UiT The Arctic University of Norway

Faculty of Biosciences, Fisheries and Economics, Department of Arctic and Marine Biology

Characterization of *Cuzd1* expression and development of *in vivo* CRISPR technology in Atlantic salmon (*Salmo salar* L.)

Sona Hakhverdyan
Master's thesis in Biology, BIO-3950, May, 2023

UiT The Arctic University of Norway

Faculty of Biosciences, Fisheries and Economics, Department of Arctic and Marine Biology

Characterization of Cuzd1 expression and development of *in vivo* CRISPR technology in Atlantic salmon (*Salmo salar*)

Sona Hakhverdyan

Master of Science in Biology – Arctic Animal Physiology

May 2023

Main supervisor

Alexander West, UiT – The Arctic University of Norway

Co-supervisor

Helge Tveiten, UiT – The Arctic University of Norway

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my main supervisor, Alex West, for giving me the opportunity to work on this project. Your support, guidance, and encouragement have been invaluable.

A huge thank you goes to my co-supervisor, Helge Tveiten, for your valuable insights, discussions and teaching me how to perform microinjects.
I am also grateful to both of my supervisors for co-parenting more than 3000 salmon embryos.

Thank you to Chandra Sekhar Ravuri for your guidance and help in the lab and to Fernando for your encouragement and support.

Special thanks go to Jan Eirik and Bjørk at Kårvika, Havbruksstasjonen in Tromsø, for taking excellent care of the salmon when I was not there.

I also want to thank you, Mariel, Silje, Sara, and Daniel, for our long hours at the office and for all the procrastination we did together. I am grateful to Magdalena for being enthusiastic and encouraging about my project.

I want to express my appreciation to Ida Amalie and Courtney for their support when I needed it most and for our amazing trip to Iceland.

I would like to extend my sincere thanks to all of my friends for their unwavering support, encouragement, and understanding throughout this project. I am deeply grateful for their love and constant motivation.

Lastly, a special thank you to my family and my partner André for the never-ending support and to my fluffy dog Nemo for always being there for cuddles and keeping me company.

Sona Hakhverdyan
Tromsø, May 2023

Abstract

The Atlantic salmon (Salmo salar L.) is a commercially important species that exhibits an anadromous life cycle involving migratory movements between freshwater and marine environments. Understanding the fundamental biology of the Atlantic salmon, particularly during key life history transitions such as smoltification, is therefore of great importance. In light of modern biotechnology and the completed reference genome for Atlantic Salmon, we can now examine the parr-smolt transition using a hypothesis-free characterization of gene expression analyses. We hypothesized that CUB and Zona Pellucida-like Domain 1 (*Cuzd1*) gene played a critical role in changing gill physiology during smoltification. To test this hypothesis, the study had two major aims: 1) to assess *Cuzd1* expression in the whole fish through ontogeny and in various tissues during smoltification; 2) to develop *in vivo* CRISPR/Cas9 tools in Atlantic salmon, including the design of novel guides to enable functional characterization of *Cuzd1*. There was a significant difference in gene expression of *Cuzd1* between the organ development stage and eyed stage, hatching, fin development, mid-point sampling, and approaching external feeding. This could provide valuable insight into the gene's developmental patterns. The gene expression during smoltification in various tissues during smoltification also varied with statistically significant differences in the muscle, gill, and brain. This finding implies that the role of *Cuzd1* during smoltification is not limited to the gill. Thus, *Cuzd1* exhibited tissue-specific expression patterns and could shed light on the mechanisms involved in the physiological changes during smoltification. A successful knockout of the positive control *Slc45a2*, a gene that affects pigmentation, was achieved. However, no evidence of *Cuzd1* knockout could be observed, and further investigation is needed to enable a proper assessment. The result of this study demonstrated strong potential of the tested methodology for improving our understanding of Atlantic salmon biology and contributing to the development of novel methods for genetic characterization. Future work could be aimed at improving the methodology, including the use of promiscuous gRNAs and more comprehensive methods for genetic characterization. This work represents a significant step towards understanding the complex developmental transition of smoltification in Atlantic salmon. Overall, this study expands our characterization of *Cuzd1* expression in Atlantic salmon and reports the first successful *in vivo* CRISPR knockouts at UiT, The Arctic University of Norway.

Keywords: Atlantic salmon (*Salmo salar*), smoltification, CRISPR/Cas9, gRNA, *Cuzd1*, *Slc45a2*.

Table of Contents

List of tables

List of figures

List of abbreviations

AC: Accessory cell

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

crRNA: CRISPR RNA

Cuzd1: CUB and Zona Pellucida-like Domain 1

DSB: Double-strand break

LP: Long photoperiod

MRC: Mitochondrion-rich cell

NDC: Non-differentiated cell

NKA: Sodium-potassium ATPase

PVC: Pavement cell

PAM: Protospacer adjacent motif

SP: Short photoperiod

Slc45a2: Solute carrier family 45, member 2

tracrRNA: Trans-activating RNA

ZFN: Zinc-finger nuclease

1 Introduction

1.1 Life history and smoltification of Atlantic salmon

The Atlantic salmon (*Salmo salar* L.) is a teleost fish and a member of the Salmoninae subfamily, which is one of three subfamilies within the Salmonidae family (order Salmoniformes) (Fricke et al., 2023). All Salmonidae subfamilies are characterized by complex life history patterns. Many species live alongside each other in diverse morphologies and follow different life history trajectories depending on the surrounding ecosystem and niche conditions (Dodson et al., 2013; Skulason & Smith, 1995). Atlantic salmon's natural distribution is restricted to regions of the North Atlantic Ocean and surrounding watersheds with temperate and subarctic climates. A wide variety of life history strategies are displayed in Salmonidae, including limnetic, marine, and anadromous. The Atlantic salmon is an anadromous (Wedemeyer, 1980) species which means that the life cycle includes periods of migratory movement between freshwater and marine environments (**Figure 1**).

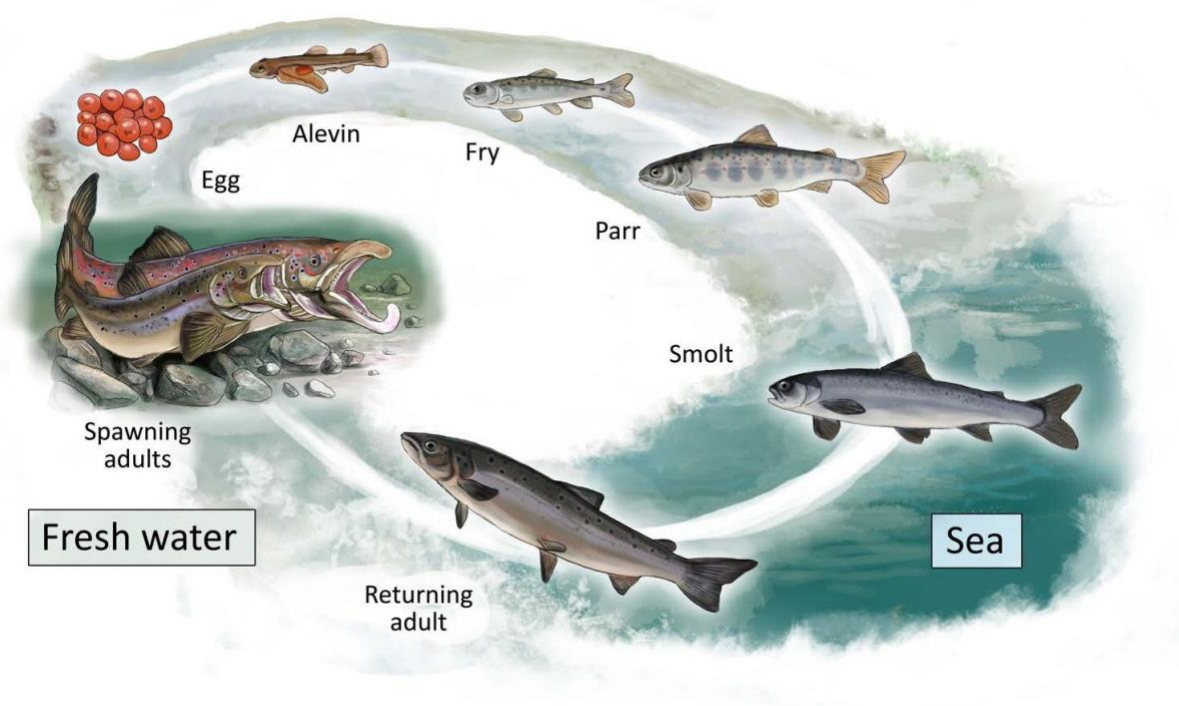

***Figure 1**. The life cycle of the Atlantic Salmon. Illustration by Jayme van Dalum.*

Freshwater streams are where the salmon eggs hatch and develop before migrating to the sea. In the late autumn, mature Atlantic salmon mate and lay their eggs in riverbed gravel, and the eggs hatch the following spring (Fleming, 1996). Before emerging from the riverbed, the alevin

feeds on the remaining yolk sack over a considerable time. At that stage, the small fish are known as fry. As the fry grows, they eventually develop into parr, characterized by the dark markings along their sides.

Wild and captive salmonid life history is fundamentally influenced by smoltification, which occurs when a benthic parr under freshwater conditions transitions to a pelagic, saltwater-ready smolt. Salmon migration between their native river and sea depends on physiological processes and environmental inputs, ensuring correct phenology. Atlantic salmon live in a region where photoperiod and temperature are subject to significant seasonal differences and are the most important environmental factors affecting smolting time (Hoar & Randall, 1988). Temperature affects the rate of smolting, and photoperiod drives the timing (McCormick, 2013; Wedemeyer, 1980). Within and among populations, the time spent as a parr varies (Klemetsen et al., 2003). Parr has a certain size threshold where they either smolt (above threshold) or stay as parr for another year (below threshold) (Thorpe, 1994). Once the parr develops beyond the size threshold, studies have shown that short (winter) photoperiods (SP), followed by an increase in photoperiod, are crucial for the pre-smolt development (Berge et al., 1995; Ebbesson et al., 2007; Handeland & Stefansson, 2001). The migration occurs later in the spring as part of the smoltification process. Smoltification is a complex multi-tissue developmental transition, and it is well characterized within the osmoregulatory capacity of the gill (McCormick, 2013).

1.2 Osmoregulation

As a euryhaline teleost, Atlantic salmon regulate their internal salt and water balances via the gill (McCormick, 2013). When salmon are in freshwater streams, their body fluids have a higher concentration of ions than the surrounding water, leading to a net movement of water into the salmon's body by osmosis. Parr deals with this environmental pressure by pumping in ions through the gill epithelium. In contrast, smolts that migrate into the ocean rely on the ability to absorb water and excrete salts to survive. The gills undergo respective changes during smoltification to allow them to adapt physiologically to the increased salt concentration of seawater. Specialized mitochondrion-rich cells (MRCs) actively transport ions out of the body to maintain their ion balance (**Figure 2**) (Evans et al., 2005; Stefansson, 2008). The best-characterized cell types that make up teleost gill epithelia are MRCs, pavement cells (PVCs), and accessory cells (ACs) (Evans et al., 2005). These cells express specific channels and ion pumps to maintain the osmotic gradient. The ability of the MRCs to excrete and take up ions is dependent on the Na^+/K^+-ATPase (NKA) pump, Na-K-Cl cotransporter (NKCC), and cystic

fibrosis transmembrane conductance regulator (CFTR) abundance. MRCs are crucial in maintaining osmoregulatory balance in both freshwater (FW) and saltwater (SW) environments. However, the mechanisms underlying the role of MRCs in osmoregulation differ in FW and SW. In FW environments, MRCs uptake ions such as Na+ and Cl- from the surrounding water, while in SW environments, MRCs excrete excess ions to maintain the osmotic balance. In the FW gill, Na^+ and Cl^- are taken up by the MRCs through specialized ion channels by the exchange of HCO_3^- and H^+ (McCormick, 2013). In the SW gill, the NKCC allows a flow of Na^+ and Cl^- into the MRC with the help of the electrochemical gradient from the NKA pump (McCormick, 2013). Chloride ions are transported out of the MRC through the CFTR channel, and Na^+ exits the gill into the saltwater through the tight junction pathways. Chloride and sodium ions flow outward to help reduce the fish's internal salt concentration (Evans et al., 2005; McCormick, 2013; Stefansson, 2008).

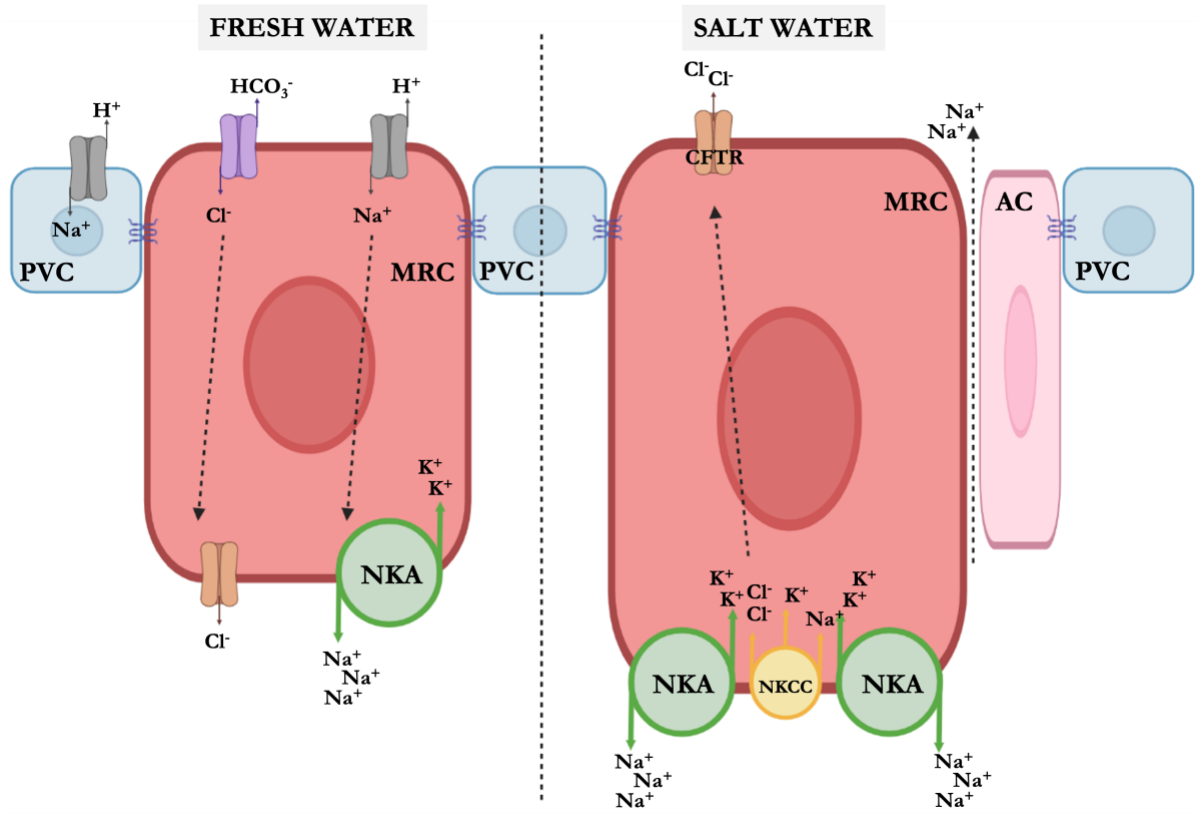

***Figure 2**. Illustration of the mitochondrion-rich cell (MRC). PVC: Pavement cell. AC: Accessory cell. Freshwater: The MRCs are associated with PVCs, forming intercellular junctions impermeable to ions, which reduces ion loss to the water. Na^+ and Cl^- are taken up by MRCs by the exchange of HCO_3^- and H^+. Saltwater: The CFTR transporter protein secretes Cl^- into the external environment. Other MRCs and ACs are seen in complexes with MRCs, and the junctions between MRCs and ACs are loose to facilitate Na^+ leakage. Figure modified from (Iversen, 2021) Ph.D., created in Biorender and Powerpoint.*

Beyond osmoregulation, the gill is essential for diverse functions, including gas exchange, nitrogenous excretion, and pH balance. The preadaptation to saltwater for juvenile salmon involves changes in behavior, morphology, and physiology (McCormick, 2013; Stefansson, 2008; West et al., 2021). The transition during smoltification that allows changes in osmoregulatory abilities has been studied extensively (Evans et al., 2005; Iversen et al., 2020; Stefansson, 2008), and it was demonstrated that the development of hypo-osmoregulatory ability as part of smoltification is an important change (Hoar & Randall, 1988). This is associated with the change in the expression of CFTR and the subunit expression of the NKA (Khaw et al., 2021). The NKA (subunits *NKAα1a* and *NKAα1b)* and CFTR channels may depend more on long photoperiod (LP) exposure than SP exposure. This distinction is important as smolts have better saltwater growth performance when given SP exposure (Iversen et al., 2020).

1.3 Recent advances and the importance of *Cuzd1* in changing gill physiology during smoltification

Early work on the molecular characterization of the gill has focused on key genes with a known role in osmoregulation (McCormick et al., 2009; Stefansson, 2008), including histological studies of the salmon gill in early development (Lubin et al., 1989; Pisam et al., 1988). In light of modern biotechnology and the completed reference genome for Atlantic Salmon (Lien et al., 2016), we can now examine the transition using a hypothesis-free characterization of gene expression analyses. Transcriptomic studies agree with previous work examining the expression of functional ion channels, NKA subunits; *NKAα1a* (freshwater ion regulation gene) and *NKAα1b* (saltwater ion regulation gene), and CFTR abundantly expressed in the gills (Houde, Akbarzadeh, et al., 2019; Houde, Gunther, et al., 2019; Iversen et al., 2020; West et al., 2021). Iversen et al. (2020) drew the distinction between the length of photoperiod change. Several genes are induced by a change from LP to SP which are likely important for the induction of other genes by a change from SP to LP. Thus, the two signals are linked to each other. Iversen et al. (2020) investigated SP-dependent gene expression, where they found that the fish had better saltwater growth performance and that LP exposure resulted in a higher growth rate. Interestingly, two studies (Iversen et al., 2020; West et al., 2021) showed that the increased expression of classical smolt markers (NKA subunits and CFTR) are normally dependent on a change from SP to LP but not necessarily dependent upon SP exposure. It was clear from the data that the expression of CFTR, *NKAα1a, and NKAα1b* (**Figure 3**) appeared to be more dependent upon LP than its exposure to SP. This is important because SP-exposed

fish have a stronger saltwater capacity (Iversen et al., 2020), and these genes are important for the marine survival of the Atlantic salmon.

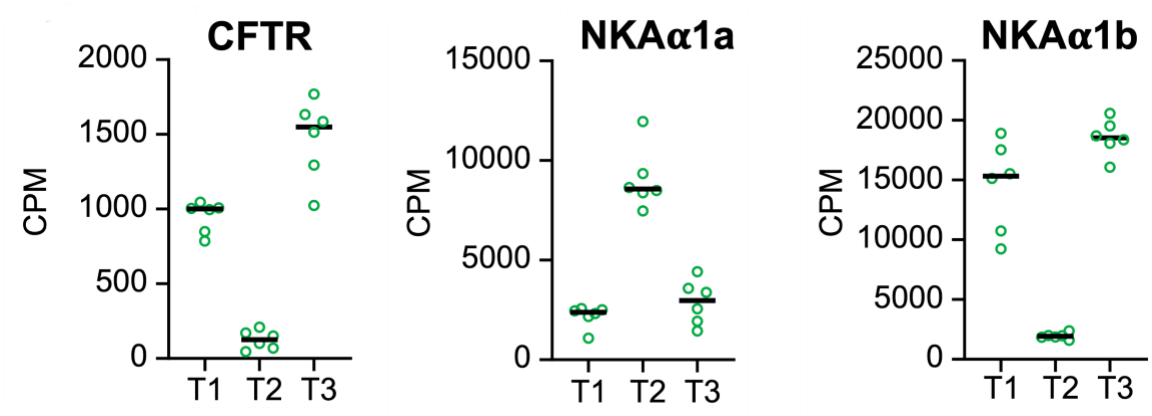

***Figure 3**. Presenting RNAseq data from West et al., 2021, showing the expression of saltwater ion regulation genes CFTR and NKAα1b and freshwater ion regulation gene NKAα1a. T1 is parr at experiment start with constant light exposure (LL), T2 is SP exposure for 6 weeks, and T3 is LL for 6 weeks.*

New transcriptomic studies identify other genes involved in smoltification. Interestingly, classical smolt markers are not SP-dependent, although there is a group that is SP dependent. West et al. (2021) identified that 9746 genes were differentially expressed during smoltification. They performed a Pearson clustering analysis, identifying a statistically significant SP-dependent cluster of 3170 genes. They focused on this cluster set to identify high-fold change expression genes, which they thought to be important in the gill during smoltification. We extracted their raw data and plotted it as the absolute expression in T3 and log fold change between T1 and T3, highlighting our gene of interest, CUB and Zona Pellucida-like Domain 1 (*Cuzd1*) (**Figure 4**). This emphasizes the importance of studying this gene further. In **Figure 4**, *Cuzd1* is singled out as an especially high fold change – high expression gene (raw data from the SP-dependent cluster of the top 100 most expressed genes in T3 and with the highest fold change between T1 and T3 are presented in **Table S 1** and **Table S 2**.

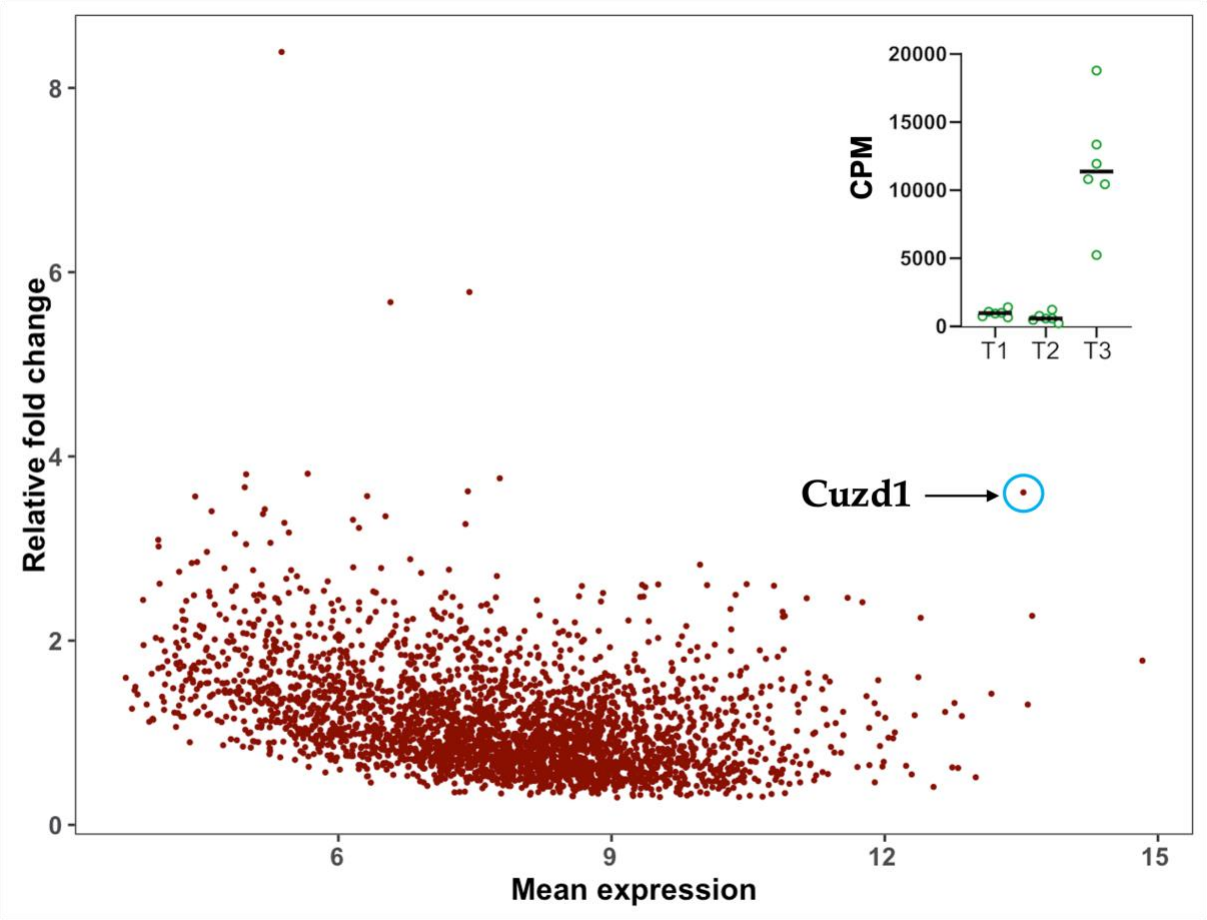

Figure 4*. MA-plot presenting 3170 genes of the SP-dependent cluster with their corresponding values of relative fold change for the difference between expression values at T1 and T3 (in logarithmic scale, Y-axis) and values of mean CPM expression at T3 (in logarithmic scale, X-axis). The point in the blue circle singles out Cuzd1 with the outstanding combination of high fold change (3.6) and high mean expression (13.5). Presented in the right corner is the RNAseq data for Cuzd1. Data obtained from West et al. (2021).*

Moreover, *Cuzd1* stands out in its expression induction and absolute expression levels. This makes a strong case for *Cuzd1* being a key player in the parr-smolt transition of the gill. The snRNAseq work shows a cell-resolution transcriptomic description of the Atlantic salmon gill **Figure 5**. The *Cuzd1* gene, in difference from other genes, demonstrated an outstanding pattern within its specific cluster in terms of its high expression in sample point T3, compared to T1 and T2 (**Figure 4**). This showed that *Cuzd1* is not induced by SP but dependent on a previous SP to be induced by a LP, showing winter dependency for the gene. Thus, suggesting the gene could be involved in the parr-smolt transition. **Figure 5A** presents the SP-dependent *Cuzd1* gene and its cluster-specific expression in non-differentiated cells (NDCs), and **Figure 5B** presents the NKA subunits and their cluster-specific expression in MRCs and ACs. NDCs are cells that have not yet developed into specific cell types and can differentiate into various cell types depending on the signals they receive. In the gills of Atlantic salmon, NDCs may

differentiate into various cell types depending on the gills' developmental stage and specific functional demands. *Cuzd1* expression may be involved in regulating the differentiation of NDCs into specific cell types in the gills (West et al., 2021). However, this is only hypothesized, and the role of NDCs and *Cuzd1* in smoltification still needs to be determined. *Cuzd1* is a valuable target for functional characterization in this study, reflecting the importance of studying this gene further.

