Norwegian University
of Life Sciences

Master's Thesis 2021 60 ECTS
Faculty of Biosciences
Department of Animal and Aquacultural Sciences

Development of a method to compare sgRNA expression driven by heterologous and endogenous U6 promoters in an Atlantic salmon head kidney cell line

Axmee Regmi
Master of Science in Aquaculture

Acknowledgement

I would like to show my sincere gratitude to my supervisor, Dr. Guro Sandvik for being such an amazing supervisor and shaping path of my thesis and guiding me throughout the period. This study would not be possible without your valuable time and suggestions. Thank you for all the gatherings and dinner at Oslo. I am so glad that I got this opportunity to work under such a positive, supportive, and inspiring supervisor. Likewise, I am extremely thankful to my co-supervisor PhD fellow Noman Reza for providing necessary guidance and supporting me in every steps.

I would also like to thank CIGENE Lab and whole lab team for providing such a positive and supportive work environment. They were always there to help and provide necessary help. Likewise, I am grateful to my friend Thomas Nesset for his wonderful company throughout the thesis. I had great time talking with you, sharing each other's problems, and supporting each other during ups and downs. Thank you for giving me such a memorable time.

I am heartly thankful to my parents and family members, especially my loving husband, Arpan, for being there at the time of high and low and loving me unconditionally forever. A big thank you to my lovely son, Aspan, who was always there to disturb and distract me from my work. He taught me how to efficiently manage time and made me realize what to prioritize in life and what not.

Axmee Regmi

31 July 2021

Abstract

Infectious diseases have negative impacts on fish welfare and sustainable aquaculture. Knowledge on host pathogen relationship and genes responsible for disease resistance could lead to development of vaccines and therapeutics and other potential commercial applications. The application of CRISPR/Cas9 methods like genome-scale CRISPR knock-out screening (GeCKO screening) has great potential in identification of functional genes for a particular trait. However, application of CRISPR in fish cell lines is still in its infancy and has several limitations. In CRISPR screens, expression of sgRNA, one of the important components of the CRISPR/Cas9 method, is driven by a U6 promoter situated in a lentiviral vector. For the success of the CRISPR screen, it is crucial that the U6 promoter is efficient in the cell line used. This has never been tested in any Atlantic salmon cell line before. The aim of this thesis was to test the efficiency of different U6 promoters – human, mouse, zebrafish, and novel uncharacterised salmon U6 promoter in the Atlantic salmon cell line, SHK-1 (Salmon head kidney-1). This was done with a sgRNA expression assay. In short, 4 different U6 promoters and GFP knockout gRNA were cloned into a single lentiviral vector system. The plasmids generated were then transfected into SHK-1 cells with electroporation. After two days, total RNA from the cells was isolated and used for cDNA synthesis. Finally, the transcription of sgRNA under each promoter was compared using qPCR. Since this has never been done before, and the methods used was new in our lab, this thesis involved a lot of method development and optimization. Due to technical difficulties and Covid-19 restrictions, I had only time to perform one single biological replicate in the SHK-1 cell line with the partially optimized protocol. I was able to demonstrate that one of the plasmids that we designed was able to express sgRNA under zebrafish U6 promoter. Although the expression level was low, it was a huge success for us as it showed overall success of plasmid design, transfection of cells, primer validation, and qPCR assay optimization. One of the major findings in this thesis is that it is crucial to digest DNA in the sample before qPCR, since we use a large amount of DNA for transfection. This thesis describes and discuss the development and optimization of a method for determining U6 promoter efficiency using a sgRNA expression assay with qPCR, from construction of vectors to optimization and analysis of the qPCR experiment. The methods optimized in this thesis can be used to assess the efficiency of multiple U6 promoters in different cell lines in the future.

Abbreviations

ASK cell line - Atlantic salmon kidney cell line

bp- base pair

cDNA- complementary DNA

CIGENE – Centre for integrative genetics

CRISPR/Cas9- Clustered regularly interspaced short palindromic repeats/ associated protein 9

crRNA- CRISPR RNA

Ct- cycle threshold

dnd gene - *dead end* gene

DSB- double strand break

DSE - distal sequence element

EGFP - Enhanced Green Fluorescence Protein

elov2 gene - ELOVL fatty acid elongase 2 gene

FBS – fetal bovine serum

GeCKO- Genome scale CRISPR knockout

gfp_ko_gRNA - GFP knockout gRNA

GFP- green fluorescent protein

mRNA- messenger RNA

NHEJ- non homologous end joining

OCT - octamer

OD- optical density

PAM- protospacer adjacent motif

Pol III - RNA polymerase type III

PSE - proximal sequence element

RINe - RNA integrity number equivalent

RNP - Ribonucleoprotein

rRNA- ribosomal RNA

RT-qPCR - reverse transcription-quantitative polymerase chain reaction,

sgRNA- single guide RNA

SHK-1 cell line - Salmon Head Kidney-1 cell line

slc45a2 gene - solute carrier family 45, member 2 gene

SNAPc - snRNA activating protein complex

snRNA- small nuclear RNA

stat2 gene - Signal transducer and activator of transcription 2 gene

TFs - transcription factors

T_m - melting temperature

tracrRNA- transactivating RNA

tRNA- transfer RNA

TSS - transcription start site

Tyr gene- tyrosinase gene

Table of Contents

1 Introduction

1.1 Background

Global demand for salmonid fish is growing every year. To meet the demand, production of farmed salmonid fish has also increased (FAO, 2018). There have been many technical advancements in aquaculture industry. However, infectious diseases still have negative impacts on fish welfare and sustainability as approximately 40% of the total potential production is lost per year (Jobling, 2011). This is one of the factors that is limiting future expansion of the aquaculture industry. One possible solution to this problem is development of vaccines and therapeutics. However, host response to the pathogens, functional genes and their variants that are responsible for disease resistance in host is largely unknown. This knowledge could potentially be used to create a healthier salmon breeding population through potential commercial applications like selective breeding (Houston, 2017; Yáñez et al., 2014), genomic selection (Houston et al., 2020; Zenger et al., 2019) and genome editing (Gratacap et al., 2019).

CRISPR/Cas9 system is a powerful, targeted genome editing tool and due to its simplicity and efficiency, it has become very popular over a short period. The application of CRISPR/Cas9 has great potential in the identification of function of a particular gene and its variant in a particular trait, for example, genes responsible for disease resistance (Staller et al., 2019). CRISPR/Cas9 has been successfully applied *in vivo* and in cell lines of various aquaculture species like rohu, grass carp, common carp, catfish, Pacific oyster, sea bream, Nile tilapia and salmonids like rainbow trout and Atlantic salmon (Gratacap et al., 2019). The methods that are currently established for *in vivo* CRISPR/Cas9 application in fish use microinjection of the CRISPR/Cas9 complex into fresh fertilized eggs in its one cell stage. Microinjection is suitable for larger cell size like embryos but not practical for high throughput application in cell lines. The establishment of CRISPR/Cas9 in cell lines require more practical high throughput delivery systems like viral transduction or transfection of plasmid. In addition, the success relies on expression of single guide RNA (sgRNA) and Cas9 which requires identification of optimal promoters. However, the methods that are successfully established for CRISPR/Cas9 in other cell lines, might not be directly applicable in fish cell lines and have some limitations that should be addressed.

The goal of our team is to establish genome-scale CRISPR knock-out (GeCKO) screening for host pathogen relationship between Atlantic salmon and infectious salmon anaemia (ISA) virus. Generally, in GeCKO screening, lentiviral transduction is used to deliver CRISPR/Cas9 components using lentiviral vector into the cells. For successful application of GeCKO screening in fish cell lines, it is important to determine optimal promoter for expression of sgRNA and Cas9. U6 RNA polymerase type III promoter for the transcription of spliceosomal U6 small nuclear RNA (snRNA) have been widely used for expression of sgRNA in cell lines. Efficient U6 promoter for zebrafish (Clarke et al., 2013) and tilapia (Hamar & Kültz, 2021) cell lines has been identified. Human and zebrafish U6 promoters has been used previously to drive the expression of sgRNA in chinook salmon cell lines (Escobar-Aguirre et al., 2019; Gratacap et al., 2020). However, efficiency of different U6 promotors for the sgRNA expression has not been tested in any salmonoid cell lines so far. The aim of this thesis is to find the efficiency of commonly used heterologous promoters (promoters that are not naturally found in salmon) – Human U6, Mouse U6 and Zebrafish U6 together with one previously uncharacterized endogenous promoter – Salmon U6 promoter in Atlantic salmon cell line using sgRNA expression assay.

1.2 CRISPR/Cas9 system, its components, mechanism, and potential use

CRISPR/Cas9 stands for Clustered regularly interspaced short palindromic repeats/ associated protein 9. The discovery of CRISPR/Cas9 has revolutionized the genome editing technology (Adli, 2018; Zhang, 2019). CRISPR/Cas9 system is adapted from the Streptococcus pyogenes bacteria's defence system (Deltcheva et al., 2011). When the bacterium is exposed to foreign genetic elements, short fragments of DNA get integrated into bacteria's CRISPR repeat-spacer array, known as protospacer sequence as a genetic record. This will defend bacteria upon invasion by same phage in future (Barrangou et al., 2007; Mojica et al., 2005). The small CRISPR RNAs (crRNA) are transcribed from the protospacer sequences (Brouns et al., 2008), which will eventually guide Cas9 endonuclease to cut viral DNA and block horizontal DNA transfer of virus (Deltcheva et al., 2011; Gasiunas et al., 2012; Jinek et al., 2012; Marraffini & Sontheimer, 2008). Gene modification through CRISPR/Cas9 requires 3 main components: Cas9 protein, sgRNA and a protospacer adjacent motif (PAM) site (Doudna & Charpentier, 2014; Jiang & Doudna, 2017; Makarova et al., 2011; Zhang, 2019).

1.2.1 Cas9 protein

Cas9 protein derived from *S. pyogenes* is a large and multifunctional DNA endonuclease that consists of 1368 amino acid. Cas9 consists of two endonuclease domains: the HNH-like endonuclease that cut the target strand and the n-terminal RuvC-like nuclease domain that cut the complementary strand of the target strand, and hence create double strand break (DSB) (Chen et al., 2014; Gasiunas et al., 2012; Jinek et al., 2012). To repair the DSB, cell activates DNA repair system, non-homologous end-joining (NHEJ), which is error prone and frequently results in disruption of the reading frame via deletions or insertions. This leads to expression of truncated, non-functional protein (Lieber, 2010; Wyman & Kanaar, 2006).

New methods for using CRISPR is developed in a great pace. For example, Cas9 protein can be catalytically inactivated, either one of its nuclease domain or both domains to create nickase Cas9 or dead Cas9 respectively, which means these Cas9 can specifically bind to the target site but cannot create DSB or is able to create a nick (Qi et al., 2013; Sapranauskas et al., 2011). Such impaired Cas9 can be used to guide effectors like gene activators or repressors (Bikard et al., 2013; Gilbert et al., 2014; Gilbert et al., 2013), base editors (Kim et al., 2017; Li et al., 2018; Nishida et al., 2016) which can modulate and modify gene and gene expression, or even visualize DNA or RNA (Chen et al., 2013; Konermann et al., 2013; Tanenbaum et al., 2014), and many more. This opens several applications of CRISPR/Cas9 and leads to next generation genome editing tools.

1.2.2 sgRNA

Cas9 is guided by short, ~20 bp long RNA sequence that is specific and complementary to target sequence (Brouns et al., 2008) followed by 5'-PAM sequence (Mojica et al., 2009). In nature, the guide RNA is comprised of two RNA pieces: the crRNA (described earlier) and the trans-activating crRNA (tracrRNA) which is 75-100 bp long and responsible for forming a scaffold that links crRNA and Cas9 and also helps in maturation of pre-crRNAs (derived from CRISPR array) to crRNA (Brouns et al., 2008; Deltcheva et al., 2011; Gasiunas et al., 2012). In 2012, Jinek et al. (2012) combined these two RNAs into one single RNA chimera, which was equally capable to direct Cas9 to target site. The RNA chimera is now known as single guide RNA (sgRNA) or guide RNA (gRNA). This simplified the use of CRISPR/Cas9 tool. A

typical sgRNA cloning vector that are generally used for CRISPR/Cas9 applications consist of 75-100 bp gRNA scaffold, restriction sites to clone ~20 bp sequence specific gRNA upstream to the scaffold, promoter to drive the expression of sgRNA and antibiotic selection marker.

1.2.3 PAM site

PAM is a 2-5 bp short, conserved sequence motif, located downstream (3') to the crRNA targeting sequence in template DNA strand and acts as binding signal for Cas9 (Mojica et al., 2009). PAM is crucial for the function of CRISPR/Cas9 system and in absence of PAM sequence Cas9 mediated DNA cleavage does not occur. PAM sequence varies between the different Cas9 variants and the PAM sequence for the most used *S. pyogenes* Cas9 is 5'NGG'3 (Adli, 2018; Doudna & Charpentier, 2014; Jiang & Doudna, 2017; Makarova et al., 2011; Mojica et al., 2009).

1.2.4 Application of CRISPR/Cas9 in aquaculture

CRISPR/Cas9 has introduced a radical change in genome editing field. Successful *in vivo* execution of CRISPR/Cas9 has been accomplished in embryo of several aquaculture species, like carp (Chakrapani et al., 2016), tilapia (Li et al., 2014), catfish (Elaswad et al., 2018) and salmon (Edvardsen et al., 2014) to name few. Most of these studies have followed established protocol in model organisms like zebrafish (Jao et al., 2013) and focus on proof-of-principle by targeting genes that result in clearly observable phenotype such as pigmentation. The *in vivo* application utilizes microinjection mode of delivery to deliver CRISPR/Cas9 components into one cell stage of the fish embryo. Recently, medaka (Liu et al., 2018), carp (Ma et al., 2018) and salmonid cell lines (Dehler et al., 2016; Escobar-Aguirre et al., 2019; Gratacap et al., 2020a; Gratacap et al., 2020) has been successfully edited with CRISPR/Cas9. Genome editing with CRISPR/Cas9 has been applied to target traits like sterility (Qin et al., 2016; Wargelius et al., 2016), growth (Cleveland et al., 2018; Kishimoto et al., 2018) and disease resistance and immunology (Chakrapani et al., 2016; Ma et al., 2018) so far.

CRISPR/Cas9 system was first applied in Atlantic salmon in 2014 (Edvardsen et al., 2014). Two genes that were responsible for pigmentation – *tyrosinase* (*tyr*) and *solute carrier family 45, member 2* (*slc45a2*) was successfully knocked out *in vivo.* In 2016, the same group of

researchers created sterile salmon by knocking out *dead end* (*dnd*) gene (Wargelius et al., 2016). Similarly in 2019, Datsomor et al. (2019) applied CRISPR/Cas9 mediated *in vivo* knockout of *ELOVL fatty acid elongase 2* (*elov2*) gene. Successful CRISPR/Cas9 editing in SHK-1 and ASK cell line via transfection of ribonucleoprotein complex (discussed later in section 1.5) has been recorded (Gratacap et al., 2020a). In addition, CRISPR/Cas9 has been successfully applied in embryonic chinook salmon cell line that stably expresses EGFP and Cas9 (Dehler et al., 2016) and CRISPR/Cas9 has been successfully applied to produce *Signal transducer and activator of transcription 2* (*stat2*) gene knock out cell line embryonic chinook salmon cell line that was resistant to viral infection (Dehler et al., 2019). Such initial success of CRISPR/Cas9 mediated genome editing both *in vivo* and in cell lines, opens new opportunities to improve production and sustainability via genetic improvement of disease resistance and other traits in Atlantic salmon as well as in other aquaculture species.

1.2.5 Genome-scale CRISPR knockout (GeCKO) screening

Recently GeCKO screening in cell cultures has been one of the promising applications of CRISPR that can be used to solve different biological questions such as identification of genes responsible for drug sensitivity or resistant, environmental toxin susceptibility, components of cellular pathways and genes responsible for diseases like genes involved in tumour growth and metastasis (Doench, 2018). The basic concept behind the CRISPR knockout screening is to knockout either every single gene present in the genome or a selection of genes of interest (illustrated in Figure 1). This involves designing of thousands of gRNAs in a library, targeting every gene in genome, synthesising these gRNA, packing them into lentiviral vector and transducing cells with virus such that single gene is knocked out per cell. Then the cells are screened by either letting them grow in normal conditions, which results in identification of genes that are important for cell growth/survival, or by challenging cells with stress or perturbation for example, drug treatment, viral infection, or toxin treatment. Over the time, this will result in limited cell population that are able to survive the conditions or even show better prevalence. Thus, the cells that were able to survive are sequenced to identify which genes' knockout is leading to cell survival and which genes' knockout led to cell death in the given condition (Adli, 2018; Doench, 2018; Zhang, 2019).

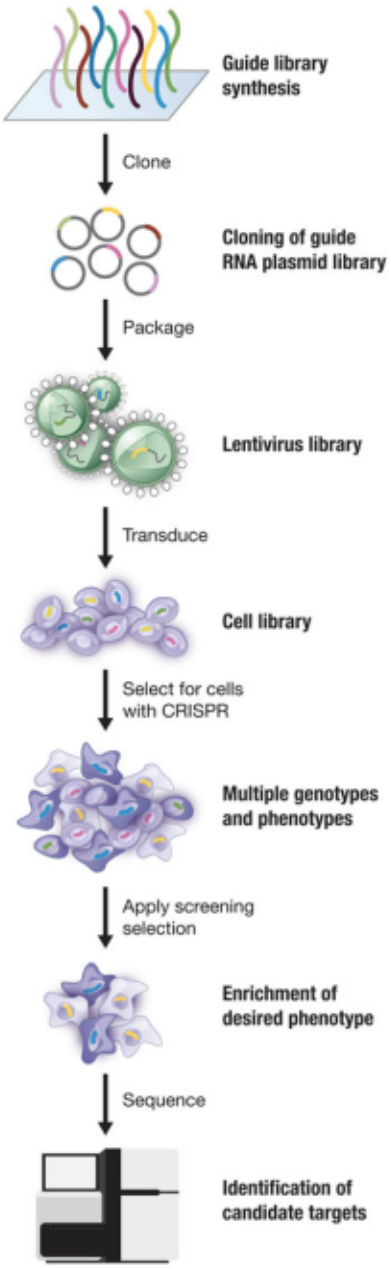

Figure 1: Illustration of steps involved in typical CRISPR screening experiment. First, gRNA library targeting different genes is designed, synthesized cloned in vector, and packed into lentivirus. The virus is then transduced into desired cells. Cells successfully transduced are selected and further selected in specific condition leading to enrichment of cells with desired phenotype. Sequencing of these cells lead to identification of genes involved in the phenotype. Figure taken from Zhang (2019)

Generally, retroviruses, mostly lentiviruses are used to deliver CRISPR/Cas9 components into the cells in CRISPR screens (Doench, 2018; Joung et al., 2017; Shalem et al., 2015). The reason behind this is that they efficiently integrate their genetic material into mammalian cell's genome regardless if it is dividing or non-dividing cell and results in long-lasting expression of sgRNA with low risk of insertional mutagenesis (Merten et al., 2016).

Delivery of Cas9 and sgRNA using lentiviral vector into the cells can be done in two ways, first using single lentiviral vector system (lentiCRISPRv2) to deliver both Cas9 and sgRNA. lentiCRISPRv2 backbone contains humanU6 promoter, gRNA scaffold, elongation factor 1a (EF1a) promoter to drive Cas9 expression, Cas9 gene and Puromycin antibiotic selection marker. Second is two vector system which uses two different vectors to deliver Cas9 (lentiCas9-Blast) and sgRNA (lentiGuide-Puro). Second method results in 10-fold higher functional viral titre over the first. However, both are equally efficient (Sanjana et al., 2014). The advantage of using two vector system is that lentiCas9-Blast can be used to generate cell lines that stably express Cas9 which can be used further for screening using lentiGuide-Puro. This method is preferred over single lentiviral vector system because of higher functional viral titre (Doench et al., 2014; Sanjana et al., 2014).

CRISPR screening led to the identification of Norovirus receptor (Orchard et al., 2016) and host

cell dependencies in dengue and Zika virus infection (Savidis et al., 2016) in human cells. Apart from human, CRISPR screening has also been applied in some model animals (Shrock & Güell, 2017) and economically important animals like cow (Tan et al., 2020) and pig (Xu et al., 2020). As mentioned earlier, infectious diseases are major barrier for sustainable aquaculture and successful application of CRISPR screening in aquacultural fish cell lines could lead to identification of genes that are responsible for disease resistance. This knowledge could be used in potential commercial application for production of disease resistance fish. However, CRISPR screening is still in its infancy in fish species. One of the major limitations that needs should be addressed for successful application of CRISPR screening in fish cell lines is lentiviral transduction as the virus is generally hosted in mammalian species and does not infect fish species. Besides, various other important aspects of CRISPR screens like genomic integration of large inserts like Cas9 gene and identification of optimal promoter to drive the expression of sgRNA and Cas9 in different fish cell lines need to be optimized.

Chinook salmon cell line is the only cell line that has been successfully transduced with lentiviral vectors (Gratacap et al., 2020), showing feasibility of lentiviral mediated GeCKO screening in salmonid cell lines. Another milestone for GeCKO screening was reached when Dehler et al. (2016) generated chinook salmon cell line that stably expresses Cas9. However, successful lentiviral transduction has not been recorded for Atlantic salmon cell lines. Similarly, promoter efficiency for expression of Cas9 and sgRNA has not been tested in any salmonid cell lines so far.

