

Condreay, J. P., Witherspoon, S. M., Clay, W. C., & Kost, T. A. (1999). Transient and stable gene expression in mammalian cells transduced with a recombinant baculovirus vector. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(1), 127–132. https://doi.org/10.1073/pnas.96.1.127

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823. https://doi.org/10.1126/science.1231143

Davidson, A. J., & Zon, L. I. (2004). The "definitive" (and 'primitive') guide to zebrafish hematopoiesis. In *Oncogene* (Vol. 23, Issue 43 REV. ISS. 6, pp. 7233–7246). Nature Publishing Group. https://doi.org/10.1038/sj.onc.1207943

Dobson, J. (2006). Gene therapy progress and prospects: Magnetic nanoparticle-based gene delivery. In *Gene Therapy* (Vol. 13, Issue 4, pp. 283–287). Nature Publishing Group. https://doi.org/10.1038/sj.gt.3302720

Dong, J., Fan, P., & Frizzell, R. A. (1996). Quantitative Analysis of the Packaging Capacity of Recombinant Adeno-Associated Virus. *Human Gene Therapy*, *7*(17), 2101–2112. https://doi.org/10.1089/hum.1996.7.17-2101

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213). https://doi.org/10.1126/science.1258096

Doyon, Y., McCammon, J. M., Miller, J. C., Faraji, F., Ngo, C., Katibah, G. E., Amora, R., Hocking, T. D., Zhang, L., Rebar, E. J., Gregory, P. D., Urnov, F. D., & Amacher, S. L. (2008). Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nature Biotechnology*, *26*(6), 702–708. https://doi.org/10.1038/nbt1409

Driever, W., Solnica-Krezel, L., Schier, A. F., Neuhauss, S. C. F., Malicki, J., Stemple, D. L., Stainier, D. Y. R., Zwartkruis, F., Abdelilah, S., Rangini, Z., Belak, J., & Boggs, C. (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development*, *123*, 37–46. https://doi.org/10.5167/uzh-215

Eisen, J. S., & Smith, J. C. (2008). Controlling morpholino experiments: don't stop making antisense. *Development*, *135*(10), 1735–1743. https://doi.org/10.1242/dev.001115

El-Brolosy, M. A., & Stainier, D. Y. R. (2017). Genetic compensation: A phenomenon in search of mechanisms. *PLOS Genetics*, *13*(7), e1006780. https://doi.org/10.1371/journal.pgen.1006780

Engeszer, R. E., Patterson, L. B., Rao, A. A., & Parichy, D. M. (2007). Zebrafish in The Wild: A Review of Natural History And New Notes from The Field. *Zebrafish*, *4*(1), 21–40. https://doi.org/10.1089/zeb.2006.9997

Fineran, P. C., & Charpentier, E. (2012). Memory of viral infections by CRISPR-Cas adaptive immune systems: Acquisition of new information. *Virology*, *434*(2), 202–209. https://doi.org/10.1016/j.virol.2012.10.003

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent andspecific genetic interferenceby double-strandedRNA in Caenorhabditiselegans. *Nature*, *391*(February), 806–811.

Foley, J. E., Yeh, J.-R. J., Maeder, M. L., Reyon, D., Sander, J. D., Peterson, R. T., & Joung, J. K. (2009). Rapid Mutation of Endogenous Zebrafish Genes Using Zinc Finger Nucleases Made by Oligomerized Pool ENgineering (OPEN). *PLoS ONE*, *4*(2), e4348. https://doi.org/10.1371/journal.pone.0004348

Franklin, R. E., & Gosling, R. G. (1995). Molecular structure of nucleic acids. Molecular configuration in sodium thymonucleate. 1953.Franklin, R. E., & Gosling, R. G. (1995). Molecular structure of nucleic acids. Molecular configuration in sodium thymonucleate. 1953. Annals of the New York Acade. *Annals of the New York Academy of Sciences*, *758*, 16–17. https://doi.org/10.1038/171740a0

Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T. G., Zimmerman, S., Richter, C., & Schier, A. F. (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS ONE*, *9*(5), 5–12. https://doi.org/10.1371/journal.pone.0098186

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(39), E2579–E2586. https://doi.org/10.1073/pnas.1208507109

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*. https://doi.org/10.1016/j.cell.2013.06.044

Goldsmith, J. R., & Jobin, C. (2012). Think Small: Zebrafish as a Model System of Human Pathology. *Journal of Biomedicine and Biotechnology*, *2012*, 1–12. https://doi.org/10.1155/2012/817341

Guo, S., & Kemphues, K. J. (1995). par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell*, *81*(4), 611–620. https://doi.org/10.1016/0092-8674(95)90082-9

Gupta, S. K., & Shukla, P. (2017). Gene editing for cell engineering: trends and applications. *Critical Reviews in Biotechnology*, *37*(5), 672–684. https://doi.org/10.1080/07388551.2016.1214557

H. Meijer, A., & P. Spaink, H. (2011). Host-Pathogen Interactions Made Transparent with the Zebrafish Model. *Current Drug Targets*, *12*(7), 1000–1017. https://doi.org/10.2174/138945011795677809

H Larson, M., A Gilbert, L., Xiaowo, W., A Lim, W., S Weissman, J., & S Qi, L. (2012). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Bone*, *23*(1), 1–7. https://doi.org/10.1038/jid.2014.371

Hall, C. J., Flores, M. V., Crosier, K. E., & Crosier, P. S. (2008). Live imaging early immune cell ontogeny and function in zebrafish Danio rerio. *Journal of Fish Biology*, *73*(8), 1833–1871. https://doi.org/10.1111/j.1095-8649.2008.01980.x

Hansson, G. K., & Edfeldt, K. (2005). Toll to be paid at the gateway to the vessel wall. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *25*(6), 1085–1087. https://doi.org/10.1161/01.ATV.0000168894.43759.47

