Norwegian University of Life Sciences

Master's Thesis 2023 60 ECTS
Faculty of Biosciences

CRISPR/dCas9 Knock-down and CRISPR/Cas9 Knock-out of *lxr* in Atlantic salmon SHK-1 cells

Marie Hellan Iversen
Biotechnology

CRISPR/dCas9 Knock-down and CRISPR/Cas9 Knock-out of *lxr* in Atlantic salmon SHK-1 cells

Marie Hellan Iversen

Supervisors:
Dr. Guro Katrine Sandvik
Dr. Simen Rød Sandve

Master thesis

Faculty of Biosciences

Norwegian University of Life Sciences

May 2023

Acknowledgements

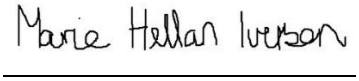
The work presented in this thesis was carried out from August 2022 to May 2023 at Centre for Integrative Genetics (CIGENE) as a part of a master's degree in biotechnology at the Faculty of Chemistry, Biotechnology and Food Science (KBM), Norwegian University of Life Sciences (NMBU). Supervision was performed by Dr. Guro Katrine Sandvik and Dr. Simen Rød Sandve.

I am grateful to my supervisor, Dr. Guro Katrine Sandvik, for the opportunity to be a part of her research group and to delve into the exciting field of CRISPR science. I appreciate her guidance and support during challenging times in the lab, and for her belief in both me and my thesis. I would also like to express my sincere thanks to Dr. Simen Rød Sandve for his honesty, continuous availability, and assistance with all the bioinformatic analyses. Furthermore, I would like to thank Postdoc Junsoung Kwak for his patience and invaluable guidance during my time in the lab. Lastly, I would like to thank all collaborators at Sandvik group that helped during my experiments. I have gained more knowledge and experience than I thought was possible in just one year, and I am truly grateful for this enriching and rewarding experience.

I am thankful to all the people who provided me with great company during my time at NMBU. I am especially grateful to Jenny, Mattias, Susanne, Mari, and Eline for making the past 5 years an unforgettable journey. Our breaks and lunches filled with quizzes, laughter, conversations, and support have been a highlight of my time at NMBU.

To my family, thank you for your constant love and support throughout my studies at NMBU. And finally, I want to express my gratitude to Gjermund for his patience, kindness, and encouragement during the tough days.

May 2023, Ås



Marie Hellan Iversen

Abstract

Understanding the genetics of Atlantic salmon (*Salmo salar*) is crucial for gaining insight into this commercially important species. The sustainable development of aquaculture requires increasing levels of plant-based ingredients in fish diets, which lack cholesterol and have less essential fatty acids that are important for consumers. The *lxr* gene, a member of the nuclear hormone receptor family, has shown to be a key regulator of lipid metabolism in vertebrates, also in Atlantic salmon (Minghetti et al., 2011). Despite its recognized importance, the precise mechanisms of action and downstream targets of *lxr* in Atlantic salmon remain incompletely understood. To address this deficiency, we have investigated the role of *lxr* in SHK-1 cells in Atlantic salmon using CRISPR/Cas9 gene editing techniques to knock-out the gene and analyze the resulting changes in mRNA levels. Furthermore, we wanted to study the potential application of CRISPR/dCas9, a modified version of CRISPR/Cas9 that can be utilized for transcriptional regulation in Atlantic salmon. To the best of our knowledge, the application of knock-down by CRISPR/dCas9 in Atlantic salmon has not been reported previously. The sample number in this thesis was too low to conclude if the CRISPR/dCas9 system can be functional in Atlantic salmon, but the results showed indications of a successful *lxr* downregulation. Thus, more tests with a larger number of samples are needed to confirm these findings. The knock-out of the *lxr* gene resulted in the upregulation of 25 genes and the downregulation of 19 genes.

Sammendrag

Kunnskap om genetikken til Atlantisk laks (*Salmo salar*) er avgjørende for å få innsikt i denne kommersielt viktige arten. Bærekraftig utvikling innen akvakultur krever økt kunnskap om plantebaserte ingredienser i fiskefôret. Fordi planter mangler kolesterol og har færre essensielle fettsyrer vil også fisken ha lavere innhold av disse komponentene. Genet *lxr*, et medlem av kjernehormonreseptorfamilien, har vist seg å være en viktig regulator for lipidmetabolismen hos virveldyr, også hos atlantisk laks (Minghetti et al., 2011). Til tross for den anerkjente betydningen, er de detaljerte mekanismene og nedstrømsmålene for *lxr* i Atlantisk laks fortsatt ufullstendig forstått. For å øke denne kunnskapen har vi undersøkt rollen til *lxr* i Atlantisk laks ved å bruke CRISPR/Cas9-genredigeringsteknikker for å slå ut genet og deretter analysere endringer i RNA-ekspresjon. Videre ønsket vi å studere mulighetene for å regulere geners transkripsjon ved bruk av CRISPR/dCas9 i Atlantisk laks. Så vidt vi vet, er studier med bruk av CRISPR/dCas9 i Atlantisk laks ikke publisert tidligere. Dette prosjektet har for få prøvereplikater for å konkludere med at CRISPR/dCas9-systemet kan være funksjonelt i Atlantisk laks. Likevel, viser våre resultater en indikasjon på suksessfull nedregulering av *lxr*, selv om ytterligere testing av systemet er nødvendig. Etter CRISPR/Cas9 knock-out av *lxr* ble 25 gener oppregulert, og 19 gener nedregulert.

Table of content

1 Introduction

The salmonids hold great significance globally, not only for their economic value, but also for their importance in scientific research (Davidson et al., 2010). The following sections intend to provide an overview of the lipid metabolism in Atlantic salmon, followed by information about the structure and function of the *lxr* gene, the properties of the SHK-1 cell line, and a summary of the Clustered regularly interspaced short palindromic repeats (CRISPR) gene editing system. Furthermore, the theoretical basis for the applied laboratory techniques will be presented upon.

1.1 Lipid metabolism in Atlantic salmon (*Salmo salar*)

The Atlantic salmon (*Salmo salar*) serves as an important part of a healthy human diet, contributing with essential omega-3 fatty acids (OM3FA). OM3FA reduce the risk related to cardiovascular diseases, enhance the reduction of cholesterol in the cells and play an important role in establishing basic, essential components like the cell membrane of humans (Elagizi et al., 2021; Rodriguez et al., 2022; Watanabe & Tatsuno, 2020). The significance of Atlantic salmon lies in its lipid composition, which could be considerably affected by changes in the feed given to farmed fish (Minghetti et al., 2011). To increase the sustainability of the salmon farming industry, and avoid over exploitation of marine resources, a significant amount of the fish feed used presently is composed of plant-based ingredients. The exchange of feed composition results in a considerable reduction of the levels of advantageous fatty acids in the salmonid (Egerton et al., 2020; Naylor et al., 2009). Plant-based oil lack cholesterol and have less healthy fatty acids, which might lead to a decrease in both quantity and quality of the fatty acids in the fish (Cruz-Garcia et al., 2009). Thus, further research in this field is necessary to optimize fish diets while preventing a reduction of essential fatty acids, which are crucial for maintaining the nutritional value of fish for consumption. Therefore, knowledge about essential genes in lipid metabolism, along with their impact on important lipid pathways, is crucial to salmon aquaculture. Consequently, gaining a better understanding could facilitate the use of sustainable plant-derived alternatives to marine products in aquaculture, leading to improved efficiency and effectiveness in the industry.

Lipid metabolism is a complex process in vertebrates that involves genes including liver X receptor (*lxr*) and the peroxisome proliferator-activated receptor (*ppar*) genes (Kidani & Bensinger, 2012; Yue et al., 2005). The interaction between PPAR and LXR transcription factors regulate other lipid metabolism genes, like fatty acid synthase (*fas*) and fatty acid

elongation (*elovl*), also in Atlantic salmon (Carmona-Antonanzas et al., 2014; Datsomor et al., 2019; Minghetti et al., 2011).

The *lxr* gene has a crucial role in regulating cholesterol metabolism, involving catabolism, storage, absorption, efflux, and transport (Cruz-Garcia et al., 2009; Wang & Tontonoz, 2018). There are two duplicated copies of the *lxr* gene, possibly originating from the salmonid whole genome duplication, one on chromosome 26 (*lxr* ID:100270809) and the other on chromosome 11 (*lxr-like* ID:106561932). Information about the *lxr-like* gene is limited, making it an interesting gene for further exploration. The *lxr* gene is a nuclear hormone receptor encoding a 462 aa long protein (Figure 1). The *lxr* gene operate as a cholesterol sensor by binding oxysterol amounts proportional to cholesterol, preventing cholesterol overload in the cell (Minghetti et al., 2011; Wojcicka et al., 2007). Additionally, *lxr* exert pivotal roles in the metabolism of phospholipids, a major component of cellular membranes, along with cholesterol (Carmona-Antonanzas et al., 2014; Wang & Tontonoz, 2018). Nuclear receptors like *lxr* can either act directly with target genes or indirectly by impacting "cross-talk" with other signaling pathways (Aranda & Pascual, 2001). The binding of *lxr* regulates the expression of genes, and the encoding proteins involved in lipid metabolism act as a response to elevated cellular levels of cholesterol (Wang & Tontonoz, 2018).

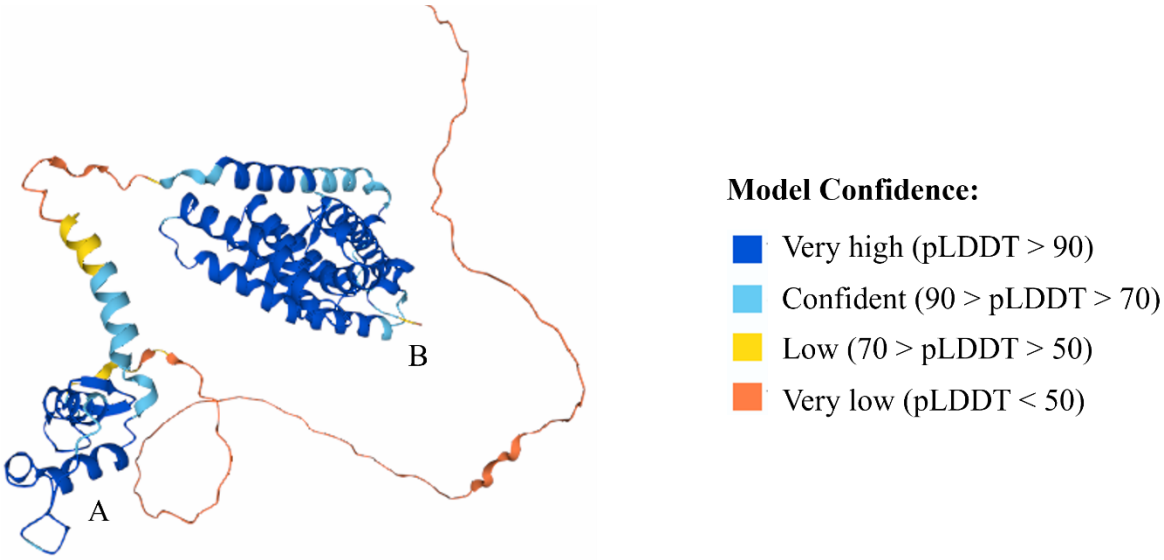


Figure 1: Model structure of the amino acids in the Lxr protein. The colors of the protein structure indicate degree of model confidence. The two regions in dark blue color indicates two regions containing zinc finger domains, A and B. AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. The structure of *lxr* is made in UniProt.

The Lxr protein contains two domains of very high model confidence (Figure 1). The first domain of the structure (A) contains a nuclear receptor DNA binding domain (DBD) that includes a zinc finger domain. The other zinc finger domain (B) in Figure 1 reveals the structure of a Ligand binding domain (LBD). The binding of the transcription factor LXR to DNA controls the expression of important lipid metabolism genes. The genes directly affected by LXR-DNA binding are described in Figure 2.

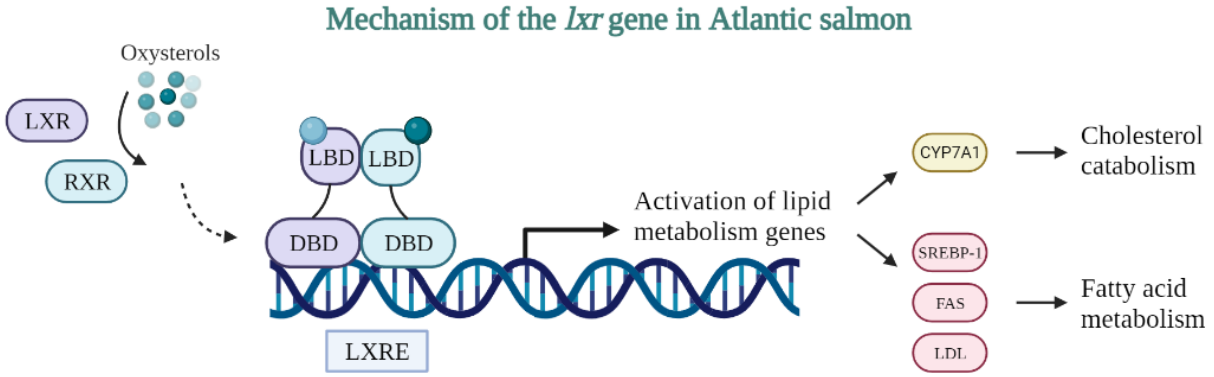


Figure 2: The *lxr* pathway in gene transcription. The liver X receptor (LXR) and retionoic X receptor (RXR) have one ligand binding domain (LBD) and one DNA binding domain (DBD), interacting with the LXR response elements (LXRE). The dimerization of LXR and RXR happens through the LBD and DBD. LXR and RXR bind oxysterols in the LBD and bind to DNA via DBD. The activation of the LXRE leads to transcription of target genes involved in lipid metabolism. Cholesterol catabolism is activated by the *cyp71a* gene, and fatty acid metabolism is activated by *srebp-1, fas* and *ldl* genes. The Figure is created in BioRender with inspiration from Cruz-Garcia et al. (2009).

As Figure 2 indicates, the *lxr* gene holds a pivotal role in establishing the balance between cholesterol catabolism and fatty acid metabolism (Cruz-Garcia et al., 2009). The LXR transcription factor creates a dimerization with the retionoic X receptor (RXR) by the DBD and LBD. The dimerization activates downstream genes involved in cholesterol catabolism (*cyp7a1*) and fatty acid metabolism (*srebp-1, fas* and *ldl*). The binding of LXR to oxysterols is directly linked to the cellular cholesterol content, resulting in the activation of cholesterol catabolism pathways as needed. Additionally, LXR binding activates genes related to fatty acid metabolism in accordance with the cellular requirements. As such, LXR acts as a regulator, maintaining optimal cellular conditions and contributing to homeostasis. Its role is critical in preserving a balanced lipid metabolism within the cell.

1.2 The whole genome duplication of Atlantic salmon

The sequencing of the Atlantic salmon genome has been pivotal in our understanding of the physiological and environmental aspects of salmon (Houston & Macqueen, 2019). The availability of high-quality assembly of the salmon genome (Ssal_v3.1) provides an adequate insight into the underlying mechanisms and the basic understanding of gene functions in Atlantic salmon. About 80 million years ago the salmon genome went through a whole genome duplication (Lien et al., 2016). The gene duplication provided salmon with two copies of a gene instead of one, which creates an opportunity for one gene to evolve while the other remains unchanged and ensures survival of the organism. Duplicates of genes (ohnologues) are often differently regulated, resulting in morphological complexity to the organism (Robertson et al., 2017). A potential result of this whole genome duplication is the emergence of the *lxr-like* gene from the *lxr* gene. The presence of *lxr* ohnologues from the gene duplication could pose a challenge in functional analyses of the genes, due to numerous similar regions in the two gene sequences. This could complicate further gene editing experiments, as it could provide difficulties in targeting the correct gene. Thus, the *lxr* ohnologues are highly intriguing to examine possible consequences of duplicated regions and warrants further exploration in this study.

1.3 Salmon head kidney (SHK-1) cell line

Recent studies investigating the regulatory control mechanism of lipid metabolism in Atlantic salmon have identified expression of the *lxr* gene in the salmon head kidney (SHK-1) cell line (Collet & Collins, 2009; Minghetti et al., 2011). The established Atlantic salmon cell line is derived from an immortalized cell line of head kidney. Advantages of using *in vitro* cell lines are the homogenous population of cells, the ease and rapid growth, continuously subculturing and ease of genetic manipulation (Kaur & Dufour, 2012). Cell lines of SHK-1 are valuable models to evaluate gene editing due to incorporation of CRISPR/Cas9 system in the cells (Gratacap et al., 2020).

1.4 Genome editing by CRISPR technology

Clustered regularly interspaced short palindromic repeats (CRISPR) technology has significantly advanced the field of gene research, allowing precision in studying gene function and regulation of intracellular signaling pathways (Li et al., 2020; Qi et al., 2013). The Cas9 endonuclease originated from *Streptococcus pyogenes* is the most frequently used in CRISPR systems (Dominguez et al., 2016; Le Rhun et al., 2019). Charpentier and Doudna received the 2020 Nobel Prize in Chemistry for their contribution to the development of CRISPR/Cas9 as a precise tool for targeted genome editing (Westermann et al., 2021). Interfering with the function of a gene is an effective mechanism to decipher its role in the organism and acquire pivotal knowledge about the function of genes (Rossi et al., 2015).

1.4.1 Gene knock-out by Cas9 protein

The CRISPR system relies on the presence of two essential components, the Cas9 endonuclease and the guide RNA (gRNA) (Dominguez et al., 2016; Le Rhun et al., 2019). The Cas9 protein is an RNA-guided DNA endonuclease responsible for cutting the targeted sequence of DNA (Barman et al., 2020; Qi et al., 2013). In the CRISPR/Cas9 system the *trans*-activating CRISPR RNA (tracrRNA) is responsible for recruitment of Cas9, while the CRISPR RNA (crRNA) is responsible for directing the complex to the target sequence (Doudna & Charpentier, 2014; Liao & Beisel, 2021). gRNA is constructed by hybridizing the 5'-end of tracrRNA with the 3'-end of crRNA (Barman et al., 2020). The single gRNA has the potential to degrade any target DNA sequence with the presence of a specific protospacer adjacent motif (PAM) (Karlson et al., 2021). PAM is a short DNA sequence Cas9 recognizes as 5'-NGG-3' (N can be any nucleotide) in *S. pyogenes*, frequently used to recognize proper target sites (Gleditzsch et al., 2019). The binding of Cas9-gRNA triggers a conformational change in the DNA, allowing crRNA to probe PAM (Le Rhun et al., 2019). Recognition of PAM leads to unwinding of DNA by breaking the hydrophobic stacking attractions in the DNA helix, separating the two strands, allowing crRNA to probe complementarity with the target DNA. Cas9 cuts approximately 3 nucleotides upstream of the PAM sequence, inducing an irreversible double-stranded break subsequent to the gRNA sequence (Barman et al., 2020).

Cellular DNA repair mechanisms are responsible for repairing the double-stranded break in DNA. In absence of a template DNA, the error-prone non-homologous end joining (NHEJ) is initiated. NHEJ causes random insertions and deletions during the repair of DNA (Li et al.,

2022). In presence of a template DNA, the high-fidelity homology-directed repair (HDR) is initiated (Yang et al., 2020). HDR uses donor templates to achieve artificially controllable insertions or deletions. The efficiency of HDR is limited due to the constant competition with the DNA repair pathway NHEJ, visualized in Figure 3 (Nambiar et al., 2019).

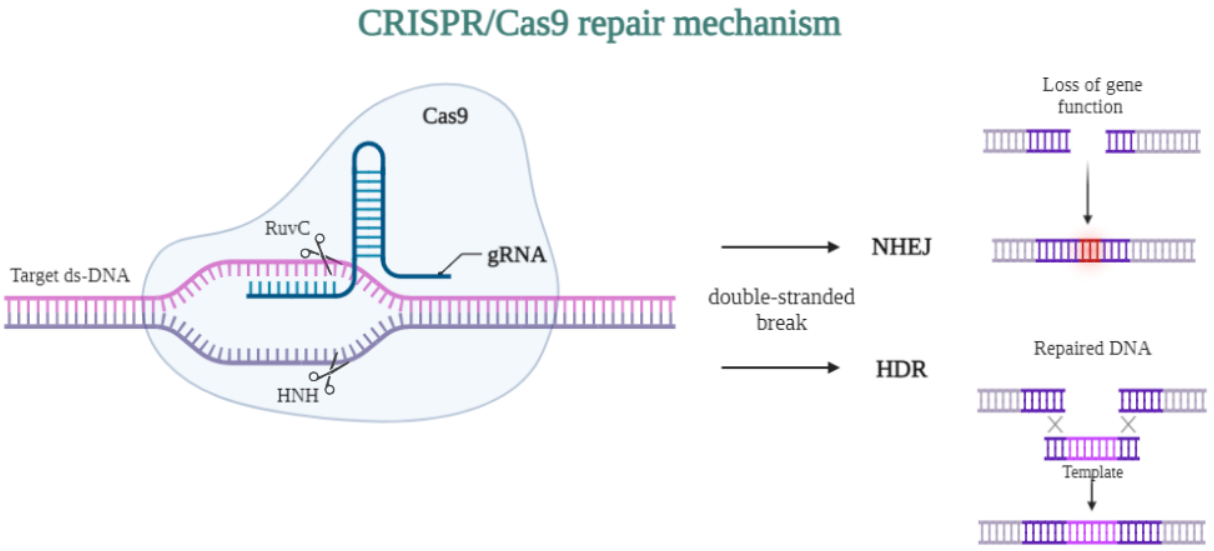


Figure 3: The mechanism of CRISPR/Cas9 gene editing followed by cellular repair mechanisms induced to mend the double-stranded cut. gRNA guides Cas9 to the target region and induces a double-stranded cut. DNA repair mechanisms, NHEJ and HDR, are initiated to mend the inserted cut. The Figure is created with BioRender.com, with inspiration from Doudna and Charpentier (2014).

In this experiment, the NHEJ repair mechanism is preferred because it often leads to genetic loss-of-function (Figure 3). This outcome is beneficial as one of the aims of the study is to investigate the pathways that are affected when a specific gene is knocked out. The genetic disruptions by NHEJ can be analyzed to determine the function of the targeted gene, as well as its role in regulating other genes and pathways.

Former studies of the CRISPR/Cas9 mechanism have revealed that Cas9 might bind to unintended regions of the genome and cleave, termed as off-target effects (Naeem et al., 2020). Off-target effects might cause regulation of non-targeted genes by cytotoxicity, genotoxicity, or chromosomal rearrangements (Hoijer et al., 2022; Zhao & Wolt, 2017). Another potential issue that could occur is the on-target unintended effects, caused by double-stranded cuts, which can lead to large deletions and rearrangements at the target site (Lackner et al., 2023). This concept refers to additional changes that could occur at the intended target site, resulting in unpredicted genomic alterations. To lower the risk of unintended effects, the CRISPR/dCas9 complex can be utilized to downregulate target genes without modifying the genes permanently, generating a reversible gene regulation (Karlson et al., 2021).

1.4.2 Gene knock-down by dCas9 protein

Since a cut in the DNA has the potential to cause unwanted effects in a cell, alternative CRISPR methods have been developed to circumvent this issue. For instance, a modified CRISPR/Cas9 system known as CRISPR interference (CRISPRi) can be utilized to decrease the expression of a specific gene (Cui et al., 2018). Mutations in the nuclease domains HNH (H840A) and RuVC1 (D10A) in Cas9 has led to the variant dead Cas9 (dCas9) containing a catalytically dead Cas9 protein (Qi et al., 2013). This altered function of the protein permits dCas9 to bind DNA without causing a cleavage, thereby providing a significant degree of flexibility to the CRISPRi system (Dominguez et al., 2016).

By using a dCas9/gRNA complex, CRISPRi can target specific sequences in DNA and modulate gene repression, visualized in Figure 4 (Hille & Charpentier, 2016). dCas9 can physically block transcription by preventing RNA polymerase (RNAP) from recognizing and binding to the promoter region (1). Knocking out the promoter involves targeting and disrupting the DNA sequence responsible for initiating transcription (Gilbert et al., 2014). Also, dCas9 has the capability to inhibit transcription even after RNAP has already bound to the target site, leading to the detachment of RNAP and terminating of the gene transcription (2). A third method is to fuse dCas9 to a transcriptional repression domain (3), like the Krüppel associated box (KRAB). The fusion accompanies a stronger and more specific inhibition of the target gene (Alerasool et al., 2020; Karlson et al., 2021; Le Rhun et al., 2019).

