Norwegian University of Life Sciences

Master's Thesis 2022 60 ECTS
Faculty of Biosciences (BIOVIT)

Functional interrogation of the *elovl5* gene regulation in Atlantic salmon

Shawline Ema
Master of Science, Genome Science

Acknowledgments

To begin with, my time at the Norwegian University of Life Sciences (NMBU) was quite valuable since I was able to study in a stimulating and supportive atmosphere. This university has given me every chance to succeed in my professional life after graduation. I appreciate everyone in the classroom who made my learning experience so smooth and enjoyable.

This thesis was supervised by Simen R. Sandve and conducted at the Center for Integrative Genetics (CIGENE) lab. It was a pleasant experience to work in this great place with fellow people. Thank you all for countless guidance and support in learning lab techniques. Specially I want to thank Thomas Nelson Harvey and Alex Datsomor for all the help and guidance during the lab work. Without your countless support, it would not be so easy to complete my lab work in time. Moreover, I want to thank my main supervisor Simen R.Sandve for helping me in every step of this study by guiding me through right way. It would have been challenging to complete my lab journey without your guidance, encouragement, and knowledge in proper understanding way. At the end I want to express my gratitude to my family for believing in me. Specially without the support of my husband and daughter, it would not be possible to finish up my work and writing in time.

Abstract

Long-chain polyunsaturated fatty acids (LC-PUFA) and cholesterol are important for different physiological pathways including immunology. As these cannot be synthesized in the teleosts including Salmon itself, food supplement is considered the prime source in this case. Two duplicated genes in salmon *elovl5a* and *elovl5b*, derived from the WGD event at about 80 Mya, are important in the synthesis of LC-PUFAs. Differential gene regulation has been shown for these genes in previous studies, including highly divergent tissue regulation. This tissue-specific expression might be the result of different binding patterns of major lipid-metabolism transcriptional regulators. According to Carmona-Antoñanzas et al., 2016, *elovl5a* equally responded to the two sterol regulatory element-binding proteins (srebp-1 and Srebp-2) while *elovl5b* responded stronger to only Srebp-2. In this thesis, we investigated the roles of different transcription factors in the gene expression for the two salmon *elovl5* gene copies. We first conducted a reporter-promoter assay to measure the effect of the Lxr-Srebp regulatory pathway on elovl5a and elovl5b regulation. Next, we conducted a CRISPR-Cas9 knockout experiment to assess the importance of predicted binding sites of Lxr, Srebp-1, Srebp-2, and NY-F for the transcriptional regulation of *elovl5b*.

Though our findings were not completely conclusive, induction of the Lxr-Srebp pathway increased the expression of both *elovl5* promoters in the SHK-1 cells. E*lovl5a* was showing stronger gene expression compared to *elovl5b*, moreover synthetic (*elovl5a* ATAC and *elovl5b* ATAC) were stronger in upregulation than the native promoters. In the CRISPR-cas knockout experiments, we found that mutating Srebp-1, Srebp-2, and NF-Y binding motifs in the elovl5 promoter decreased elovl5b expression. In addition, our experiments indicate that Srebp binding sites were potentially important for regulating *elovl5b* gene through the Lxr-pathway as knock out of these binding sites led to ablation of the Lxr-agonist effect. In conclusion, our experiments shed light on the gene regulatory mechanisms of *elovl5* genes, supporting previous findings regarding the Srebp pathway-mediated regulation in non-salmonid cell lines. Moreover, the novel approach tested for targeted CRISPR-based knock-down experiments will pave the way for more effective interrogation of gene regulatory mechanisms in the years to come.

Contents

1: Introduction:

1.1 Atlantic salmon and the Salmonidae family:

The Atlantic salmon (*Salmo salar*) is a ray-finned fish that belongs to the Salmonidae family. It is an economically important fish species of the Salmonidae family, whereas trout and char fish are also known as economically important species. Salmon is not only important economically, but also nutritionally. It is a rich source of protein, vitamins, and minerals. Salmonids has approximately 11 genera and 70 extant lineages (Lien et al., 2016). Salmon aquaculture is valued at over 8.5 billion GBP (EUR 9.7 billion) per year (FAO 2017) and contributes considerably to food, economic, and employment security in several countries, including Norway, Chile, Canada, and the United Kingdom (Houston & Macqueen, 2019). As an important fish species, the whole genome sequencing of these three species is a matter of interest in the aquaculture industry. The whole genome sequencing of Atlantic salmon is now available which makes the advanced research very much accessible for the researchers to study the evolution of salmonid species from their common ancestors and gene regulation pathway.

1.2 Gene duplication

1.2.1. Importance of gene duplication in eukaryote genome evolution

Whole-genome duplication (WGD) has been important in the evolution of eukaryotic organisms (Van de Peer et al., 2009). It is the process of forming an organism with an extra copy of its whole genome. It is also known as polyploidy. WGD form novel genes and creates extra genetic material that can be raw material for the evolution of new gene functions (Sahlström, 2021). Unbalanced crossing over, retro-position, or chromosomal (or genome) duplication can all lead to gene duplication. Gene duplication and its subsequent dynamics are critical because they give insight into genome-wide elements of evolutionary processes altering intra-specific and inter-specific genome contents, evolutionary connections, and interactions between different organisms and species (Magadum et al., 2013). Gene duplication can supply fresh genetic material for mutation, drift, and selection to operate on, resulting in novel or specialized gene functions (Magadum et al., 2013). So, it can be said that without genome duplication the variability among different species and organisms would be limited. More than 90 percent of eukaryotic genes and at least half of prokaryotic genes are the results of gene duplication, which is the driving force behind gene creation (Teichmann & Babu, 2004).

1.2.2 The evolutionary fates of gene duplicates

Following gene duplication, the newly formed duplicates can have different fates; neo-functionalization, sub-functionalization, and pseudogenization (i.e., loss). In the neo-functionalization duplication produces two copies of the gene where both would be free from the selective pressure, and one copy retains the ancestral function while other evolves a new function. Examples of neo-functionalization include a new snake venom gene i.e.; phospholipase A2 genes (Lynch, 2007). But when both copies of the duplicated gene are equally functional produces sub-functionalization. Allowing adaptation of sub-functions by splitting ancestral functions into two different genes provides an advantage in this approach (Des Marais & Rausher, 2008). On the other hand, when one copy of the gene has a detrimental effect on the population, it would be possibly deleted by the natural selection and would not add any novel variation. Such as; neurological disorders namely Pelizaeus–Merzbacher disease could be added to this list, to characterize the loss of duplicated gene function (Des Marais & Rausher, 2008).

1.3 Whole genome duplication in Atlantic salmon (*Salmon salar*):

About 100-80 million-years ago a whole-genome duplication happened in the common ancestor of salmonids and now it affords unique opportunity to learn about the evolutionary fate of duplicated vertebrate genome (Lien et al., 2016; Macqueen & Johnston, 2014). It is documented from many research that fish and mammals share a common ancestor that goes all the way back to the beginning of the vertebrate lineage (Davidson et al., 2010). Recent advent in the sequencing helps to find out the difference between many vertebrate and invertebrate species that ultimately helps to find out the basic evolutionary pattern of this groups. Salmonids retain a distinct evolutionary position in relation to fish species of the Protacanthopterygii, the most basal group of teleosts. Consequently, the salmonids offer a crucial evolutionary connection between the evolution of teleost fish and the evolution of non-teleost fish as well as other vertebrates (Davidson et al., 2010). According to Smith et al., 2013 the members of teleosts share three WGD event before divergent from jawed vertebrates and the third-round teleost specific WGD (Ts3R) occurred almost 320 million years ago (Jaillon et al., 2004). After the salmonoids split from Esociformes in a fourth cycle of WGD, the salmonid-specific auto-tetraploidization event-Ss4R took place at the base of teleost fishes about before 80 million years ago (Macqueen et al., 2014). Interestingly, the nuclear material in the salmon comprises 50–60% repeats and totals around 3 billion bases (Lien et al., 2016).

A prominent WGD type called auto-tetraploidization includes the spontaneous doubling of all chromosomes. In contrast to, allo-tetraploidization which entails the hybridization of different species there might present two set of genomes after hybridization that are different enough from each other and go to two different set of bivalents during meiosis (Otto, 2007). Additionally,

there is a process where a tetraploid genome converts back to diploidy, which is called rediploidization. The restoration of bivalent pairing in salmonids included significant structural reorganizations such as inversions and transposable element inversion (Lien et al., 2016). In the salmonoids there are a good percentage of gene remain tetraploid that might be consider as research topic now or near future to connect with evolutionary science. Still these genes are not well studied (Allendorf et al., 2015). A crucial result of rediploidization was the retention of at least half of all salmonid genes in duplicated pairs from Ss4R (Allendorf et al., 2015; Lien et al., 2016). Recent research on genome evolution showed that in salmonoids due to Ss4R event there was delayed in species radiation (Macqueen & Johnston, 2014) and ecological factors might have impact on the salmonoid diversification (Robertson et al., 2017).

1.4 Transcriptional regulation in eukaryotic cells:

Gene regulation process is sub-divided in two main categories: Transcription and Translation. Transcription is the modification of DNA to RNA while, translation is production of final protein using that RNA obtained from Transcription. Gene regulation is very important in all organisms because it controls the versatility in the living world. The product of gene regulation i.e., RNA or protein that triggers specific phenotypic characters to helps the corresponding organisms in its life cycle. For instance, gene products help the organisms to respond to the environmental actions or adapt in new kind of environment or in the growth and so on. Disruption of any aspect of this mechanism cause malfunction of gene products like RNA or protein formation and directs to the diseases.

The gene regulatory mechanism is far more intricate in eukaryotic cells than the prokaryotic. The reason behind this is mainly the gene structure, that is more complicated in eukaryotic parts. Another good reason is transcriptional regulatory proteins work in concert to control gene expression in a variety of cell types in multicellular organisms. Eukaryotic gene expression is further complicated by the packing and alteration of DNA by methylation (Andersson & Sandelin, 2020). Certain regulatory elements, including as the promoter, enhancer, RNA polymerase, transcription factors, and repressor, are involved in the control of gene transcription (Figure.1.1). Among these, promoters and enhancers play important role in transcriptional regulation (Haberle & Stark, 2018; Shlyueva et al., 2014). The promoters are residing near the transcription start site (TSS) where the transcriptional machinery is going to bind and initiate RNA transcription. Enhancers and repressors are more distal cis-regulatory elements that can impact transcription by interacting with promoters (Figure 1.1). Transcriptional factors (TFs) and coactivators attach to these open chromatin (nucleosome-free) loci in the genome to start the transcription. To regulate a gene, these elements can be positioned on the same chromosome, that called cis-regulatory element or on another chromosome that called trans-regulatory element (Chatterjee & Ahituv, 2017). A single enhancer can also control many genes, and each gene may have several enhancers that regulate it in various tissues or cell types, or even in the same tissue or cell type (Chatterjee & Ahituv, 2017). The transcription starts when the transcription machinery binds to the core promoter region, the transcription machinery comprises RNA

polymerase II, General transcription factors. Generally, Transcription start site (TSS) is the point from where the transcription might be start. The TSS is contained within a core promoter, a brief region that spans around ± 50 base pairs of the TSS. The core promoter acts as the transcriptional machinery's binding platform (Haberle & Stark, 2018).

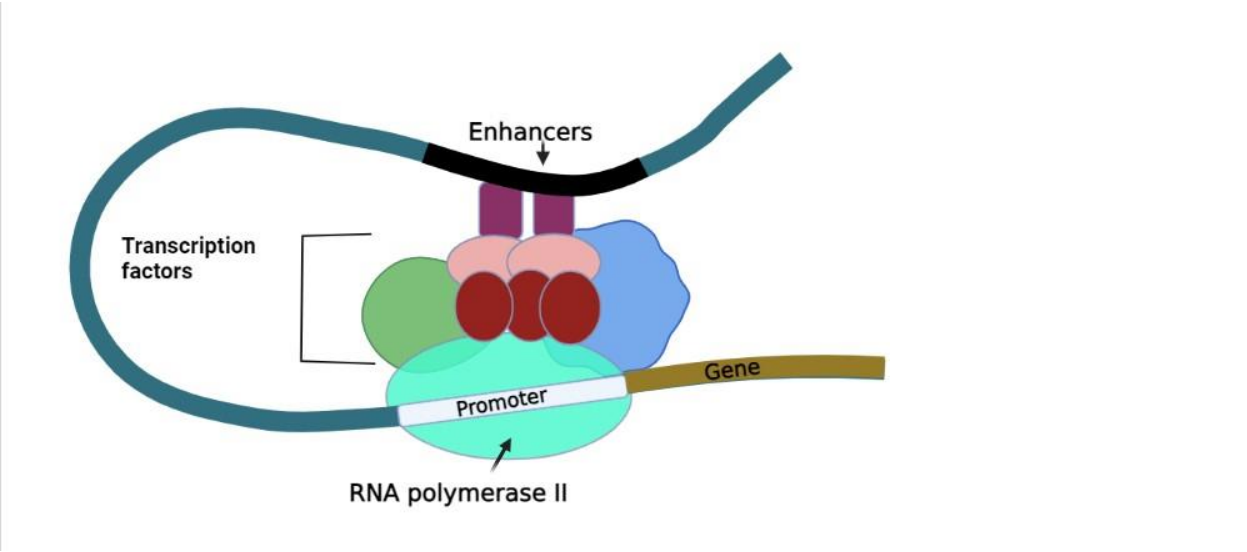


Figure 1.1: Simple outline of Transcriptional regulation in eukaryotic cell. Showing the engagement of different TFs complex, position and action of enhancer and promoter to initiate the transcription. Created in Biorender.com.

1.5 *elovl5a* and *elovl5b* gene regulation in Atlantic salmon:

Different fish species have varying abilities to biosynthesize the n-3 long-chain polyunsaturated fatty acids (LC-PUFA), of which eicosapentaenoic (EPA) and docosahexaenoic (DHA) acids are essential to the health of higher vertebrates (Morais et al., 2009b). Cholesterol and LC-PUFA are essential parts of cellular membranes and crucial precursors of bioactive lipids necessary for homeostasis, cell signaling, immunological and inflammatory responses (Simopoulos, 2008). As salmon is a healthy dietary supply for humans, LC-PUFA pathway is a recent area of vital research. Polyunsaturated fatty acids (PUFAs) are necessary and required in the food for teleosts as well as vertebrates because they cannot synthesize PUFs on their own.

Essential fatty acids (EFAs) like linolenic acid (LA) and alpha-linolenic acid (ALA) can be transformed into long-chain polyunsaturated fatty acids (LC-PUFA) (Rui, 2018). Figure 1.2 presents the LC-PUFA metabolic pathway. In salmon this pathway is directed by *elovl2, elovl4,* two duplicates of *elovl5(elovl5a* and *elovl5b).* These duplicates of *elovl5 (elovl5a* and *elovl5b)* originated in WGD event. Salmon are known as most efficient synthesizer of LC-PUFA as more than one copy of the key genes are presented here (Rui, 2018). Interestingly these gene copies

have very different regulation of transcription across tissues (Carmona-Antonanzas et al., 2016) and it is hypothesized that *elovl5a and elovl5b* genes have evolved to have different functions in the salmon LC-PUFA metabolism (Carmona-Antoñanzas et al., 2014). Studies using promoter reporter assays have found that transcription was induced equally by two sterol regulatory element-binding proteins (Srebp1 and Srebp2) under the elovl5a promoter, but that the promoter of elovl5b showed a substantially stronger response to Srebp2 (Carmona-Antonanzas et al., 2016). From the in vitro study it found that it might be the result of differential binding pattern of major TFs in these two genes.

The transcription of *elovl5a and elovl5b* genes are found regulated by liver X receptor (Lxr) and sterol regulatory element-binding proteins (Srebps) (Carmona-Antoñanzas et al., 2014). Especially, Lxr is considered as a major factor to initiate the transcription of *elovl5* genes by stimulating other dependent TFs (Carmona-Antoñanzas et al., 2014). LXRs (liver X receptors) are important mammalian regulators of lipid and cholesterol metabolism (Kalaany & Mangelsdorf, 2006). Liver X receptor is the member of nuclear hormone receptor (NR) superfamily what acts as ligand-activated transcription factors in the vertebrates (Reschly et al., 2008). Besides, targeting the lipid and cholesterol metabolism Lxr also regulate the carbohydrate and energy metabolism (Jakobsson et al., 2012) and Reschly et al., 2008 reports that it activated by the attachment of the oxysterol ligands that is a cholesterol byproduct.

Another important TF of *elovl5* genes is Sterol regulatory element-binding proteins (Srebps), that is mediated in some extent by the Lxr activating pathway in the mammalian cells (Carmona-Antoñanzas et al., 2014). Srebps is characterized by a tyrosine residue in the DNA-binding domain and a membrane-binding region that is a target for controlled proteolysis (Jakobsson et al., 2012). In mammals there are two Srebp genes, Srebp-1 and -2 found so far, where Srebp-1 regulates fatty acid metabolm and Srebp-2 involve in cholesterol metabolism (Horton et al., 2002). Another lipid metabolizing protein Fatty acid synthesis-1 (Fas-1) is a combined target of Lxr and Srebp-1 in the salmon cell (Carmona-Antoñanzas et al., 2014). These TFs are responsible to sense the status of the lipid and fatty acid in the cells by regulating the transcription of *elovl5* genes. Carmona-Antonanzas et al., 2016 observed that *elovl5a* gene has a specific Lxr response element (LXRE) at 162 nt fragment, so it showed more stimulation with Lxr ligand compare to *elovl5b.* In this study, Srebp-1 and Srebp-2 had strong expression in *elovl5b* promoter than the *elovl5a*. A tandem duplication of SRE and NF-Y cofactor binding sites in *elovl5b* might be the cause of variations in the strength of the Srebp response between these two promoters (Carmona-Antonanzas et al., 2016). All but a few *elovl5* genes have undergone transposon mobilization prior to current salmonid speciation, and this suggests that the difference in *elovl5* regulatory areas has facilitated the emergence of neofuntionalization by encouraging differential expression of these homeologous genes (Carmona-Antonanzas et al., 2016).

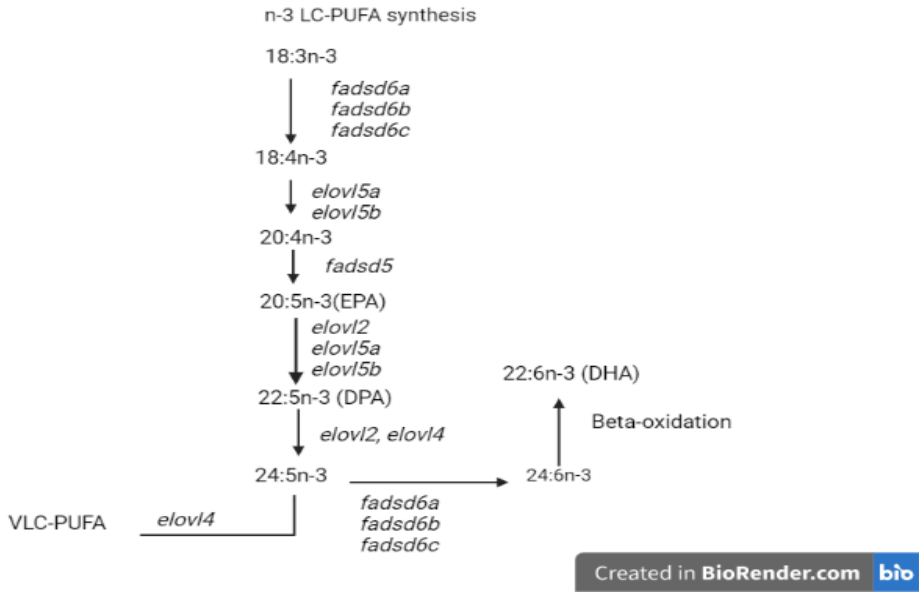


Figure 1.2: n-3:LC-PUFA (omega-3) metabolic pathway synthesize from essential fatty acids (EFAs). There also present the name of Atlantic salmon's gene that are required for synthesis of LC-PUFA. Adopted idea of the pathway from (Gillard et al., 2017). Figure created by Biorender.com.

1.6. Promoter-reporter assay (PRA)

Gene regulation is sophisticated way where several regulatory elements are connected. A DNA segment near a gene known as a promoter, contains CREs that can serve as transcription factor binding sites (TFBS). Depending on the cellular context, cell type, environment, etc., these TFs may behave as repressors, activators, or occasionally both. Promoter-reporter assay is a tool to examine the promoter of a gene, that regulates the gene expression. If the promoter of interest link to a gene it would produce an easily detectable gene product in this assay, such as the firefly luciferase; then the impact of promoter sequences on transcription could be indirectly measured (sigma-aldrich). This assay can be of different types depending on the either delivery methods of the reporter construct to the cell and signal quantification methods. i.e; in vitro or in vivo cell culture or electroporation or chemical transfection etc. In this thesis we used electroporation method to transfect the promoter of interest with firefly gene. The cis and trans regulatory regions of a gene can be analyzed using this method.

The mechanism of the PRA can be described as, the expression of luciferase gene is depended on the TFs binding ability to the cis-regulatory region. The bioluminescent produced by the assay is proportional to the gene expression (Branchini et al., 2018). Difference in transfection efficiencies, cell viability, and cell numbers in each experiment can lead to experimental variation. Therefore, cells are co-transfected with a control vector to normalize for these sources of technical variation. The Renilla luciferase value is used to normalize the firefly luciferase value. The mechanism of reporter promoter assay is in Figure 1.3**.**

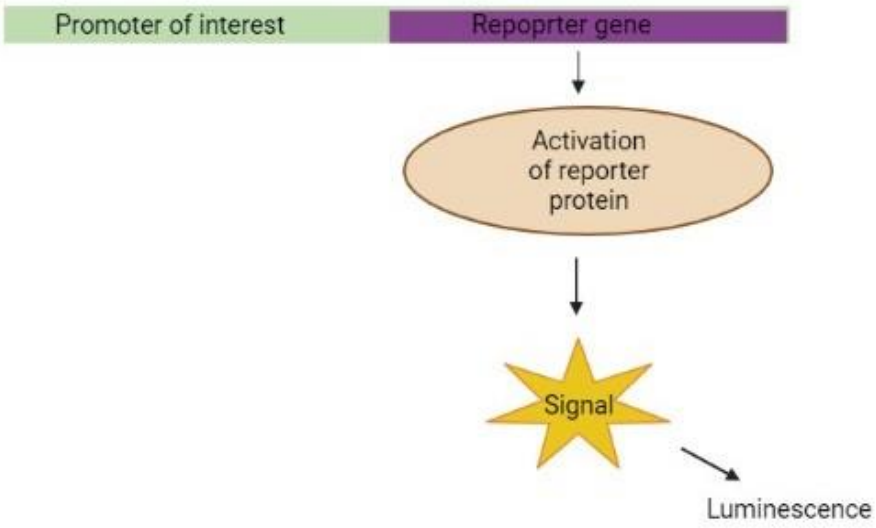


Figure 1.3: Schematic presentation of the promoter reporter assay. The luciferase experiment will show an increase in fluorescence if the promoter of interest can trigger transcription of the reporter gene. Figure created by Biorender.com.

In PRA it is essential to transfer the promoter construct into the cell with good efficiency. There are two ways to transfect the cells- electroporation and chemical transfection. For practical consideration electroporation is supposed to be more convenient and efficient. It is important to transfect high number of cells in the promoter assay to quantify the gene regulation. Less transfection efficiency will lead to difficulty to measure the gene expression in this assay. Cell viability and transfection efficiency are two important factors to perform a successful promoter assay. The foreign DNA is introducing into the cells through electroporation that will ultimately participate in the gene regulation of that cells and help to express the reporter gene transiently or permanently (Wilberg, 2020). In the PRA it is also very important to choose right vectors to transfect the cells. The desired promoter region is integrated into the plasmid vector and transfect by the electric pulse in the electroporation. This electric pulse generates electric field surrounding the cells and produces a temporary pore through which the vector is mediated into the cells and nucleus (Kim & Eberwine, 2010).

1.7 CRISPR-Cas9 technology:

CRISPR/Cas9 or Clustered regularly interspaced short palindromic repeats is a specialized, effective, and adaptable gene-editing tool that allows us to edit, delete or fix particular parts of DNA. CRISPR/Cas9 modifies genes by cutting DNA accurately and uses natural DNA repair

mechanisms (Adli, 2018). It is made up of two components: the Cas9 enzyme and a guide RNA. The CRISPR-Cas9 system evolved in bacteria and archaea as an antiviral defensive mechanism. The system works by storing (i.e., 'memorizing') DNA sequences from prior bacteriophage infection in the bacterial genome. Upon the next infection of identical bacteriophages, the prokaryotes can recognize the infection and cleave the bacteriophage DNA with Cas9 enzymes. Cas9 is an endonuclease based on crRNA that has two nuclease domains: HNH and RuvC. These domains aid in the cleavage of target and non-target DNA (Zhang et al., 2020). Double-strand breaks (DBS) are introduced at target locations by CRISPR-Cas9, which is followed by the introduction of guide RNAs to particular DNA sequences (gRNAs). Cas9 frequently connects to the PAM (protospacer adjacent motif) in Figure 1.4. One of the most prevalent methods cells utilize to mend DSBs is homology-directed repair (HDR), which is also known as non-homologous end joining (NHEJ). NHEJ introduces indels that lead to mutations with loss of function, whereas HDR fixes a mutation by introducing a template DNA sequence (Roy et al., 2018). CRISPR is presently a very popular approach in a wide range of study fields because of its potential application features with excellent efficiency and gene edit rate.

The functional gene regulation of several fish species, including salmon, has been studied using CRISPR technology. The functional significance of a specific gene or mutation in a characteristic of interest, such as infection resistance, may be tested using CRISPR/Cas editing (Staller et al., 2019). In this process the gene of interest could be knock out or transcription blocked by inactivated domain to study the expression or regulation of a particular gene. In vivo genome editing using CRISPR/Cas9 has been effective in Atlantic salmon (Edvardsen et al., 2014). The *dnd* gene, which controls the development of germline cells, was similarly knocked out using CRISPR to produce a sterile salmon (Wargelius et al., 2016). The Chinook salmon (*Oncorhynchus tshawytscha*) cell line (CHSE-EC, descended from CHSE-214) was modified to stably produce Cas9 and EGFP, and this was the first report of effective CRISPR/Cas9 editing (Dehler et al., 2016). Fish cell lines edited by CRISPR/Cas ribonucleoprotein (RNP) complexes, as has been shown with up to 62% effectiveness in medaka (*Oryzias latipes*) (Liu et al., 2018). CRISPR/Cas9 gene editing in fish embryos has been effective, but its application in cultured fish cells has been limited (Liu et al., 2018).