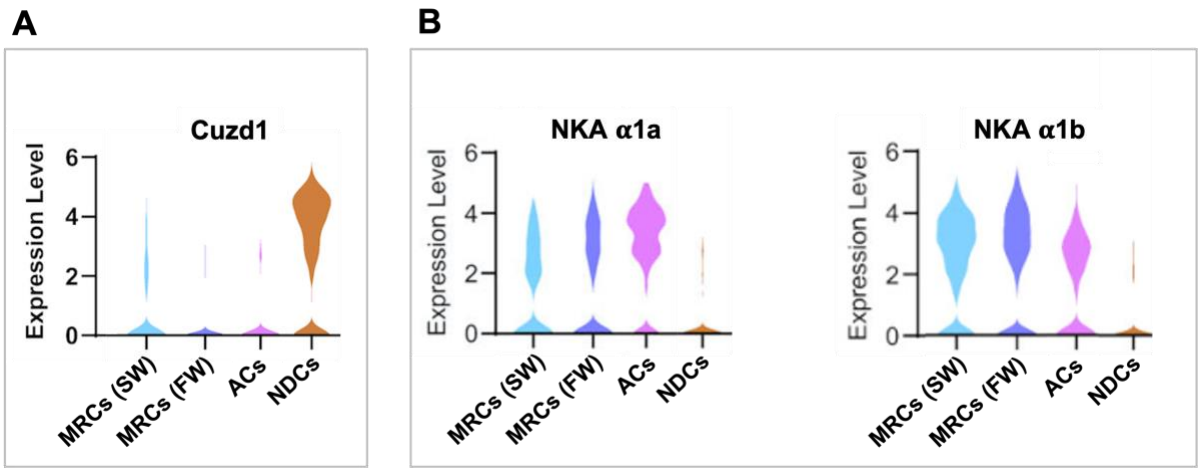

***Figure 5**. **A)** Representation of snRNA data for Cuzd1 shows the expression level for its specific cell cluster: NDCs. **B)** Representation of scaled and log1p normalized count data for NKAα1a and NKAα1b. The violin plot shows the expression level of NKAα1a and NKAα1b for two specific cell clusters: MRCs and ACs. Data obtained from West et al. (2021).*

The current information on *Cuzd1* in Atlantic salmon is sparse. However, previous work on *Cuzd1* has shown that it is associated with tumorigenesis, specifically overexpressed in the uterus and ovary of rats, and elevated levels were found in ovarian, lung, and breast cancer in human patients (Liaskos et al., 2013). *Cuzd1* also appears to play a role in mammary gland proliferation and maturation during pregnancy and lactation by expressing in the mammary ductal and alveolar epithelium (Mapes et al., 2018). In addition, the gene is associated with the branching and functionality of exocrine and endocrine organs in mammals and may be linked to the prolactin JAK/STAT signaling pathway (Mapes et al., 2018). This study investigated the difference in *Cuzd1* expression in wild-type (WT) and knockout (KO) mice, and they showed that there is a difference in the phenotype in the branching of the mammary ductal epithelium. They continue to its molecular fingerprints of the phenotype and highlight a few genes associated with epidermal growth factor (EGF) cycling, which are strongly associated with *Cuzd1* signaling. This is about an upregulation in prolactin signaling, whereas in smoltification,

it is counterintuitive; smoltification is associated with the downregulation of prolactin (Bernard et al., 2020; McCormick, 2013).

Correlative gene expression work can be informative. Nevertheless, to understand the role of *Cuzd1*, a functional analysis must be performed. Recent advances make this possible in salmon through Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 (CRISPR-associated) technology.

1.4 Genome editing and CRISPR/Cas9

Genome editing is a molecular technique with many applications, such as medicine, molecular biology, biotechnology, agriculture, and farming. With the emergence of gene editing techniques, it has become possible to create targeted mutations more precisely in recent years. The process of genome editing involves removing, adding, or modifying genetic sequences at specific locations in the DNA of an organism. Genome editing is a highly efficient method if a targeted DNA double-strand break (DSB) is achieved in the desired chromosomal sequence. Using Zinc finger nucleases (ZFN), zebrafish were used as one of the first teleosts to undergo gene editing experiments (Doyon et al., 2008; Meng et al., 2008) and later on by the transcription activator-like effectors (TALENS) (Huang et al., 2011). A study by (Yano et al., 2013) used ZFN to induce targeted mutations in the *sdY* gene of rainbow trout to confirm its role in determining sex. Moreover, ZFN technology has been applied to a catfish species to unravel the mechanisms of muscle growth (Dong et al., 2011). Since salmonids have a long generation time (2-4 years), this methodology has its disadvantages because previously targeted mutagenesis protocols cannot efficiently induce bi-allelic mutations (Edvardsen et al., 2014). Recent technological advancements in the CRISPR/Cas9 system have proven to be highly efficient at specifically inducing bi-allelic mutations in the first generation (F0) in zebrafish (Jao et al., 2013) and Xenopus (Blitz et al., 2013). With this characteristic, a non-model organism, such as salmon with a longer life span, may be adapted to this new technology with great success. Its efficiency has made CRISPR/Cas9 one of the most popular genome editing tools in recent years. This technology is proven useful and can be used to identify the role of a specific gene in a particular trait, e.g., genes involved in disease resistance (Staller et al., 2019).

Adapted from the *Streptococcus pyogenes* bacteria's defense system (Deltcheva et al., 2011), CRISPR/Cas9 has revolutionized genome editing technology (Adli, 2018; Zhang, 2019). Short fragments of DNA are integrated into the bacterium's CRISPR repeat-spacer array, known as a protospacer sequence, in response to foreign genetic elements. By doing this, bacteria will be protected against future invasion by the same phage (Barrangou et al., 2007;

Mojica et al., 2005). Small CRISPR RNAs (crRNA) are synthesized from protospacer sequences (Brouns et al., 2008), which serve as guides to the Cas9 endonuclease to cut viral DNA and prevent horizontal transmission (Deltcheva et al., 2011; Gasiunas et al., 2012; Jinek et al., 2012; Marraffini. & Sontheimer, 2008). Three main components are required to modify a gene through CRISPR/Cas9: The Cas9 protein, guide RNA (gRNA), and protospacer adjacent motif (PAM) (**Figure 6**) (Doudna & Charpentier, 2014; Jiang & Doudna, 2017; Makarova et al., 2011; Zhang, 2019).

Cas9 is a large DNA endonuclease derived from *S. pyogenes*, consisting of 1368 amino acids. Two endonuclease domains are present in Cas9: an HNH-like endonuclease and a RuvC-like nuclease that cuts the target strand and the complementary strand, respectively, which in turn causes a DSB (Chen et al., 2014; Gasiunas et al., 2012; Jinek et al., 2012). The cell has its DNA repair system to repair the DSB with non-homologous end-joining (NHEJ), which frequently leads to insertions or deletions of the nucleotide bases, and a truncated, non-functional protein is produced as a result (Lieber, 2010; Wyman & Kanaar, 2006). A short RNA sequence is used to guide Cas9, followed by a 5'-PAM sequence (Mojica et al., 2005) to identify the target sequence (Brouns et al., 2008). The gRNA is composed of two types of RNA: crRNA and trans-activating (tracrRNA), which form a scaffold linking the crRNA to Cas9 (Brouns et al., 2008; Deltcheva et al., 2011; Gasiunas et al., 2012). The crRNA and the tracrRNA are then combined into a single chimeric RNA (Jinek et al., 2012), making it equally effective at directing Cas9 to the target site. RNA chimeras are known as gRNAs. CRISPR/Cas9 gRNA cloning vectors generally contain a scaffold for gRNA, restriction sites for cloning sequence-specific gRNA, and a promoter for gene expression (**Figure 6**). The PAM sequence plays a crucial role in the CRISPR/Cas9 system. This is a short DNA sequence downstream (3') from the cut site that is required by the Cas9 nuclease to cut the DNA region targeted for cleavage (Mojica et al., 2005).

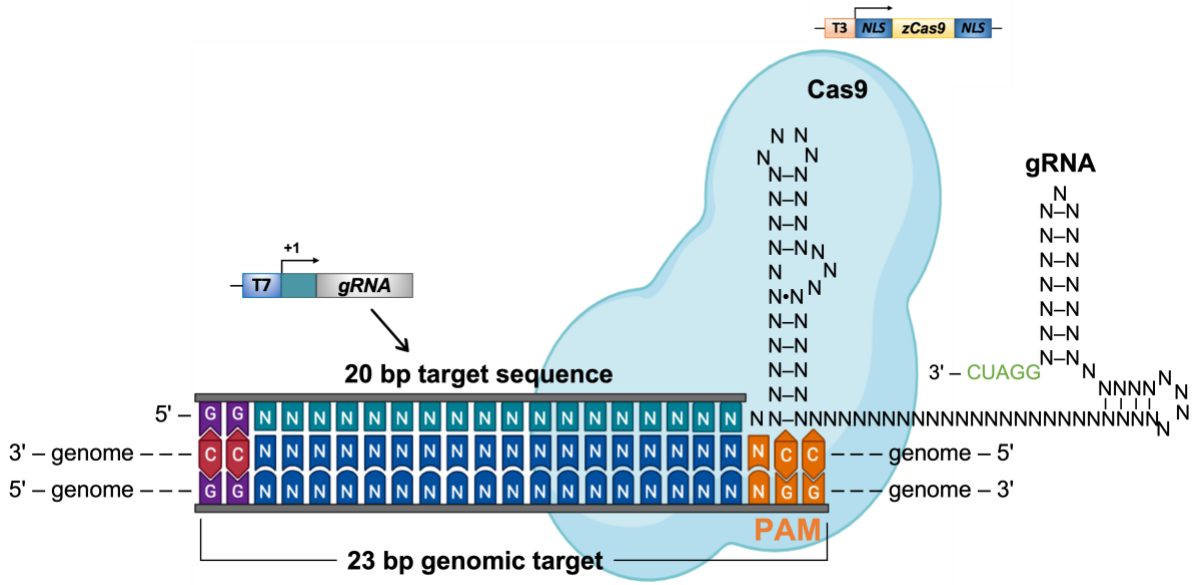

***Figure 6**. Genome editing using the CRISPR/Cas9 system. The system uses a dual codon-optimized Cas9 protein and a gRNA, which consists of a target sequence (lighter blue color) complementary to the genomic target site next to the PAM site. Both components are synthesized as RNAs by in vitro transcription from the T7 (for gRNA) and T3 (for Cas9) promoters. Figure modified from (Jao et al., 2013) and (Moreno-Mateos et al., 2015) , created in Biorender and Powerpoint.*

With the introduction of CRISPR/Cas9, genome editing has undergone a radical change. It has been demonstrated that CRISPR/Cas9 can be efficiently implemented *in vivo* in embryos of several aquaculture species, including Tilapia (Li et al., 2014), Carp (Chakrapani et al., 2016), and Atlantic salmon (Edvardsen et al., 2014). A few of these studies have employed established protocols in model organisms such as zebrafish (Jao et al., 2013) by targeting the genes that produce an observable phenotype, such as pigmentation. For applying CRISPR/Cas9 in vivo in fish, microinjecting the CRISPR/Cas9 components into freshly fertilized eggs at the one-cell stage is currently the method established. Microinjections are more suitable for large cell sizes like embryos, and the success depends on expressing both the gRNA and a Cas9 mRNA. This technology is a tractable solution for functional studies in F0 salmon.

1.4.1 CRISPR/Cas9 – a breakthrough tool for functional studies in Atlantic salmon

Genetic engineering technology allows us to inactivate or remove specific genes from an organism. It can be used to study specific gene functions in the physiology and anatomy of the organism under study. The alteration of genes can lead to phenotype alteration, revealing the specific functions of the genes of interest. It is possible to selectively remove or alter genes in embryos by using CRISPR/Cas9 so that gene loss effects on development and function can be studied. The first application of the CRISPR/Cas9 system in Atlantic salmon was executed by Edvardsen et al. (2014). The two genes successfully knocked out *in vivo* were tyrosinase (*tyr*) and solute carrier family 45, member 2 (*Slc45a2*), where both are responsible for pigmentation in Atlantic salmon. The approach was developed in Atlantic salmon, considering the known function of the target gene in pigmentation. Salmonid genomes are partially tetraploid, making the *Slc45a2* a suitable gene for gene knockout due to its single occurrence in the genome (Edvardsen et al., 2014). CRISPR/Cas9 technology was successfully used in a marine cold-water species for the first time in this study. Thus, they demonstrated that F0 fish could be used for functional studies of Atlantic salmon. Using the same technique in 2016, the same group of researchers knocked out a dead-end gene (*dnd*) to create sterile salmon (Wargelius et al., 2016). Overall, double knockout with *Slc45a2* as a screening strategy has become a standardized approach in many studies (Datsomor et al., 2019; Davis et al., 2021; Edvardsen et al., 2014; Wargelius et al., 2016). Although, there are no studies on the influence that a stand-alone *Slc45a2* knockout has on the performance/functionality of fish, except for the loss of pigmentation. However, this still provides new opportunities to study the salmon genome and improve production and sustainability via genetic improvements in Atlantic salmon utilizing CRISPR/Cas9 and double knockouts.

1.5 Significance, Purpose, and Aims

The Atlantic salmon is a fascinating animal of high commercial importance. Understanding the fundamental biology of the Atlantic salmon, particularly during key life history transitions such as smoltification, is therefore of great importance. Recent work and our own analysis show that the expression of *Cuzd1* in the smolt gill is 1) SP-dependent, 2) impressive, both in relative induction and absolute levels of expression, and 3) limited to a specific yet poorly understood cell group. Based on these data, we hypothesize that *Cuzd1* plays a critical role in changing gill physiology during smoltification. At present, however, we can infer very little from the sparse existing *Cuzd1* literature that might indicate the role that *Cuzd1* plays during smoltification. The goals of this MSc project are, therefore:

1) Conduct an expression analysis to explore the developmental regulation of *Cuzd1* and the tissue-specific expression during smoltification that will help inform the interpretation and design of subsequent projects

The particular objectives are to:

- Measure *Cuzd1* expression through ontogeny
- Measure *Cuzd1* expression in multiple tissues during smoltification

2) Develop and validate *in vivo* CRISPR/Cas9 tools in Atlantic salmon, including the design of novel guides to enable functional characterization of *Cuzd1*.

The particular objectives are to:

- Design and clone gRNAs for *Cuzd1* and *Slc45a2*
- Produce high-quality *in vitro* transcribed Cas9 mRNA and gRNAs
- Microinject the produced Cas9 mRNA and gRNAs into the single-cell stage of Atlantic salmon embryos
- Visually characterize the phenotype of the injected salmon
- Characterize possible genomic mutations created by the CRISPR/Cas9 system

2 Material and methods

Animal Welfare Statement

The Atlantic salmon knockout experiments were conducted at UiT, The Arctic University of Norway. All salmon eggs were kept at Havbruksstasjonen in Tromsø, Kårvik. The experiment was terminated before external feeding was reached, in compliance with Norwegian and European legislation on animal research. The genetically modified (GM) work was approved by Helsedirektoratet (GM application number: 22/35755-2). Parr and smolt salmon were collected for the multi-tissue array from Kårvika. These fish were collected as part of the routine, under standard aquaculture smoltification practices approved by the Norwegian Animal Research Authority.

2.1 Design of gRNA oligonucleotides

The knockout of *Cuzd1* and *Slc45a2* was generated using the CRISPR/Cas9 system (**Figure 6**). The gRNA target sequences for each gene were designed upstream from PAM with the online tool ChopChop (https://chopchop.cbu.uib.no/) and obtained commercially from Sigma-Aldrich. Three gRNAs for *Cuzd1* paralog (LOC106562604) and one target sequence for *Slc45a2* obtained from Edvardsen et al. (2014) (**Table 1**) were selected for CRISPR-induced knockouts and screened against the Atlantic salmon gene assembly. ChopChop software checks for off-target effects and potential self-complementary sites of the RNA that could affect the binding to our target gene sequence of interest. These parameters were kept to a minimum to ensure the efficiency and accuracy of the guides.

***Table 1**. Target sequences for Cuzd1 and Slc45a2 with the PAM-site underlined.*

Name	Gene	Sequence 5' – 3'	Strand	GC %
G1	*Cuzd1*	CTCGTCAATTCCAGTAACAGAGG	+	45
G2	*Cuzd1*	AATGCACGACCAGGTAAATGGGG	-	45
G3	*Cuzd1*	TCCCGTGACAGTCAAAGCCAAGG	-	55
G4	*Slc45a2**	GGGGAACAGGCCGATAAGACTGG	-	60

*Target sequence for *Slc45a2* obtained from Edvardsen et al. (2014).

2.2 Preparation of injection components

The injection components prepared for the injections were the cloned constructs from where the gRNAs were *in vitro* synthesized from the pT7-gRNA vector (Addgene, plasmid #46759) and the Cas9 mRNA which was *in vitro* synthesized from the pT3Ts-nCas9n mRNA vector (Addgene, plasmid #46757) (**Figure S 1**).

2.2.1 Incorporation of *Cuzd1* target sequences into the pT7-gRNA vector

The pT7-gRNA vector was digested with the restriction enzyme BsmBI (NEB) to insert the designed gRNAs. The digestion reaction mixture consisted of 5 µl of 10X NEBuffer, 4.42 µl of vector DNA (pT7-gRNA, 2 µg/µl), 1 µl of BsmBI-v2 RE, 39.58 µl of Nuclease-Free H_2O, with a total reaction mixture volume of 50 µl. The reaction mixture was incubated at 55 °C for 15 min followed by heat inactivation at 80 °C for 20 min. The DNA concentration was measured with NanoDrop Spectrophotometer to 18.9 ng/µl.

For annealing the gRNAs, the reaction mixture was made up of 2 µl of 100 µM oligonucleotide for each forward and reverse gRNA, 4 µl of 5X T4 DNA Ligase Buffer, and 12 µl of Nuclease-Free H_2O with total reaction mixture volume of 20 µl. The annealing cycle consisted of incubation at 95 °C for 5 min, followed by ramping down to 50 °C at 0.1 °C/s, then incubating at 50 °C at 10 min, and cooling to 4 °C at normal ramp speed (0.1 °C/s). The vector DNA was then purified according to QIAquick© PCR Purification Kit (Qiagen) following the manufacturer's protocol.

The annealed oligonucleotides (**Figure 7**) were then ligated into 50 ng (2.8 µl) of the BsmBI digested pT7-gRNA vector using 1 µl T4 DNA ligase (NEB) and 5.22 µl Nuclease-Free H_2O with a total reaction volume of 10 µl. The mixture was incubated at room temperature for 1 hour.

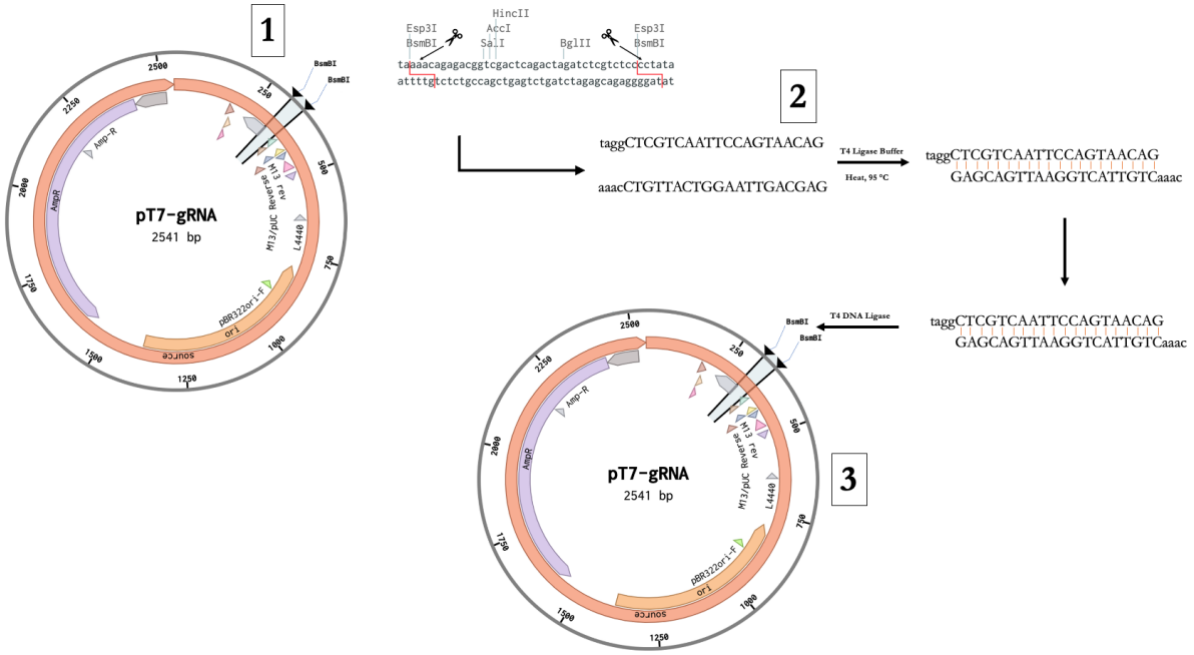

Figure 7*. Recombination of the pT7-gRNA plasmid. 1) Sticky overhang is created when BsmBI cuts the vector at the restriction site. 2) Annealed forward and reverse gRNA oligonucleotides create a heteroduplex compatible with the cutting site of the vector. 3) The heteroduplex is inserted into the vector.*

The solution for the agar plates was made by mixing 20 g of LB Broth with agar (Miller) (L3147, Sigma-Aldrich) and 500 ml of Milli-Q® (MQ) H_2O. In addition, 12.5 g of LB broth (Miller) (L3522, Sigma-Aldrich) was mixed with 500 ml of MQ H_2O. Both broths were autoclaved for 1.5 hours. Ampicillin (50 mg/ml) was added once the LB broth with agar was cooled to ~55 °C.

The products from the ligation were then transformed into NEB 5-alpha competent *E.coli* cells (NEB, Ref: C2987I), on the DH5α strain (**Figure 8**). DH5α cells were first thawed on ice, and then 2 µl of each ligation product and 25 µl of the competent DH5α cells were added to 4 different Eppendorf tubes and mixed thoroughly. The tubes were kept on ice for 20 min followed by heat-shock in a water bath at 42 °C for 30 seconds with precise timing to facilitate the uptake of the plasmids into the cells. The tubes were then incubated on ice for another 5 min. To each tube, 475 µl of SOC Outgrowth Medium (NEB, Ref: B9020) was added and then incubated at 37 °C at 250 rpm for 1 hour. The content in each tube was distributed on two LB agar plates containing 50 mg/mL Ampicillin using sterile technique. The plates were labeled according to target genes and gRNA design and incubated upside down at 37 °C overnight (o/n).