He, Y., Wang, M., Liu, M., Huang, L., Liu, C., Zhang, X., Yi, H., Cheng, A., Zhu, D., Yang, Q., Wu, Y., Zhao, X., Chen, S., Jia, R., Zhang, S., Liu, Y., Yu, Y., & Zhang, L. (2018). Cas1 and Cas2 from the type II-C CRISPR-cas system of Riemerella anatipestifer are required for spacer acquisition. *Frontiers in Cellular and Infection Microbiology*, *8*(JUN). https://doi.org/10.3389/fcimb.2018.00195

Heasman, J., Kofron, M., & Wylie, C. (2000). β-catenin signaling activity dissected in the early Xenopus embryo: A novel antisense approach. *Developmental Biology*, *222*(1), 124–134. https://doi.org/10.1006/dbio.2000.9720

Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase

activates genes from promoters and enhancers. *Nature Biotechnology*, *33*(5), 510–517. https://doi.org/10.1038/nbt.3199

Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., Cost, G. J., Zhang, L., Santiago, Y., Miller, J. C., Zeitler, B., Cherone, J. M., Meng, X., Hinkley, S. J., Rebar, E. J., Gregory, P. D., Urnov, F. D., & Jaenisch, R. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nature Biotechnology*, *29*(8), 731–734. https://doi.org/10.1038/nbt.1927

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., Collins, J. E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J. C., Koch, R., Rauch, G.-J., White, S., … Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, *496*(7446), 498–503. https://doi.org/10.1038/nature12111

Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C., & Schmid, B. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development (Cambridge)*, *140*(24), 4982–4987. https://doi.org/10.1242/dev.099085

Hustedt, N., & Durocher, D. (2017). The control of DNA repair by the cell cycle. *Nature Cell Biology*, *19*(1), 1–9. https://doi.org/10.1038/ncb3452

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Shengdar, Q., Sander, J. D., Peterson, R. T., Yeh, J. J., & Keith, J. (2013). Efficient In Vivo Genome Editing Using RNA-Guided Nucleases. *Nature Biotechnology*, *31*(3), 227–229. https://doi.org/10.1038/nbt.2501.Efficient

Hyde, J. R., Lynn, E., Humphreys, R., Musyl, M., West, A. P., & Vetter, R. (2005). Shipboard identification of fish eggs and larvae by multiplex PCR, and description of fertilized eggs of blue marlin, shortbill spearfish, and wahoo. *Marine Ecology Progress Series*, *286*, 269–277. https://doi.org/10.3354/meps286269

Invitrogen. (n.d.). *mMESSAGE mMACHINE™ SP6 Transcription Kit*. Retrieved May 17, 2020, from https://www.thermofisher.com/order/catalog/product/AM1340#/AM1340

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakatura, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isoenzyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429–5433. https://doi.org/10.1128/jb.169.12.5429-5433.1987

Janelle Espinoza, G., Michael Poland, J., & Alvarado Bremer, J. R. (2017). Genotyping live fish larvae: Non-lethal and noninvasive DNA isolation from 3-5 day old hatchlings. *BioTechniques*, *63*(4), 181–186. https://doi.org/10.2144/000114598

Jao, L. E., Wente, S. R., & Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(34), 13904–13909. https://doi.org/10.1073/pnas.1308335110

Jiang, F., & Doudna, J. A. (2017). *CRISPR – Cas9 Structures and Mechanisms*. 505–529.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., & Doudna, J. A. (2012). *A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity*. *91*(June), 1052–1055. https://doi.org/10.1002/ajh.24462

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., & Doudna, J. (2013). RNA-programmed genome editing in human cells. *ELife*, *2013*(2), 1–9. https://doi.org/10.7554/eLife.00471

Kampmann, M. (2018). CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chemical Biology*, *13*(2), 406–416. https://doi.org/10.1021/acschembio.7b00657

Kanther, M., & Rawls, J. F. (2010). Host–microbe interactions in the developing zebrafish. *Current Opinion in Immunology*, *22*(1), 10–19. https://doi.org/10.1016/j.coi.2010.01.006

Kay, S., Hahn, S., Marois, E., Hause, G., & Bonas, U. (2007). A Bacterial Effector Acts as a Plant Transcription Factor and Induces a Cell Size Regulator. *Science*, *318*(5850), 648–

651. https://doi.org/10.1126/science.1144956

Kellner, M. J., Koob, J. G., Gootenberg, J. S., Abudayyeh, O. O., & Zhang, F. (2019). SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nature Protocols*, *14*(10), 2986–3012. https://doi.org/10.1038/s41596-019-0210-2

Kelly, T. J., & Smith, H. O. (1970). A restriction enzyme from Hemophilus influenzae II: Base sequence of the Recognition site. *Journal of Molecular Biology*, *51*(2), 393–409. https://doi.org/10.1016/0022-2836(70)90150-6

Kim, E., Koo, T., Park, S. W., Kim, D., Kim, K., Cho, H. Y., Song, D. W., Lee, K. J., Jung, M. H., Kim, S., Kim, J. H., Kim, J. H., & Kim, J. S. (2017). In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. *Nature Communications*, *8*. https://doi.org/10.1038/ncomms14500

Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J. S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, *24*(6), 1012–1019. https://doi.org/10.1101/gr.171322.113

Kim, T. K., & Eberwine, J. H. (2010). Mammalian cell transfection: The present and the future. *Analytical and Bioanalytical Chemistry*, *397*(8), 3173–3178. https://doi.org/10.1007/s00216-010-3821-6

Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(3), 1156–1160. https://doi.org/10.1073/pnas.93.3.1156

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Developmental Dynamics*, *203*(3), 253–310. https://doi.org/10.1002/aja.1002030302

Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, *5*(11), 826–837. https://doi.org/10.1038/nrg1471

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., & Joung, J. K. (2016). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, *529*(7587), 490–495. https://doi.org/10.1038/nature16526

Knight, S. C., Xie, L., Deng, W., Guglielmi, B., Witkowsky, L. B., Bosanac, L., Zhang, E. T., Beheiry, M. E., Masson, J. B., Dahan, M., Liu, Z., Doudna, J. A., & Tjian, R. (2015). Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science*, *350*(6262), 823–826. https://doi.org/10.1126/science.aac6572