1.3 U6 RNA polymerase III promoter

A promoter is a DNA sequence upstream (5' to the sense strand) to a functional gene where RNA polymerase and transcription factors (TFs) bind to initiate transcription. Generally, promoters are around 100 to 1000 bp long and consists of response elements that provide stable binding site and control and regulate the transcription of gene (Feng et al., 2018). Transcripts can be either long protein encoding mRNA or short functional RNAs like tRNA, rRNA, U6 snRNA, H1 snRNA etc (Khatter et al., 2017; Willis, 1994). Generally, RNA polymerase type III (Pol III) is responsible for transcription of short RNA sequences (Arimbasseri & Maraia, 2016; Khatter et al., 2017; Willis, 1994). Pol III promotors like H1 and U6 are commonly used for expression of interference RNA and sgRNA in different biotechnological applications like

RNAi (Miyagishi & Taira, 2002; Shukla et al., 2007; Zenke & Kim, 2008) and CRISPR/Cas9 (Ranganathan et al., 2014; Schuster et al., 2016; Zheng et al., 2018). For CRISPR screens, U6 promoters are used to drive the expression of the sgRNA in the cells using lentiviral vector (Sanjana et al., 2014).

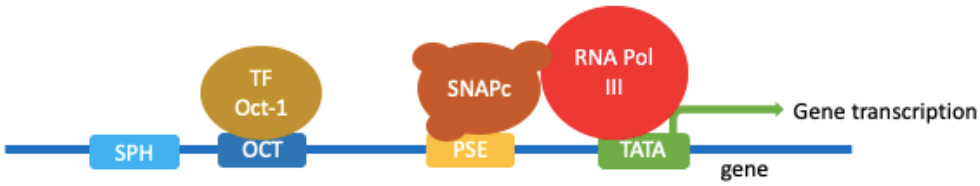

Figure 2: Illustration of U6 promoter showing different regions and binding sites for RNA pol III, different gene activators and transcription factors.

Generally, U6 Pol III promoter are around 200-300bp long and provide binding site for Pol III transcription of spliceosomal U6 snRNA, which is used for RNA splicing in eukaryotes (Arimbasseri & Maraia, 2016; Huang & Maraia, 2001; Willis, 1994). U6 Pol III promoter consists of three major regions namely, TATA box, proximal sequence element (PSE) and distal sequence element (DSE) (illustrated in Figure 2). The TATA box 5'-TATAAA-3' is situated upstream nearest to the start codon, transcription start site (TSS) which consists of RNA polymerase binding site. PSE lies upstream TATA box and consists of snRNA activating protein complex (SNAPc) binding site (Didychuk et al., 2018; Gao et al., 2018; Helbo et al., 2017; Schramm & Hernandez, 2002; White, 2011; Willis, 1994). Likewise, the distal promoter is located further upstream of PSE and consist of regulatory sequences like octamer (OCT) where TF Oct-1 binds and SPH element (Didychuk et al., 2018; Kunkel et al., 1996; Kunkel & Hixson, 1998).

The efficiency of different U6 promoters differs in organism or cell type being used. Clarke et al. (2013) showed that zebrafish U6 promoters were more effective in Zebrafish ZF4 cell line compared to mammalian cells, whereas mouse and chicken U6 promoters were more effective in mammalian cells rather than ZF4 cells. Similarly, Hamar and Kültz (2021) claimed insufficient mutation when most commonly used human and zebrafish U6 promoters were used in Tilapia brain cell line compared to endogenous U6 promotor.

However, in contrast, a study has reported that Human U6 promoter has been effective in transcription of sgRNA in the CHSE cell line (Gratacap et al., 2020). Roelz et al. (2010) showed that human U6 promoter was four times more efficient in knockdown when compared with murine U6 promoter in both murine and human cells. Likewise, Zheng et al. (2018) compared two heterologous (human and yeast) U6 promoter and one endogenous U6 promoter for the expression of sgRNA in Aspergillus Niger. All three U6 were efficient in producing sgRNA, however endogenous U6 was the most efficient one among them. These studies show the importance of testing U6 promoter efficiency before establishment of CRISPR screen in a new species like Atlantic salmon.

1.4 Fish cell lines

Primary cell culture is the culture of cells that is isolated directly from the tissue, whereas cell lines are derived primary cell culture that are immortalized to proliferate indefinitely. Generally, for research purpose, cell lines from different tissues of an organism are used as biological alternative instead of using primary cell cultures or whole animals. One of the major benefits of using cell lines is that they produce highly reproducible results compared to primary cell culture (Wolf & Quimby, 1976). Beside this, cell lines are cheap, readily available, easy to propagate, and can be sub-cultured several times. Such subcultures can be easily stored and transferred between labs. Cell lines are beneficial for studying organ or tissue- specific viral responses, close observation of viral diseases, detection and isolation of viruses, vaccine development, cell signalling pathways and so on (Genzel, 2015). Establishment of cell lines will provide a more realistic possibility for various research applications including CRISPR screening.

However, lack of suitable, well-characterized and well-tested cell line is one of the major limitations in aquacultural research. Due to this, primary cell cultures are often used. First permanent fish cell line was established in 1962 by Wolf and Quimby (1962) from gonads of rainbow trout, RTG-2. Since then, many fish cell lines derived from different fish tissues like skin, gill, heart, liver, kidney, spleen, swim bladder, brain, embryo, fin, ovary etc have been established. Until 2011, around 283 fish cell lines have been established (Lakra et al., 2011), where 3 were from Atlantic salmon, all from head kidney tissue – SHK-1 (Dannevig et al., 1995), ASK (Devold et al., 2000) and TO (Wergeland & Jakobsen, 2001).

1.4.1 SHK-1 cell line

SHK-1 cell line stands for Salmon Head Kidney-1 cell line (ECACC; 97111106) was established in 1995 (Dannevig et al., 1995). It was established from head kidney tissue of Atlantic Salmon. The cell line was established for the purpose of isolation and detection of casual virus of infectious salmon anaemia disease, which is a contagious disease in farmed Atlantic Salmon. The cells are flat and elongated in morphology and appear like fibroblast. SHK-1 cell line is of leucocytic origin having some macrophage like properties (Dannevig et al., 1997). Thus, SHK-1 cell line has been very effective tool for the study of cellular immunity mechanisms (Koppang et al., 1999; Lee et al., 2014; Lee et al., 2013; Lee et al., 2015) and viral detection and susceptibility of different salmon virus like ISA virus (Molloy et al., 2013; Opitz et al., 1999), infectious pancreatic necrosis virus (Reyes-Cerpa et al., 2012), and salmonid alphavirus-1 (Herath et al., 2009). In addition, SHK-1 cell line has been used to study host response against bacterial pathogen *Piscirickettsia salmonis* (Díaz et al., 2021; Oliver et al., 2015) and polyphenols (Santana et al., 2021).

1.5 Delivery methods for CRISPR/Cas9 system to cells in a culture

As mentioned earlier, CRISPR/Cas9 complex consist of Cas9 protein and sgRNA. This complex needs to be delivered into nucleus of the cell for gene editing to occur (Doudna & Charpentier, 2014; Gasiunas et al., 2012; Jinek et al., 2012). CRISPR/Cas9 complex can be delivered in three forms to a cell – first is, plasmid DNAs encoding both Cas9 protein and sgRNA, second is, mRNA that translates Cas9 protein together with sgRNA and third is, ribonucleoprotein (RNP) complex that consists of Cas9 protein along with sgRNA (Lino et al., 2018; Liu et al., 2020). The plasmid DNA can be delivered into the cell in two different ways, first is to pack plasmid into a virus and then infect the cells with the virus, known as viral transduction. As described above, viral transduction is generally used in CRISPR screening experiments. Second is use of transfection methods like physical transfection and chemical transfection (Lino et al., 2018). One of the commonly used physical delivery methods is electroporation. In this study, we have mainly used electroporation to deliver the lentiviral vector into the cells which is discussed in more details below.

Electroporation is used for delivering components into the cells and generally works well irrespective of cell type. In electroporation, pulsed high-voltage electrical current is applied to cells suspended in buffer which temporarily opens nanopores in cell membrane. This allows flow of intra cellular and extracellular components in and out of the cell, facilitating the entry of nucleic acids, proteins, drugs etc. Once the electric current is turned off, nanopores close and cells go back to their normal state (Chen et al., 2006; Fajrial et al., 2020; Lino et al., 2018; Liu et al., 2020).

Electroporation has been used in various salmonid cell lines in order to produce genetically engineered cell lines for example ASK cell line with an inducible gene expression system was produced in 2011 (Collet & Lester, 2011). In 2016, Chinook salmon cell line was developed that was able to stably express EGFP and Cas9, which can serve as valuable tool for functional genomics study via CRISPR screening (Dehler et al., 2016). Transfection via electroporation has been reported as an effective tool to deliver RNP complex in difficult to transduce Atlantic salmon – SHK-1 and ASK and rainbow trout RTG-2 cell lines (Gratacap et al., 2020) and RtgutGC cell line (Zoppo et al., 2021).

1.6 Quantitative polymerase chain reaction (qPCR)

In this study, the amount of sgRNA produced under each promoter in the cells was quantified using qPCR. In this method, mRNA is first converted to complementary DNA (cDNA) using reverse transcriptase enzyme and then followed by real time quantitative PCR. qPCR is most used method for the detection and quantification of gene expression. It is considered as one of the most powerful techniques in molecular biology as it has high sensitivity and specificity (Bustin et al., 2005; Taylor et al., 2010; Taylor et al., 2019).

1.6.1 Reverse transcription

After purification of mRNA from the sample, cDNA is synthesised from the mRNA template using reverse transcriptase enzyme in the reverse transcription (RT) reaction (Bustin et al., 2005; Simpson & Brown, 1995; Stahlberg et al., 2004). This step has been shown to contribute for most observed problem in qPCR experiment like variability and lack of reproducibility (Bustin et al., 2005; Taylor et al., 2019). First possible reason behind this is RT reaction

efficiency is dependent on amount of template present in the sample where significantly low efficiency is observed when the template amount is considerably low as it is negatively affected by background non-specific nucleic acids (Curry et al., 2002; Karrer et al., 1995; Stahlberg et al., 2004a).

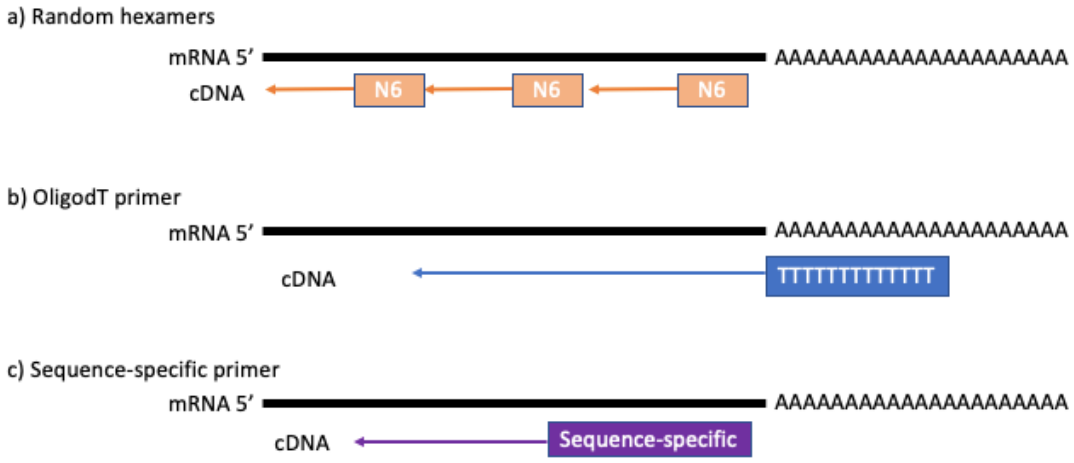

Figure 3: Illustration of three types of priming approaches for cDNA synthesis. a) Random hexamer binds non-specifically to multiple sites of RNA, b) Oligo-dT primer binds specifically to the polyA tail of mRNA, and c) sequence-specific primer binds specifically to the mRNA.

Second reason is priming approaches used for cDNA synthesis. Random hexamers, oligo-dT, target gene-specific primers or a combination of random hexamers and oligo-dT are commonly used priming approaches (Deprez et al., 2002). The description of how these primers work is illustrated in Figure 3. The drawback of using random hexamers is that it can overestimate the original mRNA copy numbers as it binds at multiple origins along the mRNA template and consequently produces more than one cDNA per original mRNA (Bustin & Nolan, 2004; Zhang & Byrne, 1999). Oligo-dT provides more specific priming compared to random hexamers. However, if the length of mRNA is very long and has RNA secondary structure, it may fail to reach extreme 5'-end. This can be a drawback if upstream primer-binding site lies extreme 5'-end (Bustin et al., 2005; Bustin & Nolan, 2004). Target specific primers produces most sensitive and specific cDNA. It requires separate priming reaction per target gene which can be drawback if the amount of sample RNA is limited (Deprez et al., 2002). It is also possible to multiplex the RT reaction in same tube, however in order to produce good quantitative data, experimental design and reaction conditions optimisation should be done

very carefully (Wittwer et al., 2001). Thus, each priming strategy has its own benefit and drawback, however to produce comparable qPCR data, it is important that same priming strategy and reaction conditions is applied (Stahlberg et al., 2004).

1.6.2 qPCR assay

A typical PCR run has three phases, an exponential phase, a linear phase, and a plateau phase (illustrated in Figure 4). In the exponential phase, the amount of PCR product increases exponentially due to high abundance of reagents. As reagents starts to be used up, a linear increase of PCR product is seen in the linear phase. Eventually, some of the reagents get depleted and there is no increase in PCR product. This is where the PCR run hits the plateau phase (Ginzinger, 2002; Yuan et al., 2006). Each reaction reaches plateau at different points because PCR reaction kinetics for each PCR reaction is different. This is why quantification of PCR product at end-point is not reliable as this does not resemble the initial quantity of the template (Ginzinger, 2002; Polz & Cavanaugh, 1998). However, during the exponential phase, under optimal PCR conditions like primer characteristics, template purity, and amplicon length, a high PCR reaction efficiency can be achieved, meaning PCR product will almost double after each cycle. Thus, by comparing the PCR cycle number by which the different reactions reach a certain level in this phase, it is possible to extrapolate back in order to find out the initial template quantity (Gibson et al., 1996; Heid et al., 1996).

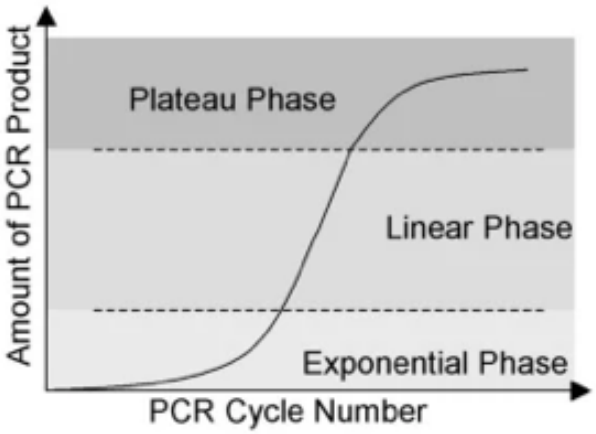

Figure 4: A typical PCR run plot obtained by plotting amount of PCR product in y-axis and cycle number in x-axis, showing 3 phases of PCR run namely exponential, linear and plateau phase. Figure from Yuan et al. (2006)

In real-time quantitative PCR, it is possible to measure the amount of PCR product generated after each cycle throughout the process. Typically, dynamics of a PCR run is visualized by using either DNA-binding dyes like SYBR green or DNA hybridization probes like molecular beacons (Strategene) and TaqMan probes (Applied Biosystems) (Bustin, 2000). The fluorescence signal is proportional to the number of amplicons in the reaction. In a typical amplification plot (Figure 5), the fluorescence signal is plotted against cycle number. For further analysis, a baseline and signal threshold are set up. During initial cycles of exponential phase, up to cycle 20 in the Figure 5, fluorescence signal is at background level which is not detected by the machine. This phase is known as baseline. After this, as amount of amplicon continues to increase exponentially, the signal coming from samples becomes significantly detectable. At this point a fluorescence signal threshold line is set up, where fluorescence signal from all samples can be compared. The cycle number at which the amplicon amount hits the threshold line is known as cycle threshold (Ct). Ct value is the observed value in a real-time PCR experiment, as it is directly proportional to the initial template amount. Thus, Ct value provides the base for quantification of DNA or RNA in any given sample (Bustin, 2000; Heid et al., 1996; Yuan et al., 2006).

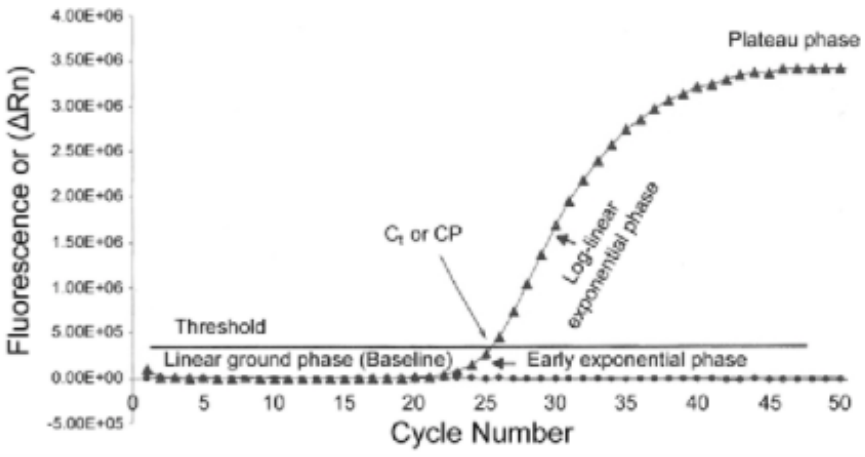

Figure 5: A typical real-time qPCR amplification plot, obtained by plotting fluorescence signal vs PCR cycle number, illustrating baseline, threshold and Ct. Figure from Ahmed (2005)

1.6.2.1 Absolute quantification

Real-time PCR data can be quantified in two ways, absolute and relative quantification. Absolute quantification also known as standard-curve quantification requires an external calibration curve of 5-fold or 10-fold serial dilution of known copy number to quantify unknown template copy number. Thus, absolute quantification helps to determine the exact copy number of the template in the sample (Boeuf et al., 2005; Bustin, 2000; Sivaganesan et al., 2010). The standard curve can be generated by plotting Ct value against known concentration of DNA or RNA molecules. Standards are generally generated from *in vitro* synthesized RNA (Boeuf et al., 2005), PCR fragment (Leong et al., 2007) or plasmid (Dhanasekaran et al., 2010).

For gene expression analysis, RNA molecules of known copy number are used as standards and are considered more accurate compared to DNA molecules as DNA standard curve assumes RT efficiency is constant and 100%. Whereas amount of cDNA produced depends on RT reaction efficiency and sequence and structure of RNA molecule. Use of RNA standard curve will even out any differences created during cDNA synthesis (Boeuf et al., 2005). The RNA standards can be generated by cloning part or whole transcript of interest into a standard cloning vector with RNA polymerase T7 or T3 or SP6 promoter. From the *in vitro* transcription, transcripts can be obtained, quantified and serial diluted solutions with desired copy number can be made, which is then reverse transcribed to generate a standard curve. However, it should be ensured that mRNA is free from any plasmid DNA contamination. RNA is more unstable than DNA so, stability of standards should be checked more often. This is the reason for DNA standards are generally preferred over RNA.

1.6.2.2 Relative quantification

Relative quantification is a quantification method in which mRNA level of gene of interest is compared with one or several control genes in same sample. This is done by normalizing the mRNA level of target gene to mRNA level of reference gene in the sample, provided that all the samples are treated the same way. The normalized ratio is then compared between different samples. Normalization will compensate for bias created due to differences in starting material amount, RNA isolation process, quality and purity of RNA, RT efficiencies etc that lead to

variation in data generated and hence more reliable and biologically significant results are produced (Bustin et al., 2005; Ginzinger, 2002). However, the expression level of target gene is dependent on expression level of reference gene and relative quantification can be biased if only one normalizing gene is used as the expression level of these genes may vary depending on cell type, experimental treatment, and condition. Thus, appropriate normalization is very critical to produce biologically relevant results. Hence, it is required to validate normalizing gene based on the experimental design (Huggett et al., 2005). To avoid the bias generated by using only one normalizing gene, it is generally recommended to use more than one normalizing gene (Vandesompele et al., 2002).

To quantify relatively, PCR efficiency of the reactions should be similar, preferable 90% or more. PCR efficiency can be determined by doing PCR of a 10-fold serial dilution of cDNA or a positive control template. By plotting Ct value against log (10) concentration of template (illustrated in Figure 6), a trend line is obtained, and the slope of the line will be the function of the PCR efficiency. 100% PCR efficiency is achieved when slope of the line is -3.32. Generally, qPCR efficiency ranging from 90% - 110% and R2 value > 0.980 is desirable for qPCR reactions. R2 value of a standard curve shows how linear the experimental data are and how well they fit in the regression line.