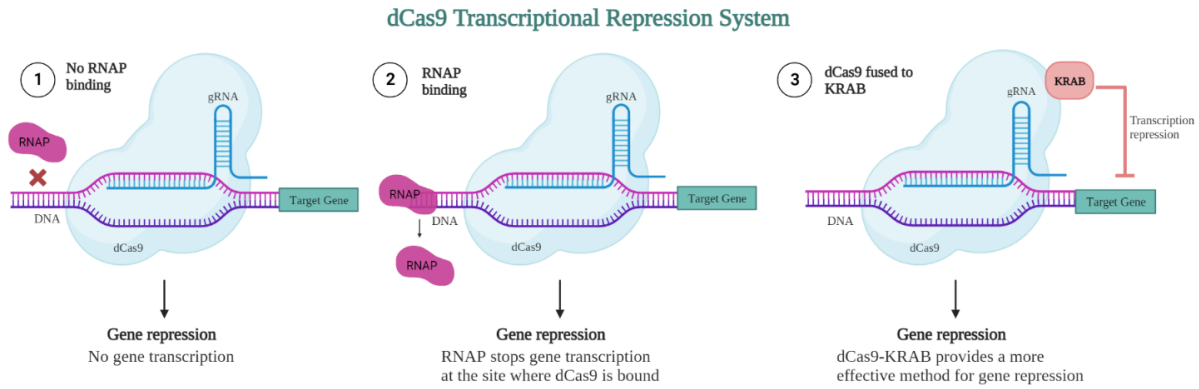


Figure 4: Mechanisms of CRISPR/dCas9 transcription repression system. Gene repression by dCas9 can block transcription physically by creating steric hindrance, or fuse dCas9 to a transcriptional repressor. Use of dCas9-gRNA induces gene repression without cleaving the DNA strands. dCas9 guided to the promoter region of the gene inhibits the RNAP to bind to the DNA, thereby preventing the transcription to initiate (1). dCas9 can physically block the transcriptional elongation of RNAP to induce gene repression by steric hindrance (2). dCas9 tethered to KRAB further strengthens the target gene repression (3). The Figure is created with BioRender.com, with inspiration from Li et al. (2019) and Ran et al. (2013)

The sequence and location of the gRNAs determines whether the dCas9 protein will induce physical blockage of RNAP binding (1) or RNAP detachment after binding (2) in Figure 4. Specifically, gRNA binding before the ATG start codon enhances promoter binding and subsequent inhibition of RNAP recruitment, while gRNA binding after the start codon induces RNAP detachment and inhibits further gene transcription. dCas9 fused to a KRAB (3) can recruit histone modifiers to induce methylation at enhancers forming heterochromatin, generating less accessibility in chromatins resulting in further silencing of gene expression (Hille & Charpentier, 2016; Thakore et al., 2015). All three dCas9-systems generates repression of the target gene.

1.4.3 Genetic compensation

Genes that share similarities may exhibit the phenomenon of genetic compensation, where knock-out of one gene result in upregulation of another similar gene (Peng, 2019; Rouf et al., 2022). Given this, a CRISPR knock-out of the *lxr* gene in Atlantic salmon could potentially trigger upregulation of the homologous *lxr-like* gene, preserving crucial lipid metabolism functions. The cellular response to loss-of-function of a gene occurs to maintain organisms viability due to disruption of important genes (Rouf et al., 2022). Until recently, scientists believed that if they removed a gene and did not see any obvious effects, the gene was not important. In 2015, Rossi et al. reported their findings that deleterious mutations induced genetic compensation by activating similar genes with homologue sequences. The recent findings have reviewed that genetic compensation occurs when the knocked out gene holds a pivotal function of the viability of the organism (Rossi et al., 2015; Rouf et al., 2022).

1.5 Delivery of CRISPR components into cells by electroporation

In 2020 Gratacap and colleagues published papers indicating electroporation of RNP complex (Cas9-gRNA) leads to efficient CRISPR incorporation in Atlantic salmon (Gratacap et al., 2020). Electroporation is a method to induce molecules into cells by using high-voltage pulses to overcome the barrier of the cellular membrane (Gehl, 2003). The external electrical field induces short rearrangements in the bilayer of phospholipids in the cell membrane, creating tiny pores open a few milli seconds (msec). DNA can be transferred through the pores by the destabilized cellular membrane. Utilizing electroporation to incorporate the RNP complex is an efficient method to investigate cell lines (Kim et al., 2014).

1.6 Measuring the success of gene editing

Evaluating the effectiveness of gene editing techniques like CRISPR/Cas9 and dCas9 requires precise measurements, that can detect genomic changes or gene expression alterations induced by the genetic manipulation. The efficacy of gene editing with CRISPR/Cas9 can be evaluated through polymerase chain reaction (PCR) and Sanger sequencing, which enables the detection and verification of the targeted genomic modifications. Meanwhile, the gene regulation mediated by dCas9 is usually quantified by measuring the changes in RNA levels using quantitative PCR (qPCR) cycles. By leveraging these techniques, researchers can obtain quantitative and qualitative data on the outcomes of their gene editing experiments.

Sanger and Coulson published a method for sequencing DNA in 1975, Sanger sequencing, determining the order of nucleic acid sequences in DNA. Sanger sequencing is a first-generation sequence method (Sanger & Coulson, 1975). DNA is annealed with a primer and extended with either of the four deoxynucleotide triphosphates (dNTP: dATP, dTTP, dCTP, dGTP) by DNA polymerase (Crossley et al., 2020). The amount of dNTP is limited and consequently DNA fragments can be distinguished by the various lengths using Sanger sequencing (Gratacap et al., 2020). Sanger is a simple and rapid sequencing method; hence it is the gold standard for monitoring the efficacy of molecular-based assays (Crossley et al., 2020).

To determine the level of gene expression in a biological sample, qPCR is utilized for the precise measurement of the amount of messenger-RNA (mRNA). This method reflects the degree of gene activity or regulation under different conditions. In qPCR, the increase in amplification products is monitored throughout the reaction (Peirson & Butler, 2007). Where the fluorescence is reaching a certain threshold is defined as the C_q value for that reaction, and that is correlated with the initial amount of DNA in the initial sample. The method $\Delta\Delta C_q$ is a way of calculating the change in mRNA levels between samples (Haimes & Kelley, 2015). This technique involves normalizing the expression level of the target gene to an endogenous reference gene, whose expression should remain unchanged throughout the experiment. The results obtained from qPCR analysis show if the target gene mRNA levels are increased, decreased, or unchanged compared to the controls.

1.7 Measuring the effect of gene editing on other genes and pathways

RNA sequencing, also known as RNA-seq, provides researchers with a level of detail about gene expression in different conditions across samples, to gain insight into how genes are regulated within cells (Corchete et al., 2020). This technology provides fine insight of the transcriptome, providing vastly higher coverage and resolution compared to Sanger sequencing (Kukurba & Montgomery, 2015; Stark et al., 2019). RNA-seq uses high-throughput next-generation sequencing (NGS) to sequence complementary DNA (cDNA) from a sample in order to reveal the presence and amount of RNA (Kukurba & Montgomery, 2015). mRNA is elucidated by reverse transcription into cDNA to understand more of the functional complexity of the transcriptome, including novel gene structures, mRNA splicing and quantification of allele-specific expression (Montgomery et al., 2010). RNA-seq is usually conducted on samples containing ensembles of millions of cells, yielding a vast amount of information about cellular responses in a cell (Haque et al., 2017). Particularly, RNA-seq is often used for analyzing differential gene expression after a gene editing (Stark et al., 2019). Information provided by RNA-seq drives the exploration within differential gene expression and cellular responses to its wide range application, high sensitivity and accuracy (Haque et al., 2017). In many instances, RNA-seq is conducted to identify differences in gene expression between two or more conditions. This is also the aim of this experiment, where RNA-seq data is utilized to detect and characterize patterns between samples comprising either *lxr* knock-out, *lxr* knock-down or control samples.

2 Research questions

It is our hypothesis that a better understanding of the molecular processes involved in lipid metabolism and regulation in Atlantic salmon will lead to increased quality and effectiveness in the salmonid farming. Ongoing research and improvement of CRISPR editing methods have the potential to significantly enhance the long-term sustainability and success of the aquaculture industry. Specifically, this is our research questions:

1) Is CRISPR/dCas9 an efficient method to conduct knock-down experiments in cell cultures of Atlantic salmon?
2) What are the intracellular signaling pathways impacted by CRISPR gene regulation of the *lxr* gene in Atlantic salmon?

3 Materials and methods

We aimed to assess dCas9 gene regulation and identify genes affected by *lxr* knock-out. The laboratory experiments were conducted to optimize the results at each step before we proceeded with further experiments. This included the choice of 12-well plates for cell culturing, selection of successful gRNA and CDS primer pairs, choice of electroporation conditions for RNP delivery, and choosing the optimal bioinformatic tools to interpret the effects of *lxr* gene editing. The process of the practical lab work is visualized in Figure 5.

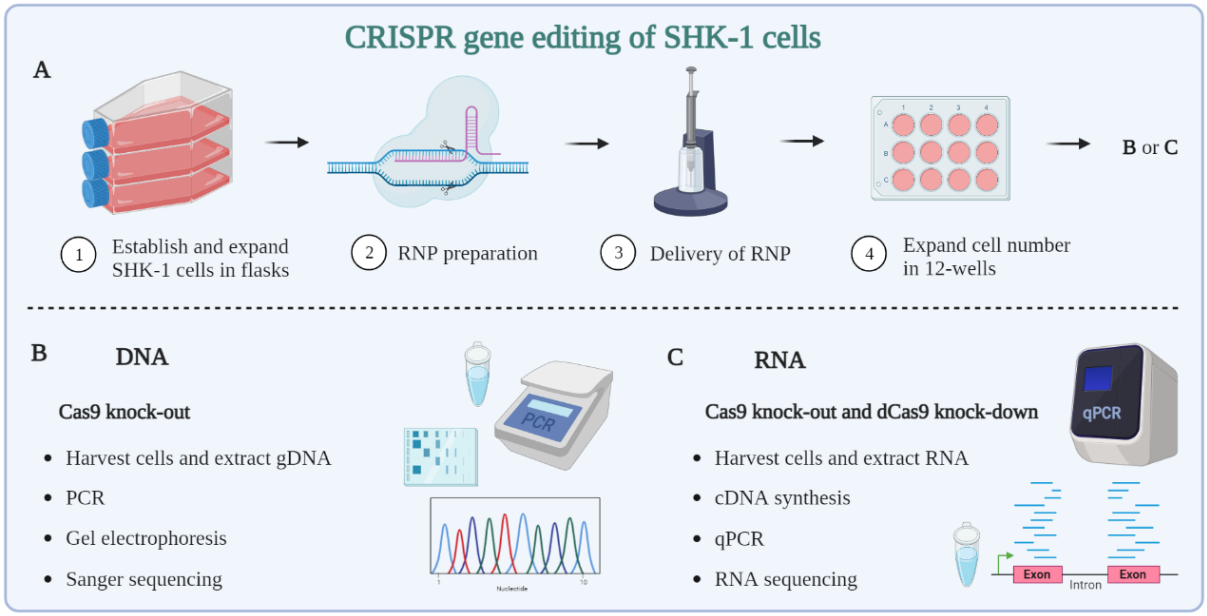


Figure 5: Schematic flow chart of the mechanisms of CRISPR/Cas9 knock-out and CRISPR/dCas9 knock-down gene editing in SHK-1 cells. A) SHK-1 cells were expanded to establish a cell culture. RNP complex (gRNA and Cas9/dCas9) and SHK-1 cells were prepared prior to electroporation. Electroporation used an electrical pulse to incorporate the RNP complex into the cells. The cells proliferate and expand in 12-wells while the CRISPR gene editing occurs. RNA from the knock-down cells was extracted after three days. DNA/RNA from the knock-out cells were extracted after one week of incubation. B) gDNA was purified prior to amplification using PCR, and DNA fragments were separated by gel electrophoresis before sending the DNA to Sanger sequencing. C) RNA was harvested and extracted from the cells prior cDNA synthesis followed by qPCR, or sent for RNA-seq. The Figure is created with BioRender.com.

The common steps for preparing cells for knock-out and knock-down are schematized in Figure 5A, while the unique steps for the specific procedures for further work with DNA or RNA are shown in Figure 5B and Figure 5C, respectively. The materials and methods used in this thesis for CRISPR knock-out were based on the protocol published by Gratacap et al. (2020).

3.1 Initial bioinformatic analyses of the *lxr* genes

To gain a basic understanding of the *lxr* genes, we performed an initial bioinformatic analysis on an existing gene expression dataset using R Project for Statistical Computing version 4.3.0 (R Core Team, 2021). We examined the gene expression of *lxr* and *lxr-like* in 15 different tissues of Atlantic salmon. Our analysis was carried out using the R BiocManager package *BgeeDB* to retrieve tissue gene expression data as normalized transcript per million (TPM) from the Short Read Archive study SRP011583. The data consists of wild-type Atlantic salmon in a sexually immature stage (BioProject: PRJNA72713). A detailed description of the bioinformatic analyses conducted in R can be found in the R Markdown file provided in Appendix G.

3.2 Culturing of the SHK-1 cell line

The SHK-1 cell line obtained from the European Collection of Authenticated Cell Cultures (#9711106) was cultured in a 75 cm^2 flask with cell culture media at 20 °C without CO_2. Cell culture media consists of 45 ml L-15 media with glutamax (Sigma-Aldrich #L1518), 5 ml (5.0 %) fetal bovine serum (FBS) (Gibco #16000044) and 500 µl Penicillin-Streptomycin (Gibco #15140122). To maintain sterility and prevent sample and cell contamination, all work with SHK-1 cells was conducted in a Laminar Flow Cabinet (LAF) with continuous ventilation.

SHK-1 cells grew rapidly and needed to be split at 80% confluency in 1:2 ratio approximately once a week to ensure optimal growth and further viability. The cell culture flask was gently washed 2 times with 5 ml phosphate buffered saline (PBS) and aspirated in between the washes. 2 ml of Trypsin-EDTA (0.05%) was added to cover the surface for 5 min to detach the cells from the plate. Cell culture media was added to the cell suspension and transferred to a tube. The cells were spun down at 200 x g for 5 minutes at room temperature, after which the supernatant was removed. The cells were then resuspended in 4 ml of fresh media and transferred to new cell culture flasks, where they were incubated at 20 °C.

3.3 Electroporation of cells to incorporate the CRISPR complex

To introduce the RNP complex and achieve CRISPR-mediated gene editing in the SHK-1 cells, we utilized the Neon Transfection System (Invitrogen #MPK5000). Preparation of the Cas9 protein and gRNA is necessary prior to RNP delivery. First, crRNA and tracrRNA, the two components of gRNA, were dissolved in nuclease-free water (100 µM each). 1 µl of crRNA and 1 µl of tracrRNA were mixed and incubated at 95 °C for 5 min to form the gRNA complex. When the mixture reached room temperature, 2 µl of Cas9 (20 µM) (NEB #M0386T) was added to a final concentration of 10 µM Cas9 and 25 µM gRNA. The RNP complex were cooled in room temperature for 15 min and put on ice (Figure 6). The same procedure and concentrations were used to create RNP complexes with dCas9 as with Cas9 (Gratacap et al., 2020).

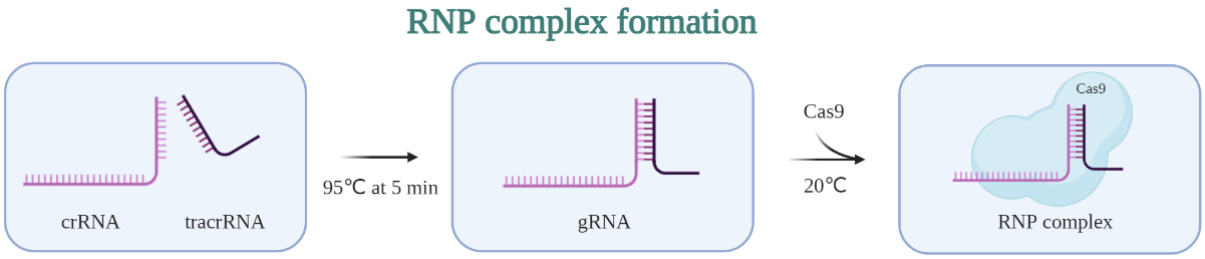


Figure 6: Fusion of gRNA and Cas9 protein to form the RNP complex. Overview of the ribonucleoprotein (RNP) complex preparation. crRNA and tracrRNA are combined at equimolar concentration and heated at 95 °C for 5 min. Once the duplex crRNA and tracrRNA is formed, Cas9 protein is added to form the RNP complex at 20 °C. The Figure is created in Biorender.com, with inspiration from Zoppo et al. (2020).

Figure 6 shows a schematic presentation of the RNP complex preparation. To obtain final concentration, 1.4 µl of the RNP complex were diluted with 2.6 µl Opti-MEM Reduced Serum Medium (Gibco #31985062), to a total solution mix of 4 µl. Before mixing the diluted RNP complex with SHK-1 cells, the cells needed to be prepared.

The SHK-1 cells were prepared for transfection according to Neon Transfection System instructions (ThermoFisher #MAN0001557). The SHK-1 cells were detached with Trypsin-EDTA (0.05%) and counted to determine cell density and viability. To count the cells, 10 µl of the SHK-1 cell solution was transferred to a 1.5 ml tube and mixed with 10 µl Trypan Blue stain (0.4%) (Gibco #15250061). The cells were counted by Countess Automated Cell Counter (Invitrogen #AMQAX2000) following the instruction manual of Invitrogen (#C10227). The required number of cell solution needed for 10^7 SHK-1 cells/ml were centrifuged at 200 x g for 5 min, followed by removal of supernatant. The cells were washed with 1 ml Dulbecco's phosphate-buffered saline (DPBS) to remove dead cells and debris, and centrifuged once again at 200 x g for 5 min. After removal of the supernatant, the cells were resuspended by adding 50

µl of Opti-MEM prior to electroporation. Tubes were prepared to electroporation with 10 µl of the cells containing Opti-MEM buffer, mixed with the diluted RNP complex (4 µl), to a total amount of 14 µl. The mixture was incubated at room temperature for 5 min. Then, 10 µl of the 14 µl RNP/cell solution was pipetted using Neon Pipette and inserted to the Neon Tube, which had already been loaded with 3 ml Electrolytic Buffer. Two electroporation settings were tested; 1600 V 10 msec and 3 pulses, and 1400 V 20 msec pulse and 2 pulses (Datsomor et al., 2023; Gratacap et al., 2020). The transfection of the SHK-1 cells is visualized in Figure 7.

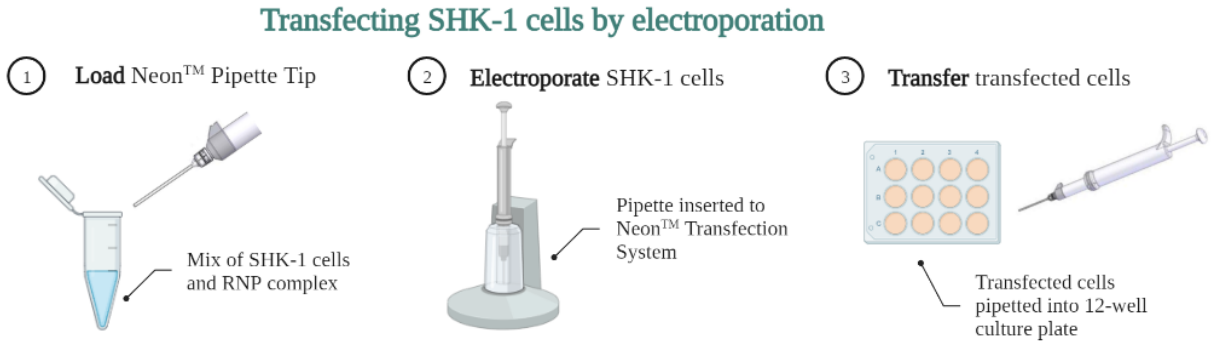


Figure 7: Electroporation of SHK-1 cells by Neon Transfection System. The Neon Transfection System provides a simple 3-step process with loading, electroporating and transferring cells. Mixture of SHK-1 cells and RNP complex was loaded in Neon Pipette Tip (1). The transfection occurs when the Neon Pipette Tip is inserted to the Neon Transfection System (2). Following electroporation, transfected cells are pipetted into 12-wells to further expansion and growth (3). The Figure is created in BioRender with inspiration from Invitrogen Neon Transfection System user guide (#MAN0001557).

The Neon transfection process consists of three steps: loading, electroporating and transferring cells (Figure 7). Achieving a successful CRISPR experiment hinges upon the accurate delivery of its components (Cas9 and gRNA) to the intended target cell. Following the electroporation, the cells were transferred to a 12-well culturing plate with 2 ml of L-15 fresh media (10% FBS, no antibiotics) per well. The 12-well plate was incubated overnight at 20 °C. The following day, we investigated the cell viability using EVOS M5000 Microscope (Invitrogen #AMF5000). Surviving cells still attached to the bottom of the flasks, while dead cells floated in the solution. After assessing the viability, the media in the 12-well plate was replaced with 2 ml of complete cell culture media containing antibiotics and incubated at 20 °C. RNA from the knock-down cells was extracted after three days, while DNA/RNA from the knock-out cells were extracted after one week of incubation.

3.4 Extraction of DNA for PCR and RNA for qPCR

Electroporated SHK-1 cells were harvested and collected when the incubated cells in the 12-well plate reached confluency. Prior gDNA extraction, the media of the cell culture was removed and 250 µl Trypsin-EDTA (0.05%) was added to the cell plate to detach the cells. 1 ml of fresh media was added to the plate, and the cell solution was transferred to a tube before centrifugation at 200 x g for 5 min. The supernatant was discarded, and the pellet was added 100 µl QuickExtract DNA Extraction Solution (Biosearch Technologies #QE09050), and we followed the steps of the protocol of Lucigen QuickExtract DNA Extraction Solution. The QuickExtract DNA Extraction Solution eliminates the need for centrifugation, toxic chemicals, and column-based methods. This is achieved through the heat treatment, which effectively lyses the cells, releases the DNA, and degrades compounds that can inhibit the amplification process. The tube containing the cell pellet dissolved in QuickExtract DNA Extraction Solution was heated at 65 °C for 6 min, 98 °C for 2 min and stored at 4 °C, ready for use for PCR.

For the extraction of RNA from the SHK-1 cells, we utilized the Monarch Total RNA Miniprep Kit Protocol (NEB #T2010) following the protocol of the manufacturer. During the initial stages of RNA extraction from the cell plates, we made some exceptions to the protocol. First, we aspirated the cell media from the wells and washed the cells with 1 ml of DPBS. Next, we added 300 µl of Lysis Buffer to the cell plate to detach the cells, which were then scraped from the bottom and collected in a gDNA removal column with a collection tube. Finally, we followed Monarch's protocol precisely for the RNA binding and elution steps. The protocol included removal of gDNA, addition of ethanol and washing of RNA.

Sample concentration of DNA and RNA were assessed using NanoDrop 8000 spectrophotometer (BioNordika #ND-8000-GL). Samples containing gDNA were amplified by PCR, whereas samples containing RNA were subjected to cDNA synthesis followed by qPCR or RNA-seq.

3.5 Sample preparation prior Sanger sequencing

gDNA from cells exposed to CRISPR/Cas9 gene editing was used to run a PCR with 15 µl DreamTaq Green PCR Master Mix 2x (ThermoFisher #K1082), 1 µl Forward Primer (10 pmol), 1 µl Reverse Primer (10 pmol), 5 µl DNA template and 8 µl distilled water (see Table 6 for primers). The following program was used: 95°C in 3 min, followed by 30 cycles of 95°C in 30 sec, 59°C in 30 sec, 72°C in 45 sec, completed by 72°C in 2 min. To visualize and separate the amplified samples from PCR, we utilized agarose gel electrophoresis. The PCR products were run on a 0.8% agarose gel. To clean the PCR product, we excised the bands from the agarose gel and performed subsequent gel extraction steps following the GeneJET Gel Extraction Kit protocol (ThermoFisher #K0691). In short, we dissolved the gel, extracted and purified the gDNA and then stored the gDNA at -20°C. We utilized NanoDrop 8000 spectrophotometer to determine the presence of the required gDNA concentration for Sanger sequencing.

The gDNA from the *lxr* knock-out samples sent for Sanger sequencing were amplified by PCR using F2/R2 primer pairs and sequenced with F3 primer. The primer sequences are attached in Table 6 (Appendix C), and the primer locations are visualized in Figure 20 (Appendix D). 5 µl of the purified PCR sample (5 ng/µl) were combined with 5 µl of the F3 primer (5 pmol/µl), following the instructions of Mix2Seq Kit (Eurofins Genomics). The CRISPR edited samples were sent to Eurofins Genomics for Sanger sequencing (Eurofins Genomics – Genomic services by experts).