Knott, G. J., & Doudna, J. A. (2018). CRISPR-Cas guides the future of genetic engineering. In *Science* (Vol. 361, Issue 6405, pp. 866–869). https://doi.org/10.1126/science.aat5011

Kok, F. O., Shin, M., Ni, C. W., Gupta, A., Grosse, A. S., VanImpel, A., Kirchmaier, B. C., Peterson-Maduro, J., Kourkoulis, G., Male, I., DeSantis, D. F., Sheppard-Tindell, S., Ebarasi, L., Betsholtz, C., Schulte-Merker, S., Wolfe, S. A., & Lawson, N. D. (2015). Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Developmental Cell*, *32*(1), 97–108. https://doi.org/10.1016/j.devcel.2014.11.018

Kotani, H., Taimatsu, K., Ohga, R., Ota, S., & Kawahara, A. (2015a). Efficient multiple genome modifications induced by the crRNAs, tracrRNA and Cas9 protein complex in zebrafish. *PLoS ONE*, *10*(5), 1–16. https://doi.org/10.1371/journal.pone.0128319

Kotani, H., Taimatsu, K., Ohga, R., Ota, S., & Kawahara, A. (2015b). Efficient multiple genome modifications induced by the crRNAs, tracrRNA and Cas9 protein complex in zebrafish. *PLoS ONE*, *10*(5). https://doi.org/10.1371/journal.pone.0128319

Kozlowski, P. (2001). Combined SSCP/duplex analysis by capillary electrophoresis for more efficient mutation detection. *Nucleic Acids Research*, *29*(14), 71e – 71. https://doi.org/10.1093/nar/29.14.e71

Kumar, M., Keller, B., Makalou, N., & Sutton, R. E. (2001). Systematic determination of the packaging limit of lentiviral vectors. *Human Gene Therapy*, *12*(15), 1893–1905.

https://doi.org/10.1089/104303401753153947

Lam, S. H., Chua, H. L., Gong, Z., Lam, T. J., & Sin, Y. M. (2004). Development and maturation of the immune system in zebrafish, Danio rerio: A gene expression profiling, in situ hybridization and immunological study. *Developmental and Comparative Immunology*, *28*(1), 9–28. https://doi.org/10.1016/S0145-305X(03)00103-4

Lambert, C. J., Freshner, B. C., Chung, A., Stevenson, T. J., Bowles, D. M., Samuel, R., Gale, B. K., & Bonkowsky, J. L. (2018). An automated system for rapid cellular extraction from live zebrafish embryos and larvae: Development and application to genotyping. *PLoS ONE*, *13*(3). https://doi.org/10.1371/journal.pone.0193180

Langheinrich, U., Vacun, G., & Wagner, T. (2003). Zebrafish embryos express an orthologue of HERG and are sensitive toward a range of QT-prolonging drugs inducing severe arrhythmia. *Toxicology and Applied Pharmacology*, *193*(3), 370–382. https://doi.org/10.1016/j.taap.2003.07.012

Li, L., Wu, L. P., & Chandrasegaran, S. (1992). Functional domains in Fok I restriction endonuclease. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(10), 4275–4279. https://doi.org/10.1073/pnas.89.10.4275

Li, S. D., & Huang, L. (2006). Gene therapy progress and prospects: Non-viral gene therapy by systemic delivery. In *Gene Therapy* (Vol. 13, Issue 18, pp. 1313–1319). Nature Publishing Group. https://doi.org/10.1038/sj.gt.3302838

Li, Y., Li, Y., Cao, X., Jin, X., & Jin, T. (2017). Pattern recognition receptors in zebrafish provide functional and evolutionary insight into innate immune signaling pathways. In *Cellular and Molecular Immunology* (Vol. 14, Issue 1, pp. 80–89). Chinese Soc Immunology. https://doi.org/10.1038/cmi.2016.50

Lieber, M. R. (2010). The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annual Review of Biochemistry*, *79*(1), 181–211. https://doi.org/10.1146/annurev.biochem.052308.093131

Liu, C. X., Li, C. Y., Hu, C. C., Wang, Y., Lin, J., Jiang, Y. H., Li, Q., & Xu, X. (2018). CRISPR/Cas9-induced shank3b mutant zebrafish display autism-like behaviors. *Molecular Autism*, *9*(1), 1–13. https://doi.org/10.1186/s13229-018-0204-x

Liu, K., Petree, C., Requena, T., Varshney, P., & Varshney, G. K. (2019). Expanding the CRISPR Toolbox in Zebrafish for Studying Development and Disease. *Frontiers in Cell and Developmental Biology*, *7*(MAR). https://doi.org/10.3389/fcell.2019.00013

Liu, Q., Yuan, Y., Zhu, F., Hong, Y., & Ge, R. (2018). Efficient genome editing using CRISPR/Cas9 ribonucleoprotein approach in cultured Medaka fish cells. *Biology Open*, *7*(8). https://doi.org/10.1242/bio.035170

Long, L., Guo, H., Yao, D., Xiong, K., Li, Y., Liu, P., Zhu, Z., & Liu, D. (2015). Regulation of transcriptionally active genes via the catalytically inactive Cas9 in C. elegans and D. rerio. In *Cell Research* (Vol. 25, Issue 5, pp. 638–641). Nature Publishing Group. https://doi.org/10.1038/cr.2015.35

Ma, Y., Zhang, L., & Huang, X. (2014). Genome modification by CRISPR/Cas9. *FEBS Journal*, *281*(23), 5186–5193. https://doi.org/10.1111/febs.13110

Maeder, M. L., & Gersbach, C. A. (2016). Genome-editing technologies for gene and cell therapy. *Molecular Therapy*, *24*(3), 430–446. https://doi.org/10.1038/mt.2016.10

Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., Charpentier, E., Cheng, D., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Scott, D., Shah, S. A., Siksnys, V., Terns, M. P., Venclovas, Č., White, M. F., Yakunin, A. F., … Koonin, E. V. (2020). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nature Reviews Microbiology*, *18*(2), 67–83. https://doi.org/10.1038/s41579-019-0299-x