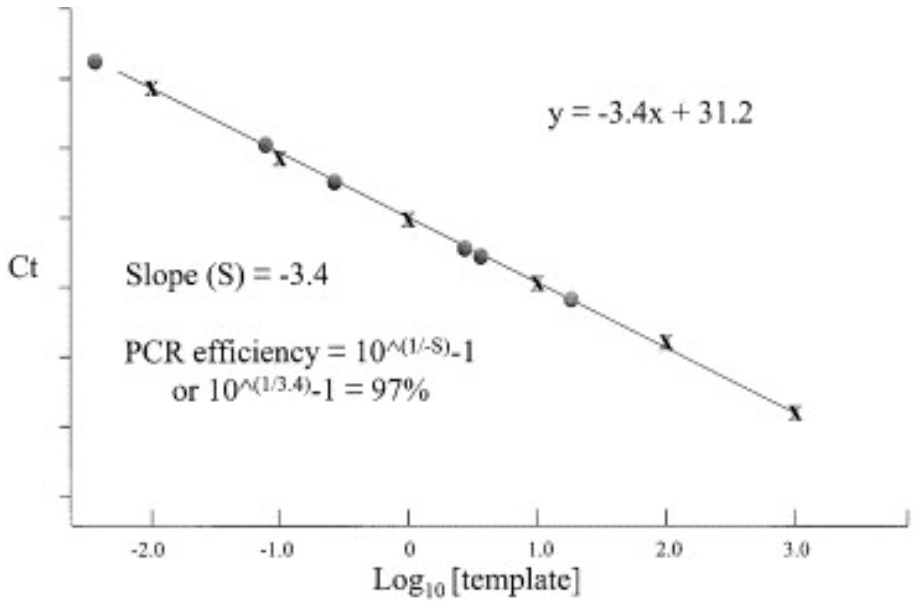

Figure 6: A typical standard curve plot for determination of PCR efficiency, obtained by plotting Ct value of 10-fold serially diluted positive control template vs log (10) template concentration. Figure from Ginzinger (2002)

Generally, relative quantification is done in two ways: comparative Ct method or ΔΔCt method and efficiency based relative quantification method. In ΔΔCt method, quantification is done by calculating the difference between Ct value of control gene and gene of interest, which is then exponent of the base 2 (as in optimal condition PCR efficiency is between 90-110%). This will represent the fold difference of template of these genes. Comparative Ct method is used only when the amplification efficiencies of both target and reference gene are comparable (Livak & Schmittgen, 2001) whereas, efficiency based relative quantification method can be used in both cases i.e., when the amplification efficiency is comparable and non-comparable (Pfaffl & Hageleit, 2001; Pfaffl et al., 2002). The amplification efficiency is affected by various contaminants and inhibitors present in the sample. Amplification efficiencies is generally compared from the standard curves that is prepared from the dilution series of each cDNA sample. This will reduce any error introduced due to different PCR efficiencies of cDNA.

1.7 Aims of study

The main objective of this study was to evaluate and compare the efficiency of different U6 promoters: human U6, mouse U6, salmon U6 and zebrafish U6 promoter in an Atlantic salmon head kidney cell line – SHK-1, using a sgRNA expression assay. The specific objectives of this study were:

- To design and construct plasmid vectors with different U6 promoter and sgRNA
- To optimize transfection protocol and transfect SHK-1 cell line with the vectors
- To design qPCR primers and optimize and validate the qPCR assay
- To quantify expression of sgRNA under different promoters with the qPCR assay

2 Materials and methods

List of all the reagents with their catalogue number and manufacturer, equipment and instruments, software, plasmids, and online tools with their links are listed in Appendix I.

2.1 Experimental outline

In this study, the overall experiment was divided into 3 main steps. First step of the experiment was to construct plasmids that contained 4 different U6 promoters and clone them with GFP knockout gRNA. The second step was to transfect SHK-1 cell line with the plasmids constructed using electroporation. The final step was to quantify the sgRNA expressed under different U6 promoter using qPCR, as shown in Figure 7.

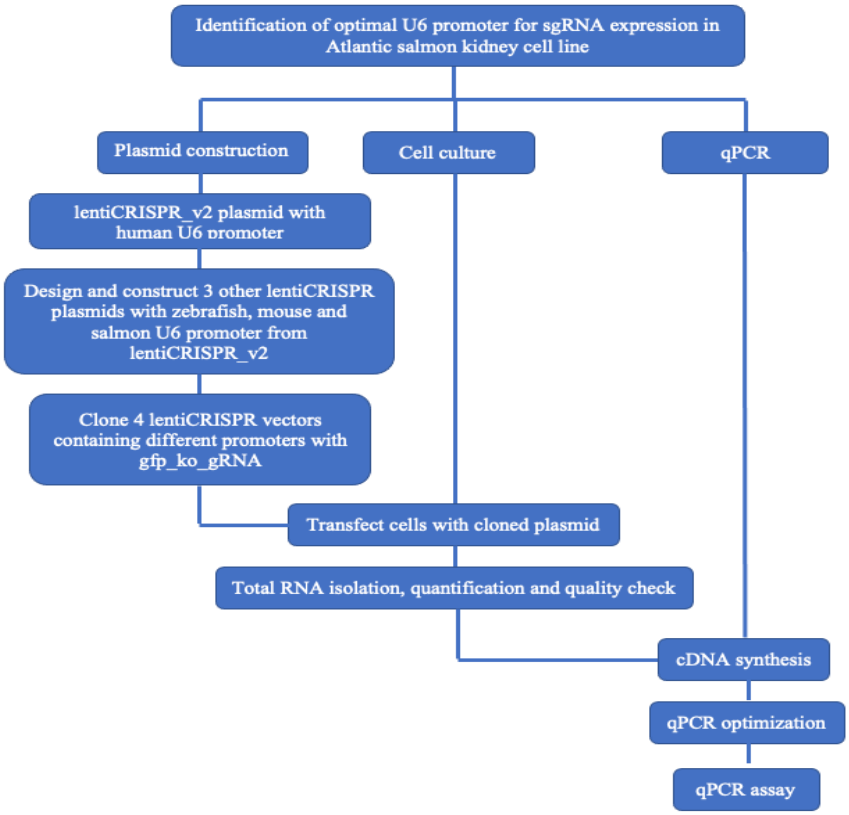

Figure 7: A systematic overview of experimental outline carried out in this study. I started with cloning of plasmid that will contain 4 different U6 promoters followed by insertion of gfp_ko_gRNA. The SHK-1 cells were then transfected with the plasmids and total RNA was isolated from the cells transfected, quantified and quality checked. cDNA was synthesised from the RNA sample and used further for qPCR optimization and qPCR assay.

2.2 Plasmid construction

2.2.1 lentiCRISPR_v2_no_U6 plasmid design for insertion of different U6 promoter

The lentiviral vector, lentiCRISPR v2 vector (Addgene plasmid # 52961), commonly used in GeCKO screening experiments was used as the expression vector. All the digital work related to designing and visualizing plasmid DNA sequence, and primers was carried out in cloud-based software Benchling: https://www.benchling.com. The size of lentiCRISPR v2 vector is 14873 bp and it contains the human U6 promoter followed by two BsmBI cut sites for sgRNA insertion, a gRNA scaffold, a Cas9 gene driven by a EF-1-alpha core promoter (see Figure 8).

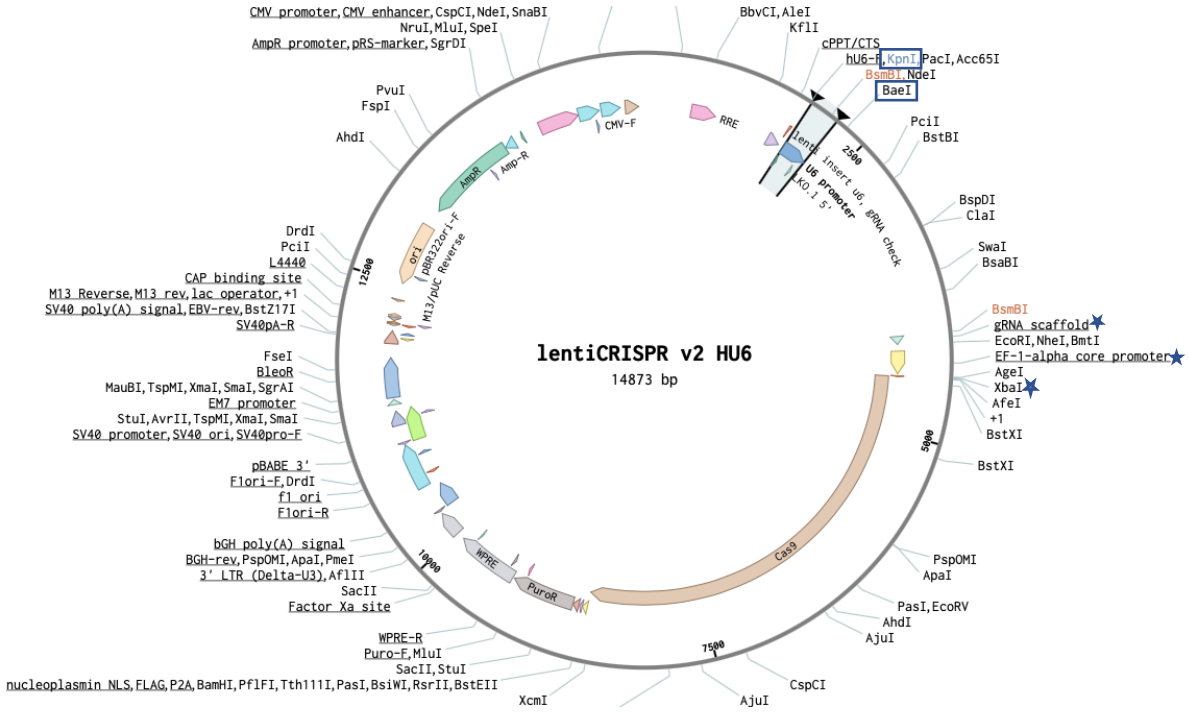

Figure 8: lentiCRISPR v2 plasmid from Addgene contain human U6 promoter, gRNA scaffold (blue starred), EF-1-alpha core promoter (blue starred) and Cas9 gene. Marked blue arrow in the plasmid shows human U6 promoter and nearby restriction sites, KpnI and BaeI (inside blue rectangle). XbaI restriction site (starred) was used later to verify lentiCRISPR_v2_no_U6

To replace the human U6 promoter with zebrafish, mouse, and salmon U6 promoter, a U6 promoter less plasmid was created to facilitate the insertion of the different promoters and named lentiCRISPR_v2_no_U6 plasmid (shown in Figure 9 step 1). To do this, an insert of 38 bp was designed containing XhoI restriction site for later insertion of U6 promoters and BsmBI

for later insertion of gRNA, and it had KpnI and BaeI overhangs for the cloning of the insert into the lentiCRISPR v2 vector (see Figure 10 and Table 1 for sequences). The primers used in this experiment were designed according to method described later in section 2.7.1. The primers and oligonucleotides including insert that were used in this experiment were ordered from Invitrogen. The forward and reverse oligos of insert were annealed using the following protocol: 1 ul of 100 uM of each forward and reverse oligo, 1 ul of 5x T4 buffer (New England Biolabs) was added to an Eppendorf tube and final solution was made up to 10 ul using nuclease free water. The mixture was incubated at 95 ºC for 5 min then cooled down to 25 ºC at 1% ramp rate.

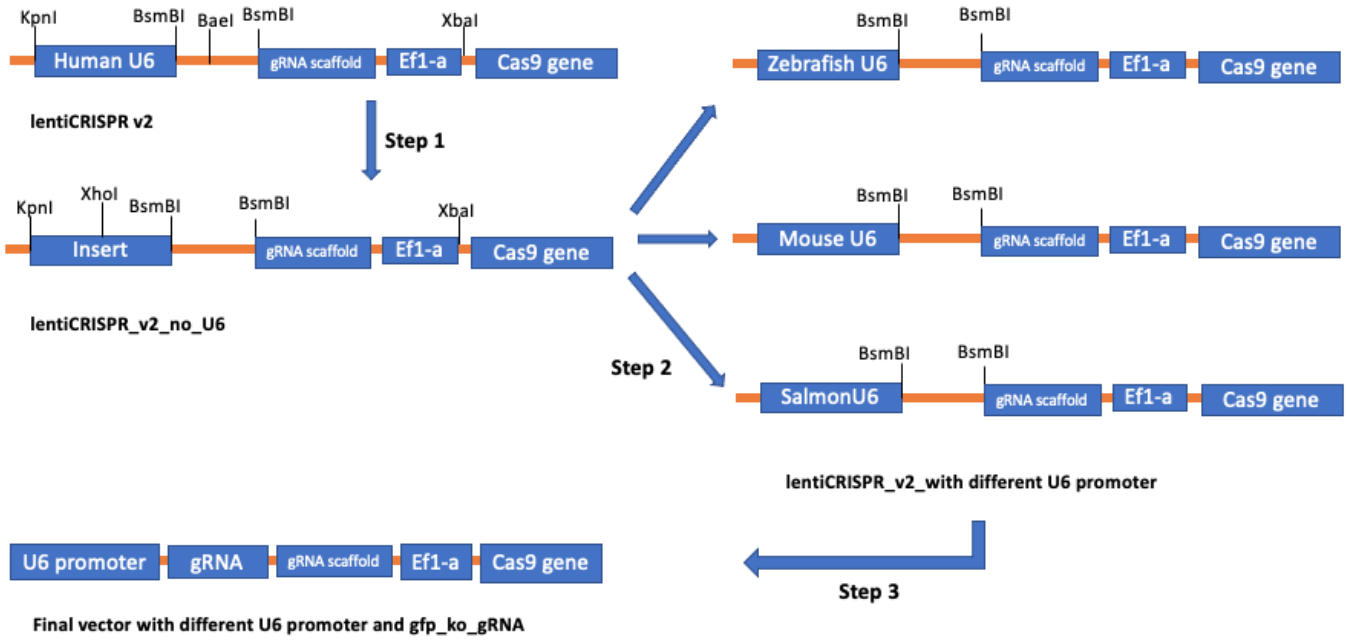

Figure 9: A systematic overview of plasmid construction. Step 1: human U6 promoter in lentiCRISPR v2 plasmid was replaced with insert that contained XhoI and BsmBI cut site for further insertion of different U6 promoter and gRNA. Step 2: cloning of mouse, zebrafish, and salmon U6 promoter in lentiCRISPR_v2_no_U6. Step 3: cloning of gfp_ko_gRNA in all 4 plasmids including lentiCRISPR v2 with human U6.

The human U6 promoter was removed by digesting lentiCRISPR v2 vector with KpnI and BaeI (New England Biolabs) and linear plasmid without human U6 was gel extracted using QIAquick Gel Extraction kit (Qiagen). Linear plasmid without promoter was then ligated with the annealed insert, following the protocol: 50 ng linear vector, 1 ul of 200 times diluted

annealed insert, 2 ul of 5x T4 buffer was added to an Eppendorf tube and final solution was made up to 10 ul with nuclease free water, then 1 ul of T4 ligase (New England Biolabs) was added to final solution and incubated at room temperature for 1 hour. Now the plasmid with insert was named lentiCRISPR_v2_no_U6 hereafter.

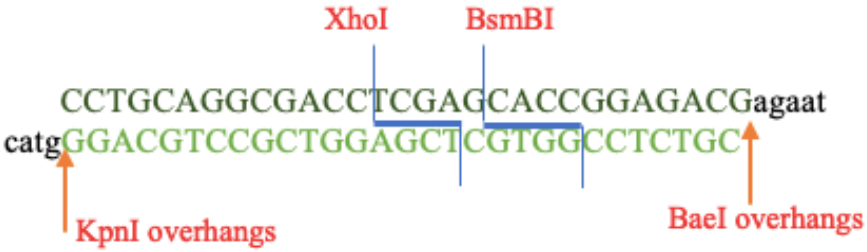

Figure 10: Insert sequence showing XhoI and BsmBI restriction site and KpnI and BaeI overhangs.

50 ul of one shot Stbl3 competent *E. coli* cells (Thermofisher Scientific) were then transformed with lentiCRISPR_v2_no_U6, using the heat shock treatment at 42 ºC for 30 sec. The cells were then kept in 250 ul of SOC medium (Invitrogen) and incubated at 37 ºC, 250 rpm for 1 hour. The cells were then inoculated in LB agar plate with 100 ug/ml Ampicillin and then incubated overnight at 37 ºC.

Random colonies from agar plate were selected and inoculated in 5 ml LB broth with 100 ug/ml ampicillin and incubated overnight at 37 ºC and 250 rpm. 1:1 glycerol stock was prepared for colonies in 50% glycerol (autoclaved), where 500 ul overnight bacterial culture was mixed with 500 ul 50% glycerol and stored in -80 degrees. Rest of the bacterial culture was used for plasmid extraction using ZR plasmid miniprep-classic kit (Zymo Research Crop). To check the success of ligation, the plasmids were then restriction digested with XhoI and XbaI (New England Biolabs) and visualized in a 1% agarose gel. The successfully ligated plasmid upon restriction digestion with XhoI and XbaI (restriction site shown in Figure 8 and 9) will produce two fragments of size 2144 bp and 12398 bp, which is easily visible in agarose gel. Thus, the plasmids that showed two clear bands were selected. Final verification of lentiCRISPR_v2_no_U6 vector was done by Sanger sequencing the plasmid with lenti_check_primer (refer to primer sequences in Table 1) which is located 26 bp upstream the KpnI cut site in the plasmid. For Sanger sequencing, the plasmids were sent to GATC

LightRun, Eurofins Genomics. The chromatograms data (in forms of .abi files) from sequencing were aligned and analysed in Benchling to assess the insertion of correct sequence in the plasmid. After verification the bacteria colony containing correct insert was kept and stored in 1:1 glycerol stock at -80 ºC for further use.

2.2.2 Insertion of different U6 promoter in lentiCRISPR_v2_no_U6

For zebrafish U6 promoter, sequence was obtained from Professor Sebastián Escobar Aguirre (Pontificia Universidad Católica de Chile), and synthetic DNA was ordered from Integrated DNA Technologies (sequences of all heterologous U6 promoters given Appendix III). The synthetic DNA included KpnI and XhoI cut sites followed by 5 random base pairs as overhangs. The zU6 promoter was then restriction digested with KpnI and XhoI (New England Biolabs) to create sticky ends, gel extracted and then ligated into the KpnI and XhoI digested linear lentiCRISPR_v2_no_U6 (as shown in Figure 9, step 2). All the protocols of linearization, ligation, and transformation was done as described in section 2.2.1. Random colonies from agar plate were selected for colony PCR using Platinum II Hot-Start Green PCR Master Mix (2X) (Invitrogen) and ZebrafishU6_fwd_KpnI and ZebrafishU6_rev_XhoI primers (refer to primer sequences in Table 1). Each reaction volume contained: 10 ul of Platinum II Green PCR Master Mix (2X), 4 ul of Platinum GC enhancer, 1 ul of each 10 uM forward and reverse primer and 4 ul of nuclease free water, total volume was 20 ul and finally each reaction mix was inoculated with selected bacterial colony. PCR conditions were: 94 ºC for 2 min, 35 cycles of 94 ºC for 15 sec, 60 ºC for 15 sec and 68 ºC for 15 sec, and finally 4 ºC. The PCR products were visualized in a 1% agarose gel. Plasmids from the colonies which PCRs were showing a clear band on the gel were extracted. For final verification, the plasmids were sent for Sanger sequencing using lenti_check_primer. Thus, created vector was then named lentiCRISPR_v2_zebrafish_U6.

Mouse U6 promoter sequence was PCR synthesized from the pSico plasmid (Addgene plasmid # 11578), using MouseU6_fwd_KpnI and MouseU6_rev_XhoI primers that included KpnI and XhoI cut sites followed by 5 random base pairs as overhangs (refer to primer sequences in Table 1). AccuPrime™ Pfx DNA Polymerase kit (Invitrogen) was used with; PCR reaction mix: 5 ul 10X AccuPrime Pfx Reaction Mix, 1.5 ul of 10 uM forward and reverse primer, 1 ul of template DNA, 0.5 ul AccuPrime Pfx DNA Polymerase and finally nuclease free water was

added to make 50 ul of final volume; and PCR condition: 95 ºC for 2 min, 35 cycles of (95 ºC for 15 sec, 54 ºC for 15 sec and 68 ºC for 15 sec) and finally 4 ºC. The PCR product was then gel extracted, restriction digested with KpnI and XhoI, gel extracted and ligated into linear lentiCRISPR_v2_no_U6. All the protocols of linearization, ligation, transformation, bacterial colony selection, colony PCR condition, plasmid extraction and final verification was done as described in section 2.2.2 for zebrafish U6 promoter insertion. The primer set used for DNA synthesis was used for colony PCR as well and lenti_check_primer was used for sequencing. The vector was named lentiCRISPR_v2_mouse_U6.

Salmon U6 promoter sequence was identified in the genome by our collaborator at CIGENE, Dr. Thomas Nelson Harvey. The salmon U6 DNA was then obtained by PCR synthesis from salmon genomic DNA (provided by Professor Simen Rød Sandve, CIGENE) with the primer SalmonU6_fwd_KpnI and SalmonU6_rev_XhoI which included KpnI and XhoI cut sites followed by 4-5 random base pairs as overhangs (refer to primer sequences in Table 1) using the same kit and same PCR condition as described in section 2.2.2 for mouse U6 promoter insertion. Like mouse and zebrafish U6, the salmon U6 promoter was also restriction digested, ligated into linear lentiCRISPR_v2_no_U6, transformed and finally verified using same method described earlier. The primers used for DNA synthesis and colony PCR were same and lenti_check_primer was used for sequencing. The vector was named lentiCRISPR_v2_salmon_U6. Finally, after verification all the bacteria colony containing correct U6 promoter sequence was kept and stored in 1:1 glycerol stock at -80 ºC for further use.