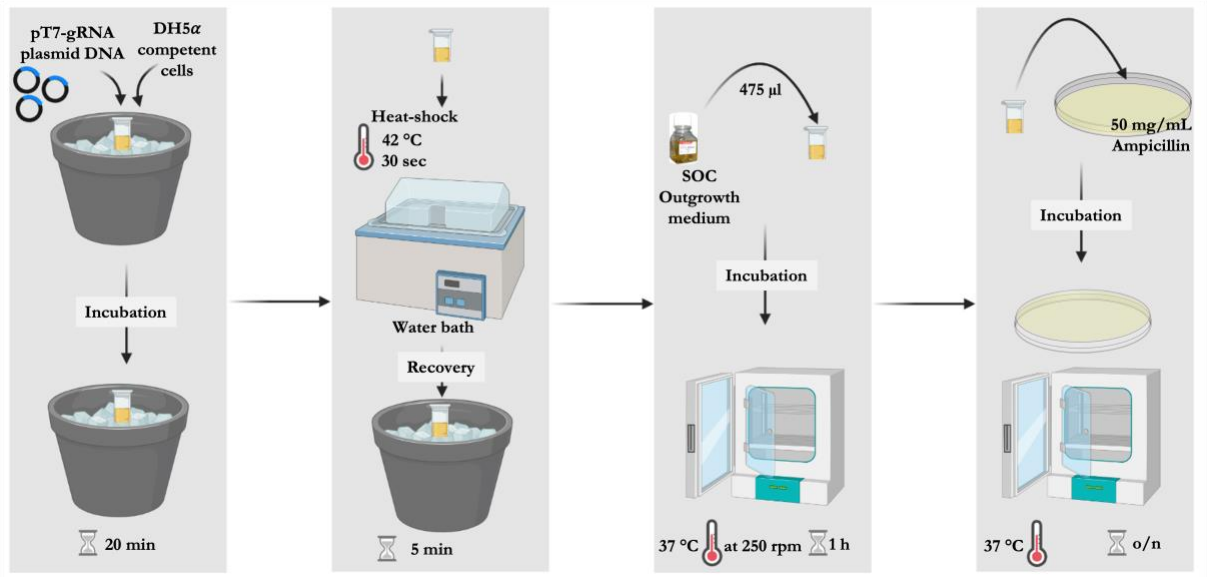

Figure 8. *Flowchart for the transformation process. Product from the ligation and DH5α cells were kept on ice for 20 min (Panel 1) followed by heat-shock at 42 °C for 30 seconds with precise timing and then incubated on ice for another 5 min (Panel 2). SOC Outgrowth medium (475 µl) was added to each tube, then incubated at 37 °C at 250 rpm for 1h (Panel 3). The content of each tube was transferred onto LB agar plates containing 50 mg/mL Ampicillin. Plates were incubated upside down at 37 °C overnight (Panel 4),*

Two colonies from each plate were selected for cultivation o/n in 5 mL of LB Broth in Falcon tubes (15 mL) with one colony per tube (total of 16 tubes). Sterile pipette tips were used to transfer the colonies. The lid of the tubes was slightly closed and secured with tape, allowing air into the tube. Thereafter, the tubes were incubated o/n at 37 °C at 200 rpm.

The o/n incubated cultures were purified using QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's protocol, and DNA concentrations were measured using NanoDrop Spectrophotometer (**Table S 3**).

To identify which clones had inserts of the predicted size, PCR was performed according to Phusion® DNA Polymerase protocol following the manufacturer's protocol. PCR Master Mix was done for 34 reactions, with a total reaction volume of 20 µl in each PCR tube. The protocol was modified; instead of the designed forward primers, our designed gRNA reverse oligonucleotides were used with the pT7-gRNA M13 forward primer (**Table 2**). Less than 250 ng of the DNA template was added.

***Table 2**. gRNA PCR reverse primers (G1-G4) and the pT7-gRNA vector M13 forward primer.*

Name	**Sequence 5' – 3'**
(G1) Cuzd1 gRNA PCR R	aaacCTGTTACTGGAATTGACGAG
(G2) Cuzd1 gRNA PCR R	aaacCATTTACCTGGTCGTGCATT
(G3) Cuzd1 gRNA PCR R	aaacTGGCTTTGACTGTCACGGGA
**(G4) Slc45a2 gRNA PCR R*	aaacGTCTTATCGGCCTGTTCC
M13 PCR F	TGTAAAACGACGGCCAGT

* Primer obtained from Edvardsen et al. (2014)

In total, 32 reactions were used in the PCR. The PCR consisted of initial denaturation (98 °C, 30 s), denaturation (98 °C, 5 s), annealing (58 °C, 10 s), extension (72 °C, 15 s), and cooling (4 °C). The denaturation/annealing/extension cycle was repeated for 40 cycles. The PCR products were analyzed by gel electrophoresis using a 2 % agarose gel (with SYBR safe) to investigate if the plasmid digestion succeeded.

Nine purified plasmids were chosen to be amplified and sent for sequencing. The plasmids were prepared for sequencing using Big Dye ™ Terminator v3.1 Cycle sequencing kit (ThermoFisher). A master mix was made for ten reactions. Each PCR reaction mixture consisted of 0.5 µl of pT7-gRNA M13 forward primer, 0.5 µl BigDye, 3 µl sequencing buffer, 15 µl Nuclease-Free H_2O and 1 µl each of the chosen purified DNA (diluted to 100 ng/µl). The total volume was 20 µl in each of the nine PCR tubes. The PCR consisted of initial denaturation (96 °C, 5 min), denaturation (96 °C, 10 s), annealing (50 °C, 5 s), extension (60 °C, 4 min), and cooling (4 °C, ∞). The denaturation/annealing/extension cycle was repeated 40 times. The PCR products were sequenced in the sequencing lab at the Medicinal Genetics Department (University Hospital in Northern Norway HF). The insertion of the target gRNAs was confirmed using Benchling (Biology software, https://benchling.com, 2022).

Four cloned construct samples of the highest concentration were chosen for isolation of the plasmids and to further yield a higher concentration according to HiSpeed Plasmid Maxi Kit (Qiagen). The concentration of DNA was measured by NanoDrop Spectrophotometer (G1 = 249.9 ng/µl; G2 = 272,5 ng/µl; G3 = 323.3 ng/µl; G4 = 282,2 ng/µl).

2.2.2 gRNA production for *Cuzd1* and *Slc45a2*

The gRNAs were produced from the above-described cloned constructs. The cloned constructs were linearized with BamHI (NEB) and incubated at 37°C for 1 hour. The linearization was confirmed with agarose gel electrophoresis. The linearized plasmid was then purified using QIAquick© PCR Purification Kit (Qiagen).

The *in vitro* transcription was set up using the MEGAshortscript T7 Kit (Amnion/Invitrogen) with ~400 ng of purified DNA, with some alterations to the manufacturer´s protocol (using between 1.2 – 1.6 µg of purified DNA for each guide). The products were purified with the mirVana miRNA isolation Kit (Invitrogen) according to the manufacturer´s protocol. The concentration of RNA was measured by NanoDrop Stectrophotometer (G1 = 171.1 ng/µl; G2 = 81.9 ng/µl; G3 = 117.1 ng/µl; G4 = 201.7 ng/µl). The purity of the RNA was analyzed with Bioanalyzer Small RNA analysis (Agilent 2100).

2.2.3 Cas9 mRNA production

The Cas9 mRNA was produced from the pT3Ts-nCas9n mRNA vector, linearized with XbaI (NEB), and then a capped nls-zCas9-nls RNA was synthesized by *in vitro* transcription using mMESSAGE mMACHINE T3 kit (Invitrogen) with some alterations to the manufacturer's protocol (reaction volume was scaled up to 100 µl). The capped nls-zCas9-nls RNA was purified using RNeasy Mini kit (Qiagen) and mirVana miRNA isolation Kit (Invitrogen) following both manufacturer's protocols. The concentration of RNA was measured by NanoDrop Spectrophotometer: Cas9 mRNA (mirVana) ≈ 3200 ng/µl; Cas9 mRNA (RNeasy MiniKit) = 447.4 ng/µl. The purity of the RNA was analyzed with TapeStation analysis (Agilent 2100).

2.3 Fertilization and injection procedure

The salmon sperm and eggs were obtained from Aquagen (Trondheim, Norway). The eggs were fertilized with sperm in freshwater (7°C) containing 1 mM reduced Glutathione according to the scheme presented in **Figure 9**.

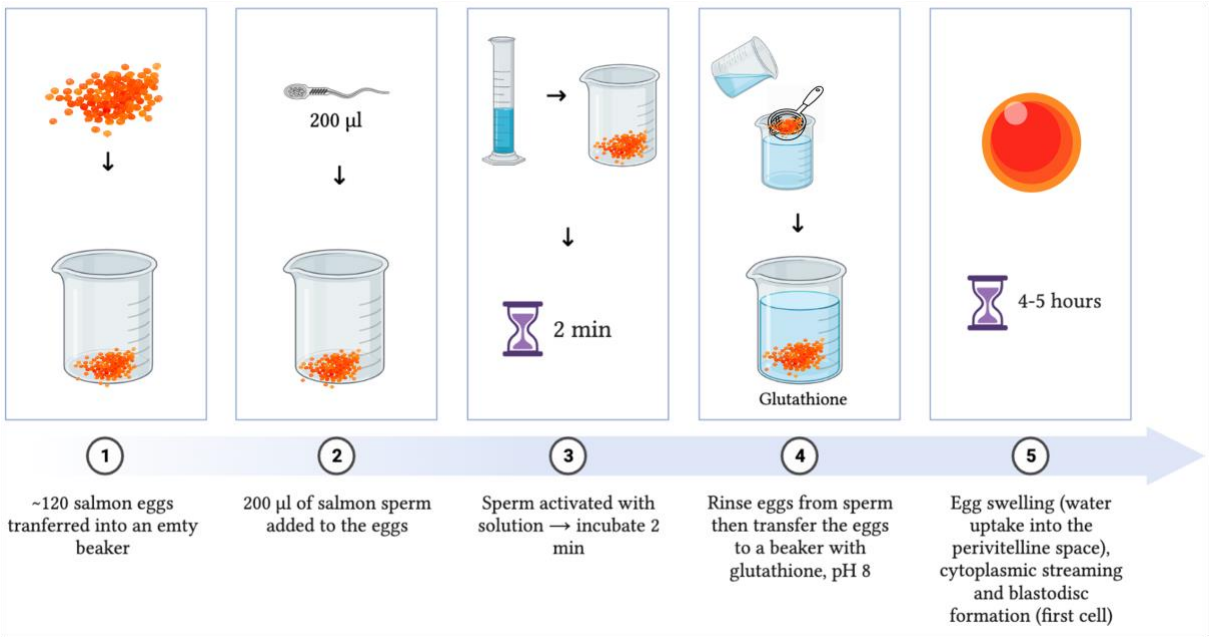

Figure 9*. The fertilization scheme presents the following steps: 1) Placement of a tablespoon of eggs (~120 eggs) into an empty beaker; 2) Additions of the Salmon sperm (200 µl) to the eggs; 3) Sperm activation and incubation at 7°C for 2 min; 4) Rinsing the eggs from the sperm and eggs placed in 1.0 mM reduced glutathione for the activation; 5) Water uptake into the perivitelline space, cytoplasmic streaming and blastodisc formation (first cell).*

The embryos were then incubated for 3 – 4 hours at 7 °C until a visible first cell, and at that point, the eggs were injected for the knockout of the genes of interest (*Slc45a2* and *Cuzd1)*. The eggs were placed on an in-house-made holder and injected group by group (**Table 4**). The injection mixture (20 µl) consisted of phenol red, Cas9 mRNA (150 ng/µl), and respective gRNA (50 ng/µl) (**Table 3**) and was loaded into the microneedle with an Eppendorf microloader pipette tip. About 30 nl, representing ~5% of the blastodic volume, was injected. All injection mixtures contained Phenol red (0.1 %) to visually ensure that the mixture was injected correctly.

***Table 3**. RNA Concentration and volume of components in the different injection mixtures.*

Component	Initial concentration (ng/µl)	Mix 1 (µl)	Mix 2 (µl)	Mix 3 (µl)	Mix 4 (µl)
Cas9 mRNA	447.4	6.7	6.7	-	-
Cas9 mRNA	1411.5	-	-	2.12	2.12
Cuzd1 G1	171.5	-	5.8	-	-
Cuzd1 G2	81.9	-	-	12.2	-
Cuzd1 G3	117.1	-	-	-	8.54
Slc45a2 G4	201.7	5	5	5	5
Phenol red	-	2	2	2	2
NF H_2O	-	6.3	0.5	-	2.3
Tot vol (µl)		20	20	20	20

***Table 4**. Number of eggs included in the control and injected groups.*

Name	Group	Number of eggs
C	Control	~3000
G1	*Cuzd1 + Slc45a2 + Cas9*	241
G2	*Cuzd1 + Slc45a2 + Cas9*	245
G3	*Cuzd1 + Slc45a2 + Cas9*	224
G4	*Slc45a2 + Cas9 (Positive control)*	240

2.4 Egg storage and sampling

After injections, the eggs were kept in a water-running tank at UiT at an ambient temperature of ~5°C. The temperature in the water tank was recorded, and the eggs were observed under the stereoscope every day to follow their development until they reached 100% epiboly (end of gastrulation). The age of the fish is always given in day degrees (DG) because development depends on the incubation temperature. DG is calculated from days since fertilization multiplied by incubation temperature (in °C). At 177 DG (32 days after fertilization), the eggs were moved to Havbruksstasjonen in Tromsø, Kårvika, where they were kept at an ambient temperature of ~8°C for the rest of the experiment before they reached external feeding at 811 DG (119 days after fertilization).

2.4.1 Sampling points throughout the experiment

Sampling was done during different development stages, according to **Table 5**. The eggs sampled for histology were stored in 4 % paraformaldehyde (PFA) and RNA later solution for later RNA extraction.

***Table 5**. Number of sampled eggs for the subsequent RNA analysis and description of respective development stages.*

Sampling point and development stage	**Date** **Group** **DG**	**RNA later solution (RNA analyses)**	**4%PFA (Histology)**
(1) End of gastrulation	231122 C28 **147 DG**	24	24
(2) Organ development	091222 C28 **250 DG**	24	24
(3) Eyed stage	191222 C28/G1/G2/G3/G4 **328 DG**	24/5/5/5/5	24/0/0/0/0
(4) Hatching	090123 C28 **491 DG**	24	24
(5) Fin development	250123 C28 **610 DG**	24	24
(6) Midpoint sampling	060223 C28 **701 DG**	24	24
(7) Approaching external feeding	210223 C28/G1/G2/G3/G4 **811 DG**	24/18/18/10/16	24/5/5/5/5

The mortality was recorded before the move to Kårvika and after that regularly. At the final sampling point, all fish were collected. The control group was collected in 4% PFA and RNA later directly from the water tank. All injected groups were collected accordingly: The fish was moved to a beaker with water and Benzoak Vet 200 mg/ml to euthanize the fish. Then the anesthetics were washed off from the fish in a clean beaker of water. The fish were then moved to the lab, counted, measured, and placed in order of gradually increasing pigmentation. The KO fish were transferred into separate tubes with RNA later solution.

Four salmon parr and smolt were obtained from Havbruksstasjonen in Tromsø, Kårvika, and dissected to harvest 12 different tissues from each fish. The different tissues were transferred to separate tubes with RNA later solution for subsequent RNA isolation.

2.5 Visual phenotyping

Visual characterization was made on the intact KO fish based on the albino phenotype. The fish were placed in order of gradually increasing pigmentation and scored accordingly.

2.6 RNA isolation and cDNA conversion

RNA was isolated from the tissues harvested from parr and smolt with a total of 102 samples. Two extraction methods were used, dependent on the tissue type: RNeasy MiniKit (Qiagen) and RNeasy Plus Universal MiniKit (Qiagen) (**Table S 4**). The RNA concentrations were measured using NanoDrop Spectrophotometer (**Table S 5**). The RNA was converted to cDNA according to High-Capacity RNA-to-cDNA™ Kit following the manufacturer's protocol.

RNA was also isolated from the control fish group (from the different sampling points, **Table 5**) using RNeasy Plus Universal MiniKit (Qiagen) and Trizol-Chloroform-Isopropanol method (see Appendix C for protocol). Triplicates were used for each sampling point, starting from organ development (2A-C – 7A-C). The RNA concentrations were measured using NanoDrop Spectrophotometer. The RNA was converted to cDNA according to High-Capacity RNA-to-cDNA™ Kit following the manufacturer's protocol.

2.7 qPCR primer design

qPCR primers (**Table 6**) were designed for the *Cuzd1* gene with the use of Benchling (Biology software, https://benchling.com, 2022) and Primer3 (https://primer3.ut.ee/) specifically to match the target gene of interest. The housekeeping gene *EF1-α* (LOC100136525) was used as a reference. The primers for the *Cuzd1* paralog (LOC106562604) were found by retrieving the cDNA from Salmobase (https://salmobase.org/) using the gene ID and thereafter copying the sequence into Benchling and highlighting the exons.

Table 6*. qPCR primers specific for Cuzd1 and EF1-alpha*

Entry	Name	Sequence 5'-3'	Annealing temp. (°C)
1	*Cuzd1 qPCR F1*	CCATAATAATGTCTGGAACGG	57
	Cuzd1 qPCR R1	GACAAGCTCCTGATAATGG	57
2	*Cuzd1 qPCR F2*	AATTCATCTTGACGCTGC	57
	Cuzd1 qPCR R2	TGGTTTCCTTGAGTTTGATG	57
*3**	*EF1-α qPCR F3*	GGTACTACGTCACAATCATT	61
	EF1-α qPCR R3	CAATCAGCCTGAGATGTAC	61

* Primer pair for *EF1-α* was tested and validated by Therese Solberg, unpublished

2.7.1 qPCR primer efficiency for *Cuzd1*

To confirm the qPCR primer efficiency for *Cuzd1*, six dilutions (**Table 7**) were made from extracted salmon gill cDNA. qPCR was performed using the GoTaq® qPCR Kit (Promega) following the manufacturer's protocol. For each primer pair, a master mix was prepared (**Table 8**). In a 96-well plate, 19 µl of the master mix was placed in each well, and 1 µl of each dilution template. The qPCR was performed using Bio-Rad CFX manager 3.1, and the thermal cycling condition was: denaturation for 2 min at 95°C followed by 40 cycles of 15 s at 95°C, annealing for 15 s at 57 °C, and 1 min extension at 60°C.

***Table 7**. Dilution series of gill cDNA.*

Dilution	cDNA concentration
1	100 %
2	50 %
3	25 %
4	12.5 %
5	6.25 %
6	3.125 %

***Table 8**. qPCR MasterMix content*

Reagents	x1	x13
GoTaq®	10 µl	130 µl
CXR Reference Dye	0.2 µl	2.6 µl
Forward primer (100 µM)	1 µl	13 µl
Reverse primer (100 µM)	1 µl	13 µl
Template DNA (cDNA)	1 µl	-
Nuclease-free water	6.8 µl	88.4 µl

Using Microsoft Excel (Office 365), cycle threshold (Ct) values were plotted against the log of the dilution concentrations. For each primer pair, the linear equation was obtained through the trendline function in Excel, and the primer efficiency was calculated through Equation 1. Primers with more than 90% efficiency were used (**Figure S 2**).

$$Effiency\ \% = \left(10^{\left(\frac{-1}{slope}\right)-1}\right) * 100 \quad (1)$$

2.7.2 Relative fold gene expression of parr and smolt tissues

qPCR was performed using the GoTaq® qPCR Kit (Promega). qPCR primer pairs F2-R2 and F3-R3 were used, respectively for *Cuzd1* and *EF1-α* (**Table 6**). A 96-well plate was used to accommodate the tissue cDNA. Four plates were used for parr and smolt tissue cDNA with two primer pairs, as outlined above. A master mix was prepared for each plate (**Table 9**), the amount of RNA was normalized to 2 µg, and all tissues were run in 2-8 replicates. The qPCR was performed using Bio-Rad CFX manager 3.1 Thermal with cycling conditions for:

Cuzd1: Denaturation for 2 min at 95°C followed by 40 cycles of 15 s at 95°C, annealing for 15 s at 57 °C, and 1 min extension at 60°C.

Ef1-α: Denaturation for 2 min at 95°C followed by 40 cycles of 15 s at 95°C, annealing for 15 s at 61 °C, and 1 min extension at 60°C.

Table 9*. qPCR MasterMix content for the 96 well plates.*

		Parr x2	**Smolt x2**
Reagents	**x1**	**x80**	**x86**
GoTaq®	10 µl	800 µl	860 µl
CXR Reference Dye	0.2 µl	16 µl	17.2 µl
Forward primer (5 µM)	1 µl	80 µl	86 µl
Reverse primer (5 µM)	1 µl	60 µl	86 µl
Template DNA (cDNA)	2 µg	-	-
Nuclease-free water	5.8 µl	464 µl	499 µl

The relative fold gene expression calculation was conducted only on samples with successful 6-8 replicates in Excel using the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001). The Ct values of the target gene, *Cuzd1*, were normalized against the housekeeping gene *EF1-α.* Relative fold gene expression was quantified by the ΔΔCt values for parr using *EF1a* as the reference gene, meaning the data was plotted for *Cuzd1* expression with smolt relative to parr.

2.7.3 Gene expression of *Cuzd1* through ontogeny

qPCR was performed using the Promega GoTaq® qPCR Kit and primer pairs F2-R2 and F3-R3, respectively for *Cuzd1* and *EF1*-α (**Table 6**). One 96-well plate was used for each primer pair. A master mix was prepared for each plate (**Table 10**), and all control fish cDNA were run in 2 replicates. The qPCR was performed using Bio-Rad CFX manager 3.1 Thermal with the same cycling conditions as described in section 2.7.2.

Table 10. *qPCR MasterMix for two 96 well plates for the control fish in every sampling point.*

Reagents	**x1**	**x38 x2**
GoTaq®	10 µl	380 µl
CXR Reference Dye	0.2 µl	7.6 µl
Forward primer (5 µM)	1 µl	38 µl
Reverse primer (5 µM)	1 µl	38 µl
Template DNA (cDNA)	2 µg	-
Nuclease-free water	5.8 µl	220 µl

2.8 Conventional PCR primer design

The PCR primers for the target gene *Cuzd1* were designed with the online tool ChopChop (https://chopchop.cbu.uib.no/) and obtained commercially from Sigma-Aldrich. The primer pair for the positive control gene, *Slc45a2,* was obtained from Edvardsen et al. (2014) (**Table 11**). The primer pairs were specifically designed to amplify within the target gene paralog, and the amplicon size was confirmed in Benchling (**Table 11**).

Table 11. *Designed PCR primers for Cuzd1 and Slc45a2.*

Name	Sequence 5' – 3'	Tm°	GC %	Annealing temp. (°C)	Amplicon size
(G1) Cuzd1 PCR F	GTTCACAGTCGGATTGTTTTGA	63.8	40.9	60	230 bp
(G1) Cuzd1 PCR R	GTGGGTGTTGTTAGAGAAGCGT	64.2	50		
(G2) Cuzd1 PCR F	CATTGTGGACGAGACGGTTC	65.5	55	60	163 bp
(G2) Cuzd1 PCR R	ACTGAACAGAGAAGATATCACCC	60.1	43.4		
(G3) Cuzd1 PCR F	GTTCACAGTCGGATTGTTTTGA	63.8	40.9	60	238 bp
(G3) Cuzd1 PCR R	CCAATGATGTGGGTGTTGTTAG	63.9	45.4		
**(G4) Slc45a2 PCR F*	TGCCACAGCCTCAGAATGTACA	67.9	50	67	371 bp
**(G4) Slc45a2 PCR R*	CAGATGTCCAGAGGCTGCTGCT	70.2	59		

* Primer pair obtained from Edvardsen et al. (2014).

2.8.1 DNA extraction from salmon gill tissue

DNA extraction was made from salmon gill tissue to investigate if the designed PCR primers amplified our target gene amplicon of interest. The tissue was homogenized with Fisherbrand ™ Bead Mill Homogenizer (aluminum beads) in Tissue Lyser II (Qiagen) at a frequency of 30/s for 1 min. The sample was centrifuged for 3 min at the highest speed (13 000 rpm). The DNA extraction was done with AllPrep DNA/RNA Mini Kit (Qiagen) following the manufacturer's protocol. The DNA concentration was measured with NanoDrop Spectrophotometer to 122.6 ng/µl.

2.8.2 PCR primer testing

The PCR was conducted using the Phusion® DNA Polymerase protocol following the manufacturer's protocol. The PCR consisted of initial denaturation (98 °C, 30 s), denaturation (98 °C, 10 s), graded annealing temperatures (60-67 °C, 15 s), extension (72 °C, 20 s), and cooling (4 °C). The denaturation/annealing/extension cycle was repeated for 40 cycles. The PCR products were analyzed on a 2% agarose gel to confirm the expected amplicon size (**Table 11**). The 1kB ladder (NEB) and 100 bp ladder (NEB) were used as references.

2.9 Analysis of mutations

DNA was extracted from our highest-ranked KO fish from each group (G1-G4) using AllPrep DNA/RNA Mini Kit (Qiagen) following the manufacturer's protocol. PCR was performed on the genomic DNA to obtain the targeted amplicon site (PCR primers in **Table 11**) with the same thermal cycling conditions as described in section 2.8.2. PCR products were analyzed on a 2% agarose gel to confirm the expected amplicon size. Thereafter, the PCR products were cleaned up with the QIAqiuck® PCR Purification Kit (Qiagen) following the manufacturer's protocol. The amplified DNA concentration was measured using NanoDrop Spectrophotometer. The cleaned-up PCR products were subsequently cloned into the TOPO-vector using the Zero Blunt® TOPO® PCR cloning kit (ThermoFisher) following the manufacturer's protocol. The ligation product was transformed into NEB 5-alpha competent *E. coli* cells following the transformation process described in section **Error! Reference source not found.**. Agar plates were made as described in the same section, except Kanamycin (50 mg/µl) was used instead of Ampicillin. The products from the transformation were purified using QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's protocol. PCR and gel electrophoresis were performed on the plasmid DNA to confirm the insertion of the target sequence. Thereafter, the samples were prepared for sequencing using Big Dye ™ Terminator v3.1 Cycle sequencing kit (ThermoFisher). A total of 35 samples were prepared for sequencing. The PCR consisted of initial denaturation (96 °C, 5 min), denaturation (96 °C, 10 s), annealing (50 °C, 5 s), extension (60 °C, 4 min), and cooling (4 °C, ∞). The denaturation/annealing/extension cycle was repeated 40 times. The PCR products were sequenced in the sequencing lab at the Medicinal Genetics Department (University Hospital in Northern Norway HF).