Mali, P., Aach, J., Stranges, B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., & Church, G. M. (2013). *CAs9 transcriptional activators for target specificity screening and paired*

nickases for cooperative genome engineering. https://doi.org/10.1038/nbt.2675

Marois, E., Van Den Ackerveken, G., & Bonas, U. (2002). The Xanthomonas type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. *Molecular Plant-Microbe Interactions*, *15*(7), 637–646. https://doi.org/10.1094/MPMI.2002.15.7.637

Mashiko, D., Young, S. A. M., Muto, M., Kato, H., Nozawa, K., Ogawa, M., Noda, T., Kim, Y. J., Satouh, Y., Fujihara, Y., & Ikawa, M. (2014). Feasibility for a large scale mouse mutagenesis by injecting CRISPR/Cas plasmid into zygotes. *Development Growth and Differentiation*, *56*(1), 122–129. https://doi.org/10.1111/dgd.12113

McCluskey, B. M., & Postlethwait, J. H. (2015). Phylogeny of zebrafish, a "model species," within Danio, a "model genus." *Molecular Biology and Evolution*, *32*(3), 635–652. https://doi.org/10.1093/molbev/msu325

McMahon, J. M., Signori, E., Wells, K. E., Fazio, V. M., & Wells, D. J. (2001). Optimisation of electrotransfer of plasmid into skeletal muscle by pretreatment with hyaluronidase - Increased expression with reduced muscle damage. *Gene Therapy*, *8*(16), 1264–1270. https://doi.org/10.1038/sj.gt.3301522

Meeker, N. D., Hutchinson, S. A., Ho, L., & Trede, N. S. (2007). Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *BioTechniques*, *43*(5), 610–614. https://doi.org/10.2144/000112619

Meijer, A. H., Gabby Krens, S. F., Medina Rodriguez, I. A., He, S., Bitter, W., Snaar-Jagalska, B. E., & Spaink, H. P. (2004). Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish. *Molecular Immunology*, *40*(11), 773–783. https://doi.org/10.1016/j.molimm.2003.10.003

Mekler, V., Kuznedelov, K., & Severinov, K. (2020). Quantification of the affinities of CRISPR-Cas9 nucleases for cognate protospacer adjacent motif (PAM) sequences. *J Biol Chem*. https://doi.org/10.1074/jbc.RA119.012239

Men, W., Dong, Q., Shi, W., & Yao, K. (2019). Serotype distribution and antimicrobial resistance patterns of invasive pneumococcal disease isolates from children in mainland China—a systematic review. *Brazilian Journal of Microbiology*. https://doi.org/10.1007/s42770-019-00198-9

Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D., & Wolfe, S. A. (2008). Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nature Biotechnology*, *26*(6), 695–701. https://doi.org/10.1038/nbt1398

Menke, A. L., Spitsbergen, J. M., Wolterbeek, A. P. M., & Woutersen, R. A. (2011). Normal Anatomy and Histology of the Adult Zebrafish. *Toxicologic Pathology*, *39*(5), 759–775. https://doi.org/10.1177/0192623311409597

Meselson, M., & Yuan, R. (1968). DNA restriction enzyme from E. coli. *Nature*, *217*(5134), 1110–1114. https://doi.org/10.1038/2171110a0

Miller, J. B., Zhang, S., Kos, P., Xiong, H., Zhou, K., Perelman, S. S., Zhu, H., & Siegwart, D. J. (2017). Non-Viral CRISPR/Cas Gene Editing In Vitro and In Vivo Enabled by Synthetic Nanoparticle Co-Delivery of Cas9 mRNA and sgRNA. *Angewandte Chemie International Edition*, *56*(4), 1059–1063. https://doi.org/10.1002/anie.201610209

Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., Dulay, G. P., Hua, K. L., Ankoudinova, I., Cost, G. J., Urnov, F. D., Zhang, H. S., Holmes, M. C., Zhang, L., Gregory, P. D., & Rebar, E. J. (2011). A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology*, *29*(2), 143–150. https://doi.org/10.1038/nbt.1755

Miller, J., McLachlan, A. D., & Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *The EMBO Journal*, *4*(6), 1609–1614. https://doi.org/10.1002/j.1460-2075.1985.tb03825.x

Mirimin, L., O'Keeffe, D., Ruggiero, A., Bolton-Warberg, M., Vartia, S., & Fitzgerald, R.

(2011). A quick, least-invasive, inexpensive and reliable method for sampling Gadus morhua postlarvae for genetic analysis. *Journal of Fish Biology*, *79*(3), 801–805. https://doi.org/10.1111/j.1095-8649.2011.03049.x

Mojica, F. J. M., Díez-Villaseñor, C., Soria, E., & Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. In *Molecular Microbiology* (Vol. 36, Issue 1, pp. 244–246). https://doi.org/10.1046/j.1365-2958.2000.01838.x

Moscou, M. J., & Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science*, *326*(5959), 1501. https://doi.org/10.1126/science.1178817

Mout, R., Ray, M., Yesilbag Tonga, G., Lee, Y.-W., Tay, T., Sasaki, K., & Rotello, V. M. (2017). Direct Cytosolic Delivery of CRISPR/Cas9-Ribonucleoprotein for Efficient Gene Editing. *ACS Nano*, *11*(3), 2452–2458. https://doi.org/10.1021/acsnano.6b07600

Mukherjee, A. S., & Beermann, W. (1965). Synthesis of Ribonucleic Acid by the X-Chromosomes of Drosophila melanogaster and the Problem of Dosage Compensation. *Nature*, *207*(4998), 785–786. https://doi.org/10.1038/207785a0

Muller, H. J. (1932). Further studies on the nature and causes of gene mutations. In *International Congress of Genetics* (Vol. 6, Issue 1, pp. 213–255). http://darwin.amnh.org/bib/bibtexbrowser.php?key=muller1932yu83j&bib=EvoLit.bib&

Nagamine, C. M., Chan, K., & Lau, Y. F. (1989). A PCR artifact: generation of heteroduplexes. *American Journal of Human Genetics*, *47*(1), 169–170. https://doi.org/10.1016/S0016-7878(61)80010-2

Nasevicius, A., & Ekker, S. C. (2000). Effective targeted gene “knockdown” in zebrafish. *Nat Genet*, *26*(2), 216–220.