2.2.3 Insertion of gfp_ko_gRNA in lentiCRISPR vector with different U6 promoter

Since we also want to do GFP knockout assay of a GFP cell line to further assess the promoter efficiency, GFP knockout gRNA (gfp_ko_gRNA) was used as the gRNA and cloned into each plasmid with different U6 promoter (as shown in Figure 9, step 3). The gRNA sequence was taken from Shalem et al. (2014) and designed with BsmBI overhangs on both sense and antisense nucleotide (refer to Table 1 for sequence and overhangs). Sense and anti-sense gfp_ko_gRNA oligonucleotides were annealed following annealing protocol mentioned in section 2.2.1. All four vectors: lentiCRISPR_v2 containing human U6, lentiCRISPR_v2_mouse_U6, lentiCRISPR_v2_zebrafish_U6 and

lentiCRISPR_v2_salmon_U6 were linearized with BsmBI (New England Biolabs) which produced two fragments of size 1885 bp and approx. 13000 bp. The bigger fragment was gel extracted. The annealed gfp_ko gRNA was ligated to each linearized U6 carrying vector and one shot Stbl3 competent *E. coli* cells were transformed separately. All the protocols of linearization, ligation, transformation, bacterial colony selection, colony PCR condition, plasmid extraction and final verification was done as discussed in section 2.2.1 and 2.2.2, except the primers used for colony PCR was gfp_ko_rev and lenti_check_primer (refer to Table 1 for primer sequence) with melting temperature 56 ºC. lenti_check_primer was used for sequencing. The vectors were named:

- lentiCRISPR_v2_human_U6_with_GFP_KO,
- lentiCRISPR_v2_zebrafish_U6_with_GFP_KO,
- lentiCRISPR_v2_mouse_U6_with_GFP_KO and
- lentiCRISPR_v2_salmon_U6_with_GFP_KO.

After verification the bacteria colony containing correct gRNA sequence was kept and stored in 1:1 glycerol stock at -80 ºC for further use. These were the final vectors that were further used for transfection into the SHK-1 cells. The plasmids were prepared in larger quantity using ZymoPURE II Plasmid Midiprep Kit, (Zymo Research Crop) following user manual provided by the manufacturer.

Table 1: List of primers and oligonucleotides and their sequences used in cloning vectors used in this experiment. The base pairs in red are respective cut site and overhangs.

Name of primers and oligonucleotides	Sequence
Insert_fwd	CCTGCAGGCGACCTCGAGCACCGGAGACGAGAAT
Insert_rev	CATGGGACGTCCGCTGGAGCTCGTGGCCTCTGC
ZebrafishU6_fwd_KpnI	TACGGGTACCCACCTCAACAAAAGCTCCTCG
ZebrafishU6_rev_XhoI	ACTACTCGAGTGGGAGTCTGGAGGACGG
MouseU6_fwd_KpnI	CGATGGTACCGATCCGACGCGCCATCTC
MouseU6_rev_XhoI	AGATCTCGAGAATTACTTTACAGTTAGGGTGAGTT
SalmonU6_fwd_KpnI	TGAGGTACCAGTGTACTTGCATATCACCCAGC
SalmonU6_rev_XhoI	TAGACTCGAGAATACAGACATAGGAGAGGCCCTC
lenti_check_primer	CGGGTTTATTACAGGGACA
gfp_ko_fwd (sgRNA)	CACCGGGCGAGGAGCTGTTCACCG
gfp_ko_rev (sgRNA)	AAACCGGTGAACAGCTCCTCGCCC

2.3 Maintaining SHK-1 Cell line

All the cell work was carried out under biosafety cabinet. The SHK-1 cells were maintained at ambient CO_2 and 20 ºC in Leibovitz's L-15 media (Gibco) with 10% Fetal Bovine Serum (Gibco), 40 uM 2Mercapto-ethanol (Gibco) and 1% penicillin-streptomycin (Gibco). First, SHK-1 cells were revived from -80 ºC liquid nitrogen storage using the protocol discussed here. At first a vial of SHK-1 cell line was taken out from liquid nitrogen and fast thawed in 20 ºC water bath. The content of vial was transferred to 9 ml of prewarmed (at 20 ºC) SHK-1 culture media. The cell in media was then centrifuged at 200 g for 5 min. Supernatant was discarded and cell pellet was resuspended in 4 ml prewarmed SHK-1 culture media. Finally, it was transferred to filter-cap T-25 flask and incubated. The cells were let grow until they were 80-90% confluent. In case of too many dead cells, media was changed in between until the desired confluency was reached. When the cells had reached 80-90% confluency level, they were sub-cultured in filter-cap T-75 flask and every time when the confluency had reached 80-90% in T-75 flask, cells were splitted into two flasks. The protocol used for sub culturing and splitting is detailed below.

The confluency level of cells was checked under microscope (Carl Zeiss), if 80-90% confluent go further. The old media from the flask was transferred into a clean falcon tube and the cells were washed twice with PBS solution (Gibco). Then, the old media was syringe filtered using 0.2 um filter and conditioned media was prepared by adding same amount of prewarmed fresh culture media as of filtered media i.e., 1:1 and mixed by inverting the tube 2-3 times. 1 ml 0.25% Trypsin EDTA (Gibco) was added to the flask and the flask was moved in circular motion so that trypsin was spread out evenly all over the cells. The flask was incubated at room temperature for 1 min and then checked under microscope. By this time, if the cells were circular, bright and some started to float on the surface, the flask was then smacked from all 3 sides to facilitate detachment of cells from the flask surface. 5 ml conditioned media was added into the trypsinated flask to stop the action of trypsin and cells were collected by pipetting up and then flushing media to all the corners of the flasks. Thus, collected cells in media was transferred to a Falcon tube and centrifuged at 200 g for 5 min. 10 ul of cells was taken before centrifuging for cell counting and mixed with 10 ul of Typhan Blue dye (Gibco). 10 ul from the mixture was added into the groove of cell counter plate (Bio-Rad) and inserted into TC-20 Automated cell counter (Bio-Rad). The supernatant from the centrifuged cells was discarded,

and cell pellet was resuspended in conditioned media. Depending on the cell count, resuspended cells were thus distributed in such a way that each T-75 flask get 2 million cells. Finally, conditioned media was added to make final volume 10 ml in each flask and incubated at 20 ºC incubator until the desired confluency was reached.

2.4 Transfection via Electroporation

Transfection was done using Neon Transfection System (Invitrogen) and Neon transfection 10 ul Kit (Invitrogen). The transfection protocol for SHK-1 cell line was experimentally optimized by my co-supervisor PhD fellow Noman Reza based on the proportion of EGFP (Enhanced Green Fluorescence Protein) expressing cells, 2 days after getting transfected. For the optimization, 24 well optimization protocol mentioned in user manual provided by the manufacturer was used. A 24 well plate and 0.5 ml plating media volume (used same media condition that is used for growing SHK-1 cell line, but without antibiotics) was used throughout the experiment. Cells with different cell densities – $0.5x10^5$, $1x10^5$ and $2x10^5$ were transfected with 1 and 2 ug of pEGFP-Puro (Addgene plasmid #45561). 24 different pulse programs mentioned in the user manual was applied to each well. Two days after transfection cells were observed for GFP expressing cells, non-GFP expressing cells and dead cells under a fluorescence microscope (Carl Zeiss). The optimal transfection protocol was selected based on the proportion of EGFP expressing cells over non-GFP expressing cells and viable cells over non-viable cells.

After optimization, cell density $2x10^5$ cells/well, and 2 ug of DNA (plasmid) amount was selected for further experiment. The pulse program with pulse voltage 1050, pulse width 30 and pulse no. 2 was used throughout the experiment except for the negative control without electroporation.

Two days before transfection desired number of cells were seeded on T-75 flask. On the day of transfection, plating media was prepared, and 0.5 ml was added to each well of 24 well plate, and pre-incubated. Then, the seeded flasks were taken out, media was removed, the cells were washed with PBS, trypsinized and all the cells were collected. 10 ul from the cell collection was used for cell count and rest was centrifuged, the supernatant was discarded, and the pelleted cells were washed with PBS. The cell pellet was resuspended in PBS and centrifuged. The

supernatant was then removed and then the cells were resuspended in Opti-MEM I Reduced Serum media (Gibco), pre-warmed at 20 ºC, and the number of cells and volume of media was adjusted to get 2 x 10^5 cells in 10 ul of media in each well. The cells were then divided into 5 groups, volume depending on number of wells for each group and mixed with 5 different plasmid DNA, namely:

- pEGFP-Puro,
- lentiCRISPR_v2_human_U6_with_GFP_KO,
- lentiCRISPR_v2_zebrafish_U6_with_GFP_KO,
- lentiCRISPR_v2_mouse_U6_with_GFP_KO and
- lentiCRISPR_v2_salmon_U6_with_GFP_KO

The pEGFP-Puro was used as controls, one negative control where cells were mixed with plasmid but not electroporated, and one positive control, where cells were mixed with plasmid and electroporated. With the help of Neon pipette, 10 ul of cell and DNA mix was taken and electroporated, following the user manual. The above-mentioned pulse program was applied. After transfection the mixture was evenly distributed into the well of the pre-incubated 24-well plate and incubated. 4 wells were used for each plasmid with different promoters.

One day after transfection, transfected cells were selected via antibiotic selection. The cells were washed twice with PBS and replaced with antibiotic selection media i.e., regular media with 1ug/ml Puromycin (Gibco). Two days after transfection, the cells were harvested for total RNA isolation.

2.5 Total RNA isolation, its quantification and quality check

RNeasy Plus Universal Kit (Qiagen) and QIAzol Lysis reagent (Qiagen) was used for the total RNA isolation from the transfected cells. First, media was removed, and cells were washed twice with PBS. 900 ul of QIAzol lysis reagent was added to each well, cells were then scraped with the help of cell scraper and collected in an Eppendorf tube. Then the cells were incubated at RT for 5 min. The cells from all four wells transfected with each plasmid were pooled into one tube and used for RNA extraction. Hereafter the protocol provided by the manufacturer

was followed. The kit uses organic phase extraction using chloroform for removal of cell debris and other residues. In addition, the kit contains gDNA eliminator solution that claims to effectively remove genomic DNA contamination during organic phase extraction. At the end, I had 4 RNA samples from cells transfected with 4 different plasmids:

- lentiCRISPR_v2_human_U6_with_GFP_KO,
- lentiCRISPR_v2_zebrafish_U6_with_GFP_KO,
- lentiCRISPR_v2_mouse_U6_with_GFP_KO and
- lentiCRISPR_v2_salmon_U6_with_GFP_KO

which were named hU6, zU6, mU6 and sU6 RNA samples respectively. Finally, isolated total RNA samples were stored at -80 ºC.

The concentration (in ng/ul) and purity of the RNA isolated was measured using NanoDrop 8000 Spectrophotometer (Thermofisher Scientific) and NanoDrop 8000 V2.1.0 software. Generally, spectrophotometer detect nucleic acid at 260 nm, protein at 280 nm and organic compounds like chaotropic salt, carbohydrates, phenols, and aromatic compounds in general at 230. Any possible contamination of protein and organic compounds can be determined in terms of optical density (OD) 260/280 ratio and OD 260/230 ratio, which was also measured using NanoDrop. The sample is considered of pure if the ratio is greater or equal to 1.8.

The RNA integrity was assessed using Agilent 4150 TapeStation system (Agilent) according to RNA ScreenTape Assay for TapeStation Systems quick guide. The TapeStation system uses capillary electrophoresis to separate the RNA strands which is detected by laser induced fluorescence detection. Generally, ribosomal RNA 18S and 28S is used to assess the integrity of total RNA in sample by generating a gel-like image and a graph of fluorescence plotted over time known as electropherogram. An intact RNA will give 2 sharp peaks of 18S and 28S rRNA in the electropherogram and 2 clear bands in gel-like image, where 28S is approximately twice in size of 18S (see Figure 11). In contrast, a degraded RNA will appear as a low molecular weight smear and lack sharp peak and band in electropherogram and gel-image respectively. Finally, the quality of RNA is given in terms of RNA integrity number equivalent (RINe) value, ranging from 1-10, 10 being the highest quality RNA. The RINe value is calculated by the software using mathematical model that make assessment based on ribosomal peak ratios,

separation, and purity of the sample. In addition to RNA integrity, ScreenTape assay can also quantify the RNA in the sample.

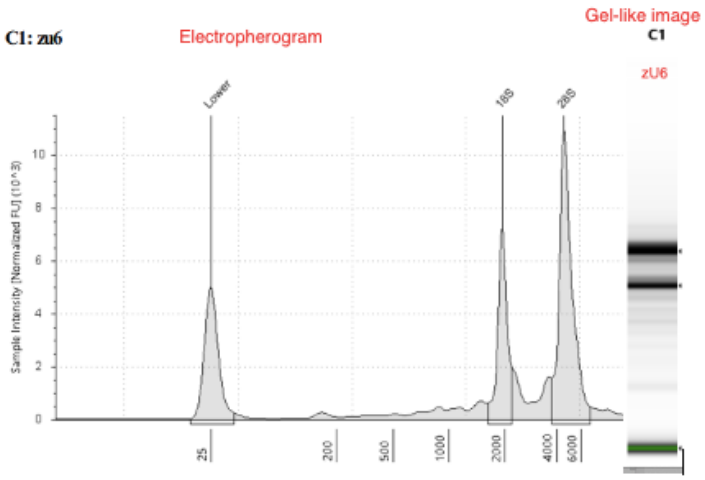

Figure 11: An electropherogram and a gel-like image of cells transfected with lentiCRISPR_v2_zebrafish_U6_with_GFP_KO generated by RNA ScreenTape assay.

2.6 cDNA synthesis

500 ng of total RNA in 20 ul of reaction volume of each RNA sample was used in the cDNA synthesis reaction using QuantiTect Reverse Transcription Kit (Qiagen) following the user manual. The kit comes with unique gDNA wipe-out buffer and claims to effectively remove genomic DNA from the sample before cDNA synthesis. In addition, the kit comes with Quantiscript Reverse Transcriptase enzyme and RT primer mix, which consists of mixture of random primers and oligo-dT primers. The synthesized cDNA was then 5 times diluted by making final volume up to 100 ul with nuclease free water. A qPCR assay was carried out with this sample (described later in section 2.7.4) with a 5x diluted negative RT control (RNA sample before reverse transcription reaction) to check if our samples were DNA contamination free. Unfortunately, Ct value showing high DNA contamination was observed. Thus, it was decided to treat the RNA samples with DNase I and repeat the qPCR assay.

To remove all DNA from the samples, 500 ng of RNA from each RNA sample was treated with DNase I, RNase-free HC (Thermofisher Scientific), according to the user manual. Since,

the concentration of total RNA in some samples was low and very little volume was left, instead of 10 ul reaction volume 14 ul was used, to have 500 ng of RNA in the reaction volume, along with 0.6 ul DNase I enzyme and 1.4 ul 10x reaction buffer. DNase I treated RNA was then directly used in the cDNA synthesis reaction using QuantiTect Reverse Transcription Kit except the step with gDNA wipe-out buffer due to reaction volume limit, as the RNA sample was already treated with DNase I. Finally, the cDNA was 5x diluted as earlier. The RNA samples after DNase I treatment was named as dhU6, dzU6, dmU6 and dsU6 RNA.

2.7 qPCR

In this experiment, absolute quantification method was used for quantification of sgRNA and Cas9 mRNA level in SHK-1 cells, where set of external standards using PCR fragment was used and amount of sgRNA and Cas9 expressed were measured in terms of copy number per ul. In addition, we did relative quantification by using absolute quantification data where Cas9 gene was taken as normalization gene. The sgRNA mRNA level was normalized with respect to Cas9, by dividing copy number sgRNA by Cas9. This kind of relative quantification from absolute quantification data was done previously, described by Leong et al. (2007).

SYBR Select Master Mix (Applied Biosystems) was used for preparing the reaction mix, which contains SYBR GreenER dye, AmpliTaq DNA polymerase UP, cNTPs with dUTP/dTTP blend, heat liable UDG, ROX passive reference dye and optimized buffer components. Instrument CFX96 Touch Real-Time PCR detection system (Bio-Rad) with 96-well BR-white optical plate was used for running the qPCR. The instrument uses CFX Maestro software to analyse the qPCR data.

2.7.1 Primer design

Online tool Oligo Calc: http://biotools.nubic.northwestern.edu/OligoCalc.html was used with values of 300 nM for primer concentration and 50 nM for salt concentration to assess the primer properties like, GC content, melting temperature (T_m), self-complementary regions and 3' complementarity region of forward and reverse primers. 2 sets of primers for both sgRNA and Cas9 was designed such that the amplicon size ranged between 75-150 bp and T_m around 60ºC.

2.7.2 Primer and annealing temperature optimization

To find out optimal primer among two primers and its annealing temperature, all four cDNA samples were pooled and used to optimize primer and annealing temperature. A reaction mix for 8 reactions per primer pair for sgRNA and Cas9 was prepared separately as: 10 ul SYBR Select Master mix, 0.8 ul 10 uM forward and reverse primer, 1 ul of pooled cDNA sample and nuclease free water was added to make 20 ul final reaction volume. PCR condition: 50 ºC 2 min, 95 ºC 2 min and 40 cycles of 95 ºC 15 sec and temperature gradient from 55 ºC to 65 ºC for 1 min, was applied. In temperature gradient, 8 temperature points between 55 ºC to 65 ºC were selected and each temperature was applied to a reaction tube per primer pair. Thus, 8 annealing temperatures were evaluated in same reaction plate. The primer pair that showed single sharp peak in the melting curve analysis and temperature that had lowest Ct value were selected and annealing temperature was set to 63 ºC for further qPCR experiments. Hereafter, PCR condition: 50 ºC 2 min, 95 ºC 2 min and 40 cycles of 95 ºC 15 sec and 63 ºC for 1 min, was applied.

Table 2: List of selected qPCR primers for sgRNA and Cas9 expression and their sequence.

Primer name	Sequence
GFP_KO_qPCR_pLenti_Fwd 2	GCGAGGAGCTGTTCACCG
GFP_KO_qPCR_pLenti_Rev1,2	ACTCGGTGCCACTTTTTCAAG
Cas9_KO_qPCR_pLenti_Fwd 1	ACCTATGCCCACCTGTTCG
Cas9_KO_qPCR_pLenti_Rev1	AGGATTGTCTTGCCGGACTG

2.7.3 Standard curve and qPCR efficiency estimation

In this experiment, a 10-fold dilution series with known concentration set of external DNA standards was prepared using PCR fragment. For this, pooled cDNA sample was PCR amplified using Platinum II Taq Hot-Start DNA Polymerase Kit (Invitrogen). 50 ul reaction volume was prepared using previously selected primer set for sgRNA and Cas9 separately and pooled cDNA sample, which included 10 ul of 5 x Platinum II PCR buffer, 1 ul of 10 uM dNTP, 1 ul of 10 uM forward and reverse primer, 10 ul of Platinum GC enhancer, 1 ul of template DNA, 0.4 ul of Taq. DNA polymerase and water was added to make volume 50 ul. The PCR program used was: 94 ºC 2 min (94 ºC, 15 sec – 63 ºC, 15 sec – 68 ºC 15 sec) X 35, 4 ºC. The PCR products were then visualized in a 2% certified low range ultra-agarose gel

(Bio-Rad). Then, the DNA band was cut, and gel extracted using QIAquick gel Extraction Kit. Finally, the concentration of the extracted PCR product for sgRNA and Cas9 was measured using NanoDrop spectrophotometer.

The copy number of the PCR products were calculated using the formula below:

$$copy\ number/ul = \frac{concentration\ of\ amplicon\left(in\ {}^{g}/_{ul}\right)}{amplicon\ length\ (bp)\ \times\ 660} \times 6.022 \times 10^{23}$$

The PCR products were then diluted to in 1 x 10^{10} copy number/ul and then serially diluted until 1 x 10^{1} copy number/ul, for both sgRNA and Cas9 PCR amplicon. 2 ul of template DNA from dilution range from 1 x 10^{6} to 1 x 10^{1} copy number/ul was taken for preparing reaction mix and PCR condition described earlier was applied. qPCR was performed in triplicate. The Ct value was calculated by the CFX Maestro program using baseline subtracted curve fit setting for baseline determination and single threshold mode for Ct determination mode. The standard curve was generated by the CFX Maestro software by plotting Ct value against log initial quantity of the 10-fold serially diluted PCR amplicons. qPCR efficiency and R^2 value was then calculated by the program, based on standard curve.

2.7.4 qPCR assay

2 ul of cDNA from all 4 cDNA samples from each promoter was used in qPCR reaction with GFP_KO_qPCR_pLenti and Cas9_KO_qPCR_pLenti primers separately, the qPCR was done in triplicates. Different standards, positive and negative controls were included in each plate. One standard curve set for both sgRNA and Cas9 was included as described in previous chapter. Two negative controls were included: 2 ul of one no cDNA template control in duplicate and one non-RT control (the RNA 5 x diluted) was run in triplicate. 2 ul of plasmid DNA that was used to transfect the cells (diluted to 1 x 10^{4} copy number/ul) was run as positive control. PCR conditions as described earlier.

2.7.5 Data analysis

The Sq value was calculated by the CFX Maestro software by extrapolating Ct value of the samples in the standard curve. The Sq value represents the initial amount of cDNA in the sample (in terms of copy number/ul). The average of Sq value and its standard deviation for sgRNA and Cas9 expression was then calculated for 4 cDNA samples from each promoter. Finally, sgRNA expression relative to Cas9 expression was also calculated by dividing Sq value of sgRNA by Cas9. All data are expressed as average + SD, n=3, where n is 3 technical qPCR replicates not biological replicate.