By examining the Sanger sequencing results, the editing efficiency of CRISPR-injected cells was determined by analyzing the interference of CRISPR edits using Synthego ICE (Synthego, 2019). By comparing control samples with treated samples, the indel percentage was calculated, indicating the success rate of the gRNAs tested. Furthermore, Sanger sequencing analysis using Synthego ICE provided insights into the most frequently occurring indels and the specific location where the Cas9 protein had cut the DNA sequence.

3.6 Measurements of changes in RNA levels of the *lxr* gene after CRISPR editing

To evaluate alternation in gene regulation, we assessed RNA measurements using both CRISPR/Cas9 knock-out and CRISPR/dCas9 knock-down samples in qPCR. RNA was extracted from the knock-out cells one week after electroporation, whereas for the knock-down cells, RNA extraction was performed three days after electroporation. Each experimental

treatment was conducted with 2 technical replicates along with their corresponding control samples. The purified RNA samples were converted into cDNA followed by qPCR measurements. To create cDNA from RNA, we used ProtoScript II First Strand cDNA Synthesis Kit (BioLabs #E6560). We followed the standard protocol provided by the manufacturer, which involved mixing RNA with Oligo(d)T, denaturing the mixture through heating, adding enzymes, and complete the cDNA synthesis. The completed cDNA samples were stored at -20 °C until use. Prior qPCR, tubes were prepared by adding 10 µl SYBR Select Master Mix (ThermoFisher #4472908), Forward and Reverse Primer (1 µl each, 5 pmol), 5 µl cDNA template and 3 µl distilled water. The CFX96 Real-Time PCR Detection System (BioRad #10000068706) was set to following program: 50 °C in 2 min, 95 °C in 2 min, followed by 40 cycles of 95 °C in 15 sec and 63 °C in 1 min. Then, completed by 65 °C in 5 sec and 95 °C in 5 sec, and set to 4 °C until further use. To ensure successful gene editing, qPCR measures the expression of reference genes concomitant as samples containing Cas9 and dCas9, making it easy to follow the alternation in gene regulation. In the current study, primer pairs for EF1A, RPL1 and 18S reference genes are used as internal control because of their previously demonstrated transcriptional stability in Atlantic salmon (Jorgensen et al., 2006). The validated primers tested by Jørgensen and colleagues can be found in Table 8, Appendix C.

To analyze the qPCR results we conducted variants of the C_q values, indicating the different amounts of mRNA levels among the samples. The qPCR machine presented the C_q values of the samples after 40 cycles of amplification. The difference in the C_q values between the gene of interest and the reference gene gives us the ΔC_q value:

$$\Delta \mathrm{Cq} = \mathrm{Cq}(\text{target gene}) - \mathrm{Cq}(\text{reference gene})$$

To find the alternation in gene regulation between treated sample and control ($\Delta\Delta C_q$ value), we utilized this equation:

$$\Delta\Delta \mathrm{Cq} = 2^{\Delta \mathrm{Cq}(\text{treatment}) - \Delta \mathrm{Cq}(\text{control})}$$

The calculations are based on the article published by Haimes and Kelley (2015). An example of the calculations of the ΔC_q and $\Delta\Delta C_q$ values is attached in Table 10, Appendix E. The calculated values obtained from qPCR analysis provided valuable insight into the impact of a gene treatment on the mRNA level of the gene. Since we had only 2 technical replicates in the qPCR measurements, we were unable to use statistical analysis to determine the significance of the $\Delta\Delta C_q$ values.

3.7 RNA-seq to investigate genes affected by *lxr* knock-out

To obtain a more comprehensive understanding of the potential genes and pathways impacted by CRISPR gene editing, samples with extracted RNA from the SHK-1 cells were submitted for RNA-seq at Novogene. A total of 12 RNA samples, comprising 4 knock-out, 4 knock-down, and 4 control samples, were chosen for RNA-seq (Table 1). The selected samples were a combination of various technical replicates (repeated measurements of the same sample) and biological replicates (samples done on different days and with different cell passage number). The RNA samples were selected based on their high amount of purified RNA (Table 11, Appendix F). Unfortunately, there were no control samples sent for RNA-seq from the same experimental day we performed the knock-down experiment. The consequences of this incident will be further discussed in results and discussion.

Table 1: Overview of the 12 RNA samples sent for RNA-seq. The samples consist of 4 knock-out, 4 knock-down and 4 control samples from various experimental parallels. The color code explains the different days of conducting the experiments. The samples were conducted in both technical and biological parallels.

Treatment	Experimental day	Replicates	
Control #1	1	Technical	Biological x2
Control #2	1		
Control #3	2	Technical	
Control #4	2		
Knock-out #1	1	Technical	Biological x2
Knock-out #2	1		
Knock-out #3	2	Technical	
Knock-out #4	2		
Knock-down #1	3	-	Biological x3
Knock-down #2	4	-	
Knock-down #3	5	Technical	
Knock-down #4	5		

To ensure the integrity of the 12 samples during transportation to RNA-seq, they were carefully packaged in Eppendorf tubes and secured with parafilm to prevent any damage. The samples sent for sequencing had RNA concentrations ranging from 58 ng/µl to 161 ng/µl, and amounts ranging from 18 µl to 23 µl (details in Table 11, Appendix F). The RNA sequencing was prepared using stranded mRNA library, and the coverage of the samples were set to 20 million reads per sample (6 Gb data per sample).

We utilized various bioinformatic and statistical software to analyze and interpret the RNA-seq data (Table 4, Appendix A). First, we trimmed the low-quality reads and adapters using *fastp* tool (S. Chen et al., 2018). Then, we used 'Salmon' to quantify gene expression levels, which implements a k-mer based gene expression quantification strategy (Patro et al., 2017). The expression of transcripts was then summarized into gene levels.

For quality control of data we performed a PCA using the prcomp() function in R. The input for this analysis was vst-normalized gene expression values obtained from the R-package DESeq2. Finally, differential gene expression analyses were carried out using the DESeq2 R-package. We conducted differential gene analyses to identify variations in gene expression across *lxr* knock-out, *lxr* knock-down, and control samples. To extract genes with statistically significant expression alternations, we utilized filtering techniques in R using adjusted p-value < 0.1. Comparison of the TPM values made it possible to examine the differential expression of *lxr* across the various treatments, as well as exploring the downstream genes affected by its regulation.

4 Results

4.1 Gene alignment of *lxr* and *lxr-like*

A duplication event of *lxr* likely resulted in occurrence of the *lxr-like* gene during the whole genome duplication of the Atlantic salmon. To explore the variations resulting from the duplication event, we performed a gene alignment of the *lxr* and *lxr-like* ohnologues (Figure 8).

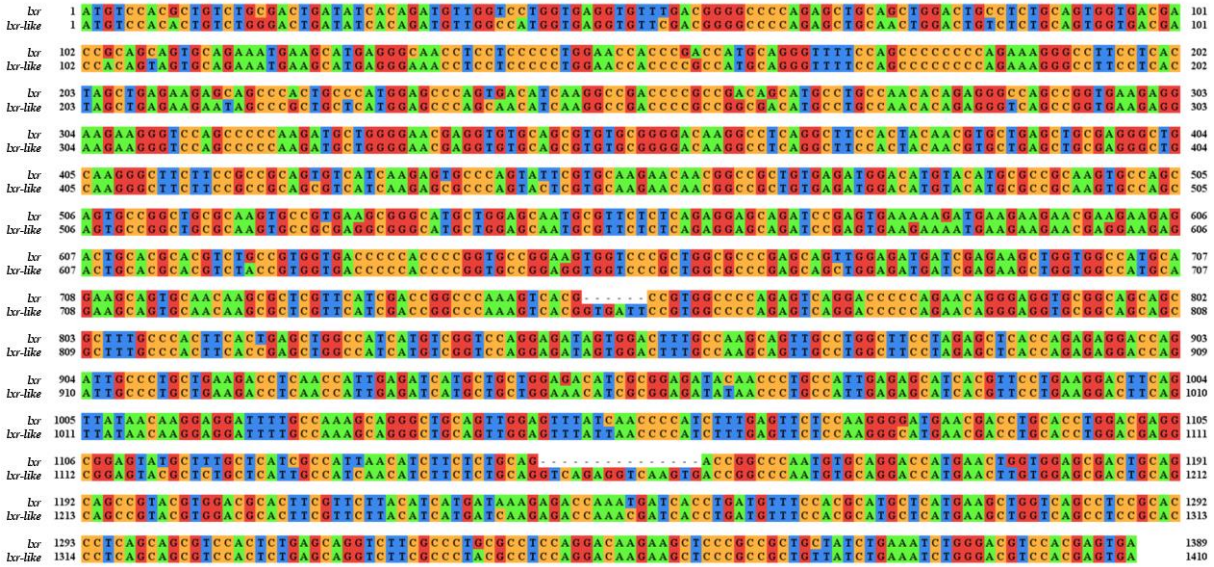


Figure 8: Alignment of the nucleotide sequences of *lxr* and *lxr-like*. The *lxr* gene consists of 1397 nucleotides, while the *lxr-like* gene consists of 1410 nucleotides, resulting in two white dashed gaps in the sequence. The four colors indicate the four different nucleotides (A, T, C, G). The gene alignment shows a high degree of similarity between the two genes. The multiple sequence alignment of *lxr* and *lxr-like* was generated using ClustalW (Larkin et al., 2007). The Figure of the gene alignment was created using Jalview (Waterhouse et al., 2009).

The *lxr* and *lxr-like* genes show high sequence similarity (Figure 8). Despite the scarcity of published information on the *lxr-like* gene, it is reasonable to assume that it performs a comparable function to the *lxr* gene, given the high degree of sequence identity shared by the two genes. The presence of the gene similarity in the *lxr* and *lxr-like* genes poses a challenge when conducting further gene editing experiments, as it makes it difficult to create distinct sequences for targeting the correct genes. Furthermore, we aimed to explore the relative expression of the *lxr* ohnologues in Atlantic salmon.

4.2 Tissue distribution of gene expression of *lxr* and *lxr-like*

Two homologous genes are often differently regulated in a cell, where one homologue is often expressed more than the other homologue. To investigate the differences in gene expression of *lxr* and *lxr-like* we conducted a gene analysis using *BgeeDB* package in R (Bastian et al., 2020). A detailed description of the bioinformatic analysis is attached in Appendix G. The expression of *lxr* gene was visibly high compared to the expression of *lxr-like* gene in all tissues examined (Figure 9).

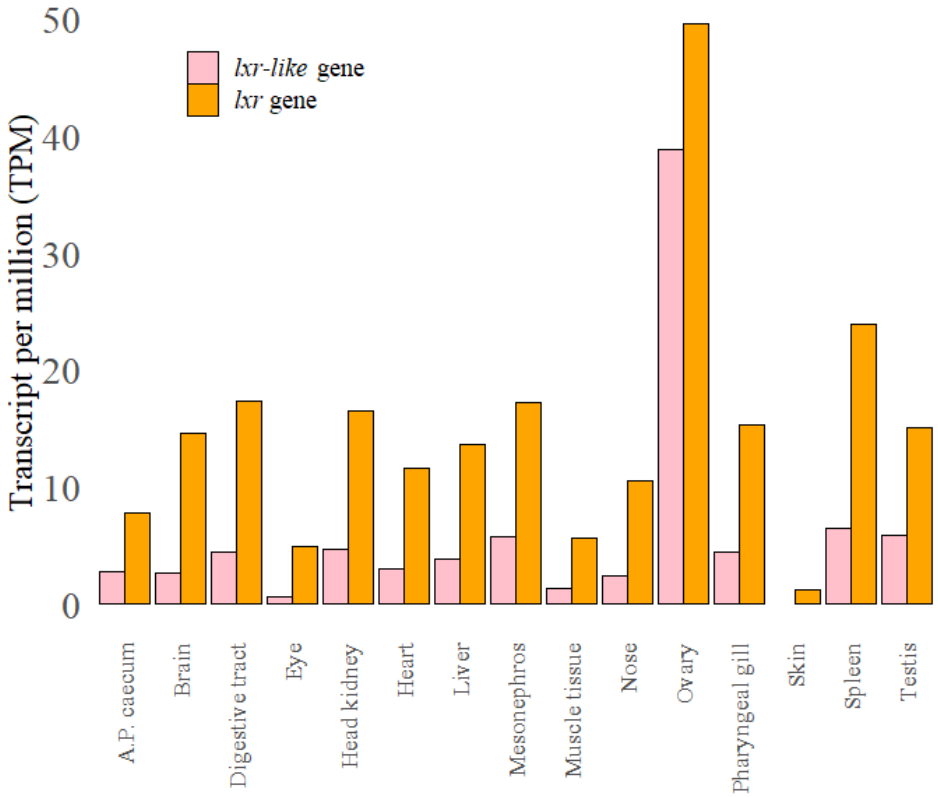


Figure 9: Differential gene expression of *lxr* and *lxr-like* in tissues of Atlantic salmon. Differentiation of gene expression in respectively *lxr* (orange) and *lxr-like* (pink) in Atlantic salmon tissues are presented with transcript per million (TPM). The *lxr* gene is consistently higher expressed compared with the *lxr-like* gene in all tissues examined. A.P. caecum; actinopterygian pyloric caecum. The Figure is created using R.

Across most tissues, *lxr* was expressed approximately three times more than *lxr-like*, including the liver and head kidney (Figure 9). Information about the fish used in the dataset can be found in section 3.1. The expression of both *lxr* and *lxr-like* genes were significantly higher in the ovary than in other tissues.

4.3 Optimalization of cell culturing

To optimize cell growth and viability we tested different plate sizes and input cell numbers of SHK-1 cells, ensuring dense cell growth and sufficient DNA/RNA yield for subsequent sequencing. Three plate sizes were tested: 6-plate, 12-plate, and 24-plate. The number of cells cultured varied between $2x10^5$ cells and $1x10^6$ cells. The cell culture optimalization experiment were conducted in two biologic parallels showing equal results. Figure 10 presents information from one of the parallels.

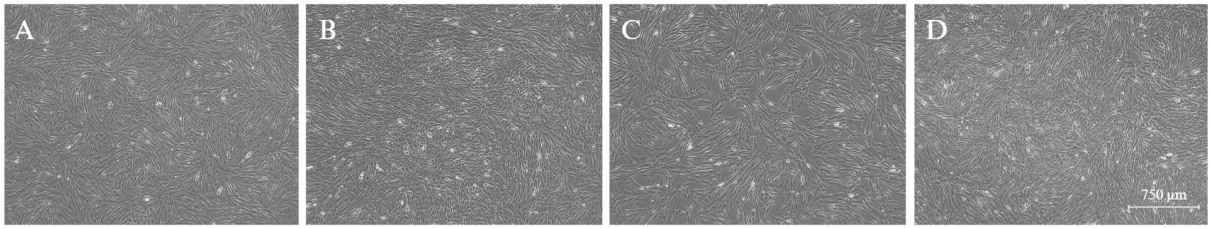


Figure 10: Photomicrograph of SHK-1 cells grown in different cell number and plate sizes. Plate A contains $1x10^6$ cells cultured in a 6-plate. Plate B and C are 12-plates containing $4.5x10^5$ cells and $3x10^5$ cells respectively. Plate D contains $2x10^5$ cells cultured in a 24-plate. All plates show dense growth of cells. The 12-plate with $3x10^5$ cells (C) has the least confluent cells. The EVOS microscope show a definition size of 750 µm with use of 4x/10x lens.

The SHK-1 cells had grown sufficient in all conditions tested (Figure 10). To obtain more information about the characteristics of the cells in the different conditions, we carried out cell counting, viability test and RNA extraction to examine the RNA concentrations (Table 2).

Table 2: Optimalization of SHK-1 cell culturing with various combinations of cells and plate sizes. The table show information about input cell number, plate size, cell count, viability, and RNA concentration after RNA extraction from the cells. Well A exhibits the highest cell count, viability, and RNA concentration.

Well	Input cell number	Plate size	Cell count	Viability	RNA concentration
A	1×10^6	6	1.56×10^6	100 %	439 ng/µl
B	4.5×10^5	12	5.65×10^5	94 %	129 ng/µl
C	3×10^5	12	4.29×10^5	93 %	103 ng/µl
D	2×10^5	24	2.83×10^5	89 %	45 ng/µl

As expected, Table 2 presents well A with the highest input cell number, as the well containing most live cells and highest RNA concentration. The large number of cells required in 6-well plates are not suitable for parallel experiments. 12-well plates with $4.5x10^5$ cells offer optimal growth and confluency. 24-well plates lack sufficient RNA for sequencing analysis. In summary, the 12-well plate with $4.5x10^5$ cells was found most optimal for this experiment and was therefore used through the rest of the procedure.

4.4 Optimalization of primer pairs

4.4.1 PCR primer pairs

To amplify the region containing the potential edit, primers on each side of the gRNA cut site were designed in Primer3, ordered by ThermoFisher, and tested using PCR and gel electrophoresis. Two primer pairs for *lxr* and four primer pairs for *lxr-like* were tested and visualized using gel electrophoresis (Figure 11).

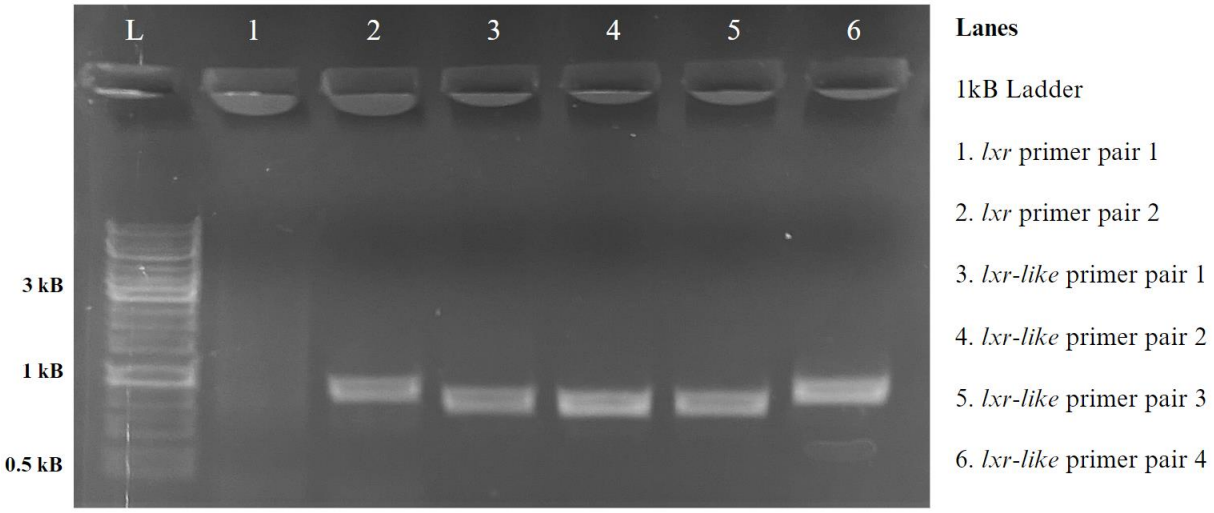


Figure 11: Gel electrophoresis to test PCR primer pairs for *lxr* gene and *lxr-like* gene. Two primer pairs were tested for *lxr* (1-2) and four primer pairs for *lxr-like* (3-6). Primer pair 2 is successful for *lxr* and all four primer pairs for *lxr-like* show successful bands at expected size. The forward and reverse primer sequences tested are attached in Table 6, Appendix C.

We tested two primer pairs for *lxr*, where only primer pair 2 showed success. The band size of primer pair 2 in Figure 11 correlates to the distance (841 bp) between the forward primer and the reverse primer. For the *lxr-like* gene we tested four primer pairs, which all four indicated successes in amplification. All primer pairs tested are attached in Table 6 in Appendix C.

4.4.2 qPCR primer pairs

To explore regulation of the *lxr* genes, we aimed to design primers for *lxr* and *lxr-like* to measure the level of RNA in the samples. The primer pairs for qPCR were designed using Primer3 web tool and ordered by ThermoFisher. Only two of the tested primer pairs were able to amplify the *lxr* gene, and all the primer pairs designed for *lxr-like* gene had C_q values above 40 or missing entirely, indicating that the primers were not working properly. As a result of the low C_q values obtained, and time constraints, we decided to exclude the *lxr-like* gene from the remaining experiments and instead focus on the *lxr* gene.

Afterwards, we realized that a mistake was made during the design of the primer pairs, that led to some of the primers being placed in an intron. The two successful primer pairs for the *lxr* gene targeted the coding sequence (CDS). So, the lack of amplification of the *lxr-like* gene was most likely due to the use of inappropriate primer pairs rather than the gene being expressed at low levels. If we had more time, we would have designed new primer pairs targeting the coding sequence of *lxr-like*. All qPCR primer pairs tested are attached in Table 7, Appendix E.

The qPCR results of control cDNA indicated that the two primer pairs designed within the coding sequence of *lxr*, hereafter referred to as CDS1 and CDS2, were successful. We designed primer pairs covering both the 5'-end (CDS1) and the 3'-end (CDS2) of the gene, to examine if primer pairs have different efficiency various places in the gene. Both CDS1 and CDS2 primer pairs showed successful amplification of the *lxr* gene with C_q values below 28 in the qPCR results (Table 9, Appendix E). The melting peak analysis had single peaks and therefore we assumed that the primer pairs were specific. Among the two specific primer pairs, CDS1 exhibited the lowest C_q-values and was selected as the most suitable primer pair for further analysis.

Reference genes were used in all qPCR plates to detect any errors during the experiment (Table 8, Appendix C). Out of three tested reference genes, EF1A primer pair showed the most constant gene expression in treated samples and controls.

4.5 Cell viability after CRISPR incorporation

To deliver the CRISPR complex to the SHK-1 cells, we utilized electroporation. This treatment can be tough for the cells. We observed that the SHK-1 cells exposed to the electroporation settings used by Gratacap et al. (2020; 1600 V 10 msec 3 pulses) led to massive cell death testing two parallels. The highest cell viability was obtained using 1400 V 20 msec 2 pulses. In one of the two wells testing 1600 V 10 msec 3 pulses, only floating dead cells were observed. In the other well, some cells showed signs of survival, although most of them were also dead. We conducted a viability test of the surviving cells using Trypan Blue stain (0.4%), revealing a survival rate of 11% by using the Countess machine. Despite this low rate, we decided to investigate the possibility that although cell viability was lower, transfection efficiency might have been higher. We extracted the surviving knock-out cells from the 1600V 10 msec 3 pulses experiment, and prepared DNA samples for further analysis to test the transfection efficiency.

The second electroporation setting, 1400 V 20 msec 2 pulses, showed high viability of SHK-1 cells. Figure 12 shows photomicrographs of *lxr* knock-out cells (A) and *lxr* knock-down cells (B) after electroporation with CRISPR components, compared to control cells (C).

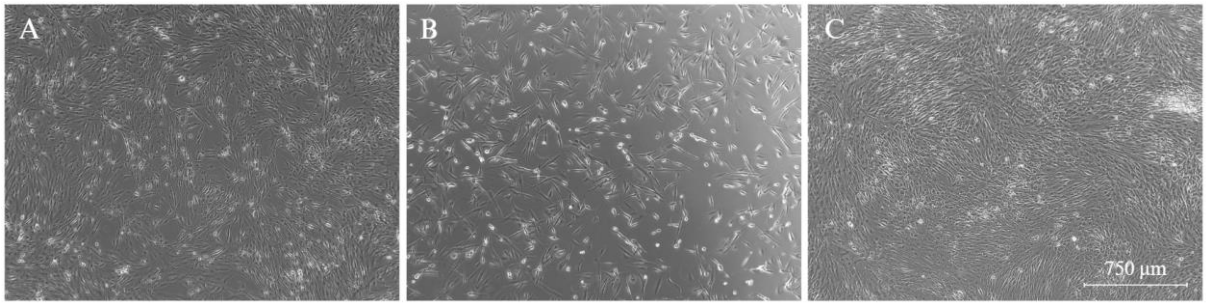


Figure 12: Photomicrograph of *lxr* knock-out, *lxr* knock-down, and control SHK-1 cells after CRISPR incorporation by electroporation. The *lxr* knock-out cells (A) show more confluent growth compared to *lxr* knock-down cells (B). The control plate (C) has the highest confluency of SHK-1 cells. The electroporation setting was 1400 V 20 msec 2 pulses. All three plates show adequate cell growth, indicating that the SHK-1 cells survived the electrical shock by the Neon transfection system. The EVOS microscope use a definition size of 750 µm with use of 4x/10x lens.

Although many cells died due to the hard treatment of electroporation, Figure 12 shows satisfactory growth of both *lxr* knock-out and *lxr* knock-down cells. The cell density in the light microscopy images indicated that the knock-out cells (A) and knock-down cells (B) grew less confluent compared to the *lxr* control cells (C). DNA from the knock-out cells and control cells was extracted and prepared for Sanger sequencing to test the transfection efficiency of the CRISPR/Cas9 complex with different gRNAs.