Neely, M. N., Pfeifer, J. D., & Caparon, M. (2002). Streptococcus-Zebrafish Model of Bacterial Pathogenesis. *INFECTION AND IMMUNITY*, *70*(7), 3904–3914. https://doi.org/10.1128/IAI.70.7.3904-3914.2002

Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., Ishitani, R., Zhang, F., & Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. https://doi.org/10.1016/j.cell.2014.02.001

Nobbs, A. H., Jenkinson, H. F., & Everett, D. B. (2015). Generic determinants of Streptococcus colonization and infection. *Infection, Genetics and Evolution*, *33*, 361–370. https://doi.org/10.1016/j.meegid.2014.09.018

Novoa, B., & Figueras, A. (2012). Zebrafish: Model for the study of inflammation and the innate immune response to infectious diseases. *Advances in Experimental Medicine and Biology*, *946*, 253–275. https://doi.org/10.1007/978-1-4614-0106-3_15

O’Callaghan, D., & Vergunst, A. (2010). Non-mammalian animal models to study infectious disease: worms or fly fishing? In *Current Opinion in Microbiology* (Vol. 13, Issue 1, pp. 79–85). Elsevier Current Trends. https://doi.org/10.1016/j.mib.2009.12.005

Ota, S., Hisano, Y., Ikawa, Y., & Kawahara, A. (2014). Multiple genome modifications by the CRISPR/Cas9 system in zebrafish. *Genes to Cells*, *19*(7), 555–564. https://doi.org/10.1111/gtc.12154

Ota, S., Taimatsu, K., Yanagi, K., Namiki, T., Ohga, R., Higashijima, S. I., & Kawahara, A. (2016). Functional visualization and disruption of targeted genes using CRISPR/Cas9-mediated eGFP reporter integration in zebrafish. *Scientific Reports*, *6*. https://doi.org/10.1038/srep34991

Ousterout, D. G., Perez-Pinera, P., Thakore, P. I., Kabadi, A. M., Brown, M. T., Qin, X., Fedrigo, O., Mouly, V., Tremblay, J. P., & Gersbach, C. A. (2013). Reading frame correction by targeted genome editing restores dystrophin expression in cells from duchenne muscular dystrophy patients. *Molecular Therapy*, *21*(9), 1718–1726. https://doi.org/10.1038/mt.2013.111

Pallarès Masmitjà, M., Knödlseder, N., & Güell, M. (2019). CRISPR-gRNA Design. *Methods*

in Molecular Biology, *1961*, 3–11. https://doi.org/10.1007/978-1-4939-9170-9_1

Prykhozhij, S. V., Fuller, C., Steele, S. L., Veinotte, C. J., Razaghi, B., Robitaille, J. M., McMaster, C. R., Shlien, A., Malkin, D., & Berman, J. N. (2018). Optimized knock-in of point mutations in zebrafish using CRISPR/Cas9. *Nucleic Acids Research*, *46*(17), 1–15. https://doi.org/10.1093/nar/gky512

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *CELL*, *152*, 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

Ramaswamy, S., & Kordower, J. H. (2012). Gene therapy for Huntington's disease. *Neurobiology of Disease*, *48*(2), 243–254. https://doi.org/10.1016/j.nbd.2011.12.030

Ramirez, C. L., Foley, J. E., Wright, D. A., Müller-Lerch, F., Rahman, S. H., Cornu, T. I., Winfrey, R. J., Sander, J. D., Fu, F., Townsend, J. A., Cathomen, T., Voytas, D. F., & Joung, J. K. (2008). Unexpected failure rates for modular assembly of engineered zinc fingers (Nature Methods (2008) vol. 5 (374-375)). *Nature Methods*, *5*(6), 575. https://doi.org/10.1038/nmeth0608-575a

Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., Zetsche, B., Shalem, O., Wu, X., Makarova, K. S., Koonin, E. V., Sharp, P. A., & Zhang, F. (2015). In vivo genome editing using Staphylococcus aureus Cas9. *Nature*, *520*(7546), 186–191. https://doi.org/10.1038/nature14299

Rath, D., Amlinger, L., Rath, A., & Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. In *Biochimie*. https://doi.org/10.1016/j.biochi.2015.03.025

Renshaw, S. A., Loynes, C. A., Trushell, D. M. I., Elworthy, S., Ingham, P. W., & Whyte, M. K. B. (2006). A transgenic zebrafish model of neutrophilic inflammation. *Blood*, *108*(13), 3976–3978. https://doi.org/10.1182/blood-2006-05-024075

Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., & Corn, J. E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature Biotechnology*, *34*(3), 339–344. https://doi.org/10.1038/nbt.3481

Rounioja, S., Saralahti, A., Rantala, L., Parikka, M., Henriques-Normark, B., Silvennoinen, O., & Rämet, M. (2012). Defense of zebrafish embryos against Streptococcus pneumoniae infection is dependent on the phagocytic activity of leukocytes. *Developmental and Comparative Immunology*, *36*(2), 342–348. https://doi.org/10.1016/j.dci.2011.05.008

Rudin, N., Sugarman, E., & Haber, J. E. (1989). Genetic and physical analysis of double-strand break repair and recombination in Saccharomyces cerevisiae. *Genetics*, *122*(3), 519–534.