3 Results

3.1 Plasmid construction

In this experiment, we constructed 4 different lentiCRISPR vectors with human, mouse, zebrafish, and salmon U6 promoter to express sgRNA for GFP knockout that will be used in the future for optimization of genome wide CRISPR knockout screening experiment. The original lentiCRISPR v2 vector contained human U6 promoter that was replaced with the insert to obtain lentiCRISPR_v2_no_U6. 3 more lentiCRISPR vector with Mouse, zebrafish, and salmon U6 promoter were then obtained by cloning the lentiCRISPR_v2_no_U6 plasmid with synthetic DNA fragment of mouse, zebrafish, and salmon U6 promoter. The final plasmids were obtained by cloning gfp_ko_gRNA into those 4 plasmids (refer to Figure 9, step 1-3). To assess the success of lentiCRISPR_v2_no_U6 plasmid, a restriction digestion analysis with XhoI and XbaI was performed which showed release of a filler fragment of size between 2-3 kb. Whereas the unsuccessfully ligated plasmid showed single band of size above 10kb, Figure 12A.

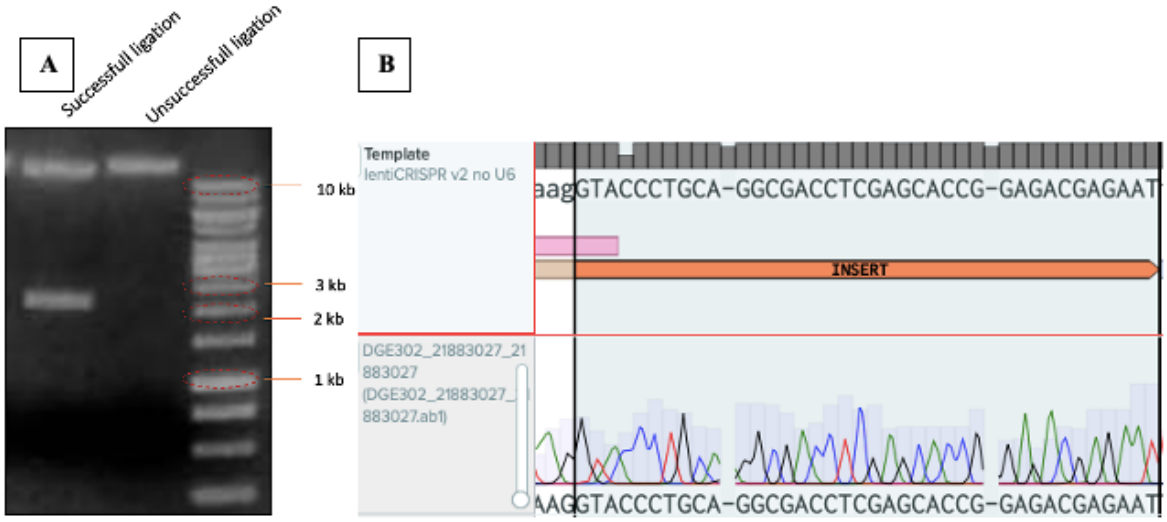

Figure 12: Confirmation of the success of lentiCRISPR_v2_no_U6 plasmid construction by replacing human U6 with insert. A: restriction digestion analysis of successfully ligated plasmid by XhoI and XbaI gave a filler fragment of size above 2 kb. B: further confirmation of correct sequence insertion via Sanger sequencing showed exact match. Two nucleotides are missing from the template and sample sequence because multiple sample sequences were aligned at the same time and some samples had insertion of two extra nucleotides that generated gap in the template.

In addition, to assess correct insert sequence being inserted, Sanger sequencing of the plasmid showed exact match of the plasmid insert sequence with the template, Figure 12B. The successful cloning of mouse, zebrafish, and salmon U6 promoter in lentiCRISPR_v2_no_U6 and gfp_ko_gRNA (as shown in Figure 13) in each plasmid was also determined via Sanger sequencing.

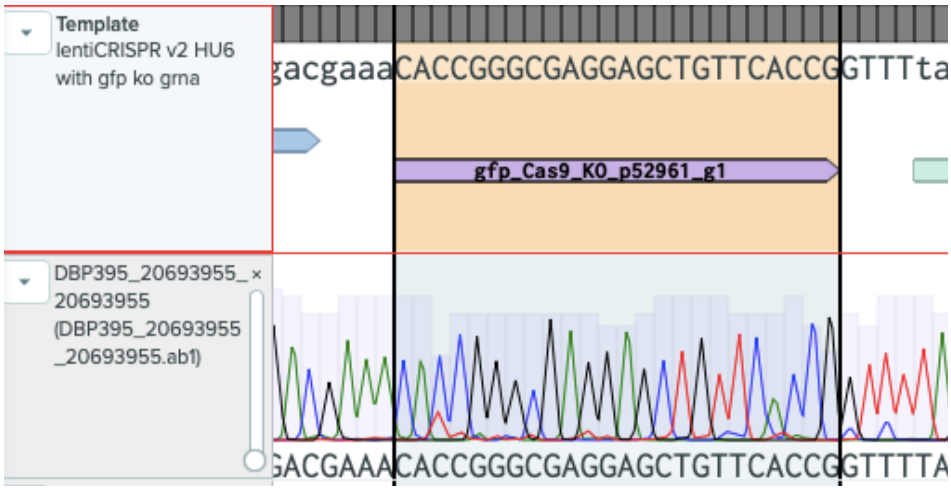

Figure 13: Confirmation of successful cloning of gfp_ko_gRNA in lentiCRISPR_v2 with Sanger sequencing, here exemplified by the lentiCRISPR_v2_human_U6_with_GFP_KO plasmid.

3.2 Electroporation of SHK-1

The success of electroporation in SHK-1 cells was determined visually for cells transfected with EGFP and with PCR followed by cDNA synthesis for cells transfected with plasmids containing 4 different U6. Cells transfected with EGFP showed bright green fluorescence with 80-90% confluency under fluorescence microscope, Figure 14B. Whereas, the cells transfected with plasmids containing 4 different U6 had lower confluency, around 50-60%, compared to cells transfected with EGFP.

Pooled cDNA from cells transfected with 4 different plasmids upon PCR amplification by sequence specific qPCR primer showed clear bands in agarose gel, Figure 14A. Amplicon size for sgRNA and Cas9 was 93 and 128 respectively. However, this result was generated with initial cDNA samples that was DNA contaminated as they were not treated with DNase I. The

success of electroporation and successful expression of the plasmids in the cells cannot be determined with this result.

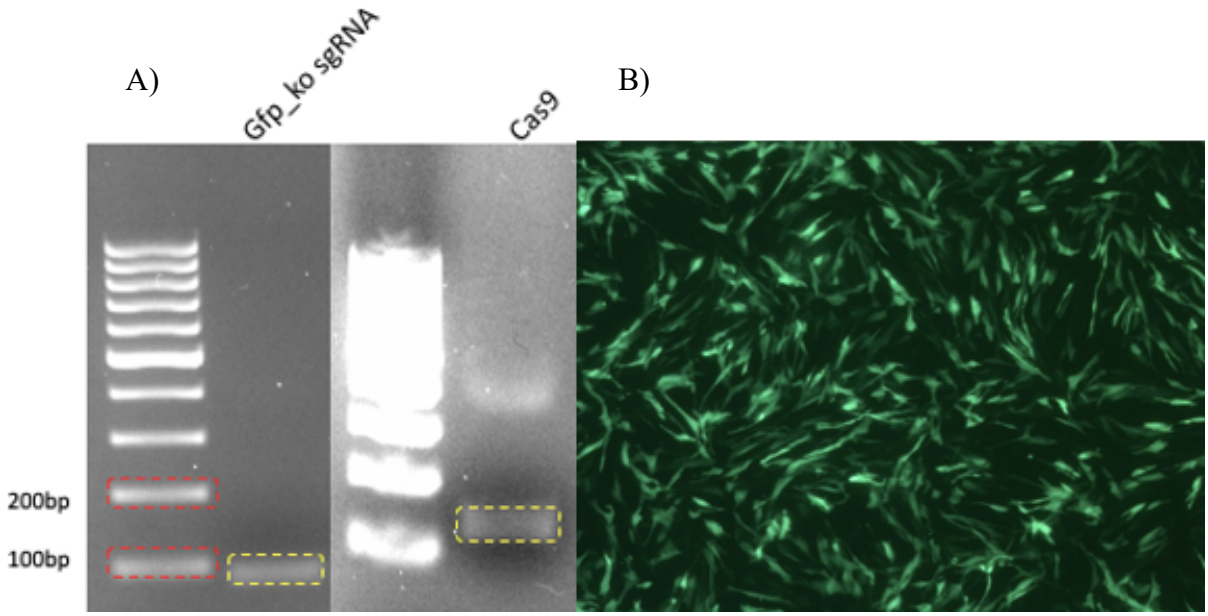

Figure 14: A) sgRNA and Cas9 PCR amplicon showing clear bands in agarose gel. B) GFP expressing cells visualized under fluorescence microscope.

3.3 RNA quantification and quality check

The concentration of total RNA in all 4 RNA samples, hU6, zU6, mU6 and sU6, was lower than 100 ng/ul when measured by NanoDrop. The ScreenTape assay showed the concentration was even lower, less than 50 ng/ul, except for zU6 whose concentration didn't change in both assessments, (different values obtained from NanoDrop and ScreenTape assay are listed in Table 3). The purity assessment of RNA samples showed the samples were free from protein contamination as the value for OD 260/280 was equal or greater than 1.8 except for mU6 RNA sample which was slightly lower than 1.8. However, most of the RNA samples were highly contaminated with organic compounds, exception zU6 that was free from organic compounds. Overall, the NanoDrop data showed that among the 4 RNA samples, zU6 RNA sample was most pure with highest concentration of total RNA isolated and mU6 was highly contaminated with lowest concentration of total RNA isolated. The RINe value showed the RNA samples were intact and free from RNA degrading compounds. The gel-like image and electropherogram also supported the RINe value by showing 2 clear bands and 2 sharp peaks

respectively, Figure 15 A-B. Due to limited sample amount and its low concentration, the quantity, purity, and quality of RNA sample after DNase I treatment couldn't be assessed.

Table 3: List of different values obtained from NanoDrop and ScreenTape assay.

	hU6	zU6	mU6	sU6
Concentration (ng/ul) from NanoDrop	55.3	78.5	42.5	54.1
OD 260/280	1.85	1.92	1.78	1.82
OD 260/230	1.27	1.80	1.09	0.87
RINe number	9.7	9.8	9.3	9.0
Concentration (ng/ul) from RNA ScreenTape assay	41.2	78.1	36.5	45.7

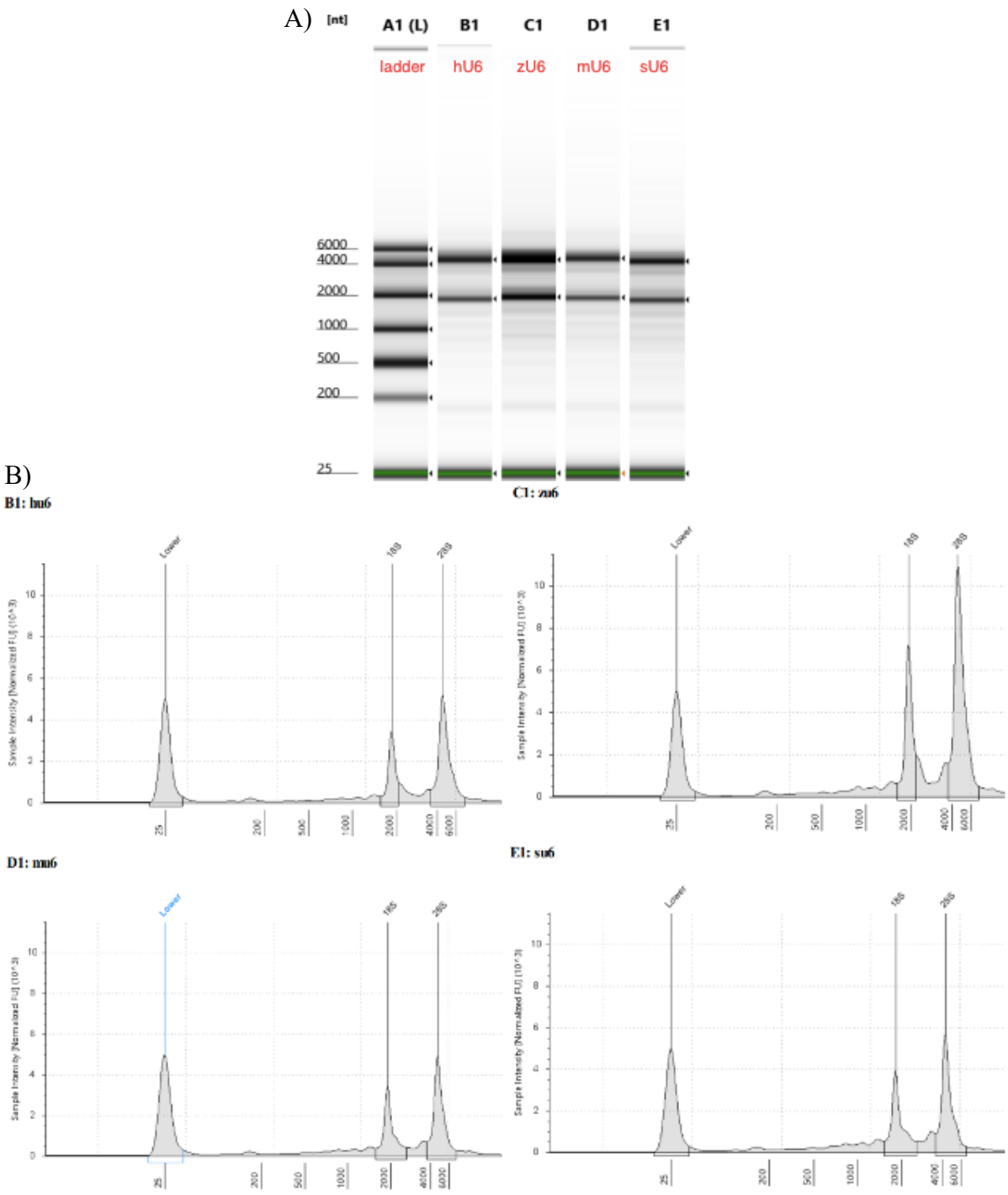

Figure 15: Gel-like image (A) and electropherogram (B) of all four RNA samples marked in the figure.

3.4 qPCR assay optimization

The melting curve analysis showed single sharp peak for both sgRNA and Cas9 expression and lowest Ct value, 28.35 and 25.79 respectively at 63 ºC annealing temperature for the primers that were selected for further qPCR assay, see Figure 16 A-B.

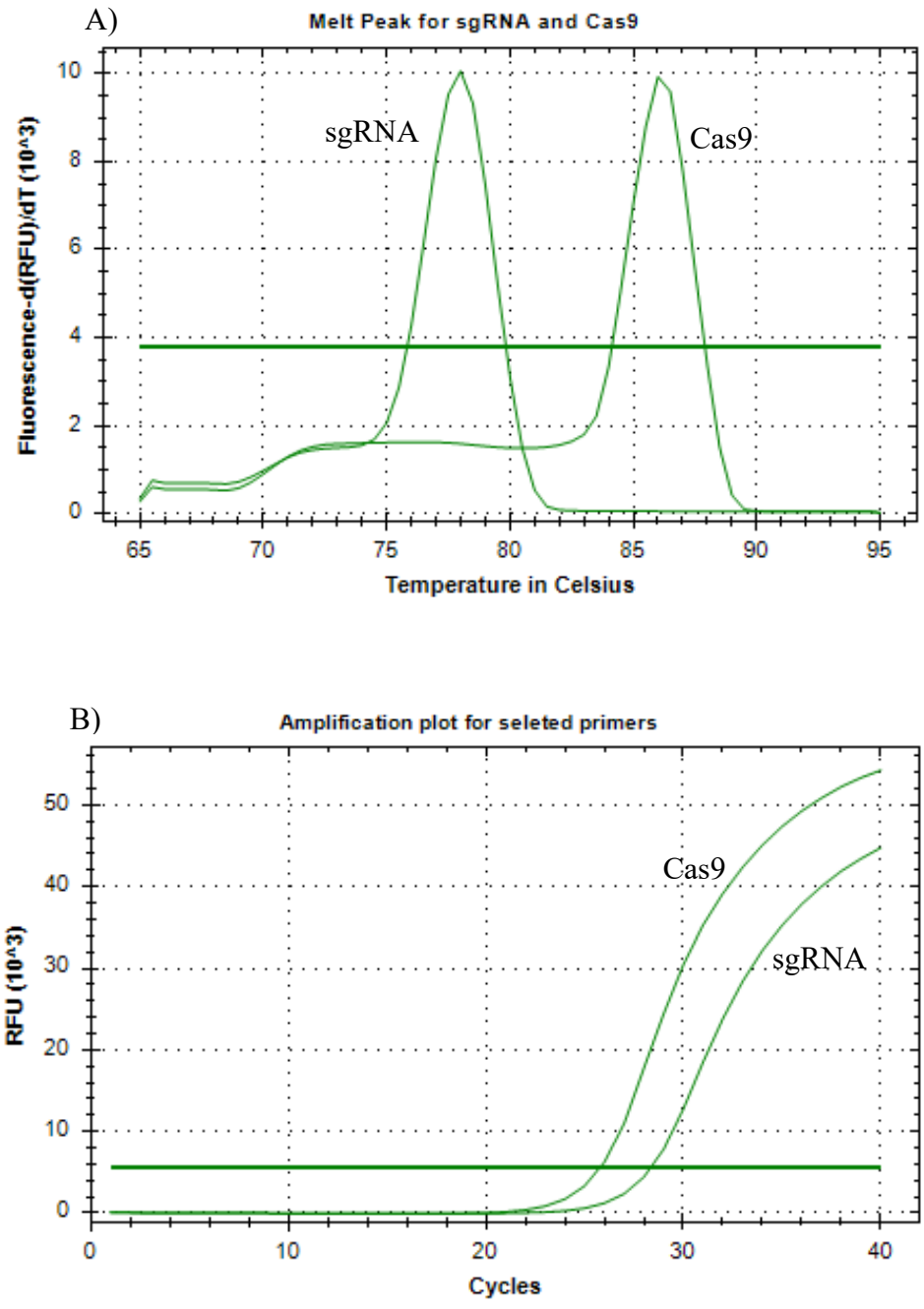

Figure 16: Melting curve (A) and amplification plot (B) for sgRNA and Cas9 expression at 63 ºC annealing temperature.

CFX Majestro software generated a standard curve made from 10-fold dilution of PCR amplicon of sgRNA and Cas9 with slope value of 3.2 and 3.3 respectively. From this curve the efficiency and R^2 value was calculated by the program itself as 104.5% and 0.99 respectively for sgRNA and 97.4% and 0.98 respectively for Cas9, Figure 17 A-B.

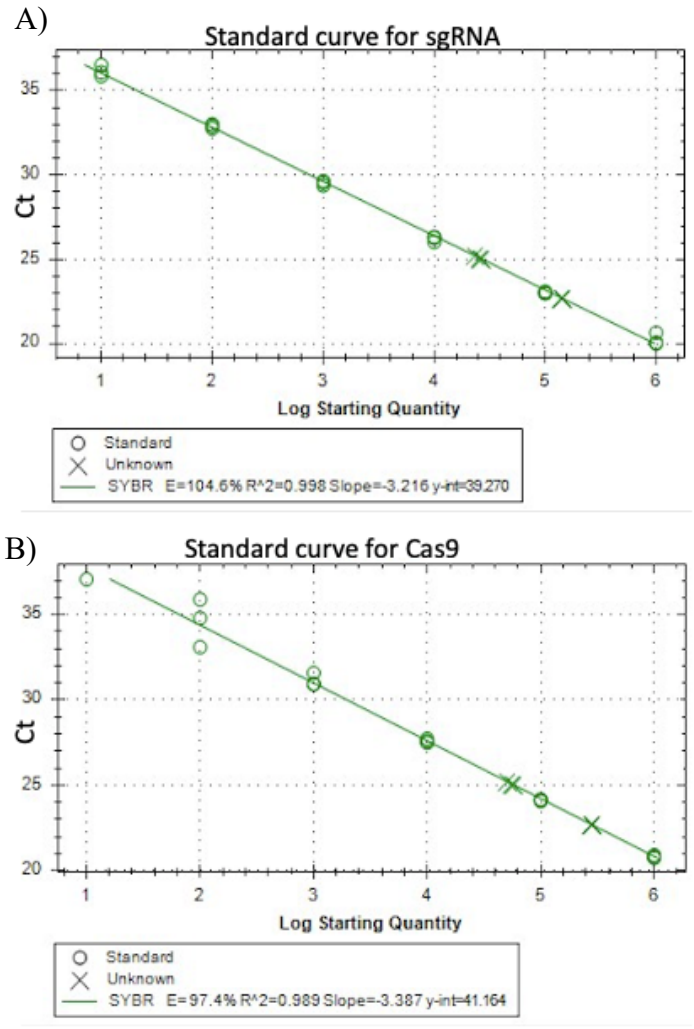

Figure 17: Standard curve for a) sgRNA and b) Cas9 generated by plotting Ct value against log of starting quantity in copy number/ul by CFX majestro program.