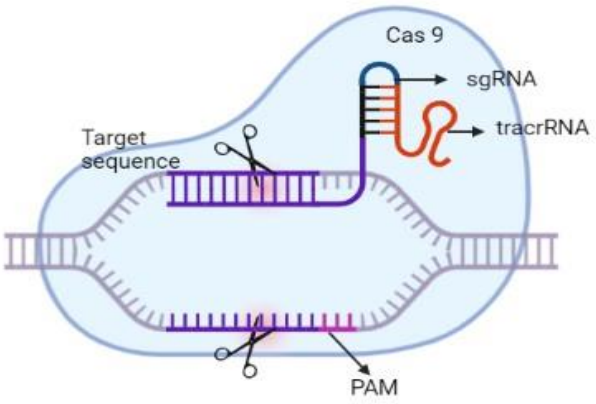


Figure1.4: The diagram showing the CRISPR-Cas9 system in action. PAM is the point at which the Cas9 adjusts to begin cutting. The sgRNA, which is made up of the tracrRNA and crRNA, directs the Cas9 to the best cutting location. Created in Biorender.com.

1.8 Aims and objectives:

Two copies of the *elovl5* gene, originating from a WGD 100-80 million years ago, are found in the Atlantic salmon genome. Tissue-specific regulation of these copies (*elovl5a* and *elovl5b*) has been demonstrated, and differences in promoter architecture/sequence have been linked to this regulatory evolutionary divergence. Previous investigations, however, were restricted to plasmid-based reporter assays and exclusively studied promoter function in non-salmonid cell lines.

The aim of this study is twofold:

(i) To assess the regulatory control of Lxr-pathway on *elovl5* duplicates in salmon cells (Salmon head kidney cells, SHK-1).

(ii) Use CRISPR-based approaches to study *elovl5* regulatory elements in the genomic context.

2. Methods:

2.1 Vector preparation:

2.1.1 Transformation of cloned vector on the competent cell:

To multiply the plasmid vector prior to transformation into salmon cells we used competent *E. Coli* cells (Invitrogen, # 18265017). The transfection of *E. coli* cells was done following the manufacturer's protocol, outlined briefly here. We thawed 50 µL of the competent cell (E. coli), which was stored at -80°C, on ice for 30 minutes. The competent cell was then combined with 3 µL of the plasmid mixture, and the mixture was put on ice immediately. The mixture was then subjected to a thermal shock treatment for 45 seconds at 42°C, followed by 2 minutes on ice. The total volume was increased to 500 µL by adding 445 µL of pre-warmed (37°C) SOC (ThermoFisher #15544034) to help the cells recover from heat shock. The mixture was then incubated at 37 °C with 220 rmp stirring for an hour. 30 µL of bacterial culture was utilized after an hour for plating on an Agar plate that had been pre-prepared with the vector-specific antibiotic (ampicillin) (see section 2.1.2). These cultured plates were then placed in the incubator for an overnight culture at 37 °C. All steps were carried out in the sterile hood.

The next day, plates were checked for bacterial colony growth. Single colonies were carefully moved from the plates to 50 ml falcon tubes with LB broth medium using a pipette tip. We then added 15 µL of ampicillin antibiotics and gently swiveled them to achieve a homogeneous mixture. Falcon tubes were stored overnight in an incubator at 37°C and 220 rmp, it was capped loosely to facilitate bacterial growth. As *E. coli* bacteria is facultative anaerobic bacteria it grows well in the presence of oxygen.

2.1.2 LB Broth medium + agar and LB broth medium:

LB agar medium was used to grow the *E. coli* bacteria with vector construct. First 17.5 mg of LB broth with agar (Sigma Aldrich, #L2897) and 10 g of LB broth base (Invitrogen, #12780052) were measured and transferred to 500 mL bottle. A magnetic stirrer was added to the LB broth agar medium bottle only. 500 mL of distilled water was added to each bottle and Shaked to dissolve the powder with the water. Then the lid was attached loosely to the bottle by covering with aluminum foil marked with autoclave tape. Later, these bottles were autoclaved for nearly 1 hours at 121^{o}C temperature. As we autoclaved the glass bottles at high temperature it was possible to break out the bottles at this high temperature pressure, so we kept the lid slightly loose. After autoclave sat bottles on the machine for a while to be cooled. The LB broth agar medium was cooled on magnetic stirrer to avoid coagulation. After cooling down added 100 mg/mL ampicillin in the LB broth agar medium only in sterile hood. For 500 mL used 500 µL of Ampicillin. Then prepared about 18-20 plates with the LB broth agar medium + ampicillin mix and labeled with antibiotic name and preparation date. These plates were stored at -4^{o}C temperature.

2.1.3 Purification of the vector from competent cells:

After overnight culture plasmids were purified from the competent cells by using the ZymoPURE™ Plasmid Miniprep Kit (Zymo Research, #D4210), following the manufacturer's protocol. We purified pEGFP-N1, pGL4.10[luc2] backbone, pGL4.10[luc2] *elovl5a* WT, pGL4.10[luc2] *elovl5a* ATAC, pGL4.10[luc2] *elovl5b* WT, pGL4.10[luc2] *elovl5b* ATAC construct with this method.

To separate the bacteria from the culture medium we first centrifuged the overnight culture in a 50 ml tube at 4000 rmp, at 4°C for 10 minutes, and discarded the supernatant. Following this, we added 500 µL ZymoPURE ™ P_1 (Red) instead of 250 µL to adjust for a larger pellet size that described in the protocol. Then we divided this mixture into two 1.5 ml tubes. Following that added 250 µL ZymoPURE ™ P_2 (Green color solution) and ZymoPURE ™ P3 (Yellow) per tube for lysis and mixed it immediately by inverting gently 6-8 times. After complete lysis of the cells, we transferred 600 µL of supernatant from each of two tubes into another clean 1.5 ml tube by not disturbing the pellet. We then added 275 µL ZymoPURE ™ Binding buffer to each of 1.5 µL tubes and mixed thoroughly and transferred the entire mixture in the 1.5 ml collection tubes. In the next steps washed the column with ZymoPURE ™ Wash 1 and 2 according to protocol and spin down at 11000 X g for 1 minute. In the following step finally transferred two columns into two clean 1.5 ml tubes and added 25 µL of hot Nuclease-free water (heated at 55°C, 15 minutes) directly in the column matrix. After the elution, the tubes were incubated for 5 minutes instead of 2 minutes to elute completely. Tubes were then spin down for 3 minutes instead of 1 minute at 11000 X g in a microcentrifuge and measured the DNA concentration using a Nanodrop spectrometer.

To isolate plasmids for transfection into salmon cells we also used the ZymoPURE II Plasmid Midiprep Kit (Cat. # D4200) following the manufacturer's protocol, it was used to extract a large amount plasmid at a time (around 200 µL). After, the DNA extraction we preserved the vector constructs at -20°C for future use. We extracted the pGL4.75[hRluc] vector (Promega #E6931, GenBank®Accession Number: AY738231.1), an internal normalized control vector, and pEGFP-N1 with this method.

2.2 Cell line culture and maintenance protocol:

2.2.1 Preparation of complete L-15 media:

For in-vitro maintenance and correct development of the SHK-1 cell line, Leibovitz's L-15 complete medium used in the current investigation. The preparation of complete L-15 media was performed in the sterile hood. At first, cleaned all the things used in the preparation with 70%

ethanol. Then thawed L-15 Medium + GlutaMAXTM Supplement (ThermoFisher Scientific, #31415029), Foetal Bovine Serum (FBS) (Sigma Aldrich, #F7524), and PenStrep (ThermoFisher Scientific, #15140-122) at 37^{o} C in the water bath. Thereafter, calculate the amount for 5% FBS, 1X PenStrep, and 40µM 2-Mercaptoethanol (Sigma Aldrich, #M3148) for a 500 ml bottle of L-15 and aseptically pipetted out 25 ml of FBS, 5 ml of PenStrep, and 1.4 µL of 2-Me to add in L-15. This mixture was used to maintain the cell line and stored at -4^{o}C.

2.2.2 Initiation of cell culture from Atlantic Salmon head kidney cell:

SHK-1 cell has been established from the primary cell culture of Atlantic salmon head kidney (sigma-aldrich). In the present study, this cell line has been used to test hypotheses about the evolution of gene expression.

The safety cabinet was initially sterilized filled with 70% ethanol, and the frozen ampoule of the SHK-1 cell (stored in liquid nitrogen, -80° C) was opened by holding with a tissue paper saturated in 70% ethanol. Only a quarter of the ampoule's cap needed to be rotated to liberate any trapped liquid nitrogen before the lid was retightened. We then thawed the cells in cryovials at 37^{o} C water bath and ensured that the water did not contact with the cap. Once the frozen cell suspension had melted, swiftly remove the lid and clean the outside with 70% ethanol. As cells were preserved in the dimethyl sulfoxide (DMSO, Sigma Aldrich, #D8418) the later steps had done very quickly as the DMSO is toxic for cells above +4^{o} C. We aseptically pipetted out 2 ml pre-warmed L-15 and mixed with the thawed cell suspension in the cryovials. The L-15 was added drop by drop initially to prevent the osmotic shock of the cells and finally pipetted up and down gently to make a homogenous mix. The whole mixture was then transferred to a 75 cm^{2} flask (SARSTEDT, #83.3911.002), and an additional 8 ml pre-warmed L-15 medium was added. This flask has specialized filtered cap to allow O_2 into the flask to help in cell growth. Then this flask was incubated for 24 hours at 20^{o}C temperature.

2.2.3 Splitting/ sub-culturing protocol of SHK-1 cells:

To maintain good condition and growth of cells the density of cell needs to be controlled along removal of waste product and dead cells from the media also very important. Every 7-10 days cell cultures were therefore split in half and replenished with fresh L-15 medium to make total volume of 10 ml. The replacement of the medium was conducted in the sterile hood aseptically by did not disturb the cell mass on the flask floor and supplied medium slowly to the side of the flask.

When the cell culture was 70% to 90% confluent, the cell line was split up. To begin, prepare the sterile hood by cleaning it with 70% ethanol and loading it with all the necessary instruments.

After thawing in the water bath, 0.25 percent trypsin-EDTA (Gibco, #25200-072) was aliquoted together with 1X PBS (ThermoFisher Scientific #10010015) in accordance with the necessity/number of flasks to be sub-cultured. All these substances were then stored at room temperature until they were used.

To split the cells at first removed all the previous medium to a waste tube. Then cells were washed with 10 ml PBS two times, here PBS (phosphate-buffered saline) was used to wash the cell to remove the Fetal bovine serum (FBS). As FBS contains calcium and magnesium ions and can inhibit the action of trypsin, it is necessary to remove FBS prior to Trypsin treatment (sigma-aldrich). PBS was added along the side of the flask slowly, not to disturb the cell monolayer directly, and tilted several times to wash. Thereafter we added 2 ml of Trypsin-EDTA to free the cells from the monolayer. The EDTA is a calcium chelator used to remove any remaining cations from the cell's solution. After applying the trypsin-EDTA we left the cells for 3 minutes until the cells formed a round shape, which was confirmed visually in a microscope. Next, we added four times as much L-15 complete medium to quickly inactivate the trypsin because cells are extremely sensitive to it. The solution was gently pipetted up and down to wash all the cells and transferred to a 15 ml tube. 10 ml of cell solution and 10 ml of trypan blue (Sigma Aldrich #T8154) was taken in a 1.5 ml tube and mixed gently. 10 ml of the mixture were loaded on a cell plate and cells were counted using a Bio-Rad TC20 automated cell counter (Bio-Rad, #1450102). Seeding the cells at $5x10^4$cells/cm^2 or 200000 cells per ml. All flasks were labeled according to cell type, passage number, and date of splitting. In this study, no flasks were used more than three times to avoid contamination and split every 7-10 days.

2.3. Experiment 1: Optimization of the appropriate dose of agonist:

One hypothesis is that the differences in liver gene expression of the *elovl5b* gene duplicates has evolved thorough divergence in the Srebp-1 TF binding affinity in the promoters of these gene duplicates. One prediction from this hypothesis is that increasing Srebp-1 levels will induce expression of *elovl5b* more than *elovl5a*. To test this prediction, we first performed an experiment to find an optimal dose of an Lxr (Liver X Receptor) agonist (GW 3965, Sigma-Aldrich™, # G6295), a chemical known to induce transcription of Srebp-1(Sterol regulatory element-binding protein) TF. The overview of Dose-response experiment is in Figure 2.1.

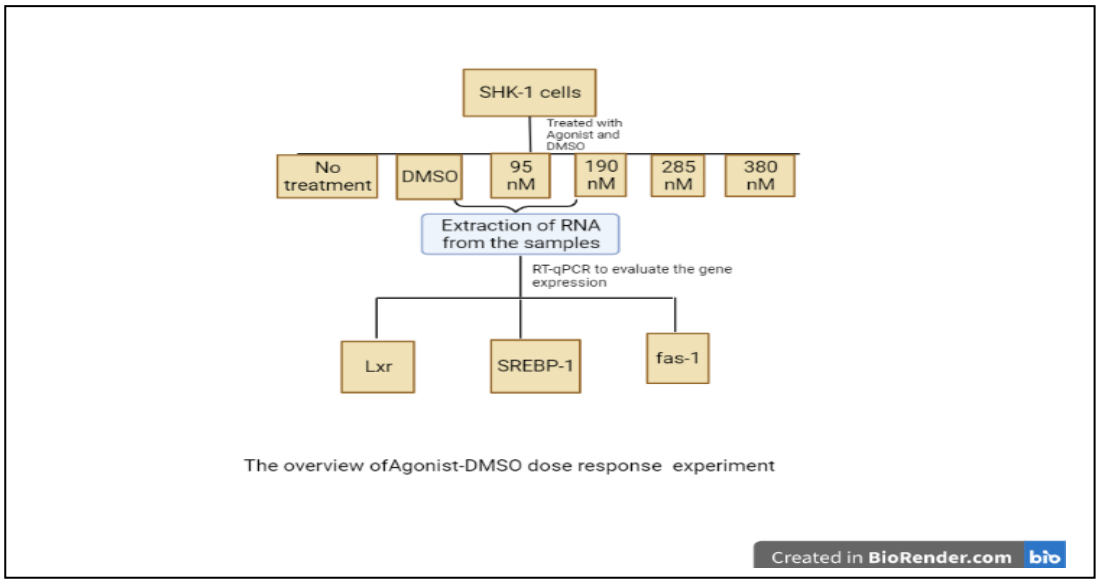


Figure2.1. Overview of Dose-response experiment with agonist and DMSO. We used two negative controls (no treatment and DMSO) as well as four concentrations of the Lxr-agonist GW3965. Gene expression quantification of the direct target gene (Lxr) and two downstream targets of the Lxr pathway (Srebp-1, Fas-1) was carried out using RT-qPCR. Figure created in Biorender.com

2.3.1 Preparing the Agonist (GW 3965) stock and solution:

We dissolved the GW3965 the Lxr-Agonist (3-phenylacetic acid hydrochloride) in DMSO according to the manufacturers protocol, preparing stock concentration of 100 nM to use in future. For preparing 100 mM stock concentration we calculated the amount of DMSO as outlined below:

Mass, m= 5 mg.

Molecular weight M_w = 618.51 g/mol

Stock concentration = 100 mM

So, mass concentration P = $C \times M_w$

$= 100 \times 10^{-3}\ M \times 618.51\ g/mol$

$= 61.85\ g/L$

Volume, V = m/P

$$= \frac{5 \times 10^3\ g}{61.851 g/L}$$

= 0.0000808 L

= 80.8

In total we added 80.8 µL DMSO to 5 mg agonist powder. It was preserved for future use in -4^{o} C temperature.

After preparing the GW3965 stock solution, we prepared test concentrations according to the EC_{50} value of GW3965. EC_{50} is the concentration of a drug that gives half of the maximal response of that drug. We tested four different concentrations: One with EC_{50} (190 nM), as well as 95 nM, 285 nM, and 380 nM.

2.3.2 Cell seeding and agonist treatment:

After dividing the cells as indicated in section 2.2.3, the total number of cells in each ml of fluid was determined using a Bio-Rad TC20 automatic cell counter (Bio-Rad, #1450102). We used Trypan blue staining solution (Sigma Aldrich #T8154) to count the cell. We got 2.56 x 10^5 cells/mL with 92% viability. For this experiment 600,000 cells were optimal for 6 well plates, so calculated the volume of cell solution to get 600,000 cells per well like bellow:

Total cell count x percentage of live cells = live cell count

$2.56 \times 10^5 \times 0.98 = 235{,}520$ cells/mL.

$$V = \frac{N_{cells}}{C_{cells}}$$

$$= \frac{600{,}000\ cells}{235{,}520\ cells/mL}$$

=2.5 mL.

Here the N_{cells} is the number of cells per well and C_{cells} is the number of cells per volume unit. As there were triplicates for each sample (per agonist dose), 2.5 x 3 = 7.5 ml cell solution was needed per sample. After aliquoted the preferred amount in a 50 ml tube. 2.5 ml of solution was gently poured into each well. Then the plates were incubated in the incubator for 24 hours at 20^{O} C. The next day we changed the old medium and supplied 2 ml of complete L-15 medium per wells. In this L-15 medium added specific amount of Lxr agonist according to the doses. Plates were then gently tilted to mix the agonist with medium and incubated for 48 hours at 20^{o} C.

2.3.3 RNA extraction:

After the incubation for 48 hours, we extracted RNA from agonist treated and control cells by using RNeasy Plus Universal Mini Kit (50) from QIAGEN (Cat. No. / ID: 74034) following the manufacturers protocol except for a few minor changes. At first, the hood was cleaned with RNA-free water to remove any possible debris RNA particle or contamination. We washed the cells with ice-cold PBS two times and added 900 µl QIAzol lysis reagent per sample. Following the addition of reagent, the cells were scraped with care so that the cells from each sample were not mixed. A separate scraper was used each time. We then transferred the homogenate in a 2 ml tube and kept in for 2-3 minutes at room temperature to promote dissociation of nucleoprotein complexes. Following this we added 100 µl of gDNA eliminator solution and mixed by vortexing for 15 seconds to allow the elimination of gDNA from the solution. After adding 180 µl of chloroform placed in benchtop for 2-3 minutes to settle down the homogenate. The top aqueous solution was transferred to a fresh microcentrifugation tube after centrifugation. 1 volume of 70 % ethanol was mixed thoroughly by pipetting up and down. In the last step eluted the RNA by centrifugation at 12000xg with 30 µl of RNase free water. To prevent degradation of RNA quality by contamination, we did not mix flow through with the extraction column. Following that RNA extraction tubes containing RNA were kept on ice for RNA concentration measurements using Nanodrop (ThermoFisher #ND-8000-GL).

2.3.4 RT-qPCR for dose response experiment:

Real time qPCR (RT-qPCR) is used for the detection of nucleic acid and quantifying these by reverse transcriptase reaction. In this experiment, we used RT-qPCR for the measurement of gene expression for Lxr, Srebp-1, and Fas-1 transcription factors and to find out a perfect agonist dose that would be related to the increase of these gene expressions. The RT-qPCR allows us to estimate transcript levels by the detection of fluorescent signals. Here, we applied this method to determine the effect of the Lxr agonist on Lxr, Srebp-1, and Fas-1 transcription in the SHK-1 cell line using the elf-1α housekeeping gene to normalize the differences in RNA-concentrations between samples.

A Qiagen QuantiTect Reverse Transcription Kit (50) (Ref # 205311) was used to synthesize cDNA from the isolated RNA in section 2.3.3 using the manufacturer's instructions. RNA samples were stored in -20^{o}C and was thawed for 30 minutes on ice before starting the work. We used 1.5 µg of RNA sample per reaction and prepared a SYBR green master mix per target gene. We adopted the oligonucleotide primer sequences from (Carmona-Antoñanzas et al., 2011) paper that is in Table 1 to target the desired gene.

Table 1: qPCR primer sequences and annealing temperatures used in RT-qPCR.

Transcript	Forward Sequence	Reverse sequence	Annealing temperature
Lxr	GCCGCCGCTATCTGAAATCTG	CAATCCGGCAACCAATCTGTAGG	58^{o}C
Srebp-1	GCCATGCGCAGGTTGTTTCTTCA	TCTGGCCAGGACGCATCTCACACT	63^{o}C
Fas-1	ACCGCCAAGCTCAGTGTGC	CAGGCCCCAAAGGAGTAGC	60^{o}C
elf-1 α	CTGCCCCTCCAGGACGTTTACAA	CACCGGGCATAGCCGATTCC	59^{O}C

Prepared working solution of all the primers of Table 1 by adding RNAse-free water according to calculation, these were coming in powder form. We used 10 µl of both reverse and forward primer (5µM concentration and in 1:20 ratio) and in the master, mix used 10 µl of SYBR green. The rest of the part was like the protocol. We did each sample in triplicates and also added three wells without any cDNA samples called NTC (No template control) to check if there were any DNA contamination in the samples. Annealing temperatures were set accordingly (Carmona-Antonanzas et al., 2016). Then conduct the qPCR cycle in the Bio-Rad-96 real time qPCR machine according to protocol. We sat the annealing temperature at 60^{o}C.

Table 2: Thermal cycling protocol (adapted from iTaq™ Universal green one-step kit).

Reverse transcription reaction	Polymerase activation and DNA denaturation	Amplification			Melt-curve analysis
		Denaturation	Annealing/extension	cycles	
			10-30 seconds		
10 minutes at 50ºC	1 minute at 95ºC	10 Seconds	15-30 seconds		
			60 seconds	35-40	65-95ºC
		15 seconds	10-30 seconds		
			60 seconds		
			30 seconds		

2.4 Experiment 2: Evaluation of the Lxr-Srebp-1 pathway in driving expression under elovl5 gene promoters

To test if the agonist treatment triggers the transcription in the *elovl5a* and *elovl5b* promoters we conducted a promoter-reporter luciferase assay. Here used four promoter constructs: *elovl5a* WT, *elovl5a* ATAC, *elovl5b* WT, *elovl5b* ATAC, one positive control pGL4.10.luc_2 and a pEGFP-N1 (Addgene plasmid # 60360). The purpose of this study was to determine if variations in Srebp-1 binding patterns between the two copies of *elovl5* can explain the variations of gene expression in liver. Previously, from the qPCR experiment, we got a suitable dose of the Lxr agonist that was 95 nM and used this in the reporter assay. pEGFP-N1 was utilized to determine the transfection efficiency. To estimate the gene expression, we used Dual-Glo®luciferase assay. The experiment was conducted in triplicate form. We repeated luciferase assay three times in the same experimental condition to reduce the experimental variation.

2.4.1 Transfection of plasmid constructs:

To perform the Dual-Glo®luciferase assay in the SHK-1 cells we transfected the plasmid constructs into the SHK-1 cells with Neon Transfection System Kit (ThermoFisher Scientific,

#MPK1096). In this experiment we used pGL4.10[luc2] (Promega #E6651, Genebank association number-AY738222) as the vector to transfect our desired promoters in the SHK-1 cells. The detail structure of this vector can be found in Figure 2.2.

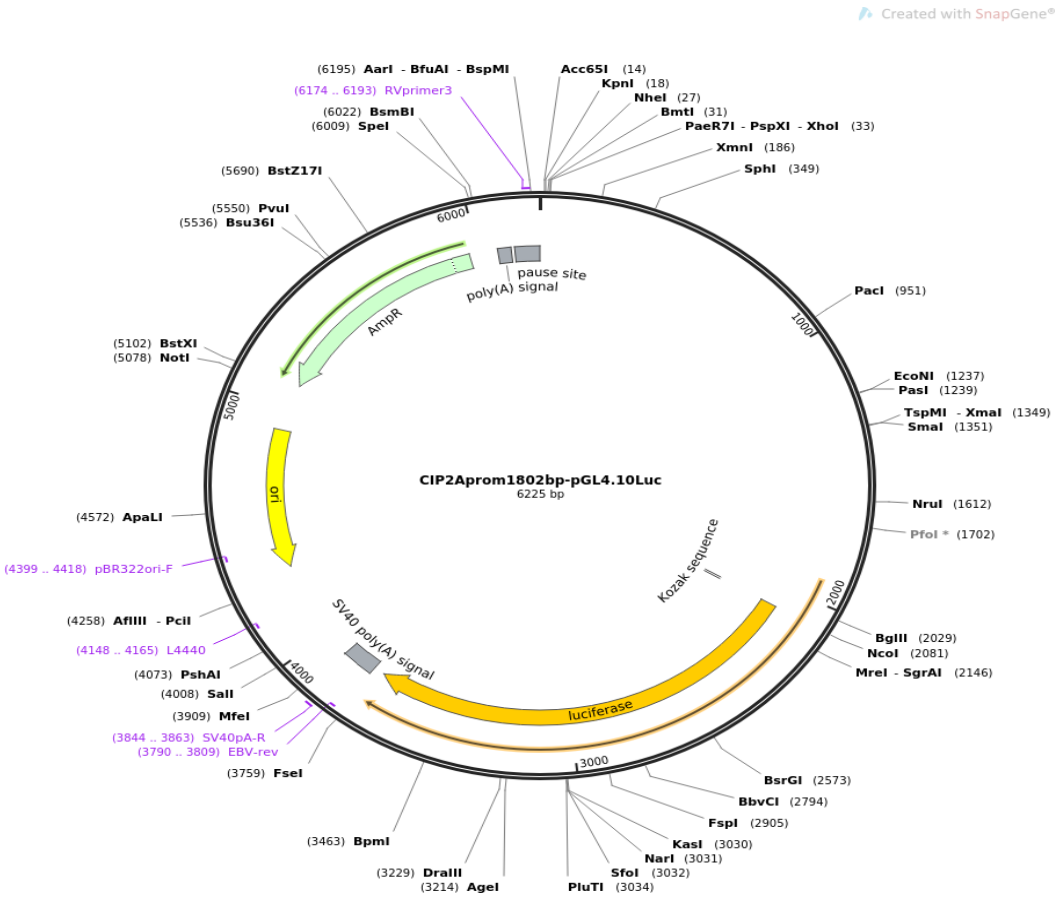


Figure 2.2: Overview of pGL4.10[luc2] (Promega #E6651) vector. boxes represent different coding regions and open reading frames of the vector including luciferase region. Names and lines on the outside of the vector marks different restriction enzyme cut sites.

The pGL4.10 [luc2] vector encodes a luciferase reporter gene named luc2 from *Photinus pyralis* to get high expression and reduced anomalous transcription in the mammalian cells (Promega). It is an optimized vector with fewer consensus regulatory sequences and a synthetic gene in it. This vector is does not have a promoter sequence itself but allows cloning of multiple promoter region on it. It also has a synthetic poly A tail, a transcriptional pause site in upstream of cloning site (Promega, https://no.promega.com/resources/protocols/product-information-sheets/a/pgl410-vector-protocol/). This vector backbone was used to cloning desired promoter sequences of

elovl5a and *elovl5b* genes in it to transfect SHK-1 cells. In figure 2.3 showing the wild type of promoter and synthetic promoter vector outline in pGL4.10[luc2].

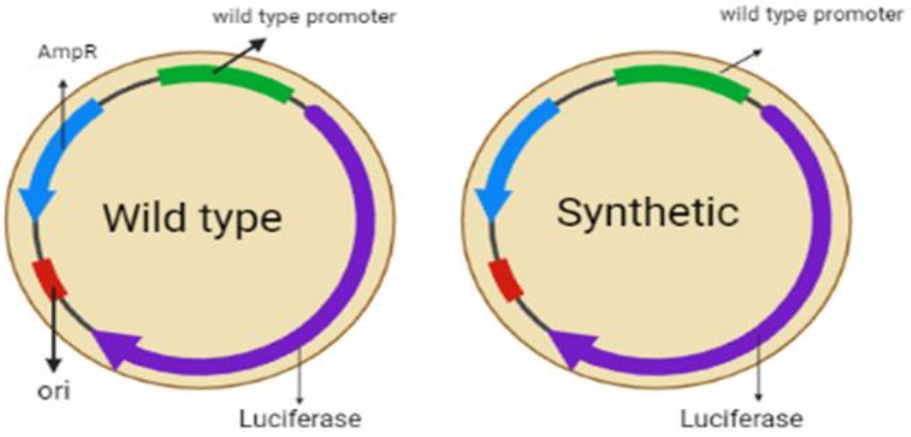


Figure 2.3: *elovl5* promoter constructs with wild type and synthetic. The *elovl5* WT and synthetic promoter constructs in the pGL4.10[luc2] basal vector. The promoter's sequences were cloned on the upstream of the luciferase gene (purple). Figure created by Biorender.com

We also used pGL4.75 vector (Promega #E6931, GenBank®Accession Number: AY738231.1) as an internal control vector (Figure 2.4). The pGL4.75 vector was co-transfected into each cell to measure biomass/cells number. Since the Renilla enzyme binds to a different substrate other than the firefly enzyme, bioluminescence may be detected independently of the firefly enzyme in the same cells. The internal control vector becomes a baseline for having the insight about the of transfection efficiency and to normalize the samples.

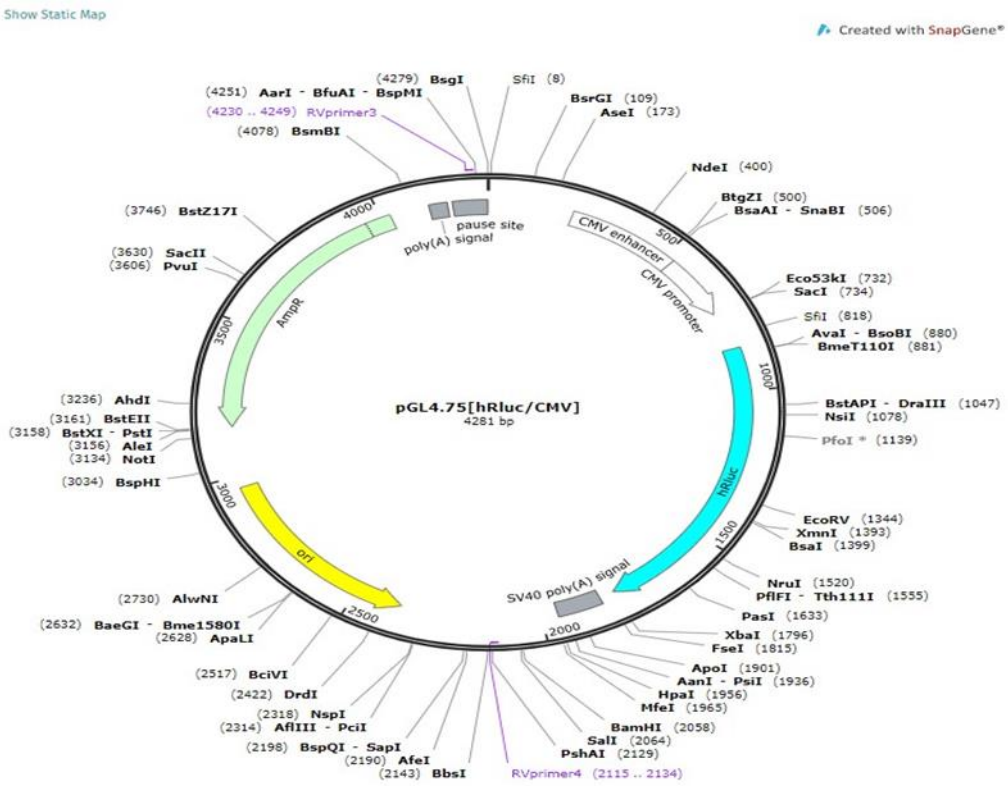


Figure 2.4: Overview of pGL4.75 [hRrluc/CMV] (Promega #E6931) vector. Here the boxes are representing coding regions and open reading frames of the vector including luciferase region and outside marking different restriction enzymes and cut sites. The light green region is Ampicillin resistant region and sky blue color box is renilla luciferase gene.

pEGFP-N1 was utilized to determine whether the transfection was effective or not. pEGFP-N1 plasmid contains a green fluorescent protein coding sequence from *Aequorea coerulescens* that produces fluorescent light in the mammalian cells (RealGene, https://www.realgenelabs.com/pegfp-n1-mammalian-expression-vector-514013.). In Figure 2.5 presents the overview of the pEGFP-N1vector.

To deliver the plasmid constructs complex into the SHK-1 cell line with Neon Transfection System Kit we followed the manufacturer's protocol. The electroporation setting was 1400 Volt, 20 ms, 2pulses according to the protocol. For transfection we first prepare the L-15 medium without antibiotic, as the transfected cells are very sensitive to antibiotic. For each transfection we aimed at 3.87 X10 ^{5} cells. Using the BioRad TC20 cell counter we then estimated the volume

needed to get this number of cells using the following calculation. So, for 39 wells we took approximately 18 ml cell solution.

$$\frac{39\times1.5x105}{3.87\times105} = 16.5 \text{ ml (we took 18 ml)},$$

Next, we centrifuged the samples of cells in medium at 200×g for 5 minutes in two 15 ml tube and discarded the supernatant. Following that washed with PBS (3-5 ml) to completely remove the L-15 medium not disturbing the pellet.

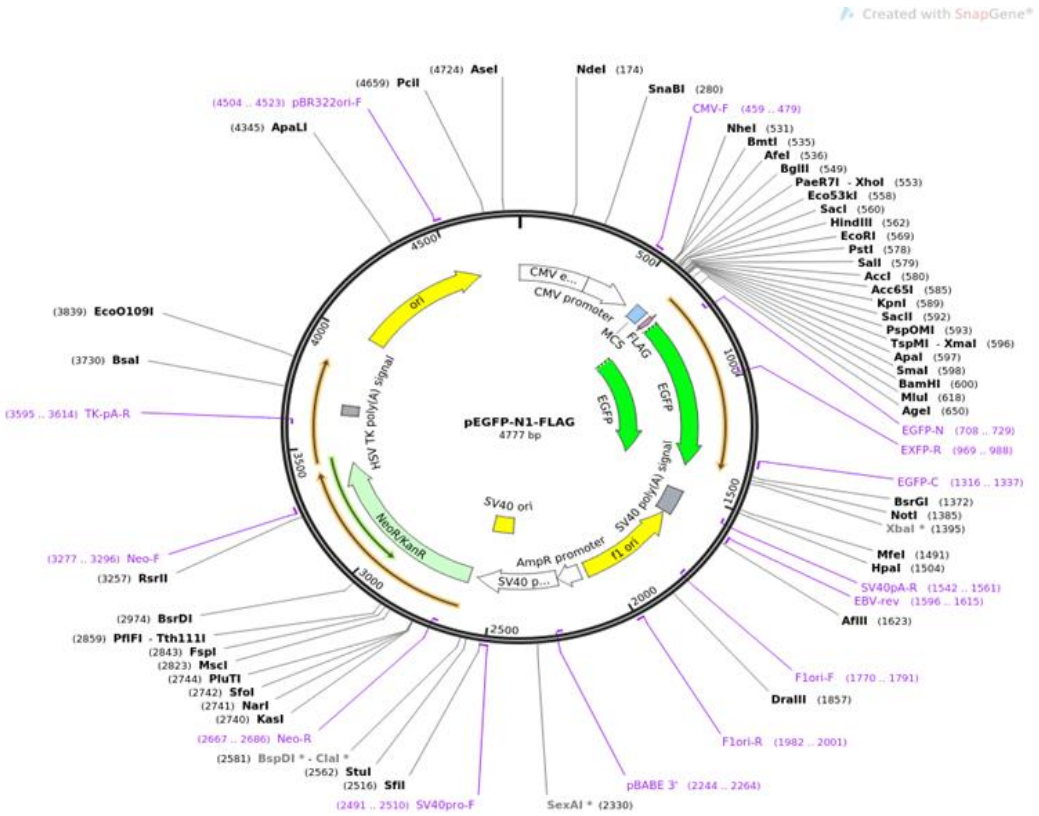


Figure 2.5: pEGFP-N1 vector backbone. Green regions represent the EGFP gene. This gene will produce green, fluorescent light if the transfection would be successful in the mammalian cells.

Resuspended the pellet after removing the PBS carefully in R buffer that came with the Neon transfection kit. We used 7 constructs (*elovl5a* WT, *elovl5a* ATAC, *elovl5b* WT, *elovl5b* ATAC, pGL4.10.luc_2, a pEGFP-N1 and non-transfected control) with three replicates each for Agonist and DMSO. We used 1.5 µg of desired plasmid construct and 0.5 µg of Renilla plasmid as internal control, only exception was for pEGFP-N1, for which we used 2 µg of only pEGFP-N1

plasmid and no Renilla plasmid. In each well we aliquoted 1 ml L-15 medium before starting the transfection. E buffer was placed into the transfection tube after the Neon transfection machine had been set up according to protocol. We used 10 µL pipette to transfect from the Neon transfection system. After 24 hours we treated the cells with agonist and DMSO. To treat the cells with agonist and DMSO we first prepared a mixture of L-15 medium by mixing with Lxr-Agonist/DMSO in two different tubes according to the specific dose. Then replace the old medium that was without antibiotic with L-15 and agonist/DMSO mixture. We aliquoted 1 ml mixture in each well. Then cells were left in incubator again for 48 hours at 37°C (5% CO2). After 48 hours we started for Dual-Glo®luciferase assay according to section 2.4.2. We checked the pEGFP-N1 cells under the fluorescent and white light to monitor transfection efficiency after 24 and 48 hours.

2.4.2 Dual-Glo®luciferase assay:

The Dual-Glo Luciferase Assay is a quick and simple quantification of a stable luminous signal from two reporter genes in a sample. It can reduce false-positive and false-negative results produced by nonspecific variables to minimize sample variability. The renilla luciferase values were used to normalize the variation that occurred due to different transfection efficiency. The reading was noted in the unit RLU (Relative light unit). The Dual-Glo®Luciferase (Promega, #E2920) assay was conducted according to manufacturer's protocol in this study except in step 3, where we increased the time before signal intensity measurement to 45 minutes instead of 10 minutes to get a better signal intensity. After 24 hours, we withdrew the L-15 medium from the transfected cells since it might lower the luciferase assay's effectiveness. Dulbecco's Modified Eagle Medium (DMEM, ThermoFisher, #31331093) was used in place of L-15.

2.4.3 RT-qPCR with promoter construct:

The second RT-qPCR experiment was designed to confirm that Lxr agonist (GW3965) boosted gene expression in the SHK-1 cells by the presence of *elovl5a* and *elovl5b* promoter. SHK-1 cells were transiently transfected with the required promoter constructs in a separate experiment to detect it. After transfection, the cells were treated with 95 nM Lxr-agonist and DMSO dose 24 hours later. The DMSO was used as a negative control in this experiment since Lxr-agonist

solution was prepared using DMSO. Here for RT-qPCR we used iTaq^{TM} Universal SYBR Green One Step kit (Cat. # 172-5150) according to manufacturer's protocol. Primers and PCR program were identical to what we describe in section 2.3.5 and Table 1. We used 100 nM RNA per sample in a total volume of 20 µl. Sat the thermal cycle according to protocol (like section 2.3.5) and fixed the annealing temperature at 60^{o}C. The qPCR analysis was conducted in a Bio-Rad-96 RT-qPCR machine. See section 2.3.4 for details.

2.5. Experiment 3: Testing the transcriptional regulatory regions of *elovl5b* by CRISPR-Cas9 technology

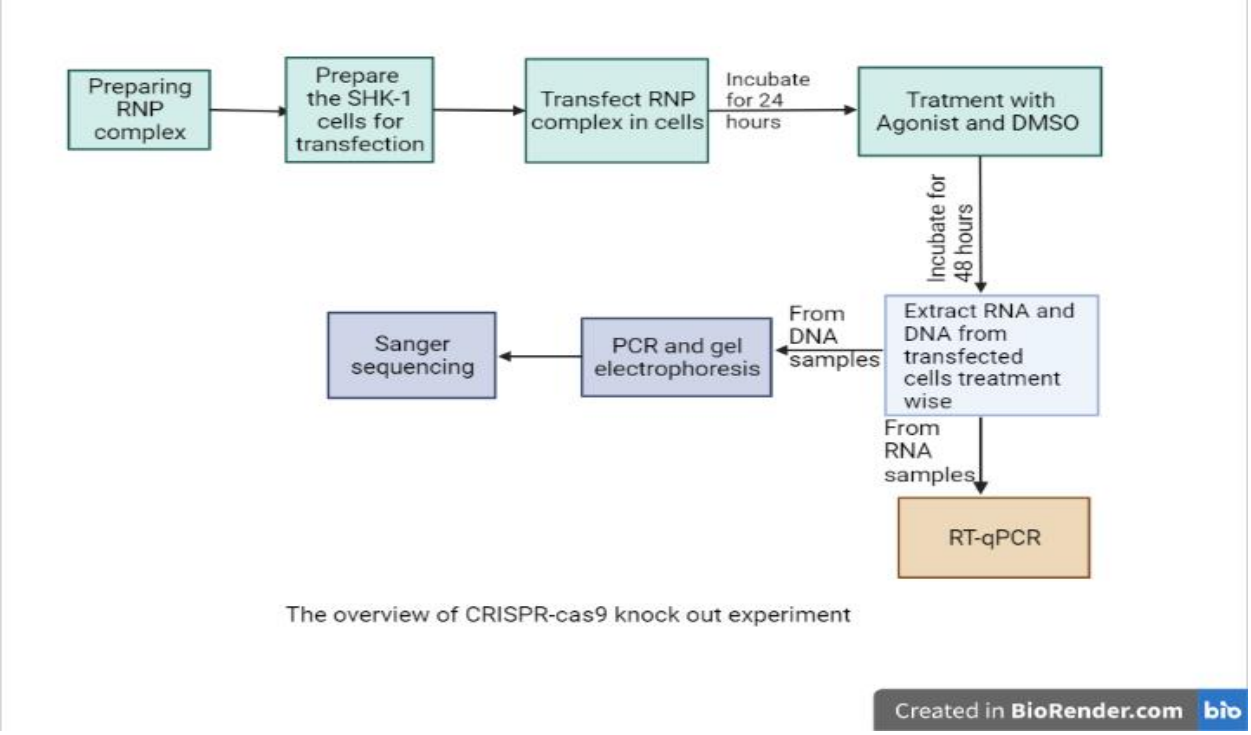


Figure 2.6: Overview of CRISPR-Cas9 knock out experiment using RNP transfection method. Sanger sequencing were used to evaluate the knockout efficiency of the TF binding sites while RT-qPCR was used to measure the impact of TF binding site KO on elovl5b expression. Created in Biorender.com.

This CRISPR-Cas9 knockout experiment aimed to characterize transcriptional regulatory components involved in regulating lipid metabolism and the elovl5b gene in Atlantic salmon. We have targeted two different TF binding motifs for CRISPR-Cas9 based knock out, that are known to be involved in lipid metabolism in salmon (Carmona-Antonanzas et al., 2016; Carmona-

Antoñanzas et al., 2014). These sites are Nuclear transcription factor Y (NF-Y) and sterol regulatory element-binding proteins (Srebps). In addition, we targeted the transcription start site (TSS) as a positive control. We also designed a second positive control named Ssal_tp53 gene to check the success of the CRISPR knock out. If these sites are involved in *elovl5b* regulation through the Lxr-pathway, the gene expression would be less in knocked out cells compare to normal cells. Here, we used direct delivery of CRISPR-Cas9 system as ribonucleoprotein (RNP) complex. The RNP complex is delivered directly in the desired cells along with a single guided RNA (sgRNA). This sgRNA guides the Cas9 protein to the target site to direct the knockout process. The guide RNA was designed using the Snapgene^{R}viewer 6.0 and Salmobase.V3 platform (https://salmobase.org/). We used Salmobase to mark the TFBS in the *elovl5b* gene of salmon. Figure 2.7 is overview of *elovl5b* gene.

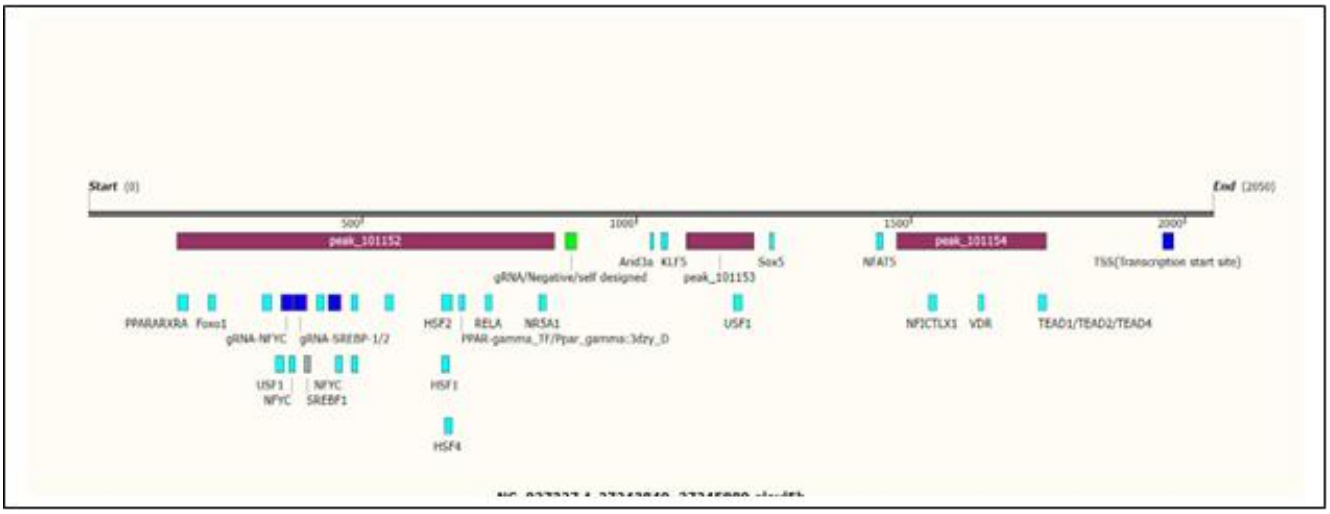


Figure-2.7: The overview of the elovl5b gene promoter. The open chromatin regions are marked in purple color, sky blue color presents different TF binding site and Dark blue color indicates the gRNS sites for NY-F, Srebp-1/2 and TSS. Figure in SnapgeneRviewer 6.0.

Guide RNAs are listed below in table 3 that were used in knockout experiment. We targeted Srebp-1 and Srebp-2 and the guideRNA sequences were different here (Table 3).

Table 3: Guide RNA sequences for the CRISPR KO target sites.

Name of guide RNA	Sequences
NY-C	**GAGTGATGCAGTGATTTGATTGG**
Srebp-1	**CTCCACTGACGAGAAATGAATGG**
Srebp-2	**CTCCACTGATGCGAAATGAATGG**
TSS	**GCTGTTTCCACTGACCGACGAGG**
Ssal_tp53	**GCTCGTACCACTGCCCCAGG**

2.5.1 Preparing RNP complex:

The overview of the CRISPR-Cas9 experiment can be found in Figure 2.6. We first generated the RNP complex with the guide RNA sequence (crRNA), chosen from Table 3, and a Cas9 nuclease-recruiting sequence, tracerRNA, for cellular delivery. We used Alt-R® CRISPR-Cas9 system (Ref# 1072554) to prepare the RNP complex. At first, we split the SHK-1 cells 2-3 days before the electroporation according to section 2.2.4. The optimal cell density for electroporation was 1 x 10^5 to 1 x 10^6 cells/mL. After that, we mixed the crRNA and tracrRNA oligos in equimolar concentrations in a sterile micro centrifuge tube to a final duplex concentration of 44 µM according to the calculation in the table 4:

Table 4: Calculation of tracrRNA and crRNA complex mixture.

Component	**Amount (µL)**
200 µM Alt-R CRISPR-Cas9 crRNA	1.76
200 µM Alt-R CRISPR-Cas9 tracrRNA	1.76
Nuclease-Free IDTE Buffer	4.5
Total volume	**8**

We prepared 8 ml tracrRNA and crRNA complex for DMSO and Agonist treated cells separately. As the Cas9 enzyme supplied in the concentration of 62 µM, we diluted it in 36 µM by adding resuspension buffer R from Neon Transfection System Kit (ThermoFisher Scientific,

#MPK1096) according to the calculation in Table 5. We prepared about 15 µL of Cas9 diluted solution as we will need about 4 µL to prepare RNP complex in total.

Table 5: Calculation of Cas9 protein dilution.

Component	**Amount (µL)**
Alt-R Cas9 enzyme (62 µM stock)	8.7
Resuspension Buffer R (from Neon System Kit)	6.3
Total volume	**15**

Next, we mixed the ca9 nuclease and tracrRNA and crRNA complex gently by pipetting up and down in a 2ml microcentrifuge tube according to Table 6. It was recommended to add 0.5 µL of Alt-R guide RNA (crRNA:tracrRNA duplex) and 0.5 µL of Diluted Alt-R Cas9 enzyme to get the perfect RNP complex for CRISPR Knock out experiment. Then we incubated the mixture for 10-20 minutes in room temperature to prepare the Neon transfection system.

Table 6: Calculation of gRNA solution preparation

Component	Amount
Alt-R guide RNA (crRNA:tracrRNA duplex)	4µL
Diluted Alt-R Cas9 enzyme	4µL
Total volume	8µL

.

2.5.2 Preparation of Neon Transfection System:

To deliver the RNP complex into the SHK-1 cell line we used Neon Transfection System Kit (ThermoFisher Scientific, Ref# MPK1096) according to manufacturer's protocol. The electroporation setting was 1400 Volt for 20 ms in 2 pulses. For transfection we first prepare the L-15 medium without antibiotic, as the transfected cells are very sensitive to antibiotic. After regular splitting procedure we counted the cell number and calculated appropriate volume of cell solution. After counting on the Bio-Rad TC20 automated cell counter (Bio-Rad, #1450102) we

got 6.59 X10 ^{5} cells per ml with 91% live cells. So, for 20 wells we took approximately 22.5 ml cell solution.

$$\frac{37\times4x105}{6.59\times105} = 22.5 \text{ ml},$$

Then we did the transfection according to section 2.4.1. After transfection carefully spread the cells in marked wells according to treatment and guide RNA sequences. After 24 hours we treated the cells with Lxr-agonist and DMSO by replacing the old L-15 medium that was without antibiotic. We aliquoted 1 ml of L-15 and Lxr-agonist/DMSO mixture in each well. After 72 hours extracted RNA from these cells using the same protocol from section 2.3.3. We also extracted DNA for sanger sequencing and PCR experiment according to the manufacturer's protocol (Section2.5.3).

2.5.3 DNA extraction:

After 72 hours of transfection with CRISPR RNPs we extracted DNA from the agonist and DMSO treated cells to conduct sanger sequencing to estimate CRISPR editing efficiency. To extract DNA, we used DNeasyR Blood & Tissue kit (Ref# 69581) following the manufacturer's protocol. After extraction, DNA was stored in the -80^{o} C.

2.5.4 qPCR analysis of CRISPR knockout cell's RNA:

To evaluate the gene expression of the CRISPR knockout cells we used the same procedure in section 2.4.3. We designed the qPCR experiment for TSS, Srebp1 and 2, NF-Y and one control from cells transfected without RNP ('No-RNP'). qPCR primers for Lxr, Srebp-1 and 2, Fas-1 and elf-1 α (both forward and reverse) from the (Carmona-Antonanzas et al., 2016). We selected Annealing temperature and thermal cycle of qPCR reaction accordingly the Table 7 adopted from (Carmona-Antonanzas et al., 2016) .

Table 7: qPCR primers and their annealing temperature.

Transcript	Primer sequence	Annealing temperature
Elovl5b forward	5'- ACAAAAAGCCATGTTTATCTGAAAGA -3'	60^{o}C
Elovl5b reverse	5'- AAGTGGGTCTCTGGGGCTGTG -3'	60^{o}C

2.5.5 PCR of elovl5b promoter and sanger sequencing:

We designed two PCR amplifications; one PCR amplifying the *elovl5b* promoter so that all the TF binding sites for NF-Y, Srebp-1, Srebp-2 as well as the TSS was included in the amplicon and another PCR amplifying the region covering the CRISPR target site in the P_{53} gene. PCR primers were designed using Benchling.com. Table 8 has listed the PCR primers for *elovl5b* and P_{53} gene.

Table 8: PCR primers of *elovl5b* and Ssal_tp53 gene.

Name of gene	Forward sequence	Reverse sequence
elovl5b	CCGGTGTGGTGATACTGATGG	AGCCCAACTCTTATCGCACA
Ssal_tp53	AGAGTCAGACAAGAACAATGGG	CTGTCTCAGAGTGTTACCATCC

To perform the PCRs, we used Invitrogen™ Platinum™ II Hot-Start Green PCR Master Mix (2X) (Ref # 14001-012) according to manufacturer's protocol. Before Starting PCR reaction, we thawed all the frozen ingredients in the ice and prepared the PCR master mix according to Table 9. We have prepared 32X for final concentration and combined master mix, PCR primers, and nuclease-free water according to Table 9. We used 50 ng/ µL DNA per sample.

Table 9: Calculation of Platinum™ II Hot-Start Green PCR Master Mix

Component	Volume for 50- µL rxn	Final concentration
Water nuclease-free	23 µL	736 µL
Platinum™ II Hot-Start Green PCR Master Mix (2X)	25 µL	800 µL
10 µM forward primer	1 µL	32 µL
10 µM reverse primer	1 µL	32 µL

Following the addition of the master mix and aliquots of DNA, the mixture was incubated in the PCR machine as per Table 10. PCR products were stored in -20^{o}C for gel electrophoresis.

Table 10: Temperature and duration of PCR cycle.

Step	Temperature	Time
Initial denaturation	94^{o}C	2 seconds
Denaturation	98^{o}C	5 seconds
Anneal	60^{o}C	30 seconds
Extend		
Hold	4^{o}C	Hold

To visually inspect the PCR products, we made 1 % agarose gel for gel electrophoresis by microwaving 0.6 g of agarose gel with 60 ml of buffer TAE until we had a clear gel solution. As our both amplicons were around 1kb bases, we choose 1kb GeneRular ladder. The gel was run for 1 hour at 100V.

2.5.6 Sanger Sequencing:

We used Sanger sequencing method to sequence the PCR products to be able to estimate KO efficiency of TF binding target sites. For Sanger sequencing we used Mix2Seq Kit from Eurofin Genomics and prepared the samples for sequencing according to manufacturer's protocol by mixing 5 µL of DNA sample with 5 µL of Primer for each sample. The reverse PCR primer was used in the sequencing reactions for all targets. LightRun (Eurofins Genomics) provided a barcode for use on the sample tubes, which were subsequently sent to Eurofin Genomics for sequencing.

3. Results

3.1 Dose-Response of SHK-1 cells:

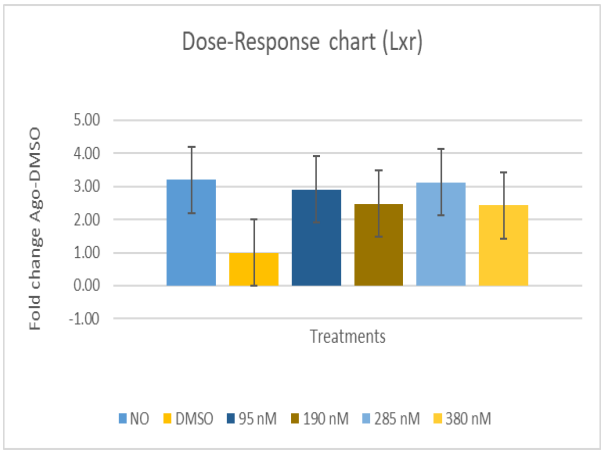
A dosage response experiment was done in the SHK-1 cell line using RT-qPCR methodology in order to examine if Lxr-agonist (GW3965) could trigger the Lxr regulatory pathway, and an optimum dose was selected from this experiment that can triggers the transcriptional regulation of Lxr TF. Lxr-agonist is a well-known ligand to the researcher to mimic the Lxr regulated gene expression by attaching steroid receptor coactivator 1 to human Lxr-α (selleckchem.com). To examine how Lxr agonist administration control the gene expression in SHK-1 cells, we target two TFs related to the lipid metabolism (Lxr, Srebp-1) and a downstream Lxr dependent enzyme (Fas-1) in the *elovl5a* and *elovl5b* genes.

In this work, it was crucial to determine a dosage of Lxr-agonist that maximized gene expression without harming cell lines (SHK-1). Comparing RT-qPCR Cq values measured Lxr, Srebp-1, and Fas-1 expression. Lower Cq values indicate a larger starting number of target RNA molecules in the sample, while higher Cq values suggest the opposite.

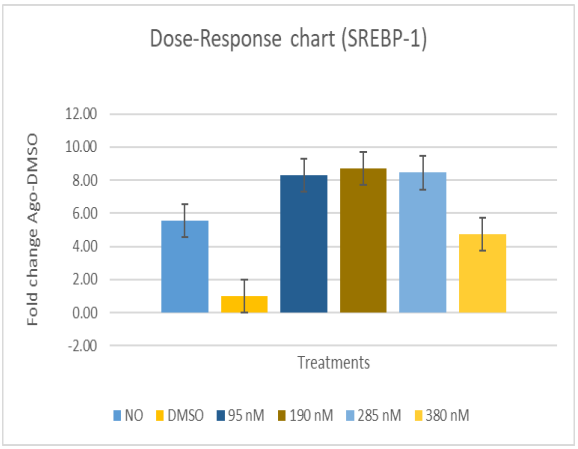
Dose-response graphs for Lxr, Srebp-1, and Fas-1 in the SHK-1 cells are shown in Figure 3.1. Almost similar trend was observed from the Srebp-1 and Fas-1 qPCR graph as observed increased the gene expression, but we don't observe any visual induction of gene expressionin Lxr compared to No treated control samples. DMSO treatment was used as a second negative control to check that the gene expression was not impacted by DMSO treatment itself. DMSO showed lower expression compared to Lxr-agonist treatments and no treated cells. That means DMSO decreased the gene expression. All agonist doses increased gene expression for Srebp-1 and Fas-1 compared to DMSO control and non-treated cells (Figure 3.1). For Srebp-1 we observed nearly 8-8.5-fold gene expression from all the agonist doses compared to DMSO, where non-treated cells showed about 5.8-fold expression. For Fas-1 we observed about 40-, 30-, 35- and 25-fold gene expression respectively from 95nM, 190 nM, 285 nM and 380 nM doses. For the Lxr gene this effect was less pronounced, but the trend was similar, with agonist increasing expression of the Lxr gene. As, 95nM was a small dose of the GW3965 agonist and showed comparatively increased gene expression, we choose 95 Nm to treat the SHK-1 cells and

plasmid constructs after the transfection. As in the results variations (error bar) are quite large, so it's quite unpredictable to give conclusion about the gene expression trend but there is an upregulation in gene expression from the Lxr-agonist. Furthermore, only clear upregulation observed from Srebp-1 and Fas-1 and not from Lxr compared to No treated samples, it makes the gene expression data quite complicated to explain.

A



B



C

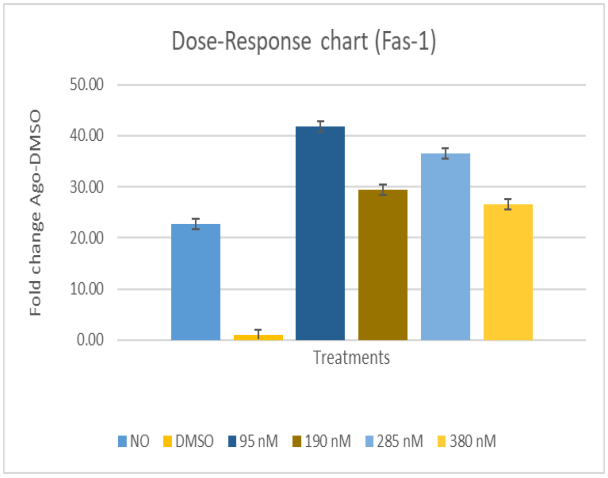
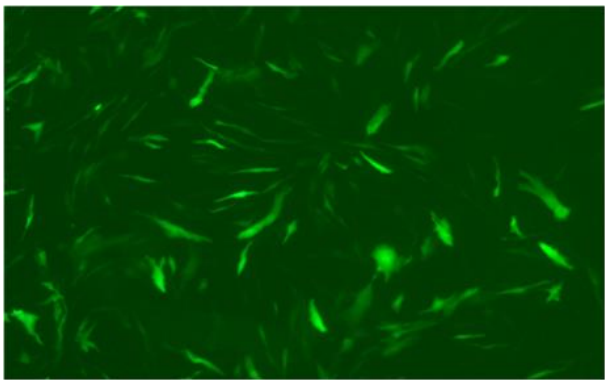


Figure 3.1: The effect of the Lxr-agonist GW3965 on gene expression in SHK-1 cells. Concentrations of 95nM, 190 nM, 285nM, and 380 nM were tested in addition to two negative controls (No agonist nor DMSO treatment = No, and DMSO). A) Effect on Lxr gene expression B) Effect on Srebp-1 expression C) Effect on Fas-1 expression. Along the X-axis plotted the treatments and Y-axis contains fold-change comparing with DMSO.

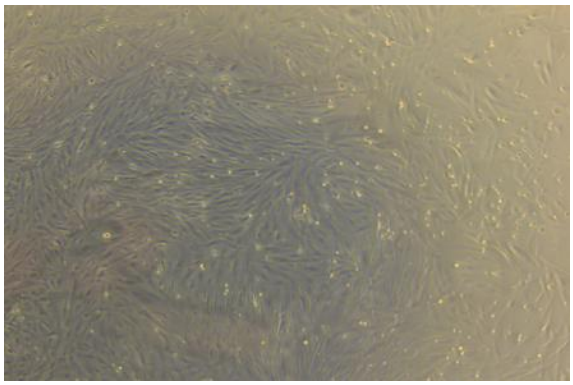
3.2 Effect of LRX induction on elovl5 promoter-activity in the SHK-1 cell line

We hypothesized that tissue expression differences between the two duplicates of the *elovl5* gene are due to differential promoter binding affinity of key liver-centric transcription factors involved in lipid metabolism. Specifically, we wanted to test the prediction that *elovl5b* has a higher liver gene expression compared to *elovl5a* because of differences in promoter binding of the Srebp transcription factors. To test this, we performed luciferase reporter assays with both native promoters of *elovl5* duplicates as well as synthesized promoter sequences containing only the accessible chromatin regions of the promoters (referred to as ATAC-promoters). These constructs were tested with and without Lxr-pathway induction by Lxr-agonist to assess cis-regulatory divergence of elovl5 duplicate promoters.

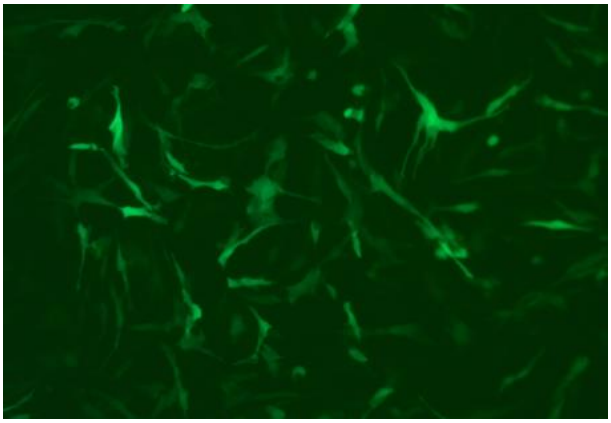
Before starting the luciferase assay, we checked transfection efficiency with co-transfected pEGFP-N1 vector. The quantitative assessment of the GFP in pEGFP-N1 transfected cells was done under the fluorescent and white light and about 30%-40% cells were transfected in Experiment-1, experiment-2 and expreminet-3 (Figure 3.2). All images of pEGFP wells under fluorescent and white light can be found in the appendix Appendix-3.



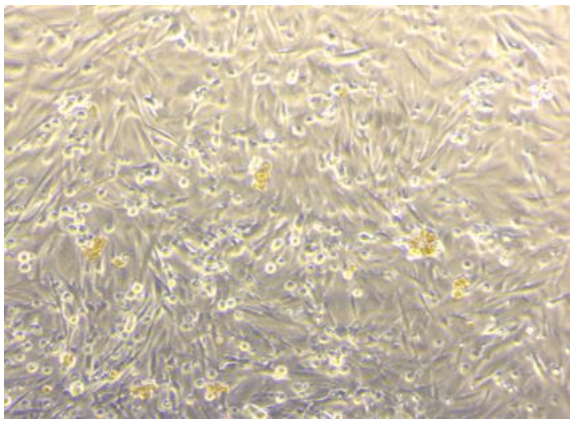
a) pEGFPN1 Well: fluorescent light (Experiment-1)



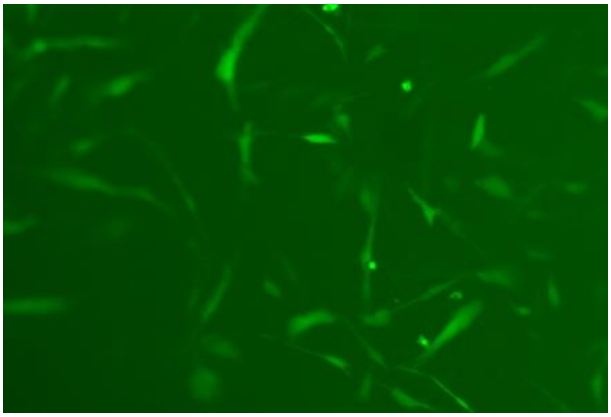
b) pEGFPN1 Well: white light (Experiment-1)



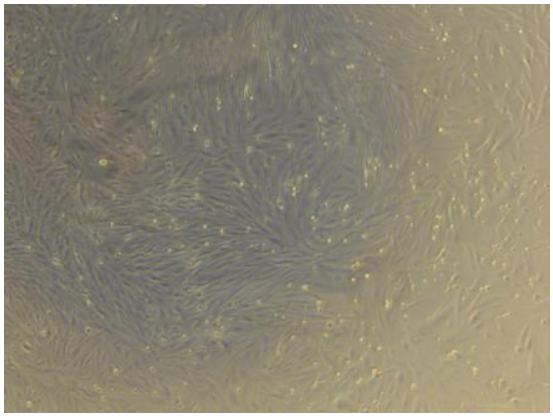
c) pEGFPN1 Well: fluorescent light (Experiment-2)



d) pEGFPN1 Well: white light (Experiment-2)



e) pEGFPN1 Well: fluorescent light (Experiment-3)

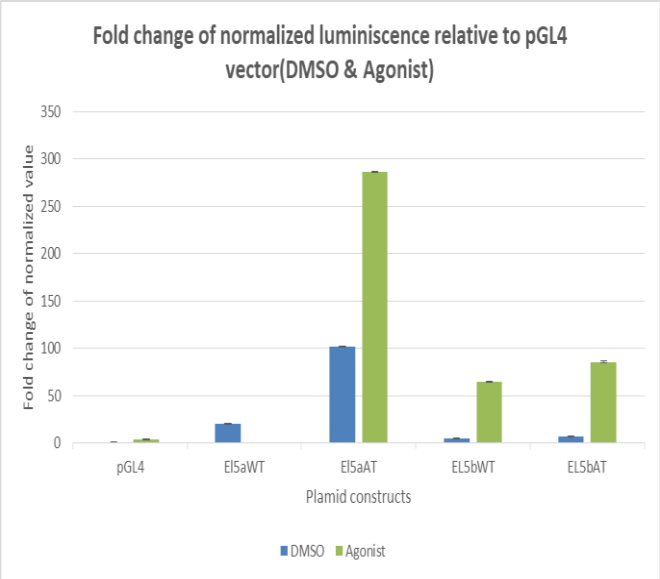


f) pEGFPN1 Well: white light (Experiment-3)

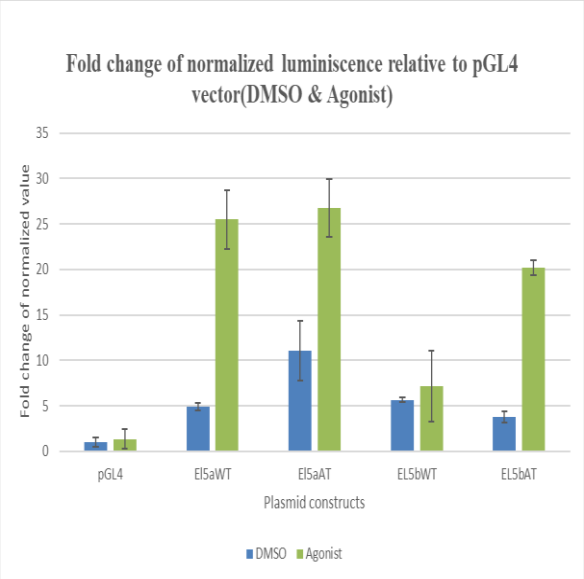
Figure 3.2: The images of transfected SHK-1 cells with pEGFP—N1 vector (in a,c,e) and non-transfected cells (in b,d,f) under fluorescent and white light after 48 hours.

The Srebp-1 regulated gene expression is dependent on the Lxr activity (Carmona-Antonanzas et al., 2016). If Lxr agonist induce Srebp-1 gene expression, and *elovl5* promoters have functional binding sites for the Srebp-1 TF, we expect increased luciferase signals in the Lxr-agonist treated samples. From the Figure 3.3, we observe a general trend in luciferase signals obtained after agonist treatment on SHK-1 cells. We observed increased luciferase signals from all the promoter constructs compared to control sample (pGL4.10. luc2). It indicates that the *elovl5a* and *elovl5b* promoter constructs induced the luciferase gene expression in this SHK-1 cell line. We performed this luciferase reporter assays with both native promoters *of elovl5*a and *elovl5b*

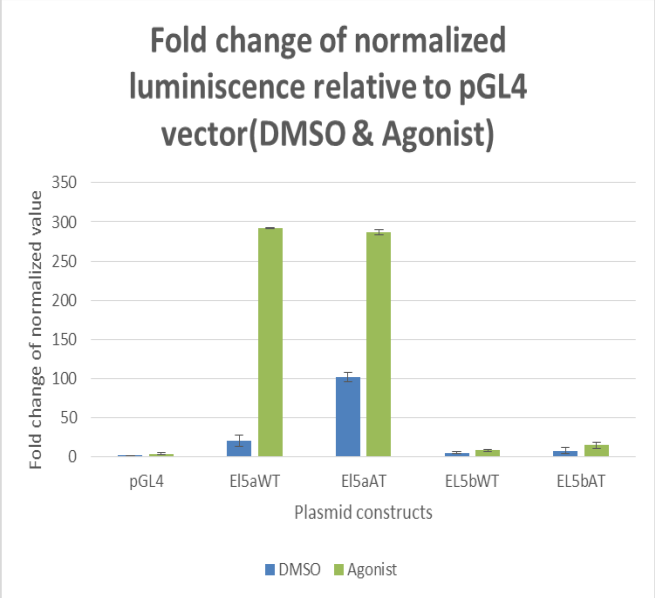
as well as synthesized promoter sequences containing only the accessible chromatin regions of the promoters (referred to as ATAC-promoters). In previous study (Wilberg, 2020), they obtained that *elovl5a* was expressed more than *elovl5b* in liver cells that was opposite of the finding of (Carmona-Antonanzas et al., 2016). To make sure that this difference was not because the native promoter contained regions with suppressors that were in closed chromatin, we constructed these promoters based on the promoter regions found in accessible chromatin only. These results show that all the promoter constructs were induced by the Lxr-agonist compared to DMSO (negative control) (Figure 3.3). But native promoter constructs were comparatively less effective in driving luciferase gene expression compared to the ATAC-promoters. Moreover, constructs with *elovl5a* promoters had consistently higher luciferase signals compared to *elovl5b*. In Figure 3.3 (a) and (c) observe that *elovl5a* ATAC induced about 280-fold gene expression where in graph 3.3 (b) shows it only drive about 27-fold more expression. Additionally, *elovl5a* WT drives about 25-fold more gene expression in Figure 3.3 (b) and about 280-fold in Figure 3.3(c). Finally, we find that transcriptional regulation of constructs *elovl5a* and *elovl5b* promoters was induced by the Lxr- agonist as we get increased luciferase signals from agonist treated cells compare to DMSO treated cells. In Figure 3.3 (a), a green graph bar (agonist treatment) is absent because transfection efficiency was lower in this sample and we did not have RNA to conduct luciferase assay for this well.



a) Luciferase assay- Experiment:1



b) Luciferase assay- Experiment:2



c) Luciferase assay- Experiment:3

Figure 3.3: Results from Dual-Glo luciferase assay. This figure represents three bar graphs from Dual- luciferase assay with same experimental conditions. The data was normalized by internal control vector renilla. Blue bars present negative control (DMSO) and green bars present Agonist treatment. These graphs show the fold change compared to negative control (DMSO) and the gene expression patterns of SHK-1 cells treated with Lxr-agonist and DMSO. In the X-axis plotted the plasmid constructs and Y-axis plotted the fold change normalized value

3.3 CRISPR-Cas9 interrogation of Srebp-1 and NF-Y binding sites in *elovl5b* promoter

Previous studies on the regulation of *elovl5* genes has indicated that Srebp-1, NF-Y TFs are involved in the regulation of these genes, however for the Atlantic salmon *elovl5* genes have only been evaluated experimentally in cell lines from another species (Carmona-Antonanzas et al., 2016). Moreover, these experiments have been based on testing regulatory sequences in vectors which often can yield different results compared to when the same regulatory sequences are found in the native genomic context (Inoue et al., 2017). To this end, we wanted to further interrogate the function of the *elovl5* gene cis-regulatory landscape using a CRISPR-Cas9 based approach.

For the CRISPR-Cas9 knock out experiment, the main targets were Srebp-1 and NY-F TF binding sites in the promoter of the *elovl5b*, as well as the TSS as a control. To assess the importance of these TF binding sites on *elovl5b* regulation we compared RT-qPCR of *elovl5b* between cells not transfected with RNP ("No-RNP") and cells with mutations in the TF binding sites (Srebp-1 and Srebp-2, NY-F). From Figure 3.5, the Sanger sequencing results, we observed that CRISPR knock out was successful in the SHK-1 cell line for all the target TFs binding site in *elovl5b* promoter.

3.4 Evaluation of CRISPR Knockout success by Sanger sequencing:

PCR amplification:

To evaluate the success of knock out of TF binding sites we performed Sanger sequencing of PCR products from CRISPR target regions in cells transfected with and without CRISPR-Cas9 constructs. Gel electrophoresis (Figure 3.4) shows that amplified bands for 12 PCR amplicons (P_{53}-1, P_{53}-2, TSS-1, TSS-2, NF-Y-1, NF-Y-2, S_1-1, S_1-2,S_2-1,S_2-2 respectively) matched length expectations. Here P_{53} amplicon was 850 bp length, so its amplicon was less than 1kb (1000 bp) in the GeneRuler ladder.

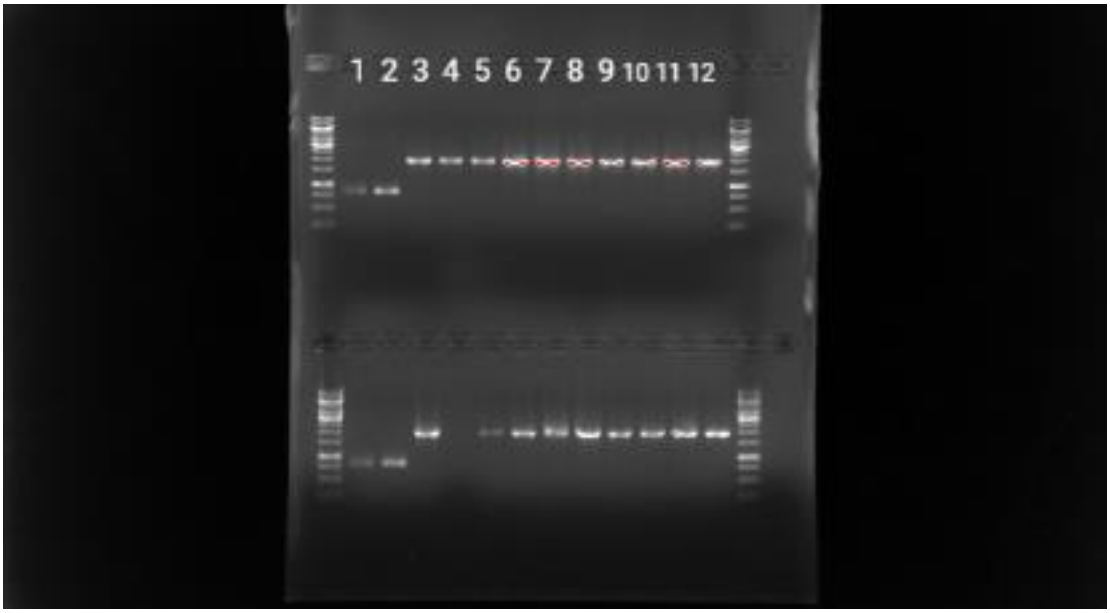


Figure 3.4: 1% Gel electrophoresis run result. Here 12 samples were run on 1% gel electrophoresis to verify the PCR amplicons with 1kb GeneRular ladder. Except P_{53} all other samples were reach to 1kb ladder as expected. P_{53} was less than 1kb so it was in lower position of 1kb ladder. In the Figure well 1and 2 contained P_{53} samples, 3 and 4 had TSS-1 and TSS-2, 5 and 6 had NF-Y-1, NF-Y-2, 7,8,9 and 10 had S_1-1, S_1-2,S_2-1,S_2-2 respectively, 11 and 12 had No RNP-1 and No RNP-2 samples.

Sanger sequencing of the PCR products (Figure 3.5) verified knockout of correct genomic regions. Using Benchling.com we estimated knock out efficiency as well as positional and size distribution of CRISPR-Cas9 induced mutations. From the sequencing result, successful knockout for each specific gRNA was observed compared to the No RNP samples. Only exception was SRE-2-DMSO-1, we did not find any knockout success from this sample nor any signal. It might be result of some technical problem or less amount of DNA sample in the reaction. We also have missing alignment in TSS-DMSO-1, as we did not have successful transfection in these wells during transfection, so we did not have any DNA sample for sequencing. From the Figure 3.5 (a,b and c) it also observed unequally distributed aberrant sequence signal in the knockout site for NFY-1-Agonist-1, S-1-Agonist-1 and TSS-1-Agonist-1 compared to No RNP sequencing that indicate successful knockout. No RNP has equally distributed signals that indicated regular signals and no indel formation. But when knockout started in the gRNA specific samples there induced indel that made the sequence signals unequal in the downstream due to deletion. In the Figure 3.5 we compared the No RNP, NF-Y, S-1 and TSS-1 Sanger sequencing with gRNA reference sequence upside. All the Sanger sequencing alignment could be found in Appendix 1.

a) NF-Y-Agonist-1

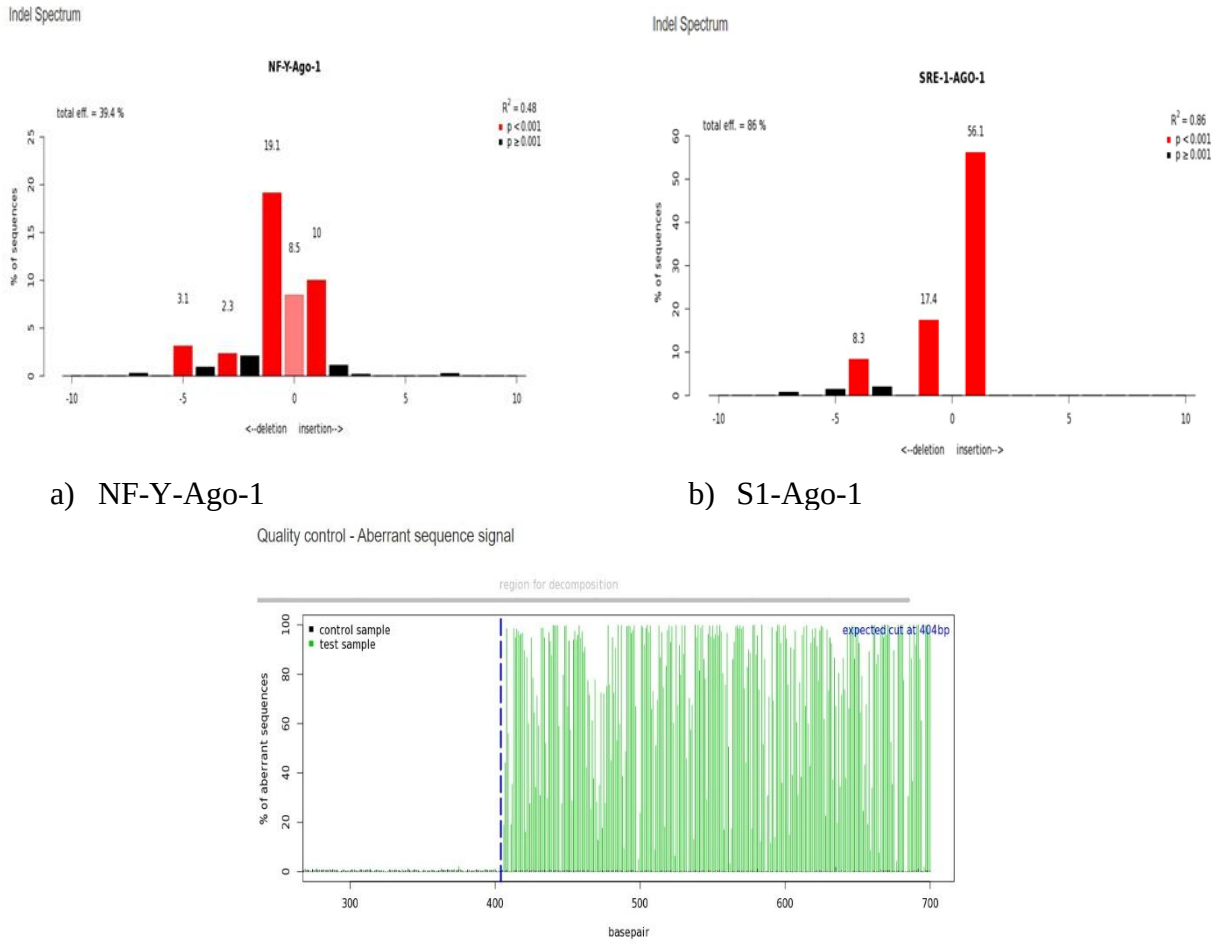
b) S-1-Agonist-1

c) TSS-Agonist-1

Figure 3.5: Example of Sanger sequencing alignment. Alignment was done in Benching.com compared to the *elovl5b* sequence reference sequence (most upper reference sequence). a) presents the Sanger sequence alignment of NY-F-Agonist-1 and b) S-1-Agonist-1, c) TSS-Agonist-1. Here Orange color bars present desired gRNA that we used to target (NF-Y and S-1. TSS). Successful Knockout observed in this gene sequence. The unequally distributed aberration signal represents the knockout site.

The percentage of successful knockout by CRISPR-Cas9 was quantified by comparing the knockout percentage of each sample to the No RNP percentage from Sanger sequencing result. To compare these results used Tide (https://tide.nki.nl/), a web computational tool that can precisely determine the spectrum of mutation generated by CRISPR-Cas9 tool by comparing with the control sequence. It can estimate the frequency of small indels (insertion or deletion) generated by genome editing tools (CRISPR-Cas9, TALENs and ZFNs). All the figures of Tide (https://tide.nki.nl/) with indel frequency graph and quality control windows will be found in Appendix 2. Although some example figures are also presents in Figure 3.5. In this figure quality-control window is the scale to verify that the sequences were of good quality or not.

There is a breaksite (the dot line in the Figure 3.5 c) line in the quality-control window, that is the expected target site. In the good quality sequences the control sample (black) would be a low and equally distributed aberrant sequence signal whereas in the test sample (green) it would be a low signal before the breaksite and a higher signal downstream of the breaksite (https://tide.nki.nl/). From quality-control graph we observed the same as expected from the good quality sequences.



a) NF-Y-Ago-1

b) S1-Ago-1

c) S1(Srebp-1)-Ago-1 Quality control graph

Figure 3.5: Frequency and quality graph from Tide (https://tide.nki.nl/). a) shows the indel frequency of NF-Y-Ago-1 sample, here left site presents indel frequency by deletion and right site is insertion. Red bars are the value less than 0.001 and black bars are presenting value higher than 0.001. b) shows indel spectrum from S1-Ago-1 samples. c) presents Quality-control graph for S1-Ago-1 samples. Here we got low and equally distributed aberrant signal from black line (control) and higher signal downstream from green line (Test sample), that indicates good quality sequence.

A graph is formulated from the percentage of the successful CRISPR knockout (Figure 3.6), that shows that NF-Y (both Agonist and DMSO) got about 38% indel frequency and 19.1 % significant indel frequency was obtained by the deletion (Appendix 2). About 81% and 85% indel frequency were obtained from SRE1(Srebp-1) (both Agonist and DMSO) and SRE-2 (Srebp-2) (both Agonist and DMASO) respectively, that is the highest indel frequency among all

the samples. Here most of the indel frequency were obtained from insertion but deletion also contributed to give indel spectrum (Appendix 2). TSS also shows same trend in the indel spectrum as SRE-1(Srebp-1) and SRE-2 (Sprebp-2) that is about 80-81%.

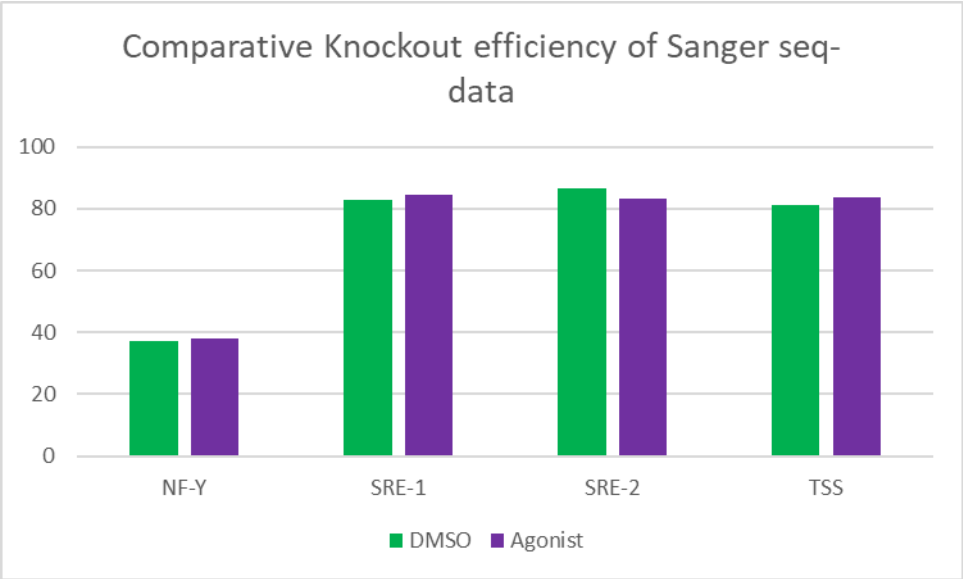


Figure 3.6: Comparative knockout efficiency of Sanger sequencing data. Green graphs are DMSO, purples are Agonist sample bar. From the sequencing result observed about 40% knockout success for NF-Y site and about almost 80% success for SRE-1 (Srebp-1), SRE-2 (Sprebp-2) and TSS site.

Moreover, from the RT-qPCR result (Figure 3.7) of this knock out experiment, it is observed that the most successful knock out were observed on NY-F and Srebp-2 target sites, then TSS and Srebp-1 were most successful respectively. Here Blue bars represents DMSO treated cells data and Orange is for Agonist treated. After transfection process, we treated the cells with Lxr-agonist and DMSO to check that after knocking out of TF binding sites is these ligands able to induce gene expression or not in *elovl5b* promoter. As we observed that the gene expression is downregulated by the knockout, it can be assumed that these TF binding sites are responsible for the transcriptional regulation in the *elovl5b* gene. We plot the $2^{\wedge-\Delta\Delta CT}$ value on the y-axis and the treatment along the x-axis. $2^{\wedge-\Delta\Delta CT}$ value was calculated after normalizing the data by a house keeping gene (elf-α) and comparing with the negative control DMSO.

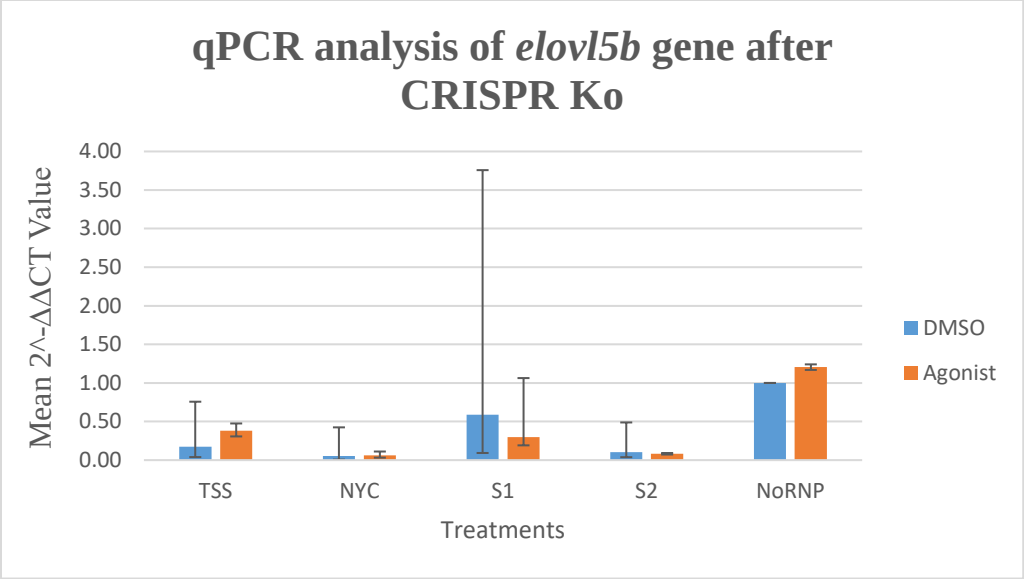


Figure 3.7: RT-qPCR analysis of elovl5 gene in cells with and without CRISPR-Cas9 knock out of TF binding sites. Blue bars present 2^-ΔΔCT value of DMSO and Orange bars for Lxr-Agonist (GW3965). In the qPCR data the gene expression quantifies on the basis of cq value. Less cq value means present of more RNA in the samples where more cq value is of its opposite. If gene expression increased due to certain treatment qPCR will give less cq value and vice-versa. Here error bars were calculated on the basis of upper and lower $2^{\wedge-\Delta\Delta CT}$ value by adjusting with SD values. X-axis contains Treatments and Y-axis contains $2^{\wedge-\Delta\Delta CT}$ value.

4. Discussion:

To acquire a comprehensive picture of the WGD event in the salmonoid family, the *elovl5* gene of the Atlantic salmon has been investigated extensively (Macqueen & Johnston, 2014). Due to differences in the binding pattern of key transcription factors, two duplicates, *elovl5a* and *elovl5b* are expressed differently in various tissues of Atlantic salmon. The aim of this study is to assess which TFs are responsible with this differential expression in these two duplicates as well as the transcriptional regulation of specifically the *elovl5b* gene. Luciferase reporter assay and CRISPR-Cas9 knockout technology has been using to evaluate Lxr induced gene expression with some *elovl5a* and *elovl5b* promoter constructs in SHK-1 cell line. Promoter-reporter assay and CRISPR-Cas9 knock out were two methods used to get insight about the regulatory pathway of these two duplicates.

We will begin by discussing the technical aspects of the studies, and then move on to the analysis of the promoter-reporter assay, the dose response experiment of the Lxr agonist (GW3965), the transfection efficiency, and the RT-qPCR results. The results of RT-qPCR and Sanger sequencing experiments that performed after the knockout of core TF binding sites in *elovl5b* will be discussed in the next section.

4.1 Technical standpoint of the experiments:

In this section, some technical errors will be discussed with some possible explanations. Firstly, in the transfection results, some of the renilla and luciferase values seemed to be quite low compared to other wells. As we have used same number of cells and experimental conditions in the transfection and luciferase assays, there should be similar values from each well. There was probably an issue with the experiment during transfection of SHK-1 cells. It is possible that the amount of solutions added to each sample was incorrectly pipetted, which might account for different values from same sample. To normalize the difference in transfection efficiency we co-transfected with a vector containing renilla luciferase as an internal control.

Another concern was the differential tissue-specific expression of *elovl5* duplicates was recorded in salmon liver cells; however, in this thesis, we employed SHK-1 cells instead of liver cells to study the entire experiment (head kidney cells of salmon). Use of liver cells might be more optimistic but working with this kind of cells is very difficult and transfection efficiency can vary due to differences in proportions of other cell types such as red blood cells (Wilberg, 2020). Hence, we used the SHK-1 cell line as it was available to us and suitable to study gene regulation in-vitro. Several studies have been conducted using SHK-1 cell line to explore the gene regulatory pathway of *elovl5* gene (Carmona-Antoñanzas et al., 2014; Minghetti et al., 2011).

In the reporter promoter test, we have performed three assays using the identical experimental settings, however in the third assay, we did not receive any decent luciferase signals from the *elovl5a* WT construct because of low transfection efficiency. In addition, the TSS-DMSO-1 well failed to provide any knockdown results during the sanger sequencing. However, we were able to obtain data from another TSS well (experimental replicate) for further research. This is most likely due to a technical fault in the experiment like improper transfection of the RNP complex into the cells. But we obtain successful knockout from other replicate.

4.2 Lxr-agonist (GW3965) induced gene expression in SHK-1 cell for transcription binding site in *elovl5a* and *elovl5b* promoters:

Lxr-agonist dosage to induce gene expression:

In this research, gene regulation in SHK-1 cells was successfully induced by Lxr-agonist as observed by upregulation in expression of Srebp-1 and Fas-1. Previous research (Carmona-Antoñanzas et al., 2014) indicated that GW3965 was a suitable ligand to initiate this pathway, hence it was used to investigate the roles of Lxr and other transcription factors (Srebp-1 and Srebp-2, NF-Y) in lipid metabolism. In the Carmona-Antoñanzas et al., 2014 they used high doses of GW3965 (10 μM), but we used lower doses in our study (95 nM, 190 nM, 285nM and 380 nM) to ensure cell viability. From the dose response experiment we found that all the doses were able to induce the gene expression but 95 nM was more promising to induce expression consistently for all TFs (Lxr, srebp-1) and Fas-1 that is in line with the finding of (Collins et al., 2002), they observed upregulation in gene expression with 190 nM and 30 nM of Lxr-agonist doses. Hence 95 nM was the preferred dose for rest of the experiment.

Comparison of gene expression between two duplicates of *elovl5* gene:

Experiments with the SHK-1 cell line using promoter constructs for *elovl5a* and *elovl5b* have shown differential gene regulation between two copies of *elovl5*. Generally, in three-reporter assays, all promoter constructs increased the luciferase gene expression with some differences in the degree of fold changes. In the previous studies *elovl5b* shower higher expression in liver and brain than it's ohnologue *elovl5a* (Morais et al., 2009a). Interestingly our study *elovl5a* produced higher luciferase expression than the *elovl5b* in the SHK-1 cells. Some technical error with the experiment might have occurred, such as pipetting error. Another possible reason could be that there might be other transcription factor binding site or enhancers outside of that promoter region in *elovl5b* that we omitted in our constructs for this gene. It is evident in a previous study that over a vast chromosomal area, enhancers have the ability to activate many neighboring genes (West & Fraser, 2005). Sometimes gene regulation can be different in the ohnologues due to divergence in the intronic regions that situated within the transcribed region and control the gene regulation in concert of other machineries like splicing, transcription, polyadenylation,

mRNA export (Rose, 2019). Moreover, the presence of an enhancer element or sequences that accelerate translation may be found in certain introns (Akua & Shaul, 2013; Kim et al., 2006). Our finding however in agreement with (Wilberg, 2020) where they also found higher gene expression in the *elovl5a* synthetic promoter in SHK-1 cells. Another possible explanation could be the gene regulation is different in SHK-1 cells than the hepatic cells for these two duplicates, as (Sahlström, 2021) obtained higher expression from the *elovl5b* in hepatic cells. Although there were some technical and experimental differences among these studies, for example (Sahlström, 2021) only studied the first ATAC peak region of the *elovl5b* gene, whereas we compared the Lxr pathway only for Srebp-1, Srebp-2, NF-Y and TSS site in *elovl5a and elovl5b*. Hence, the reason for differences in gene expression between these two studies cannot be determined without further investigation.

TF binding site in the *elovl5a* and *elovl5b* promoter:

It also observed from the RT-qPCR result that Lxr-agonist induced the Lxr, Srebp-1 regulated gene expression including Fas-1, but it showed comparatively less upregulation for Lxr pathway. From this point of view, it can be explained that there could be some technical issues when setting for RT-qPCR experiment. Even though we utilized the recommended amount of RNA for RT-qPCR based on the manufacturer's instructions, the amount used in the setup might have varied due to the handling purpose. In the previous study Srebp-1 is observed as a direct target of Lxr directed regulation (Repa et al., 2000) and we observed an upregulation of Srebp-1 and Fas-1 compared to control. Though in the results there were large variation among the data it can be assumed that Lxr-agonist has induced the Lxr regulated gene regulation. Hence, the combined observation of dose response experiment and reporter-promoter luciferase assay is that possibly Lxr, Srebp-1 are related to the gene regulation of *elovl5a* and *elovl5b* gene. This is in agreement with the finding of Carmona-Antonanzas et al., 2016. We find involvment of some more transcription factors from the CRISPR-Cas9 experiment in *elovl5a* and *elovl5b* gene regulation that will be discuss in next section.

4.3 Study of the regulatory elements of *elovl5b* gene by CRISPR-Cas9 based approach:

The aim of the CRISPR-Cas9 knockout experiment in this thesis was to get insight about the regulatory regions of transcription factor binding sites within the *elovl5b* gene promoter. The results of RT-qPCR and Sanger sequencing suggested that due to knockout of the main transcription factor binding sites (TFBS), gene expression of *elovl5b* was downregulated as expected. The primary regulators of lipogenesis are the Lxr, srebp-1, and Srebp-2, which are regulated by dietary sterols and fatty acids. (Espenshade, 2006). CRISPR-Cas9 experiments reveal that Srebp-1 and Srebp-2 play a role in *elovl5b* gene regulation. The expression of the *elovl5b* gene was reduced in comparison to control samples after Srebp-1 and Srebp-2 were knocked out. Furthermore, in the knockout experiment, we also targeted co-factor named NF-Y

site, which is close to the Srebp-1 and Srebp-2 areas and reduced gene expression after deletion. In the knockout, the highest downregulation was obtained from NF-Y and Srebp-1 knockout site. Srebp-2 knockout showed more downregulation than srebp-1. It was also evident that Sterol regulatory element-1 (SRE) and NF-Y regulated activity is necessary for the highest levels of Srebp-dependent activity (Carmona-Antoñanzas et al., 2014) and we observed less gene expression from both Srebp1, Srebp-2 and NF-Y knockout samples. In our experimental data there was a high variability, so it is not possible to conclude anything here, but the gene expression was reduced after knockout the main TFBS in all samples. The transcription starting site was targeted as the positive control and it also showed down regulation in the *elovl5b* expression but still showed some stimulation (2^-Δ ΔCq value was 0.38). It might be that after knockout of the TSS there were still other TFBS present upstream, that caused some stimulation for agonist treatment.

However, the previous study also found the same observation by the mutation of NF-Y, Srebp1 and Srebp-2 in *elovl5b* gene (Carmona-Antonanzas et al., 2016). They obtained less Srebp-1 and Srebp-2 dependent activity due to mutation in the *elovl5b,* where Srebp-2 regulated agonist stimulation was more prominent than Srebp-1. There were some experimental differences between the two studies. For example, they did all the work using plasmid constructs in a different species of fish (Fathead minnow, FHM cells), whereas we used salmon cells and directly mutated the genomic DNA by CRISPR-Cas9. However, our finding is in line with the findings of (Carmona-Antoñanzas et al., 2014).

We planned the CRISPR knockout in only *elovl5b* gene as this gene showed upregulation in the liver and we aimed to find out that which cis-regulatory regions are responsible for this upregulation. It would be interesting to compare the cis-regulatory regions of *elovl5a* gene in the same way, but due to short time constrain we could only test the *elovl5b* gene. Carmona-Antonanzas et al., 2016 observed that mutation in the TFBS in *elovl5a* gene showed no significant change in Srebp-1 and Srebp-2 dependent gene regulation. Because of large variation in our data it is not possible to conclude that the Srebp-1 and NF-Y dependent divergence causes tissue-specific differences between two copies of *elovl5* gene, but our data show that SREBP-1 and NF-Y are likely responsible for the Lxr mediated transcriptional control in the *elovl5b* gene.

Sanger sequencing also helps us to interprate the qPCR data. If knockout efficiency is low then we would expect less gene expression in qPCR data because fewer of the cells are modified. In Sanger sequencing Srebp-1 and Srebp-2 showed almost same knockout efficiency (Figure-3.6). Furthermore, for both Srebp-1 and Srebp-2 obtained more than 80% knockout efficiency where for NF-Y site obtained about 40%. From the RT-qPCR finding (Figure- 3.7) we observed less gene expression from NF-Y and Srebp-2, while the Lxr-agonist treated TSS and Srebp-1 obtained little stimulation in gene expression (The 2^-Δ ΔCq value were about 0.38 and 0.30 respectively). Though Srebp-1 showed good knockout efficiency in Sanger sequencing, but in RT-qPCR data it showed some stimulation in gene regulation. This finding is bit contradictory to each other. It might be due to some experimental error during RT-qPCR. This result could not be happened due to off target effect because we extracted RNA and DNA from the identical samples following the knockout experiment to carry out RT-qPCR and Sanger sequencing.

However, we observed an overall trend of downreagulation in gene expression due to knockout of Srebp and NF-Y dependent TFBS in this experiment. Other studies have shown that Srebp is a main factor in the LC-PUFAs regulatory pathway that influenced by it's nearest co-factor NF-Y to upregulate the gene expression (Carmona-Antonanzas et al., 2016; Espenshade, 2006). From the promoter-reporter assay we observed upregulation in gene expression from the Srebp-1 site and also from CRISPR-Cas9 knockout findings, so it can be said that possibly Srebp-1 is involved in the LC-PUFA pathway, though we can not conclude that from our data due to large variation that it is directly involced in this pathway. However, we observed less gene expression after knockout of Srebp-1, Srebp-2 and NF-Y site that is in line with (Carmona-Antonanzas et al., 2016) finding.

5. Conclusion:

The objective of this thesis was to investigate the Lxr induced gene regulation in two ohnologues of the *elovl5* gene. Srebp-dependent divergence was identified in previous research (Carmona-Antonanzas et al., 2016) that was thoroughly examined using plasmid-based reported assays in non-salmonoid cells. In our study, we observed upregulation in gene expression for *elovl5a* and *elovl5b* promoter constructs in SHK-1 cells by applying Lxr-agonist. This demonstrates the relationship between Lxr and other dependent TFs to lipid metabolism of Atlantic salmon. Moreover, we observed lower gene expression from the *elovl5b* gene in the cells compared to the *elovl5a,* that was not in agreement with some previous findings, but was similar to Wilberg, R. 2020. Large variation in the data that we observed in the knockout experiment makes it difficult to draw conclusions about Lxr dependent gene regulation in *elovl5b*. We did however observe a trend of lower gene expression from Lxr mediated regulation when major TF binding sites in *elovl5b* promoters were knocked out. Though large variation among the data makes it impossible to give a definitive conclusion, the overall trend was in line with previous studies (Carmona-Antonanzas et al., 2016). However, these experiments should be repeated to confirm the results. Furthermore, an experiment can be performed to knockout of the core binding sites of *elovl5a* and *elovl5b* gene in parallel to demonstrate the Srebp dependent regulation in these two ohnologues. Interestingly in this thesis we were successful with the knockout protocol in the SHK-1 cells using an RNP directed CRISPR-Cas9 approach. This will guide future functional studies using similar approaches in this cell line.

References:

Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nature Communications*, 9 (1): 1911. doi: 10.1038/s41467-018-04252-2.

Akua, T. & Shaul, O. (2013). The Arabidopsis thaliana MHX gene includes an intronic element that boosts translation when localized in a 5′ UTR intron. *Journal of experimental botany*, 64 (14): 4255-4270.

Allendorf, F. W., Bassham, S., Cresko, W. A., Limborg, M. T., Seeb, L. W. & Seeb, J. E. (2015). Effects of crossovers between homeologs on inheritance and population genomics in polyploid-derived salmonid fishes. *Journal of Heredity*, 106 (3): 217-227.

Andersson, R. & Sandelin, A. (2020). Determinants of enhancer and promoter activities of regulatory elements. *Nature Reviews Genetics*, 21 (2): 71-87.

Branchini, B. R., Southworth, T. L., Fontaine, D. M., Kohrt, D., Florentine, C. M. & Grossel, M. J. (2018). A firefly luciferase dual color bioluminescence reporter assay using two substrates to simultaneously monitor two gene expression events. *Scientific reports*, 8 (1): 1-7.

Carmona-Antonanzas, G., Zheng, X., Tocher, D. R. & Leaver, M. J. (2016). Regulatory divergence of homeologous Atlantic salmon elovl5 genes following the salmonid-specific whole-genome duplication. *Gene*, 591 (1): 34-42.

Carmona-Antoñanzas, G., Monroig, Ó., Dick, J. R., Davie, A. & Tocher, D. R. (2011). Biosynthesis of very long-chain fatty acids (C> 24) in Atlantic salmon: Cloning, functional characterisation, and tissue distribution of an Elovl4 elongase. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 159 (2): 122-129.

Carmona-Antoñanzas, G., Tocher, D. R., Martinez-Rubio, L. & Leaver, M. J. (2014). Conservation of lipid metabolic gene transcriptional regulatory networks in fish and mammals. *Gene*, 534 (1): 1-9. doi: 10.1016/j.gene.2013.10.040.

Chatterjee, S. & Ahituv, N. (2017). Gene regulatory elements, major drivers of human disease. *Annual review of genomics and human genetics*, 18: 45-63.

Collins, J. L., Fivush, A. M., Watson, M. A., Galardi, C. M., Lewis, M. C., Moore, L. B., Parks, D. J., Wilson, J. G., Tippin, T. K. & Binz, J. G. (2002). Identification of a nonsteroidal liver X receptor agonist through parallel array synthesis of tertiary amines. *Journal of medicinal chemistry*, 45 (10): 1963-1966.

Davidson, W. S., Koop, B. F., Jones, S. J., Iturra, P., Vidal, R., Maass, A., Jonassen, I., Lien, S. & Omholt, S. W. (2010). Sequencing the genome of the Atlantic salmon (Salmo salar). *Genome biology*, 11 (9): 1-7.

Dehler, C. E., Boudinot, P., Martin, S. A. & Collet, B. (2016). Development of an efficient genome editing method by CRISPR/Cas9 in a fish cell line. *Marine Biotechnology*, 18 (4): 449-452.

Des Marais, D. L. & Rausher, M. D. (2008). Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature*, 454 (7205): 762-765.

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.

Espenshade, P. J. (2006). SREBPs: sterol-regulated transcription factors. *Journal of cell science*, 119 (6): 973-976.

Gillard, G., Harvey, T. N., Gjuvsland, A., Jin, Y., Thomassen, M., Lien, S., Leaver, M., Torgersen, J. S., Hvidsten, T. R. & Vik, J. O. (2017). Life stage associated remodeling of lipid metabolism regulation in the duplicated Atlantic salmon genome. *bioRxiv*: 140442.

Haberle, V. & Stark, A. (2018). Eukaryotic core promoters and the functional basis of transcription initiation. *Nature reviews Molecular cell biology*, 19 (10): 621-637.

Horton, J. D., Goldstein, J. L. & Brown, M. S. (2002). SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *The Journal of clinical investigation*, 109 (9): 1125-1131.

Houston, R. & Macqueen, D. (2019). Atlantic salmon (Salmo salar L.) genetics in the 21st century: taking leaps forward in aquaculture and biological understanding. *Animal Genetics*, 50 (1): 3-14.

Inoue, F., Kircher, M., Martin, B., Cooper, G. M., Witten, D. M., McManus, M. T., Ahituv, N. & Shendure, J. (2017). A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity. *Genome Res*, 27 (1): 38-52. doi: 10.1101/gr.212092.116.

Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C. & Bernot, A. (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature*, 431 (7011): 946-957.

Jakobsson, T., Treuter, E., Gustafsson, J.-Å. & Steffensen, K. R. (2012). Liver X receptor biology and pharmacology: new pathways, challenges and opportunities. *Trends in pharmacological sciences*, 33 (7): 394-404.

Kalaany, N. Y. & Mangelsdorf, D. J. (2006). LXRS and FXR: the yin and yang of cholesterol and fat metabolism. *Annu. Rev. Physiol.*, 68: 159-191.

Kim, M. J., Kim, H., Shin, J. S., Chung, C.-H., Ohlrogge, J. B. & Suh, M. C. (2006). Seed-specific expression of sesame microsomal oleic acid desaturase is controlled by combinatorial properties between negative cis-regulatory elements in the SeFAD2 promoter and enhancers in the 5'-UTR intron. *Molecular Genetics and Genomics*, 276 (4): 351-368.

Kim, T. K. & Eberwine, J. H. (2010). Mammalian cell transfection: the present and the future. *Analytical and bioanalytical chemistry*, 397 (8): 3173-3178.

Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., Hvidsten, T. R., Leong, J. S., Minkley, D. R. & Zimin, A. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, 533 (7602): 200-205.

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.

Lynch, V. J. (2007). Inventing an arsenal: adaptive evolution and neofunctionalization of snake venom phospholipase A2 genes. *BMC evolutionary biology*, 7 (1): 1-14.

Macqueen, D. J. & Johnston, I. A. (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proceedings of the Royal Society B: Biological Sciences*, 281 (1778): 20132881.

Magadum, S., Banerjee, U., Murugan, P., Gangapur, D. & Ravikesavan, R. (2013). Gene duplication as a major force in evolution. *Journal of genetics*, 92 (1): 155-161.

Minghetti, M., Leaver, M. J. & Tocher, D. R. (2011). Transcriptional control mechanisms of genes of lipid and fatty acid metabolism in the Atlantic salmon (Salmo salar L.) established cell line, SHK-1. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1811 (3): 194-202.

Morais, S., Monroig, O., Zheng, X., Leaver, M. J. & Tocher, D. R. (2009a). Highly Unsaturated Fatty Acid Synthesis in Atlantic Salmon: Characterization of ELOVL5- and ELOVL2-like Elongases. *Marine Biotechnology*, 11 (5): 627-639. doi: 10.1007/s10126-009-9179-0.

Morais, S., Monroig, O., Zheng, X., Leaver, M. J. & Tocher, D. R. (2009b). Highly unsaturated fatty acid synthesis in Atlantic salmon: characterization of ELOVL5-and ELOVL2-like elongases. *Marine Biotechnology*, 11 (5): 627-639.

Otto, S. P. (2007). The evolutionary consequences of polyploidy. *Cell*, 131 (3): 452-462.
Repa, J. J., Liang, G., Ou, J., Bashmakov, Y., Lobaccaro, J. M., Shimomura, I., Shan, B., Brown, M. S., Goldstein, J. L. & Mangelsdorf, D. J. (2000). Regulation of mouse sterol regulatory element-binding protein-1c gene (SREBP-1c) by oxysterol receptors, LXRalpha and LXRbeta. *Genes Dev*, 14 (22): 2819-30. doi: 10.1101/gad.844900.
Reschly, E. J., Ai, N., Welsh, W. J., Ekins, S., Hagey, L. R. & Krasowski, M. D. (2008). Ligand specificity and evolution of liver X receptors. *The Journal of steroid biochemistry and molecular biology*, 110 (1-2): 83-94.
Robertson, F. M., Gundappa, M. K., Grammes, F., Hvidsten, T. R., Redmond, A. K., Lien, S., Martin, S. A., Holland, P. W., Sandve, S. R. & Macqueen, D. J. (2017). Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. *Genome biology*, 18 (1): 1-14.
Rose, A. B. (2019). Introns as gene regulators: a brick on the accelerator. *Frontiers in genetics*, 9: 672.
Roy, B., Zhao, J., Yang, C., Luo, W., Xiong, T., Li, Y., Fang, X., Gao, G., Singh, C. O. & Madsen, L. (2018). CRISPR/Cascade 9-mediated genome editing-challenges and opportunities. *Frontiers in Genetics*, 9: 240.
Rui, E. (2018). *Insights into the regulation of LC-PUFA biosynthesis in wild and aquaculture Atlantic salmon in response to a vegetable oil diet*: Norwegian University of Life Sciences, Ås.
Sahlström, H. (2021). Transposable Elements as a Source of Novel Gene Regulation in Atlantic Salmon (Salmo salar).
selleckchem.com, A. a. Available at: https://www.selleckchem.com/products/gw3965.html#s_ref.
Shlyueva, D., Stampfel, G. & Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. *Nature Reviews Genetics*, 15 (4): 272-286.
sigma-aldrich. Available at: https://www.sigmaaldrich.com/NO/en/product/sigma/luc1?gclid=CjwKCAjw9NeXBhAMEiwAbaY4lqBzCfXnWoM-r22FE_5S0HPjbDCykEo3sQLZTCFM3QriDxLW8qQHZRoCaSkQAvD_BwE.
sigma-aldrich. Available at: https://www.sigmaaldrich.com/NO/en/product/sigma.
Simopoulos, A. P. (2008). The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Experimental biology and medicine*, 233 (6): 674-688.
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): e00217-19.
Teichmann, S. A. & Babu, M. M. (2004). Gene regulatory network growth by duplication. *Nature genetics*, 36 (5): 492-496.
Van de Peer, Y., Maere, S. & Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics*, 10 (10): 725-732. doi: 10.1038/nrg2600.
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.
West, A. G. & Fraser, P. (2005). Remote control of gene transcription. *Human molecular genetics*, 14 (suppl_1): R101-R111.
Wilberg, R. (2020). *Optimization of transfection of primary hepatocytes from Atlantic salmon for functional studies*: Norwegian University of Life Sciences, Ås.
Zhang, D., Hussain, A., Manghwar, H., Xie, K., Xie, S., Zhao, S., Larkin, R. M., Qing, P., Jin, S. & Ding, F. (2020). Genome editing with the CRISPR-Cas system: an art, ethics and global regulatory perspective. *Plant biotechnology journal*, 18 (8): 1651-1669.

Appendixes

Appendix 1: Sanger sequencing alignment comparing to control sequence (No RNP) (Section 4.6):

NF-Y-1 Agonist

NF-Y-2 Agonist

NF-Y-1 DMSO

NF-Y-2 DMSO

SRE-1-Agonist-1

SRE-1-Agonist-2

SRE-1-DMSO-1

SRE-1-DMSO-2

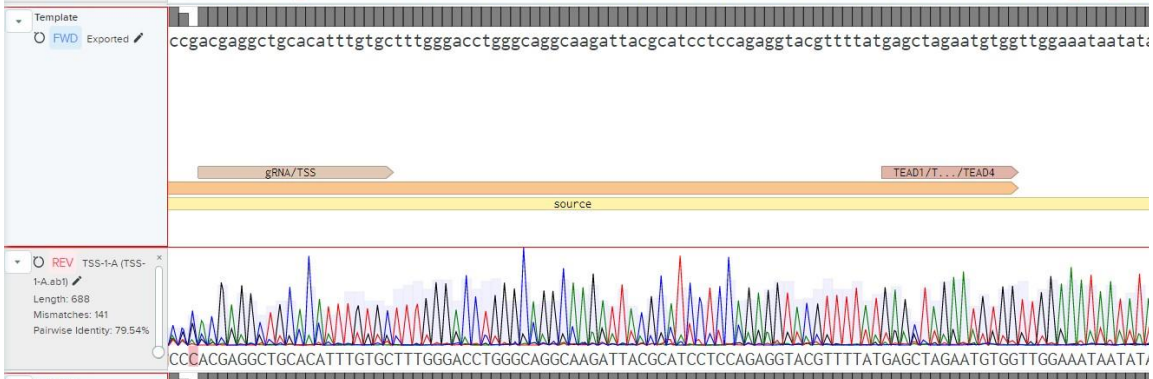
SRE-2-Agonist-1

SRE-2-Agonist-2

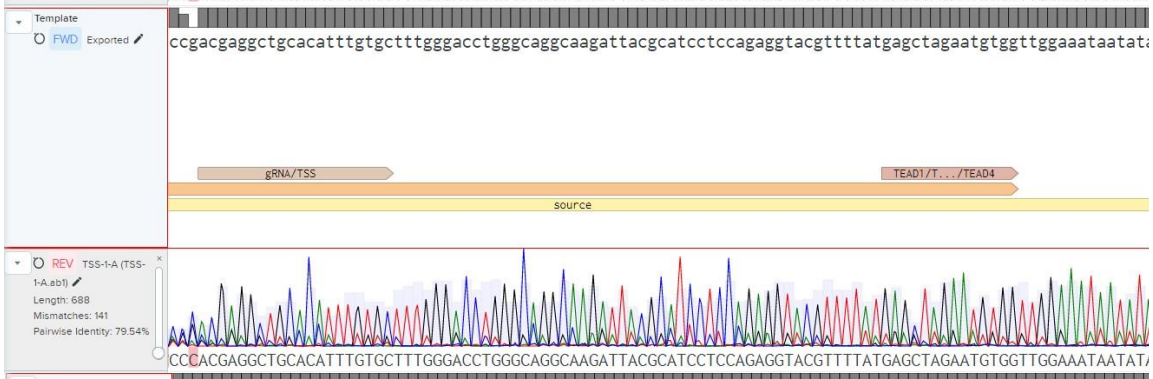
SRE-2-DMSO-1

SRE-2-DMSO-2

TSS-Agonist-1

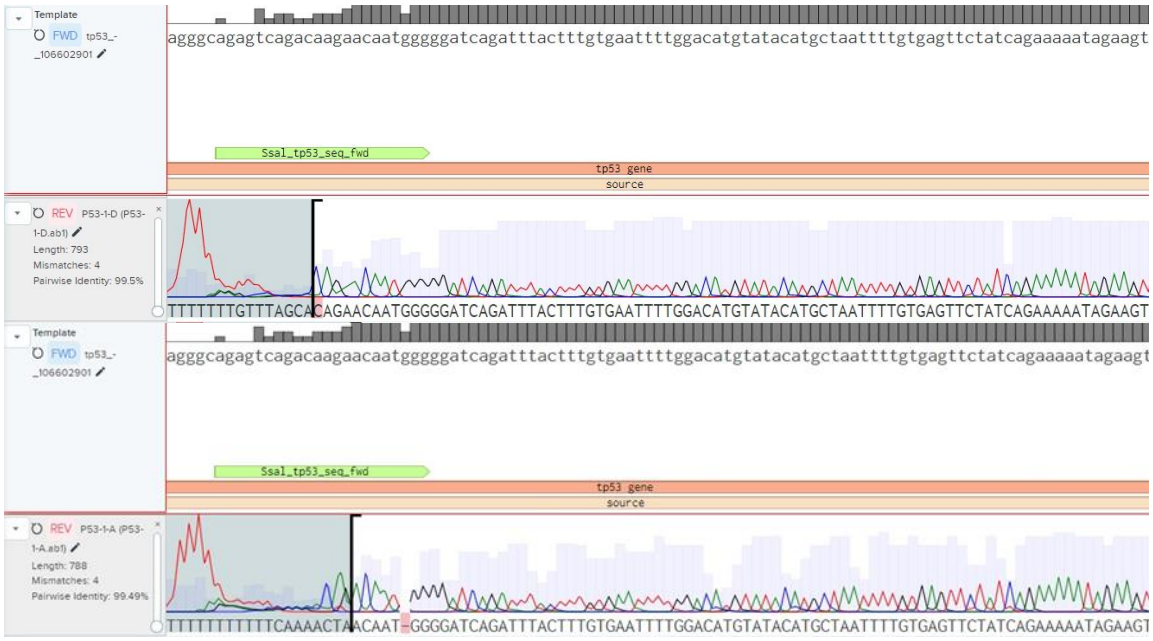


TSS-Agonist-2



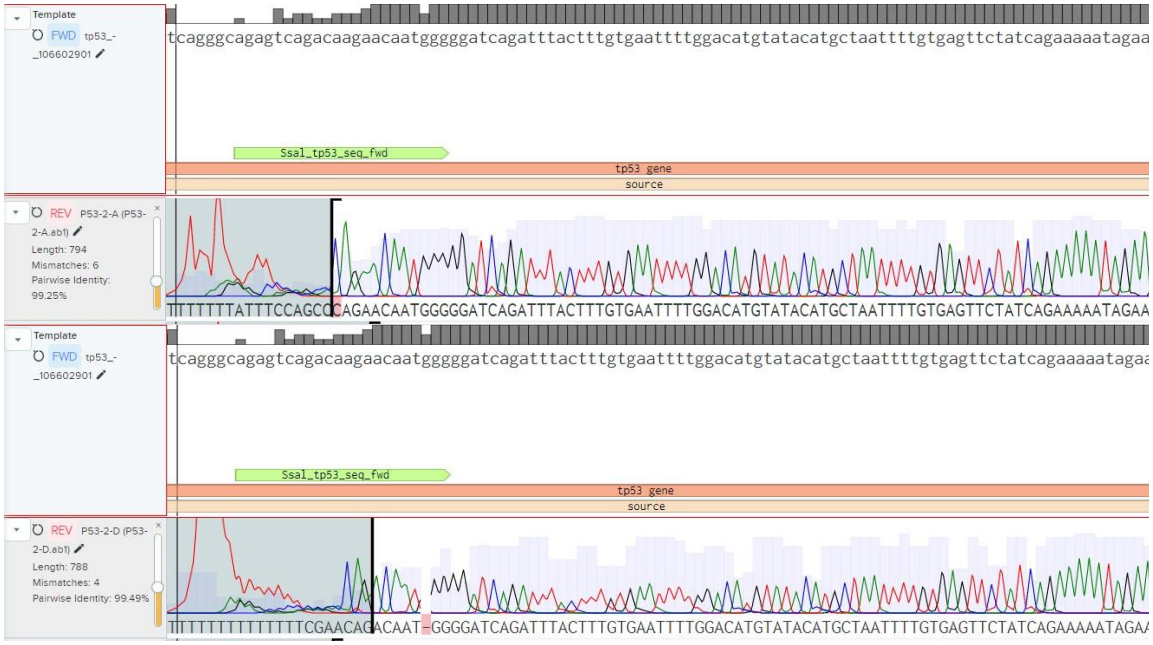
TSS-DMSO-1

P53-Agonist-1



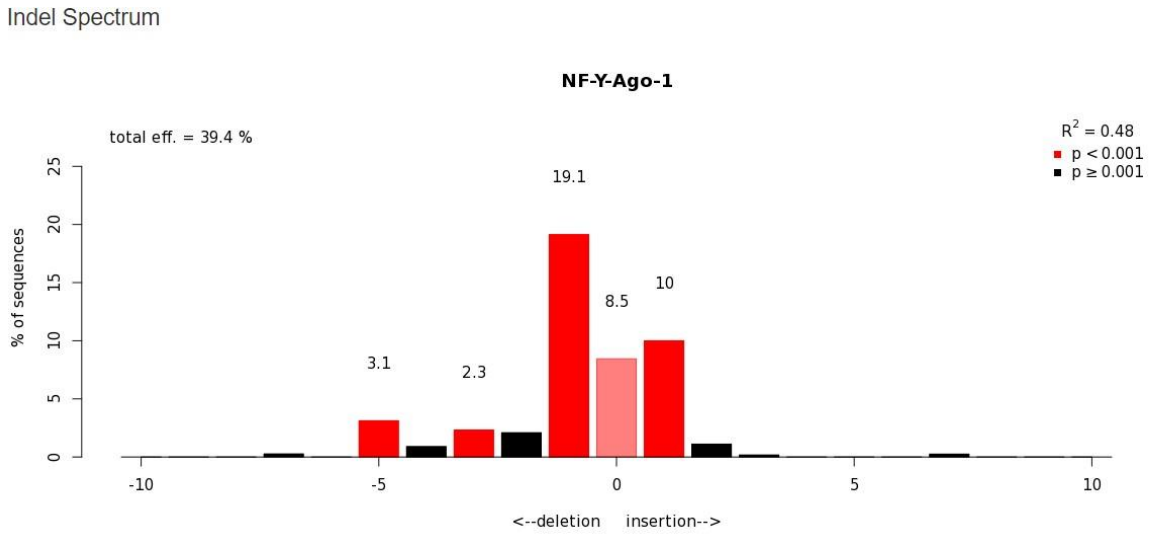
P53-Agonist-2

P53-DMSO-1

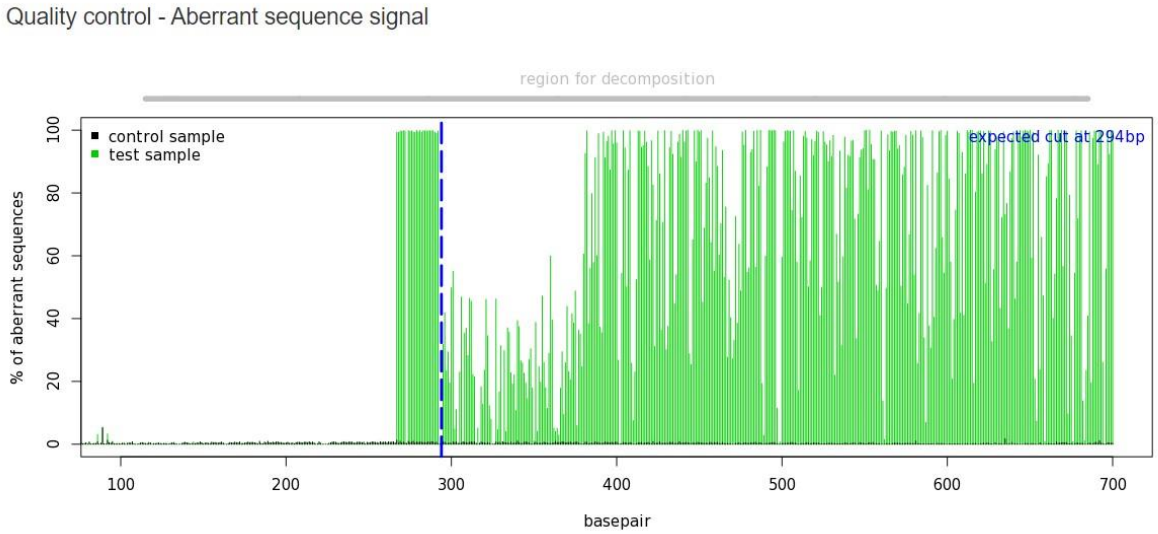


P53-DMSO-2

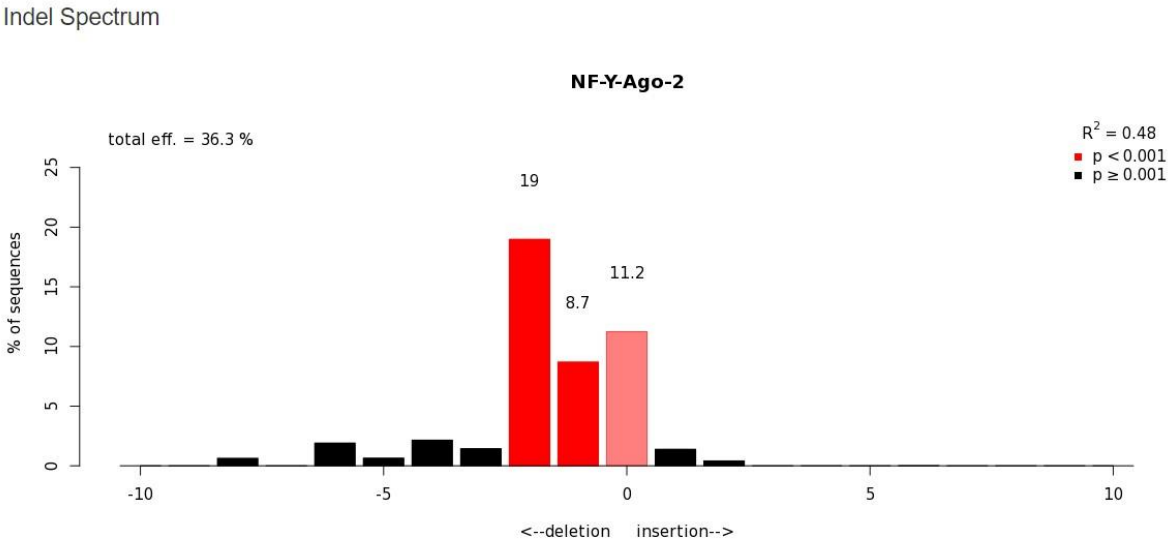
Appendix 2: CRISPR-Cas9 knockout success evaluation by Tide (Web tool, https://tide.nki.nl/**),** Section 4.6



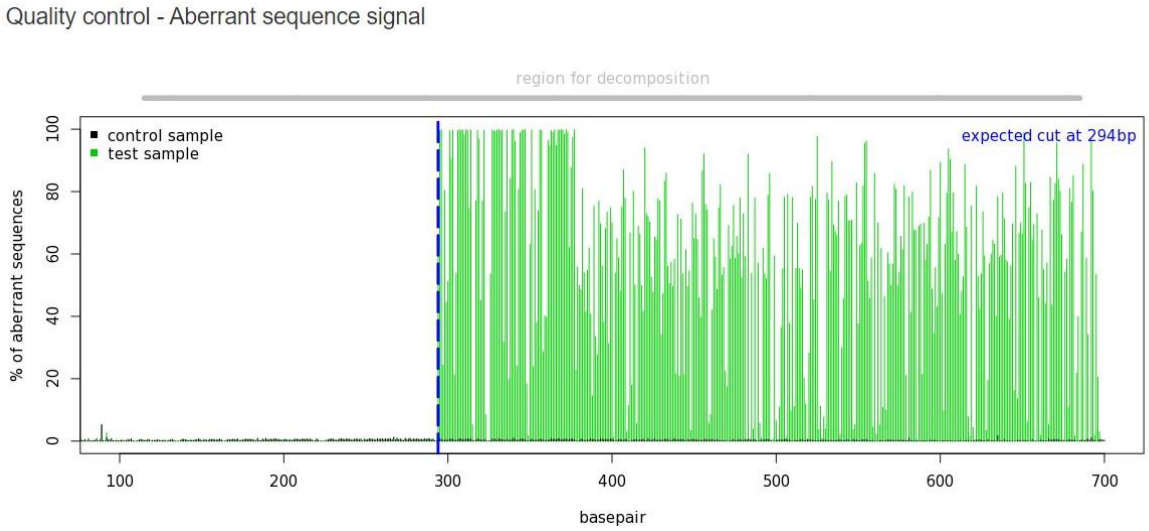
a) NF-Y-Ago-1



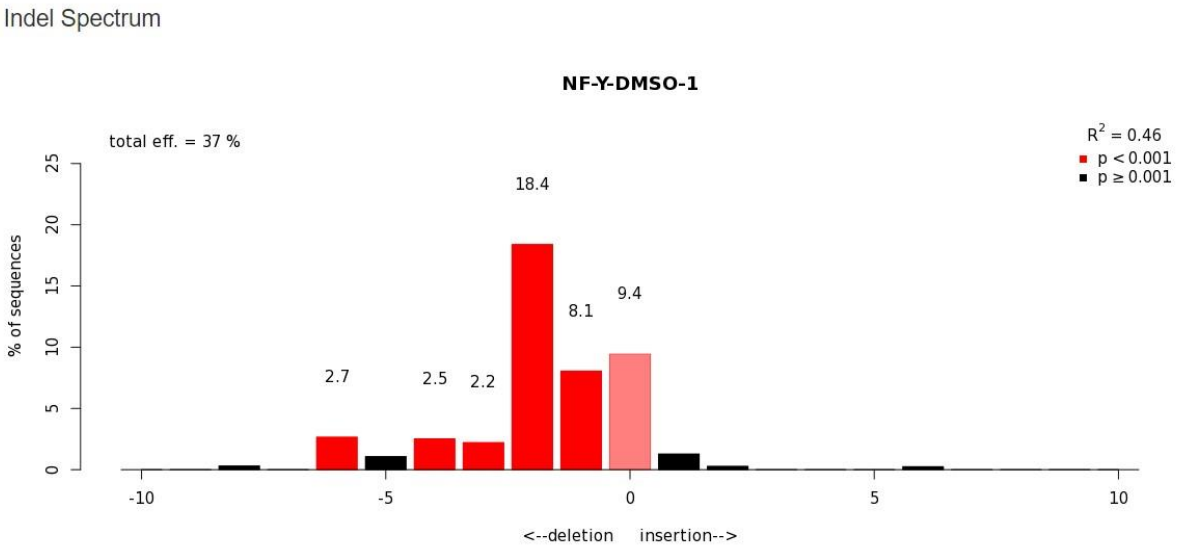
b) NF-Y-Ago-1-QUALITY control graph



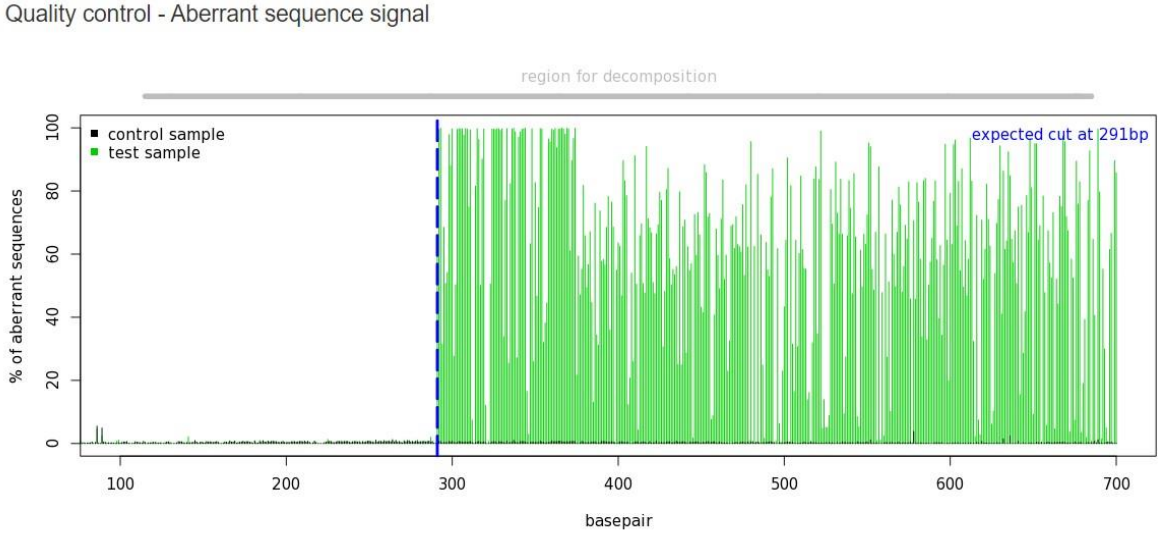
c) NF-Y-Ago-2



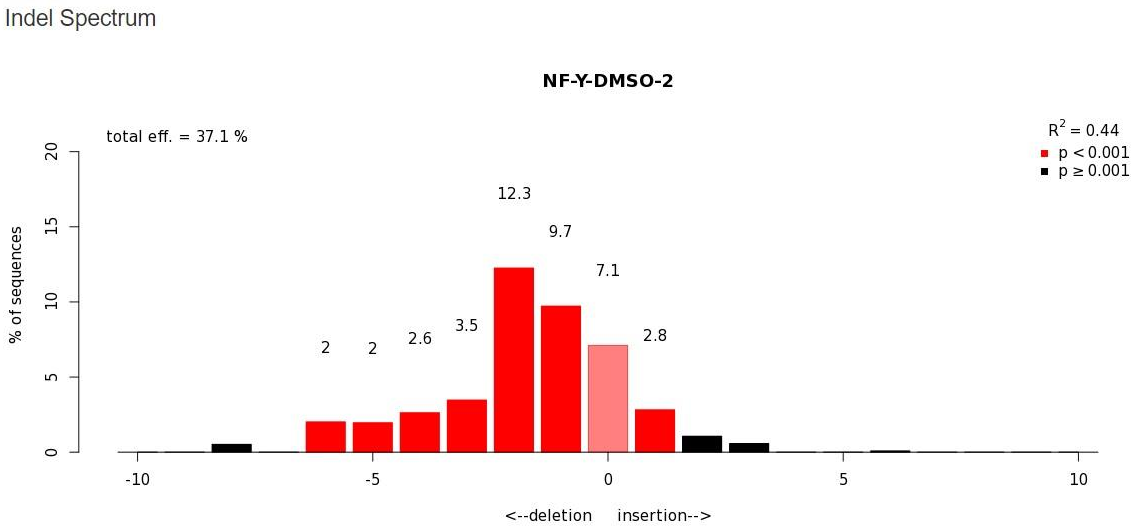
d) NF-Y-Ago-2-QUALITY control graph



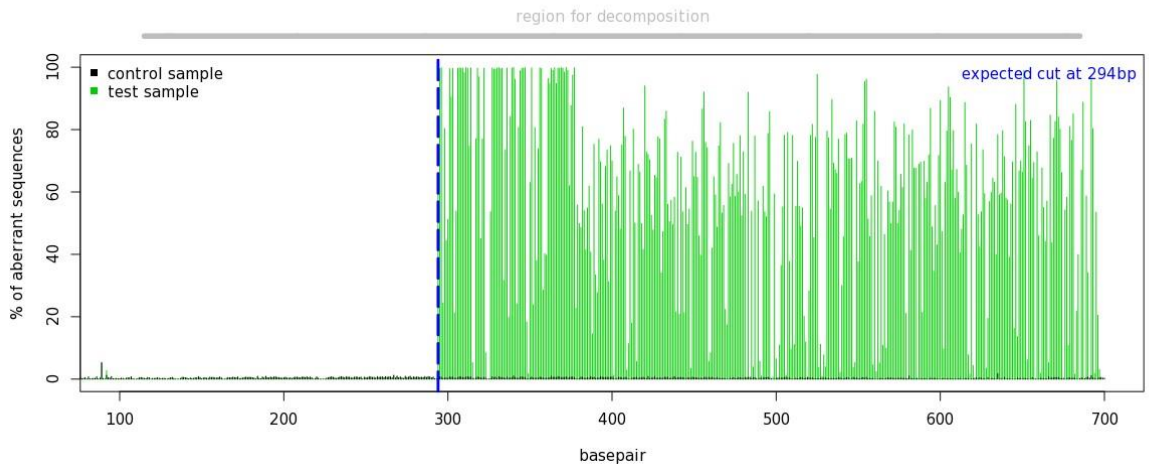
e) NF-Y-DMSO-1



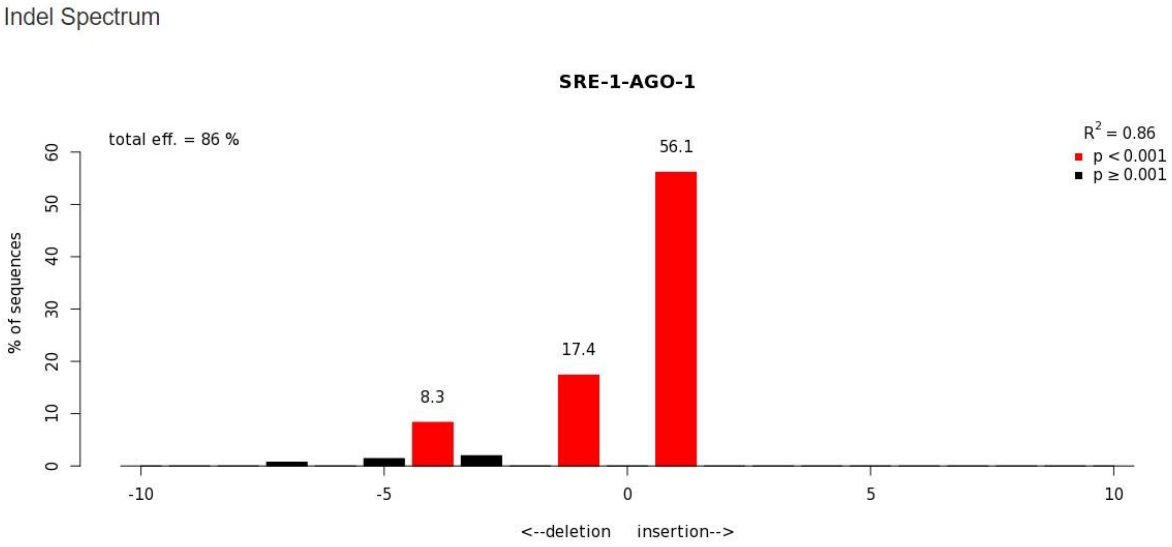
f) NF-Y-DMSO-1-QUALITY



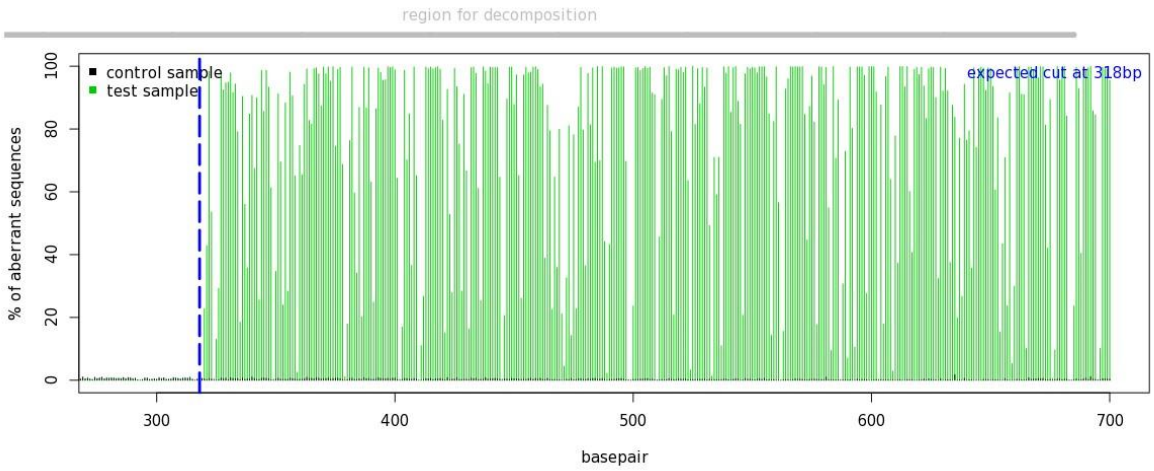
g) NF-Y-DMSO-2



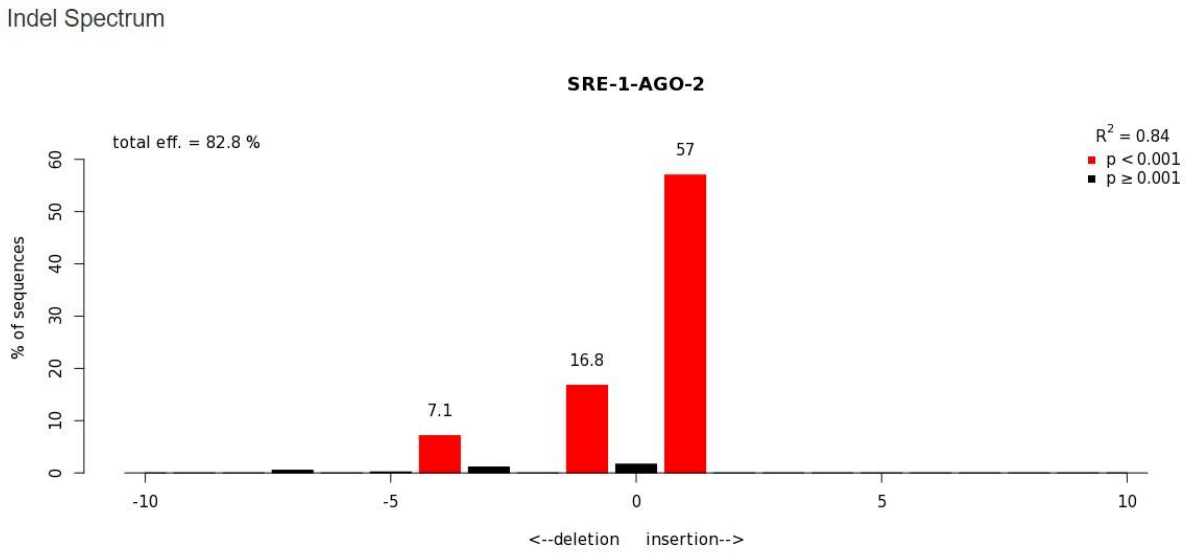
h) NF-Y-DMSO-2-QUALITY control graph



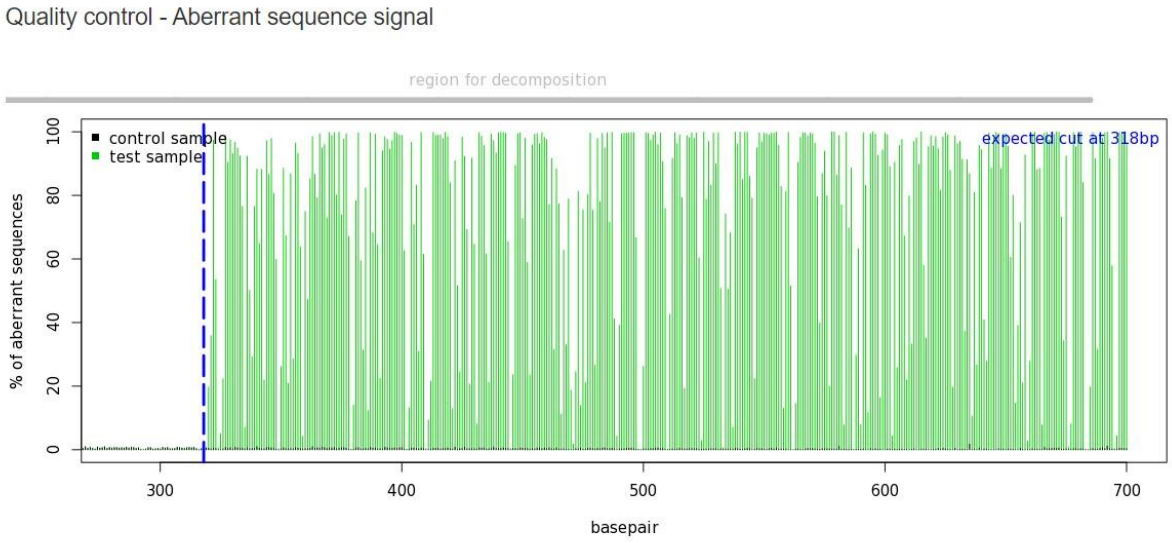
i) SRE1-Ago-1



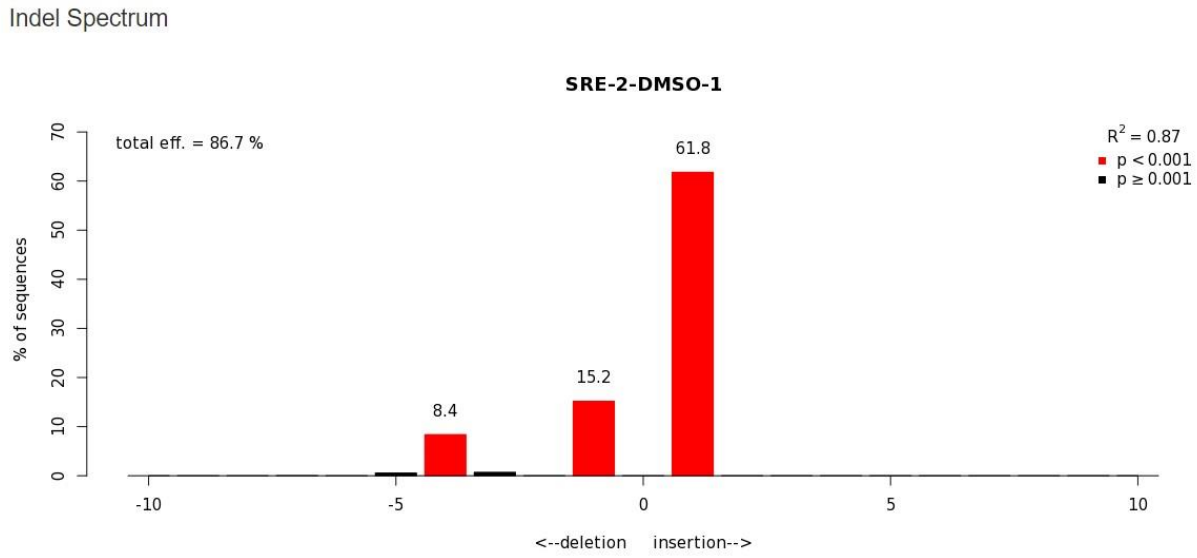
j) SRE1-Ago-1-Quality control graph



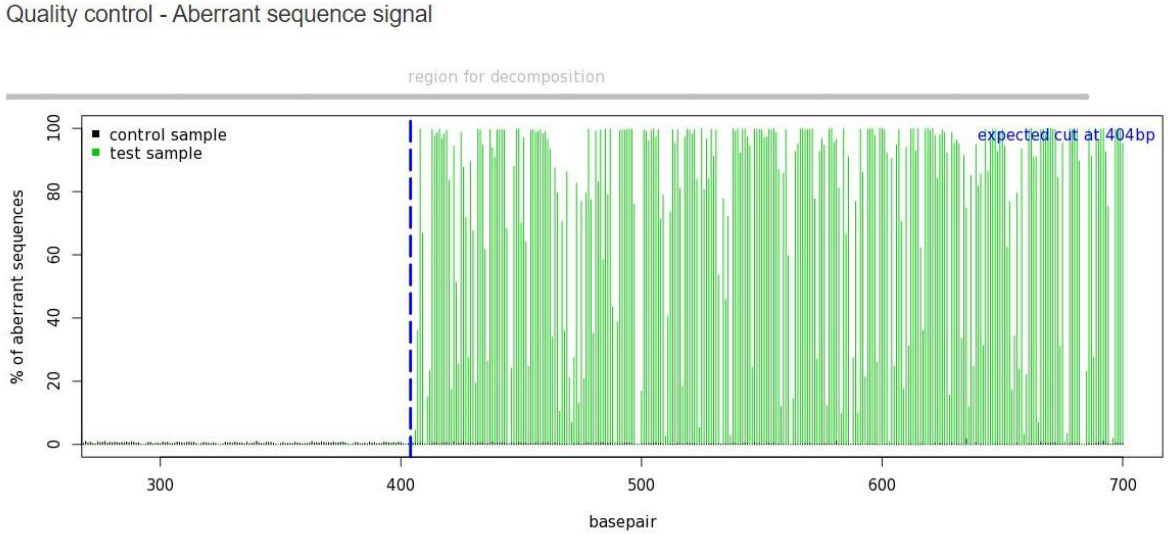
K) SRE1-Ago-2



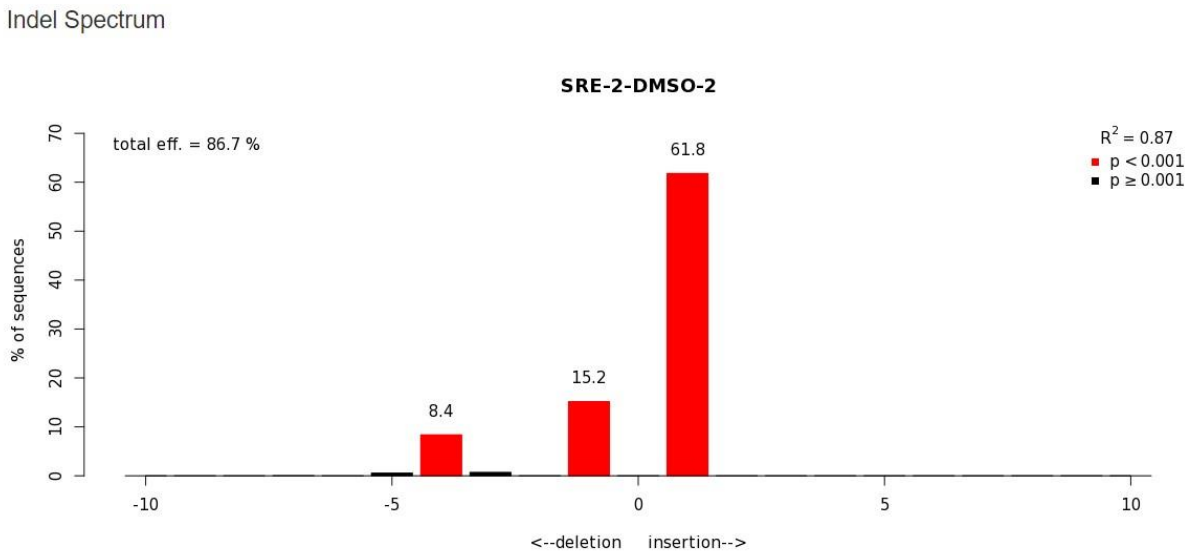
I) SRE2-Ago-2-Quality control graph



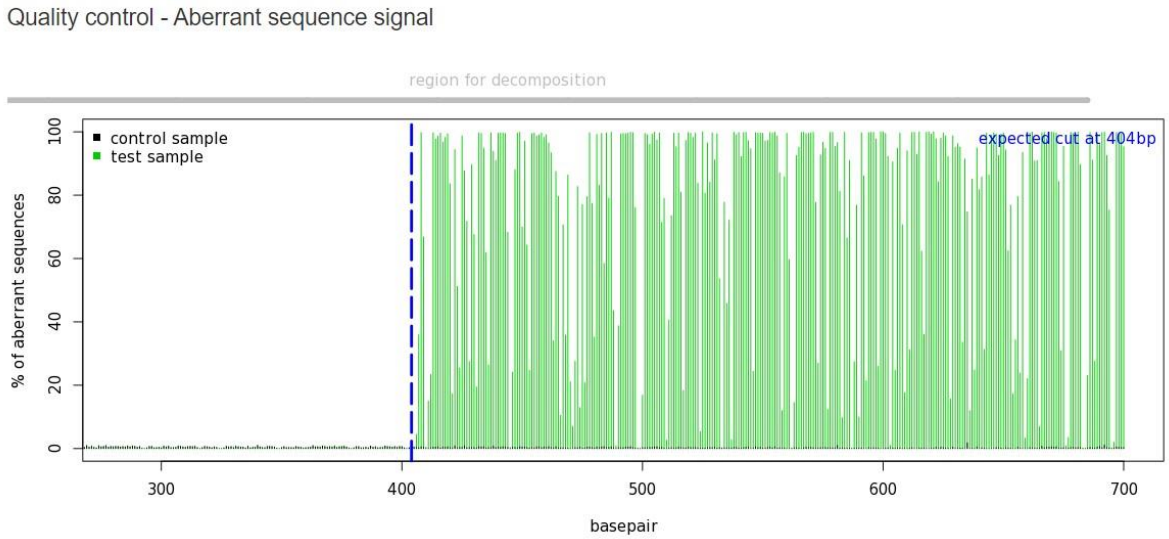
m) SRE2-DMSO-1



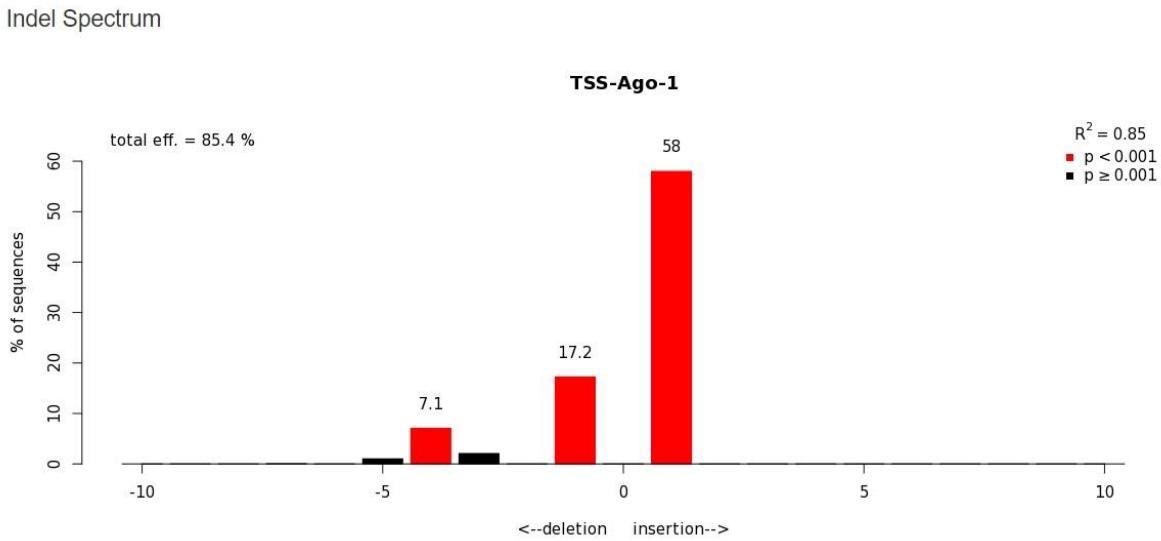
n) SRE2-DMSO-1-Quality control graph



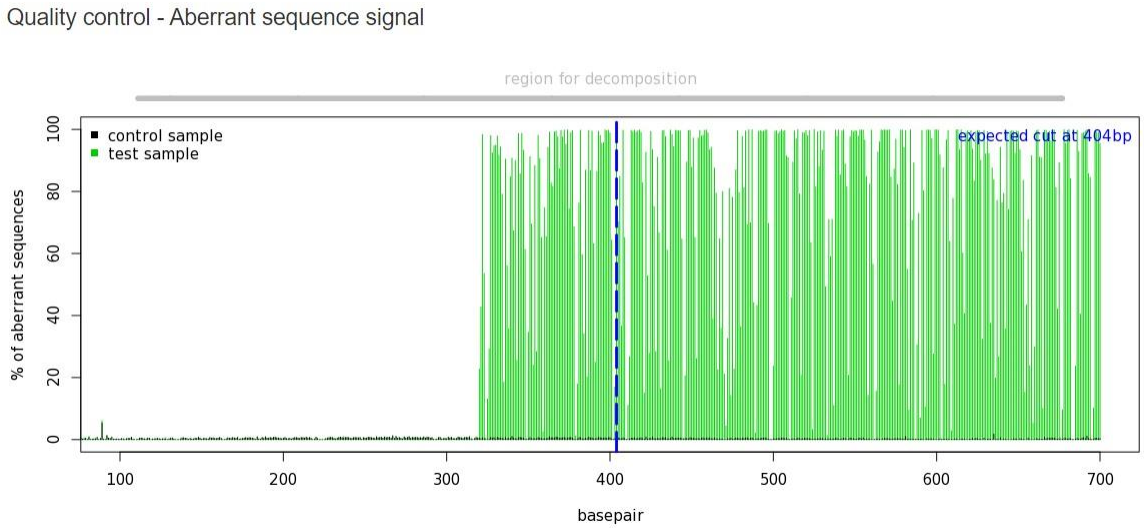
o) SRE2-DMSO-2



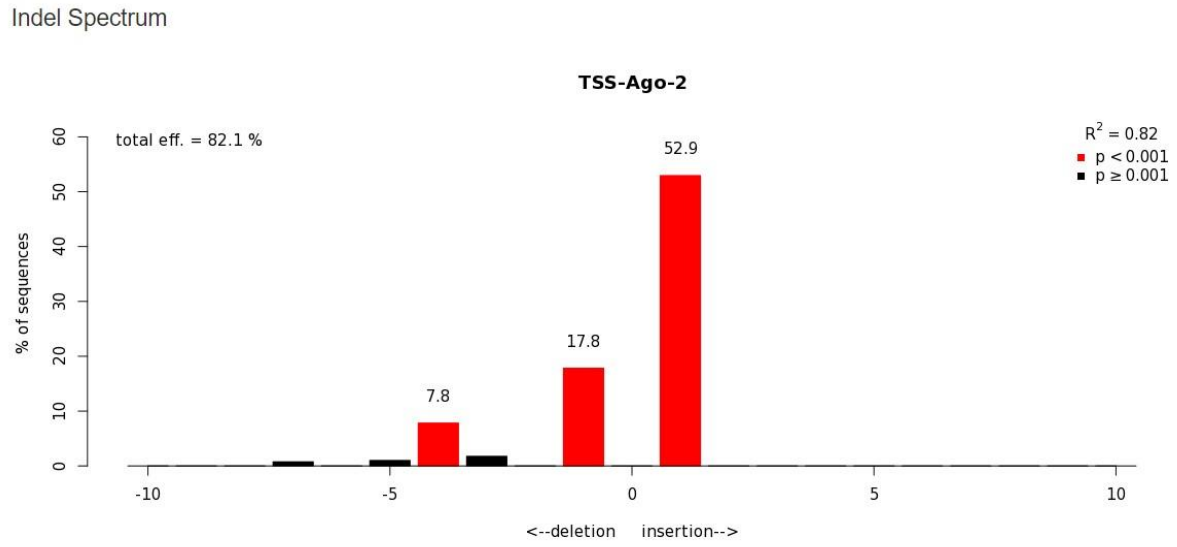
p) SRE2-DMSO-2-Quality control graph



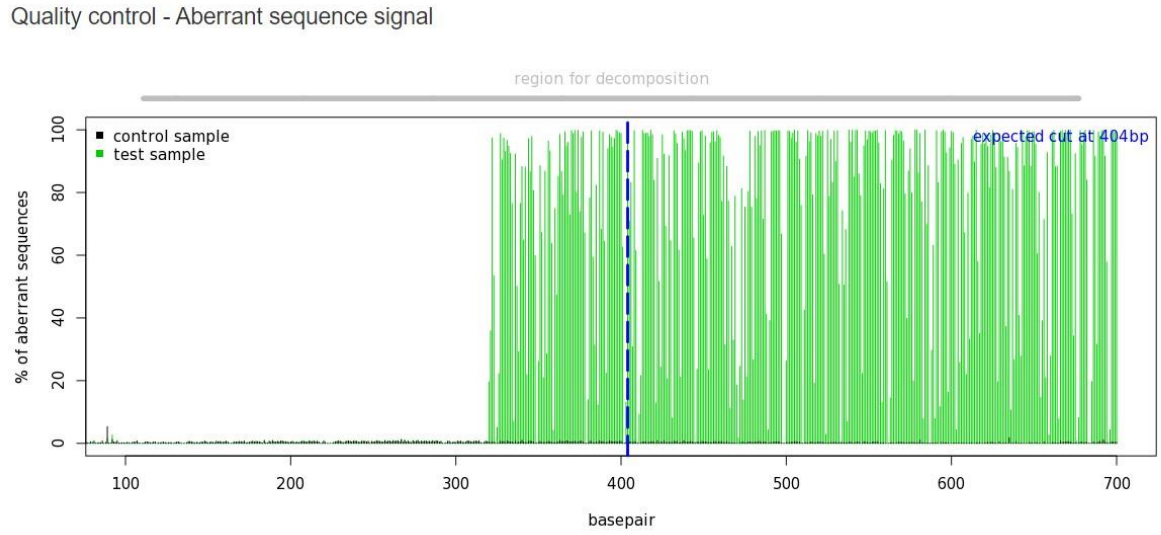
q) TSS-Ago-1



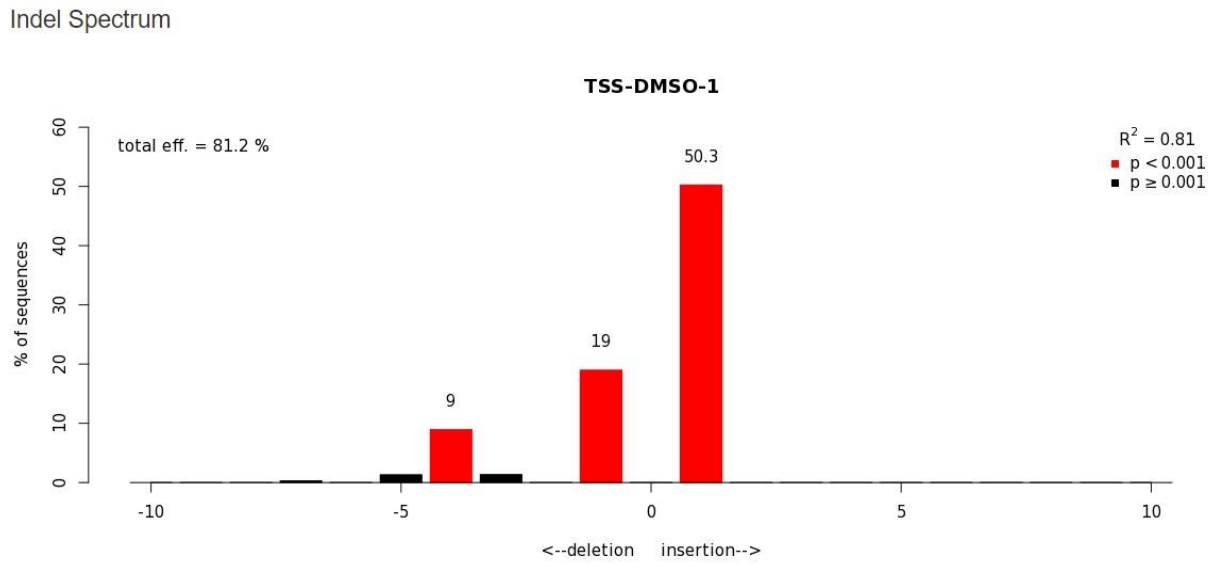
r) TSS-Ago-1-QUALITY control graph



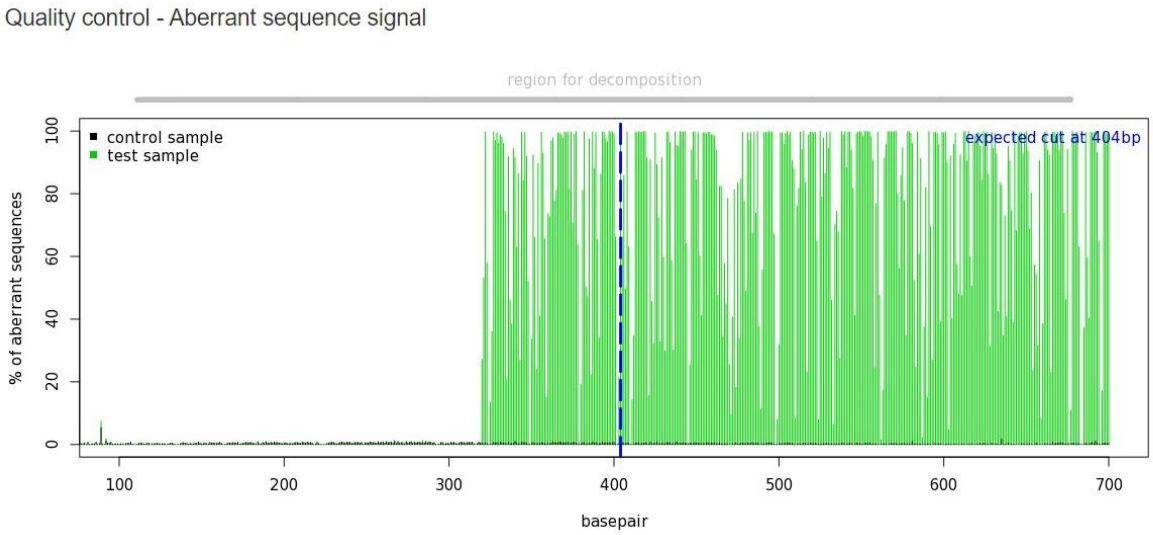
s) TSS-Ago-2



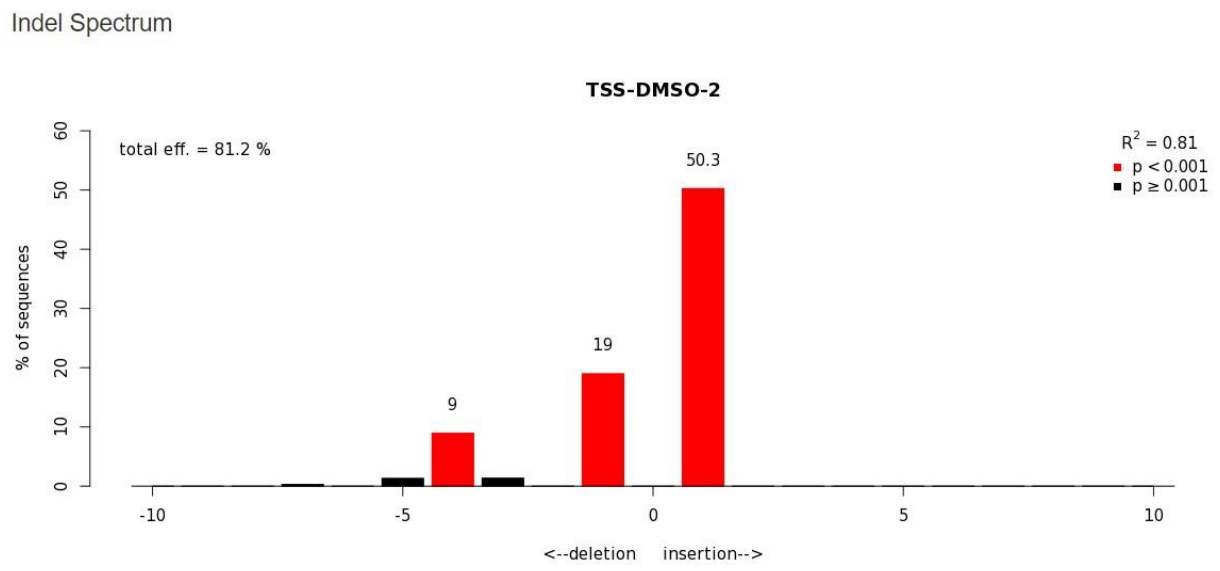
t) TSS-Ago-2-QUALITY



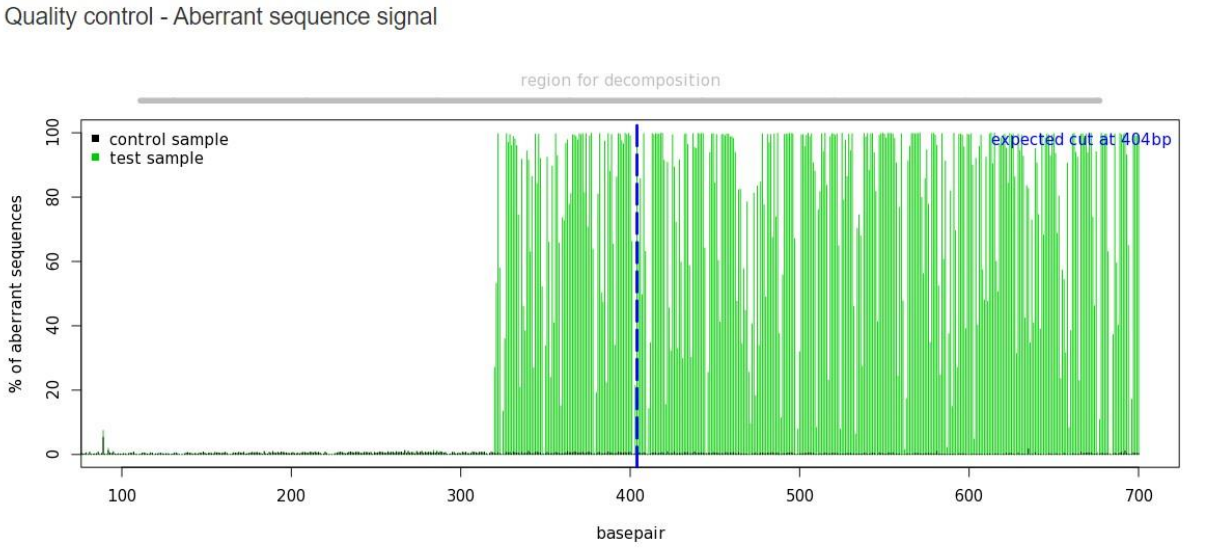
u) TSS-DMSO-1



v) TSS-DMSO-1-QUALITY coality control graph

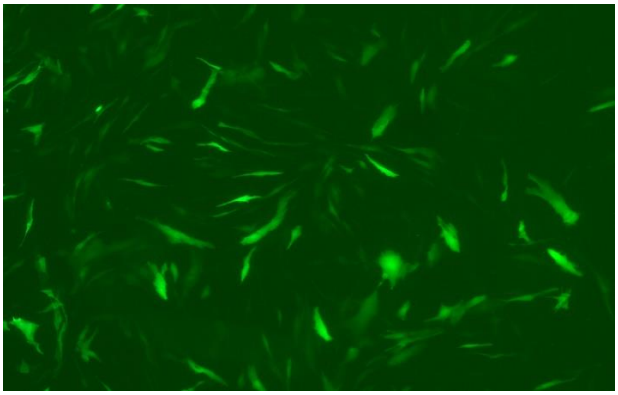


w) TSS-DMSO-2

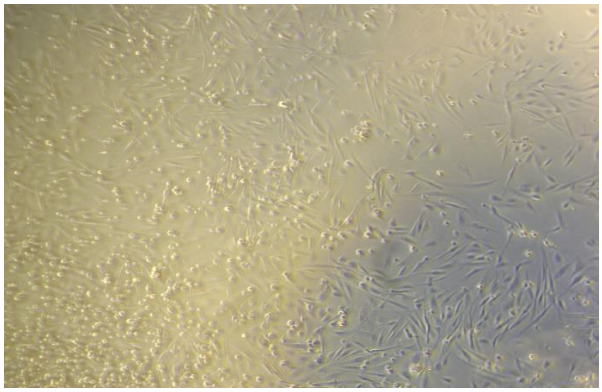


x) TSS-DMSO-2-QUALITY control graph

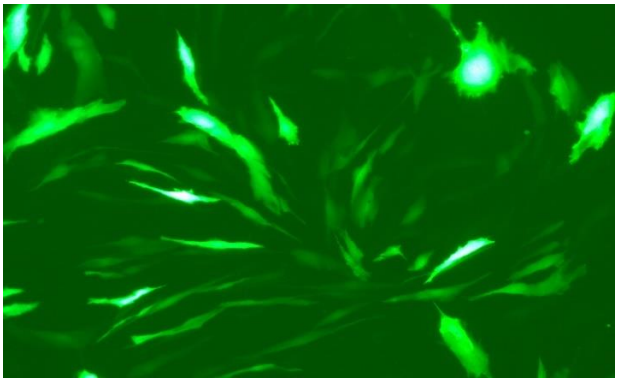
Appendix 3: Images of pEGFP under fluorescent and white light (5x and 10x resulation):



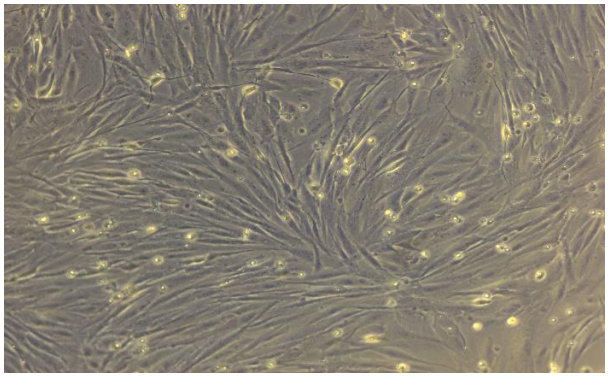
pEGFPN1 Well: fluorescent light (Experiment-1),5x



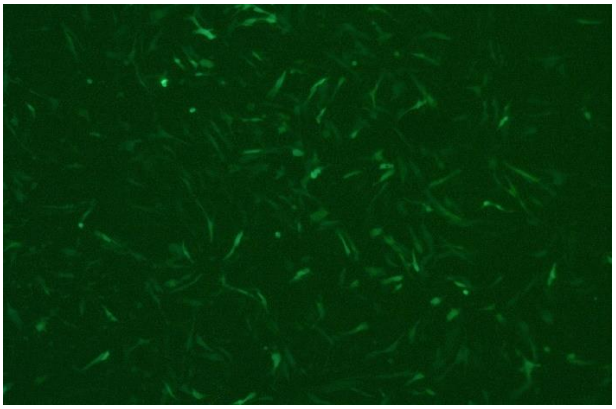
pEGFPN1 Well: white light (Experiment-1),5x



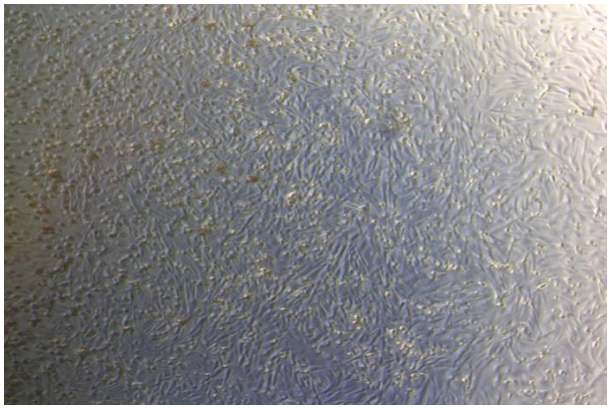
pEGFPN1 Well: fluorescent light (Experiment-1), 10 x



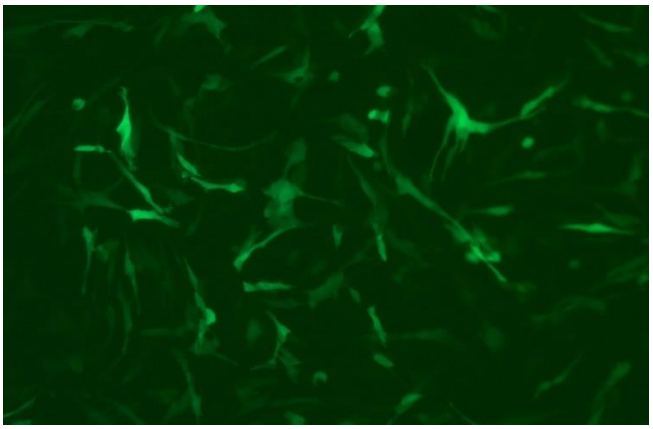
pEGFPN1 Well: white light (Experiment-1), 10x



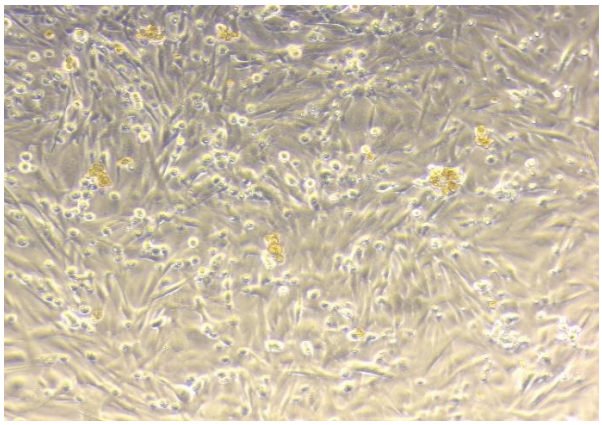
pEGFPN1 Well: fluorescent light (Experiment-2),5x



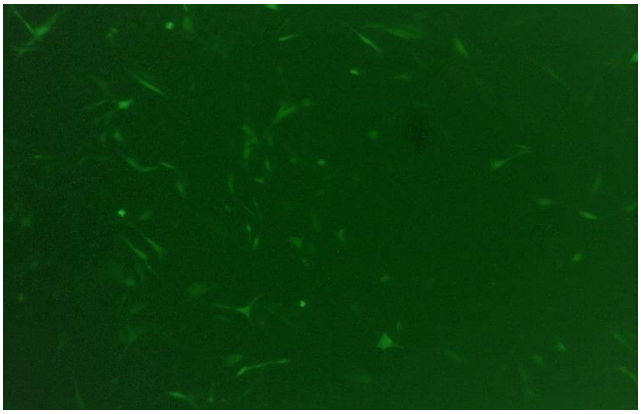
pEGFPN1 Well: white light (Experiment-2),5x



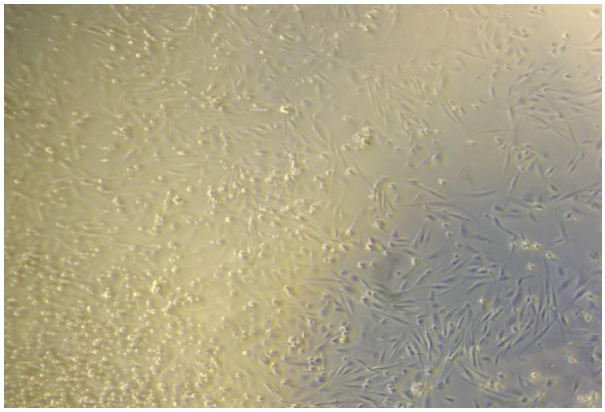
pEGFPN1 Well: fluorescent light (Experiment-2),10x



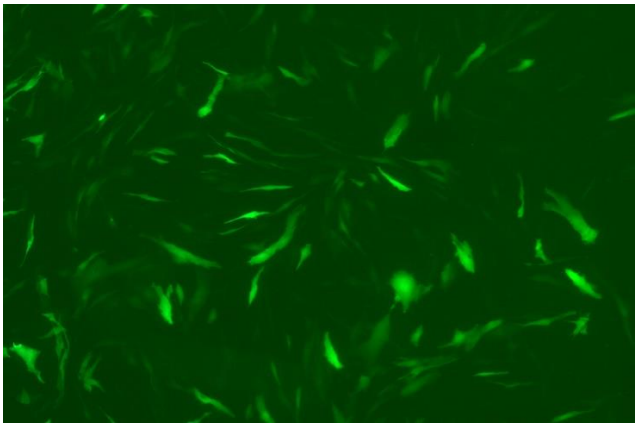
pEGFPN1 Well: white light (Experiment-2),10x



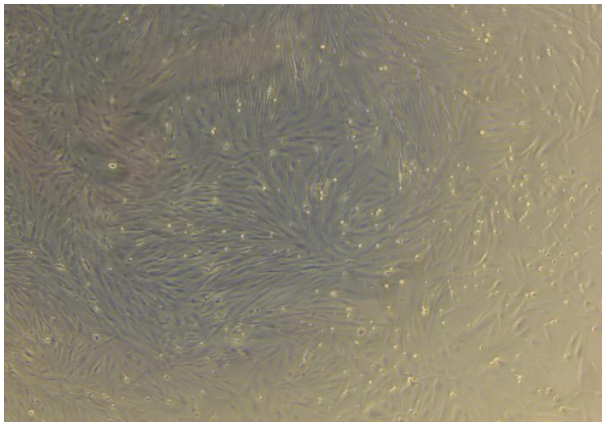
pEGFPN1 Well: fluorescent light (Experiment-3),5x



pEGFPN1 Well: white light (Experiment-3),5x



pEGFPN1 Well: fluorescent light (Experiment-3),10x



pEGFPN1 Well: white light (Experiment-1),10x

Appendix 4: Dual-Glo Luciferase Assay (Raw Data)

Experiment-1

Renilla-luciferase value

fluc	1	2	3	4	5	6	7	8	9	10	11	12
	No trans-1	No trans-2	No trans-3	pGL4.10 -1	pGL4.10 -2	pGL4.10 -3	El5aWT -1	El5aWT -2	El5aWT -3	El5aAT -1	El5aAT -2	El5aAT -3
A	4	6	5	14	5	2				92	112	60
	El5bWT -1	El5bWT -2	El5bWT -3	El5bAT-1	El5bAT-2	El5bAT-3						
B	102	80	60	50	24	12	**Agonist**					
	No trans-1	No trans-2	No trans-3	pGL4.10 -1	pGL4.10 -2	pGL4.10 -3	El5aWT -1	El5aWT -2	El5aWT -3	El5aAT -1	El5aAT -2	El5aAT -3
C	26	38	26	12	13	4	55	69	25	1383	1444	101
	El5bWT -1	El5bWT -2	El5bWT -3	El5bAT-1	El5bAT-2	El5bAT-3						
D	69	13	16	122	65	41	**DMSO**					
E												

Renilla-luciferase value

rluc	1	2	3	4	5	6	7	8	9	10	11	12
	No trans-1	No trans-2	No trans-3	pGL4.10 -1	pGL4.10 -2	pGL4.10 -3	El5aWT -1	El5aWT -2	El5aWT -3	El5aAT -1	El5aAT -2	El5aAT -3
A	1929	1062	905	2656	356	258				120	113	139
	El5bWT -1	El5bWT -2	El5bWT -3	El5bAT-1	El5bAT-2	El5bAT-3						
B	690	400	404	109	206	150	**Agonist**					
	No trans-1	No trans-2	No trans-3	pGL4.10 -1	pGL4.10 -2	pGL4.10 -3	El5aWT -1	El5aWT -2	El5aWT -3	El5aAT -1	El5aAT -2	El5aAT -3
C	12493	38135	26728	7647	3330	1846	708	1353	955	11294	3232	483
	El5bWT -1	El5bWT -2	El5bWT -3	El5bAT-1	El5bAT-2	El5bAT-3						
D	4383	926	1988	8668	5441	1384	**DMSO**					

Experiment-2

Firefly-luciferase value

fluc	1	2	3	4	5	6	7	8	9	10	11	12
	No trans-1	No trans-2	No trans-3	pGL4.10-1	pGL4.10-2	pGL4.10-3	El5aWT-1	El5aWT-2	El5aWT-3	El5aAT-1	El5aAT-2	El5aAT-3
A	15	10	10	28	21	17	221	236	88	338	517	217
	El5bWT-1	El5bWT-2	El5bWT-3	El5bAT-1	El5bAT-2	El5bAT-3		DMSO				
B	286	108	305	177	416	273						
	No trans-1	No trans-2	No trans-3	pGL4.10-1	pGL4.10-2	pGL4.10-3	El5aWT-1	El5aWT-2	El5aWT-3	El5aAT-1	El5aAT-2	El5aAT-3
C	6	9	6	29	11	19	9	23	20	196	110	10
	El5bWT-1	El5bWT-2	El5bWT-3	El5bAT-1	El5bAT-2	El5bAT-3		Agonist				
D	139	215	76	324	135	6						

Renilla-luciferase value

rluc	1	2	3	4	5	6	7	8	9	10	11	12
	No trans-1	No trans-2	No trans-3	pGL4.10-1	pGL4.10-2	pGL4.10-3	El5aWT-1	El5aWT-2	El5aWT-3	El5aAT-1	El5aAT-2	El5aAT-3
A	304	22	15	23588	14368	5614	24279	28952	8495	21291	26906	7801
	El5bWT-1	El5bWT-2	El5bWT-3	El5bAT-1	El5bAT-2	El5bAT-3		DMSO				
B	25942	9860	30161	23997	50696	46317						
	No trans-1	No trans-2	No trans-3	pGL4.10-1	pGL4.10-2	pGL4.10-3	El5aWT-1	El5aWT-2	El5aWT-3	El5aAT-1	El5aAT-2	El5aAT-3
C	38	80	39	5912	6991	14921	560	1057	709	8003	3750	570
	El5bWT-1	El5bWT-2	El5bWT-3	El5bAT-1	El5bAT-2	El5bAT-3		Agonist				
D	16033	21463	3464	36568	18284	570						

Experiment-3

Firefly-luciferase value

fluc	1	2	3	4	5	6	7	8	9	10	11	12
	No trans-1	No trans-2	No trans-3	pGL4.10-1	pGL4.10-2	pGL4.10-3	El5aWT-1	El5aWT-2	El5aWT-3	El5aAT-1	El5aAT-2	El5aAT-3
A	4	6	5	14	5	2	190	104	187	145	112	80
B	21	11	14	12	18	15						
C	26	38	26	12	13	4	55	69	25	1383	1444	101
D	69	13	16	122	65	41						

Renilla-luciferase value

rluc	1	2	3	4	5	6	7	8	9	10	11	12
	No trans-1	No trans-2	No trans-3	pGL4.10-1	pGL4.10-2	pGL4.10-3	El5aWT-1	El5aWT-2	El5aWT-3	El5aAT-1	El5aAT-2	El5aAT-3
A	1929	1062	905	2656	356	258	1929	1062	1929	1930	1297	1125
B	946	664	672	462	445	352						
C	12493	38135	26728	7647	3330	1846	1867	1353	1490	18294	32325	1890
D	4383	926	1988	8668	5441	1384						

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

Postboks 5003
NO-1432 Ås
Norway