2.9.1 Data analysis

The gene expression data from the ontogeny analysis was not normally distributed and thus needed to be log-transformed. A one-way ANOVA with a posthoc test (Tukey) was performed on the log-transformed ontogeny data. Statistical analyses for the tissue distribution during smoltification were performed using an unpaired t-test assuming unequal variance. All statistical analysis was performed in R using the base package (R Core Team, 2023). All plots were created using ggplot2 package in R. Sequences of KO fish were analyzed in Benchling.

3 Results

3.1 Stage 1: RNA expression of *Cuzd1* through ontogeny and during smoltification

To date, our knowledge of *Cuzd1's* role in Atlantic salmon is limited in the parr-smolt transition in the gill. To better understand the potential involvement during development and in different tissues, we measured the expression of *Cuzd1* RNA in 1) whole eggs and fish during early development and; 2) a suite of tissues taken during the parr-smolt transition.

3.1.1 Developmental record

The salmon groups were sampled and photographed at specific developmental stages, as **Figure 10** shows. We aimed to observe how wild-type and KO salmon embryos develop and capture images of their morphological changes over time. The end of gastrulation (147 DG) was an essential stage in development to record, as it symbolizes the formation of the three germ layers developing the salmon tissues. In the organ development stage (250 DG), the heart, liver, and digestive systems are forming and are critical to the organism's growth. Sampling at that stage provided insights into the formation and maturation of the control and KO fish. The eyed stage (328 DG) is where the eyes become clearly visible due to increased pigmentation. This is where we expected to gain information about the pigmentation patterns in the KO fish. When the fish hatched at 491 DG, they were in the initial stages of development. The fish grew rapidly during this period as they adapted to their new environment outside the egg. As the fish reached 610 DG, we characterized them to be in the fin development stage. Fin development and maturation are important aspects of locomotion and stability in the water, so sampling at this stage provided useful information. 701 DG marked the midpoint in the development of the fish. The fish were sampled at this stage to assess their overall developmental progress and pigmentation. A major milestone in the development of the fish and this project was reached when the fish were at 811 DG, which was approaching external feeding. The salmon was transitioning from feeding on yolk sacs to external food sources at this stage. We gained insight into the phenotype, development and maturation, and overall health and survival during this last sampling point.

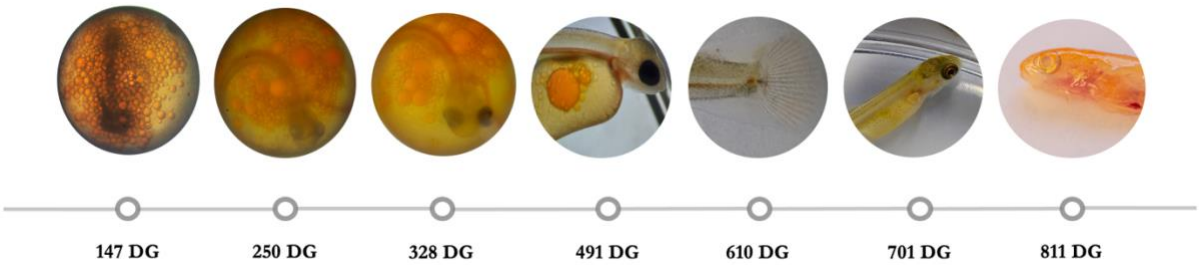

Figure 10*. Timeline of sampling points at specific day degrees (DG) and development stages. Timeline of sampling points at specific DG and development stages. 147 DG: 100 % epiboly (end of gastrulation), 250 DG: Organ development, 328 DG: Eyed stage, 491 DG: Hatching, 610 DG: Fin development, 701 DG: Mid-point sampling and 811 DG: Approaching external feeding.*

3.1.2 RNA expression of *Cuzd1* through ontogeny

The relative fold change gene expression of *Cuzd1* was calculated using *EF1a* as a reference gene and organ development stage as a calibrator. The expression of *Cuzd1* demonstrated an increasing pattern of expression over time. **Figure 11** presents the relative fold change for various developmental stages in WT salmon. The fold change increased with time at the point of approaching external feeding over 140 times compared to the organ development stage. Variation amongst individuals was quite low, as reflected in the error bar size. A one-way ANOVA was performed to test whether there was a significant difference in the expression level of *Cuzd1* across different developmental stages. The results showed a significant effect of developmental stage on gene expression ($F(5, 12) = 34.23$, $p < 0.001$) (see Appendix F). The *p-value* indicates that the null hypothesis of equal means could be rejected, showing that at least one mean is different. The result from the posthoc analysis showed that gene expression levels were significantly higher in the approaching external feeding stage compared to the eyed stage ($p = < 0.0001$) and hatching ($p = 0.008$). Fin development had significantly higher gene expression levels compared to the eyed stage ($p = 0.001$). Mid-point sampling had significantly higher gene expression levels compared to the eyed stage ($p = 0.001$). Organ development had significantly lower gene expression levels compared to the eyed stage ($p = 0.04$), hatching ($p = 0.0003$), fin development ($p = < 0.0001$), mid-point sampling ($p = < 0.0001$), and approaching external feeding ($p = < 0.0001$).

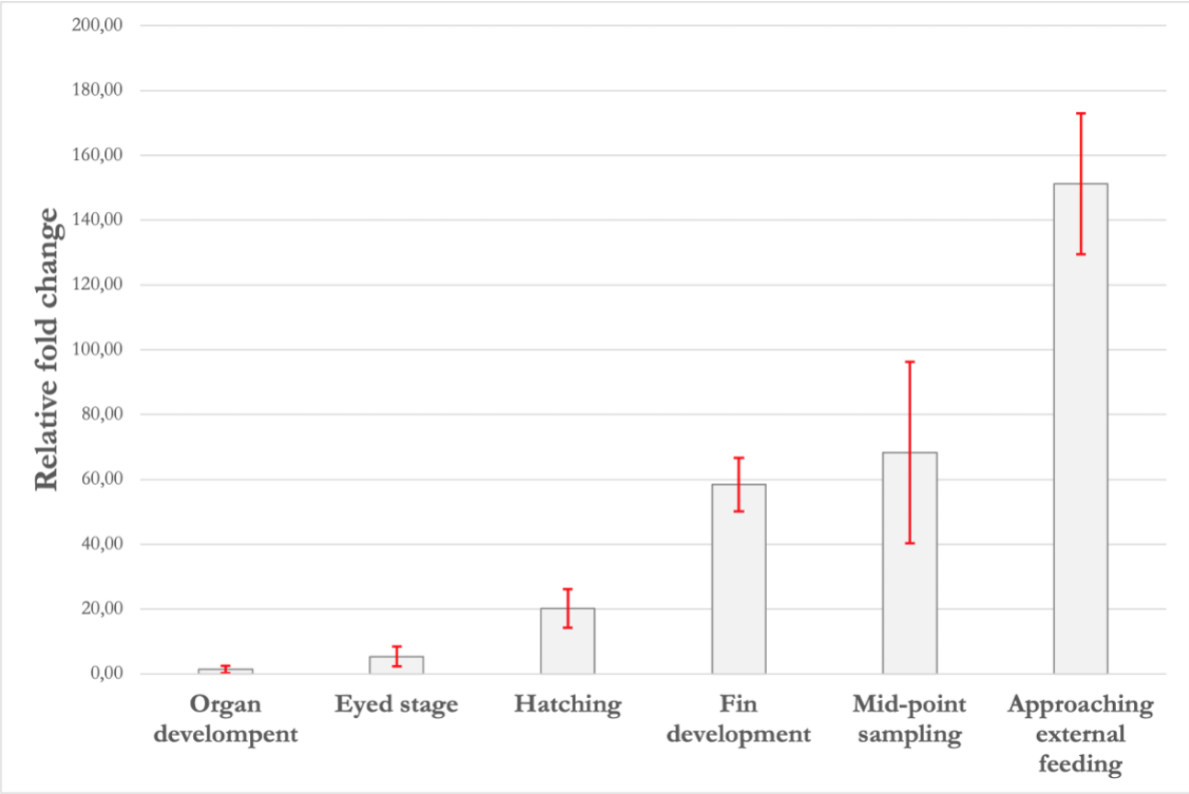

Figure 11. *Relative fold gene expression of Cuzd1 through ontogeny in absolute values (all other stages relative to organ development). A significant difference in expression was found between organ development and eyed stage (p = 0.04), hatching (p = 0.0003), fin development (p = < 0.0001), mid-point sampling (p = < 0.0001), and approaching external feeding (p = < 0.0001). The error bars represent the standard deviation of the data.*

3.1.3 Multi-tissue RNA expression of *Cuzd1* during smoltification

Gill, brain, heart, muscle, pancreas, fin, and spleen samples were collected from representative parr and smolt Atlantic salmon (n = 4). The relative fold gene expression of *Cuzd1* for tissue distribution in parr and smolt was calculated using *EF1a* as a reference gene. The expression of *Cuzd1* demonstrated a variable pattern in tissue distribution. **Figure 12** presents the relative fold change for various tissues in parr and smolt. The largest differences in *Cuzd1* expression were observed in the muscle, gill, and brain with *p*-values of 0.02, 0.04, and 0.02, respectively. There was a clear difference in the expression of *Cuzd1* in the spleen. However, the statistical analysis resulted in a non-significant *p*-value (0.05) (see Appendix F).

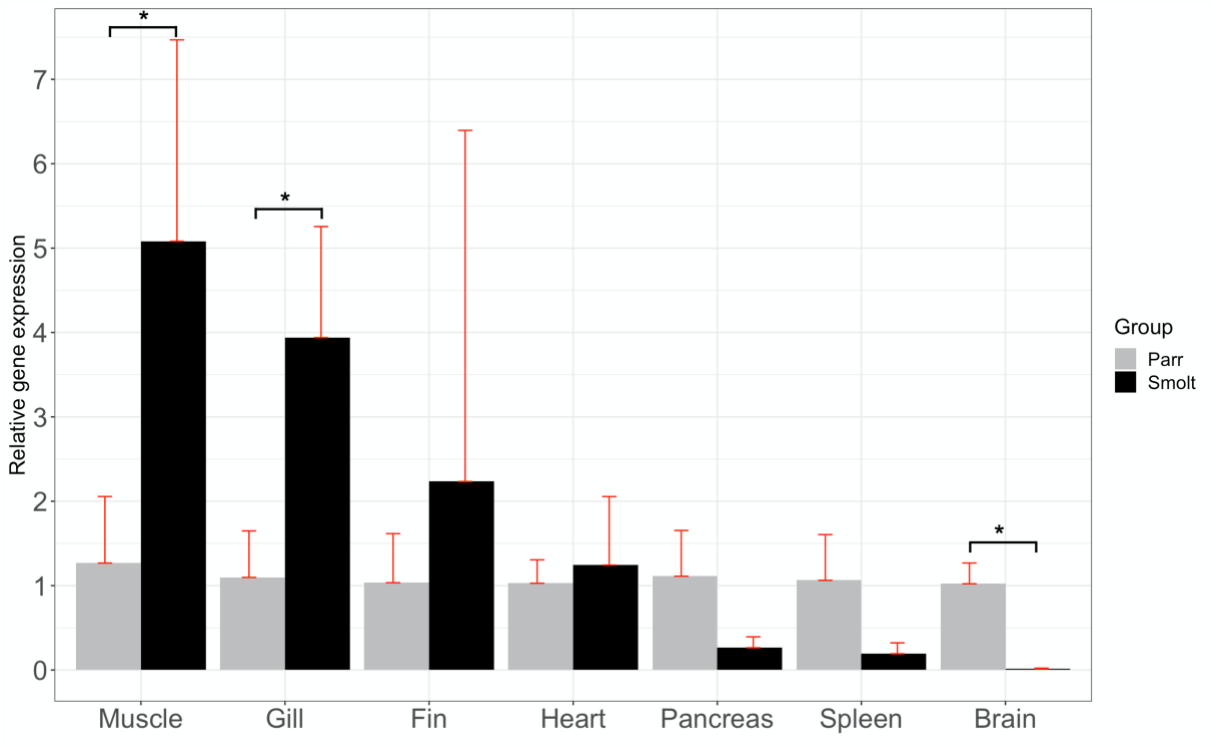

Figure 12. *Relative fold gene expression (smolt relative to parr) of Cuzd1 in parr and smolt tissues (n=3 for brain, heart, fin, pancreas, and spleen; n=4 for gill and muscle). A significant change in expression was found in the muscle (p = 0.02), gill (p = 0.04), and brain (p = 0.02). The error bars represent the standard deviation of the data.*

3.2 Stage 2: gRNA and Cas9 mRNA design, production, and RNA quality

We used the bioinformatics analysis tool (ChopChop) to test gRNAs for specificity to the target site in order to avoid off-target effects as best as possible. Designing specific gRNAs was challenging due to the near overlap between the two *Cuzd1* paralogs (clustal alignment in **Figure S 3**). The gRNA target sequences cloned into the pT7-gRNA vector were confirmed through sequence alignment (Benchling, Biology software, https://benchling.com, 2022). The GC content for each target sequence (**Table 1**) was: G1 (45 %), G2 (45 %), G3 (55 %), and G4 (60 %). The gRNAs were designed upstream of a PAM sequence so that Cas9 would recognize the cut site. **Figure 13** presents the target sequences for *Cuzd1*: G1 was designed within exon 5 with the cut-site in the exon, G2 was designed on the 7th exon boundary with the cut-site within the intron and G3 on the 5th exon boundary with the cut-site within the exon. Thorough characterization of CRISPR mutations shows that many indels are wider than 5 bp. Therefore, we reconciled that G2 might cause a mutation within the exon despite the cut site location being

in the adjacent intron (Shen et al., 2018). **Figure 14** presents a schematic overview of the target and PAM sequence for *Slc45a2* within exon 6 (target sequence obtained from Edvardsen et al. (2014)).

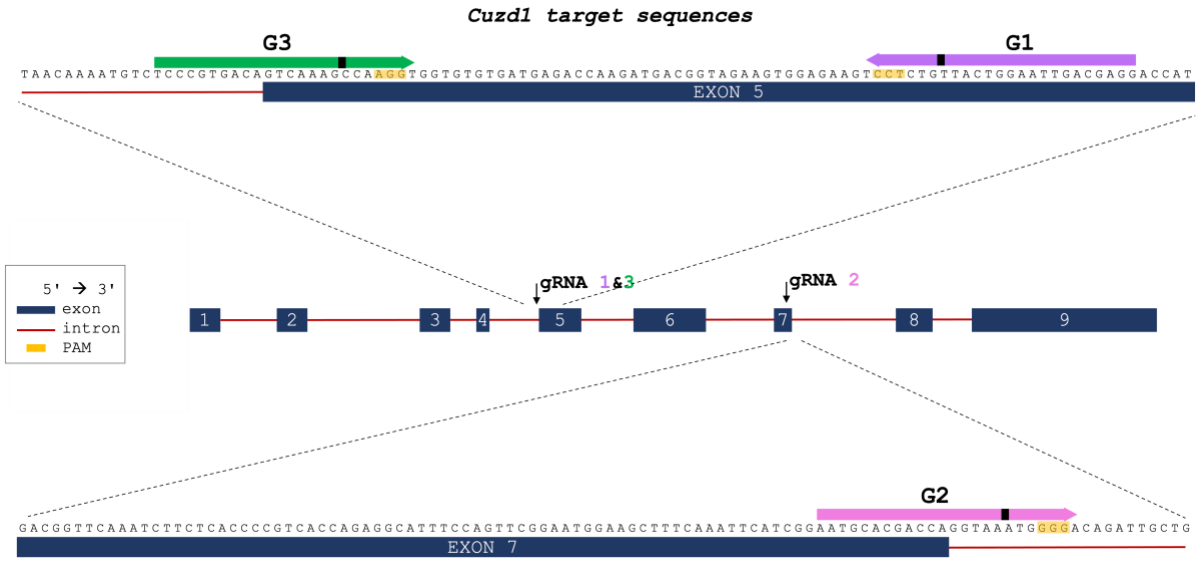

***Figure 13**. Target sequences for Cuzd1 gRNAs. G1 (purple) within exon 5, G3 (green) at the 5th exon boundary, and G2 (pink) at the 7th exon boundary. PAM sequences for each guide is marked in orange.*

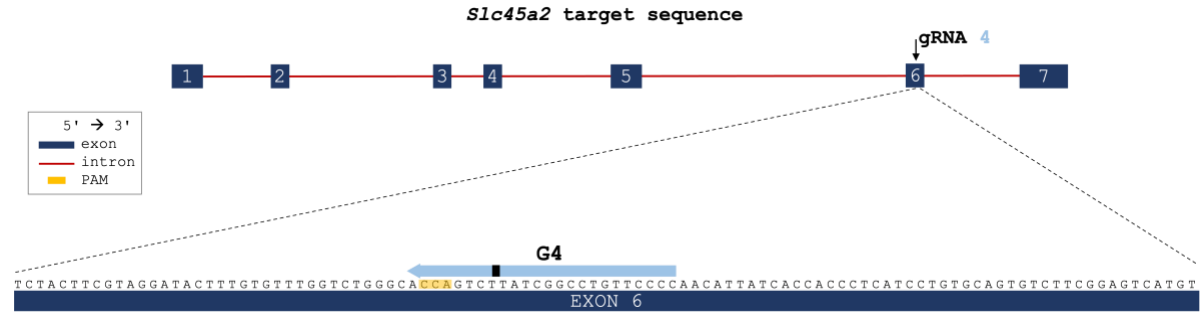

***Figure 14**. Target sequence for Slc45a2 gRNA G4 within the exon. PAM sequence marked in orange.*

3.2.1 Analysis of RNA quality

Examination of the electropherograms (**Figure 15**) from the Bioanalyzer Small RNA analysis and the TapeStation analysis showed clear single peaks corresponding to the expected size of the gRNAs and the Cas9 mRNA, ~95 nt and ~5000 nt, respectively. These results confirm the high purity of the *in vitro* transcribed RNAs.

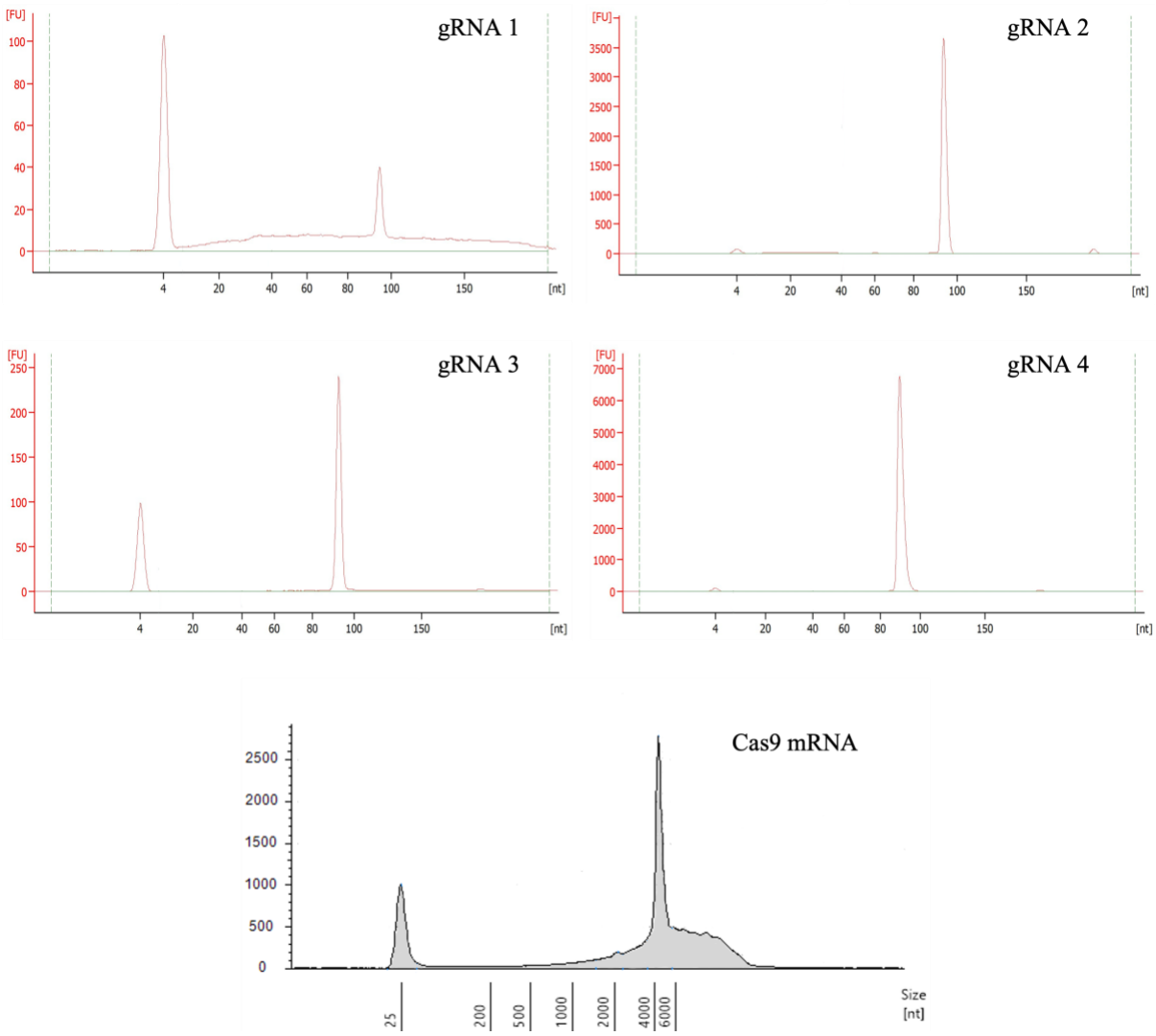

Figure 15*. Electropherograms for gRNA 1, gRNA2, gRNA 3, and gRNA 4 from the Bioanalyzer readout and Cas9 mRNA from the TapeStation readout. The peak on the right shows RNA fragments in size ~95 nt for gRNA 1-4 and ~5000 nt for the Cas9 mRNA. The peak on the left at ~4 nt and ~25 nt is the lower marker (internal control).*

3.2.2 Mortality records

Salmon embryos successfully developed and hatched without a high mortality rate. In the KO groups, embryonic development and hatching were successfully completed, which could suggest that *Cuzd1* KO is not a lethal mutation. Here we evaluated the mortality of salmon embryos injected with different gene KO constructs. A total of 245, 224, 240, and 241 eggs were injected with *Cuzd1-Slc45a2* (G1), *Cuzd1-Slc45a2* (G2), *Cuzd1-Slc45a2* (G3), and *Slc45a2* (G4) constructs, respectively. The control group consisted of 3000 non-injected eggs (**Table 12**). The results show that the mortality was lowest in group G4 (5.4 %), followed by the G3 group (7.9 %), and groups G1 and G2 (9.4 % and 10.7 %). The data is also presented in survival rate over time in **Figure 16**. Based on these results, *Slc45a2* KO alone is suggested to have a minimal effect on embryonic mortality, while KO of *Cuzd1* and *Slc45a2* may contribute to a slight increase in mortality. However, all experimental groups had lower mortality rates than the control group (14.5%), suggesting that overall embryonic survival was not negatively affected by the microinjections gene KO constructs. These findings provide important insight into these genes' role in salmon embryonic development and emphasize that mortality rates should be considered when evaluating gene KO constructs.