3.5 qPCR assay

3.5.1 Initial qPCR results showed high contamination of DNA

The absolute quantification of sgRNA in different cDNA samples showed that sgRNA level from human U6 promoter was comparatively more than other promoters, lowest being salmon U6 promoter, see Figure 18A (see Appendix II for raw Sq and Ct values). However, these values are not corrected for possible differences in transfection efficiency, which we can assess with the measurement of Cas9 mRNA level. Indeed, absolute quantification of Cas9 in different samples also showed that the mRNA level of Cas9 in the cells that where transfected with lentiCRISPR_v2_human_U6_with_GFP_KO was highest, Figure 18B, indicating that the transfection efficiency was highest for this plasmid. In contrast, when sgRNA level in the samples was normalized to Cas9, the expression of sgRNA by salmon U6 promoter was observed to be highest, followed by zebrafish, mouse, and human, see Figure 19. These results show the necessity of a reference to normalize the qPCR results.

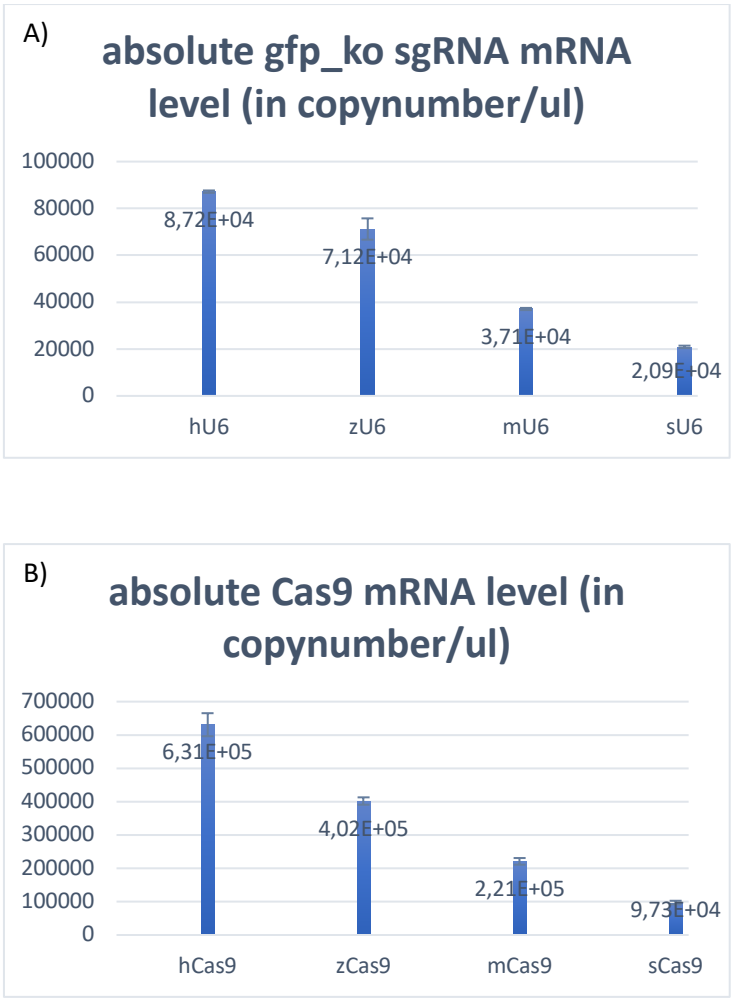

Figure 18: Absolute quantification of sgRNA (A) and Cas9 (B) in different cDNA samples from cells transfected with different plasmids. Data are expressed as average +SD, n=3, where n is technical qPCR replicate.

Furthermore, despite the use of a gDNA wipe-out buffer in our cDNA synthesis kit, the hU6 RNA sample which was run as a negative RT control showed high level of DNA contamination in our RNA sample. The Ct value 28.28 and 27.63 for sgRNA and Cas9 expression respectively was close to the values of the hU6 cDNA sample, Ct value 27.18 and 24.46 for sgRNA and Cas9 mRNA levels (see Appendix II for raw Sq and Ct values), respectively.

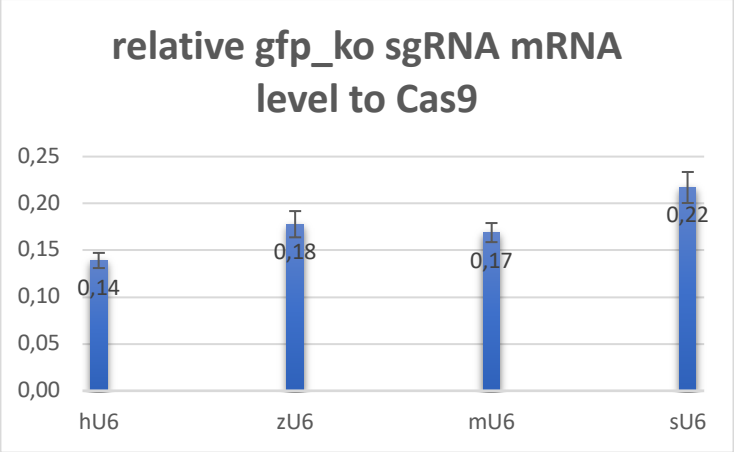

Figure 19: sgRNA levels normalized to Cas9 in different cDNA samples from cells transfected with different plasmids. Data are expressed as average +SD, n=3, where n is technical qPCR replicate.

3.5.2 DNase I treated samples

After an additional DNase I treatment, no Ct values was observed for negative RT controls, indicating that the samples were free from DNA contamination this time. Low level of sgRNA expression, 13.1 copy number/ul, was seen in case of dzU6 cDNA sample. Regarding Cas9 mRNA level, 14.7 copy number/ul was seen in case of dzU6 cDNA sample (see Figure 20 and Appendix II for raw Sq values). In the dhU6 and dsU6 samples, the Cas9 mRNA level was at the detection limit; 2 out of 3 technical replicates showed low Cas9 expression, 5.4 copies/ul for dhU6 sample and 1 technical replicate showed Cas9 expression for dsU6 sample, 3.9 copies/ul. No Ct values for neither sgRNA nor Cas9 was observed in case of dmU6 RNA sample.

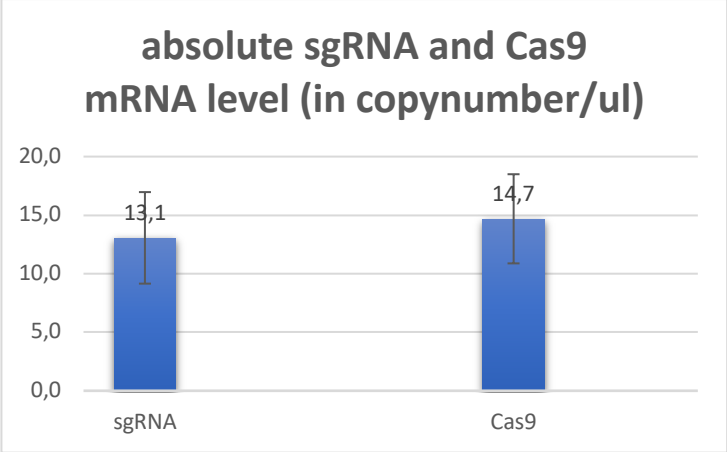

Figure 20: Absolute quantification of sgRNA and Cas9 in dzU6 sample after DNase I treatment. Data are expressed as average +SD, n=3, where n is technical qPCR replicate.

Despite the Ct values were higher than 35 (see Figure 21 A-B and Appendix II for raw Ct values), we can trust this result, since the negative template control showed no Ct value. This was a great achievement for us as this showed that the plasmid that we designed worked and zebrafish U6 was able to produce sgRNA. However, no sgRNA expression was detected in dhU6, dmU6 and dsU6 samples.

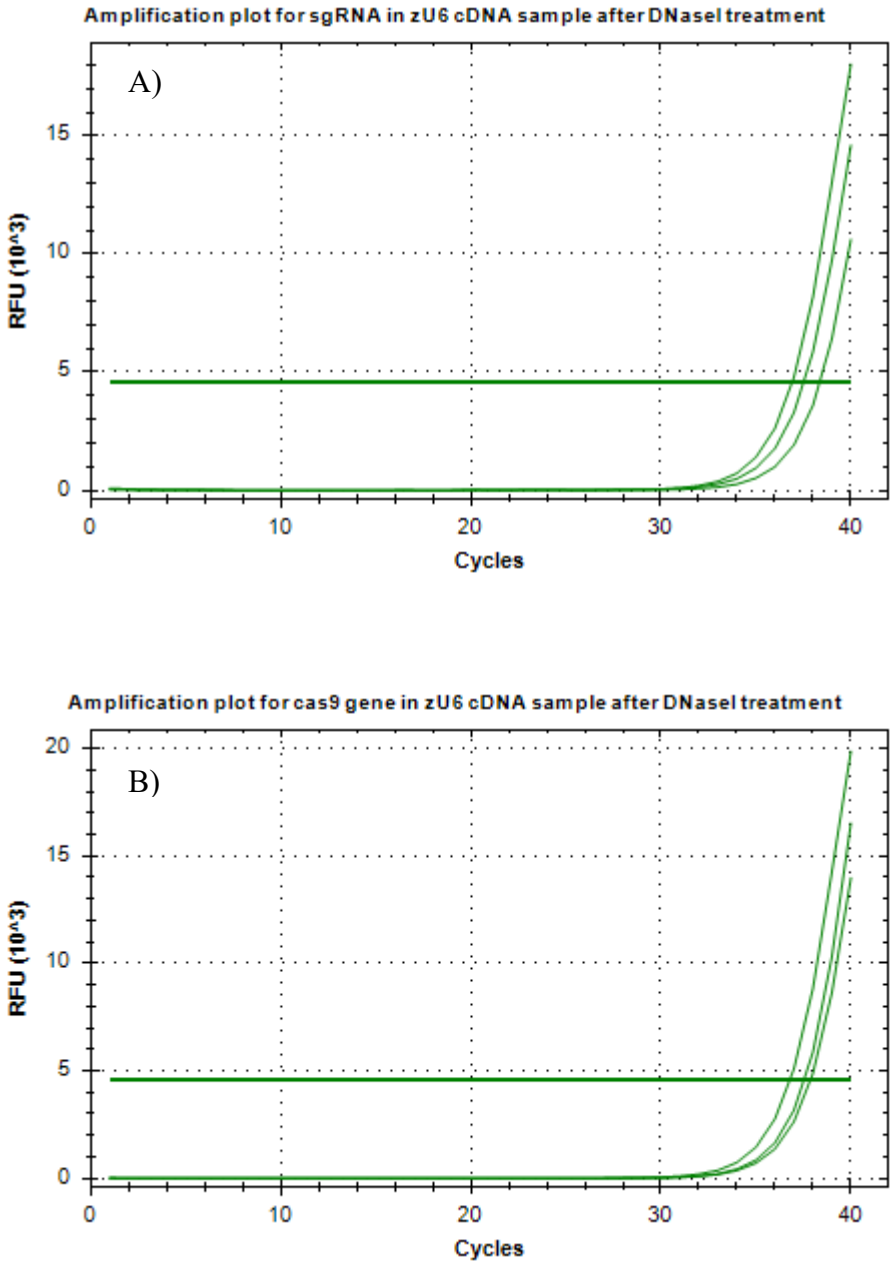

Figure 21: Amplification plot for sgRNA(A) and Cas9 (B) gene expression in dzU6 cDNA sample after DNase I treatment.

4 Discussion

The main goal of this experiment was to test and compare the efficiency of different U6 promoters using sgRNA expression assay. Unfortunately, I could not reach the main aim during this master thesis, because the time was used on trial and errors and set-up of new methods in our lab. Another major reason behind this was Covid-19 restriction to lab and limitation in availability of different lab resources for the experiments. Nevertheless, this thesis involved a lot of method development and optimization which has brought us much closer to our goal. Different aspects of the method optimization and suggestions for improvements will be discussed in this chapter.

4.1 Plasmid design

The first step of this experiment was to clone different U6 promoters and gfp_ko gRNA into lentiCRISPR vector. The success of cloning was confirmed via Sanger sequencing of the plasmids, as discussed in section 3.1 and Figure 12-13. As mentioned in section 3.5.2, Figure 20 and 21, the qPCR assay after DNase I treatment of the samples showed the successful expression of sgRNA under zebrafish U6 promoter. This shows the overall success of our plasmid design, transfection protocol, cDNA synthesis and primer validation and qPCR assay.

4.2 Cell culture and cell line selection

The SHK-1 cell line that we used in this study is a slow growing cell line and takes around 10-14 days to reach confluency (Dannevig et al., 1997; Dannevig et al., 1995). It took time to accumulate the desired number of cells for transfection. However, we were able to solve this problem later by changing the media composition from 5% FBS to 10% FBS, which improved the growth of the cells. Besides, we faced considerable contamination problems during this master's project, and had to discard all the cells and revive cells twice during the experiment. Due to Covid-19 restriction to access the cell lab, I was unable to observe cells daily and change media whenever necessary. I had to subculture cells in same flask for more than 3-4 times due to shortage of T-75 flasks and other necessary accessories during pandemic. All this led to contamination.

Studies have shown that SHK-1 cell line is difficult to maintain compared to another Atlantic salmon head kidney cell line, ASK cell line. In contrast to SHK-1 cell line, ASK cells are fast growing, easy to maintain and more adaptable to standard cell culture routines (Rolland et al., 2005; Rolland et al., 2003). ASK cell line was also developed to isolate ISA virus by Devold et al. (2000) and has been used to study immune response against ISA virus (Andresen et al., 2020), salmonid alphavirus (SAV) (Munir et al., 2020), *Renibacterium salmoninarum* (Bethke et al., 2019), cellular pathway in response to ISA virus (Jørgensen et al., 2006) and bisphenol-a toxicity response (Yazdani et al., 2016). This shows ASK cell line is equally competent and could be a better choice for this experiment in future as well as for overall CRISPR screening experiment which demands large number of cells.

4.3 qPCR assay

One of the major source of errors in qPCR is nucleic acid contamination (Taylor et al., 2019). In a qPCR where the RNA is extracted from cells or tissue, genomic DNA is also extracted and needs to be removed before the cDNA synthesis. As a safety, the qPCR primers should be placed such as they are separated by a large intron, or placed directly at the exon/exon border, so that the reaction can only take place at the cDNA strands, not the genomic DNA (Ginzinger, 2002). However, in this case, where we were doing qPCR on RNA from intron-less genes in a plasmid, DNase I treatment of RNA sample is even more crucial to avoid DNA contamination (Ginzinger, 2002).

In the initial experiment, we trusted the RNA isolation kit and cDNA synthesis kit, which claimed to wipe out genomic DNA. In addition, NanoDrop and TapeStation do not show whether the RNA sample is free from nucleic acid contamination or not, the only way is to do qPCR on the RNA samples before the cDNA synthesis (Taylor et al., 2019). When we did this, we could see that our qPCR results showed high DNA contamination in our RNA samples (refer to section 3.5.1 for Ct values). The source of DNA contamination could be the residual of plasmids that were still present inside the transfected cells which came along RNA during RNA isolation process. Further, we observed that almost everything we thought was mRNA expression observed prior to the DNase I treatment was due to the DNA contamination, as there was very low level or no expression after the samples were treated with DNase I. The optimal way to remove genomic DNA contamination is to treat the RNA sample with DNase

I followed by phenol extraction before cDNA synthesis (Naderi et al., 2004), and these steps should be included in the future protocol.

However, after DNase I treatment, we succeeded to remove DNA contamination and observed no or very low level of sgRNA and Cas9 expression, with only zebrafish U6 showing sgRNA levels above the detection limit. The possible reasons for the presence of low level of our target RNAs could be low transfection efficiency, poor RNA quality and low RT efficiency for cDNA synthesis. These reasons are discussed in detail below:

4.3.1 Transfection efficiency

As mentioned in section 2.4, the optimal electroporation protocol was determined visually based on proportion of EGFP expressing cells over non-EGFP expressing cells. The transfection efficiency and cell viability of lentiCRISPR plasmids was lower than that of pEGFP-Puro plasmid, as confluency of cells transfected with lentiCRISPR plasmids was lower compared to cells transfected with pEGFP-Puro, as mentioned in section 3.2. In transfection using electrical pulse, the transfection efficiency and cell viability after transfection are highly affected by physical size of plasmid being used, larger plasmid size result in lower transfection efficiency and cell viability (Lesueur et al., 2016). The size of pEGFP- Puro plasmid is 5 kbp whereas, the size of lentiCRISPR vector is big and our final plasmids became around 13 kbp after cloning, which is 2.5 times bigger than pEGFP- Puro. This could result in lower transfection efficiency, lower cell viability and eventually lower transgene expression in cells with lentiCRISPR vector compared to cells with EGFP.

We should consider reoptimizing transfection protocol for a large plasmid size, as Lesueur et al. (2016) find that recovery time up to 45 min after electroporation has increased the transfection efficiency and cell survival rate that is similar to small plasmid. Or even better if we could have created a transgenic Atlantic salmon cell line that stably express Cas9 and EGFP and transfect cells with a smaller plasmid without Cas9, then both the cell line and the plasmid could be used for CRISPR screening experiment in future. This kind of cell line has been already developed in Chinook salmon cell line (CHSE-EC) (Dehler et al., 2016) and has been used to study antiviral immunity (Dehler et al., 2019).

4.3.2 Quality and quantity of RNA

Another possible reason for low initial RNA level is poor RNA quality (Taylor et al., 2019). The NanoDrop data showed that, among the 4 samples, zU6 RNA sample was highest in concentration and lowest in contamination level, whereas mU6 RNA sample was lowest in concentration and highest in contamination level (see Table 3). Likewise, we observed low level of sgRNA and Cas9 expression in dzU6 sample whereas no expression was seen in case of dmU6 sample. This shows the purity and concentration of RNA samples are important to produce significant qPCR data.

We observed that OD 260/230 was very low which means the sample is contaminated with organic compounds (see Table 3). The possible contaminant could be guanidine salt that is present in Qiazol and RWT buffer which came along with sample during organic phase extraction. Such contaminants could affect downstream RT-qPCR application by interfering with primer annealing and polymerase efficiency making data generated unreliable (Taylor et al., 2019). However, Qiagen claims that concentration of guanidine salt up to 100mM in RNA sample doesn't compromise the reliability of RT-qPCR (Qiagen, 2010). Beside this, the quality and integrity of RNA sample was good according to the ScreenTape analysis (see Figure 15 A-B).

The concentration of total RNA isolated from cells transfected with lentiCRISPR was low. This means the amount of our target mRNA was even lower which was supported by the low Sq values that we observed in DNase I treated samples, (refer to the section 3.5.2). The lower concentration of target mRNA (< 10 copies/ul or < 100 copies per reaction) could lead to both systemic error via stochastic amplification and measurement uncertainty and random error via subsampling error (which we saw in our result after DNase I treatment, refer to Figure 21 and raw qPCR data for DNase I treated samples in Appendix II), hence increasing the gross error (Taylor et al., 2019). To minimize the error and data variability, we could have included larger volume of sample and increased the number of technical replicates. More importantly, we should consider transfecting more cells than we did in this experiment to increase the amount of target RNAs in the samples and isolate total RNA more carefully to produce high quality RNA samples.

4.3.3 cDNA synthesis

Another important possible reason for having low template level in cDNA sample is low RT efficiency of target RNA due to either contamination present (discussed earlier) or priming strategy. As mentioned in section 1.6.1, cDNA synthesis has been shown to be a crucial step in qPCR, with the possibility of introducing substantial variability. It is very important to apply same priming strategy and reaction conditions in order to normalize sgRNA expression with Cas9 expression (Stahlberg et al., 2004). The primer mix that we used for cDNA synthesis contained mixture of random hexamer and oligo-dT. Our target sgRNA doesn't contain any polyA tail at the end, whereas Cas9 contains polyA tail. However, Cas9 is transcribed together with nuclear localization signal (NLS) and Puromycin, and cleaved after translation. The polyA tail is located 3.6 kbp far away from our target Cas9 amplicon. If the distance between target amplicon of mRNA and polyA tail is very long and has RNA secondary structure, it may fail to synthesize the target amplicon. However, if oligo-dT primer succeed to synthesise Cas9 cDNA (which we don't know), this will generate a biased result that are not comparable.

In addition, the amount of target mRNA compared to background nonspecific nucleic acid can have major impact on RT efficiency of target mRNA (Curry et al., 2002; Karrer et al., 1995; Stahlberg et al., 2004a; Taylor et al., 2019). The cDNA synthesis of nonspecific mRNA will reduce the efficiency of our target mRNA even more since we are using primer mix. The use of primer combination is a drawback in this case where the concentration of template RNA is low and more importantly only one of our target gene has polyA tail. We should consider using a kit that uses random hexamers as primer for cDNA synthesis, as this will produce the least bias in cDNA synthesis and treat all the mRNAs same compared to oligodT and sequence specific primers (Ginzinger, 2002).

4.4 Optimization and validation of qPCR

In this experiment, optimal primer set, and its annealing temperature was selected based on melting curve analysis that showed one sharp peak and amplification plot that showed lowest Ct value (refer to Figure 16). In addition, we observed the efficiency of the PCR reaction for sgRNA and Cas9 standards were high enough and the data generated was linear (see Figure 17 A-B). This shows that the standards we prepared were valid and reaction efficiency and primer

annealing temperature are optimal. Primer validation is a way to assess optimal annealing temperature and dilution factor in order to get accurate Ct value that represent template amount in the sample regardless of contaminants present (Taylor et al., 2019). However, dilution of cDNA, 1:5, was done randomly without optimization. Appropriate dilution factor is essential to ensure minimal presence of contaminants that could affect primer efficiency and hence reliability and accuracy of data. In case of this experiment, where we saw our sample was contaminated with organic compounds, optimization of dilution factor could have given us more reliable data. We should consider optimization of dilution factor in our future experiments to produce biologically significant data.