Ryan, R. P., Vorhölter, F.-J., Potnis, N., Jones, J. B., Van Sluys, M.-A., Bogdanove, A. J., & Dow, J. M. (2011). Pathogenomics of Xanthomonas: understanding bacterium–plant interactions. *Nature Reviews Microbiology*, *9*(5), 344–355. https://doi.org/10.1038/nrmicro2558

Sander, J. D., Cade, L., Khayter, C., Reyon, D., Peterson, R. T., Joung, J. K., & Yeh, J. J. (2011). Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nature Biotechnology*, *29*(8), 697–698. https://doi.org/10.1038/nbt.1934

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (n.d.). *The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli*. https://doi.org/10.1093/nar/gkr606

Saralahti, A., Piippo, H., Parikka, M., Henriques-Normark, B., Rämet, M., & Rounioja, S. (2014). Adult zebrafish model for pneumococcal pathogenesis. *Developmental and Comparative Immunology*, *42*(2), 345–353. https://doi.org/10.1016/j.dci.2013.09.009

Scherer, S., & Davis, R. W. (1979). Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proceedings of the National Academy of Sciences of the*

United States of America, *76*(10), 4951–4955. https://doi.org/10.1073/pnas.76.10.4951

Schmid, B., Hruscha, A., Hogl, S., Banzhaf-Strathmann, J., Strecker, K., Van Der Zee, J., Teucke, M., Eimer, S., Hegermann, J., Kittelmann, M., Kremmer, E., Cruts, M., Solchenberger, B., Hasenkamp, L., Van Bebber, F., Van Broeckhoven, C., Edbauer, D., Lichtenthaler, S. F., & Haass, C. (2013). Loss of ALS-associated TDP-43 in zebrafish causes muscle degeneration, vascular dysfunction, and reduced motor neuron axon outgrowth. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(13), 4986–4991. https://doi.org/10.1073/pnas.1218311110

Schmidt, F., & Grimm, D. (2015). CRISPR genome engineering and viral gene delivery: A case of mutual attraction. *Biotechnology Journal*, *10*(2), 258–272. https://doi.org/10.1002/biot.201400529

Schulte-Merker, S., & Stainier, D. Y. R. (2014). Out with the old, in with the new: Reassessing morpholino knockdowns in light of genome editing technology. *Development (Cambridge)*, *141*(16), 3103–3104. https://doi.org/10.1242/dev.112003

Segal, D. J., Beerli, R. R., Blancafort, P., Dreier, B., Effertz, K., Huber, A., Koksch, B., Lund, C. V., Magnenat, L., Valente, D., & Barbas, C. F. (2003). Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. *Biochemistry*, *42*(7), 2137–2148. https://doi.org/10.1021/bi026806o

Senís, E., Fatouros, C., Große, S., Wiedtke, E., Niopek, D., Mueller, A. K., Börner, K., & Grimm, D. (2014). CRISPR/Cas9-mediated genome engineering: An adeno-associated viral (AAV) vector toolbox. *Biotechnology Journal*, *9*(11), 1402–1412. https://doi.org/10.1002/biot.201400046

Shah, A. N., Davey, C. F., Whitebirch, A. C., Miller, A. C., & Moens, C. B. (2016). Rapid Reverse Genetic Screening Using CRISPR in Zebrafish. *Zebrafish*, *13*(2), 152–153. https://doi.org/10.1089/zeb.2015.29000.sha

Simpson, R. J., Wilding, C. S., & Grahame, J. (1999). Simple protocol for extracting nuclear DNA from single embryos of a marine snail. *BioTechniques*, *26*(6), 1050–1052. https://doi.org/10.2144/99266bm08

Smith, H. O., & Welcox, K. W. (1970). A Restriction enzyme from Hemophilus influenzae. I. Purification and general properties. *Journal of Molecular Biology*, *51*(2), 379–391. https://doi.org/10.1016/0022-2836(70)90149-X

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., & Kucherlapati, R. S. (1985). Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. *Nature*, *317*(6034), 230–234. https://doi.org/10.1038/317230a0

Sprague, J. (2006). The Zebrafish Information Network: the zebrafish model organism database. *Nucleic Acids Research*, *34*(90001), D581–D585. https://doi.org/10.1093/nar/gkj086

Stainier, D. Y. R., Raz, E., Lawson, N. D., Ekker, S. C., Burdine, R. D., Eisen, J. S., Ingham, P. W., Schulte-Merker, S., Yelon, D., Weinstein, B. M., Mullins, M. C., Wilson, S. W., Ramakrishnan, L., Amacher, S. L., Neuhauss, S. C. F., Meng, A., Mochizuki, N., Panula, P., & Moens, C. B. (2017). Guidelines for morpholino use in zebrafish. *PLOS Genetics*, *13*(10), e1007000. https://doi.org/10.1371/journal.pgen.1007000

Sternberg, S. H., Richter, H., Charpentier, E., & Qimron, U. (2016). Adaptation in CRISPR-Cas Systems. *Molecular Cell*, *61*(6), 797–808. https://doi.org/10.1016/j.molcel.2016.01.030

Sullivan, C., & Kim, C. H. (2008). Zebrafish as a model for infectious disease and immune function. *Fish and Shellfish Immunology*, *25*(4), 341–350. https://doi.org/10.1016/j.fsi.2008.05.005

Summerton, J. (2007). Morpholino, siRNA, and S-DNA Compared: Impact of Structure and Mechanism of Action on Off-Target Effects and Sequence Specificity. *Current Topics in Medicinal Chemistry*, *7*(7), 651–660. https://doi.org/10.2174/156802607780487740

Summerton, J., & Weller, D. (1997). Morpholino Antisense Oligomers: Design, Preparation, and Properties. *Antisense and Nucleic Acid Drug Development*, *7*(3), 187–195. https://doi.org/10.1089/oli.1.1997.7.187

Sun, Z. (2004). A genetic screen in zebrafish identifies cilia genes as a principal cause of cystic kidney. *Development*, *131*(16), 4085–4093. https://doi.org/10.1242/dev.01240

Sung, Y. H., Kim, J. M., Kim, H. T., Lee, J., Jeon, J., Jin, Y., Choi, J. H., Ban, Y. H., Ha, S. J., Kim, C. H., Lee, H. W., & Kim, J. S. (2014). Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Research*, *24*(1), 125–131. https://doi.org/10.1101/gr.163394.113

Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D. J., & Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nature Biotechnology*, *25*(7), 786–793. https://doi.org/10.1038/nbt1317

Tampakaki, A. P., Fadouloglou, V. E., Gazi, A. D., Panopoulos, N. J., & Kokkinidis, M. (2004). Conserved features of type III secretion. *Cellular Microbiology*, *6*(9), 805–816. https://doi.org/10.1111/j.1462-5822.2004.00432.x

Tautz, D. (1992). Problems and paradigms: Redundancies, development and the flow of information. *BioEssays*, *14*(4), 263–266. https://doi.org/10.1002/bies.950140410

Terns, M. P., & Terns, R. M. (2012). CRISPR-Based Adaptive Immune Systems. *Bone*, *23*(1), 1–7. https://doi.org/10.1038/jid.2014.371

Thakore, P. I., D'Ippolito, A. M., Song, L., Safi, A., Shivakumar, N. K., Kabadi, A. M., Reddy, T. E., Crawford, G. E., & Gersbach, C. A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature Methods*, *12*(12), 1143–1149. https://doi.org/10.1038/nmeth.3630

Tobin, D. M., May, R. C., & Wheeler, R. T. (2012). Zebrafish: A see-through host and a fluorescent toolbox to probe host-pathogen interaction. *PLoS Pathogens*, *8*(1), 8–10. https://doi.org/10.1371/journal.ppat.1002349

Tobin, D. M., Vary, J. C., Ray, J. P., Walsh, G. S., Dunstan, S. J., Bang, N. D., Hagge, D. A., Khadge, S., King, M. C., Hawn, T. R., Moens, C. B., & Ramakrishnan, L. (2010). The lta4h Locus Modulates Susceptibility to Mycobacterial Infection in Zebrafish and Humans. *Cell*, *140*(5), 717–730. https://doi.org/10.1016/j.cell.2010.02.013

Torraca, V., & Mostowy, S. (2018). Zebrafish Infection: From Pathogenesis to Cell Biology. *Trends in Cell Biology*, *28*(2), 143–156. https://doi.org/10.1016/j.tcb.2017.10.002

Trede, N. S., Langenau, D. M., Traver, D., Look, A. T., & Zon, L. I. (2004). The use of zebrafish to understand immunity. *Immunity*, *20*(4), 367–379. https://doi.org/10.1016/S1074-7613(04)00084-6

Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., Goodwin, M. J., Aryee, M. J., & Joung, J. K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature Biotechnology*, *32*(6), 569–576. https://doi.org/10.1038/nbt.2908

Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, *183*, 93–102. https://doi.org/10.1016/j.biocon.2014.11.017

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics*, *11*(9), 636–646. https://doi.org/10.1038/nrg2842

van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M., & Brouns, S. J. J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends in Biochemical Sciences*, *34*(8), 401–407. https://doi.org/10.1016/j.tibs.2009.05.002

Vargas, R. A., Sarmiento, K., & Vásquez, I. C. (2015). Zebrafish (Danio rerio): A Potential Model for Toxinological Studies. *Zebrafish*, *12*(5), 320–326. https://doi.org/10.1089/zeb.2015.1102

Villamizar, N., Ribas, L., Piferrer, F., Vera, L. M., & Sánchez-Vázquez, F. J. (2012). Impact of Daily Thermocycles on Hatching Rhythms, Larval Performance and Sex Differentiation of Zebrafish. *PLoS ONE*, *7*(12), e52153. https://doi.org/10.1371/journal.pone.0052153

Villarreal, M. A., Biediger, N. M., Bonner, N. A., Miller, J. N., Zepeda, S. K., Ricard, B. J., García, D. M., & Lewis, K. A. (2017). Determining Zebrafish Epitope Reactivity to Commercially Available Antibodies. *Zebrafish*, *14*(4), 387–389. https://doi.org/10.1089/zeb.2016.1401

Vishwanath, W. (2010). *Danio rerio*. https://www.iucnredlist.org/species/166487/6219667

Watson, J. D., & Crick, F. H. (1953). The structure of DNA. *Cold Spring Harbor Symposia on Quantitative Biology*, *18*, 123–131. https://doi.org/10.1101/SQB.1953.018.01.020

Wells, D. J. (2004). Gene therapy progress and prospects: Electroporation and other physical methods. In *Gene Therapy* (Vol. 11, Issue 18, pp. 1363–1369). Nature Publishing Group. https://doi.org/10.1038/sj.gt.3302337

WHO. (2013). *WHO | Estimated Hib and pneumococcal deaths for children under 5 years of age, 2008*. World Health Organization. https://www.who.int/immunization/monitoring_surveillance/burden/estimates/Pneumo_hib/en/

Wigler, M., Silverstein, S., Lee, L. S., Pellicer, A., Cheng, Y. chi, & Axel, R. (1977). Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell*, *11*(1), 223–232. https://doi.org/10.1016/0092-8674(77)90333-6

Wilkinson, R. N., Elworthy, S., Ingham, P. W., & van Eeden, F. J. M. (2013). A method for high-throughput PCR-based genotyping of larval zebrafish tail biopsies. *BioTechniques*, *55*(6), 314–316. https://doi.org/10.2144/000114116

Wolfe, S. A., Nekludova, L., & Pabo, C. O. (2009). DNA recognition by Cys2His2 zinc finger proteins. *Biophysical Structure*, *25*(1), 22–29. https://doi.org/10.1093/bioinformatics/btn580

Woods, I. G., Kelly, P. D., Chu, F., Ngo-Hazelett, P., Yan, Y.-L., Huang, H., Postlethwait, J. H., & Talbot, W. S. (2000). A Comparative Map of the Zebrafish Genome. *Genome Research*, *10*(12), 1903–1914. https://doi.org/10.1101/gr.10.12.1903

Wu, Z., Yang, H., & Colosi, P. (2010). Effect of Genome Size on AAV Vector Packaging. *Molecular Therapy*, *18*(1), 80–86. https://doi.org/10.1038/mt.2009.255

Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X., Zhu, Z., Lin, S., & Zhang, B. (2013). Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Research*, *41*(14), 1–11. https://doi.org/10.1093/nar/gkt464

Xing, L., Quist, T. S., Stevenson, T. J., Dahlem, T. J., & Bonkowsky, J. L. (2014). Rapid and efficient zebrafish genotyping using PCR with high-resolution melt analysis. *Journal of Visualized Experiments*, *84*. https://doi.org/10.3791/51138

Yue, G. H., & Orban, L. (2001). Rapid isolation of DNA from fresh and preserved fish scales for polymerase chain reaction. *Marine Biotechnology*, *3*(3), 199–204. https://doi.org/10.1007/s10126-001-0010-9

ZFIN The Zebrafish Information Network. (n.d.). Retrieved March 22, 2020, from https://zfin.org/

Zhang, F., Abudayyeh, O. O., Gootenberg, J. S., Sciences, C., & Mathers, L. (2020). A protocol for detection of COVID-19 using CRISPR diagnostics. *Bioarchive*, 1–8.

Zilberman, N., Reikhav, S., Hulata, G., & Ron, M. (2006). High-throughput genomic DNA extraction protocol from tilapia's fin tissue. *Aquaculture*, *255*(1–4), 597–599. https://doi.org/10.1016/j.aquaculture.2005.11.042

Zon, L. I., & Peterson, R. T. (2005). In vivo drug discovery in the zebrafish. *Nature Reviews Drug Discovery*, *4*(1), 35–44. https://doi.org/10.1038/nrd1606

Zou, M., Friedrich, R. W., & Bianco, I. H. (2016). *Targeted Electroporation in Embryonic, Larval, and Adult Zebrafish* (K. Kawakami, E. E. Patton, & M. Orger (Eds.); Vol. 1451, pp. 259–269). Springer New York. https://doi.org/10.1007/978-1-4939-3771-4_17

6 ANNEXES

6.1 FIGURES

Figure 1: Genome engineering with DSB. This figure gives an overview of the different kinds of genome engineering. At first, a (targeted) DSB is inflicted in the target DNA, this can be done through targeted nucleases such as ZFN, TALENs, and CRISPR/Cas. In reaction A, the DSB is repaired with the NHEJ mechanism resulting in indels (shown in red) and subsequent gene disruption or knockout. In reaction B and reaction C, the DSB is repaired by HDR. In mechanism B, the donor DNA consists of the wild type (WT), but with the correct base sequence. This is done for genetic diseases where a mutation distorts the gene activity. In reaction C the donor DNA is a transgenic DNA, this can be used for gene knock-ins. A: NHEJ, knockout; B: HDR, Gene correction; C: HDR, Knock-in.............. 13

Figure 2: Adult female zebrafish. The blue stripes running across the sides of the zebrafish clearly show how the fish got its name. The fish in this figure is a female fish. This can be determined by looking at the abdomen, but one usually needs some practice before being able to clearly distinguish between sexes. The abdomen (located between the pectoral fin and pelvic fin) of the female fish is bigger than a males' abdomen. 15

Figure 3: Graph of search results per year. This graph shows the amount of hits on PubMed search when using the search query 'zebrafish'. While there are a modest 657 hits in the year 2000, this number rose to a staggering 3882 hits in 2019. Data retrieved from pubmed.ncbi.nlm.nih.gov (02-05-2020) 16

Figure 4: Developmental stages of Danio rerio (Kimmel et al., 1995). In this figure, the developmental stages of the zebrafish are pictured. Starting from the one-cell stage (0,2 hours post fertilization (hpf)) up until the formation of the protruding mouth (72 hpf). It is clear that in 72 hours, the zebrafish embryo develops. The rate at which the embryo develops is pictured at 27,5°C. These rates can be slowed down by using cold water of 24°C (Villamizar et al., 2012)... 17

Figure 5: Human Zinc Finger protein structure. (Q9NZL3, ZN224_HUMAN). On the left the structure from Uniprot. The blue dot in the middle (and the purple dot of the right structure) represents the zinc ion. The alpha helix can be seen on the left-hand side and the double beta sheet on the right. The right structure is from SMR. The residues in yellow depict the cysteine residues. The alpha Helix can be found at the C-terminus while the beta sheets are connected to the N-terminus. 21

Figure 6: Structure of ZFN dimer bound to target DNA. This figure clearly shows the modular mechanisms of ZFN. The FokI nuclease cleavage domain is shown in the middle, bound to ZFP. There are four ZFP on each side. Every single zinc finger has a recognition site of 3 bp, making this particular example a ZFN with a recognition site of 24 bp. (Urnov et al., 2010) 22

Figure 7: TAL-effector domain. representation of the TAL protein. In the middle section, the tandem repeat array can be found. The first repeat has been enlarged. 'R' represents an amino acid residue. The 34 amino acids residues can be seen; The residues on position 12 and 13 are the hypervariable residues (represented by 'V') that make up the RVD and define the DNA binding specificity. Closer to the C-terminus, Nuclear localization signals (NLS) can be found as well as an acidic activator domain (AD) 23

Figure 8: Morpholino oligomer backbone. One repeating unit of the Morpholino oligomer backbone. Where B represents one of the nucleobases (A,C,G,T). Each subunit contains one base bound to the morpholine ring structure. The phosphorodiamidate linkages connect the subunits. 24

Figure 9: CRISPR Locus organization. (simplified representation). Depicted here is a type II CRISPR locus organization. The cas operon genes can be found on the left, followed by the leader and on the right the CRISPR array. The blue rectangles represent the repeats, while the filled green rectangles represent the unique spacers, derived from protospacers or invading MGEs. 25

Figure 10: Processing of CRISPR array to crRNA. Expression stage of CRISPR immunity response. Figure A shows the CRISPR locus. The CRISPR array with its spacers and repeats is transcribed. Figure B shows the pre-crRNA, with hairpin loops. The acquired spacers are shown in green. The CRISPR transcript is then processed by Cas proteins and cut, resulting in crRNA. crRNA can be seen in figure C. The crRNA from figure C is later combined with the Cas9 protein to form the RNP. 27

Figure 11: Comparison between the interference stage of the CRISPR-Cas9 systems found in nature (left) and the CRISPR-Cas9 system used in the laboratories, using a single chimeric RNA (sgRNA) (right). Figure adapted from Jinek et al., 2012. 28

Figure 12: Representation of microinjection in single cell stage of zebrafish embryos. The single cell stage is shown at 0.2 hpf, the 2-cell stage is shown at 0.75 hpf. A is the yolk; B is the single cell; C is the needle used for injection (usually borosilicate), and D is the chorion. 33