***Table 12**. Data presenting the number of eggs/fish in each group from the start and their mortality in percentage.*

Group	**Gene** KO	**No. of eggs/fish**	**Mortality %**
G1	*Cuzd1 + Slc45a2*	245	9.4
G2	*Cuzd1 + Slc45a2*	224	10.7
G3	*Cuzd1 + Slc45a2*	240	7.9
G4	*Slc45a2*	241	5.4
Control	-	3000	14.5

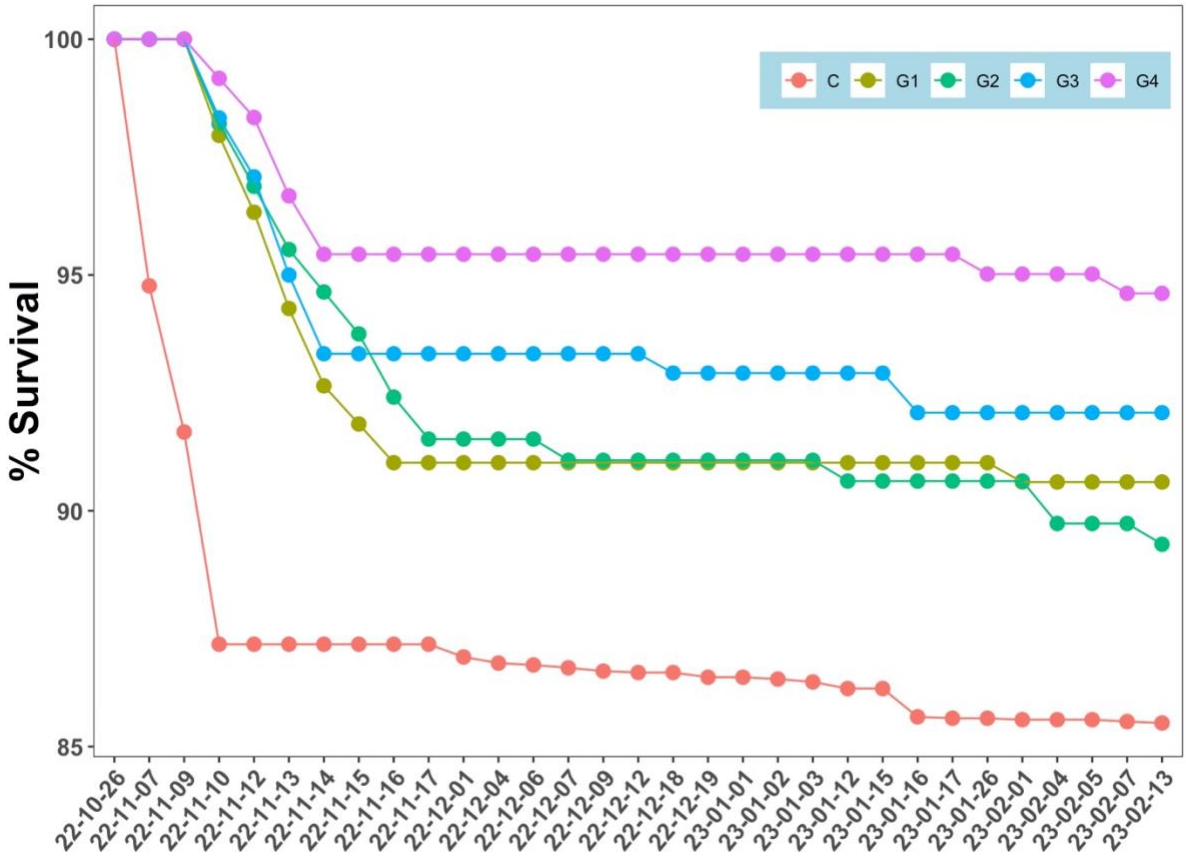

***Figure 16**. Survival rate of each group. Control group (red): 85.5 %. G1 (moss green): 90.6 %. G2 (green): 89.3 %. G3 (blue): 92.1 %. G4 (pink): 94.6 %.*

3.2.3 Visual identification of CRISPR*Slc45a2*Cas9 mutations

Fish from KO groups were visually screened to identify those showing total and partial pigmentation loss. These fish were then collected in RNA later for subsequence genetic analysis (**Table 13**). Examples of KO fish were placed in the order of gradually increasing pigmentation and shown in **Figure 17** (left column; **Figure S 5**). The difference between the KO and control fish in their appearance is prominent. **Figure 17** presents the fish from KO group G1 (left column), where the first three appear to have a complete KO of *Slc45a2*, thus total pigmentation loss, while the rest of the KO fish demonstrated partial pigmentation. In contrast, the appearance of WT fish is homogenously pigmented in all individuals. The groups with the least knockouts were G2 and G3. The groups that displayed only partial pigmentation were G3 and G4, and the group with no visible KO was G2 (**Table 13**).

***Table 13**. Number and percentage of KO fish in each group based on visual characterization.*

Group	Gene KO	No of eggs	No. KO	% KO
G1	*Cuzd1 + Slc45a2*	245	9	3.6
G2	*Cuzd1 + Slc45a2*	224	0	0
G3	*Cuzd1 + Slc45a2*	240	3	1.3
G4	*Slc45a2*	241	5	2.1
Control	-	3000	-	-

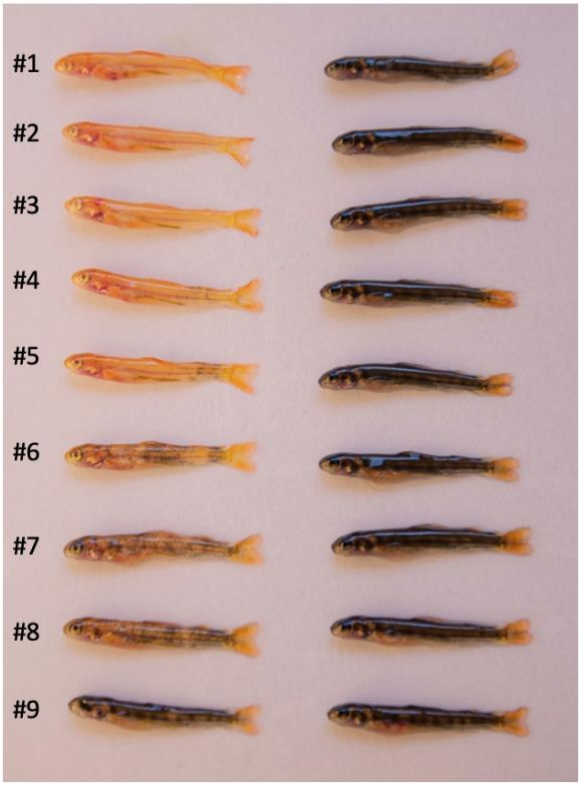

***Figure 17**. Photograph of group G1 (left column) with 3.6 % KO, placed in the order of gradually increasing pigmentation and wild-type group with intact pigmentation (right column).*

3.2.4 Genotyping of *Slc45a2* and *Cuzd1* knockouts

We performed a target sequence analysis on the highest-ranked fish from each group (G1-G4) to verify the knockouts and validate the accuracy of the phenotype's visual characterization. For the *Slc45a2* gene, 20 clones were sequenced in total from all groups. For the *Cuzd1* gene, 15 clones were sequenced in total from groups G1, G2 and G3 (**Table 14**).

***Table 14**. Number of plasmids sequenced for Cuzd1 and Slc45a2*

Group	**Gene** KO	**No. of plasmids sequenced (*Cuzd1*)**	**No. of plasmids sequenced (*Slc45a2*)**
G1	*Cuzd1 + Slc45a2*	5	5
G2	*Cuzd1 + Slc45a2*	5	5
G3	*Cuzd1 + Slc452*	5	5
G4	*Slc45a2*	-	5

3.2.4.1 Identification of CRISPR*Slc45a2*Cas9 mutations

Group G1 displayed 21 deletions in one sequence, suggesting a partial KO of *Slc45a2*. DNA sequences from G2 showed total homology (**Figure 18**) to the *Slc45a2* reference genome. The lack of indels or mismatches in the sequences suggests low or absence of KO of *Slc45a2* in this group, which corresponds with the visual scoring (**Figure S 5**). Group G3 displayed 5 deletions and 21 deletions in two sequences, suggesting a partial KO of *Slc45a2*. The positive control group G4 displayed 26 deletions in one of the sequences with an expected Cas9 cut near the PAM site. This result confirmed the phenotype from the visual scoring.

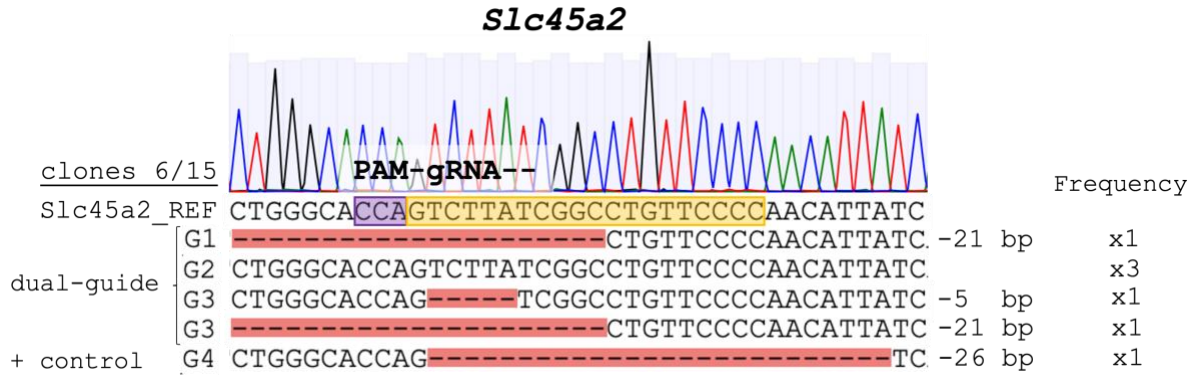

***Figure 18**. Validation of Atlantic salmon Slc45a2 mutations. Individual sequence analysis of whole KO fish for G1, G2, G3, and G4 compared to the reference gene Slc45a2. PAM-site is colored purple, and the target sequence is orange. Various deletions were validated for G3 and G4. No mutations in G2.*

3.2.4.2 Identification of CRISPR*Cuzd1*Cas9 mutations

Sequence data for the *Cuzd1* locus for the G2 group displayed a mismatch for 4 sequences and homology for one sequence with the *Cuzd1* reference genome (**Figure 19**). Comparison of the mismatched sequences to the Atlantic salmon genome revealed that homology with the second (non-smoltification related) *Cuzd1* paralog (LOC106587759) (**Figure S 4**), meaning that the original PCR reaction was not specific for the *Cuzd1* target gene and amplified both *Cuzd1* paralogs. Group G3 showed the same pattern from the sequencing, also with an off-target effect for two sequences, and one sequence showed complete homology with the target sequence.

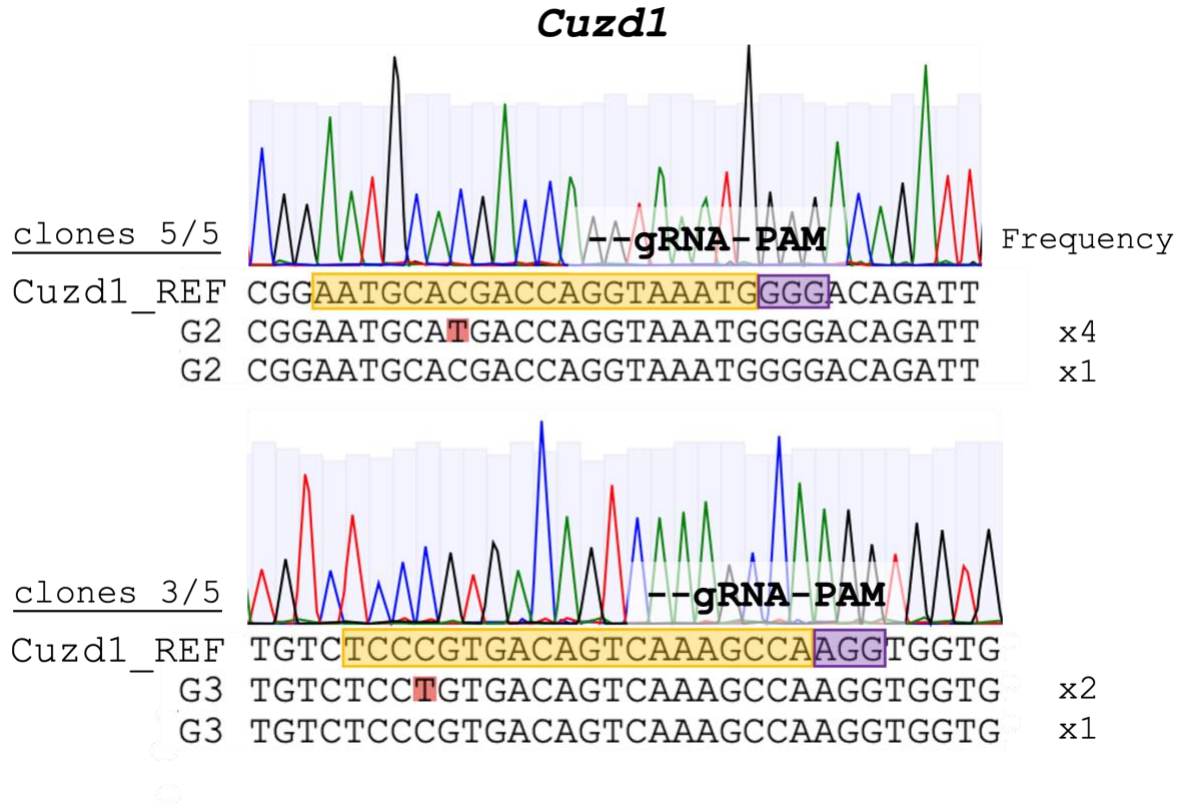

***Figure 19**. Validation of Atlantic salmon Cuzd1 mutations. Sequencing results for G2 show one mismatch for four of the sequences and no mutation in one sequence. Sequencing results for G3 show one mismatch for two sequences and one with no mutation.*

4 Discussion

The Atlantic salmon is highly commercially important and has a fascinating life history. It is of utmost importance to understand the mechanism of its adaptation to the environment. In particular, exploring the function of genes that might be responsible for the adaptation might provide fundamental knowledge and new opportunities to study the salmon genome. The classical smolt markers, NKA subunits, and CFTR have been shown to be expressed in the gill (Houde, Akbarzadeh, et al., 2019; Houde, Gunther, et al., 2019; Iversen et al., 2020; West et al., 2021). These markers had expected expression patterns where they are induced at LP and suppressed at SP (Houde, Akbarzadeh, et al., 2019; Houde, Gunther, et al., 2019). *Cuzd1*, in contrast, follows SP exposure and shows winter dependency; it has a high expression and high fold change in the gill and is expressed in a single-cell type with little characterization (West et al., 2021). Existing information on *Cuzd1* expression in Atlantic salmon is sparse. Unraveling its function through the ontogeny and distribution of its expression and conducting knockouts is important for understanding gill remodeling in smoltification. The overall aim of this project was twofold: To investigate the regulation of *Cuzd1* through ontogeny and smoltification and establish *in vivo* CRISPR methodology in Atlantic salmon.

4.1 Stage 1: RNA expression of *Cuzd1*

The known expression of *Cuzd1* is limited. Evidence in mammals shows that *Cuzd1* plays a key role in several tissues (Huynh et al., 2001; Leong et al., 2004; Leong et al., 2007; Liaskos et al., 2013; Mapes et al., 2018). Until now, *Cuzd1* has a known and high expression in humans in the exocrine pancreas and has been found to act as a marker of ovarian cancer (Leong et al., 2007). Moreover, the rat ortholog (UO-44) was found to be expressed in the ovaries and uterus (Huynh et al., 2001; Leong et al., 2004). Mapes et al. (2018) recorded data showing *Cuzd1* is expressed in mice's mammary ductal and alveolar epithelium. The characterization of *Cuzd1* is limited in Atlantic salmon and teleosts in general. Only one publication (West et al., 2021) reported *Cuzd1* expression in the Atlantic salmon. However, the research was conducted in the gill but not in other tissues. Thus, there is a need to investigate the expression of *Cuzd1* in multiple tissues during smoltification and throughout development.

4.1.1 RNA expression of *Cuzd1* through ontogeny

The expression of *Cuzd1* was investigated during early developmental stages in Atlantic salmon. According to our findings, the expression increased in whole embryos/fish over time (**Figure 11**). There was a significant difference between the organ development stage and eyed stage, hatching, fin development, mid-point sampling, and approaching external feeding. Many genes do not appear until later in development, although *Cuzd1* should be expressed fairly early if it is related to the gills. Although we did not investigate *Cuzd1* expression at the tissue level, it would be interesting to explore the spatial distribution of *Cuzd1* expression in early developmental stages, especially in relation to the gills. Since *Cuzd1* is believed to be involved in gill development and function, our results may suggest a potential role of *Cuzd1* in early gill development. Expression of *Cuzd1* during early developmental stages could shed light on its role in the formation of various tissues, and it provides valuable insight into the gene's developmental patterns.

4.1.2 Multi-tissue RNA expression of *Cuzd1* during smoltification

The expression of *Cuzd1* was found in the muscle, gill, fin, heart, pancreas, spleen, and brain (**Figure 12**). It was compared between parr and smolt with statistically significant differences for the muscle, brain, and gill. The expression of *Cuzd1* was enhanced during smoltification in the muscle, thus suggesting that *Cuzd1* could have a potential role in the growth during smoltification. Despite the lack of statistical significance in the difference of *Cuzd1* expression levels in the spleen between parr and smolt, a noticeable trend towards a decreased expression of *Cuzd1* during smoltification was observed. The decreased expression levels in the smolt spleen could suggest that *Cudz1* has a potential role in immune and disease function. One interesting finding was the suppression of *Cuzd1* in the smolt brain. To the best of our knowledge, there are no studies about the expression of *Cuzd1* in the brain in any organism. It can be interpreted as *Cuzd1* being involved in the neuroendocrine regulation of the smoltification process and needs to be downregulated in the brain, considering that Mapes et al. (2018) found *Cuzd1* being involved in mammary gland development and lactation. It is possible to speculate that *Cuzd1* may be involved in the signaling pathway downstream of prolactin. Since prolactin is downregulated during smoltification (Bernard et al., 2020; McCormick, 2013), this could be interesting to investigate further. A possible future considerations could be the protein characterization of *Cuzd1* using immunohistochemistry (IHC) on gill tissues at different points during smoltification. Using bromodeoxyuridine (BrdU)

staining (Mapes et al., 2018) to investigate the differences in genes involved in cell proliferation could tell us if *Cuzd1* is associated with the proliferation of the gill epithelium cells. An interesting future perspective would be to conduct BrdU staining in the gill of Atlantic salmon during smoltification to determine whether it co-localizes with the cells expressing *Cuzd1*. Since BrdU is a method where an alternative base can be incorporated into the DNA, we could examine which cells have divided after a period of time (Mapes et al., 2018). Consequently, through the prolactin JAK/STAT signaling pathway, we could investigate if *Cuzd1* is controlling cellular proliferation. To fully understand the significant finding of suppression of *Cuzd1*, further studies could investigate different brain regions or signaling pathways in the brain. Our analyses suggest that *Cuzd1* expression exhibits tissue-specific patterns, as evidenced by the differences in expression levels across various tissue types. We speculate that this tissue-specificity may be influenced by changing hormone profiles, as exogenous treatment with growth hormone (GH) and cortisol (McCormick, 2001) may drive a "smolt-like" change in *Cuzd1* expression. GH and cortisol, which act synergistically during the adaptation to seawater in regulating NKA activity, may also play a role in driving *Cuzd1* expression.

The significant difference in gene expression of *Cuzd1* in the gill could shed light on the mechanisms involved in the physiological changes during smoltification. The relative fold gene expression during smoltification provides an important insight into *Cuzd1* tissue-specific patterns. Further research is warranted to understand the function and regulation of *Cuzd1* fully.

4.2 Stage 2: Methodology and analysis of knockout fish

To investigate the function and regulation of *Cuzd1*, we used *in vivo* CRISPR technology to target *Cuzd1* and *Slc45a2*. This study establishes the technique at UiT for the first time, and for its implementation, it was required to design and clone gRNAs for *Cuzd1* and *Slc45a2* and produce high-quality *in vitro* transcribed gRNAs and Cas9 mRNA. The Cas9 mRNA and the gRNAs were injected into the single cell-stage embryos of Atlantic salmon. The results from the CRISPR knockouts were confirmed by phenotyping and genotyping.

4.2.1 Challenges of *Cuzd1* gRNA design

As outlined above, this project aimed to develop *in vivo* CRISPR tools for genome editing in Atlantic salmon. Crucial to the CRISPR-Cas9 system was designing specific gRNAs targeting our gene of interest, *Cuzd1*.

In the process of designing our gRNA for genome editing, we had two imperatives: 1) specificity, 2) efficiency. Obtaining a gene-specific gRNA can be challenging, particularly when using computational tools to predict potential target sites. In this study, we faced this difficulty as all of the guides identified by ChopChop had a low uniqueness score because of the high similarity between the *Cuzd1* paralogs. This limited our options as the guides could bind with no mismatches, potentially leading to off-target effects, or contain one or two mismatches, reducing their specificity for the target site. To address this challenge, we carefully evaluated the potential target sites and selected the ones we believed would be unique for our gene of interest (**Figure 13**; **Figure 14**). However, due to the limited availability of gene-specific guides, we had to use gRNAs with large overlap (**Figure S 3**) with other genomic regions or cut sites outside the exons. Nonetheless, we believe that our choices helped us to maximize the efficiency of Cas9 cutting at the right site. To achieve this, a PAM sequence was incorporated intended to enhance binding between the gRNA and Cas9. The second imperative of designing a gRNA was to avoid secondary structures. We tried to avoid target site segments with a high GC content or other structural motifs that might interfere with the gRNA-Cas9 binding. One potential approach for future studies is to utilize promiscuous gRNAs that can target both paralogs of *Cuzd1* simultaneously.

The specificity of the CRISPR/Cas9 system can be affected by the number of mismatches in the gRNA sequence, resulting in an off-target effect (**Table 15**). Even a single mismatch can lead to an unintended cut in the genome. However, not all mismatches are detrimental to the specificity. In the case of designing our gRNAs through ChopChop, the online tool is designed to minimize off-target effects, and by default, it searches a little bit outside the exons. This accounts for wider deletions that could overlap with the exon (Shen et al., 2018). Although these precautions were considered, it was not possible to completely eliminate off-target effects.

Table 15*. Possible mismatches (marked in red) for the designed gRNAs G1, G2, and G3.*

Guide	Location	Sequence (with mismatches in red)	No. of mismatches
G1	NC_027324.1:36806886	CTC**ac**CAAT**a**CCAGTAACAGGGG	3
G1	NC_027325.1:36158026	CCTCTGTTACTGGAATT**c**ACGAG	1
G2	NC_027325.1:36159307	AATGCA**t**GACCAGGTAAATGGGG	1
G3	NC_027325.1:36157961	TCC**t**GTGACAGTCAAAGCCAAGG	1

Since Atlantic salmon has been through four whole genome duplications (Macqueen & Johnston, 2014), many similar paralogs are present. This is a significant challenge in designing gRNAs with high specificity towards a certain paralog. While our efforts to design specific and efficient gRNAs were aimed to minimize off-target effects, an alternative approach could be to design guides that do not discriminate between the *Cuzd1* paralogs. This could involve selecting gRNAs that target both *Cuzd1* paralogs, potentially resulting in a double knockout. Although this compromise may not achieve the desired outcome of specifically targeting one gene, it could provide a viable alternative to using suboptimal guides for a single target. Further studies could explore this option and compare the efficacy and specificity of double versus single gene targeting for genome editing purposes.

4.2.2 Cas9 mRNA and gRNA delivery

In selecting the appropriate concentration of Cas9 mRNA for our experiment, we followed established protocols based on the best practices of previous studies. Specifically, we referred to the work of Edvardsen et al. (2014) and Jao et al. (2013), which have demonstrated relatively high success rates using similar methods. By relying on established protocols, we aimed to increase the reproducibility and reliability of our results and ensure that our experimental conditions were consistent with prior research in the field. The Cas9 mRNA could have been injected at a higher concentration, but there was a limit to how high the concentration could be for the injections. Li et al. (2014) investigated the optimal concentrations of gRNA and Cas9 mRNA delivery. The study showed that the optimal concentration for delivery was 50/500 ng/µl of gRNA and Cas9 mRNA. This resulted in the optimal mutation rate, although it resulted in a lower survival rate than concentrations of 50/100 ng/µl for gRNA and Cas9 mRNA. As a result of injection procedures and the degree of knockout, we cannot rule out a concentration-dependent effect. Furthermore, we used a zebrafish codon-optimized Cas9-system based on a previous study (Jao et al., 2013). What can be done in the future is to re-codon optimize the Cas9-system for Atlantic salmon. Nevertheless, there are ongoing efforts in the labs at UiT to develop a cold-adapted Cas9.

4.2.3 Microinjections

The aim was to microinject the produced Cas9 mRNA and gRNAs into the single-cell stage of Atlantic salmon embryos. The injections were performed in a cold room (7 °C), which is favorable when working with cold-water marine species. The injection procedure could be a

plausible explanation for the low success rate in knockouts. Since freshly fertilized salmon eggs are darker in color and the chorion is hard, injecting directly into the first cell is challenging. Although, we took precautions and fertilized the eggs in glutathione to keep the chorion from hardening. Moreover, the non-trained eye of the operator could also be an explanation for the success rate of the microinjections and, consequently, the low rate of embryo mutations caused by failed injections. Although the microinjections could be a factor affecting the success rate, it could be more reflected in the poor efficiency of the CRISPR/Cas9 system (discussed in the previous section) in Atlantic Salmon and in marine cold-water species in general. This may have to do with both Cas9 mRNA translation, but also its cutting efficiency under low-temperature conditions since the Cas9 used in this project was adapted to 37 °C.

4.2.4 Monitoring development and mortality

All groups of salmon were sampled and photographed through ontogeny to observe the development over time. All stages displayed no macroscopic deformities during the development. However, these evaluations were not very detailed, and we could not exclude microscopic or non-visual/molecular phenotypes. Although, this provided insights into the formation and maturation, the morphological changes over time, and the survival of the fish.

Both control and injected salmon embryos exhibited 50% epiboly (halfway in gastrulation), thus indicating no arrest in the development (Gorodilov, 1996) at an early stage (10 days after fertilization).

4.2.5 Visual identification of CRISPR*Slc45a2*Cas9 mutants

All groups (G1-G4) (**Table 13**) were injected with the *Slc45a2* gRNA for potential visual screening. The order of the groups rated according to the acquired phenotype was graded as follows: G1 > G4 > G3 > G2 (**Figure 17**; **Figure S 5**). Visual characterization of *Slc45a2* mutants resulted in a phenotype scoring of 3.6 %, 0 %, 1.3 %, and 2.1% in G1, G2, G3, G4, respectively. The group with the least visible phenotypes was G2, which displayed the absence of the albino phenotype. We found that the overall visual scoring in our study was lower than observed in a previous study on Atlantic salmon (8.7%) (Edvardsen et al., 2014). Moreover,

there is no clear reason, based on the starting material, that would suggest why some of the groups were less successful in terms of the knockouts.