4.5 qPCR data analysis

Despite that these data are unreliable because of the presence of DNA contamination in the RNA samples, I chose to analyze them because of the lack of valid data. In this way, I could go through the pipeline for qPCR data analysis and optimize how we should set up and analyze the qPCR experiment in the future. I am showing the results as an example on how it could look like when we get data from RNA samples with high quality and quantity.

Cas9 gene was used as normalization gene because of these reasons: first, the target gene and normalization gene are present in same plasmid to be transfected into the cells. This will normalize the gene of interest only to the gene that is expressed in transfected cells. Second, Cas9 gene has been exposed to same experimental condition as target gene form the beginning of the experiment and both genes are transiently expressed in the cells. Thus, normalization using Cas9 will significantly reduce the standard deviation within biological replicates as it will even out any differences in transfection efficiencies and will generate reproducible, valid and interpretable data (Jiwaji et al., 2010). Relative quantification of sgRNA normalized with Cas9 showed higher expression of sgRNA by salmon U6 promoter followed by zebrafish, mouse and human. Since both genes are exposed to same experimental condition throughout the experiment, normalization to Cas9 will even out the bias created due to presence of contamination, different RT efficiencies and different transfection efficiencies.

Absolute quantification gives the exact copies of mRNA present in sample and the data generated are independent of any normalizing gene. However, absolute quantification is

dependent on set of external standards, and results are highly reproducible with same set of standard and varies if different set of standards are used. In addition, errors can be easily introduced during measurement of standard and hence the exact number of copies present in standard is questionable (Ginzinger, 2002). Similarly, other source of errors like errors in experimental design, contaminants, and inhibitors present, RT efficiency, sampling and handling errors are not considered which could be present in sample but not in the standard. Hence, the result may not reflect the exact amount of template present in sample.

In contrast, relative quantification mRNA level of target gene is normalized to mRNA level of reference gene in the same sample, provided that all the samples are treated the same way. The normalized ratio is then compared between different samples. Normalization will compensate for any bias created due to differences in starting material amount, RNA isolation process, quality and purity of RNA, RT efficiencies etc that lead to variation in data generated and hence more reliable results are produced (Bustin et al., 2005; Ginzinger, 2002). However, the relative quantification method we used in this experiment may not consider these differences as the relative quantification calculation was based on copy number of the genes which was calculated based on external standard that were free from the inhibitors and contaminants, possibly present in sample. To address this, we should maybe consider using RNA standard, as this will consider different RT efficiencies (described in section 1.6.2.1). Even better, if we consider preparing standards from the pool of cDNA and apply relative quantification method (described in section 1.6.2.2) which will even out most of the sources of errors mentioned above.

5 Concluding remarks

This experiment was an initial attempt to find the most efficient U6 promoter for the expression of sgRNA in Atlantic salmon head kidney cell line, SHK-1 cell line. Since this has never been done before, and the methods used was new in our lab, this thesis involved a lot of method development and optimization. Due to technical difficulties and Covid-19 restrictions, I had only time to perform one single biological replicate in the SHK-1 cell line with the partially optimized protocol. The experiment will be further carried out by my co-supervisor PhD fellow Noman Reza. Although I was unable to reach the main goal of the thesis, the methods that I developed and optimized has brought us much closer towards our goal. I was able to demonstrate that one of the plasmids that we designed was able to express sgRNA under zebrafish U6 promoter. Although the expression level was low, it was a huge success for us as it showed overall success of plasmid design, transfection of cells, primer validation, and qPCR assay optimization. One of the major findings in this thesis is that it is crucial to digest DNA in the sample before qPCR, since we use a large amount of DNA for transfection. In addition, it is important to have high quality RNA samples in higher quantity to produce significant qPCR data. The methods optimized in this thesis can be used to assess the efficiency of multiple U6 promoters in different cell lines in the future.

6 References

Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nature communications*, *9*(1), 1-13.

Ahmed, F. (2005). Quantitative Real-time RT-PCR: Application to Carcinogenesis. *Cancer genomics & proteomics*, *2 6*, 317-332.

Andresen, A. M. S., Boudinot, P., & Gjøen, T. (2020). Kinetics of transcriptional response against poly (I:C) and infectious salmon anemia virus (ISAV) in Atlantic salmon kidney (ASK) cell line. *Developmental & Comparative Immunology*, *110*, 103716. https://www.sciencedirect.com/science/article/pii/S0145305X19306470

Arimbasseri, A. G., & Maraia, R. J. (2016). RNA polymerase III advances: structural and tRNA functional views. *Trends in biochemical sciences*, *41*(6), 546-559.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709-1712.

Bethke, J., Arias-Muñoz, E., Yáñez, A., & Avendaño-Herrera, R. (2019). Renibacterium salmoninarum iron-acquisition mechanisms and ASK cell line infection: Virulence and immune response. *Journal of fish diseases*, *42*(9), 1283-1291.

Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., & Marraffini, L. A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic acids research*, *41*(15), 7429-7437.

Boeuf, P., Vigan-Womas, I., Jublot, D., Loizon, S., Barale, J.-C., Akanmori, B. D., Mercereau-Puijalon, O., & Behr, C. (2005). CyProQuant-PCR: a real time RT-PCR technique for profiling human cytokines, based on external RNA standards, readily automatable for clinical use. *BMC Immunology*, *6*(1), 5. https://doi.org/10.1186/1471-2172-6-5

Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V., & Van Der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, *321*(5891), 960-964.

Bustin, S. (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology*, *25*(2), 169-193. https://doi.org/10.1677/jme.0.0250169

Bustin, S. A., Benes, V., Nolan, T., & Pfaffl, M. W. (2005). Quantitative real-time RT-PCR – a perspective. *Journal of Molecular Endocrinology*, *34*(3), 597-601. https://doi.org/10.1677/jme.1.01755

Bustin, S. A., & Nolan, T. (2004). Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal of biomolecular techniques: JBT*, *15*(3), 155.

Chakrapani, V., Patra, S. K., Panda, R. P., Rasal, K. D., Jayasankar, P., & Barman, H. K. (2016). Establishing targeted carp TLR22 gene disruption via homologous recombination using CRISPR/Cas9. *Developmental & Comparative Immunology*, *61*, 242-247.

Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G. W., Park, J., Blackburn, E. H., Weissman, J. S., Qi, L. S., & Huang, B. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, *155*(7), 1479-1491. https://doi.org/10.1016/j.cell.2013.12.001

Chen, C., Smye, S. W., Robinson, M. P., & Evans, J. A. (2006). Membrane electroporation theories: a review. *Medical and Biological Engineering and Computing*, *44*(1), 5-14. https://doi.org/10.1007/s11517-005-0020-2

Chen, H., Choi, J., & Bailey, S. (2014). Cut site selection by the two nuclease domains of the Cas9 RNA-guided endonuclease. *Journal of Biological Chemistry*, *289*(19), 13284-13294.

Clarke, B. D., Cummins, D. M., McColl, K. A., Ward, A. C., & Doran, T. J. (2013). Characterization of zebrafish polymerase III promoters for the expression of short-hairpin RNA interference molecules. *Zebrafish*, *10*(4), 472-479.

Cleveland, B. M., Yamaguchi, G., Radler, L. M., & Shimizu, M. (2018). Editing the duplicated insulin-like growth factor binding protein-2b gene in rainbow trout (Oncorhynchus mykiss). *Scientific reports*, *8*(1), 1-13.

Collet, B., & Lester, K. (2011). Establishment of an Atlantic salmon kidney cell line with an inducible gene expression system. *Journal of Biotechnology*, *154*(4), 209-211. https://www.sciencedirect.com/science/article/pii/S0168165611002690

Curry, J., McHale, C., & Smith, M. T. (2002). Low efficiency of the Moloney murine leukemia virus reverse transcriptase during reverse transcription of rare t (8; 21) fusion gene transcripts. *Biotechniques*, *32*(4), 768-775.

Dannevig, B. H., Brudeseth, B. E., Gjøen, T., Rode, M., Wergeland, H. I., Evensen, Ø., & Press, C. M. (1997). Characterisation of a long-term cell line (SHK-1) developed from

the head kidney of Atlantic salmon (Salmo salarL.). *Fish & shellfish immunology*, *7*(4), 213-226. https://www.sciencedirect.com/science/article/pii/S1050464896900765

Dannevig, B. H., Falk, K., & Namork, E. (1995). Isolation of the causal virus of infectious salmon anaemia (ISA) in a long-term cell line from Atlantic salmon head kidney. *Journal of General Virology*, *76*(6), 1353-1359.

Datsomor, A. K., Zic, N., Li, K., Olsen, R. E., Jin, Y., Vik, J. O., Edvardsen, R. B., Grammes, F., Wargelius, A., & Winge, P. (2019). CRISPR/Cas9-mediated ablation of elovl2 in Atlantic salmon (Salmo salar L.) inhibits elongation of polyunsaturated fatty acids and induces Srebp-1 and target genes. *Scientific reports*, *9*(1), 1-13.

Dehler, C. E., Boudinot, P., Martin, S. A. M., & Collet, B. (2016). Development of an Efficient Genome Editing Method by CRISPR/Cas9 in a Fish Cell Line. *Marine Biotechnology*, *18*(4), 449-452. https://doi.org/10.1007/s10126-016-9708-6

Dehler, C. E., Lester, K., Della Pelle, G., Jouneau, L., Houel, A., Collins, C., Dovgan, T., Machat, R., Zou, J., & Boudinot, P. (2019). Viral resistance and IFN signaling in STAT2 knockout fish cells. *The Journal of Immunology*, *203*(2), 465-475.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602-607.

Deprez, R. H. L., Fijnvandraat, A. C., Ruijter, J. M., & Moorman, A. F. (2002). Sensitivity and accuracy of quantitative real-time polymerase chain reaction using SYBR green I depends on cDNA synthesis conditions. *Analytical biochemistry*, *307*(1), 63-69.

Devold, M., Krossøy, B., Aspehaug, V., & Nylund, A. (2000). Use of RT-PCR for diagnosis of infectious salmon anaemia virus (ISAV) in carrier sea trout Salmo trutta after experimental infection. *Diseases of aquatic organisms*, *40*(1), 9-18.

Dhanasekaran, S., Doherty, T. M., & Kenneth, J. (2010). Comparison of different standards for real-time PCR-based absolute quantification. *Journal of Immunological Methods*, *354*(1), 34-39. https://www.sciencedirect.com/science/article/pii/S0022175910000062

Díaz, R., Troncoso, J., Jakob, E., & Skugor, S. (2021). "Limiting access to iron decreases infection of Atlantic salmon SHK-1 cells with bacterium Piscirickettsia salmonis". *BMC Veterinary Research*, *17*(1), 155. https://doi.org/10.1186/s12917-021-02853-6

Didychuk, A. L., Butcher, S. E., & Brow, D. A. (2018). The life of U6 small nuclear RNA, from cradle to grave. *RNA*, *24*(4), 437-460.

Doench, J. G. (2018). Am I ready for CRISPR? A user's guide to genetic screens. *Nature Reviews Genetics*, *19*(2), 67.

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B. L., Xavier, R. J., & Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nature biotechnology*, *32*(12), 1262-1267.

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213).

Edvardsen, R. B., Leininger, S., Kleppe, L., Skaftnesmo, K. O., & Wargelius, A. (2014). Targeted mutagenesis in Atlantic salmon (Salmo salar L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation. *PloS one*, *9*(9), e108622.

Elaswad, A., Khalil, K., Ye, Z., Liu, Z., Liu, S., Peatman, E., Odin, R., Vo, K., Drescher, D., & Gosh, K. (2018). Effects of CRISPR/Cas9 dosage on TICAM1 and RBL gene mutation rate, embryonic development, hatchability and fry survival in channel catfish. *Scientific reports*, *8*(1), 1-17.

Escobar-Aguirre, S., Arancibia, D., Escorza, A., Bravo, C., Andrés, M. E., Zamorano, P., & Martínez, V. (2019). Development of a Bicistronic Vector for the Expression of a CRISPR/Cas9-mCherry System in Fish Cell Lines. *Cells*, *8*(1), 75. https://www.mdpi.com/2073-4409/8/1/75

Fajrial, A. K., He, Q. Q., Wirusanti, N. I., Slansky, J. E., & Ding, X. (2020). A review of emerging physical transfection methods for CRISPR/Cas9-mediated gene editing. *Theranostics*, *10*(12), 5532-5549. https://doi.org/10.7150/thno.43465

FAO, R. (2018). FAO yearbook: Fishery and aquaculture statistics, 2016.

Feng, Y., Xie, Z., Jiang, X., Li, Z., Shen, Y., Wang, B., & Liu, J. (2018). The Applications of Promoter-gene-Engineered Biosensors. *Sensors (Basel)*, *18*(9). https://doi.org/10.3390/s18092823

Gao, Z., Herrera-Carrillo, E., & Berkhout, B. (2018). RNA polymerase II activity of type 3 Pol III promoters. *Molecular Therapy-Nucleic Acids*, *12*, 135-145.

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*, *109*(39), E2579-E2586.

Genzel, Y. (2015). Designing cell lines for viral vaccine production: where do we stand? *Biotechnology journal*, *10*(5), 728-740.

Gibson, U. E., Heid, C. A., & Williams, P. M. (1996). A novel method for real time quantitative RT-PCR. *Genome research*, *6*(10), 995-1001.

Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., Guimaraes, C., Panning, B., Ploegh, H. L., Bassik, M. C., Qi, L. S., Kampmann, M., & Weissman, J. S. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, *159*(3), 647-661. https://doi.org/10.1016/j.cell.2014.09.029

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*, *154*(2), 442-451. https://doi.org/10.1016/j.cell.2013.06.044

Ginzinger, D. G. (2002). Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. *Experimental hematology*, *30*(6), 503-512. https://www.sciencedirect.com/science/article/pii/S0301472X02008068

Gratacap, R. L., Jin, Y. H., Mantsopoulou, M., & Houston, R. D. (2020a). Efficient Genome Editing in Multiple Salmonid Cell Lines Using Ribonucleoprotein Complexes. *Marine Biotechnology*, *22*(5), 717-724. https://doi.org/10.1007/s10126-020-09995-y

Gratacap, R. L., Regan, T., Dehler, C. E., Martin, S. A. M., Boudinot, P., Collet, B., & Houston, R. D. (2020). Efficient CRISPR/Cas9 genome editing in a salmonid fish cell line using a lentivirus delivery system. *BMC Biotechnology*, *20*(1), 35. https://doi.org/10.1186/s12896-020-00626-x

Gratacap, R. L., Wargelius, A., Edvardsen, R. B., & Houston, R. D. (2019). Potential of genome editing to improve aquaculture breeding and production. *Trends in Genetics*, *35*(9), 672-684.

Hamar, J., & Kültz, D. (2021). An efficient vector-based CRISPR/Cas9 system in an Oreochromis mossambicus cell line using endogenous promoters. *Scientific reports*, *11*(1), 1-17.

Heid, C. A., Stevens, J., Livak, K. J., & Williams, P. M. (1996). Real time quantitative PCR. *Genome research*, *6*(10), 986-994.

Helbo, A. S., Lay, F. D., Jones, P. A., Liang, G., & Grønbæk, K. (2017). Nucleosome positioning and NDR structure at RNA polymerase III promoters. *Scientific reports*, *7*(1), 1-11.

Herath, T., Costa, J. Z., Thompson, K., Adams, A., & Richards, R. (2009). Alternative cell line for the isolation of salmonid alphavirus-1. *Icelandic Agricultural Sciences*, *22*, 19-27.

Houston, R. D. (2017). Future directions in breeding for disease resistance in aquaculture species. *Revista Brasileira de Zootecnia*, *46*(6), 545-551.

Houston, R. D., Bean, T. P., Macqueen, D. J., Gundappa, M. K., Jin, Y. H., Jenkins, T. L., Selly, S. L. C., Martin, S. A., Stevens, J. R., & Santos, E. M. (2020). Harnessing genomics to fast-track genetic improvement in aquaculture. *Nature Reviews Genetics*, *21*(7), 389-409.

Huang, Y., & Maraia, R. J. (2001). Comparison of the RNA polymerase III transcription machinery in Schizosaccharomyces pombe, Saccharomyces cerevisiae and human. *Nucleic Acids Res*, *29*(13), 2675-2690. https://doi.org/10.1093/nar/29.13.2675

Huggett, J., Dheda, K., Bustin, S., & Zumla, A. (2005). Real-time RT-PCR normalisation; strategies and considerations. *Genes & Immunity*, *6*(4), 279-284.

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*, *110*(34), 13904-13909.

Jiang, F., & Doudna, J. A. (2017). CRISPR–Cas9 structures and mechanisms. *Annual review of biophysics*, *46*, 505-529.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816-821.

Jiwaji, M., Daly, R., Pansare, K., McLean, P., Yang, J., Kolch, W., & Pitt, A. R. (2010). The Renilla luciferase gene as a reference gene for normalization of gene expression in transiently transfected cells. *BMC Molecular Biology*, *11*(1), 103. https://doi.org/10.1186/1471-2199-11-103

Jobling, M. (2011). JS Lucas and PC Southgate (eds): Aquaculture—Farming aquatic animals and plants. In (pp. 214-228). J. Wiley & Sons.

Jørgensen, S. M., Lyng-Syvertsen, B., Lukacs, M., Grimholt, U., & Gjøen, T. (2006). Expression of MHC class I pathway genes in response to infectious salmon anaemia virus in Atlantic salmon (Salmo salar L.) cells. *Fish & shellfish immunology*, *21*(5), 548-560.

Joung, J., Konermann, S., Gootenberg, J. S., Abudayyeh, O. O., Platt, R. J., Brigham, M. D., Sanjana, N. E., & Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature Protocols*, *12*(4), 828-863. https://doi.org/10.1038/nprot.2017.016

Karrer, E. E., Lincoln, J. E., Hogenhout, S., Bennett, A. B., Bostock, R. M., Martineau, B., Lucas, W. J., Gilchrist, D. G., & Alexander, D. (1995). In situ isolation of mRNA from individual plant cells: creation of cell-specific cDNA libraries. *Proceedings of the National Academy of Sciences*, *92*(9), 3814-3818.

Khatter, H., Vorländer, M. K., & Müller, C. W. (2017). RNA polymerase I and III: similar yet unique. *Curr Opin Struct Biol*, *47*, 88-94. https://doi.org/10.1016/j.sbi.2017.05.008

Kim, Y. B., Komor, A. C., Levy, J. M., Packer, M. S., Zhao, K. T., & Liu, D. R. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat Biotechnol*, *35*(4), 371-376. https://doi.org/10.1038/nbt.3803

Kishimoto, K., Washio, Y., Yoshiura, Y., Toyoda, A., Ueno, T., Fukuyama, H., Kato, K., & Kinoshita, M. (2018). Production of a breed of red sea bream Pagrus major with an increase of skeletal muscle mass and reduced body length by genome editing with CRISPR/Cas9. *Aquaculture*, *495*, 415-427.

Konermann, S., Brigham, M. D., Trevino, A. E., Hsu, P. D., Heidenreich, M., Cong, L., Platt, R. J., Scott, D. A., Church, G. M., & Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, *500*(7463), 472-476.

Koppang, E. O., Dannevig, B. H., Lie, Ø., RØNningen, K., & Press, C. M. (1999). Expression of Mhc class I and II mRNA in a macrophage-like cell line (SHK-1) derived from Atlantic salmon, Salmo salar L., head kidney. *Fish & shellfish immunology*, *9*(6), 473-489. https://www.sciencedirect.com/science/article/pii/S1050464899902139

Kunkel, G. R., Cheung, T. C., Miyake, J. H., Urso, O., McNamara-Schroeder, K. J., & Stumph, W. E. (1996). Identification of a SPH element in the distal region of a human U6 small nuclear RNA gene promoter and characterization of the SPH binding factor in HeLa cell extracts. *Gene Expr*, *6*(2), 59-72.

Kunkel, G. R., & Hixson, J. D. (1998). The distal elements, OCT and SPH, stimulate the formation of preinitiation complexes on a human U6 snRNA gene promoter in vitro. *Nucleic Acids Res*, *26*(6), 1536-1543. https://doi.org/10.1093/nar/26.6.1536

Lakra, W., Swaminathan, T. R., & Joy, K. (2011). Development, characterization, conservation and storage of fish cell lines: a review. *Fish physiology and biochemistry*, *37*(1), 1-20.

Lee, P. T., Zou, J., Holland, J. W., Martin, S. A. M., Collet, B., Kanellos, T., & Secombes, C. J. (2014). Identification and characterisation of TLR18-21 genes in Atlantic salmon (Salmo salar). *Fish & shellfish immunology*, *41*(2), 549-559. https://www.sciencedirect.com/science/article/pii/S1050464814003805

Lee, P. T., Zou, J., Holland, J. W., Martin, S. A. M., Kanellos, T., & Secombes, C. J. (2013). Identification and characterization of TLR7, TLR8a2, TLR8b1 and TLR8b2 genes in Atlantic salmon (Salmo salar). *Developmental & Comparative Immunology*, *41*(2), 295-305. https://www.sciencedirect.com/science/article/pii/S0145305X13001407

Lee, P. T., Zou, J., Holland, J. W., Martin, S. A. M., Scott, C. J. W., Kanellos, T., & Secombes, C. J. (2015). Functional characterisation of a TLR accessory protein, UNC93B1, in Atlantic salmon (Salmo salar). *Developmental & Comparative Immunology*, *50*(1), 38-48. https://www.sciencedirect.com/science/article/pii/S0145305X14003231

Leong, D. T., Gupta, A., Bai, H. F., Wan, G., Yoong, L. F., Too, H.-P., Chew, F. T., & Hutmacher, D. W. (2007). Absolute quantification of gene expression in biomaterials research using real-time PCR. *Biomaterials*, *28*(2), 203-210. https://www.sciencedirect.com/science/article/pii/S0142961206007836

Lesueur, L. L., Mir, L. M., & André, F. M. (2016). Overcoming the specific toxicity of large plasmids electrotransfer in primary cells in vitro. *Molecular Therapy-Nucleic Acids*, *5*, e291.