4.6 Estimation of CRISPR knock-out efficiency

We assessed the transfection efficiency in *lxr* knock-out cells by analyzing Sanger sequencing data using the Synthego ICE web tool. We tested two different electroporation settings and compared the results (Table 3). Unfortunately, the sample quality of the cells exposed to 1600 V 10 msec 3 pulses was too low to obtain any meaningful results in the Synthego ICE tool. However, the analysis of cells exposed to 1400 V 20 msec 2 pulses showed a high transfection efficiency. Given the high cell viability and successful sequencing results obtained with this electroporation setting, we decided to continue using it and terminated testing of the other one (1600 V 10 msec 3 pulses).

Table 3: Sanger sequencing to test indel efficiency of gRNAs for *lxr* knock-out cells using Synthego ICE tool. The gRNA sequences are targeting either the (+) strand or the (-) strand. gRNA1 achieved successful indel efficiencies, while gRNA2 failed due to poor sample quality.

Gene	gRNA (included PAM)	Indel efficiency (%)		
		Sample #1	Sample #2	Sample #3
lxr	GCTGCGCAAGTGCCGTGAAGCGG (+)	34 %	59 %	60 %
lxr	CGTTGTTCTTGCACGAATACTGG (-)	-	-	-

We conducted *lxr* knock-out experiments using 1400 V 20 msec 2 pulses testing two different gRNAs, each performed in 3 biological replicates (Table 3). Analyses of CRISPR editing efficiency from Sanger sequencing reads using Synthego ICE revealed a maximum indel efficiency of 60% in the SHK-1 cells using gRNA1. However, gRNA2 was unsuccessful in generating indels due to the low quality of the samples obtained. The Sanger sequencing revealed insertions and deletions of varying lengths and positions using gRNA1. The analysis indicated a cut by Cas9 3 bases upstream of the PAM site, illustrated in Figure 13.

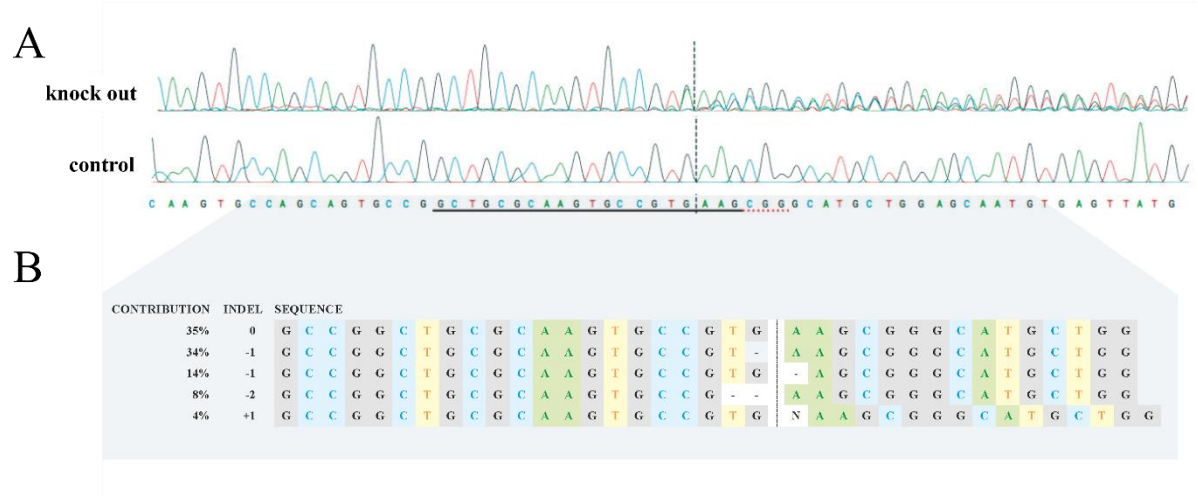


Figure 13: Sanger chromatogram and indel efficiency after CRISPR/Cas9 gene editing. The chromatogram (A) shows the gRNA sequence (black underline), the 'CGG'-PAM site (red dotted underline) and the cut side of Cas9 (vertical dotted line). The distribution of indels (B) show a total indel efficiency of 60%. The most common indel (34%) from the gene editing is one deletion of a guanin 3 bases upstream of PAM site. The chromatogram is made using ICE Synthego web tool.

The *lxr* knock-out chromatogram (Figure 13A) shows peaks with several colors and heights after Cas9 cut site. Each peak in the chromatogram has a color indicating a specific nucleic acid. Downstream of the Cas9 cut site, the chromatogram peaks are lower and double, indicating that the Sanger sequencing is reading several bases at the same time due to different indels in different cells. According to the ICE analysis (Figure 13B) the most frequent indel is a deletion of a guanin base, leading to a -1 frameshift in the gene resulting in disruption of the gene function.

Our attempts to test the efficiency of the gRNAs for knock-out of *lxr-like* using Sanger sequencing analysis have been unsuccessful. Despite using four gRNAs and testing them in 5 biological replicates, we were unable to obtain any positive results. The quality file provided by Eurofins Genomics displayed a chromatogram characterized by low and double peaks, lacking a distinct pattern in the bases. Moreover, none of the gRNAs passed the Synthego ICE test due to their poor sample quality. All gRNAs tested are attached in Table 5, Appendix B.

4.7 Both CRISPR knock-out and knock-down led to lowered mRNA levels of *lxr*

To compare the mRNA levels of *lxr* in knock-out cells and knock-down cells we ran qPCR with the CDS1 primer pair (Table 7, Appendix C). The rate of fold change ($\Delta\Delta C_q$) is presented in Figure 14. The CRISPR gene editing seemed to result in reduced levels of mRNA of the *lxr* gene compared to the control. Our test represents an initial assessment with a single biological replicate, highlighting the potential of the experiment rather than providing a conclusive outcome.

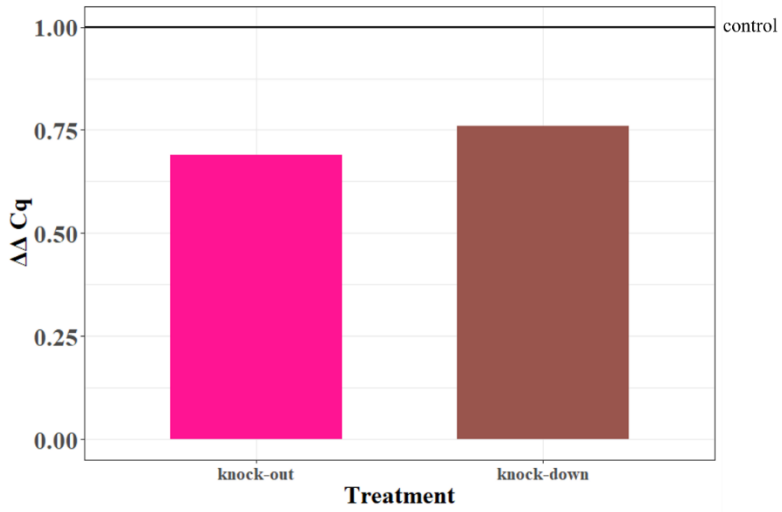


Figure 14: $\Delta\Delta C_q$-values indicated efficient downregulation of *lxr* mRNA levels in SHK-1 cell line by electroporation of CRISPR/Cas9 and CRISPR/dCas9 complex. The Figure presents mRNA levels of *lxr* measured during 40 cycles of qPCR presented with the $\Delta\Delta C_q$ fold change. The horizontal black line presents the control samples being 1. The knock-out (pink) and knock-down (brown) bars indicate decrease in mRNA levels of the *lxr* gene. Gene editing by CRISPR/Cas9 (knock-out) led to approximately 31% downregulation of *lxr* mRNA levels, while CRISPR/dCas9 knock-down resulted in approximately 24% downregulation compared to the control. CDS1 successfully amplified the gene of interest during the qPCR cycles. The Figure lack error bars since the bars are conducted from single samples and not parallels (n =1). The reference gene used was *ef1a*. Exact C_q-values from the qPCR cycles are attached in Table 9 (Appendix E).

The qPCR results in Figure 14 highlight two main findings. Firstly, an indication of that both knock-out and knock-down led to decreased mRNA levels of *lxr*. Secondly, comparing the knock-out and knock-down results indicated that knock-out is the most efficient approach to reduce mRNA levels of *lxr* in a cell culture sample. The qPCR measurements were primarily conducted to determine the rate of success of the CRISPR knock-down gene editing by dCas9.

In the original experiment, we had 3 biological replicates per group (knock-out and knock-down) and included 3 technical replicates each. However, we were unable to obtain successful qPCR results in these experiments due to non-functional primer pairs. Due to time constraints, we were only able to perform qPCR measurements using the functional primer pairs on one sample per group to assess any potential change in *lxr* mRNA levels (see Table 9).

If the knock-down samples indicated successful repression of *lxr* mRNA levels using qPCR measurements, we assumed that the remaining samples also had similar outcomes. Based on the encouraging qPCR results, we decided to send 12 samples for RNA-seq (Table 11, Appendix F), despite we only tested 4 of them using the functional primer pairs in qPCR.

4.8 The impact of *lxr* knock-out in SHK-1 cells

In this study, we utilized RNA-seq to explore the effect of *lxr* gene editing on gene expression of other genes in the SHK-1 cells. Knock-out, knock-down, and control samples were sent to Novogene for sequencing. Upon receiving the sample results, the RNA-seq files with raw data were processed into a table of gene-level counts with 55,845 genes. The experiments with knock-out and knock-down samples were conducted on different days, along with their respective control samples. Unfortunately, we only sent controls from the knock-out experiment and tried to utilize these control samples to compare with the knock-down samples. This approach could have introduced a confounding factor that could impact the outcomes, given that day to day variation can significantly affect the results.

To investigate if using the same control samples for both knock-down and knock-out experiments had any detrimental effects on the overall experiment, we performed a principal component analysis (PCA). This allowed us to gain additional insights and information from the RNA-seq (Figure 15).

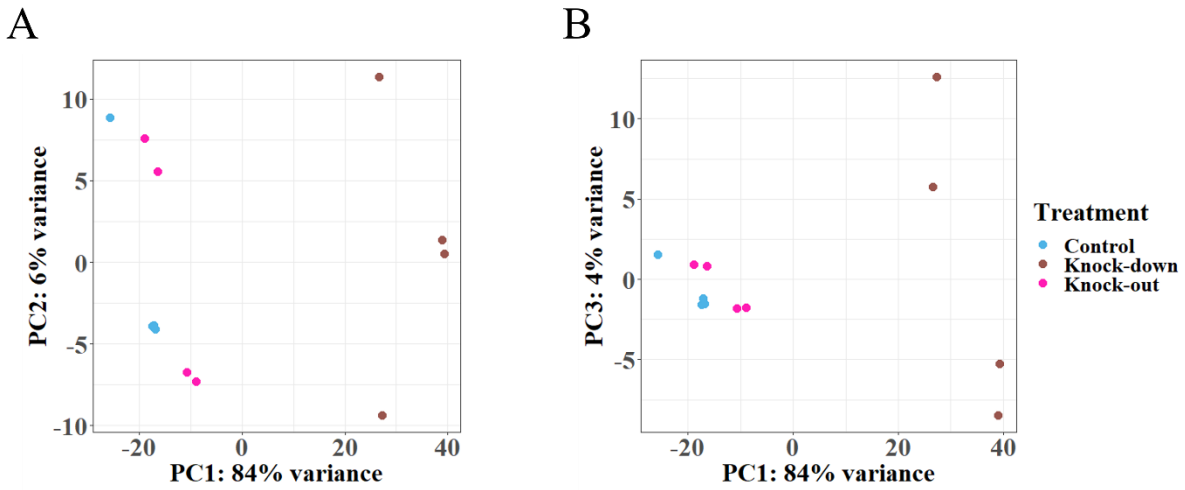


Figure 15: Two-dimensional PCA plots indicating patterns in variance between *lxr* knock-out, *lxr* knock-down and control samples. The PCA plot displays 12 samples along PC1 and PC2 (A), and PC1 and PC3 (B). The PC1 reveals 84% of the variance, PC2 6% of the variance and PC3 4% of the variance within the expression data set. The knock-down samples (brown) show no correlation with the control samples (blue), neither the knock-out samples (pink). The figure is created in R with data from the RNA-seq.

We performed a PCA using normalized gene expression quantification of knock-out, knock-down and control samples (Figure 15). As anticipated, the suboptimal knock-down controls in the analysis posed challenges in interpreting the RNA-seq results. Plotting the samples along PC1 and PC2 demonstrated that the knock-down samples were very different from both control and knock-out, separated on the PC1 axis which explained 84% of the variance. A similar pattern was seen plotting PC3 vs. PC1. It is highly unlikely that we can make significant comparisons between the knock-down samples and the knock-out control samples. Taken together, the PCA analysis suggest that the control samples cannot be used to perform further differences in gene expression analyses with knock-down cells. Thus, the subsequent RNA-seq data analysis will focus on successful *lxr* knock-out samples compared to their respective control samples.

We conducted a differential expression gene analysis to gain more information of the regulated genes after *lxr* knock-out. To obtain genes with significantly altered expression following a reduction in *lxr* expression, we utilized filtering techniques with adjusted p-value < 0.1 using R. The filtering process yielded a concise list of 44 differentially expressed genes after a *lxr* knock-out (Figure 16).

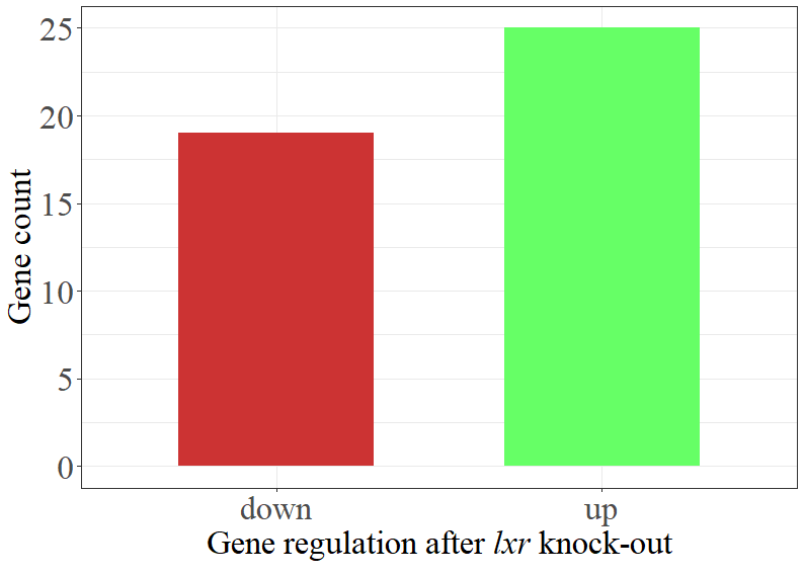


Figure 16: Distribution of significant differentially expressed genes after a *lxr* knock-out. The gene knock-out of *lxr* lead to 44 significant differentially expressed genes. The distribution of upregulated (green) and downregulated (red) genes is relatively even, albeit with a slight predominance of upregulated genes. To ensure that only genes significantly regulated were included, the Figure utilized an adjusted p-value threshold of < 0.1. The Figure is created in R.

Figure 16 shows that 19 genes were downregulated, and 25 genes were upregulated as a result of the *lxr* knock out. To gain insight into which genes were regulated differentially in the *lxr* knock-out, and of the effect size of expression change we manually curated the list of 44 differentially expressed genes and visualized the fold changes between knock-out and control salmon (Figure 17).

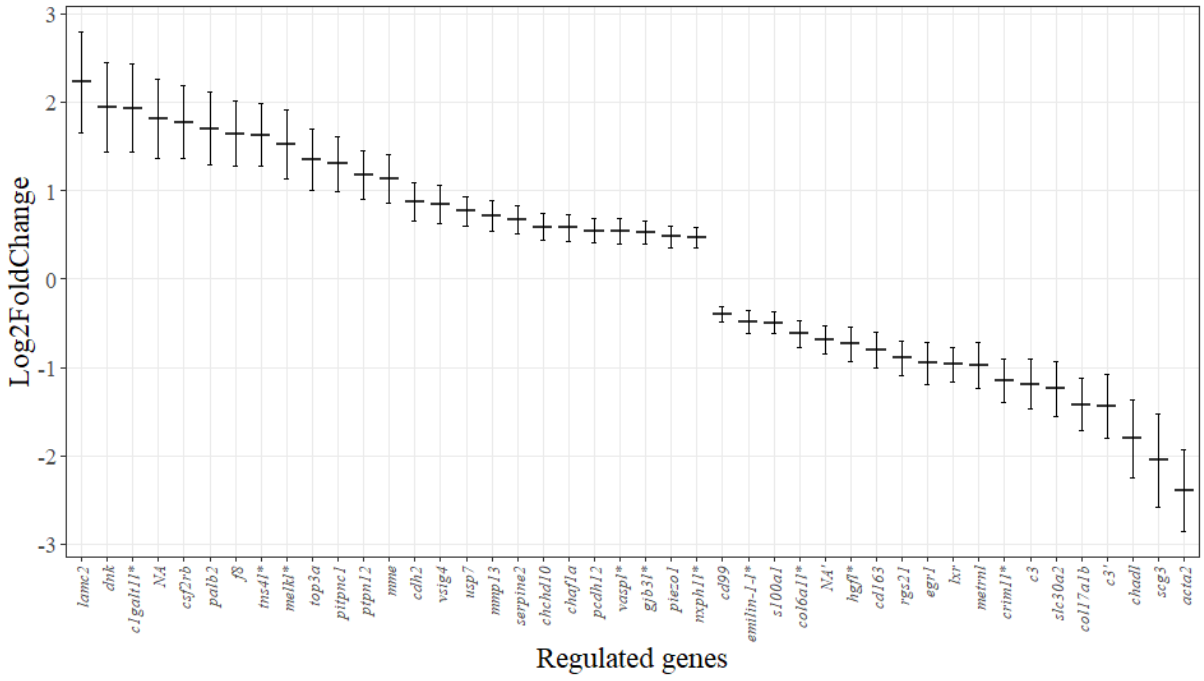


Figure 17: Significant differential expressed genes after a *lxr* knock-out in SHK-1 cells. The log2FoldChange values are visualized in all genes differentially expressed compared to the control samples. Values > 0 indicates upregulation (25 genes), and values < 0 indicates downregulation (19 genes). The expression of the *lxr* gene was downregulated with a log2FoldChange value of -0.97. The full name of the genes and the exact gene ID are described in Table 12 and Table 13 in Appendix F. Some of the gene nomenclature used in this Figure is marked with an asterisk (*) to denote its non-public status as a variant of a published gene. As such, the suffix "l" (for "like") has been appended to indicate its non-standardized nature. It should be emphasized that these designations are intended solely for internal usage within this thesis and are not intended to function as public gene names. The gene IDs marked with an (‘) shares similar names.

Figure 17 shows that the genes *lamc2*, *dnk* and *c1galt1l** are most upregulated, while the genes *acta2*, *scg3* and *chadl* are most downregulated. Similar to what was shown with qPCR, the mRNA level of the *lxr* gene was decreased.

To test if the differential expression analysis were enriched for genes linked to a particular function or pathway we performed a gene ontology enrichment-analysis. No clear association was found among the downregulated genes, while the upregulated genes showed enrichment for some functions (Figure 18).

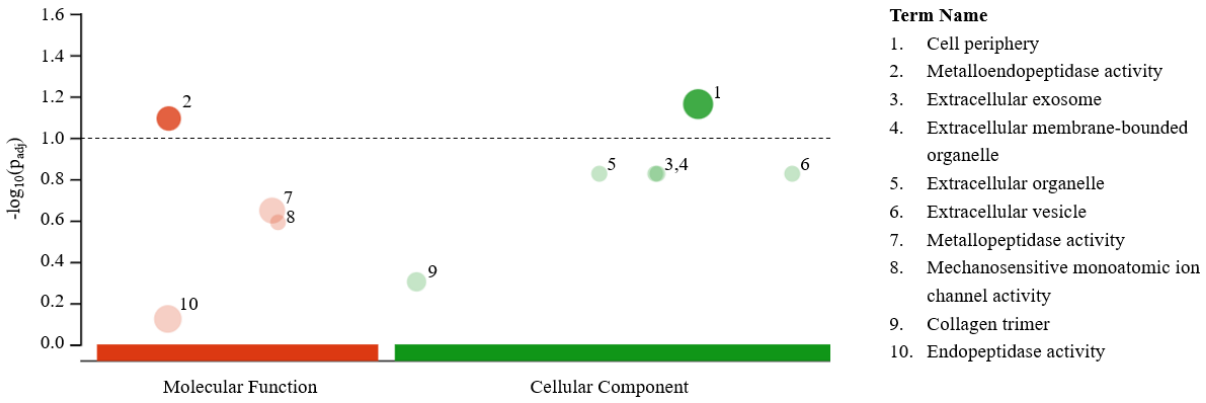


Figure 18: Gene ontology of the upregulated genes after a CRISPR knock-out of the *lxr* gene. The genes were divided into different molecular functions (red) and cellular components (green). The dotted line visualizes the p-adjusted value = 0.1. The cellular component 'cell periphery' and the molecular function 'metalloendopeptidase activity' were showing significant upregulation in the knock-out samples. The Figure is created in g:Profiler (Raudvere et al., 2019).

Figure 18 indicates that some of the genes upregulated after the *lxr* knock-out engage in cell-to-cell communication and extracellular interactions. The gene ontology analysis groups several of the upregulated genes to functions associated with extracellular organelles, cell periphery and membrane bounded activities.

4.9 Gene expression levels of lipid pathway genes in SHK-1 cells

Intriguingly, the regulated genes (Figure 17) did not include the essential genes implicated in salmon lipid metabolism such as *srebp*, *elovl*, *fads*, and *ppar*. To verify whether the selected genes had expression levels that was close to the threshold of statistical significance (p-adjusted < 0.1), an analysis of the genes of interest was conducted to assess changes in gene expression after knock-out of the *lxr* gene. The differential gene expression of the lipid metabolism genes examined resulted in logFC values close to 0, and adjusted p-value =1, indicating no significant change in gene expression. A detailed list of the genes evaluated is provided in Table 14 (Appendix F). Looking more in depth into the expression levels of the genes in the lipid metabolism pathway known to be targets of *lxr* (Table 15), revealed that the *srebp* and *elovl* genes are highly expressed in SHK-1 cells. On the contrary, one of the primary interactions for *lxr*, the members of the *ppar* gene family were not highly expressed in SHK-1 cells.

Hence, to investigate whether the lack of impact of *lxr* knock-out on lipid metabolism genes was due to low *lxr* expression levels in SHK-1 cells, we plotted the mean distribution of gene expression levels for all genes, and for *lxr* specifically (Figure 19).

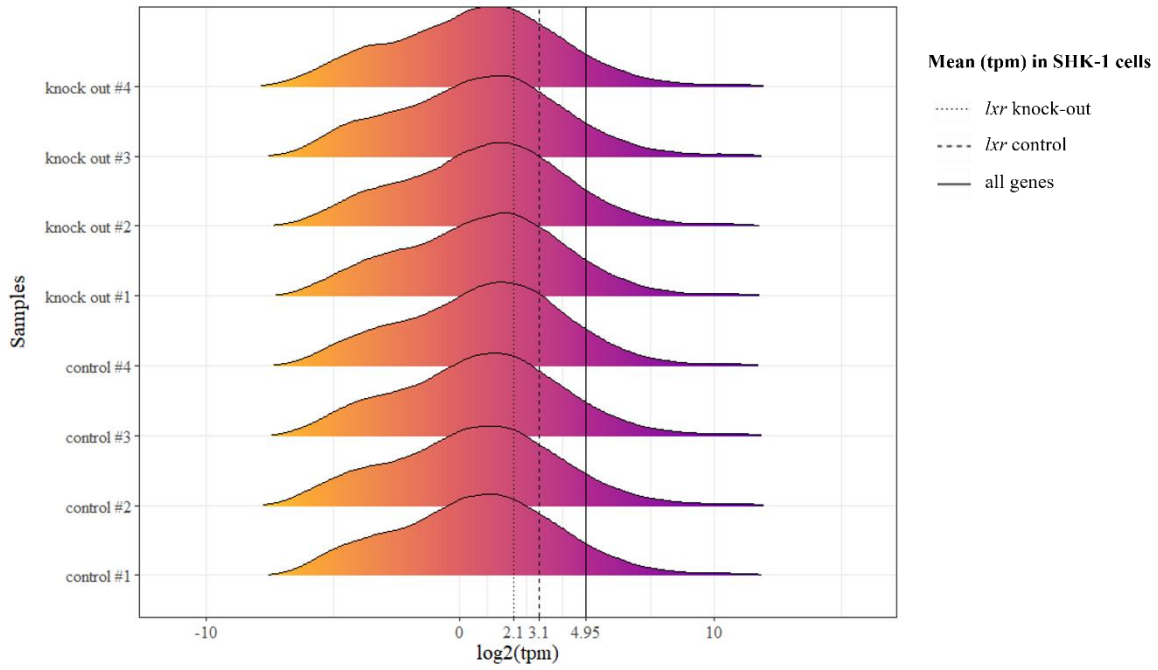


Figure 19: Mean TPM values for all genes in SHK-1 cells compared to the mean TPM value of the *lxr* gene in Atlantic salmon. The dotted line is the mean TPM value for *lxr* knock-out cells (2.1), the dashed line is the mean TPM value for *lxr* control cells (3.1), and the solid line is the mean TPM values for all genes in the SHK-1 cells (4.95). The x-axis presents TPM values as log2, and the y-axis presents the four replicates of the knock-out and control samples. The Figure is created in R with data from the RNA-seq.

We assessed the expression level of *lxr* relative to all other genes expressed in SHK-1 cells to gain a better understanding of the extent of gene expression in these cells, as they are not liver cells (Figure 19). Our analysis showed that while the expression levels of *lxr* in SHK-1 cells were not remarkably high, they were not among the lowest expressed genes either. The first dotted line represents the mean *lxr* value for the knock-out cells (2.1), while the second, dashed line represents the mean *lxr* value for the control cells (3.1). The solid line in Figure 19 was positioned far to the right because of the expression values of the most highly expressed genes, resulting in a relatively high mean expression value (4.95). The high peaks in the eight distribution plots represents the median expression of all genes. Figure 19 indicates that the expression level of *lxr* was higher than the median expression of all genes, but less than the mean gene expression of all genes. To determine the quantile placement of the *lxr* gene, we performed a quantile analysis of all genes. The analysis revealed that *lxr* was situated in the fourth quantile (Appendix G), which indicates that *lxr* is one of the top 25% most highly expressed genes in SHK-1 cells. In other words, *lxr* is expressed at a moderate level, but other genes are expressed at much higher levels.

5 Discussion

The aim of this project was two-fold: 1) to evaluate the effectiveness of the CRISPR/dCas9 system in the Atlantic salmon SHK-1 cell line, and 2) to deduce downstream targets of *lxr* by looking at the impact of knock-out and knock-down on gene expression. Although the experimental set-ups had weaknesses which leads to uncertain data, the qPCR measurements demonstrated indications of successful downregulation of the *lxr* gene using the CRISPR/dCas9 complex (Figure 14). Knock-out of *lxr* was successful and resulted in regulation of several genes, even if the identity and number of regulated genes was lower than expected. In the subsequent sections, we will discuss possibilities and development of dCas9 downregulation of gene transcription. Furthermore, we will investigate potential causes behind the unsuccessful measurement of *lxr-like*, and analyze possible patterns observed in the genes regulated after the knock-out of *lxr* (Figure 17).

5.1 CRISPR/dCas9 gene editing

To the best of our knowledge, this experiment is the first reported indication of successful CRISPR/dCas9 gene editing of a fish cell line. The regulatory pattern observed for the *lxr* gene (Figure 14) by qPCR indicates a decrease in mRNA levels three days after electroporation. This is only a preliminary indication, as we conducted the experiment with a sample size of n=1, and this result could be a result of pure coincident. It is crucial to perform qPCR experiments in multiple biological and technical replicates to really know if we succeeded.

In addition to conducting multiple parallels, it would be interesting to perform additional experiments to explore the potential for enhancing the efficacy of dCas9 gene editing. Several aspects could be optimized to explore the features of dCas9, including the amount and concentration of dCas9 and a more extensive time-course analysis. One interesting factor to study further is the timing of the RNA extraction from the dCas9-treated cells, as it could have influenced the level of gene repression observed. RNA extraction was only conducted from SHK-1 cells three days post-treatment (dpt), while the original plan was to extract RNA from cells at multiple time points (1 dpt, 3 dpt, 7 dpt, and 10 dpt). This approach would have allowed us to investigate the duration of RNP complex activity and binding, as well as the optimal time point for the maximum gene repression. Perhaps the timing we conducted the RNA extraction from the dCas9-treated cells did not correspond to the point of maximum repression of the *lxr* gene. A second approach that could have been used to enhance the efficiency of dCas9, is the

utilization of an additional domain, specifically KRAB (Figure 4). Former studies by researchers have revealed that a fusion of the KRAB domain to the dCas9 protein makes the repression of the target gene more efficient (Long et al., 2015). A third approach could involve conducting two separate dCas9 experiments, comparing the efficacy of gRNA binding to the promoter region versus the coding region of the *lxr* gene. As illustrated in Figure 4, the gRNA target determines whether the dCas9 protein blocks the binding of RNAP or causes RNAP detachment after binding. This investigation would provide valuable insights into whether gRNA binding to the promoter region or the coding region is more effective in inhibiting gene transcription. Given more time, we could have conducted these additional investigations to gain further insights into the mechanisms of successful CRISPR/dCas9 gene editing in Atlantic salmon.

When performing knock-down experiments using dCas9, it is not possible to measure the degree of gene editing with Sanger sequencing because there is no change in the DNA. One possible way of measuring gene editing efficiency is by co-transfecting the cells with an additional gRNA (Chong et al., 2021), such as the gene which encodes for albinism, solute carrier family 45, member 2 (*slc45a*). Another option is to use the green fluorescent protein (GFP) cells and use a gRNA targeting the GFP (Gratacap et al., 2020). The *slc45a* gene can be sequenced or the GFP protein observed visually, and these can be used to indicate the efficiency of transfection.

Our study aimed to investigate gene regulation *in vitro* using the CRISPR/dCas9 complex. Using *in vitro* cells in the laboratory is relatively easy, cost-effective, and controlled, as the cells grow in a regulated environment (Kaur & Dufour, 2012). However, when using a cell line that does not naturally express the target genes of interest, it may be difficult to obtain realistic results. One could also consider using primary cells for gene editing, but they often pose a challenge in terms of maintaining their stability over an extended period, making the editing process more difficult. Thus, considerations should be taken when using cell lines, and experiments where key findings are confirmed in primary cultures should always be included (Kaur & Dufour, 2012). This study seeks to contribute to a greater understanding of the molecular mechanisms underlying *in vitro* CRISPR/dCas9 editing in fish, as *in vitro* is less explored compared to *in vivo* in aquaculture. *In vivo* studies are better suited to assess the impact of CRISPR gene editing on the entire organism and offer an opportunity to explore the role of the immune system. Experiments using *in vivo* cells pose challenges such as difficulty in delivery method and ethical considerations (Dai et al., 2016). There have been published some

studies exploring the potential use of CRISPR/dCas9 in other fish species such as zebrafish and medaka (Fukushima et al., 2019; Long et al., 2015). In these studies, researchers have utilized CRISPR/dCas9 *in vivo* injection of embryos for targeted gene regulation with success. As far as we are aware, neither dCas9 *in vivo* nor *in vitro* experiments in Atlantic salmon have been published yet. It would be of great value to explore the efficacy of dCas9 *in vivo* and compare its results with *in vitro* gene editing in Atlantic salmon.

5.2 Mechanisms for degrading mRNA after CRISPR/Cas9 knock-out

CRISPR/Cas9 gene editing can lead to an unfunctional gene in several ways. Firstly, one common way is through the introduction of frameshift mutations and thereby introduction of premature stop codons (PTC), which prevents the synthesis of a functional protein if it is early in the gene (Shi et al., 2015). Furthermore, the Cas9 nuclease could induce deletion of critical exons or regulatory regions. Another mechanism is through the introduction of small insertions or deletions that can disrupt gene splicing or protein folding (D. Chen et al., 2018). Lastly, it is worth mentioning that the binding of the Cas9 protein to the DNA can sterically block the RNA polymerase from transcribing the target gene (Qi et al., 2013). This interference in transcription can lead to decreased mRNA levels, resembling a gene knock-out effect. The specific mechanism underlying the decrease in mRNA levels of the *lxr* gene in this study remains unknown. However, we suggest the explanation for the *lxr* knock-out could be the occurrence of frameshift mutations and the induction of PTCs, which are commonly associated with genetic disruption by CRISPR/Cas9 knock-out. By understanding the various mechanisms by which gene knock-out can lead to decreased mRNA levels, researchers can optimize their experimental design and interpret their results more accurately.

5.3 Genes regulated as a response to *lxr* knock-out

A knock-out of *lxr* in SHK-1 cells resulted in the significant upregulation of 25 genes (Figure 15), and notably, Figure 18 shows that a proportion of these genes were related with functions involved in extracellular functions. The gene ontology analysis indicated that cell periphery is the most significant cellular component, and the metalloendopeptidase is the most significant molecular function (Figure 18). The molecular function metalloendopeptidase degrades extracellular matrix proteins (Toriseva et al., 2007). Previous studies have shown that *lxr* activation can repress some metalloendopeptidases (Calkin & Tontonoz, 2010; Castrillo et al.,

2003; Lo Sasso et al., 2010). Since metalloendopeptidase was upregulated when the *lxr* gene was knocked out, we suggested that *lxr* typically inhibits the expression of certain metalloendopeptidases. Upon further examination of the metalloendopeptidase function, it becomes evident that it plays a pivotal role in numerous pathways and cellular functions (Cerda-Costa & Gomis-Ruth, 2014). Given that the cleavage of peptide bonds by metalloendopeptidase is essential for most physiological processes, it is difficult to specify a credible reason why this cellular function was upregulated.

To investigate if the highly upregulated genes after *lxr* knock-out (Figure 17) had any direct function involved in the extracellular matrix (Figure 18), we examined their respective gene functions. Among two of the most upregulated genes (*lamc2* and *c1galt1l*), subsequent investigation revealed that the *lamc2* gene is involved in the signaling molecules of the extracellular matrix receptor in salmon (Kanehisa & Goto, 2000; KEGG), while the *c1galt1l* gene is involved in the glycosylation process in the extracellular matrix in mice (Brockhausen et al., 2009; Brockhausen et al., 2022; Sun et al., 2021). The RNA-seq findings suggest a potential relationship between these two genes and extracellular matrix-related functions, but further investigation is certainly needed to confirm and understand these results.

Despite the crucial role of the *lxr* gene in regulating the lipid metabolism, the absence of this gene surprisingly did not result in significant changes in the regulation of other important lipid-related genes (Table 14, Appendix F). To determine whether the non-alternation of lipid metabolism genes was due to low expression levels of the genes in general, RNA-seq was used to examine gene counts (Table 15, Appendix F). The low expression of genes involved in the *ppar* pathway, which is the primary pathway for *lxr* in Atlantic salmon, could explain why the knock-out of *lxr* had limited impact on other lipid metabolism genes. The interaction between PPAR and LXR receptors is an important element in lipid gene regulation. When *ppar* genes are expressed at low levels, the binding between PPAR and LXR receptor is weakened, regardless of *lxr* expression. Consequently, a decrease in *lxr* gene expression caused by CRISPR gene editing may not be adequate to have impact on other lipid genes due to low expression levels of the *ppar*-related genes. Considering the lack of significant alterations observed in the lipid genes, investigating the consequences of *lxr* knock-out in cell types like adipocytes, known for their involvement in lipid metabolism, or in the ovary, where the *lxr* gene exhibits high expression (Figure 9), is of great interest. It is possible that the lipid genes regulated by the PPAR pathway could be more affected in a different cell type with higher

expression level of *lxr* and *ppar*-genes. According to Table 14 and Table 15, the use of the SHK-1 cell line may not be optimal for investigating lipid metabolism in Atlantic salmon.

5.4 Investigation of *lxr-like* gene expression and genetic compensation

The *lxr-like* gene amplification during qPCR was unsuccessful, most likely due to the use of suboptimal primer pairs (Table 7). As earlier mentioned, the *lxr-like* qPCR primers were designed using the full gene sequence. Consequently, four of the primers were targeting introns, obviously not providing successful results in the qPCR analysis. On the other hand, the remaining primers targeted exons, but they may have unintentionally targeted genomic regions that were either sequestered by nucleosomes or were otherwise inaccessible. It remains unclear whether the difficulties in measuring *lxr-like* was solely attributable to the primer design. Another contributing factor could be the relatively low expression levels of *lxr-like* in SHK-1 cells (Figure 19). Given the successful amplification of the samples containing the *lxr* gene, it is probable that the qPCR conditions, template concentration, or technical errors did not pose any issues during the lab work, as the *lxr-like* samples were executed concurrently with the *lxr* samples. Therefore, the success of CRISPR/Cas9 editing of the *lxr-like* gene remains unknown as mRNA levels could not be measured by qPCR.

We have delved into the concept of genetic compensation by analyzing the gene alteration of *lxr-like* following the knock-out of *lxr*. Given the central role of *lxr* in cellular processes, we hypothesized that *lxr-like* would become upregulated in response to *lxr* knock-out (Figure 9). Notably, the fold change rate of the *lxr-like* gene did not significantly increase after the *lxr* knock-out, suggesting that genetic compensation did not occur by *lxr-like* (Table 14). This lack of compensation may be attributed to a potential less pivotal role of the *lxr* gene in SHK-1 cells, as its expression is lower compared to the mean expression of other genes (Figure 19). It would be intriguing to explore the concept of genetic compensation in cell types with specifically high expression of *lxr* to observe whether the *lxr-like* gene is triggered and upregulated as a response.

5.5 Potential for dCas9 in salmon

The flexible nature of dCas9 gene editing offers immense potential for gene regulation, making it an attractive tool for future use. The 2018 report from the Norwegian Biotechnology Advisory Board emphasizes the importance of establishing a regulatory framework that enables technological progress while ensuring government supervision and control. The board

presented the rules for gene editing by a three-tiered system covered by gene modification organism (GMO) regulations (Norwegian Biotechnology Advisor Board, 2018). What is particularly promising is that gene editing with dCas9 fits into the exempted category of organisms with temporary, non-heritable changes. The benefits of dCas9 become increasingly evident, and it allows for less strict rules in gene editing. The reversible nature of this gene editing technique could potentially facilitate its approval for societal implementation. It is particularly promising for maximizing the potential benefits of salmon, and thus, it is crucial to explore and develop new methods that can increase the nutritional value, sustainability, and versatility of this remarkable fish. The safety and precision of dCas9 make it a less invasive tool for gene editing, and an important asset in the development of new salmon farming practices.

5.6 Utilize CRISPR to achieve a more sustainable aquaculture

Knowledge of lipid metabolism in Atlantic salmon could enable researchers to develop sustainable alternatives in aquaculture (Cruz-Garcia et al., 2009). Our hypothesis is that by increasing our understanding of the molecular mechanisms involved in lipid metabolism and regulation in fish, we can better utilize sustainable plant-based alternatives to marine products in aquaculture. Therefore, more research in this field is required to optimize the fish feed to not decrease the important fatty acids in the fish. Knowledge about *lxr*, along with its impact on important pathways in lipid metabolism, is crucial to aquaculture and manufacturing feeds.

The utilization of CRISPR/dCas9 presents various potential applications in the fish industry, encompassing the enhancement of health, productivity, and increased yields in farmed fish. This approach enables the selective regulation of gene expression associated with growth, disease resistance, and genes associated with viral or bacterial infections (Ahmed et al., 2019; Brander, 2007; Okoli et al., 2022). Furthermore, the technique allows for the engineering of fish to adapt to broader environmental conditions, such as increased tolerance to saltwater concentration or temperature variations (Miller et al., 2018; Muñoz et al., 2015). Successful implementation of these traits through selective breeding may result in increased survival rates of the fish in harsh environments, thereby enabling survival in less restrictive conditions. Considering the ongoing climate change, this outcome holds considerable significance, as it can potentially mitigate the deleterious effects of environmental disturbances on the fish industry.

6 Conclusion

CRISPR/dCas9 gene editing of *lxr* showed indications of successful downregulation. Further work is needed to conclude if gene editing with dCas9 is an effective method to use in cell lines of Atlantic salmon, because this study had too low sample number. Gene modification by dCas9 display promising results, although further research using dCas9-mediated transcription is needed to expand the possibilities of RNA-guided mechanisms in aquaculture (Dominguez et al., 2016). Exploring the potential of dCas9 gene editing is particularly interesting as it falls outside the scope of GMO regulations, which creates new opportunities to use CRISPR technology in sustainable aquaculture practices.

The successful CRISPR/Cas9 knock-out of *lxr* altered the expression of 44 other genes, however no distinct trend in the significantly altered genes was observed. Some upregulated genes were related to functions within metalloendopeptidases and the extracellular matrix, but the experiment is not sufficient to draw conclusive associations.

7 Future perspectives

Our indications of successful application of CRISPR/dCas9 gene editing in Atlantic salmon, as indicated by the qPCR analysis, provides valuable insights that can further advance research on lipid metabolism. In hindsight, we would rather have tested all twelve parallels (4 knock-down, 4 knock-out and 4 controls) in qPCR, before sending them to RNA-seq. Since we got positive results on two samples containing both knock-down and knock-out, we assumed all samples were successful, and sent the rest of the purified RNA samples for RNA-seq. Additionally, we should have tested the quality and reliability of the RNA samples before sending the samples for RNA-seq, by using for example the TapeStation system. However, it is evident that including control samples from the day of the knock-down experiment in the RNA-seq analysis would have been crucial for ensuring the reliability of the data. Despite that we did send the appropriate samples for RNA-seq analysis, we unfortunately did not receive the results in time to incorporate them into the final submission of the master's thesis. Substantiating the results from the initial qPCR analysis with RNA-seq data would be novel, as CRISPR/dCas9 gene regulation in Atlantic salmon has not been published before. Contributing to the development of new gene editing technologies in salmonids is an exciting prospect, and this finding could serve as a starting point for further significant research in the coming years.

To gain a deeper insight into the CRISPR experiment, the initial plan was to explore the differences in gRNA targeting. gRNAs targeting the 5'-UTR regions of the *lxr* gene could enhance promoter binding and subsequent inhibition of RNAP recruitment, whereas gRNA binding after the start codon could induce RNAP detachment and inhibit further gene transcription. Designing gRNAs targeting either the promoter or the coding sequence could provide valuable insights into the difference in gene editing efficiency between the two methods.

Our work represents an advancement towards a novel understanding of the transcriptional regulation of lipid metabolism in Atlantic salmon. The utilization of dCas9 in CRISPR gene editing presents a promising approach to perform gene edits in salmon, with the added advantage of potentially being more permissible under current GMO regulations. By using this technology, we can make significant progress in providing sustainable and nutritious salmon that can satisfy the needs of an expanding global population.

8 References

Ahmed, N., Thompson, S., & Glaser, M. (2019). Global Aquaculture Productivity, Environmental Sustainability, and Climate Change Adaptability. *Environ Manage*, *63*(2), 159-172. https://doi.org/10.1007/s00267-018-1117-3

Alerasool, N., Segal, D., Lee, H., & Taipale, M. (2020). An efficient KRAB domain for CRISPRi applications in human cells. *Nat Methods*, *17*(11), 1093-1096. https://doi.org/10.1038/s41592-020-0966-x

Aranda, A., & Pascual, A. (2001). Nuclear hormone receptors and gene expression. *Physiol Rev*, *81*(3), 1269-1304. https://doi.org/10.1152/physrev.2001.81.3.1269

Barman, A., Deb, B., & Chakraborty, S. (2020). A glance at genome editing with CRISPR-Cas9 technology. *Curr Genet*, *66*(3), 447-462. https://doi.org/10.1007/s00294-019-01040-3

Bastian, F. B., Roux, J., Niknejad, A., Comte, A., Fonseca Costa, Sara S., de Farias, T. M., Moretti, S., Parmentier, G., de Laval, V. R., Rosikiewicz, M., Wollbrett, J., Echchiki, A., Escoriza, A., Gharib, W. H., Gonzales-Porta, M., Jarosz, Y., Laurenczy, B., Moret, P., Person, E., . . . Robinson-Rechavi, M. (2020). The Bgee suite: integrated curated expression atlas and comparative transcriptomics in animals. *Nucleic Acids Research*, *49*(D1), D831-D847. https://doi.org/10.1093/nar/gkaa793

Brander, K. M. (2007). Global fish production and climate change. *Proceedings of the National Academy of Sciences*, *104*(50), 19709-19714. https://doi.org/doi:10.1073/pnas.0702059104

Brockhausen, I., Schachter, H., & Stanley, P. (2009). O-GalNAc Glycans. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). https://www.ncbi.nlm.nih.gov/pubmed/20301232

Brockhausen, I., Wandall, H. H., Hagen, K. G. T., & Stanley, P. (2022). O-GalNAc Glycans. In A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, & P. H. Seeberger (Eds.), *Essentials of Glycobiology* (4th ed., pp. 117-128). https://doi.org/10.1101/glycobiology.4e.10

Calkin, A. C., & Tontonoz, P. (2010). Liver x receptor signaling pathways and atherosclerosis. *Arterioscler Thromb Vasc Biol*, *30*(8), 1513-1518. https://doi.org/10.1161/ATVBAHA.109.191197

Carmona-Antonanzas, 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. https://doi.org/10.1016/j.gene.2013.10.040

Castrillo, A., Joseph, S. B., Marathe, C., Mangelsdorf, D. J., & Tontonoz, P. (2003). Liver X receptor-dependent repression of matrix metalloproteinase-9 expression in macrophages. *J Biol Chem*, *278*(12), 10443-10449. https://doi.org/10.1074/jbc.M213071200

Cerda-Costa, N., & Gomis-Ruth, F. X. (2014). Architecture and function of metallopeptidase catalytic domains. *Protein Sci*, *23*(2), 123-144. https://doi.org/10.1002/pro.2400

Chen, D., Tang, J. X., Li, B., Hou, L., Wang, X., & Kang, L. (2018). CRISPR/Cas9-mediated genome editing induces exon skipping by complete or stochastic altering splicing in the migratory locust. *BMC Biotechnol*, *18*(1), 60. https://doi.org/10.1186/s12896-018-0465-7

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890. https://doi.org/10.1093/bioinformatics/bty560

Chong, Z. X., Yeap, S. K., & Ho, W. Y. (2021). Transfection types, methods and strategies: a technical review. *PeerJ*, *9*, e11165. https://doi.org/10.7717/peerj.11165

Collet, B., & Collins, C. (2009). Comparative gene expression profile in two Atlantic salmon cell lines TO and SHK-1. *Vet Immunol Immunopathol*, *130*(1-2), 92-95. https://doi.org/10.1016/j.vetimm.2008.12.022

Corchete, L. A., Rojas, E. A., Alonso-Lopez, D., De Las Rivas, J., Gutierrez, N. C., & Burguillo, F. J. (2020). Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. *Sci Rep*, *10*(1), 19737. https://doi.org/10.1038/s41598-020-76881-x

Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., Clement, T., & Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring. *J Vet Diagn Invest*, *32*(6), 767-775. https://doi.org/10.1177/1040638720905833

Cruz-Garcia, L., Minghetti, M., Navarro, I., & Tocher, D. R. (2009). Molecular cloning, tissue expression and regulation of liver X receptor (LXR) transcription factors of Atlantic salmon (Salmo salar) and rainbow trout (Oncorhynchus mykiss). *Comp Biochem Physiol B Biochem Mol Biol*, *153*(1), 81-88. https://doi.org/10.1016/j.cbpb.2009.02.001

Cui, Y., Xu, J., Cheng, M., Liao, X., & Peng, S. (2018). Review of CRISPR/Cas9 sgRNA Design Tools. *Interdiscip Sci*, *10*(2), 455-465. https://doi.org/10.1007/s12539-018-0298-z

Dai, W. J., Zhu, L. Y., Yan, Z. Y., Xu, Y., Wang, Q. L., & Lu, X. J. (2016). CRISPR-Cas9 for in vivo Gene Therapy: Promise and Hurdles. *Mol Ther Nucleic Acids*, *5*(8), e349. https://doi.org/10.1038/mtna.2016.58

Datsomor, A. K., Wilberg, R., Torgersen, J. S., Sandve, S. R., & Harvey, T. N. (2023). Efficient transfection of Atlantic salmon primary hepatocyte cells for functional assays and gene editing. *G3 (Bethesda)*, *13*(4). https://doi.org/10.1093/g3journal/jkad039

Datsomor, A. K., Zic, N., Li, K., Olsen, R. E., Jin, Y., Vik, J. O., Edvardsen, R. B., Grammes, F., Wargelius, A., & Winge, P. (2019). CRISPR/Cas9-mediated ablation of elovl2 in Atlantic salmon (Salmo salar L.) inhibits elongation of polyunsaturated fatty acids and induces Srebp-1 and target genes. *Sci Rep*, *9*(1), 7533. https://doi.org/10.1038/s41598-019-43862-8

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 Biol*, *11*(9), 403. https://doi.org/10.1186/gb-2010-11-9-403

Dominguez, A. A., Lim, W. A., & Qi, L. S. (2016). Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol*, *17*(1), 5-15. https://doi.org/10.1038/nrm.2015.2

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

Egerton, S., Wan, A., Murphy, K., Collins, F., Ahern, G., Sugrue, I., Busca, K., Egan, F., Muller, N., Whooley, J., McGinnity, P., Culloty, S., Ross, R. P., & Stanton, C. (2020). Replacing fishmeal with plant protein in Atlantic salmon (Salmo salar) diets by supplementation with fish protein hydrolysate. *Sci Rep*, *10*(1), 4194. https://doi.org/10.1038/s41598-020-60325-7

Elagizi, A., Lavie, C. J., O'Keefe, E., Marshall, K., O'Keefe, J. H., & Milani, R. V. (2021). An Update on Omega-3 Polyunsaturated Fatty Acids and Cardiovascular Health. *Nutrients*, *13*(1). https://doi.org/10.3390/nu13010204

Fukushima, H. S., Takeda, H., & Nakamura, R. (2019). Targeted in vivo epigenome editing of H3K27me3. *Epigenetics Chromatin*, *12*(1), 17. https://doi.org/10.1186/s13072-019-0263-z

Gehl, J. (2003). Electroporation: theory and methods, perspectives for drug delivery, gene therapy and research. *Acta Physiol Scand*, *177*(4), 437-447. https://doi.org/10.1046/j.1365-201X.2003.01093.x

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

Gleditzsch, D., Pausch, P., Muller-Esparza, H., Ozcan, A., Guo, X., Bange, G., & Randau, L. (2019). PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. *RNA Biol*, *16*(4), 504-517. https://doi.org/10.1080/15476286.2018.1504546

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

Haimes, J. D., & Kelley, M. L. (2015). Demonstration of a ΔΔCq Calculation Method to Compute Relative Gene Expression from qPCR Data.

Haque, A., Engel, J., Teichmann, S. A., & Lonnberg, T. (2017). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med*, *9*(1), 75. https://doi.org/10.1186/s13073-017-0467-4

Hille, F., & Charpentier, E. (2016). CRISPR-Cas: biology, mechanisms and relevance. *Philos Trans R Soc Lond B Biol Sci*, *371*(1707). https://doi.org/10.1098/rstb.2015.0496

Hoijer, I., Emmanouilidou, A., Ostlund, R., van Schendel, R., Bozorgpana, S., Tijsterman, M., Feuk, L., Gyllensten, U., den Hoed, M., & Ameur, A. (2022). CRISPR-Cas9 induces large structural variants at on-target and off-target sites in vivo that segregate across generations. *Nat Commun*, *13*(1), 627. https://doi.org/10.1038/s41467-022-28244-5

Houston, R. D., & Macqueen, D. J. (2019). Atlantic salmon (Salmo salar L.) genetics in the 21st century: taking leaps forward in aquaculture and biological understanding. *Anim Genet*, *50*(1), 3-14. https://doi.org/10.1111/age.12748

Jorgensen, S. M., Kleveland, E. J., Grimholt, U., & Gjoen, T. (2006). Validation of reference genes for real-time polymerase chain reaction studies in Atlantic salmon. *Mar Biotechnol (NY)*, *8*(4), 398-408. https://doi.org/10.1007/s10126-005-5164-4

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, *28*(1), 27-30. https://doi.org/10.1093/nar/28.1.27

Karlson, C. K. S., Mohd-Noor, S. N., Nolte, N., & Tan, B. C. (2021). CRISPR/dCas9-Based Systems: Mechanisms and Applications in Plant Sciences. *Plants (Basel)*, *10*(10). https://doi.org/10.3390/plants10102055

Kaur, G., & Dufour, J. M. (2012). Cell lines: Valuable tools or useless artifacts. *Spermatogenesis*, *2*(1), 1-5. https://doi.org/10.4161/spmg.19885

KEGG. https://www.genome.jp/entry/sasa:106599435

Kidani, Y., & Bensinger, S. J. (2012). Liver X receptor and peroxisome proliferator-activated receptor as integrators of lipid homeostasis and immunity. *Immunol Rev*, *249*(1), 72-83. https://doi.org/10.1111/j.1600-065X.2012.01153.x

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

Kukurba, K. R., & Montgomery, S. B. (2015). RNA Sequencing and Analysis. *Cold Spring Harb Protoc*, *2015*(11), 951-969. https://doi.org/10.1101/pdb.top084970

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., & Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res*, *47*(W1), W171-W174. https://doi.org/10.1093/nar/gkz365

Lackner, M., Helmbrecht, N., Paabo, S., & Riesenberg, S. (2023). Detection of unintended on-target effects in CRISPR genome editing by DNA donors carrying diagnostic substitutions. *Nucleic Acids Res*, *51*(5), e26. https://doi.org/10.1093/nar/gkac1254

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, *23*(21), 2947-2948. https://doi.org/10.1093/bioinformatics/btm404

Le Rhun, A., Escalera-Maurer, A., Bratovic, M., & Charpentier, E. (2019). CRISPR-Cas in Streptococcus pyogenes. *RNA Biol*, *16*(4), 380-389. https://doi.org/10.1080/15476286.2019.1582974

Li, H., Yang, Y., Hong, W., Huang, M., Wu, M., & Zhao, X. (2020). Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal Transduct Target Ther*, *5*(1), 1. https://doi.org/10.1038/s41392-019-0089-y

Li, X., Sun, B., Qian, H., Ma, J., Paolino, M., & Zhang, Z. (2022). A high-efficiency and versatile CRISPR/Cas9-mediated HDR-based biallelic editing system. *J Zhejiang Univ Sci B*, *23*(2), 141-152. https://doi.org/10.1631/jzus.B2100196

Li, Z., Xiong, X., & Li, J.-F. (2019). The working dead: repurposing inactive CRISPR-associated nucleases as programmable transcriptional regulators in plants. *aBIOTECH*, *1*(1), 32-40. https://doi.org/10.1007/s42994-019-00003-z

Liao, C., & Beisel, C. L. (2021). The tracrRNA in CRISPR Biology and Technologies. *Annu Rev Genet*, *55*, 161-181. https://doi.org/10.1146/annurev-genet-071719-022559

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., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R. A., von Schalburg, K., Rondeau, E. B., Di Genova, A., Samy, J. K., . . . Davidson, W. S. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, *533*(7602), 200-205. https://doi.org/10.1038/nature17164

Lo Sasso, G., Celli, N., Caboni, M., Murzilli, S., Salvatore, L., Morgano, A., Vacca, M., Pagliani, T., Parini, P., & Moschetta, A. (2010). Down-regulation of the LXR transcriptome provides the requisite cholesterol levels to proliferating hepatocytes. *Hepatology*, *51*(4), 1334-1344. https://doi.org/10.1002/hep.23436

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

Miller, D. D., Ota, Y., Sumaila, U. R., Cisneros-Montemayor, A. M., & Cheung, W. W. L. (2018). Adaptation strategies to climate change in marine systems. *Glob Chang Biol*, *24*(1), e1-e14. https://doi.org/10.1111/gcb.13829

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. *Biochim Biophys Acta*, *1811*(3), 194-202. https://doi.org/10.1016/j.bbalip.2010.12.008

Montgomery, S. B., Sammeth, M., Gutierrez-Arcelus, M., Lach, R. P., Ingle, C., Nisbett, J., Guigo, R., & Dermitzakis, E. T. (2010). Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature*, *464*(7289), 773-777. https://doi.org/10.1038/nature08903

Muñoz, N. J., Farrell, A. P., Heath, J. W., & Neff, B. D. (2015). Adaptive potential of a Pacific salmon challenged by climate change. *Nature Climate Change*, *5*(2), 163-166. https://doi.org/10.1038/nclimate2473

Naeem, M., Majeed, S., Hoque, M. Z., & Ahmad, I. (2020). Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing. *Cells*, *9*(7). https://doi.org/10.3390/cells9071608

Nambiar, T. S., Billon, P., Diedenhofen, G., Hayward, S. B., Taglialatela, A., Cai, K., Huang, J. W., Leuzzi, G., Cuella-Martin, R., Palacios, A., Gupta, A., Egli, D., & Ciccia, A. (2019). Stimulation of CRISPR-mediated homology-directed repair by an engineered RAD18 variant. *Nat Commun*, *10*(1), 3395. https://doi.org/10.1038/s41467-019-11105-z

Naylor, R. L., Hardy, R. W., Bureau, D. P., Chiu, A., Elliott, M., Farrell, A. P., Forster, I., Gatlin, D. M., Goldburg, R. J., Hua, K., & Nichols, P. D. (2009). Feeding aquaculture in an era of finite resources. *Proc Natl Acad Sci U S A*, *106*(36), 15103-15110. https://doi.org/10.1073/pnas.0905235106

Norwegian Biotechnology Advisor Board, T. (2018). A forward-looking regulatory framework for GMO.

Okoli, A. S., Blix, T., Myhr, A. I., Xu, W., & Xu, X. (2022). Sustainable use of CRISPR/Cas in fish aquaculture: the biosafety perspective. *Transgenic Res*, *31*(1), 1-21. https://doi.org/10.1007/s11248-021-00274-7

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*, *14*(4), 417-419. https://doi.org/10.1038/nmeth.4197

Peirson, S. N., & Butler, J. N. (2007). Quantitative polymerase chain reaction. *Methods Mol Biol*, *362*, 349-362. https://doi.org/10.1007/978-1-59745-257-1_25

Peng, J. (2019). Gene redundancy and gene compensation: An updated view. *J Genet Genomics*, *46*(7), 329-333. https://doi.org/10.1016/j.jgg.2019.07.001

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173-1183. https://doi.org/10.1016/j.cell.2013.02.022

R Core Team, K. (2021). R: Regulatory Compliance and Validation Issues.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*, *8*(11), 2281-2308. https://doi.org/10.1038/nprot.2013.143

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*, *47*(W1), W191-W198. https://doi.org/10.1093/nar/gkz369

Robertson, F. M., Gundappa, M. K., Grammes, F., Hvidsten, T. R., Redmond, A. K., Lien, S., Martin, S. A. M., Holland, P. W. H., Sandve, S. R., & Macqueen, D. J. (2017). Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. *Genome Biol*, *18*(1), 111. https://doi.org/10.1186/s13059-017-1241-z

Rodriguez, D., Lavie, C. J., Elagizi, A., & Milani, R. V. (2022). Update on Omega-3 Polyunsaturated Fatty Acids on Cardiovascular Health. *Nutrients*, *14*(23). https://doi.org/10.3390/nu14235146

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Holper, S., Kruger, M., & Stainier, D. Y. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*, *524*(7564), 230-233. https://doi.org/10.1038/nature14580

Rouf, M. A., Wen, L., Mahendra, Y., Wang, J., Zhang, K., Liang, S., Wang, Y., Li, Z., Wang, Y., & Wang, G. (2022). The recent advances and future perspectives of genetic compensation studies in the zebrafish model. *Genes & Diseases*. https://doi.org/10.1016/j.gendis.2021.12.003

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol*, *94*(3), 441-448. https://doi.org/10.1016/0022-2836(75)90213-2

Shi, M., Zhang, H., Wang, L., Zhu, C., Sheng, K., Du, Y., Wang, K., Dias, A., Chen, S., Whitman, M., Wang, E., Reed, R., & Cheng, H. (2015). Premature Termination Codons Are Recognized in the Nucleus in A Reading-Frame Dependent Manner. *Cell Discov*, *1*, 15001-. https://doi.org/10.1038/celldisc.2015.1

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat Rev Genet*, *20*(11), 631-656. https://doi.org/10.1038/s41576-019-0150-2

Sun, X., Zhan, M., Sun, X., Liu, W., & Meng, X. (2021). C1GALT1 in health and disease. *Oncol Lett*, *22*(2), 589. https://doi.org/10.3892/ol.2021.12850

Synthego. (2019). *ICE Analysis: Get NGS quality results with Sanger data in seconds*. In https://ice.synthego.com/#/

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

Toriseva, M. J., Ala-aho, R., Karvinen, J., Baker, A. H., Marjomäki, V. S., Heino, J., & Kähäri, V. M. (2007). Collagenase-3 (MMP-13) enhances remodeling of three-dimensional collagen and promotes survival of human skin fibroblasts. *J Invest Dermatol*, *127*(1), 49-59. https://doi.org/10.1038/sj.jid.5700500

Wang, B., & Tontonoz, P. (2018). Liver X receptors in lipid signalling and membrane homeostasis. *Nat Rev Endocrinol*, *14*(8), 452-463. https://doi.org/10.1038/s41574-018-0037-x

Watanabe, Y., & Tatsuno, I. (2020). Prevention of Cardiovascular Events with Omega-3 Polyunsaturated Fatty Acids and the Mechanism Involved. *J Atheroscler Thromb*, *27*(3), 183-198. https://doi.org/10.5551/jat.50658

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189-1191. https://doi.org/10.1093/bioinformatics/btp033

Westermann, L., Neubauer, B., & Kottgen, M. (2021). Nobel Prize 2020 in Chemistry honors CRISPR: a tool for rewriting the code of life. *Pflugers Arch*, *473*(1), 1-2. https://doi.org/10.1007/s00424-020-02497-9

Wojcicka, G., Jamroz-Wisniewska, A., Horoszewicz, K., & Beltowski, J. (2007). Liver X receptors (LXRs). Part I: structure, function, regulation of activity, and role in lipid metabolism. *Postepy Hig Med Dosw (Online)*, *61*, 736-759. https://www.ncbi.nlm.nih.gov/pubmed/18063918

Yang, H., Ren, S., Yu, S., Pan, H., Li, T., Ge, S., Zhang, J., & Xia, N. (2020). Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks. *Int J Mol Sci*, *21*(18). https://doi.org/10.3390/ijms21186461

Yue, L., Ye, F., Gui, C., Luo, H., Cai, J., Shen, J., Chen, K., Shen, X., & Jiang, H. (2005). Ligand-binding regulation of LXR/RXR and LXR/PPAR heterodimerizations: SPR technology-based kinetic analysis correlated with molecular dynamics simulation. *Protein Sci*, *14*(3), 812-822. https://doi.org/10.1110/ps.04951405

Zhao, H., & Wolt, J. D. (2017). Risk associated with off-target plant genome editing and methods for its limitation. *Emerg Top Life Sci*, *1*(2), 231-240. https://doi.org/10.1042/ETLS20170037

Appendix A: Bioinformatic web tools

Table 4: Software and online resources used in the thesis. The table includes name of the software/online tool, supplier, and application.

SOFTWARE/ONLINE TOOL	SUPPLIER	APPLICATION
Benchling	Benchling	Visualization of gene sequences
BioRender	BioRender	Online illustration tool
CHOPCHOP	(Labun et al., 2019)	Design gRNA sequences
ClustalW	ClustalW (Larkin et al., 2007)	Generate multiple sequence alignment file of *lxr* and *lxr-like*
fastp	FASTQ (S. Chen et al., 2018)	Trimmed the low-quality reads and adapters from RNA-seq
g:Profiler	(Raudvere et al., 2019)	Gene ontology analysis
Jalview	(Waterhouse et al., 2009)	Visualization of gene alignment of *lxr* and *lxr-like*
KEGG	Kyoto encyclopedia of genes and genomes (Kanehisa & Goto, 2000)	Investigating genes interactions and pathways
Primer3web	Primer3	Design primer pairs for use in PCR and qPCR
R	R Desktop	Interpretation of RNA-seq data analysis to make figures
Salmon	Salmon software (Patro et al., 2017)	Quantify transcripts from RNA-seq
Synthego ICE	Synthego ICE	Interpretation of gene editing by Sanger sequencing
UniProt	UniProt Consortium	Visualize the structure of *lxr*

Appendix B: gRNA sequences

Table 5: gRNA designed to match various regions in the *lxr* gene and *lxr-like* gene to obtain knock-out and knock-down of gene expression. The gRNA sequences are designed in the CRISPR gRNA design tool CHOPCHOP and ordered from IDT.

GENE	CRISPR TARGET	gRNA (INCLUDED PAM)
lxr	After ATG	GCTGCGCAAGTGCCGTGAAGCGG (+)
lxr	After ATG	CGTTGTTCTTGCACGAATACTGG (-)
lxr-like	5'-UTR	AGAGTACTCTTTGAATCCCTTGG (+)
lxr-like	5'-UTR	CGTTTGAACAAAGCTGATGTAGG (+)
lxr-like	After ATG	TTCGCAGTTGCACGAATGTAGGG (+)
lxr-like	After ATG	GCTGCAGACCTACCAAAAATTGG (-)

Appendix C: Primer pairs for PCR and qPCR

Table 6: List of primers used in PCR for CRISPR knock-out and knock-down of *lxr* and *lxr-like* in the thesis. The primers were designed in Primer3 and ordered by ThermoFisher.

GENE	SEQUENCE 5' → 3'	ORIENTATION	ID
lxr	CTGTGTGTTCAGTGACGTCG	FWD	lxr_F1
lxr	CCAGGAACACAATCAGCACA	REV	lxr_R1
lxr	GACTGGTCCAAGCATCAGG	FWD	lxr_F2
lxr	GTGCGTGCAGTCTCTTCTTC	REV	lxr_R2
lxr	CGAGGTGTGCAGCGTGTG	FWD	lxr_F3
lxr-like	TACTCAAGGGCACATCGACA	FWD	lxr_like_TSS_F1
lxr-like	AAGTGTCTCAGCGCAATTCT	REV	lxr_like_TSS_R1
lxr-like	ACTTGTGCCACTCAGACGAA	FWD	lxr_like_TSS_F2
lxr-like	AAGGGCACGTCGTAATAACTT	REV	lxr_like_TSS_R2
lxr-like	TATGTGTTCCTCCTGGCGTT	FWD	lxr_like_exon_F1
lxr-like	CCACTATCTCCTGGACCGAC	REV	lxr_like_exon_R1
lxr-like	TAGAGATGGCTTGGCAAGGG	FWD	lxr_like_exon_F2
lxr-like	TCTCCTGGACCGACATGATG	REV	lxr_like_exon_R2

Table 7: List of primers used in qPCR studying mRNA levels of *lxr* and *lxr-like* in the thesis. The primer pairs were targeting either the coding sequence (CDS) or both the exogenic and intergenic regions. The primers were ordered by ThermoFisher.

GENE	SEQUENCE 5' → 3'	ORIENTATION	ID
lxr	TGGAGCCCAGTGACATCAAG	FWD	CDS1-F
lxr	CCCTTCTTCCTCTTCACCGG	REV	CDS1-R
lxr	ATGAACGACCTGCACCTGG	FWD	CDS2-F
lxr	AGTCGCTCCACCAGTTCATG	REV	CDS2-R
lxr	TTGAGGGCTGAGGGTTTGTC	FWD	LXR1-F
lxr	AGTGTTGACAAGCAGAAGGGT	REV	LXR1-R
lxr	CAGTGACATCAAGGCCGACC	FWD	LXR2-F
lxr	CAAGATGCTGGGGAACGAGG	REV	LXR2-R
lxr	GCTCCACTCTCAGATCCACA	FWD	LXR3-F
lxr	CCTGTGTTGCAATGGGATAGG	REV	LXR3-R
lxr-like	GTCATAACTTCCACGAGACAGA	FWD	LXR-like1-F
lxr-like	CCAGGGTAGCAGCCACGT	REV	LXR-like1-R
lxr-like	GAAATTAGTCCGAGGGGCCC	FWD	LXR-like2-F
lxr-like	CAGCATGGTTCCAGACATCT	REV	LXR-like2-R
lxr-like	AGGGTTGGTGGTCTTCTATTTGT	FWD	LXR-like3-F
lxr-like	TGTCTTTGATGGCGCTGACT	REV	LXR-like3-R

Table 8: List of primers used in qPCR studying mRNA levels of the reference genes *18s, ef1a* and *rpl1* in the thesis. The primers are designed and verified by Jorgensen et al. (2006).

GENE	SEQUENCE 5' → 3'	ORIENTATION
18s	TGTGCCGCTAGAGGTGAAATT	FWD
18s	GCAAATGCTTTCGCTTTCG	REV
ef1a	CACCACCGGCCATCTGATCTACAA	FWD
ef1a	TCAGCAGCCTCCTTCTCGAACTTC	REV
rpl1	ACTATGGCTGTCGAGAAGGTGCT	FWD
rpl1	TGTACTCGAACAGTCGTGGGTCA	REV

Appendix D: Location of gRNAs and PCR primers in the *lxr* gene

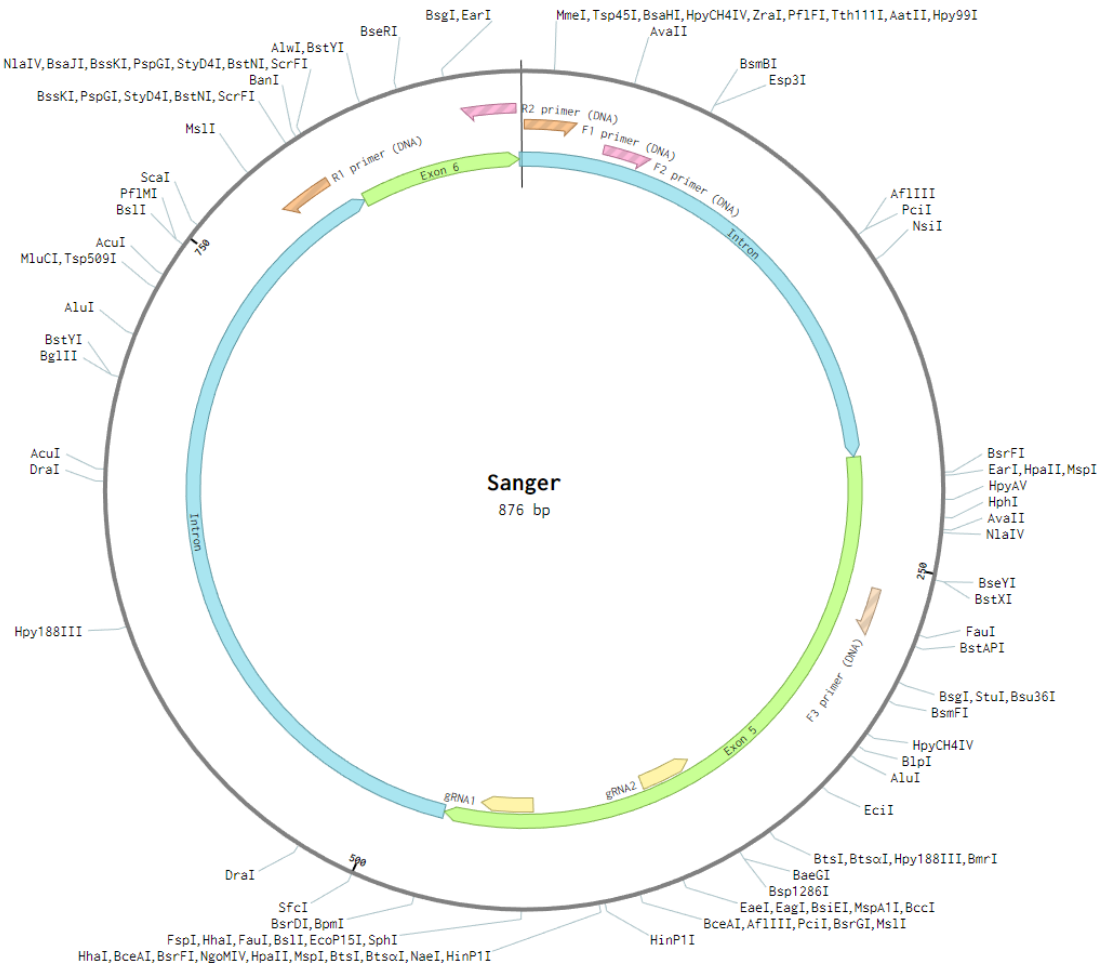


Figure 20: Location of primer pairs and gRNAs in *lxr* gene. Primer pairs and gRNA were designed to the regions within exon 5 and exon 6 (green), with introns (blue) in between. The gRNAs (yellow), *lxr* primer pair 1 (orange), *lxr* primer pair 2 (pink) and the F3 primer for Sanger sequencing (beige) are presented in the gene map. Primer sequences are attached in Table 6 (Appendix C) and gRNA sequences in Table 5 (Appendix B). The Figure is created in Benchling.

Appendix E: qPCR

Table 9: C_q-values from qPCR measurements after *lxr* gene editing. Knock-down, knock-out and control samples during 40 cycles of qPCR. EF1A primer pair was used as reference gene.

SAMPLE	PRIMER PAIR	TREATMENT (C_q)	CONTROL (C_q)	$\Delta\Delta C_q$
lxr knock-out	EF1A	21,53	20,36	1
	CDS1	28,52	26,81	0,69
	CDS2	28,23	26,48	0,67
lxr knock-down	EF1A	18,38	18,15	1
	CDS1	25,08	24,45	0,76
	CDS2	25,12	24,68	0,86

Table 10: Example of calculation of ΔC_q and $\Delta\Delta C_q$ using *lxr* knock-out sample with CDS1 primer pair. The calculation for ΔC_q value is (C_q(treatment) - C_q(control)) and ΔC_q value is $2^{(\Delta Cq(treatment) - \Delta Cq(control))}$. A detailed calculation is attached below the table.

SAMPLE	REFERENCE GENE	GENE OF INTEREST	ΔC_q	$\Delta\Delta C_q$
Control	20,36	26,81	6,45	1,00
Treated	21,53	28,52	6,99	0,69

ΔC_q equation: $C_q^{(treatment)} - C_q^{(control)}$

$\Delta C_q^{(control)}$: 26,81 – 20,36 = 6,45

$\Delta C_q^{(treated)}$: 28,52 – 21,53 = 6,99

$\Delta\Delta C_q$ equation: $2^{(\Delta Cq(treatment) - \Delta Cq(control))}$

$\Delta\Delta C_q^{(control)}$: $2^{(6,45-6,45)}$ = 1,00

$\Delta\Delta C_q^{(treated)}$: $2^{(6,99-6,45)}$ = 0,69

Appendix F: RNA-seq

Table 11: RNA concentrations and volumes of the *lxr* knock-out, *lxr* knock-down and control samples sent for RNA-seq. All samples sent for RNA-seq had a total amount of RNA > 1000 ng.

SAMPLE NAME	CONCENTRATION (ng/µl)	VOLUME (µl)	TOTAL AMOUNT (ng)
Knock-out #1	107	18	1926
Knock-out #2	157	20	3140
Knock-out #3	114	23	2622
Knock-out #4	131	23	3013
Knock-down #1	156	21	3276
Knock-down #2	58	18	1044
Knock-down #3	111	19	2109
Knock-down #4	67	23	1541
Control #1	137	23	3151
Control #2	143	23	3289
Control #3	147	19	2793
Control #4	161	20	3220

Table 12: Upregulated genes as a result of *lxr* knock-out. In total, 25 genes were upregulated in SHK-1 cells due to knock-out of *lxr*. The table presents the log2FoldChange for all genes, along with standard error (lfcSE), adjusted p-value, and gene ID, including the full name of the gene abbreviation. The gene nomenclature employed in this thesis is marked with an asterisk (*) to denote its non-public status as a variant of a published gene. As such, the suffix "L" (for "like") has been appended to indicate its non-standardized nature. It should be emphasized that these designations are intended solely for internal usage within this thesis and are not intended to function as public gene names.

genename	log2FoldChange	lfcSE	padj	geneID	complete gene name
lamc2	2,2233742	0,5733866	0,07623333	106599435	Laminin subunit gamma-2
dnk	1,9368746	0,5065722	0,08549006	106603918	Deoxynucleoside kinase
*c1galt1l**	1,9324951	0,4949528	0,07623333	123740261	Glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1-like
NA	1,8146867	0,4460141	0,05825056	123738345	*NA*
csf2rb	1,7711297	0,4117477	0,03582472	106600897	Colony stimulating factor 2 receptor subunit beta
palb2	1,6960235	0,412921	0,05337023	106593073	Partner and localizer of BRCA2
f8	1,6417465	0,3705429	0,02645176	106612066	Coagulation factor VIII
*tns4l**	1,6254608	0,3517055	0,01929361	106587857	Tensin-4-like
*melkl**	1,5215543	0,389655	0,07623333	106594008	Maternal embryonic leucine zipper kinase-like
top3a	1,3439277	0,3476604	0,07623333	106594146	DNA topoisomerase 3-alpha-like
pitpnc1	1,3044536	0,3105307	0,047965	106564871	Cytoplasmic phosphatidylinositol transfer protein 1
ptpn12	1,1736478	0,2689046	0,02995677	100380432	Tyrosine-protein phosphatase non-receptor type 12
mme	1,1345952	0,2731664	0,047965	106581480	Neprilysin
cdh2	0,8726624	0,215796	0,05825056	106560349	Cadherin 2, type 1, N-cadherin (neuronal)
vsig4	0,8421666	0,2236082	0,0993135	106608944	Immunoglobulin domain-containing protein
usp7	0,7643759	0,1716688	0,02645176	106608838	Ubiquitin carboxyl-terminal hydrolase 7
mmp13	0,7154201	0,1726236	0,047965	100195495	Collagenase 3
serpine2	0,6714074	0,157595	0,03978813	100196662	Serine (or cysteine) proteinase inhibitor clade E, nexin, plasminogen activator inhibitor type 1-like
chchd10	0,587485	0,1477843	0,06361669	106570486	Coiled-coil-helix-coiled-coil-helix domain containing 10
chaf1a	0,5805447	0,1500439	0,07623333	106613808	Chromatin assembly factor 1 subunit A
pcdh12	0,5431956	0,1405599	0,07623333	106604573	Protocadherin-12
*vaspl**	0,5411668	0,1440692	0,0993135	106603798	Vasodilator-stimulated phosphoprotein-like
*gjb3l**	0,5238114	0,129335	0,05825056	106570709	Gap junction beta-3 protein-like
piezo1	0,4752707	0,1246527	0,08705788	106562059	Piezo-type mechanosensitive ion channel component 1
*nxph1l**	0,4650873	0,1195826	0,07623333	106606905	Neurexophilin-1-like

Table 13: Downregulated genes as a result of *lxr* knock-out. In total, 19 genes were downregulated in SHK-1 cells due to knock-out of *lxr*. The table presents the log2FoldChange for all genes, along with standard error (lfcSE), adjusted p-value, and gene ID, including the full name of the gene abbreviation. The gene nomenclature employed in this thesis is marked with an asterisk (*) to denote its non-public status as a variant of a published gene. As such, the suffix "L" (for "like") has been appended to indicate its non-standardized nature. It should be emphasized that these designations are intended solely for internal usage within this particular thesis and are not intended to function as public gene names.

gene name	log2FoldChange	lfcSE	padj	geneID	complete gene name
acta2	-2,3993034	0,4618374	0,005184	106578961	Actin alpha 2
scg3	-2,0484873	0,52689	0,076233	106562528	Secretogranin-3
chadl	-1,8080242	0,4355932	0,047965	106589269	Chondroadherin-like
c3	-1,4418884	0,3600389	0,058251	106605127	Complement C3
col17a1b	-1,4217592	0,2981092	0,014759	106577015	Collagen type XVII alpha 1 chain
slc30a2	-1,2408517	0,3092306	0,058251	106570053	Zinc transporter 2
c3	-1,1882454	0,2849633	0,047965	106597357	Complement C3
*crim1l**	-1,1467131	0,2428195	0,014759	106613828	Cysteine-rich motor neuron 1 protein-like
metrnl	-0,9788675	0,2554271	0,084664	106606275	Meteorin-like protein
lxr	-0,9701698	0,1958753	0,009259	100270809	Liver x receptor
egr1	-0,9521489	0,2361233	0,058251	106611930	Early growth response protein 1
rgs21	-0,8949319	0,1974376	0,021078	106560359	Regulator of G-protein signaling 21
cd163	-0,8080292	0,2012865	0,058251	106570622	Scavenger receptor cysteine-rich type 1 protein M130-like
*hgfl**	-0,7348801	0,1954363	0,099313	106576483	Hepatocyte growth factor-like
na	-0,6889342	0,1580111	0,029957	123744600	LOC123744600
*col6a1l**	-0,6184452	0,1532798	0,058251	106561261	Collagen alpha-1(VI) chain-like
s100a1	-0,4939337	0,1255859	0,073294	100136538	S100 Calcium Binding Protein A1
*emilin-1-l**	-0,4870979	0,1284182	0,091967	106611378	Elastin Microfibril Interfacer 1
cd99	-0,3976399	0,0871511	0,021078	106575202	Cell surface glycoprotein

Table 14: Differential gene expression of essential genes involved in lipid metabolism. Information about genes involved in lipid metabolism in Atlantic salmon after a *lxr* knock-out. The table presents the log2FoldChange for all genes, along with standard error (lfcSE), p-value, adjusted p-value, and gene ID, including the full name of the gene abbreviation.

gene name	log2FoldChange	lfcSE	pvalue	padj	geneID	full gene name
lxr	-0.9716793	0.1955447	6.72×10^{-7}	0.000481938	100270809	Liver X receptor
lxr-like	0.07661415	0.1491964	0.6075931	1	106561932	Liver X receptor-like
rxr	0.02991317	0.167474	0.8582409	1	106588394	Retinoid X receptor
srebpf1	0.08674523	0.2414657	0.7194122	1	100502556	SREBP regulating gene protein factor 1
srebpf2	0.1678681	0.142505	0.2388046	1	100502557	SREBP regulating gene protein factor 2
srebpl	-0.003784591	0.3571243	0.9915447	1	106585235	SREBP regulating gene protein-like
elovl5	-0.09257731	0.282789	0.7433862	1	100192340	ELOVL fatty acid elongase 5
elovl5a	0.03473227	0.1676945	0.8359191	1	100136433	Polyunsaturated fatty acid elongase
elovl8a	-1.427588	3.812965	0.708104	*NA*	106569295	ELOVL fatty acid elongase 8a
elovl6	-0.1908628	0.5730644	*NA*	*NA*	106604515	ELOVL fatty acid elongase 6
elovl1b	0.09296032	0.1275131	0.4659861	1	100286454	ELOVL fatty acid elongase 1b
elovl6l	0.1954284	0.2641218	0.4593499	1	106606039	ELOVL family member 6, elongation of long chain fatty acids like
elovl2	0.05009241	0.3756393	0.8939147	1	100192341	ELOVL fatty acid elongase 2
hmdh	0.07635904	0.118969	0.5209773	1	100380797	3-hydroxy-3-methylglutaryl-coenzyme A reductase
acox3	0.1386952	0.122736	0.2584641	1	100306814	Acyl-CoA oxidase 3, pristanoyl
facr1	0.1700378	0.1257575	0.1763406	1	100195056	Fatty acyl-CoA reductase 1
lpl	0.6266273	0.3492876	0.07281077	1	106573544	Lipoprotein lipase
acsl4	0.09793911	0.2232217	0.6608408	1	100380405	Long-chain-fatty-acid--CoA ligase 4
fadsd6	0.1130984	0.2763418	0.6823415	1	100136441	Delta-6 fatty acyl desaturase
fadsd5	-0.05247948	0.16415	0.7491924	1	100136383	Delta-5 fatty acyl desaturase
pparaa	*NA*	*NA*	*NA*	*NA*	106575923	peroxisome proliferator-activated receptor alpha a
pparg	-0.2073708	0.7161165	0.7721399	1	106583865	Peroxisome proliferator-activated receptor gamma [product]
pparb2b	0.2368757	0.4495469	0.5982484	1	100136536	Peroxisome proliferator-activated receptor beta2B
apof	*NA*	*NA*	*NA*	*NA*	100196604	Apolipoprotein F
apo	*NA*	*NA*	*NA*	*NA*	106586541	Apolipoprotein A-I
apol	0.3730856	0.7327647	0.6106485	1	106577499	Apolipoprotein A-I-like
apoeb	-0.2710783	0.5080676	0.5936545	1	106577506	Apolipoprotein Eb

Table 15: TPM measurements of important lipid metabolism genes in SHK-1 cells. The genes have four replicates of controls and four replicates of *lxr* knock-out samples. The table shows gene ID, gene name and TPM measurement values for the eight samples.

GeneID	GeneName	control #1	control #2	control #3	control #4	knock-out #1	knock-out #2	knock-out #3	knock-out #4
100270809	*lxr*	6,92	7,37	8,44	10,87	3,26	4,06	5,35	5,04
106561932	*lxr-like*	4,42	3,58	4,61	4,84	5,63	5,34	4,29	4,13
106566994	*ppardb*	0,81	0,46	1,12	0,98	1,06	0,77	1,20	0,87
106575923	*pparaa*	-	-	-	-	-	-	-	-
106583865	*pparg*	1,29	1,52	2,34	7,12	4,61	5,96	2,44	1,30
106608616	*ppargc1a*	-	0,03	0,04	0,02	0,02	-	0,09	0,05
100136536	*pparb2b*	0,33	1,01	0,44	0,25	0,42	0,41	0,30	0,47
106566994	*ppardb*	0,81	0,46	1,12	0,98	1,06	0,77	1,20	0,87
100502556	*srebf1*	44,56	47,64	53,20	75,40	67,66	64,48	58,18	54,26
100502557	*srebf2*	31,27	30,38	34,22	49,60	46,48	48,83	39,01	35,95
106604515	*elovl6*	151,09	153,80	150,93	199,18	165,35	176,92	128,25	131,09
106606039	*elovl6l*	1,83	1,94	1,44	2,46	2,05	1,76	2,68	2,58
100192340	*elovl5*	4,93	3,89	5,58	4,77	3,87	4,19	6,39	4,27
100192341	*elovl2*	17,75	17,10	17,46	28,51	23,55	30,23	17,56	16,02
100286454	*elovl1b*	14,23	14,43	15,82	19,04	17,79	19,06	18,77	15,80
100500788	*elovl4*	0,04	0,06	-	0,09	0,09	-	0,08	0,03
106569295	*elovl8a*	0,04	0,03	-	-	-	-	-	-

Appendix G: R Markdown file of R analyses

The R code used for the bioinformatic data analysis is provided in the R Markdown file.

Masters thesis

Marie Hellan Iversen

2023-05-15

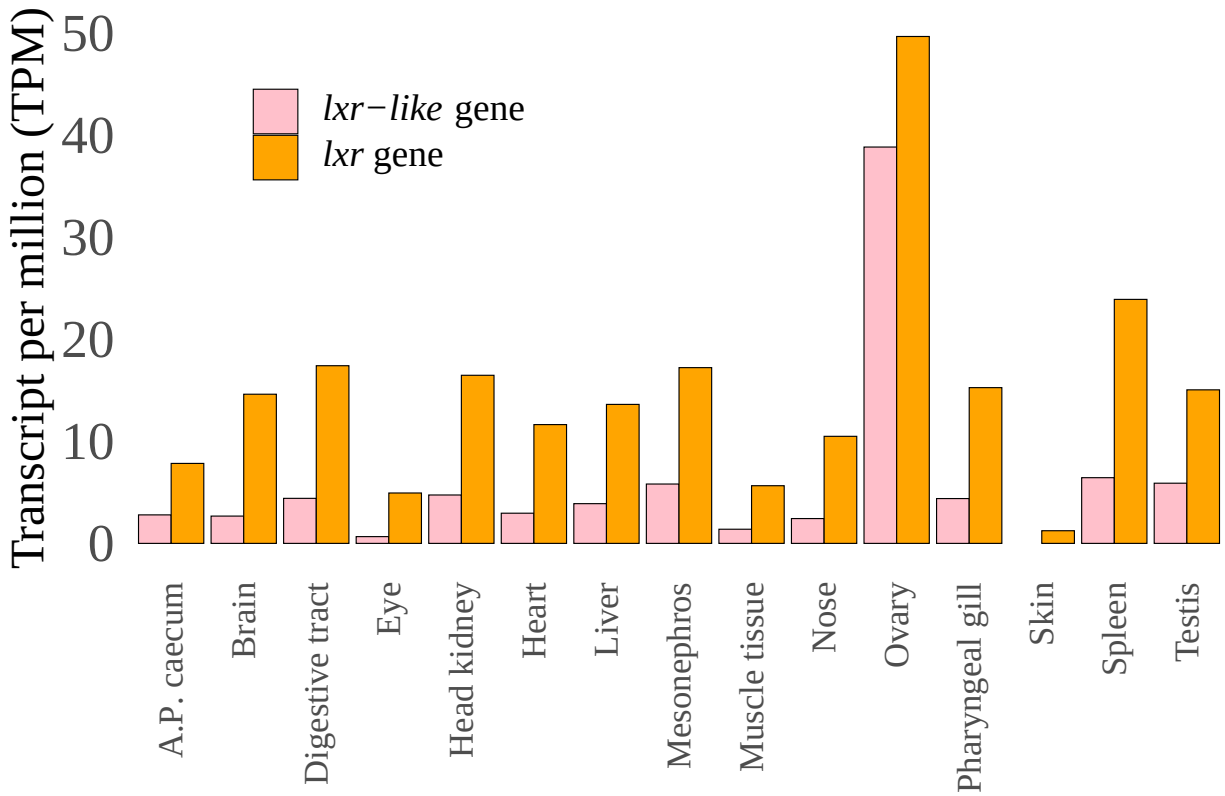
Bioinformatic analyses in R

Tissue distribution of gene expression of *lxr* and *lxr-like*

Barplot illustrating the different regulation of the *lxr* and *lxr-like* genes in Atlantic salmon.

```
if (!require("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("BgeeDB")
library(BgeeDB)

#Generate tpm table from diverse tissues
bgee <- Bgee$new(species = "Salmo_salar", dataType = "rna_seq")
data_bgee <- getData(bgee, experimentId = "SRP011583")
data_bgee <- data_bgee[! data_bgee$Anatomical.entity.name %in%
                         "\"sexually immature organism\"",]
gene.expression.tpm <- formatData(bgee, data_bgee, callType = "present", stats = "tpm")
tpm.tissues <- gene.expression.tpm@assayData$exprs
colnames(tpm.tissues) <- gsub(" ", "_", gsub("\"", "",
                          data_bgee$Anatomical.entity.name[match(colnames(tpm.tissues),
                          data_bgee$Library.ID)]))

#Create a data frame for ggplot
tpm_df <- data.frame(Tissue = c("Pharyngeal gill","Digestive tract","Head kidney",
                                "Heart","Mesonephros", "Liver","Muscle tissue",
                                "Nose","A.P. caecum","Skin","Spleen","Brain",
                                "Eye","Ovary","Testis"),
                     Gene1 = tpm.tissues["ENSSSAG00000065747", ],
                     Gene2 = tpm.tissues["ENSSSAG00000074043", ])

#Melt the data frame
library(reshape2)
tpm_df <- melt(tpm_df, id.vars = "Tissue", variable.name = "Gene")

#Create the ggplot
library(ggplot2)
ggplot(tpm_df, aes(x = Tissue, y = value, fill = Gene)) +
  geom_bar(stat = "identity", position = "dodge", color = "black", size = 0.2) +
  scale_fill_manual(values = c("pink", "orange"), name = "",
                    labels = c(bquote(paste(italic("lxr-like"), " gene")),
                               italic("lxr")~"gene")) +
```

```
  labs(x = "", y = "Transcript per million (TPM)") +
  scale_y_continuous(limits = c(0, 50)) +
  theme_minimal() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        legend.justification = c(0, 1), legend.position = c(0.1, 1),
        text = element_text(size = 18, family="serif"),
        axis.text = element_text(size = 20),
        axis.text.x = element_text(size = 13, angle = 90, vjust = 0.5, hjust=1))
```

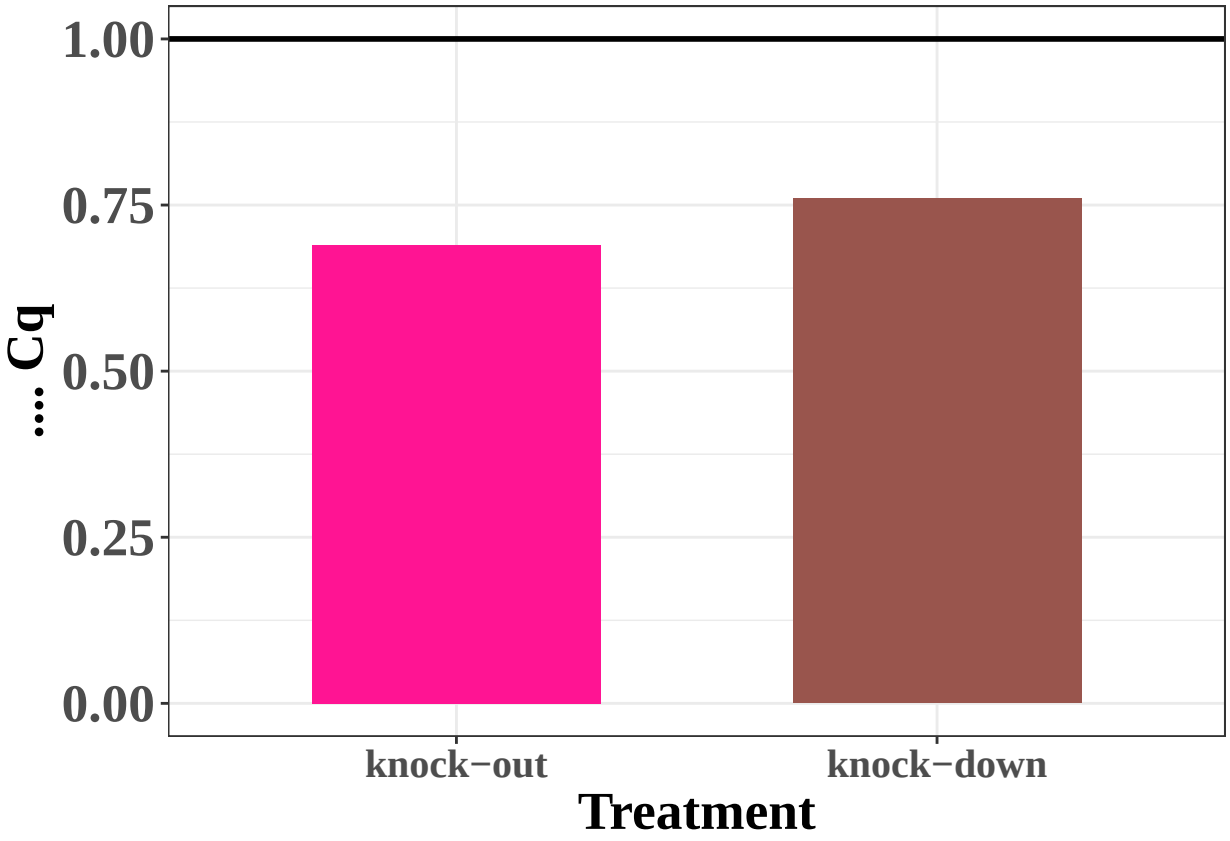


Delta-delta-Cq of *lxr* knock-out and *lxr* knock-down experiments

Barplot showing delta-delta-Cq values indicating efficient repression of mRNA levels of the *lxr* gene in both knock-out and knock-down samples.

```
#Make data frame with ΔΔCq of lxr knock-out and knock-down experiments
df.qPCR <- data.frame(Genename=c("knock-out", "knock-down"),
                      Cq=c(0.69,0.76),
                      Treatment=c("knock-out", "knock-down"))

df.qPCR$Genename <- factor(df.qPCR$Genename, levels = c("knock-out", "knock-down"))

library(ggplot2)
ggplot(data=df.qPCR, aes(x=Genename, y=Cq)) +
  geom_bar(stat="identity", position = position_dodge(0.9), width=0.6,
```

```
          fill = c("#FF1493","#99554D")) +
  theme_bw() +
  theme(text = element_text(size = 20, family="serif", face = "bold"),
        axis.text = element_text(size = 25, face = "bold", family="serif"),
        axis.text.x = element_text(size = 15, family="serif", face = "bold"),
        axis.text.y = element_text(size = 20)) +
  labs(x="Treatment", y = "ΔΔ Cq", color = "Treatment", size = 25) +
  geom_hline(yintercept = 1, linetype = "solid", color = "black", size = 1)
```



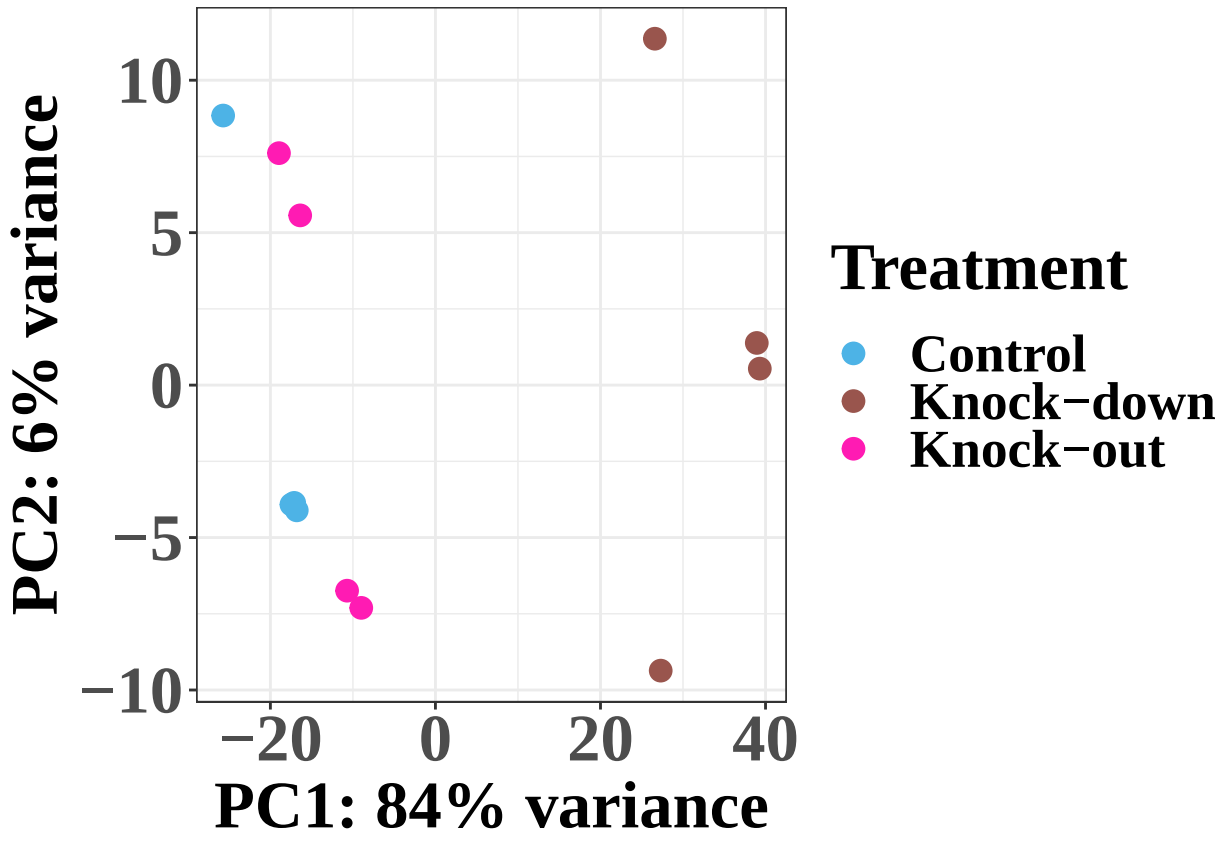
The impact of *lxr* knock-out in SHK-1 cells

Two-dimensional PCA plots indicating patterns in variance between *lxr* knock-out, *lxr* knock-down and control samples.

```
#Preparing counts table from RNA-seq data analysis
library(readr)
library(DESeq2)
library(dplyr)

data <- read_tsv("C:/Users/marie/Dropbox/PC/Downloads/counts.tsv")
genes <- paste0(data$gene_id, "_", 1:nrow(data))
data <- as.data.frame(data)
rownames(data) <- genes
counts <- data[, -c(1,2)]
```

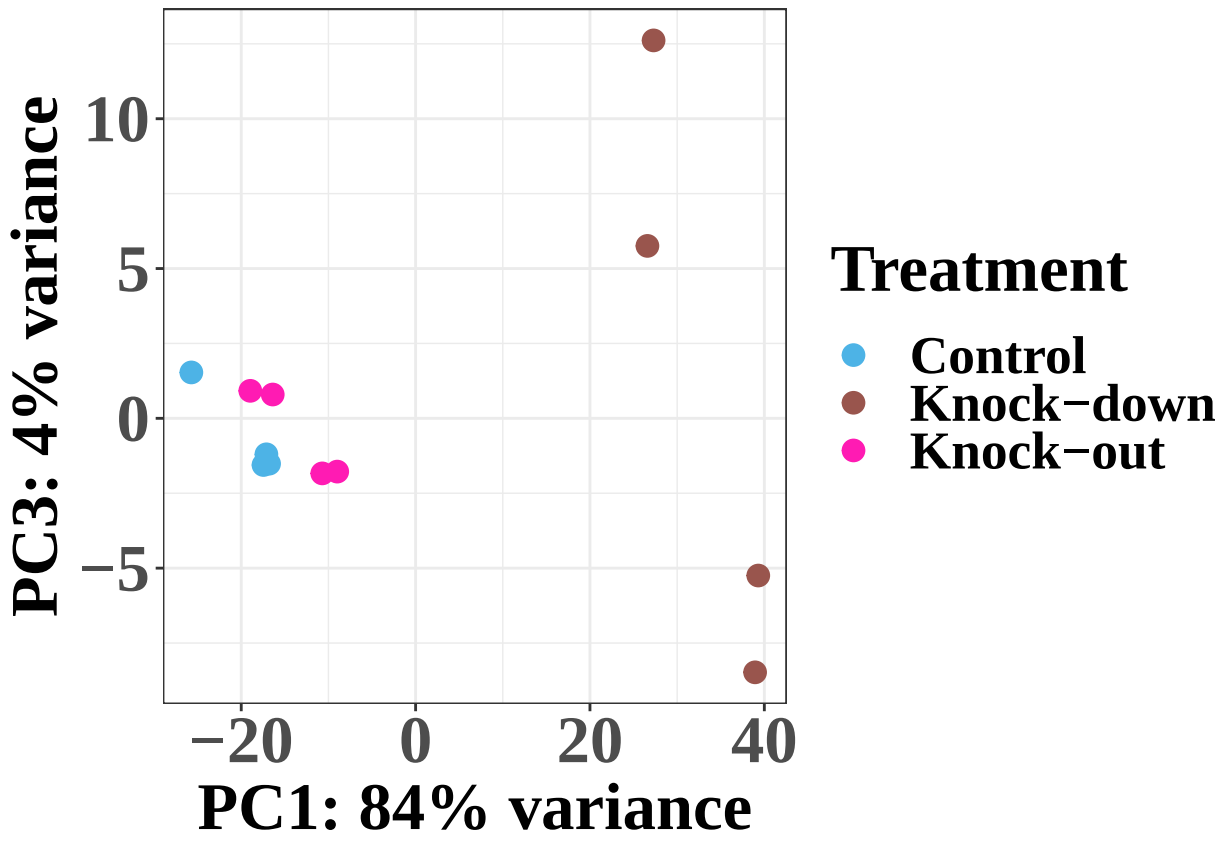
```
#Preparing coldata
coldata = data.frame(condition = gsub('[0-9]', '', colnames(counts)),
                     row.names = colnames(counts))

dds <- DESeqDataSetFromMatrix(countData = round(counts),
                              colData = coldata,
                              design = ~condition)
dds <- DESeq(dds)

#Extracting transformed values
vsd <- vst(dds, blind=FALSE)
rld <- rlog(dds, blind=FALSE)

#Function for plotPCAmodify
plotPCAmodify <- function (object, intgroup = "condition", ntop = 500,
                           returnData = FALSE, nPC=2)
{
  rv <- rowVars(assay(object))
  select <- order(rv, decreasing = TRUE)[seq_len(min(ntop,
                                                     length(rv)))]
  pca <- prcomp(t(assay(object)[select, ]))
  percentVar <- pca$sdev^2/sum(pca$sdev^2)
  if (!all(intgroup %in% names(colData(object)))) {
    stop("the argument 'intgroup' should specify columns of colData(dds)")
  }
  intgroup.df <- as.data.frame(colData(object)[, intgroup,
                                               drop = FALSE])
  group <- if (length(intgroup) > 1) {
    factor(apply(intgroup.df, 1, paste, collapse = ":"))
  }
  else {
    colData(object)[[intgroup]]
  }
  #d <- data.frame(PC1 = pca$x[, 1], PC2 = pca$x[, 2],
        #group = group, intgroup.df, name = colnames(object))
  d <- cbind(pca$x[,seq_len(min(nPC, ncol(pca$x)))], drop = FALSE],
             data.frame(group = group, intgroup.df, name = colnames(object)))
  if (returnData) {
    attr(d, "percentVar") <- percentVar[1:nPC] # 1:2 ==> 1:nPC
    return(d)
  }
  ggplot(data = d, aes_string(x = "PC1", y = "PC2", color = "group")) +
    geom_point(size = 3) + xlab(paste0("PC1: ", round(percentVar[1] *
                                                        100), "% variance")) +
    ylab(paste0("PC2: ", round(percentVar[2] *
}

#PC1 vs PC2
pcaData <- plotPCAmodify(vsd, intgroup=("condition"), nPC=2, returnData = T)

pc1pc2 <- pcaData %>% dplyr::select(PC1, PC2, group, condition, name)

library(ggplot2)
```

```
percentVar <- round(100 * attr(pc1pc2, "percentVar")[c(1,2)])
ggplot(pc1pc2, aes(PC1, PC2, color=condition)) +
  geom_point(size=3.5) +
  xlab(paste0("PC1: ",percentVar[1],"% variance")) +
  ylab(paste0("PC2: ",percentVar[2],"% variance")) +
  scale_color_manual(values = c("#4DB3E6","#99554D","#FF1BB3"),
                     labels = c("Control", "Knock-down", "Knock-out")) +
  theme_bw() +
  theme(text = element_text(size = 25, family="serif", face = "bold"),
        axis.text = element_text(size = 25),
        axis.text.x = element_text(size = 25), axis.text.y =
          element_text(size = 25)) +
  labs(color = "Treatment")
```




```
#PC1 vs PC3
pcaData <- plotPCAmodify(vsd, intgroup=("condition"), nPC=3, returnData = T)
pc1pc3 <- pcaData %>% select(PC1, PC3, group, condition, name)

library(ggplot2)

percentVar <- round(100 * attr(pc1pc3, "percentVar")[c(1,3)])
ggplot(pc1pc3, aes(PC1, PC3, color=condition)) +
  geom_point(size=3.5) +
  xlab(paste0("PC1: ",percentVar[1],"% variance")) +
  ylab(paste0("PC3: ",percentVar[2],"% variance")) +
```

```
  scale_color_manual(values = c("#4DB3E6","#99554D","#FF1BB3"),
                     labels = c("Control", "Knock-down", "Knock-out")) +
  theme_bw() +
  theme(text = element_text(size = 25, family="serif", face = "bold"),
        axis.text = element_text(size = 25),
        axis.text.x = element_text(size = 25), axis.text.y =
          element_text(size = 25)) +
  labs(color = "Treatment")
```

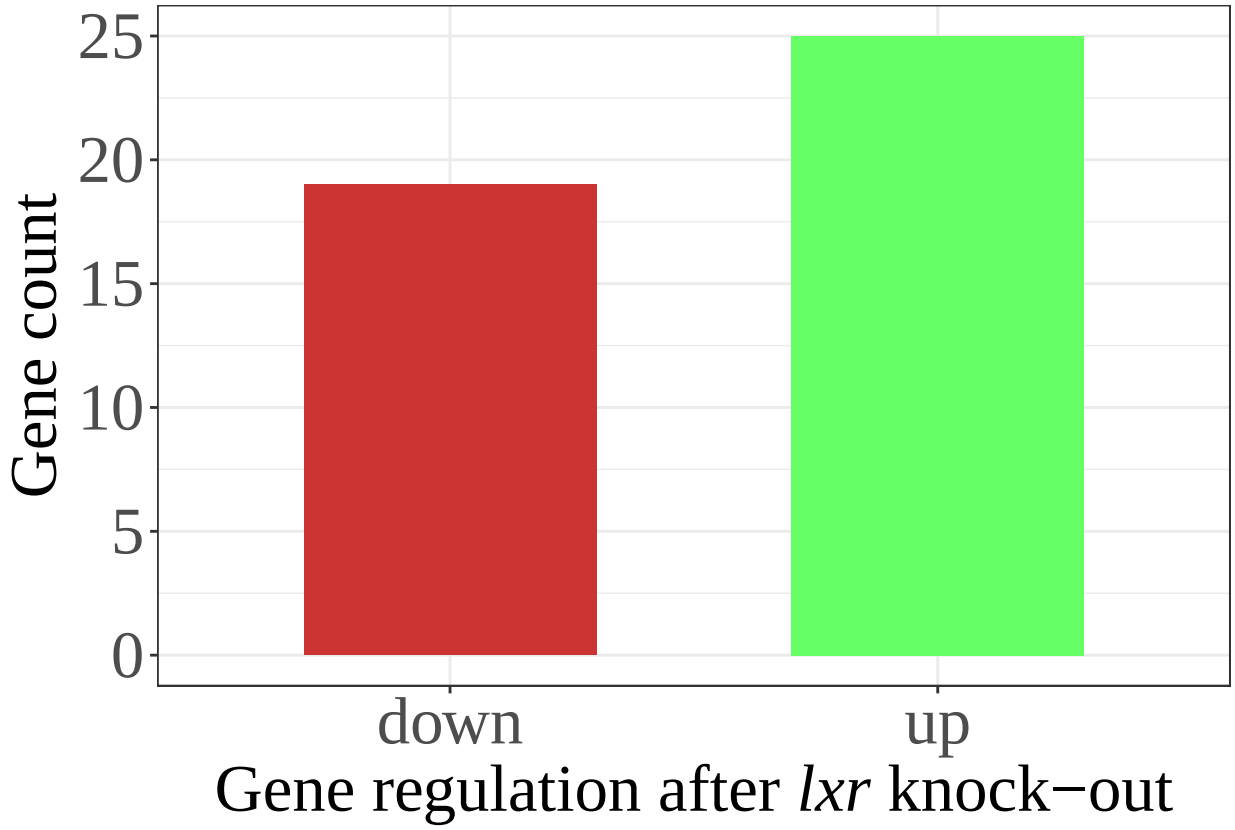


Distribution of significant differentially expressed genes after a *lxr* knock-out

Barplot showing how many genes were regulated differently after the *lxr* knock-out.

```
resKD <- results(dds, contrast=c("condition","CTRL","KD"))
resKO <- results(dds, contrast=c("condition","CTRL","KO"))

#Make barplot with regulated genes after lxr knock out
library(tibble)
resKO %>% as.tibble() %>% filter(padj<0.1) %>%
  mutate(direction=ifelse(log2FoldChange>0, "down", "up")) %>%
  ggplot(aes(x=direction)) +
  geom_bar(position = position_dodge(0.9), width=0.6, fill = c("#CC3333", "#66FF66")) +
  theme_bw() +
  theme(text = element_text(size = 25, family="serif"),
```

```
        axis.text = element_text(size = 25),
        axis.text.x = element_text(size = 25),
        axis.text.y = element_text(size = 25)) +
  labs(x=expression("Gene regulation after" ~
                      italic("lxr") ~ "knock-out"), y = "Gene count")
```

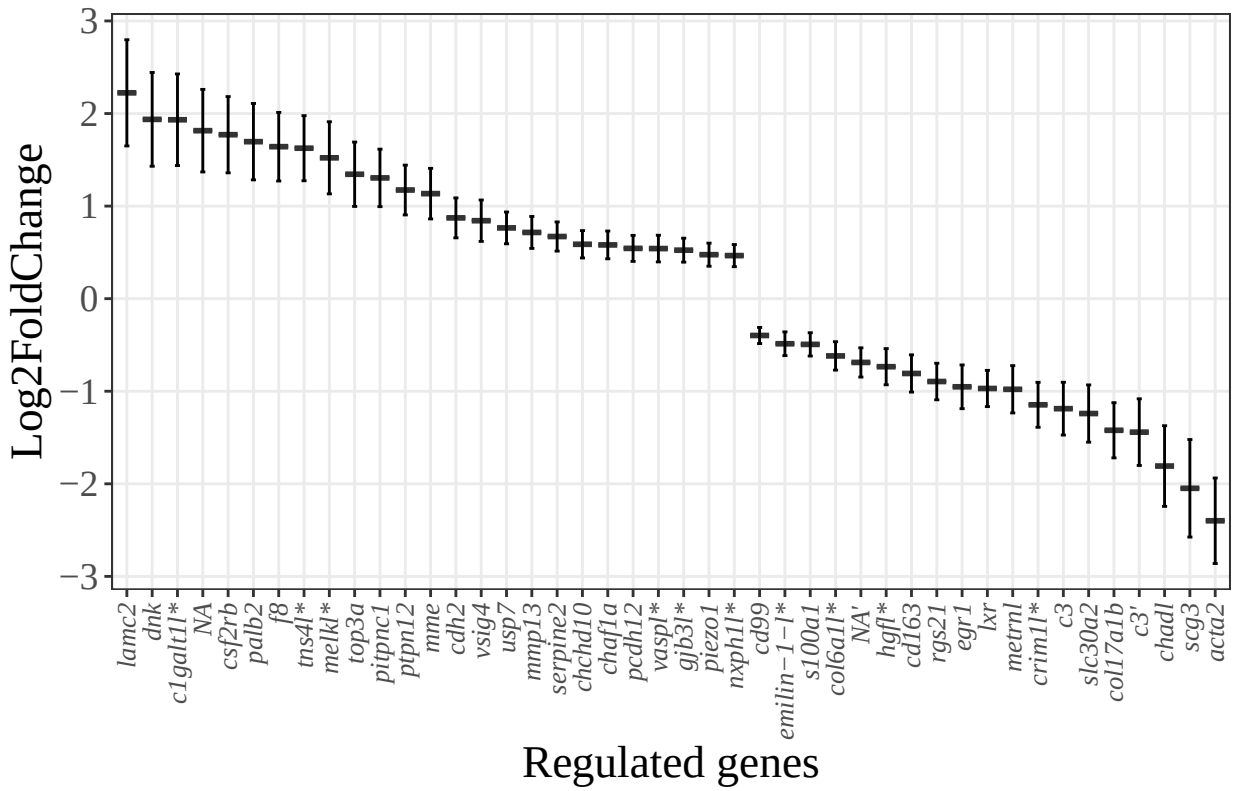


Investigation of the upregulated and downregulated genes

Boxplot showing significant differential expressed genes after a *lxr* knock-out in SHK-1 cells.

```
#Find significant genes
resKO %>% as.tibble() %>% mutate(geneID = rownames(resKO)) %>%
  filter(padj<0.1) %>%
  arrange(padj) #44 genes

resKD %>% as.tibble() %>% mutate(geneID = rownames(resKD)) %>%
  filter(padj<0.1) %>%
  arrange(padj) #13745 genes

#Create boxplot with regulted genes after lxr knock-out
library(readxl)
Diff_genes <-
  read_excel("C:/Users/marie/Dropbox/PC/Downloads/GeneID_KO_significant__genes.xlsx",
                          sheet = 1)
```

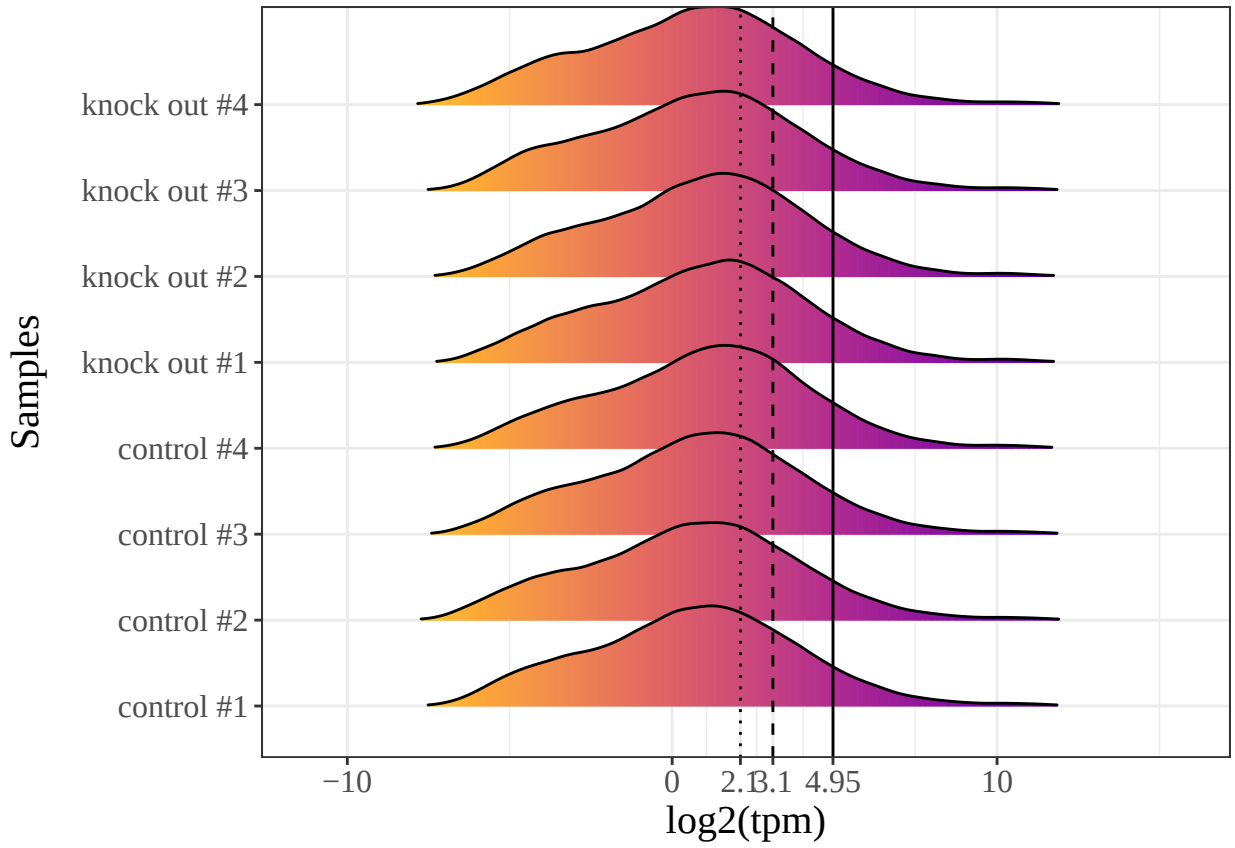
```
ggplot(Diff_genes,
       aes(x = factor(genename,
                      level=c("lamc2", "dnk", "c1galt1l*","NA","csf2rb","palb2","f8",
                              "tns4l*","melkl*","top3a","pitpnc1","ptpn12","mme",
                              "cdh2","vsig4","usp7","mmp13","serpine2",
                              "chchd10","chaf1a","pcdh12","vaspl*",
                              "gjb3l*","piezo1","nxph1l*",
                              "cd99","emilin-1-l*","s100a1","col6a1l*","NA'","hgfl*",
                              "cd163","rgs21","egr1","lxr","metrnl","crim1l*",
                              "c3","slc30a2","col17a1b","c3'","chadl","scg3","acta2")),
                    y = log2FoldChange)) + geom_boxplot(position =
                                                          position_dodge(0.9)) +
                    theme_bw() +
                    theme(panel.grid.minor = element_blank(),
                    text = element_text(size = 17, family="serif"),
                    axis.text.x = element_text(face = "italic", size = 10,
                                               angle = 90,
                                               vjust = 0.5, hjust=1)) +
   geom_errorbar(aes(ymin=log2FoldChange-lfcSE, ymax=log2FoldChange+lfcSE),
                 width=.2,
                position=position_dodge(.9)) +
   labs(title="", x="Regulated genes", y = "Log2FoldChange")
```

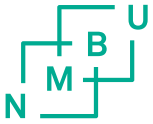


Expression of *lxr* gene compared to all other genes expressed in SHK-1 cells

Density plot presenting mean *lxr* values and mean TPM values for all genes in SHK-1 cells of Atlantic salmon.

```
library(tidyverse)
library(ggridges)
tpm_data <- read_table("C:/Users/marie/Dropbox/PC/Downloads/TPM_KO.txt")

#Normalization of data
norm_tpm_data <- tpm_data %>% select(-c(1:2)) %>%
  mutate_if(is.character, as.numeric) %>%
  cbind(tpm_data[,1:2],.) %>%  as_tibble() %>%
  pivot_longer(3:10, names_to = "samples", values_to="tpm") %>%
  mutate(tpm = ifelse(is.infinite(tpm), NA, tpm)) %>%
  drop_na() %>%
  filter(tpm>0)

#Make density plot
tpm_data %>% select(-c(1:2)) %>%
  mutate_if(is.character, as.numeric) %>%
  cbind(tpm_data[,1:2],.) %>%  as_tibble() %>%
  pivot_longer(3:10, names_to = "samples", values_to="tpm") %>%
  mutate(tpm = ifelse(is.infinite(tpm), NA, tpm)) %>%
  drop_na() %>%
  filter(tpm>0) %>%
  ggplot(aes(y=samples, x=log2(tpm))) +
  geom_density_ridges_gradient(aes(fill = ..x..), scale = 1.2,
                               rel_min_height = 0.01) +
  scale_fill_viridis_c(name = "log2(tpm)", option = "C", direction = -1,
                       guide = "none") +
  theme_bw() +
  theme(text = element_text(family = "serif", size = 15),
        axis.title.y = element_text(margin = margin(r = 13))) +
  labs(x="log2(tpm)", y = "Samples", size = 25) +
  scale_y_discrete(labels = c("control #1", "control #2", "control #3", "control #4",
                              "knock out #1", "knock out #2",
                              "knock out #3", "knock out #4")) +
  geom_vline(xintercept = 3.1, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 4.95, linetype = "solid", color = "black") +
  geom_vline(xintercept = 2.1, linetype = "dotted", color = "black") +
  scale_x_continuous(breaks = c(-10, 0, 2.1, 3.1, 4.95, 10),
                     labels = c("-10", "0", "2.1", "3.1", "4.95", "10"))
```


Samples
knock out #4
knock out #3
knock out #2
knock out #1
control #4
control #3
control #2
control #1
−10
0
2.1
3.1
4.95
10
log2(tpm)



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

Postboks 5003
NO-1432 Ås
Norway