4.2.6 Genotyping of *Slc45a2* and *Cuzd1* knockouts

The aim was to characterize possible mutations created by the CRISPR/Cas9 system. Recent developments in CRISPR-Cas9 technology generated double allelic mutations in the F0 generation of species, including Xenopus, zebrafish, tilapia, and Atlantic salmon, enabling a relatively rapid study of functional studies (Edvardsen et al., 2014; Jao et al., 2013; Li et al., 2014; Straume et al., 2020; Wargelius et al., 2016). As previously described, CRISPR-Cas9-mediated mutations of pigmentation genes *Slc45a2* and *tyr* can produce fish with various degrees of mosaicism, demonstrating the degree of loss of function (Edvardsen et al., 2014).

4.2.6.1 *Slc45a2*

Injected embryos in G2 showed total homology (**Figure 18**) to the *Slc45a2* reference genome. This could be due to several reasons described in sections 4.2.1 and 4.2.2. Sequences from group G3 displayed 5 deletions and 21 deletions in two sequences, suggesting a partial KO of *Slc45a2*. G1 displayed the exact same 21 deletions as G3. There are different factors to consider why the sequences were different. The timing of the Cas9 cut can vary depending on the cell stage during embryonic development, which can result in mosaicism. Although we injected the Cas9 and the gRNAs at the single-cell stage, the actual cut may occur later in development, after the cells have divided, which is a potential cause of mosaicism. Even at the single-cell stage, the presence of two alleles implies that each of them has the potential to be independently targeted and cleaved by the CRISPR/Cas9 system, which may result in different indel patterns. When the RNA was extracted from whole fish, the different cells came from various parts of the fish, which can contribute to the heterogeneity in the expression. The positive control group G4 displayed 26 deletions in one of the sequences with an expected Cas9 cut near the PAM site. This result confirmed the phenotype from the visual scoring. A similar pattern in deletion variants of *Slc45a2* was previously observed by Edvardsen et al. (2014). We predicted that with the same gRNA in the same strain, we would get some similarities in the deletions, as in G3 and G4. However, we would need additional sequencing results for a more detailed comparison. In hindsight, it is imperative to commence the investigation with a Miseq analysis to achieve a comprehensive characterization of the experimental outcomes, particularly in terms of evaluating the efficacy of the guides. Further elaboration on this aspect is provided in the subsequent section.

4.2.6.2 Cuzd1

The sequencing results for G2 showed one sequence, which amplified our target *Cuzd1* paralog, and four sequences with one mismatch, as presented in (**Figure 19**; **Table 15**). This showed that this specific mismatch led to the amplification of the *Cuzd1* paralog. As outlined in section 4.2.1, the Atlantic salmon have many paralogs. A challenge that we faced was the presence of the non-target *Cuzd1* paralog. To avoid amplifying the paralog, we tried to design specific PCR primers that targeted our gene of interest. Our approach was to align the two paralog sequences (**Figure S 3**) and identify the regions where the primers could anneal specifically to the target gene. To ensure adequate coverage of our target regions, it was necessary to design additional primers for the gRNAs. However, due to the time constraints of the project, we were unable to pursue this approach within the given timeframe. Despite this limitation, we endeavored to maximize the efficiency and specificity of our experimental design within the scope of available resources.

There are two main challenges regarding off-target effects when using CRISPR technology. Firstly, ensuring the specificity of the gRNA for the intended target gene. Secondly, designing specific PCR primers for the target region. The immediate solution could involve the development of an alternative set of cloning primers. However, due to constraints of time within the scope of this project, this approach is not feasible. As outlined above, designing an appropriate target for Atlantic salmon is challenging due to the high number of paralogs. One strategy could have been to target both *Cuzd1* paralogs with a single guide that matches 100% to increase the success of knockout. However, in this study, we only chose to target one of the *Cuzd1* paralogs to investigate its potential functionalization. It was preferable from a technical point of view because the different paralogs may have distinct functions.

The same efficiency could not be expected for every KO fish because the gRNA needs access to the gene to guide the Cas9 where to cut. When working with the F0 generation, it means that the CRISPR components are introduced at a very early stage, in our case, in the single-cell stage. Injecting CRISPR components in F0 embryos can be challenging since several factors can affect the efficiency of the CRISPR/Cas9 system. These factors are the timing of the injections, potential off-target effects, and the mosaicism that can occur. Therefore, we aimed to plan and optimize the knockout experiment carefully to achieve a reasonable degree of knockout efficiency. We did not expect to get a 100% KO of either *Slc45a2* or *Cuzd1*, although we hoped it would still be possible to analyze and identify the CRISPR-edited fish through visual scoring and genotyping.

4.3 Limitations and future considerations

Two main challenges were encountered in the project: 1) low efficiency of the gRNAs and Cas9 mRNA, 2) sequence characterization. To address this, we discuss several possible options to improve the efficiency and obtain a more complete characterization of the sequences in the future. To address the first challenge, several options are possible to improve the efficiency of the system, including promiscuous gRNAs that can simultaneously target both *Cuzd1* paralogs. Optimization of the concentration and type of Cas9 used, and the microinjection procedures can be changed to increase the chances of generating KOs. Regarding the second challenge, we encountered difficulties characterizing the genetic changes. We used a cloning technique to assess the genetic KO, which was a cost-effective method but limited in terms of sample and clone numbers. We ended up targeting the non-smoltification related *Cuzd1* paralog, which could have affected the reliability of our results. In future studies, it may be beneficial to use more comprehensive methods for genetic characterization, such as Miseq sequencing. Miseq sequencing is an excellent option for characterizing the genetic effects of CRISPR-Cas9 editing, owing to its ability to provide extensive data coverage. Although Miseq analysis may require additional time and resources, its increased accuracy and comprehensiveness justify its use in certain experimental settings. In our study, we could have benefited from Miseq analysis to characterize further the effects of CRISPR-Cas9 editing on our target gene. In addition, by obtaining RNAseq and snRNAseq data, we could have investigated if *Cuzd1* is expressed in NDCs and compared that data to the findings of West et al. (2021).

5 Conclusions

This study expands our characterization of *Cuzd1* expression in Atlantic salmon and reports the first successful use of *in vivo* CRISPR knockout at UiT. We developed a characterization of *Cuzd1* expression through ontogeny and provided a muti-tissue perspective through smoltification – a developmental transition known to involve dynamic control of *Cuzd1* expression in the gill. Our evidence suggests that the RNA expression of *Cuzd1* increased during early development providing valuable insights into the gene's developmental patterns. We also measured a clear difference in relative gene expression of *Cuzd1* during smoltification in the muscle, brain, and gill – suggesting that the role of *Cuzd1* during smoltification is not limited to the gill. We also provide visual and genetic evidence for successful CRISPR targeting in Atlantic salmon *in vivo* for *Slc45a2*. While further genetic characterization is required to confirm the impact of the novel *Cuzd1* CRISPR guides, this work ultimately establishes an important technique within the research group.

References

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

Barrangou, F. C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science (American Association for the Advancement of Science)*, *315*(5819). https://doi.org/https://doi.org/10.1126/science.1138140

Berge, A. I., Berg, A., Fyhn, H. J., Barnung, T., Hansen, T., & Stefansson, S. O. (1995). Development of Salinity Tolerance in Underyearling Smolts of Atlantic Salmon (Salmo-Salar) Reared under Different Photoperiods. *Canadian Journal of Fisheries and Aquatic Sciences*, *52*(2), 243-251. https://doi.org/DOI 10.1139/f95-024

Bernard, B., Leguen, I., Mandiki, S. N. M., Cornet, V., Redivo, B., & Kestemont, P. (2020). Impact of temperature shift on gill physiology during smoltification of Atlantic salmon smolts (Salmo salar L.). *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology*, *244*. https://doi.org/ARTN 110685
10.1016/j.cbpa.2020.110685

Blitz, I. L., Biesinger, J., Xie, X., & Cho, K. W. (2013). Biallelic genome modification in F(0) Xenopus tropicalis embryos using the CRISPR/Cas system. *Genesis*, *51*(12), 827-834. https://doi.org/10.1002/dvg.22719

Brouns, J., M. M., Lundgren, M., R., W. E., J., S. R., P., S. A., J., D. M., S., M., V., K. E., & Van der Oost, J. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science (American Association for the Advancement of Science)*(321), 5891 https://doi.org/https://doi.org/10.1126/science.1159689

Chakrapani, V., Patra, S. K., Panda, R. P., Rasal, K. D., Jayasankar, P., & Barman, H. K. (2016). Establishing targeted carp TLR22 gene disruption via homologous recombination using CRISPR/Cas9. *Dev Comp Immunol*, *61*, 242-247. https://doi.org/10.1016/j.dci.2016.04.009

Chen, H., Choi, J., & Bailey, S. (2014). Cut site selection by the two nuclease domains of the Cas9 RNA-guided endonuclease. *J Biol Chem*, *289*(19), 13284-13294. https://doi.org/10.1074/jbc.M113.539726

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

Davis, A. E., Castranova, D., & Weinstein, B. M. (2021). Rapid Generation of Pigment Free, Immobile Zebrafish Embryos and Larvae in Any Genetic Background Using CRISPR-Cas9 dgRNPs. *Zebrafish*, *18*(4), 235-242. https://doi.org/10.1089/zeb.2021.0011

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602-607. https://doi.org/10.1038/nature09886

Dodson, J. J., Aubin-Horth, N., Theriault, V., & Paez, D. J. (2013). The evolutionary ecology of alternative migratory tactics in salmonid fishes. *Biological Reviews*, *88*(3), 602-625. https://doi.org/10.1111/brv.12019

Dong, Z., Ge, J., Li, K., Xu, Z., Liang, D., Li, J., Li, J., Jia, W., Li, Y., Dong, X., Cao, S., Wang, X., Pan, J., & Zhao, Q. (2011). Heritable targeted inactivation of myostatin gene in yellow catfish (Pelteobagrus fulvidraco) using engineered zinc finger nucleases. *PLoS One*, *6*(12), e28897. https://doi.org/10.1371/journal.pone.0028897

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

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

Ebbesson, L. O. E., Ebbesson, S. O. E., Nilsen, T. O., Stefansson, S. O., & Holmqvist, B. (2007). Exposure to continuous light disrupts retinal innervation of the preoptic nucleus during parr-smolt transformation in Atlantic salmon. *Aquaculture*, *273*(2-3), 345-349. https://doi.org/10.1016/j.aquaculture.2007.10.016

Edvardsen, R. B., Leininger, S., Kleppe, L., Skaftnesmo, K. O., & Wargelius, A. (2014). Targeted mutagenesis in Atlantic salmon (Salmo salar L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation. *PLoS One*, *9*(9), e108622. https://doi.org/10.1371/journal.pone.0108622

Evans, D. H., Piermarini, P. M., & Choe, K. P. (2005). The multifunctional fish gill: dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. *Physiol Rev*, *85*(1), 97-177. https://doi.org/10.1152/physrev.00050.2003

Fleming, I. A. (1996). Reproductive strategies of Atlantic salmon: Ecology and evolution. *Reviews in Fish Biology and Fisheries*, *6*(4), 379-416. https://doi.org/Doi 10.1007/Bf00164323

Fricke, Eschmeyer, & Fong. (2023). Eschmeyer's catalogue of fishes: Species by family/subfamily [Retrieved from http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp].

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A*, *109*(39), E2579-2586. https://doi.org/10.1073/pnas.1208507109

Gorodilov, Y. N. (1996). Description of the early ontogeny of the Atlantic salmon, Salmo salar, with a novel system of interval (state) identification. *Environmental biology of fishes*, *47*(2), 109–127. https://doi.org/https://doi.org/10.1007/BF00005034

Handeland, S. O., & Stefansson, S. O. (2001). Photoperiod control and influence of body size on off-season parr-smolt transformation and post-smolt growth. *Aquaculture*, *192*(2-4), 291-307. https://doi.org/Doi 10.1016/S0044-8486(00)00457-9

Hoar, W. S., & Randall, D. J. (1988). *The physiology of smolting salmonids* (Vol. 14). Academic Press.

Houde, A. L. S., Akbarzadeh, A., Gunther, O. P., Li, S. R., Patterson, D. A., Farrell, A. P., Hinch, S. G., & Miller, K. M. (2019). Salmonid gene expression biomarkers indicative of physiological responses to changes in salinity and temperature, but not dissolved oxygen. *Journal of Experimental Biology*, *222*(13). https://doi.org/ARTN jeb198036 10.1242/jeb.198036

Houde, A. L. S., Gunther, O. P., Strohm, J., Ming, T. J., Li, S. R., Kaukinen, K. H., Patterson, D. A., Farrell, A. P., Hinch, S. G., & Miller, K. M. (2019). Discovery and validation of candidate smoltification gene expression biomarkers across multiple species and ecotypes of Pacific salmonids. *Conservation Physiology*, *7*. https://doi.org/ARTN coz051 10.1093/conphys/coz051

Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S., & Zhang, B. (2011). Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol*, *29*(8), 699-700. https://doi.org/10.1038/nbt.1939

Huynh, H., Ng, C. Y., Lim, K. B., Ong, C. K., Ong, C. S., Tran, E., Nguyen, T. T. T., & Chan, T. W. M. G. (2001). Induction of UO-44 gene expression by tamoxifen in the rat uterus and ovary. *Endocrinology*, *142*(7), 2985-2995. https://doi.org/DOI 10.1210/en.142.7.2985

Iversen, M. (2021). Photoperiodic history-dependent preadaptation of the smolting gill. Novel players and SW immediate response as markers of growth and welfare. . In U. T. A. U. o. Norway (Ed.).

Iversen, M., Mulugeta, T., Gellein Blikeng, B., West, A. C., Jorgensen, E. H., Rod Sandven, S., & Hazlerigg, D. (2020). RNA profiling identifies novel, photoperiod-history dependent markers associated with enhanced saltwater performance in juvenile Atlantic salmon. *PLoS One*, *15*(4), e0227496. https://doi.org/10.1371/journal.pone.0227496

Jao, L. E., Wente, S. R., & Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci U S A*, *110*(34), 13904-13909. https://doi.org/10.1073/pnas.1308335110

Jiang, F. G., & Doudna, J. A. (2017). CRISPR-Cas9 Structures and Mechanisms. *Annual Review of Biophysics, Vol 46*, *46*, 505-529. https://doi.org/10.1146/annurev-biophys-062215-010822

Jinek, C., K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA—Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (American Association for the Advancement of Science)*, *337*(6096), 816–821. https://doi.org/https://doi.org/10.1126/science.1225829

Khaw, H. L., Gjerde, B., Boison, S. A., Hjelle, E., & Difford, G. F. (2021). Quantitative Genetics of Smoltification Status at the Time of Seawater Transfer in Atlantic Salmon (Salmo Salar). *Front Genet*, *12*, 696893. https://doi.org/10.3389/fgene.2021.696893

Klemetsen, A., Amundsen, P. A., Dempson, J. B., Jonsson, B., Jonsson, N., O'Connell, M. F., & Mortensen, E. (2003). Atlantic salmon Salmo salar
L., brown trout Salmo trutta
L. and Arctic charr Salvelinus alpinus
(L.): a review of aspects of their life histories. *Ecology of Freshwater Fish*, *12*(1), 1-59. https://doi.org/10.1034/j.1600-0633.2003.00010.x

Leong, C. T., Ng, C. Y., Ong, C. K., Ng, C. P., Ma, Z. S., Nguyen, T. H., Tay, S. K., & Huynh, H. (2004). Molecular cloning, characterization and isolation of novel spliced variants of the human ortholog of a rat estrogen-regulated membrane-associated protein, UO-44. *Oncogene*, *23*(33), 5707-5718. https://doi.org/10.1038/sj.onc.1207754

Leong, C. T., Ong, C. K., Tay, S. K., & Huynh, H. (2007). Silencing expression of UO-44 (CUZD1) using small interfering RNA sensitizes human ovarian cancer cells to cisplatin in vitro. *Oncogene*, *26*(6), 870-880. https://doi.org/10.1038/sj.onc.1209836

Li, M., Yang, H., Zhao, J., Fang, L., Shi, H., Li, M., Sun, Y., Zhang, X., Jiang, D., Zhou, L., & Wang, D. (2014). Efficient and heritable gene targeting in tilapia by CRISPR/Cas9. *Genetics*, *197*(2), 591-599. https://doi.org/10.1534/genetics.114.163667

Liaskos, C., Rigopoulou, E. I., Orfanidou, T., Bogdanos, D. P., & Papandreou, C. N. (2013). CUZD1 and anti-CUZD1 antibodies as markers of cancer and inflammatory bowel diseases. *Clin Dev Immunol*, *2013*, 968041. https://doi.org/10.1155/2013/968041

Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem*, *79*, 181-211. https://doi.org/10.1146/annurev.biochem.052308.093131

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. A., . . . Davidson, W. S. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, *533*(7602), 200-+. https://doi.org/10.1038/nature17164

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods*, *25*(4), 402-408. https://doi.org/10.1006/meth.2001.1262

Lubin, R. T., Rourke, A. W., & Bradley, T. M. (1989). Ultrastructural Alterations in Branchial Chloride Cells of Atlantic Salmon, Salmo-Salar, during Parr-Smolt Transformation and Early Development in Sea-Water. *Journal of Fish Biology*, *34*(2), 259-272. https://doi.org/DOI 10.1111/j.1095-8649.1989.tb03307.x

Macqueen, D. J., & Johnston, I. A. (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc Biol Sci*, *281*(1778), 20132881. https://doi.org/10.1098/rspb.2013.2881

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J., & Koonin, E. V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol*, *9*(6), 467-477. https://doi.org/10.1038/nrmicro2577

Mapes, J., Anandan, L., Li, Q., Neff, A., Clevenger, C. V., Bagchi, I. C., & Bagchi, M. K. (2018). Aberrantly high expression of the CUB and zona pellucida-like domain-containing protein 1 (CUZD1) in mammary epithelium leads to breast tumorigenesis. *J Biol Chem*, *293*(8), 2850-2864. https://doi.org/10.1074/jbc.RA117.000162

Marraffini., & Sontheimer, E. J. (2008). CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science (American Association for the Advancement of Science)*, *322*(5909), 1843–1845. https://doi.org/https://doi.org/10.1126/science.1165771

McCormick, S. D. (2001). Endocrine Control of Osmoregulation in Teleost Fish. *American Zoologist*, *41*(4), 781–794. https://doi.org/https://doi.org/10.1093/icb/41.4.781

McCormick, S. D. (2013). Smolt Physiology and Endocrinology. In *Euryhaline Fishes* (pp. 199-251). https://doi.org/10.1016/b978-0-12-396951-4.00005-0

McCormick, S. D., Regish, A. M., & Christensen, A. K. (2009). Distinct freshwater and seawater isoforms of Na+/K+-ATPase in gill chloride cells of Atlantic salmon. *Journal of Experimental Biology*, *212*(24), 3994-4001. https://doi.org/10.1242/jeb.037275

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

Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol*, *60*(2), 174-182. https://doi.org/10.1007/s00239-004-0046-3

Moreno-Mateos, M. A., Vejnar, C. E., Beaudoin, J. D., Fernandez, J. P., Mis, E. K., Khokha, M. K., & Giraldez, A. J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods*, *12*(10), 982-988. https://doi.org/10.1038/nmeth.3543

Pisam, M., Prunet, P., Boeuf, G., & Rambourg, A. (1988). Ultrastructural Features of Chloride Cells in the Gill Epithelium of the Atlantic Salmon, Salmo-Salar, and Their

Modifications during Smoltification. *American Journal of Anatomy*, *183*(3), 235-244. https://doi.org/DOI 10.1002/aja.1001830306

Shen, M. W., Arbab, M., Hsu, J. Y., Worstell, D., Culbertson, S. J., Krabbe, O., Cassa, C. A., Liu, D. R., Gifford, D. K., & Sherwood, R. I. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature*, *563*(7733), 646-651. https://doi.org/10.1038/s41586-018-0686-x

Skulason, S., & Smith, T. B. (1995). Resource Polymorphisms in Vertebrates. *Trends in Ecology & Evolution*, *10*(9), 366-370. https://doi.org/Doi 10.1016/S0169-5347(00)89135-1

Staller, E., Sheppard, C. M., Neasham, P. J., Mistry, B., Peacock, T. P., Goldhill, D. H., Long, J. S., & Barclay, W. S. (2019). ANP32 Proteins Are Essential for Influenza Virus Replication in Human Cells. *Journal of Virology*, *93*(17). https://doi.org/https://doi.org/10.1128/JVI.00217-19

Stefansson, S. O., Bjørnsson, B. T., Ebbesson, L. O., & McCormick, S. D. (2008). Smoltification. In B. T. B. I S. O. Stefansson, L. O. Ebbesson, & M (Ed.), *Fish Larval Physiology* (Vol. Vol. 1, pp. ss. 639-681). CRC Press.

Straume, A. H., Kjaerner-Semb, E., Ove Skaftnesmo, K., Guralp, H., Kleppe, L., Wargelius, A., & Edvardsen, R. B. (2020). Indel locations are determined by template polarity in highly efficient in vivo CRISPR/Cas9-mediated HDR in Atlantic salmon. *Sci Rep*, *10*(1), 409. https://doi.org/10.1038/s41598-019-57295-w

Thorpe, J. E. (1994). An Alternative View of Smolting in Salmonids. *Aquaculture*, *121*(1-3), 105-113. https://doi.org/Doi 10.1016/0044-8486(94)90012-4

Wargelius, A., Leininger, S., Skaftnesmo, K. O., Kleppe, L., Andersson, E., Taranger, G. L., Schulz, R. W., & Edvardsen, R. B. (2016). Dnd knockout ablates germ cells and demonstrates germ cell independent sex differentiation in Atlantic salmon. *Sci Rep*, *6*, 21284. https://doi.org/10.1038/srep21284

Wedemeyer, G. A., Saunders, R. L., & Clarke, W. C. (1980). Environmental factors affecting smoltification and early marine survival of anadromous salmonids. *Marine Fisheries Review*.

West, A. C., Mizoro, Y., Wood, S. H., Ince, L. M., Iversen, M., Jorgensen, E. H., Nome, T., Sandve, S. R., Martin, S. A. M., Loudon, A. S. I., & Hazlerigg, D. G. (2021). Immunologic Profiling of the Atlantic Salmon Gill by Single Nuclei Transcriptomics. *Front Immunol*, *12*, 669889. https://doi.org/10.3389/fimmu.2021.669889

Wyman, C., & Kanaar, R. (2006). DNA double-strand break repair: all's well that ends well. *Annu Rev Genet*, *40*, 363-383. https://doi.org/10.1146/annurev.genet.40.110405.090451

Yano, A., Nicol, B., Jouanno, E., Quillet, E., Fostier, A., Guyomard, R., & Guiguen, Y. (2013). The sexually dimorphic on the Y-chromosome gene (sdY) is a conserved male-specific Y-chromosome sequence in many salmonids. *Evol Appl*, *6*(3), 486-496. https://doi.org/10.1111/eva.12032

Zhang, F. (2019). Development of CRISPR-Cas systems for genome editing and beyond. *Quarterly Reviews of Biophysics*, *52*. https://doi.org/10.1017/s0033583519000052

R Core Team (2023) R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna).

Appendix A: Supplementary tables

Raw data from the SP-dependent cluster presented in two tables of the top 100 most expressed genes in T3 and with the highest fold change between T1 and T3.

Top 100 most expressed genes in T3

Table S 1. *Raw data retrieved from West et al. (2021) showing top 100 genes with the highest mean CMP expression in T3.*

Gene ID	Locus ID	Protein ID	Product Name	HGNC ID	Cluster ID	Log2 Mean T3
gene51191:106589632	LOC106589632	XP_014035319.1	collagen alpha-1(I) chain-like	COL1A1	1	14,829
rna133441					1	13,618
gene39471:106578378	LOC106578378	XP_014012593.1	lipocalin-like	PTGDS	1	13,570
gene23459:106562604	LOC106562604	XP_013983028.1	CUB and zona pellucida-like domain-containing protein 1	CUZD1	1	13,522
gene1587:106610502	LOC106610502	XP_014065361.1	collagen alpha-1(I) chain-like	COL1A1	1	13,171
gene21114:106560296	LOC106560296	XP_013978524.1	regulator of G-protein signaling 5-like	RGS5	1	12,999
gene27388:100380643	LOC100380643	XP_013989724.1	filamin-A	FLNA	1	12,847
gene47671:106586441	abca12	XP_014029202.1	ATP-binding cassette%2C sub-family A (ABC1)%2C member 12	ABCA12	1	12,805
gene32350:106571195	LOC106571195	XP_013999449.1	ammonium transporter Rh type A-like	RHAG	1	12,765
gene16096:106609224	LOC106609224	XP_014063239.1	caldesmon%2C smooth muscle-like		1	12,741
gene3802:100195786	nca11	XP_014011390.1	Neural cell adhesion molecule 1-A		1	12,663

Table S 1 *(continued)*

gene13827:100380669	LOC100380669	XP_014059052.1	myosin%2C heavy chain 9%2C non-muscle	MYH9	1	12,535
gene14999:106608163	LOC106608163	XP_014061410.1	thrombospondin-1-like	THBS1	1	12,394
gene42824:106581666	LOC106581666	XP_014019342.1	collagen alpha-3(VI) chain-like	COL6A3	1	12,368
gene10601:100380654	LOC100380654	XP_014053853.1	periostin		1	12,328
gene30121:106568844	LOC106568844	XP_013995040.1	integrin beta-1-like	ITGB1	1	12,295
gene7715:106601336	LOC106601336	XP_014048942.1	myosin-9-like	MYH9	1	12,234
gene3031:106571564	LOC106571564	XP_014000233.1	tropomyosin beta chain		1	12,110
gene37176:106576070	LOC106576070	XP_014008408.1	non-muscle caldesmon-like	CALD1	1	12,092
gene8947:106602517					1	12,040
gene28986:106567903	LOC106567903	XP_013993256.1	solute carrier organic anion transporter family member 2B1-like	SLCO2B1	1	12,005
gene6356:106599471	LOC106599471	XP_014046199.1	extended synaptotagmin-2-B-like	ESYT2	1	11,989
gene15609:106608771	LOC106608771	XP_014062372.1	cornifelin homolog B-like		1	11,974
gene8552:106602213					1	11,947
gene23860:100286444	cld5	NP_001139855.1	Claudin-5		1	11,928
gene25460:100380663	LOC100380663	XP_013986237.1	myosin-11	MYH11	1	11,926
gene26044:100380407	ical	XP_013987214.1	Calpastatin	CAST	1	11,890
gene4879:106589071	LOC106589071	XP_014034226.1	myosin-11-like	MYH11	1	11,887
gene36068:106574999	LOC106574999	XP_014006603.1	translationally-controlled tumor protein homolog	TPT1	1	11,831
gene28148:106567044	LOC106567044	XP_013991345.1	serine/threonine-protein kinase receptor R3-like	ACVRL1	1	11,828
gene50911:106589401	LOC106589401	XP_014034805.1	inositol polyphosphate multikinase-like	IPMK	1	11,800
gene69333:106591834	LOC106591834	XP_014038573.1	collagen alpha-2(VI) chain-like	COL6A2	1	11,755
gene3029:100380665	LOC100380665	XP_014000191.1	talin-1	TLN1	1	11,700
gene35872:106574791	LOC106574791	XP_014006348.1	fibronectin-like	FN1	1	11,592

Table S 1 *(continued)*

gene34346:106573225	LOC106573225	XP_014003538.1	spondin-1-like	SPON1	1	11,545
gene18113:106611173	LOC106611173	XP_014066569.1	ral GTPase-activating protein subunit alpha-1-like	RALGAPA1	1	11,542
gene36182:106575113	app	XP_014006793.1	amyloid beta (A4) precursor protein	APP	1	11,520
gene4595:106585544	LOC106585544	XP_014027359.1	collagen alpha-6(VI) chain-like	COL6A6	1	11,461
gene23263:100194903	mmp2	NP_001133404.1	matrix metalloproteinase 2	MMP2	1	11,446
gene35364:106574335	LOC106574335	XP_014005668.1	amyloid beta A4 protein-like		1	11,408
gene10713:106604151					1	11,395
gene34375:106573197	LOC106573197	XP_014003457.1	G1/S-specific cyclin-D2	CCND2	1	11,375
gene42150:106581002	LOC106581002	XP_014018122.1	complement C4-B-like	C4B	1	11,364
gene20046:106612965	LOC106612965	XP_014070173.1	Friend leukemia integration 1 transcription factor-like	FLI1	1	11,348
gene16224:100329176	gata3	XP_014063328.1	GATA-binding protein 3	GATA3	1	11,344
gene3991:106580162	LOC106580162	XP_014016392.1	ras-related and estrogen-regulated growth inhibitor-like protein		1	11,328
gene5097:106591110	LOC106591110	XP_014037777.1	cyclic AMP-dependent transcription factor ATF-4-like	ATF4	1	11,317
gene30851:106569732	LOC106569732	XP_013996781.1	collagen alpha-1(VI) chain-like	COL6A1	1	11,313
gene25509:106564685	LOC106564685	XP_013986400.1	myosin-9-like		1	11,246
gene47933:106586687	LOC106586687	XP_014029701.1	amyloid beta A4 protein-like	APP	1	11,236
gene28630:100194831	kct2	XP_013992235.1	Keratinocytes-associated transmembrane protein 2	C5orf15	1	11,225
gene206:106611189	dync1h1	XP_014066736.1	dynein%2C cytoplasmic 1%2C heavy chain 1	DYNC1H1	1	11,215
gene45141:106583923	LOC106583923	XP_014024069.1	U1 small nuclear ribonucleoprotein 70 kDa-like	SNRNP70	1	11,184
gene9086:106602664	LOC106602664	XP_014050893.1	T-cell ecto-ADP-ribosyltransferase 2-like		1	11,179
gene50859:100196695	htra1	NP_001135189.1	Serine protease HTRA1	HTRA1	1	11,165
gene19830:106612870	LOC106612870	XP_014069949.1	complement C4-like	C4B	1	11,157
gene5750:106596905	LOC106596905	XP_014043630.1	myosin heavy chain%2C fast skeletal muscle-like	MYH8	1	11,143

Table S 1 *(continued)*

gene5966:106597294	LOC106597294	XP_014043980.1	collagen alpha-1(VI) chain-like	COL6A1	1	11,141
gene28025:106566779	LOC106566779	XP_013990641.1	monoglyceride lipase-like	MGLL	1	11,116
gene10657:106604087	LOC106604087	XP_014053950.1	plastin-3-like	PLS3	1	11,070
gene31073:106570046	LOC106570046	XP_013997449.1	syndecan-2-A-like	SDC2	1	11,069
gene10199:106603674	LOC106603674	XP_014053111.1	junctional adhesion molecule C-like	JAM3	1	11,054
gene24250:106563533	LOC106563533	XP_013984690.1	CD81 antigen-like	CD81	1	11,045
gene2246:106563403	LOC106563403	XP_013984411.1	dentin sialophosphoprotein-like		1	11,044
gene14809:106608002	LOC106608002	XP_014061066.1	utrophin-like	UTRN	1	11,035
gene13865:106607103	LOC106607103	XP_014059180.1	epithelial membrane protein 2-like		1	11,027
gene37307:106576211	LOC106576211	XP_014008678.1	transcription factor GATA-3-like	GATA3	1	11,009
gene22693:106561856	LOC106561856	XP_013981599.1	von Willebrand factor-like	VWF	1	10,996
gene10630:106604062	LOC106604062	XP_014053890.1	inhibitor of apoptosis protein-like	BIRC2	1	10,990
gene43881:106582716	LOC106582716	XP_014021552.1	transcriptional enhancer factor TEF-5-like	TEAD3	1	10,953
gene38702:106577588	LOC106577588	XP_014011241.1	mucin-5AC-like		1	10,944
gene40662:100380435	irf9	NP_001167190.1	interferon regulatory factor 9		1	10,904
gene39071:106577947	LOC106577947	XP_014011901.1	sterile alpha motif domain-containing protein 9-like		1	10,902
gene10745:106604170	LOC106604170	XP_014054065.1	uncharacterized LOC106604170	TSC22D3	1	10,898
gene21065:106560254	LOC106560254	XP_013978395.1	palladin-like	PALLD	1	10,889
gene14040:100380767	LOC100380767	XP_014059730.1	myoferlin	MYOF	1	10,888
gene36806:100380696	LOC100380696	XP_014007833.1	fibronectin	FN1	1	10,882
gene37146:100196821	timp3	NP_001135315.1	Metalloproteinase inhibitor 3	TIMP3	1	10,879
gene26209:106565346	LOC106565346	XP_013987820.1	peptidyl-prolyl cis-trans isomerase FKBP5-like	FKBP5	1	10,878
gene19407:106612383	LOC106612383	XP_014068937.1	bone morphogenetic protein 2-like		1	10,873

Table S 1 *(continued)*

gene21351:100306840	ivns1abp	XP_013979000.1	influenza virus NS1A binding protein		1	10,866
gene45446:106584117	pdgfra	XP_014024468.1	platelet-derived growth factor receptor%2C alpha polypeptide	PDGFRA	1	10,865
gene39165:100195470	hmgb3	NP_001133971.1	High mobility group protein B3	HMGB3	1	10,865
gene44233:106583134	LOC106583134	XP_014022443.1	matrix-remodeling-associated protein 8-like	MXRA8	1	10,864
gene46240:106585199	sepp1	XP_014026611.1	selenoprotein P%2C plasma%2C 1	SEPP1	1	10,851
gene51829:106590282	LOC106590282	XP_014036597.1	collagen alpha-1(XVIII) chain-like	COL15A1	1	10,825
gene34051:106572902	ubr4	XP_014002962.1	ubiquitin protein ligase E3 component n-recognin 4	UBR4	1	10,815
gene9992:106603571	LOC106603571	XP_014052906.1	pigment epithelium-derived factor-like	SERPINF1	1	10,814
gene19592:106612674	LOC106612674	XP_014069553.1	tumor necrosis factor alpha-induced protein 2-like	EXOC3L2	1	10,800
gene16790:106609880	LOC106609880	XP_014064379.1	ATP-binding cassette sub-family A member 1-like	ABCA1	1	10,793
gene3723:106577511	LOC106577511	XP_014011079.1	apolipoprotein C-I-like	APOC1	1	10,783
gene40403:106579136	ptrf	XP_014014203.1	polymerase I and transcript release factor	PTRF	1	10,770
gene30022:106568918	LOC106568918	XP_013995225.1	uncharacterized LOC106568918		1	10,770
gene2919:100380773	sptan1	XP_013998412.1	spectrin%2C alpha%2C non-erythrocytic 1	SPTAN1	1	10,757
gene28134:106567029	LOC106567029	XP_013991308.1	structural maintenance of chromosomes protein 1A	SMC1A	1	10,757
gene3796:106577754	LOC106577754	XP_014011551.1	annexin A7-like		1	10,754
gene50897:106589387	LOC106589387	XP_014034767.1	actin%2C aortic smooth muscle	ACTA2	1	10,742
gene48857:106587617	LOC106587617	XP_014031653.1	epithelial splicing regulatory protein 2-like	ESRP1	1	10,736
gene16038:106609127	LOC106609127	XP_014063058.1	flocculation protein FLO11-like	PROSER2	1	10,725

Top 100 genes with the highest fold change between T1 and T3.

***Table S 2**. Raw data retrieved from West et al. (2021) shows the top 100 genes with the highest fold change.*

Gene ID	Locus ID	Protein ID	Product Name	HGNC ID	Cluster ID	LogFC T1vsT3
gene39273:106578169	LOC106578169	XP_014012273.1	fibrinogen alpha chain-like	FGA	1	8,391
gene51540:106589985	LOC106589985	XP_014035938.1	calpain-2 catalytic subunit-like	CAPN2	1	5,784
gene40102:106578986	LOC106578986	XP_014013824.1	glutamate receptor ionotropic%2C delta-1-like	GRID1	1	5,674
gene22493:106561651	LOC106561651	XP_013981286.1	troponin I%2C cardiac muscle-like	TNNI3	1	3,812
gene67963:106595856	LOC106595856	XP_014042681.1	E3 ubiquitin-protein ligase RNF216-like	RNF216	1	3,806
gene16683:106609768	LOC106609768	XP_014064253.1	cytidine monophosphate-N-acetylneuraminic acid hydroxylase-like		1	3,762
gene2720:106568282					1	3,665
gene23154:106562311	LOC106562311	XP_013982581.1	tryptophan 5-hydroxylase 1-like	TPH1	1	3,621
gene23459:106562604	LOC106562604	XP_013983028.1	CUB and zona pellucida-like domain-containing protein 1	CUZD1	1	3,609
gene12239:100194549	telt	NP_001133107.1	titin-cap	TCAP	1	3,569
gene24852:106563984	LOC106563984	XP_013985415.1	mucin-12-like		1	3,565
gene29578:106568518	esm1	XP_013994415.1	endothelial cell-specific molecule 1	ESM1	1	3,426
gene19346:106612328	LOC106612328	XP_014068854.1	golgin subfamily A member 6-like protein 22		1	3,405
gene79197:106592881	LOC106592881	XP_014039706.1	sialidase-like		1	3,374
gene38359:106577235	LOC106577235	XP_014010628.1	secreted frizzled-related protein 5-like	SFRP5	1	3,350
gene57723:106595184	LOC106595184	XP_014042038.1	trichohyalin-like		1	3,312
gene68801:106596609	LOC106596609	XP_014043356.1	trichohyalin-like		1	3,279
gene25926:106565075					1	3,265
gene57722:106595182	LOC106595182	XP_014042037.1	trichohyalin-like		1	3,224

Table S 2 *(continued)*

gene67948:106599144	LOC106599144	XP_014045720.1	golgin subfamily A member 6-like protein 22		1	3,173
gene9259:106602836	LOC106602836	XP_014051236.1	max-binding protein MNT-like		1	3,160
gene54303:106597201	LOC106597201	XP_014043896.1	trichohyalin-like		1	3,094
gene44998:106583739	LOC106583739	XP_014023758.1	stonustoxin subunit beta-like		1	3,061
gene49581:106588125					1	3,047
gene16123:106609198	LOC106609198	XP_014063190.1	leiomodin-2-like	LMOD2	1	3,022
gene11057:106604357	LOC106604357	XP_014054387.1	thymosin beta	TMSB10	1	2,963
gene2586:100194645	LOC100194645	NP_001133202.1	phosphoglycerate mutase 2-2 (muscle)	PGAM2	1	2,883
gene64345:106592191	LOC106592191	XP_014038983.1	vegetative cell wall protein gp1-like		1	2,853
gene35827:106574743	LOC106574743	XP_014006249.1	striated muscle preferentially expressed protein kinase-like		1	2,842
gene36271:106575240	LOC106575240	XP_014007101.1	matrix-remodeling-associated protein 5-like	MXRA5	1	2,824
gene3920:106579424	LOC106579424	XP_014014798.1	brain-specific angiogenesis inhibitor 1-like	ADGRB1	1	2,796
gene38115:106576998	pcsk2	XP_014010102.1	proprotein convertase subtilisin/kexin type 2	PCSK2	1	2,788
gene30725:106569611	LOC106569611	XP_013996621.1	nicotinamide riboside kinase 2-like		1	2,786
gene2197:100194501	ckmt2	XP_013979849.1	creatine kinase%2C mitochondrial 2 (sarcomeric)	CKMT2	1	2,770
gene51549:106589996	LOC106589996	XP_014035950.1	protein jagged-1b-like		1	2,765
gene14255:100195995	hsp11	NP_001134496.1	Heat shock protein Hsp-16.1/Hsp-16.11		1	2,765
gene11894:106605379	LOC106605379	XP_014056402.1	collagen alpha-1(XXVIII) chain-like	COL28A1	1	2,748
gene38976:106577856	LOC106577856	XP_014011731.1	alpha-actinin-3-like	ACTN2	1	2,735
gene25:106612874	LOC106612874	XP_014070387.1	lysyl oxidase homolog 3-like	LOXL2	1	2,701
gene57639:106591863	LOC106591863	XP_014038608.1	E3 ubiquitin-protein ligase RNF216-like	RNF216	1	2,699
gene30913:106569791	LOC106569791	XP_013996879.1	alpha-N-acetylgalactosaminide alpha-2%2C6-sialyltransferase 5-like	ST6GALNAC5	1	2,672
gene5224:106591906	LOC106591906	XP_014038655.1	tuberin-like	TSC2	1	2,643

Table S 2 *(continued)*

gene51557:106590003	LOC106590003	XP_014035958.1	AT-rich interactive domain-containing protein 4B-like		1	2,618
gene13645:106606883	LOC106606883	XP_014058782.1	glycine-rich cell wall structural protein-like		1	2,613
rna133452					1	2,610
gene25577:106564730	LOC106564730	XP_013986500.1	actin%2C alpha skeletal muscle 2-like	ACTC1	1	2,605
gene43211:106582044	LOC106582044	XP_014020211.1	spore wall protein 2-like		1	2,604
gene40782:100136486	LOC100136486	XP_014015174.1	myosin-6	MYH6	1	2,602
gene3723:106577511	LOC106577511	XP_014011079.1	apolipoprotein C-I-like	APOC1	1	2,595
rna133453					1	2,593
gene19145:106612124	LOC106612124	XP_014068489.1	serine/arginine repetitive matrix protein 2-like		1	2,591
gene40496:106579065	LOC106579065	XP_014014035.1	sarcoplasmic/endoplasmic reticulum calcium ATPase 1	ATP2A1	1	2,581
gene23997:106563158	LOC106563158	XP_013983900.1	matrix metalloproteinase-17-like	MMP17	1	2,569
gene9210:106602726	slc25a4	XP_014051019.1	solute carrier family 25 (mitochondrial carrier%3B adenine nucleotide translocator)%2C member 4	SLC25A5	1	2,541
gene62178:106592913	LOC106592913	XP_014039737.1	RNA-binding protein 47-like		1	2,539
gene34191:106573043	LOC106573043	XP_014003141.1	L-lactate dehydrogenase A chain-like	LDHB	1	2,535
gene9044:100136395	alp	NP_001117019.1	alkaline phosphatase	PDLIM3	1	2,534
gene37178:100196284	tnni1	XP_014008283.1	troponin I%2C slow skeletal muscle	TNNI3	1	2,532
gene58603:106592849	LOC106592849	XP_014039673.1	rho GTPase-activating protein SYDE1-like	SYDE1	1	2,521
gene2255:106563467	LOC106563467	XP_013984557.1	heat shock protein beta-11-like		1	2,519
gene47312:100194537	LOC100194537	NP_001133095.1	LOC394070 protein-like		1	2,518
gene31107:106570020	LOC106570020	XP_013997382.1	trichohyalin-like		1	2,517
gene5008:106590279	LOC106590279	XP_014036592.1	sarcalumenin-like	SRL	1	2,502
gene7090:106600734	atp2a1	XP_014047794.1	ATPase%2C Ca++ transporting%2C cardiac muscle%2C fast twitch 1	ATP2A1	1	2,498

Table S 2 *(continued)*

gene33815:106572662	LOC106572662	XP_014002509.1	low-density lipoprotein receptor-related protein 1-like	LRP1	1	2,495
gene5076:106590850	LOC106590850	XP_014037504.1	synaptogyrin-1-like	SYNGR1	1	2,492
gene15218:106608400	LOC106608400	XP_014061784.1	CST complex subunit CTC1-like		1	2,482
gene43441:100194700					1	2,482
gene71045:106593429	LOC106593429	XP_014040237.1	pyruvate dehydrogenase E1 component subunit beta%2C mitochondrial-like	PDHB	1	2,479
gene4719:106587486	LOC106587486	XP_014031391.1	actin%2C alpha skeletal muscle 2-like	ACTC1	1	2,475
gene47244:100194619	fbp2	NP_001133176.1	fructose-1%2C6-bisphosphatase 2		1	2,475
gene74301:100194636	LOC100194636	NP_001133193.1	enolase 3-2	ENO3	1	2,473
gene24506:106563660	LOC106563660	XP_013984904.1	glycogen phosphorylase%2C muscle form-like	PYGM	1	2,470
gene7293:106600825	LOC106600825	XP_014048002.1	nuclear receptor subfamily 1 group D member 1-like	NR1D1	1	2,468
gene35872:106574791	LOC106574791	XP_014006348.1	fibronectin-like	FN1	1	2,465
gene31753:106570612	LOC106570612	XP_013998549.1	peptidyl-prolyl cis-trans isomerase FKBP9-like	FKBP9	1	2,461
gene5750:106596905	LOC106596905	XP_014043630.1	myosin heavy chain%2C fast skeletal muscle-like	MYH8	1	2,460
gene17394:106610444	LOC106610444	XP_014065298.1	C-type lectin domain family 4 member F-like	CLEC4E	1	2,454
gene31674:106570544					1	2,441
gene40050:106578942	LOC106578942	XP_014013706.1	neurexin-1a-like		1	2,440
gene35313:106574207	LOC106574207	XP_014005354.1	dermatopontin-like	DPT	1	2,438
gene48735:106587496	LOC106587496	XP_014031410.1	calcitonin-1-like	CALCB	1	2,435
gene16934:100194655	LOC100194655	NP_001133211.1	solute carrier family 25-2	SLC25A5	1	2,431
gene6533:106599743	LOC106599743	XP_014046518.1	myosin light chain kinase 2%2C skeletal/cardiac muscle-like	MYLK4	1	2,428
gene51845:106590239	LOC106590239	XP_014036514.1	myosin-6-like		1	2,425
gene15495:106608688	LOC106608688	XP_014062262.1	neurexin-2-like	NRXN2	1	2,417
gene51129:100137050	prvb	NP_001117189.1	parvalbumin beta		1	2,417

Table S 2 *(continued)*

gene69333:106591834	LOC106591834	XP_014038573.1	collagen alpha-2(VI) chain-like	COL6A2	1	2,417
gene62858:106592374	LOC106592374	XP_014039191.1	myosin-binding protein C%2C fast-type-like	MYBPC2	1	2,413
gene3083:106572051	LOC106572051	XP_014001283.1	protein Shroom3-like		1	2,402
gene31660:106570522	LOC106570522	XP_013998414.1	3-hydroxyisobutyrate dehydrogenase%2C mitochondrial-like	HIBADH	1	2,397
gene60674:106591971	LOC106591971	XP_014038724.1	desmin-like		1	2,392
gene18297:100194689	LOC100194689	XP_014066983.1	beta-taxilin		1	2,391
gene22534:106561692	LOC106561692	XP_013981373.1	filamin-C-like	FLNC	1	2,378
gene20503:106613505					1	2,373
gene41928:106580775	LOC106580775	XP_014017651.1	ubiquitin carboxyl-terminal hydrolase 2-like		1	2,364
gene51556:106590000	LOC106590000	XP_014035956.1	protein jagged-1b-like		1	2,361
gene5698:106596436					1	2,352
gene40523:100194624	LOC100194624	XP_014014713.1	aldolase a%2C fructose-bisphosphate 1	ALDOA	1	2,343

Appendix B: Plasmids used for *in vitro* transcription and [RNA]

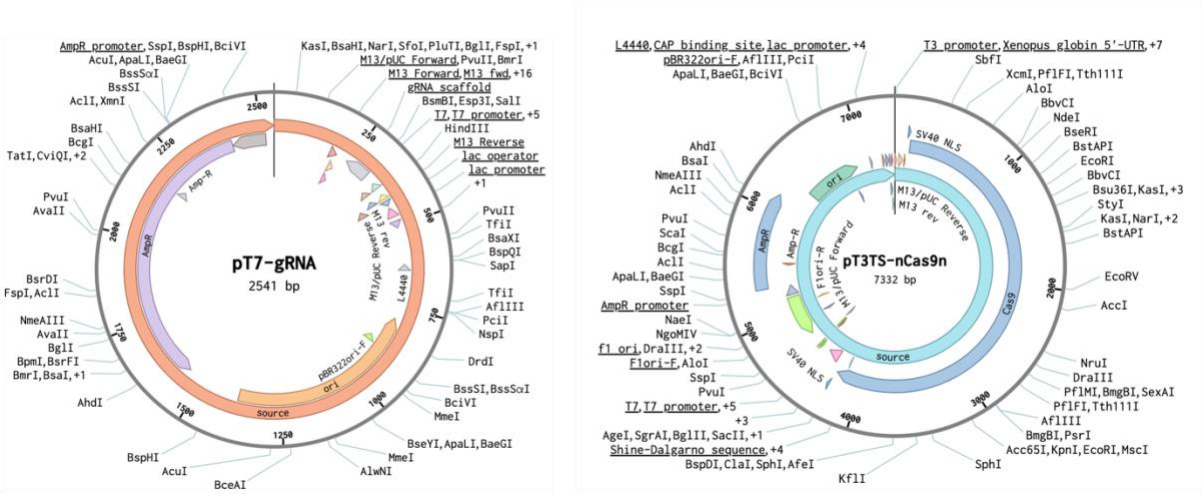

Figure S 1. *Left panel: pT7-gRNA plasmid (Addgene, #46759). Used to create cloned constructs and the gRNAs. Right panel: pT3TS-nCas9n plasmid (Addgene, #46757) used to synthesize Cas9 mRNA.*

Table S 3. *DNA concentration for each purified plasmid product, measured with NanoDrop.*

Gene	[DNA] ng/µl	Gene	[DNA] ng/µl
(4.1) Cuzd1	243	*(7.1) Cuzd1*	184.8
(4.2) Cuzd1	187.5	*(7.2) Cuzd1*	91.9
(4.3) Cuzd1	284.9	*(7.3) Cuzd1*	96.7
(4.4) Cuzd1	308.7	*(7.4) Cuzd1*	190.8
(5.1) Cuzd1	74.8	*(A1) Slc45a2*	138.5
(5.2) Cuzd1	149.8	*(A2) Slc45a2*	185.2
(5.3) Cuzd1	179.5	*(A3) Slc45a2*	184.2
(5.4) Cuzd1	188.1	*(A4) Slc45a2*	112.5

Appendix C: RNA extraction methods from tissues and concentrations

Extractions methods

***Table S 4**. Extraction method for the different tissues from parr and smolt.*

Parr		Smolt	
Tissue	**Extraction method**	**Tissue**	**Extraction method**
Gill	Qiazol	Gill	Qiazol
Liver	Qiazol/MiniKit	Liver	MiniKit
Brain	Qiazol/MiniKit	Brain	Qiazol
Heart	Qiazol	Heart	Qiazol
Eye	Qiazol	Eye	Qiazol/MiniKit
Muscle	Qiazol	Muscle	Qiazol
Intestine	MiniKit	Intestine	Qiazol/MiniKit
Pancreas	MiniKit	Pancreas	MiniKit
Head Kidney	MiniKit	Head Kidney	Qiazol
Kidney	MiniKit	Kidney	Qiazol
Fin	Qiazol	Fin	Qiazol
Spleen	MiniKit	Spleen	Qiazol

RNA isolation with Trizol

The frozen samples (stored in RNA later) were added to a 2 mL tube (on ice) with 2 beads (Fisherbrand ™ Bead Mill Homogenizer) and 1000 µl TRIzol (cooled to 4°C before addition). The samples were homogenized immediately for 2 min at 30.0 I/s twice, then incubated at room temperature for 5 min and centrifuged at 12 000 rpm for 10 min at 4°C. The supernatant was then transferred to a new tube, with further addition of 500 µl pre-cooled (4°C) chloroform. The samples were then vortexed and incubated at room temperature for 5 min, followed by centrifugation at 12 000 rpm for 20 min at 4°C. After centrifugation, the samples separate into a lower red phenol-chloroform phase, interphase, and a colorless upper aqueous phase. The aqueous phase (~500 µl) was transferred to a new tube without disturbing the interphase, followed by the addition of 500 µl TRIzol and 200 µl chloroform to the aqueous phase. The samples were then vortexed and incubated at room temperature for 5 min, followed by centrifugation at 12 000 rpm for 15 min at 4°C. The samples were separated into three phases again, and the aqueous phase was again transferred (~500 µl) to a new tube. To the aqueous phase, 500 µl of pre-cooled (-20°C) isopropanol (2-propanol) was added. The samples were mixed by shaking vigorously by hand until the mixture became turbid and then incubated at -20°C for 4-6 hours. After incubation, the samples were centrifuged at 12 000 rpm for 10 min at 4°C, then the isopropanol was carefully removed without disturbing the pellet. The pellet was then washed in 1000 µl 80% EtOH by vortexing after the addition of EtOH. The samples were then centrifuged at 10 000 rpm for 5 min at room temperature. After centrifugation, the EtOH was carefully removed without disturbing the pellet. The pellet was then air-dried for ~30 min under all EtOH was evaporated. The pellet was then dissolved in 30 µl Nuclease-Free water and incubated on a heat block at 55°C for 10 min (to increase the solubility). After incubation, the pellet was DNase-treated with 3 µl (10 % of sample volume) Turbo DNase and incubated at 37 °C for 15 min. RNA concentration was then measured with NanoDrop and later stored at -80°C.

RNA concentrations

***Table S 5**. RNA concentrations measured with NanoDrop from respective tissues of parr and smolt.*

Parr			Smolt		
#	**Organ**	**RNA [ng/µl]**	**#**	**Organ**	**RNA [ng/µl]**
1P-1	Gill	1104,6	**1S-1**	Gill	1549
1P-2	Gill	1482,8	**1S-2**	Gill	1102,7
1P-3	Gill	1053,9	**1S-3**	Gill	1365,3
1P-4	Gill	698,6	**1S-4**	Gill	3191,3
4P-1	Brain	972,2	**4S-1**	Brain	439,7
4P-2	Brain	921,6	**4S-2**	Brain	428,1
4P-3	Brain	1767,6	**4S-3**	Brain	299
4P-4	Brain	161,9	**4S-4**	Brain	256,3
5P-1	Heart	1098,6	**5S-1**	Heart	1042,6
5P-2	Heart	1333,8	**5S-2**	Heart	1458,8
5P-3	Heart	876,3	**5S-3**	Heart	1883,4
5P-4	Heart	447,6	**5S-4**	Heart	3776,1
7P-1	Muscle	1623,9	**7S-1**	Muscle	2999,3
7P-2	Muscle	864,2	**7S-2**	Muscle	2699,3
7P-3	Muscle	870,3	**7S-3**	Muscle	2889,2
7P-4	Muscle	850,5	**7S-4**	Muscle	3044,7
9P-1	Pancreas	1193,8	**9S-1**	Pancreas	306,5
9P-2	Pancreas	1288,8	**9S-2**	Pancreas	207,7
9P-3	Pancreas	467,1	**9S-3**	Pancreas	202,6
-	-	-	**9S-4**	Pancreas	140
12P-1	Fin	1294,1	**12S-1**	Fin	3988,2
12P-2	Fin	1234,5	**12S-2**	Fin	3165,9
12P-3	Fin	2481,2	**12S-3**	Fin	3772,9
12P-4	Fin	2388,7	**12S-3**	Fin	9245,9
13P-1	Spleen	312	**13S-1**	Spleen	1229
13P-2	Spleen	233	**13S-2**	Spleen	1192,4
13P-3	Spleen	110,3	**13S-3**	Spleen	692,9
-	-	-	**13S-4**	Spleen	629,9

Appendix D: qPCR primer efficiency

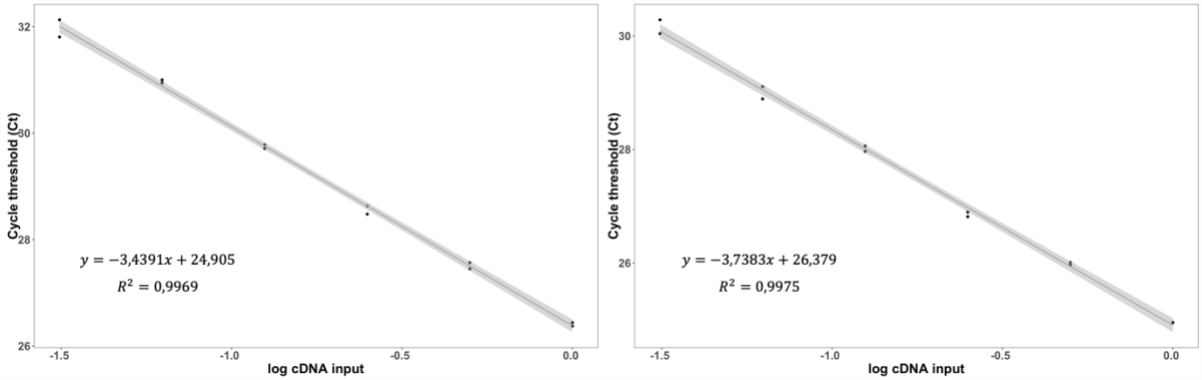

Figure S 2 *qPCR primer efficiency. Left panel shows F1R1 with a primer efficiency of 85 %. Right panel shows F2R2 with a primer efficiency of 95 %.*

Appendix E: Clustal alignment of *Cuzd1* paralogs

```
                    5577                                                                                5658
Target Cuzd1        attggacaatatgctgtggcgtgtgattggacaatatgctgtggcatgtgattggacaatatgctgtggcatgttattggac
Paralog Cuzd1       attggacaatatgctgtggtgtgtgattggacaatatgctgtgctgtgtgattggacaatatgctgtggcgtgtggattggac

                    5659                                                                                5740
                    aatatgctgtggtgtgtgattgtataattgtgcactttgATAACAGATGAAGCTGTGAATACTGTATGATACATACTATGCC
                    aatatgctgtggcgtgggattgtataattgtgcactttgATAACAGATGAAGCTGTGAATACTGTATGATACATACTACACC

                    5741                                                                                5822
                    TTGACAACACATCTGAAATCACCAAACATGAAGGAAAGATAAAAATGAAACTCAAAATGACCCCAACCAACTCAAATCATGG
                    TAGACAACACATCTGAAATCACCAAACATGAAGGAAAGATAAAAATGAAACTCGAAATGACCCCAACCAACTCAAATCGTGG

                    5823                                                                  Guide 3       5904
                    TTCACAGTCGGATTGTTTTGATTAAGATATGTATGCAAAAGATAGGACGAAATACAACAATAACAAAATGTCTCCCGTGACA
                    TTCACAGCCGGATTGTTTTG-TTAAGATATGTATGCAAAAGATAGGAAGAAATACAACAATAACCAAATGTCTCCTTGTGACA

                    5905                                                        Guide 1                 5986
                    GTCAAAGCCAAGGTGGTGTGTGATGAGACCAAGATGACGGTAGAAGTGGAGAAGTCCTCTGTTACTGGAATTGACGAGGACC
                    GTCAAAGCCAAGGTGGTGTGTGATGAGACCAAGATGACGGTAGAAGTGGAGAAGTCCTCTGTTACTGGAATTCACGAGGACC

                    5987                                                                                6068
                    ATCTCCGCCTCAACGATCCCAGTAACTCTGCCTGCGACCTGCAACGCTTCTCTAACAACACCCACATCATTGGAGTCATCCC
                    ATCTCCGCCTCAACGATCCCAGTAACTCTGCCTGCGACCTGCAACGCTTCTCTAACAACACCCACATCATTGGAGTCATCCC

                    6069                                                                                6150
                    CCTCAATGCCTGTGGCACTCAGATAGAGGTGAGGCCTCCGAGATGTCAAATATTACAGTGGGATGACACATGTACAGTATTC
                    CCTCAATGCCTGTGGCACTCAGATAGAGGTGAGGCCTCCGAGATGTGAAATATTACAGTGGGATGACACATGTACAGTATTC

                    6151                                                                                6232
                    CTCTTTGTGTGAGACCCATTCACTCCTGGGTGTTTGTCTGCGTTCTCTATATTGAACCTGGTTCGAGGGGATATTCACCCCA
                    CTCTTTGTGTGCGACCCATTCACTCCTGGGTGTTTGTCTGCGTTCTCTATATTGAACCTGGTTCGAGGGGATATTCACCCCA
                                                                    :
                                                                    :
                    7381                                                                                7462
                    GTGCATTGTGGACGAGACGGTTCAAATCTTCTCACCCCGTCACCAGAGGCATTTCCAGTTCGGAATGGAAGCTTTCAAATTC
                    GTGCATTGTGGACGAGACGGTTCAAATCTTCTCACCCCGTCACCAGAGGCATTTCCAGTTCGGAATGGAAGCTTTCAAATTC

                    7463          Guide 2                                                               7544
                    ATCGGAATGCACGACCAGGTAAATGGGGACAGATTGCTGGATCATTCAAGGGTCATATCAAGGGTGATATCTTCTctgttca
                    ATCGGAATGCATGACCAGGTAAATGGGGACAGATTGCTGGATCATTCAAGGGTCATATCAAGGGTGATATCTTCT-------

                    7545                                                                                7626
                    gtaacctttgaaccctctggctataggtgtacatcagctgttcagtgacctttaaaccctctggctataggtgtacatcagc
                    ----------------------------------------------------------------------------------
```

Figure S 3. *Clustal aligment of Cuzd1 paralogs with target sequences highlited*

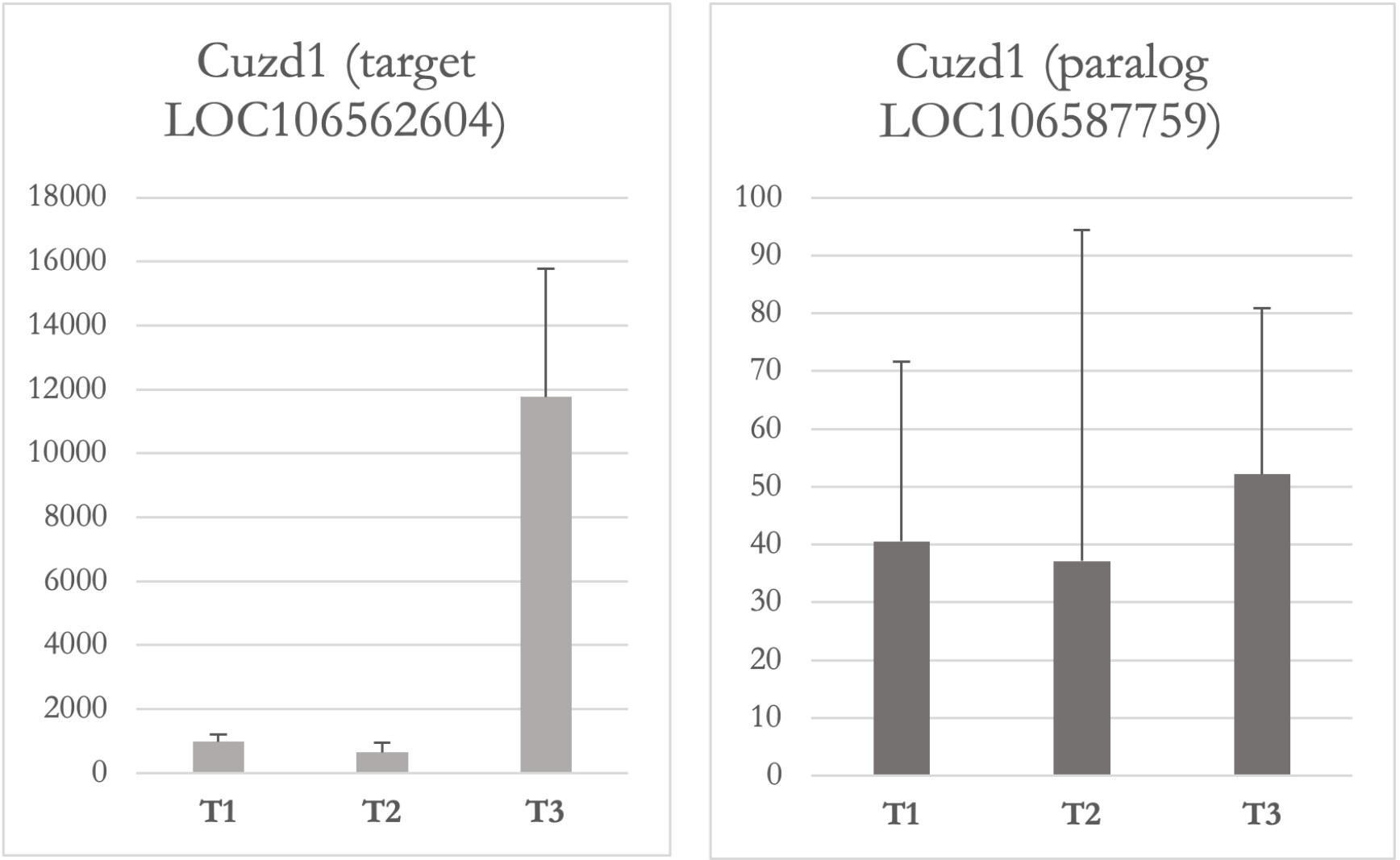

***Figure S 4**. Comparison of the two Cuzd1 paralogs in gene expression counts. Left panel: Our target Cuzd1 paralog showing winter-dependency. Right panel: non-target Cuzd1 paralog clearly shows it is not smoltification-related.*

Appendix F: Supplementary figure for visual identification

Figure S 5. *Photographs of knockout groups compared next to the control group A) G2 with no albino phenotype; B) G3 with partial pigmentation; C) G4 with partial pigmentation.*

Appendix G: Subset of target sequences for *Slc45a2* and *Cuzd1*

Part of the *Slc45a2* targe gene (Contig: AGKD01080285) (target sequence marked below + PAM sequence in grey)

CCCTCCTTCCCTCTCTCTCTGATGTGTCTGCAGATGTCCAGAGGCTGCTGCTGCCTTACATCGGTCT
GAAGGGGCTCTACTTCGTAGGATACTTTGTGTTTGGTCTGGGCACCAGTCTTATCGGCCTGTTCCCC
AACATTATCACCACCCTCATCCTGTGCAGTGTCTTCGGAGTCATGTCCAGCACCTTGTACACCATCC
CCTTCAACCTCATCTCAGAGTACCACCGTGAGGAGGAGGTCAGTACCCTCTCTTTCATAGCCTGGTC
CCAGATCTGGTTGTGCCACTACAATGACCATAGGAGTTGGCAAGACTGCACAAACAGATCTGGGA
CCAGGCTACTCTCTCACACCTCAGATAGTTGGAACTGATAGTGGATTGTACATTCTGAGGCTGTGG
CAGCTTACGTTACAAGCTAATAATGTGTAGTCGTTGCGAACCTCATCCACCACCTGTCTGTAGCTCA
TACATCGAGTGATGCACATCTGTAACTTGAATACTCACCATGTGTCATGTATCACCGCTCCCCTTTT
CGCCTTCATGAAGGCGTAGAGTTACTTTCTAATTCTATTTATTGATTTAGTCTTTATTCTGTTTTCTG
TCCAACGAGAGTGTCATGAATCACTTGTAATTCACATTGCATAGGTAATCCCACTATAGGTTGCTTT
CTCTGGAATACTTGAATGTCATTCAGGGAAAACAACTCTATACCCGCATTTCTTTTGAATTGATTCC
CACACTGTCTAAATCAATACCCATCACCCTTCCAAAGAAGAGCGTATTATAAGACAAGGAGTGGTG
AGGATTGGCATCCAAATGTGTCTTAATACAGTGAGATAATGACAGAGTGGACCCACGTGCATCACT
AGAGCTGCTTCTGATTACGTTACCACTTCCCTTCACCAACAGATGGGTGACCTTCTGTCCTCCCAAC
CCCCCGTCTCCCCACTCCCCATGAGCATGACTATTTGCCAGGCCATAGAGCCCAGCTATCTGTTGAG
CATGTCAATGTTGAGTGGTCAACGCTGTGTTGGTCAGCATGGGCCCCATGCGAAGAGCAGGTGTGC
CCTTCAAAACAGACAGGCCTCCCAAACACCCTCCCCTGTCTCTACAGACCGATCGGACAACCTGTT
CCTTTCCATAATGCTGTTGCTCAACAGGCCTCGCTTAGACAAGCTGCTGTTAGTGCTTTGGCTAAGC
AGTAAACACCCTTTCATTATGTAAAAAGGGAAAGGCTTGGCTTTGGATGGAAAAAAGGGATTAAT
GAGATTCAGTTTAGCCAATCTGCACTTCACTTTGGATTCAACACTTTCTGTCAGCTTCGGTAAGGGC
ATCTCTGGACAGCGCTCCATGTCTAATTGTGAAAATAGTAGACGTATATTCTATAATCTCAGGACG

Part of the *Cuzd1* target gene (LOC106562604) (target sequences marked below; green = G3; yellow = G1; pink = G2; PAM sequence in grey)

ATTGGACAATATGCTGTGGCATGTTATTGGACAATATGCTGTGGTGTGTGATTGTATAATTGTGCAC
TTTGATAACAGATGAAGCTGTGAATACTGTATGATACATACTATGCCTTGACAACACATCTGAAAT
CACCAAACATGAAGGAAAGATAAAAATGAAACTCAAAATGACCCCAACCAACTCAAATCATGGTT
CACAGTCGGATTGTTTTGATTAAGATATGTATGCAAAAGATAGGACGAAATACAACAATAACAAA
ATGTCTCCCGTGACAGTCAAAGCCAAGGTGGTGTGTGATGAGACCAAGATGACGGTAGAAGTGGA
GAAGTCCTCTGTTACTGGAATTGACGAGGACCATCTCCGCCTCAACGATCCCAGTAACTCTGCCTG
CGACCTGCAACGCTTCTCTAACAACACCCACATCATTGGAGTCATCCCCCTCAATGCCTGTGGCACT
CAGATAGAGGTGAGGCCTCCGAGATGTCAAATATTACAGTGGGATGACACATGTACAGTATTCCTC
TTTGTGTGAGACCCATTCACTCCTGGGTGTTTGTCTGCGTTCTCTATATTGAACCTGGTTCGAGGGG
ATATTCACCCCAAGGCACCAGCTAGGCTGACAGATCTTGAATTGGTTTCCTCAGAGAAACAGTCAA
GATGTAGTGTTTTGGTGAGGGGTGGCTTCATGCTGCCTGTGGGAAGATCTGCCCTCAGAACAGTTT
TGCATCAAAAGAGAAAGCAACATCTTGGTGATGTGTGTAAAGTTCAGTGCTGGCAGTGCCCACAA
ATGTAACTGTCTAGATGATGTTCCTCTCCATCTCTCTCAATCCAGGAGGACGACGACAACCTCATCT
TCAAGAACGAAATCACCACCTTCGACAACCCCAACGACATCATCACCAGGCACCACCAGGTGGAG
ATCCAGTTCTACTGCCAGTACGCCAAACGTGGCAACGTGTCCCTGGGCTTCAGTGCACACAGGGAC
AGCATCACGGTGGTGGAGAAAGGCTTCGGCACGTTCACCTACCAGTTTGAGTTCTACCAGACCAGC
CAGTTCACCAATATGGTGGATTCCCGTGACTACCCGTTAGACGTGGTGGTGAAGCAGATGATCTAC
ATGGACATTGAGGCCGTGTCCACTGTGAACAACACTGAGCTCTTCGTGGAGTCCTGCAGAGCGGCG
CCGTACGACAACCCCAACTACCACCCCACCTACCCCATCATAGAAAACGGGTGAGACTGATCACAC

AGATTCAGATCAGCTTTATTAGCATGAGTAGCAAGGCTACTGTTGGTAAAGCAAAGTGAGAGAAA
GACATCAACAATAACATTCAACATAATAATAATAGTAATGGTGATTAGAGAAAGGAAACACACAT
ACATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA
TATATATATATATATATATATATATATATATTTAAAATATAAACACTGAGGCATTCATCAACCTCAT
GGTGGACACTTGTCAGTACTGTGAGATGAGTAATAGATGCTGGGGAATATTATAGACTAAAAGATT
GTTTTACTGTATGTACTTTACATATTTACTGTACTCTAGGTTTTAACTGTCTGTTTTTTTACCAGGTG
CATTGTGGACGAGACGGTTCAAATCTTCTCACCCCGTCACCAGAGGCATTTCCAGTTCGGAATGGA
AGCTTTCAAATTCATCGGAATGCACGACCAGGTAAATGGGGACAGATTGCTGGATCATTCAAGGGT
CATATCAAGGGTGATATCTTCTCTGTTCAGTAACCTTTGAACCCTCTGGCTATAGGTGTACATCAGC
TGTTCAGTGACCTTTAAACCCTCTGGCTATAGGTGTACATCAGCTGTTCAGTGACCTTTGAACCCTC
TGGCTATAGGTGTACATCTGCTATTCAGTAACCTTTAAACCCTCTGGCTATAGGTGTACATCAGCTG
TTCAGTGACCTTTAATCCCTCTGGCTATAGGTGTACATCTGCTATTCAGTGACCTTTAAACGCTCTG
GCTATAGGTGTACATCAGCTGTTCAGTGACCTTTAAACCCTCTGGCTAT

Appendix F: Statistical outputs from the t-tests, one-way ANOVA and Tukey

T-test output from multi-tissue data

Welch Two Sample t-test

```
data:  PS_t$GillP and PS_t$GillS
t = -3.9866, df = 4.0221, p-value = 0.01614
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.8180121 -0.8671451
sample estimates:
mean of x mean of y
 1.097542  3.940120
```

Welch Two Sample t-test

```
data:  PS_t$BrainP and PS_t$BrainS
t = 7.0634, df = 2.0033, p-value = 0.01937
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.3945416 1.6193271
sample estimates:
 mean of x  mean of y
1.02152759 0.01459322
```

Welch Two Sample t-test

```
data:  PS_t$MuscleP and PS_t$MuscleS
t = -3.035, df = 3.6476, p-value = 0.04351
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -7.4398921 -0.1878736
sample estimates:
mean of x mean of y
 1.267726  5.081609
```

```
Welch Two Sample t-test

data:  PS_t$HeartP and PS_t$HeartS
t = -0.43407, df = 2.4606, p-value = 0.6994
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.007419  1.577117
sample estimates:
mean of x mean of y
 1.028332  1.243483

Welch Two Sample t-test

data:  PS_t$FinP and PS_t$FinS
t = -0.88951, df = 2.0782, p-value = 0.4645
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -12.226016   7.911105
sample estimates:
mean of x mean of y
 1.860639  4.018094

Welch Two Sample t-test

data:  PS_t$PancreasP and PS_t$PancreasS
t = 2.6367, df = 2.2271, p-value = 0.1063
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.408939  2.104879
sample estimates:
mean of x mean of y
1.1118169 0.2638469

Welch Two Sample t-test

data:  PS_t$SpleenP and PS_t$SpleenS
t = 4.1849, df = 2, p-value = 0.05263
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.02983432  2.15144472
sample estimates:
  mean of x   mean of y
1.063145274 0.002340071
```

ANOVA and Tukey output from ontogeny data

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stage	5	52.77	10.553	34.23	1.06e-06 ***
Residuals	12	3.70	0.308		

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = logFC ~ Stage, data = ontogeny2)

$Stage
                  diff        lwr         upr     p adj
Feed-Eye     3.4449415  1.9220412  4.96784179 0.0000725***
Fin-Eye      2.4941746  0.9712743  4.01707490 0.0014608***
Hatch-Eye    1.4128650 -0.1100353  2.93576530 0.0748499
Mid-Eye      2.6039229  1.0810226  4.12682321 0.0010053***
Organ-Eye   -1.5665126 -3.0894129 -0.04361233 0.0425464*
Fin-Feed    -0.9507669 -2.4736672  0.57213342 0.3495728
Hatch-Feed  -2.0320765 -3.5549768 -0.50917619 0.0075598**
Mid-Feed    -0.8410186 -2.3639189  0.68188172 0.4702037
Organ-Feed  -5.0114541 -6.5343544 -3.48855382 0.0000014***
Hatch-Fin   -1.0813096 -2.6042099  0.44159070 0.2350762
Mid-Fin      0.1097483 -1.4131520  1.63264860 0.9998526
Organ-Fin   -4.0606872 -5.5835875 -2.53778693 0.0000135***
Mid-Hatch    1.1910579 -0.3318424  2.71395820 0.1637090
Organ-Hatch -2.9793776 -4.5022779 -1.45647733 0.0002953***
Organ-Mid   -4.1704355 -5.6933358 -2.64753524 0.0000102***
```