Li, M., Yang, H., Zhao, J., Fang, L., Shi, H., Li, M., Sun, Y., Zhang, X., Jiang, D., & Zhou, L. (2014). Efficient and heritable gene targeting in tilapia by CRISPR/Cas9. *Genetics*, *197*(2), 591-599.

Li, Y., Ma, S., Sun, L., Zhang, T., Chang, J., Lu, W., Chen, X., Liu, Y., Wang, X., Shi, R., Zhao, P., & Xia, Q. (2018). Programmable Single and Multiplex Base-Editing in Bombyx mori Using RNA-Guided Cytidine Deaminases. *G3 (Bethesda)*, *8*(5), 1701-1709. https://doi.org/10.1534/g3.118.200134

Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry*, *79*, 181-211.

Lino, C. A., Harper, J. C., Carney, J. P., & Timlin, J. A. (2018). Delivering CRISPR: a review of the challenges and approaches. *Drug Delivery*, *25*(1), 1234-1257. https://doi.org/10.1080/10717544.2018.1474964

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), bio035170.

Liu, Z., Liao, Z., Chen, Y., Han, L., Yin, Q., & Xiao, H. (2020). Application of Various Delivery Methods for CRISPR/dCas9. *Molecular Biotechnology*, *62*(8), 355-363. https://doi.org/10.1007/s12033-020-00258-8

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods*, *25*(4), 402-408.

Ma, J., Fan, Y., Zhou, Y., Liu, W., Jiang, N., Zhang, J., & Zeng, L. (2018). Efficient resistance to grass carp reovirus infection in JAM-A knockout cells using CRISPR/Cas9. *Fish & shellfish immunology*, *76*, 206-215.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., & Yakunin, A. F. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, *9*(6), 467-477.

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, *322*(5909), 1843-1845.

Merten, O.-W., Hebben, M., & Bovolenta, C. (2016). Production of lentiviral vectors. *Molecular Therapy-Methods & Clinical Development*, *3*, 16017.

Miyagishi, M., & Taira, K. (2002). U6 promoter–driven siRNAs with four uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. *Nature biotechnology*, *20*(5), 497-500.

Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, *155*(3), 733-740.

Mojica, F. J., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of molecular evolution*, *60*(2), 174-182.

Molloy, S., Thomas, E., Hoyt, K., & Bouchard, D. (2013). Enhanced detection of infectious salmon anaemia virus using a low-speed centrifugation technique in three fish cell lines. *Journal of fish diseases*, *36*(1), 35-44.

Munir, D., Munro, E. S., Secombes, C. J., & Dooley, H. (2020). Atlantic salmon kidney (ASK) cells are an effective model to characterise interferon (IFN) and IFN-induced gene expression following salmonid alphavirus infection. *Fish Shellfish Immunol*, *106*, 792-795. https://doi.org/10.1016/j.fsi.2020.08.043

Naderi, A., Ahmed, A. A., Barbosa-Morais, N. L., Aparicio, S., Brenton, J. D., & Caldas, C. (2004). Expression microarray reproducibility is improved by optimising purification steps in RNA amplification and labelling. *BMC Genomics*, *5*(1), 9. https://doi.org/10.1186/1471-2164-5-9

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K. Y., Shimatani, Z., & Kondo, A. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science*, *353*(6305). https://doi.org/10.1126/science.aaf8729

Oliver, C., Valenzuela, K., Silva, H., Haro, R., Cortés, M., Sandoval, R., Pontigo, J., Álvarez, C., Figueroa, J., & Avendaño-Herrera, R. (2015). Effectiveness of egg yolk immunoglobulin against the intracellular salmonid pathogen Piscirickettsia salmonis. *Journal of Applied Microbiology*, *119*(2), 365-376.

Opitz, H., Bouchard, D., Anderson, E., Blake, S., B.L, N., & Keleher, W. (1999). A Comparison of Methods for the Detection of Experimentally Induced Subclinical Infectious Salmon Anaemia in Atlantic Salmon. *Bull Eur Assoc Fish Pathol*, *20*.

Orchard, R. C., Wilen, C. B., Doench, J. G., Baldridge, M. T., McCune, B. T., Lee, Y.-C. J., Lee, S., Pruett-Miller, S. M., Nelson, C. A., & Fremont, D. H. (2016). Discovery of a proteinaceous cellular receptor for a norovirus. *Science*, *353*(6302), 933-936.

Pfaffl, M. W., & Hageleit, M. (2001). Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. *Biotechnology Letters*, *23*(4), 275-282.

Pfaffl, M. W., Horgan, G. W., & Dempfle, L. (2002). Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic acids research*, *30*(9), e36-e36.

Polz, M. F., & Cavanaugh, C. M. (1998). Bias in template-to-product ratios in multitemplate PCR. *Applied and environmental Microbiology*, *64*(10), 3724-3730.

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*(5), 1173-1183.

Qiagen. (2010). Qiagen Newsletter: Gene Expression and Function Studies. 8. https://www.qiagen.com/us/resources/faq?id=c59936fb-4f1e-4191-9c16-ff083cb24574&lang=en

Qin, Z., Li, Y., Su, B., Cheng, Q., Ye, Z., Perera, D. A., Fobes, M., Shang, M., & Dunham, R. A. (2016). Editing of the luteinizing hormone gene to sterilize channel catfish, Ictalurus punctatus, using a modified zinc finger nuclease technology with electroporation. *Marine Biotechnology*, *18*(2), 255-263.

Ranganathan, V., Wahlin, K., Maruotti, J., & Zack, D. J. (2014). Expansion of the CRISPR–Cas9 genome targeting space through the use of H1 promoter-expressed guide RNAs. *Nature communications*, *5*(1), 1-8.

Reyes-Cerpa, S., Reyes-López, F. E., Toro-Ascuy, D., Ibañez, J., Maisey, K., Sandino, A. M., & Imarai, M. (2012). IPNV modulation of pro and anti-inflammatory cytokine expression in Atlantic salmon might help the establishment of infection and persistence. *Fish & shellfish immunology*, *32*(2), 291-300. https://www.sciencedirect.com/science/article/pii/S1050464811004281

Roelz, R., Pilz, I. H., Mutschler, M., & Pahl, H. L. (2010). Of mice and men: human RNA polymerase III promoter U6 is more efficient than its murine homologue for shRNA expression from a lentiviral vector in both human and murine progenitor cells. *Experimental hematology*, *38*(9), 792-797.

Rolland, J. B., Bouchard, D., Coll, J., & Winton, J. R. (2005). Combined use of the ASK and SHK-1 cell lines to enhance the detection of infectious salmon anemia virus. *Journal of veterinary diagnostic investigation*, *17*(2), 151-157.

Rolland, J. B., Bouchard, D. A., & Winton, J. R. (2003). Improved diagnosis of infectious salmon anaemia virus by use of a new cell line derived from Atlantic salmon kidney tissue. Miller, Otis; Cipriano, Rocco C., tech. coords. International response to infectious salmon anemia: prevention, control, and eradication: proceedings of a symposium; 3–4 September 2002; New Orleans, LA. Tech. Bull. 1902. Washington, DC,

Sanjana, N. E., Shalem, O., & Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nature Methods*, *11*(8), 783-784. https://doi.org/10.1038/nmeth.3047

Santana, P. A., Jara-Gutiérrez, C., Mellado, M., Forero, J. C., Guzmán, F., Barriga, A., Albericio, F., & Álvarez, C. A. (2021). Effects of elderflower extract enriched with polyphenols on antioxidant defense of salmon leukocytes. *Electronic Journal of Biotechnology*, *52*, 13-20. https://www.sciencedirect.com/science/article/pii/S0717345821000191

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic acids research*, *39*(21), 9275-9282.

Savidis, G., McDougall, W. M., Meraner, P., Perreira, J. M., Portmann, J. M., Trincucci, G., John, S. P., Aker, A. M., Renzette, N., & Robbins, D. R. (2016). Identification of Zika virus and dengue virus dependency factors using functional genomics. *Cell reports*, *16*(1), 232-246.

Schramm, L., & Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. *Genes & development*, *16*(20), 2593-2620.

Schuster, M., Schweizer, G., Reissmann, S., & Kahmann, R. (2016). Genome editing in Ustilago maydis using the CRISPR–Cas system. *Fungal Genetics and Biology*, *89*, 3-9.

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., & Doench, J. G. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, *343*(6166), 84-87.

Shalem, O., Sanjana, N. E., & Zhang, F. (2015). High-throughput functional genomics using CRISPR–Cas9. *Nature Reviews Genetics*, *16*(5), 299-311. https://doi.org/10.1038/nrg3899

Shrock, E., & Güell, M. (2017). CRISPR in animals and animal models. *Progress in molecular biology and translational science*, *152*, 95-114.

Shukla, V., Coumoul, X., & Deng, C.-X. (2007). RNAi-based conditional gene knockdown in mice using a U6 promoter driven vector. *International journal of biological sciences*, *3*(2), 91.

Simpson, C. G., & Brown, J. W. (1995). Applications of RT-PCR. In *Plant Gene Transfer and Expression Protocols* (pp. 257-269). Springer.

Sivaganesan, M., Haugland, R. A., Chern, E. C., & Shanks, O. C. (2010). Improved strategies and optimization of calibration models for real-time PCR absolute quantification. *Water Research*, *44*(16), 4726-4735.

Stahlberg, A., Hakansson, J., Xian, X., Semb, H., & Kubista, M. (2004). Properties of the reverse transcription reaction in mRNA quantification. *Clinical chemistry*, *50*(3), 509-515.

Stahlberg, A., Kubista, M., & Pfaffl, M. (2004a). Comparison of reverse transcriptases in gene expression analysis. *Clinical chemistry*, *50*(9), 1678-1680.

Staller, E., Sheppard, C. M., Neasham, P. J., Mistry, B., Peacock, T. P., Goldhill, D. H., Long, J. S., & Barclay, W. S. (2019). ANP32 proteins are essential for influenza virus replication in human cells. *Journal of virology*, *93*(17), e00217-00219.

Tan, W. S., Rong, E., Dry, I., Lillico, S., Law, A., Whitelaw, B., & Dalziel, R. (2020). Genome-wide CRISPR knockout screen reveals membrane tethering complexes EARP and GARP important for Bovine Herpes Virus Type 1 replication. *bioRxiv*.

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., & Vale, R. D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, *159*(3), 635-646. https://doi.org/10.1016/j.cell.2014.09.039

Taylor, S., Wakem, M., Dijkman, G., Alsarraj, M., & Nguyen, M. (2010). A practical approach to RT-qPCR—publishing data that conform to the MIQE guidelines. *Methods*, *50*(4), S1-S5.

Taylor, S. C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M., & Fenrich, J. (2019). The Ultimate qPCR Experiment: Producing Publication Quality, Reproducible Data the First Time. *Trends in Biotechnology*, *37*(7), 761-774. https://www.sciencedirect.com/science/article/pii/S0167779918303421

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*, *3*(7), 1-12.

Wargelius, A., Leininger, S., Skaftnesmo, K. O., Kleppe, L., Andersson, E., Taranger, G. L., Schulz, R. W., & Edvardsen, R. B. (2016). Dnd knockout ablates germ cells and demonstrates germ cell independent sex differentiation in Atlantic salmon. *Scientific reports*, *6*(1), 1-8.

Wergeland, H. I., & Jakobsen, R. A. (2001). A salmonid cell line (TO) for production of infectious salmon anaemia virus (ISAV). *Diseases of aquatic organisms*, *44*(3), 183-190.

White, R. J. (2011). Transcription by RNA polymerase III: more complex than we thought. *Nature Reviews Genetics*, *12*(7), 459-463. https://doi.org/10.1038/nrg3001

Willis, I. M. (1994). Rna polymerase iii. *EJB Reviews 1993*, 29-39.

Wittwer, C. T., Herrmann, M. G., Gundry, C. N., & Elenitoba-Johnson, K. S. (2001). Real-time multiplex PCR assays. *Methods*, *25*(4), 430-442.

Wolf, K., & Quimby, M. (1962). Established eurythermic line of fish cells in vitro. *Science*, *135*(3508), 1065-1066.

Wolf, K., & Quimby, M. (1976). Primary monolayer culture of fish cells initiated from minced tissues. *TCA manual/Tissue Culture Association*, *2*(4), 445-448.

Wyman, C., & Kanaar, R. (2006). DNA double-strand break repair: all's well that ends well. *Annu. Rev. Genet.*, *40*, 363-383.

Xu, J., Zheng, Z., Du, X., Shi, B., Wang, J., Gao, D., Zhu, Q., Chen, X., & Han, J. (2020). A cytokine screen using CRISPR-Cas9 knock-in reporter pig iPS cells reveals that Activin A regulates NANOG. *Stem cell research & therapy*, *11*(1), 1-16.

Yáñez, J. M., Houston, R. D., & Newman, S. (2014). Genetics and genomics of disease resistance in salmonid species. *Frontiers in genetics*, *5*, 415.

Yazdani, M., Andresen, A. M. S., & Gjøen, T. (2016). Short-term effect of bisphenol-a on oxidative stress responses in Atlantic salmon kidney cell line: a transcriptional study. *Toxicology mechanisms and methods*, *26*(4), 295-300.

Yuan, J. S., Reed, A., Chen, F., & Stewart, C. N. (2006). Statistical analysis of real-time PCR data. *BMC Bioinformatics*, *7*(1), 85. https://doi.org/10.1186/1471-2105-7-85

Zenger, K. R., Khatkar, M. S., Jones, D. B., Khalilisamani, N., Jerry, D. R., & Raadsma, H. W. (2019). Genomic selection in aquaculture: application, limitations and opportunities with special reference to marine shrimp and pearl oysters. *Frontiers in genetics*, *9*, 693.

Zenke, K., & Kim, K. H. (2008). Novel fugu U6 promoter driven shRNA expression vector for efficient vector based RNAi in fish cell lines. *Biochemical and biophysical research communications*, *371*(3), 480-483.

Zhang, F. (2019). Development of CRISPR-Cas systems for genome editing and beyond. *Quarterly Reviews of Biophysics*, *52*.

Zhang, J., & Byrne, C. D. (1999). Differential priming of RNA templates during cDNA synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase PCR. *Biochemical Journal*, *337*(2), 231-241.

Zheng, X., Zheng, P., Sun, J., Kun, Z., & Ma, Y. (2018). Heterologous and endogenous U6 snRNA promoters enable CRISPR/Cas9 mediated genome editing in Aspergillus niger. *Fungal biology and biotechnology*, *5*(1), 1-9.

Zoppo, M., Okoniewski, N., Pantelyushin, S., vom Berg, J., & Schirmer, K. (2021). A ribonucleoprotein transfection strategy for CRISPR/Cas9-mediated gene editing and single cell cloning in rainbow trout cells. *Cell & Bioscience*, *11*(1), 103. https://doi.org/10.1186/s13578-021-00618-0

7 Appendix

Appendix I: Reagents, equipment, plasmids, and software

Reagents: product, catalogue number and manufacturer

1. 2-Mercapto-ethanol, 31350010, Gibco
2. AccuPrime™ Pfx DNA Polymerase kit, 12344024, Invitrogen
3. Certified low range ultra-agarose gel, 1613107, Bio-Rad
4. DNase I, RNase-free HC, EN0523, Thermofisher Scientific
5. Fetal Bovine Serum, 26400044, Gibco
6. Leibovitz's L-15 media, 11415064, Gibco
7. Neon™ Transfection System 10 µL Kit, MPK1025, Invitrogen
8. One shot Stbl3 competent *E. coli* cells, C737303, Thermofisher Scientific
9. Opti-MEM I Reduced Serum medium, 31985062, Gibco
10. PBS, pH 7.4, 10010023, Gibco
11. Penicillin-streptomycin, 15140122, Gibco
12. Platinum II Hot-Start Green PCR Master Mix (2X), 14001012, Invitrogen
13. Platinum II Taq Hot-Start DNA Polymerase Kit, 1496600, Invitrogen
14. Puromycin Dihydrochloride, A1113803, Gibco
15. QIAquick Gel Extraction kit, 28706X4, Qiagen
16. QIAzol Lysis reagent, 79306, Qiagen
17. QuantiTect Reverse Transcription Kit, 205311, Qiagen
18. RNeasy Plus Universal Kit, 73404, Qiagen
19. SOC medium, 15544034, Invitrogen
20. SYBR Select Master Mix, 4472908, Applied Biosystems
21. Trypsin-EDTA, 25200056, Gibco
22. Typhan Blue solution, 15250061, Gibco
23. ZR plasmid miniprep-classic kit, D4054, Zymo Research Crop
24. ZymoPURE II Plasmid Midiprep Kit, D4201, Zymo Research Crop

Equipment: product, catalogue no, manufacturer

1. Agilent 4150 TapeStation system, G2992AA, Agilent
2. CFX96 Touch Real-Time PCR detection system, Bio-Rad
3. Fluorescence microscope, 3849000909, Carl Zeiss
4. NanoDrop 8000 Spectrophotometer, ND-8000-GL, Thermofisher Scientific
5. Neon Transfection System, MPK5000, Invitrogen
6. TC-20 Automated cell counter, 1450102, Bio Rad

Plasmids

1. lentiCRISPR v2 - Addgene plasmid # 52961
2. pEGFP-Puro - Addgene plasmid # 45561
3. pSico plasmid - Addgene plasmid # 11578

Software

1. Agilent 4150 TapeStation system software
2. CFX Maestro software
3. NanoDrop 8000 V2.1.0 software

Free software

1. Benchling: https://www.benchling.com.
2. Oligo Calc: http://biotools.nubic.northwestern.edu/OligoCalc.html

Appendix II: Raw data from qPCR assay

Initial qPCR results

sample	Ct values	Sq values	Mean Sq value, n=3	Standard deviation, n=3	Mean U6/Cas9 Sq value, n=3	Standard deviation, n=3
hU6	27.17	8,78E+04	8,72E+04	5,09E+02	1,39E-01	0,008
	27.20	8,62E+04				
	27.17	8,76E+04				
zU6	27,43	7,24E+04	7,12E+04	4,55E+03	1,78E-01	0,014
	27,63	6,28E+04				
	27,32	7,85E+04				
mU6	28,36	3,67E+04	3,71E+04	4,63E+02	1,69E-01	0,010
	28,37	3,65E+04				
	28,32	3,80E+04				
sU6	29,10	2,14E+04	2,09E+04	5,20E+02	2,17E-01	0,017
	29,21	1,99E+04				
	29,10	2,15E+04				
hCas9	24.49	6,18E+05	6,31E+05	3,45E+04		
	24.33	6,96E+05				
	24.58	5,79E+05				
zCas9	25,16	3,81E+05	4,02E+05	1,11E+04		
	25,03	4,19E+05				
	25,07	4,05E+05				
mCas9	25,85	2,30E+05	2,21E+05	9,99E+03		
	25,84	2,32E+05				
	26,03	2,01E+05				
sCas9	27,09	9,27E+04	9,73E+04	5,44E+03		
	26,88	9,10E+04				
	27,12	9,10E+04				

Note: here, n is 3 technical qPCR replicates not biological replicate.

DNase I treated sample

sample	Ct values	Sq values	Mean Sq value, n=3	Standard deviation, n=3	Mean U6/Cas9 Sq value, n=3	Standard deviation, n=3
zU6	38,35	6,52	13,05	3,91	0,526	1,14
	37,50	12,61				
	36,90	20,04				
zCas9	36,77	22,12	14,68	3,81		
	36,77	12,42				
	37,86	9,53				
hCas9	38,03	8,33	5,48			
	39,52	2,63				
sCas9	39,52	3,92	3,92			

Note: here, n is 3 technical qPCR replicates not biological replicate.

Appendix III: U6 promoter sequences

Human U6 promoter:

gagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaag
atattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttacc
gtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggac

Zebrafish u6 promoter:

cacctcaacaaaagctcctcgatgtcacacaggaagttcaggaacttatccaatcactctaaagaaacggcctgtttccttcgcatacgc
ttacagctccaaaactctacggtaaacctacataaactgctggttttcaaattttaaagaatttaagggtttacaggtttactactacacagtg
atttactgacacatgtaggtgtaaatgagttgaataagtaagtaagctatataccacacatgaaacacatacccagaagtcactggtatat
atagccgtcctccagactccca

Mouse u6 promoter:

gatccgacgcgccatctctaggcccgcgccggccccctcgcacagacttgtgggagaagctcggctactcccctgccccggttaattt
gcatataatatttcctagtaactatagaggcttaatgtgcgataaaagacagataatctgttctttttaatactagctacattttacatgatagg
cttggatttctataagagatacaaatactaaattattatttttaaaaaacagcacaaaaggaaactcaccctaactgtaaagtaatt

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway