Norwegian University
of Life Sciences

Master's Thesis 2022   30 ECTS
Faculty of Biosciences

Sequence analysis and *in vitro* genome editing of Atlantic salmon MHC-I-F10 gene in the Atlantic salmon macrophage cell line TO cells

Kobiowu Abdulhadi Akande

European Master in Animal Breeding and Genetics (EMABG)

SEQUENCE ANALYSIS AND *IN VITRO* GENOME EDITING OF ATLANTIC SALMON MHC-I-F10 GENE IN ATLANTIC SALMON MACROPHAGE CELL LINE TO CELLS

MASTER OF SCIENCE (MSc) THESIS

KOBIOWU, ABDULHADI AKANDE

A THESIS SUBMITTED TO THE DEPARTMENT OF ANIMAL AND AQUACULTURAL SCIENCE,

FACULTY OF BIOSCIENCES,

NORWEGIAN UNIVERSITY OF LIFE SCIENCES, NORWAY

JUNE 2022

EUROPEAN MASTER IN ANIMAL BREEDING AND GENETICS

Sequence Analysis and In Vitro Genome Editing of Atlantic Salmon Mhc-I-F10 Gene In Atlantic Salmon Macrophage Cell Line TO Cells

Kobiowu Abdulhadi

June 2022

Main supervisor: Øystein Evensen (NMBU)

Co-supervisor(s):

Jens Tetens (UGOE)

Amr Ahmed Abdelrahim Gamil (NMBU)

DEDICATION

This research work is dedicated to Almighty Allah, the lord of incomparable majesty, for His help, guidance and mercy throughout this program. And my parents, spouse and siblings for their support and encouragement.

ACKNOWLEDGEMENTS

Table of Contents

List of Tables

List of Figures

ABSTRACT

During the past few decades, the aquaculture industry has contributed immensely to food security in the face of the continuous growth in the global human population—and farming of Atlantic salmon has significantly contributed to this growth. Consequently, the rapid growth of the industry has led to intensification, which in turn has brought about increased disease incidences; with cardiomyopathy syndrome (CMS), a viral disease caused by piscine myocarditis virus (PMCV), being among the top diseases of economic importance in salmon farming in Norway. There are currently no known effective vaccines against the disease, and several studies have linked resistance against CMS to different variants of the major histocompatibility complex genes. However, no known functional study has pinpointed the mechanisms involved. Therefore, this study explored the sequence distribution and diversity in skin, spleen, heart, gill, liver and head kidney tissues and between infected and uninfected fish for the MHC-I-F10 gene. While no big differences were observed in the MHC-I-F10 gene in the different tissues, some insertions and deletions were only observed in the heart samples. Furthermore, the alignment of clean sequences from these tissues, except the heart, indicated a single nucleotide polymorphism at nt 533, which results in an amino acid change from cysteine to tryptophan. In contrast, the sequences from the heart showed messy chromatograms with double peaks in proximity to the N terminus. Therefore, the heart samples were sequenced with both forward and reverse primers—where the latter returned clean and contrasting results to the former. Moreover, we have developed CRISPR-mediated editing of the MHC gene in the TO cells, a permanent cell line of Atlantic salmon origin, via transfection and electroporation methods and demonstrated the possibility of successful editing of the MHC-I-F10 gene in TO cells, which lay the foundation for further functional studies of this gene related to PMCV infection. The editing efficiency was low which might have emanated from many factors. The most significant one is the short duration set aside for these studies, which has

influenced the overall results. However, the foundation laid for editing of the MHC-I-F10 gene in this study can serve as a component of a toolbox for future functional genetics and immunological studies in Atlantic salmon.

Chapter 1 Introduction

The world human population is projected to be over 9 billion by the year 2050 (United Nations, 2019), and this increase needs robust strategies to meet the population's nutritional requirements. One of such requirements is the daily intake of animal proteins of about 20 g per person per day (FAO, 2014). As a result, livestock had been implicated to be contributors to the global climate change with the emission of greenhouse gases either directly from the enteric fermentation or indirectly, through activities used to feed them (Cassandro, 2020). Therefore, the need to look at other sources to meet the animal protein requirement without jeopardising the survival and sustainability of the human population has been eminent. One important sector that can contribute to reducing the greenhouse gas effect is aquaculture. Aquaculture has consequently become increasingly important as a sector that contributes to global food security and income generation as it accounts for 46 percent of the total output and 52 percent of the consumption of fish by humans (FAO, 2020). About 17 percent of the global intake of animal proteins came from fish consumption in 2017 (FAO, 2020). One important fish species contributing to the growth of aquaculture is the farmed Atlantic salmon (*Salmo salar* L.). Atlantic salmon is a salmonid fish species (including trout, char, grayling and others) characterised by a dorsal, adipose fin posterior to the main dorsal fin and a laterally compressed body (Verspoor et al., 2007).

Since 2013, salmonids have been the most valuable commodity traded, accounting for over 19 percent of the entire value of internationally traded fish products in 2018 (FAO, 2020). They are important contributors to the improvement in the aquaculture sector and the continuous increase in the world's per capita fish consumption. The farmed Atlantic salmon has become a flexible and popular seafood item that matches trends in modern consumer preferences. Therefore, it is not strange that the farmed Atlantic salmon has become the most significant single fish product by value for several years now, as it makes up the largest portion of the

global export revenue compared to other salmonids (FAO, 2020). The industry is known for its well-coordinated international marketing strategies and swift product innovation. All these and more have contributed to the success of the Atlantic salmon aquaculture, as the industry prides itself among the world's most lucrative and technologically advanced fish culture industries. The leading countries at the core of these achievements for the Atlantic salmon culture are Norway, closely by Chile, UK and Canada. These countries account for the chief proportion of farmed Atlantic salmon sold worldwide.

In addition, the salmonid industry has benefited generously from the advancement in selective animal breeding, using the application of recent tools and techniques, such as markers assisted selection, genomic selection and prediction of breeding values of selection candidates—which are made parents and founders for the next generations (Yáñez et al., 2014).

However, despite these advancements and successes, the industry's intensification has resulted in a series of infectious disease outbreaks, threatening the profitability and sustainability of the industry. These diseases impact health and welfare, with substantial economic losses and environmental consequences such as cross-species disease transmission to wild fauna—with the widespread use of antibiotics and chemicals in disease control efforts (Barrett et al., 2020). One of the most critical diseases bedevilling the farmed Atlantic salmon industry is the Cardiomyopathy Syndrome (CMS) (Sommerset et al., 2020).

1.1 Cardiomyopathy Syndrome (CMS)

Cardiomyopathy syndrome (CMS) has been described as a severe cardiac disease that occurs in the grow-out sea phase of farmed Atlantic salmon. Moreso, field reports have indicated the manifestation of CMS in younger fish with a time range between five to six months after sea transfer (Fritsvold et al., 2021; Wiik-Nielsen et al., 2016). As the largest producer of farmed Atlantic salmon, Norway has witnessed approximately 100 outbreaks of CMS annually (Su et

al., 2021a). The disease has consequently caused a lot of losses in the industry due to the incessant outbreaks that threaten its prosperity and sustainability. Due to the economic importance of this disease, several research efforts have been geared toward understanding the underlying mechanisms of pathogenicity. This has eventually led to the identification and characterisation of the causative agent, the Piscine Myocarditis Virus (PMCV) (Haugland et al., 2011).

Currently, there is no known effective vaccine available to control CMS—this is partly due to the difficulty of culturing PMCV in cell lines. In the absence of vaccines, strategies to control viral infections rely mainly on stimulating and potentiating the fish immune system to ward off the viral infection. Several studies have been conducted to assess the tolerance and resistance levels of the farmed Atlantic salmon to the PMCV infection. These studies can provide insights as to how to utilise several tools, such as selective breeding, genomic selection and many others, to enhance the resistance, survivability, and even phasing out the menace of the virus.

In previous studies, the spatio-temporal regulation of the different arms of immunity during CMS was characterised using transcriptome analysis of immune responses in fish, which developed the strongest pathology and infection (Timmerhaus et al., 2011). The peak of cardiac pathology and viral load coincided with a cardiac-specific induction of T cell response genes and a splenic induction of complement genes. These responses were activated before viral load and pathology were reduced, implying that they may be important for viral clearance and recovery. In another study, the resistance of Atlantic salmon to CMS was found to be heritable, with values ranging between 0.12 and 0.46 (Boison et al., 2019). Furthermore, some research groups have conducted a handful of genome-wide association studies (GWAS) to detect the genomic regions linked to resistance against CMS in Atlantic salmon. One of these studies observed some regions in the salmon genome that may contribute to resistance against CMS (Boison et al., 2019). In addition, another study has confirmed the presence of major

quantitative trait loci (QTLs) detected from previous GWAS studies on the resistance of farmed salmon to PMCV infection (Hillestad et al., 2020). The regions implicated in these studies, the major histocompatibility complex (MHC), contain genes involved in the induction of adaptive immunity against viral and bacterial diseases.

1.2 Major Histocompatibility Complex (MHC)

The MHC is an important molecule that plays a crucial role in the generation of adaptive responses. Professional antigen-presenting cells (APC) present foreign peptides in association with MHC to T cells, leading to the activation of the specific immune responses (Dionne et al., 2009). There are two types of MHC in vertebrates, including the Atlantic salmon. These two are referred to as classes I and II MHC. The MHC class I molecules are expressed by all nucleus-containing cells. On the other hand, the MHC class II molecules are explicitly produced by APCs such as the dendritic cells, macrophages and B cells. Both classes of the MHC are essential in the adaptive immune system of Atlantic salmon. Therefore, in-depth knowledge of the underlying mechanism of the MHC, with the integration of various breeding and molecular tools, could be pivotal in combating the PMCV infection.

1.3 Modern tools in enhancing disease resistance in Aquaculture

As mentioned before, disease outbreaks inflict significant economic losses and affect the welfare of the fish in the aquaculture industry. About 15% of the global production value of cultured shellfish and fish accounts for the losses—with nearly 9 billion USD annually (Zilong Tan, 2007). In light of these losses and accompanying welfare problems, myriads of control measures have been devised to combat infectious diseases in the aquaculture industry. These include management practices and biosecurity measures, grouped into treatment and preventive measures (Lafferty et al., 2015). However, the incessant use of chemicals with these

measures, such as vitamins, disinfectants, antibiotics, and others, have directly or indirectly contributed to environmental problems and the ever-increasing antibiotic resistance in bacteria (Karunasagar et al., 1994). In addition, there had been the use of probiotics—non-harmful bacteria for protection against infection.

As a result of the challenges, high costs, and the transient effects of the measures mentioned above, selective breeding has been proposed and used as a natural and sustainable means of improving resistance to diseases in fish (Gjerde et al., 2004; Gjerde, 2006). Selective breeding methods have recorded immense successes in various livestock and aquaculture species and have helped in improving various economic traits of interest.

Selective breeding methods have gone through phases since their inception; from the mass selection of animals to selection based on relatives (family selection), the use of molecular markers, known as markers assisted selection (MAS) and the recent, advanced use of the genome-wide markers, popularly known as genomic selection (GS), based on single nucleotide polymorphism (SNPs). Most of the breeding programs that consider disease resistance in Atlantic salmon have been using family selection approach with relatively slow progress (long generation interval). This is mainly because the estimated breeding values (EBVs) of selection candidates were achieved with low accuracy due to the dearth of information from the selection candidates—as they are not used in challenge tests (Falconer and Mackay, 1996). However, the advent of new technologies, such as the next-generation sequencing and high throughput genotyping platforms, has heralded a new era at the frontiers of breeding for disease resistance in salmonid species. This is because of the unravelling of the regions in the genome, known as genetic variants, that influence resistance against infectious diseases (Goddard & Hayes, 2009). With these technologies, different methods have been used to obtain records on the traits of interest and the accompanying genetic variability among the fish. In Atlantic salmon and other salmonids, these methods have been used in several studies (Fevolden et al., 1999; Fevolden

& Roedb, 1990; Gjedrem & Gjøen, 1995; Gjedrem et al., 1991; Houston et al., 2010; Wetten et al., 2007) to study immunological and physiological parameters associated with specific infectious diseases in the field and in challenge experiments, where fish are exposed to pathogens in controlled environments. For example, studies on PMCV resistance in Atlantic salmon have resulted in detection of genetic variants underlying resistance to the virus (Boison et al., 2019; Hillestad et al., 2020). Heritability estimates, pointing to significant genetic variation in susceptibility to CMS were detected in several populations, giving room for genetic improvement of resistance to the disease, and of note, genes clustering in the MHC-region were found to be associated with resistance to infection. However, functional validation of the importance of these genes is required and could be done using the novel genome-editing tool such as CRISPR cas9.

1.4 Objectives of the study

As a result of the importance of the CMS disease in the Atlantic salmon industry, several research efforts have been geared towards understanding the underlying molecular mechanism and how the immune system build resistance against infection / disease. In their study in 2019, Boison and colleagues detected some QTLs associated with resistance to CMS infection in Atlantic salmon. In the same year, a study by Hillestad and others observed the same QTLs. After that, another study (Hillestad et al., 2020) confirmed the presence of the QTLs in association with resistance to CMS disease. In addition, these studies detected part of MHC genes associated with resistance to infection. There have been no known functional studies to assess the genes implicated therein. Consequently, this study aims to understand the distribution and sequence diversity of the MCH-1-F10 genes (LOC106588401 ssa27; 10508991- 10512119, LOC106588402 ssa27; 10122008- 10149393) in different tissues of the Atlantic salmon. In addition, the study also aims to establish and develop tool to study the functional importance of these genes using the novel genome editing tool approach in cell lines.

Chapter 2 Literature Review

2.1 Overview of Atlantic salmon (*Salmo salar*)

Atlantic salmon is part of the fish family Salmonidae, an essential part of the Teleosts—the modern bony fishes with the presence of a homocercal tail (Nelson, 1984). The family consists of about seven genera: Salmo (the Atlantic and trout species) and others (Hucho, Oncorhynchus, Salvelinus, Salmothymus Brachymystax, and Salvethymus). They are characterised by possessing a laterally compressed body with a dorsal, adipose fin posterior to the main dorsal fin (Verspoor et al., 2007). In addition, they are primarily found in temperate water with characteristic complex biology—spending their early life in freshwater and their later life in the sea (Storebakken, 2009).

https://aquatechcenter.com/species/atlantic-salmon/

Figure 1. Atlantic salmon (Salmo salar), by Aquatechc enter, CC BY 4.0.

2.2 Atlantic salmon culture

The contribution of farmed Atlantic salmon to ensure global food security is due to the intensification in the industry and the substantial contributions from the major producers like Norway, Chile, Canada, and the UK (Boudry et al., 2021). During the 1970s, the aquaculture of Atlantic salmon kicked off in Norway (Boudry et al., 2021) and has since become the most valued fish species in recent times (FAO, 2020). As has been indicated earlier, the lifecycle of farmed Atlantic salmon (as well as the wild) begins in the freshwater, where they are raised in pens in freshwater lakes (Chile/UK), on land-based hatcheries running on flow-through or recirculating water systems (most other salmon producing countries). Afterwards, they go through different growth phases and are transferred into the seawater for on-growth to reach harvest size (3-6kg). These different phases and practices in farmed Atlantic salmon include egg and milt collection, fertilisation, incubation, hatching, transfer to the sea and slaughter when fish reach marketable size.

Source:https://thefishsite.com/articles/cultured-aquatic-species-atlantic-salmon
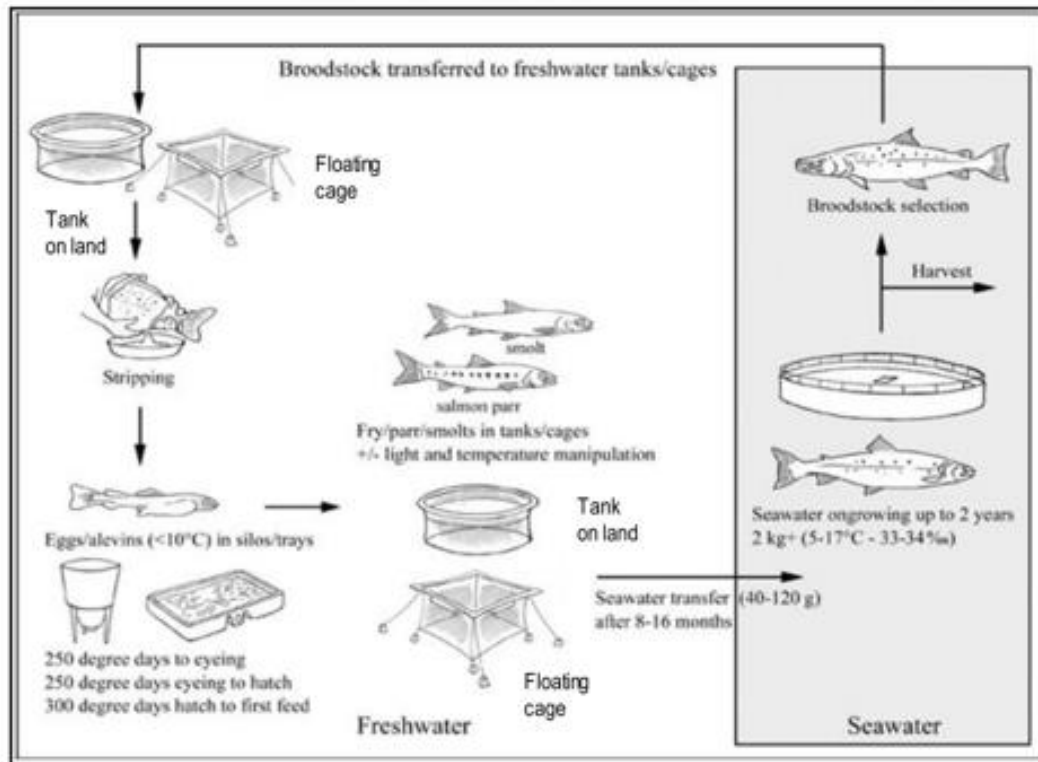
Figure 2. Atlantic salmon production, 2010, by the fish site. (https://thefishsite.com/articles/cultured-aquatic-species-atlantic-salmon). CC-BY 2.0

Spawn

Brood – Parr - Smolt

Transfer to sea

Growth phase in sea

Slaughtering
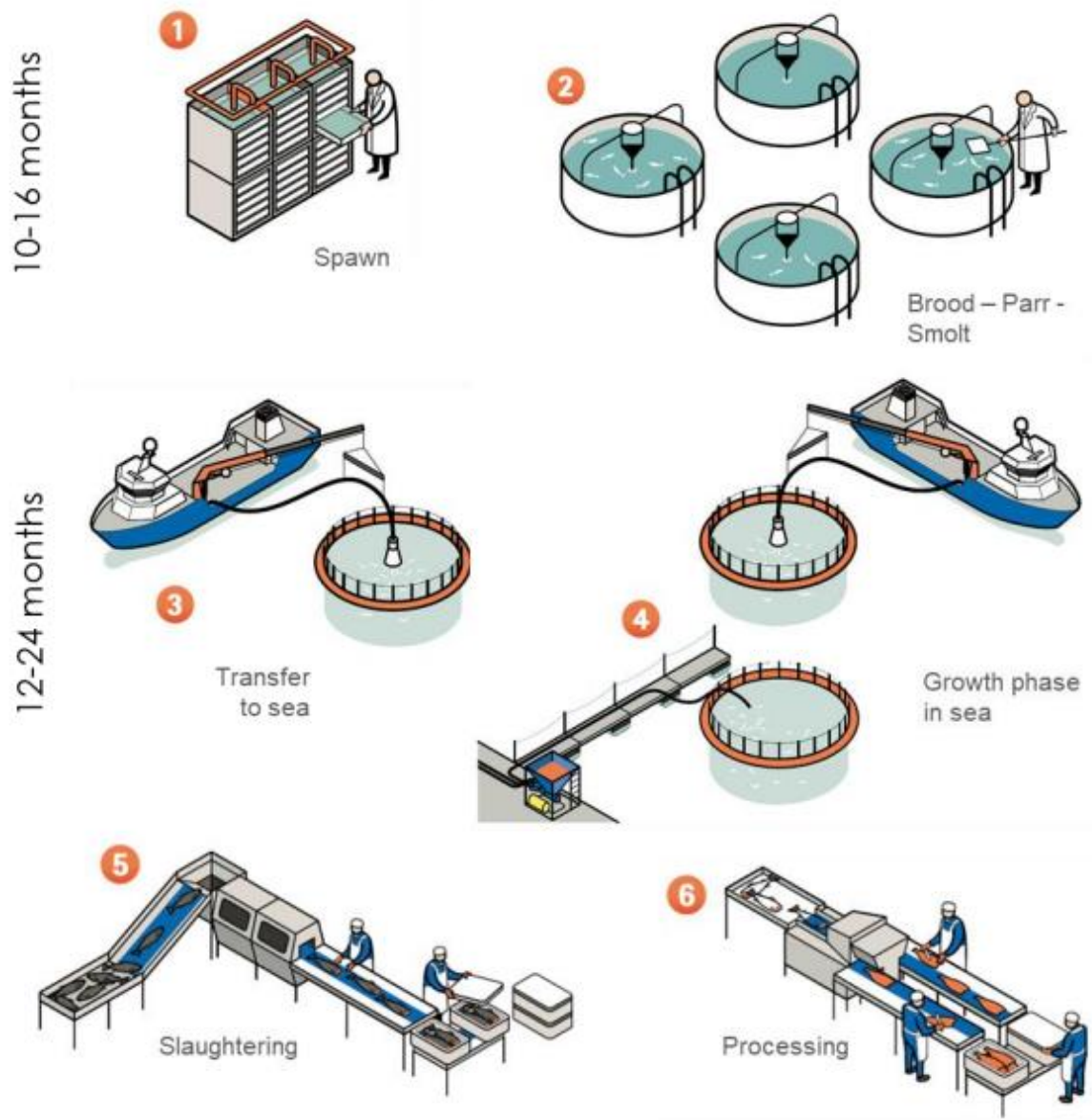
Processing

Figure 3. Atlantic salmon production cycle, including processing. From *Salmon Farming Industry Handbook*, 2020, by MOWI. Reprinted with permission.

2.2.1 Broodstock breeding and rearing

Norway began its first Atlantic salmon broodstock programs in 1971 by spawning wild Atlantic salmon from several rivers (Storebakken, 2009). This has since been advanced using various selective breeding methods and other modern molecular technologies, including family selection and mass selection, marker-assisted selection, and the recent genomic selection. Using these different methods, parental broodstocks have been selected and are kept either in the freshwater for their entire life, in sea water cages, or a combination. The age range of the broodstock is between three to seven years. However, care is usually taken to avoid using old fish due to the generation interval and inherent costs incurred (Storebakken, 2009).

2.2.2 Stripping and Fertilisation and Hatchery

Shorter photoperiod and dropping water temperature induce spawning; the timing and the relative change in daylight hours are crucial in stimulating the oogenesis, vitellogenesis, as well as gamete maturation and release (Storebakken, 2009). Furthermore, with technological advancement in the aquaculture industry spawning is synchronized through light, water temperature and hormones to make fertilized eggs available throughout the year. Stripping refers to the removal or harvesting of eggs or milt from the female or male broodstock, respectively. Before stripping, the male and female are kept in conducive conditions without disturbance. The collected eggs and milt are mixed, with the addition of water, to bring about fertilisation (Vassvik, 2000). In the hatchery, the incubation period is dependent on the prevailing water temperature. The recommended water temperature is around 8 $^{\circ}$C, and hatching takes about 60 days (Vassvik, 2000).

2.2.3 Incubation

During incubation, the fertilised eggs are left untouched due to their delicate conditions but can subsequently be handled after reaching the eyed egg stage—which takes about 220-250 degree days (Storebakken, 2009). The degree days is calculated as the product of the number of days in incubation with the prevailing temperature at that particular time (Vassvik, 2000). After that, the dead eggs are sorted from the live ones, as the two black dots can be seen in the live eggs.

2.2.4 Alevin Stage

The succeeding stage after hatching is known as the alevin stage. In this stage, the young fish possess an attached yolk on the underside, serving as a reservoir of food for continuous nourishment until their first feeding. The duration of the stage is similarly contingent on the prevailing water temperature. This stage lasts for some weeks as the nutrients in the yolk is used for growth and development with minimum movement.

2.2.5 First Feeding

This happens when most of the attached yolk in the previous stage is almost used up and is subsequently called the fry. Here, the fries are transferred from the hatchery into fish tanks. The fries are much more active as they search for food. They are provided with feed particles commensurate with their feeding stage as they grow during this phase (Vassvik, 2000). Various stocking densities are used during this stage, and as large as 10 000 fish per square meter with a water temperature range of 10-14 $^{\circ}$C can be used (Vassvik, 2000). This period takes about six weeks, after which fry are sorted into bigger tanks according to their various sizes. The light conditions and the water temperature are essential for the development of the salmon parr—an advanced fry with several dark parallel lines on the back, into smolt. Subsequently, the sorted fish are vaccinated against diseases before transferring them to the sea.

2.2.6 Smolting

Smolting can be described as the processes and series of changes that transform young Atlantic salmon parr in the freshwater to life in seawater. These changes are behavioural, morphological and physiological in nature and prepare them for seawater entry (Mccormick, 2013). Increased salinity tolerance with accompanied alteration in kidney, gill and gut function has been described as one of the most essential and amenable changes during this phase (Mccormick, 2013). Other changes include the disappearance of the parr marks and the emergence of silvery colours on the side and underside of the body. Furthermore, some visible changes occur in with the gills as the parr finally advances to a smolt (Vassvik, 2000). During this phase, the smolts are sometimes gradually acclimatised to life in the sea before final transfer. They are usually about one year in age before transfer to the sea, but can be transferred earlier using light manipulation.

2.2.7 Smolt Transfer

The smolts are transferred to the sea by boats and trucks. In addition, the storage area of these trucks and boats is filled with seawater to make the availability of fresh seawater to the smolt easy.

2.2.8 Grow-out

The remaining life of farmed Atlantic salmon is spent in the sea pens or cages up until harvesting. This period is usually referred to as the grow-out phase. In this phase, the fish face several challenges as they live in a new environment with less control from the farmers. One of these challenges is the outbreak of diseases. However, due to the vaccination before the sea phase, fish are protected against many diseases. Nevertheless, effective vaccines are yet to develop against some diseases, which continue to cause significant losses to the salmon industry. One of these diseases is the CMS, caused by PMCV, which is known to adversely

affect the health and welfare of farmed Atlantic salmon in the second year of the grow-out phase (Svendsen et al., 2019).

2.3 Piscine Myocarditis Virus: Bane of farmed Atlantic salmon

The recent intensification and expansion of global trade in the aquaculture industry have led to incidences of diseases that have impacted the health and welfare of the fish and caused economic losses. PMCV is a small unencapsulated icosahedral virus that possesses a double-stranded RNA (dsRNA) genome with three open reading frames (ORFs) (Haugland et al., 2011). Two of the three ORFs encode for presumptive proteins that may be part of the 91 kDa capsid and an RNA-dependent RNA polymerase. In addition, the third ORF in the PMCV genome encodes a putative toxin, which may be equivalent to the killer toxins whose genes carried in satellite virus particles (Dunn et al., 2013). The disease has been observed in both farmed and wild Atlantic salmon that have been exposed to PMCV (Su et al., 2021b). This virus was tentatively assigned to the Totiviridae family (Haugland et al., 2011; Løvoll et al., 2010). In general totiviruses have been mainly linked with latent infections in fungi and protozoan. that is transmitted through cell division, sporogenesis or cell fusion (Adams et al., 2014). However, PMCV differs from other Totiviridae in several ways, including the choice of a vertebrate host and the mode of extracellular transmission (Haugland et al., 2011).

Even though CMS was first discovered in 1985 (Amin & Trasti, 1988), knowledge about the pathogenic mechanisms and immune responses required for protection are lacking. The disease is now widespread along the Norwegian coast and a few other countries where salmon farming occurs. Studies using quantitative polymerase chain reaction (qPCR or real-time PCR) analysis showed a strong correlation between virus levels in the heart and the severity of cardiac lesions (Haugland et al., 2011; Timmerhaus et al., 2011). PMCV's biophysical properties are still not known and have proven to be a hard nut to crack in cell culture experiments. As a result, developing a vaccine against the virus has been difficult, and there is currently no known

effective vaccine on the market. Therefore, protection against the disease is mainly through management practices and biosecurity measures, leaving the Atlantic salmon to use its immune system to ward off the infection by the PMCV when an outbreak occurs. Recent studies have implicated some genomic regions associated with the major histocompatibility complex (MHC) in Atlantic salmon to be associated with resistance to the CMS (Boison et al., 2019; Hillestad et al., 2020).

2.4 MHC: An essential pillar in vertebrates' immune systems

The histocompatibility antigen (or H2-locus) was initially defined as a gene with a significant effect on tissue rejection in mice. However, when it became clear that several closely linked genes influenced rejection, the region was dubbed the major histocompatibility complex or MHC for short (Grimholt, 2016). The MHC has been found to consist of primarily two classes of molecules—MHC class I and II. The bid to get a clear picture of how the MHC molecules work took some few decades after their discovery. The MHC classes I and II were observed to comprise proteases and MHC-like chaperones that aid them in presenting peptides to the immune system (Rock et al., 2016). These were only made possible due to concerted efforts and several studies in the field to understand the MHC molecules. Therefore, many details have been revealed in recent years, providing new insights into their control and specificity (Neefjes et al., 2011).

The MHC classes I and II are similar and, at the same time, different from each other in their features and mode of antigen presentation. Therefore, the antigen presentation mechanism utilised by MHC genes to combat invading viruses, bacteria, and other infectious agents necessitates a deeper comprehension of the underlying processes involved. The MHC class I molecules present intracellular peptides—mostly of cytosolic and nuclear origin, to CD8+ T

cells on the cell surface, while the class II molecules present exogenous peptides to CD4+ T cells on the cell surface (Vyas et al., 2008). In addition, MHC class I molecules consist of one alpha chain and β2-microglobulin—with class II having one alpha and one beta chain (Grimholt et al., 2003). Furthermore, while MHC class II is mainly present in antigen-presenting cells (APCs), which include the dendritic cells, macrophages, and B cells and are used to present peptides to activate T cells, the MHC I is present in all nucleated cells and present cellular peptides to the surveying immune cells (Rock et al., 2016). Consequently, cells that present foreign or pathogens associated peptides are identified and eliminated by activated T cytotoxic cells. The MHC class I gene is of interest to this study and will therefore be focused on.

2.4.1 Antigen Presentation by MHC class I

The understanding of antigen presentation by the MHC class I molecules of the adaptive immune system has been deeply enriched due to recent advances in the field. These advances contributed to the recent observations that revealed the complexities of various steps in the antigen presentation pathway of MHC class I molecules (Neefjes et al., 2011). Prior to antigen presentation, the peptides involved in this pathway are processed from various sources in the cell that include the defective proteins, known as defective ribosomal products (DRiPs) and other antigens degraded by the cytosolic and nuclear proteasomes (Neefjes et al., 2011; Rock et al., 2016). To access MHC class I molecules, the consequent peptides are then translocated to the endoplasmic reticulum (ER) by the transporter associated with antigen presentation (TAP). A polymorphic heavy chain and a light chain, known as β2-microglobulin (β2m) of the MHC class I are used in assembling heterodimer in the ER (Neefjes et al., 2011). To maintain stability, peptides of 8–9 amino acids must enter the MHC class I peptide-binding groove, which can only accommodate peptides of this length. However, without the peptides, the stability of the MHC class I molecules is ensured by chaperone proteins of the ER, such as the

protein disulphide isomerase (PDI), calreticulin, and the dedicated chaperone, tapasin (Neefjes

et al., 2011; Rock et al., 2016).
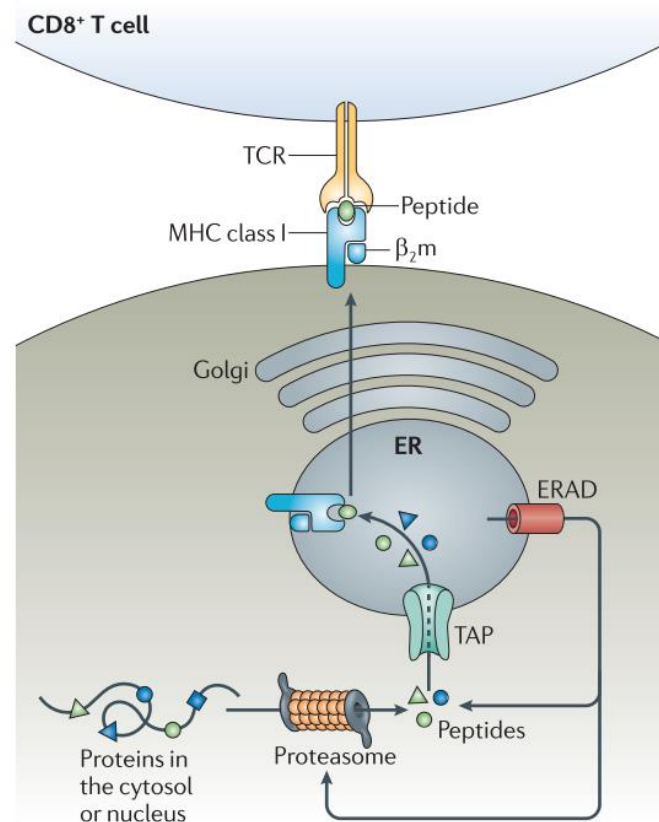
Figure 4. Basic depiction of MHC class I antigen presentation. A series of events are involved in the presentation of antigens in the intracellular spaces. Foremost, degradation of antigen proteins in the proteasome and the transportation of the ensuing peptides into the endoplasmic reticulum (ER) lumen via the transporter associated with antigen presentation (TAP) and loaded onto MHC class I molecules. The ER releases peptide–MHC class I complexes, which are then transported to the plasma membrane by the Golgi apparatus and used to present antigen to CD8+ T cells. From *Towards a systems understanding of MHC class I and MHC class II antigen presentation*, by Neefjes et al., 2016, *Nature Reviews Immunology*, 11(12), p.824. Reprinted with permission.

After loading the peptides onto the MHC I molecule in the ER, the peptide-MHC class I

complexes leave the ER to be presented on the cell surface (Vyas, 2008). On the contrary,

peptides and MHC class I molecules that cannot bind in the ER are returned to the cytosol for

degradation. In summary, the antigen presentation pathway of the MHC class I molecules can be divided into three main stages; antigen processing, the peptide-MHC class I complexes formation with modifications and antigen presentation.

<h3>2.4.2 MHC classes, polymorphism, and disease resistance in Atlantic salmon</h3>

Marsh et at., 2000, described the MHC genes as the most polymorphic genes known to date, with multiple loci and a large number of alleles at each locus. The proteins responsible for this characteristic are the human leucocyte antigens (HLAs). They are expressed in three highly polymorphic genomic regions in MHC class I (HLA-A, HLA-B, HLA-C) and MHC class II (HLA-DR, HLA-DP, HLA-DQ). The resulting allelic variability primarily influences the nature and constitution of the peptide-binding groove and, consequently, the peptide or antigen fragment repertoire displayed on the surface (Wieczorek et al., 2017). This diversity enables each individual to bind and present a diverse array of peptide ligands, which has direct functional implications for immune responses. Heterozygous advantage, over-dominant selection and frequency-dependent or balancing selection have all been proposed as possible explanations for the MHC gene's extensive polymorphism (Parham & Ohta, 1996).

In Atlantic salmon, previous studies have shown that genotypes of MHC classes I and II confer resistance to infectious diseases. For example, Langefors et al., 2001, established an association between MHC class II B and resistance to *Aeromonas salmonicida*. In the same vein, there were associations between MHC class I and class II A and resistance to Aeromonas salmonicida and Infectious salmon anaemia virus (ISAV) (Grimholt et al., 2003). Nevertheless, Miller and colleagues (2004) reported that the MHC class I and class II loci are associated with susceptibility to infectious haematopoietic necrosis virus (IHNV). A similar association has recently been found between the MHC class I genes and the CMS in farmed Atlantic salmon

(Boison et al., 2019; Hillestad et al., 2020; Hillestad & Moghadam, 2019). From animal breeding and genetics perspective, these associations between MHC genes and infectious diseases in salmon could be exploited further to develop selective breeding programmes that increase the fish population's resistance to these diseases and minimise their impact. A further investigation into the underlying mechanisms of this association is therefore warranted. One of the ways to study this association is the use of high-density single nucleotide polymorphism (SNP) chips in genotyping both challenged fish and their relatives, known as the selection candidates, for the QTL associated with the disease resistance—and making use of these groups of fish as parents of the next generation, a method known as genomic selection (GS). Furthermore, due to the rapid advancement in molecular techniques and tools, there are many other opportunities to improve resistance against infectious diseases—one of which is through novel genome editing technology.

2.5 Genome Editing Tools

The capacity to alter the building blocks of living organisms—the DNA bases, is extremely valuable in molecular biology, agriculture, medicine, and the general field of life sciences. Subsequently, a long-held goal in molecular biology is the ability to make desired changes to the genomes which came with homologous recombination (Adli, 2018; Capecchi, 1989). In addition, the unprecedented power to characterise the functional roles of various genes in model organisms provided by such targeted gene integration into the genome was revolutionary. However, this approach had some drawbacks, including the significantly low rate of integration of exogenous DNA copy (Capecchi, 1989). Furthermore, the rate of integration varied according to the cell type and state. Finally, the genome editing approach may result in arbitrary integration of the exogenous DNA into undesirable genomic loci at a rate comparable to or greater than that of the genomic area of interest (F. L. Lin et al., 1985).

Several research efforts have been geared to provide alternative approaches to overcome the abovementioned problems. An early breakthrough was the discovery that introducing a double-stranded break (DSB) at a target site increases the frequency of targeted gene integration by several manifolds. One of the approaches used in introducing the DSBs in the genome is through the use of uncommon cutting endonuclease enzymes—examples of which include an 18 base pairs I-Scel cutter that introduce the DSBs into the mouse genome (Rouet et al., 1994). However, there were two shortcomings of this method: the uniqueness of the meganucleases to recognise sequences made it challenging and lowered the chances of finding a suitable one for a locus of interest, and the use of the error-prone non-homologous end joining (NHEJ) in repairing the DSBs in the target area of interest (Adli, 2018). Thereafter, the discovery of zinc finger nucleases (ZFN) and transcription activator-like effector nucleases (TALEN) brought enormous improvement to the frontiers of genome editing due to their improved capabilities as compared to previous tools. For ZFN, they recognise three base pairs of the target area of interest in the genome, while the TALEN can precisely target and edit one base pair. Despite their stellar performances and advances they brought with applications in therapeutics and other areas of life sciences, re-designing or re-engineering of a new set of proteins was required to target different sites in the genome. ZFNs and TALENs have not been widely adopted by the scientific community due in part to their difficulty in cloning and protein engineering (Adli, 2018).

Subsequently, the advent of the novel clustered regularly interspaced short palindromic repeats (CRISPR) has transformed the field because its editing capabilities are on par with, if not superior to, the existing tools. CRISPR, with its associated nuclease, Cas, have been observed to be initially part of the prokaryotic adaptive immune systems against invading pathogens such as the phages— by binding and cutting foreign nucleic acids (Barrangou et al., 2007; Brouns et al., 2008; Marraffini & Sontheimer, 2008). In the molecular biology community, the

most commonly used type II CRISPR system comprises two parts: Cas9 nuclease and synthetic

single-guide RNA (sgRNA), which is a merger of a CRISPR RNA (crRNA) and trans-

activating crRNA (tracrRNA) (Deltcheva et al., 2011; Jinek et al., 2012). When Cas9-sgRNA

recognises an NGG—where N might be A, T, C, or G, proto- spacer-adjacent motif (PAM)

sequence, the sgRNA spacer pairs with the target DNA strand to form an R-loop. The Cas9

nuclease cleaves the DNA strands and creates a blunt-end DSB 3 bp upstream of the PAM

(Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013).
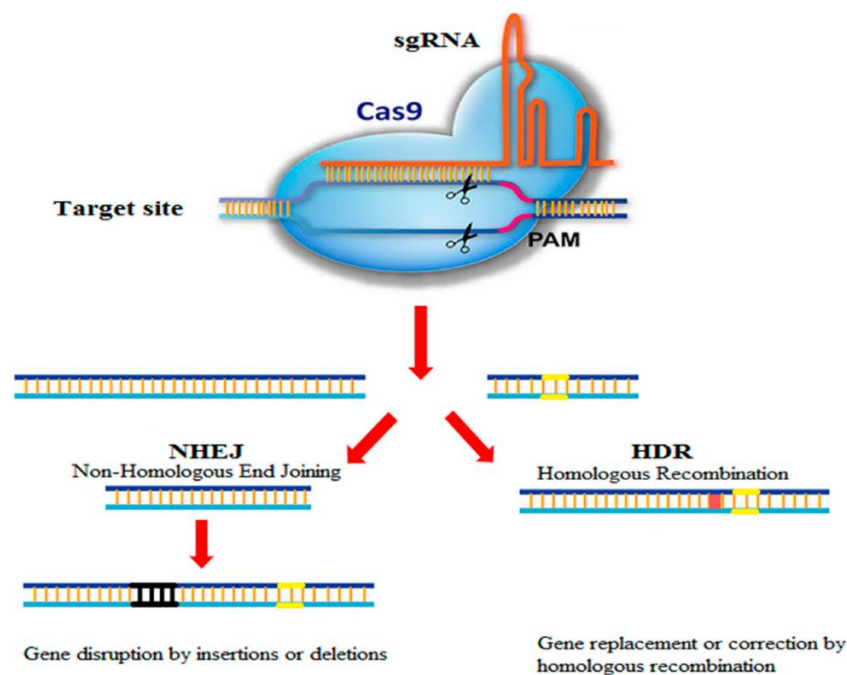
Figure 5 above depicts and sums up the typical process of CRISPR-mediated editing via non-homologous end-joining and homology-directed repair. Firstly, the guide (sgRNA) sequences attach to the target sequence, which is about 20 nucleotides away from the PAM (red) and precisely cut through the site. The disruption caused by this incident is repaired either by NHEJ, which causes insertion or deletion or by HDR, by adding a donor template. From *Applications of CRISPR-Cas9 as an advanced genome editing system in life sciences,* by Tavakoli et al., 2021, *BioTech*, 10(3), p.3. CC-BY 4.0

In aquaculture, the CRISPR Cas system has many potentials to improve disease resistance—

one of which is the fixation of alleles at existing loci. This is done through , removal, promotion

or fixation of these functional alleles of interest, segregating within the broodstock at one or

many loci in the population (Gratacap, Wargelius, et al., 2019). This could fast-track the rate of genetic improvement with disease resistance in Atlantic salmon. Also, beneficial variants from distinct populations, species, or strains could be introduced to improve uncommon traits through introgression-by-editing (Gratacap, Wargelius, et al., 2019). This method has been described as having prospects in providing new opportunities to avoid traditional introgression and the accompanying drawbacks of linkage drag— and it permits access to genetic variation in other strains and species, which would not be possible with traditional selective breeding techniques (Gratacap, Wargelius, et al., 2019). Finally, the novel CRISPR Cas technology could also be of immense help in improving disease resistance against diseases, including the CMS, by generating and employing de novo beneficial alleles that are unknown elsewhere. In this context, using the CRISPR Cas system, novel alleles can be created based on prior knowledge of the biology of the trait of interest or from approaches that use genome-wide genetic perturbation to identify candidate genes that influence the trait of interest (Gratacap, Wargelius, et al., 2019).

Chapter 3 Materials and Methods

The experiments were carried out at the Lab Nord at the *Host-pathogen interactions in infections and immunity* group, Department of Paraclinical Sciences, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Norway. In addition, the complementary deoxyribonucleic acid (cDNA) of TO cells and other cells used in this study was acquired from other experiments involving RNA extraction from different tissues of farmed Atlantic salmon. The cDNA was then further used for other downstream methods in the study.

3.1 Cell culture

3.1.1 Cell line

The cell line used in this study was the TO cell line developed from the head kidney leucocytes of the Atlantic salmon (Wergeland & Jakobsen, 2001). All cells were grown as a monolayer in L15 media (Gibco, Thermofisher) and supplemented with the heat-inactivated foetal serum (FBS) (Gibco, Waltham, USA). Cell incubation occurred at 20 °C without carbon dioxide (CO_2).

3.2 RNA isolation and cDNA synthesis

Total RNA was isolated from TO cells and different Atlantic salmon tissues. All the tissues, except the heart were obtained from fish with similar genetic backgrounds. On the other hand, the heart samples were obtained from fish with different backgrounds, including fish infected with PMCV. For TO cells, RNA was extracted using RNeasy Plus mini kit (Qiagen) as recommended by the manufacturer. On the other hand a combination of Trizol and RNeasy mini kit procedures was used as previously described (Guo et al., 2015). After extraction, the RNA concentrations were determined by spectrophotometry using Nanodrop ND1000 (Thermo Scientific). 250 ng TO cells RNA or 1µg tissue RNA was used for cDNA synthesis using Transcription First Strand cDNA Synthesis Kit (Roche) using both oligo (dT) and random hexamer primers, as recommended by the manufacturer.

3.3 Polymerase Chain Reactions (PCR)

To amplify the gene sequences and prepare them for downstream methods, such as sequencing, the cDNA from cell lines and different tissues, namely the heart, liver, spleen, head kidney, skin, and gill, were subjected to PCR reactions. Each PCR reaction, contained Phusion master mix, (Thermofisher Scientific) and the two genes' (LOC106588401 and LOC106588402) forward primers (5' GGATCGTATGTCTGCCTGGA 3', 5' ACGGGCTAACGACGATATGA 3') and the reverse primers (5' CCCGCTCGACCCAATTAAGG 3', 5' AGGTGTCTGTCCATGGATGC 3'), respectively. In addition, all PCR reactions were carried out with a program as follows; initial denaturation at 98°C for 30 secs, followed by 39 cycles at 98°C for 10 secs, then 64.2°C for 30 secs and 72°C for 1 minute. The final extension was set at 72°C for 10 minutes with a Biorad Thermocycler (model and version).

3.4 Gel Electrophoresis

After PCR amplification, the PCR products were examined using gel electrophoresis. 1.5% agarose was prepared in a corresponding volume of 1X TAE buffer. The mixture was then heated in a microwave until a clear solution was achieved. Afterwards, a SYBRsafe gel stain was added, and the gel was casted into mold with combs and left to solidify. After proper solidification, the gel mold was placed into the electrophoresis apparatus (Bio-Rad PowerPac Basic) and the coms removed. 1X TAE buffer was added to cover the gel completely. For the loading into the wells, 10 µl 1 kb DNA ladder was added to the first well in the gel, while a mixsture of 10 µl of samples and 2 µl loading dye was loaded into the remaining wells. The electrophoresis instrument is set to 90 volts of electric current at 400 mAh for 45 minutes. Finally, the visualisation of the bands and capturing of the gel picture was done with a UV box (Safe Imager^{TM}, Invitrogen) and an Azure c300 Fluorescence and Chemiluminescence Imager.

3.5 DNA purification from agarose gel

The obtained PCR products were excised from the gels and weighed with a digital analytical balance (Model: LPW-2103i, VWR). The PCR products were then purified with Qiagen QIAquick Gel Extraction Kit using a spin protocol recommended by the manufacturer. Briefly, about 3 volumes of the Buffer QG were added to the excised gels in the microcentrifuge tubes and incubated for 10 minutes at 50°C with intermittent vortexing every 2-3 minutes manufacturer instructions. This is done to ensure the homogenous dissolution of the gel with a clear yellowish colouration confirming it. The resulting solution was then transferred into a QIAquick spin column, placed into a 2 ml collecting tube, and centrifuged for 1 minute. All flow-throughs were discarded from the collecting tubes and 500 µl Buffer QG was added to the QIAquick spin column to remove all traces of agarose. At this stage, 750 µl Buffer PE was added to the QIAquick spin column and allowed to stand for 5 minutes before centrifuging for 1 minute. This is done to ensure that the resulting DNA is amenable to sequencing at the end of the extraction. All flow-throughs were discarded from the collecting tubes. After this, the QIAquick spin column was then placed into a clean 1.5 ml microcentrifuge tube and eluted by adding 40-50 µl Buffer EB (10 mM Tris.Cl, pH 8.5) or RNase free water carefully at the center of the QIAquick membrane and the columns were centrifuged for 1 minute. The DNA quality was checked with the Nanodrop (name and model), and samples were prepared and sent for Sanger sequencing at Eurofins Genomics. All centrifugation steps were carried out at 13 000 rpm in a tabletop microcentrifuge (name and model) at room temperature.

3.6 CRISPR Cas9 Genome Editing of TO Cells

To culture the required amount of cells for the experiment, old media were removed from the growing cells in the cell culture lab using the aspirator. After that, two rounds to washing with 10 ml of PBS buffer was performed in order to wash off any remaining media from the cells. The cells were then detached from the cell culture flask by trypsinisation and checked under the microscope to ensure successful trypsinisation. This was followed by adding about 3 ml of the media to stop the action of trypsin. Thereafter, cells were counted using the T-20 automated cell counter (Bio-Rad) and plated into a 24-well plate using a cell density of 50 000-90 000 per well.

3.6.1 Cas9 Ribonucleoprotein (RNP) complex assemblies

To test and optimise the Cas9 RNP in this study, the major histocompatibility complex (MHC) gene, MCH-1-F10 (NCBI accession LOC106588402 ssa27; 10122008- 10149393) of the Atlantic salmon was targeted with multiple single guide RNA (sgRNAs) through transfection of the TO cell line.

The sgRNAs were designed with the CRISPR Design Tool (Synthego Inc., Menlo Park, USA) and then ordered from Synthego Inc. (details of all sgRNAs are given in Table 1). The RNP complex was assembled as follows: the sgRNAs (1.5 nmol) tubes were briefly centrifuged to ensure the RNA pellets were collected at the bottom of the tubes and were carefully rehydrated using 15 µl nuclease-free 1X TE buffer to make 100 15 µM (100 pmol/ µl) stock solution. After that, the solution was vortexed for 30 seconds to ensure that the sgRNAs were fully resuspended and stored at -20°C. Afterwards, the working concentration was produced by diluting the rehydrated sgRNAs using nuclease-free water in a sterile microcentrifuge tube. For example, to make a 30 µM (30 pmol/ µl) working stock, 6 µl of 100 µM sgRNA is added to 14 µl of nuclease-free water. In this study, two conditions were used for the incorporation of the sgRNAs; one with the recommended volume (1259 ng/7.5 pmol) of each sgRNA (3X) and

the other with one-third of each sgRNA (1X) in each reaction. In order to form the RNP complex, 25 µl Opti-MEM^{TM} I Medium (Thermo Scientific), 2490 ng (7.5 pmol) TrueCut^{TM} Cas9 Protein (Thermo Scientific) and lastly, 2.5 µl Lipofectamine^{TM} Cas9 PLUS^{TM} Reagent (Thermo Scientific). The resulting solution was mixed and incubated for 5 minutes at room temperature to allow the formation of the RNP complex. Table 1 below shows the CRISPR effector, guide sequences and the primers used for the editing and onward sequencing of the products thereof.

Table 1. Guide RNA sequences and primers used for amplification and sequencing of MHC-I-F10 gene

Table 1 Guide RNA sequences and primers used for amplification and sequencing of MHC-I-F10 gene

CRISPR effector, target	Guide sequences (5'-3')	Primers (5'-3')	Size (bp)	Tm(°C)	PAM
spCas9, MHC-1-F10	TCTTTCATCCAGTCCTGTCG	F- ACGGGCTAACGACGATATGA	500	55	GGG
		R- GCGTTGGGGCGATACATTTC			
spCas9, MHC-1-F10	TGTTTCTATTCCAGTAATCA	F- ACGGGCTAACGACGATATGA	500	55	GGG
		R- GCGTTGGGGCGATACATTTC			
spCas9, MHC-1-F10	TGGGAAGTCTGGTATTCCTG	F- ACGGGCTAACGACGATATGA	500	55	AGG
		R- GCGTTGGGGCGATACATTTC			

F/R forward and reverse primers for the MHC-I-F10 gene.

3.6.2 Transfection Solution

The transfection solution was constituted in another clean microcentrifuge tube with 25 µl Opti-MEM^{TM} I Medium (Thermo Scientific) and 1.5 µl Lipofectamine^{TM} CRISPRMAX^{TM} Reagent (Thermo Scientific) and incubated for 1 minute at room temperature. The resulting solution was not allowed to sit for more than 3 minutes to ensure optimal transfection efficiency.

3.6.3 Transfection Complex and Incubation

The transfection solution was added to each of the 3X and 1X reactions from the RNP complex and mixed well by pipetting up and down. Afterwards, the mixed solution was incubated for 20 minutes at room temperature.

3.6.4 Cell Transfection

A 50 µl (per well) transfection complex was added to 4-9 x 10^4 cells in the 24-well plate on transfection and incubated for 48 hours in a 20 °C incubator. The transfection was performed either once, twice or three times, based on the experimental groups (Fig. 6).

Figure 6. Cell transfection with RNP complexes. Tubes 1 and 2 contain the RNP complex and the transfection solution, respectively. Addition of the two gives the RNP complex-transfection solution, or the transfection complex. This is followed by the addition of 50 µ µl of the transfection solution per well with about 50 000 cells plated in each well which are then subsequently incubated at 20 °C for 48 hours. Transfection was performed once for the wells labelled 1 and 2 in the plates, twice for 3 and 4 and three times for 5 and 6.

3.6.5 Electroporation Method

The transfection solution was added to each of the 3X and 1X reactions from the RNP complex and mixed well by pipetting up and down. Afterwards, the mixed solution was incubated for 20 minutes at room temperature. After incubation, the cells and the transfection complex were electroporated with the Neon system (Invitrogen, Carlsbad, USA) according to manufacturer's instructions. The mixture was electroporated with 10 µl tips at 1200 V, 1 pulse with 40 ms speed, and dispensed in 1 ml of fresh media on a 6-well plate.

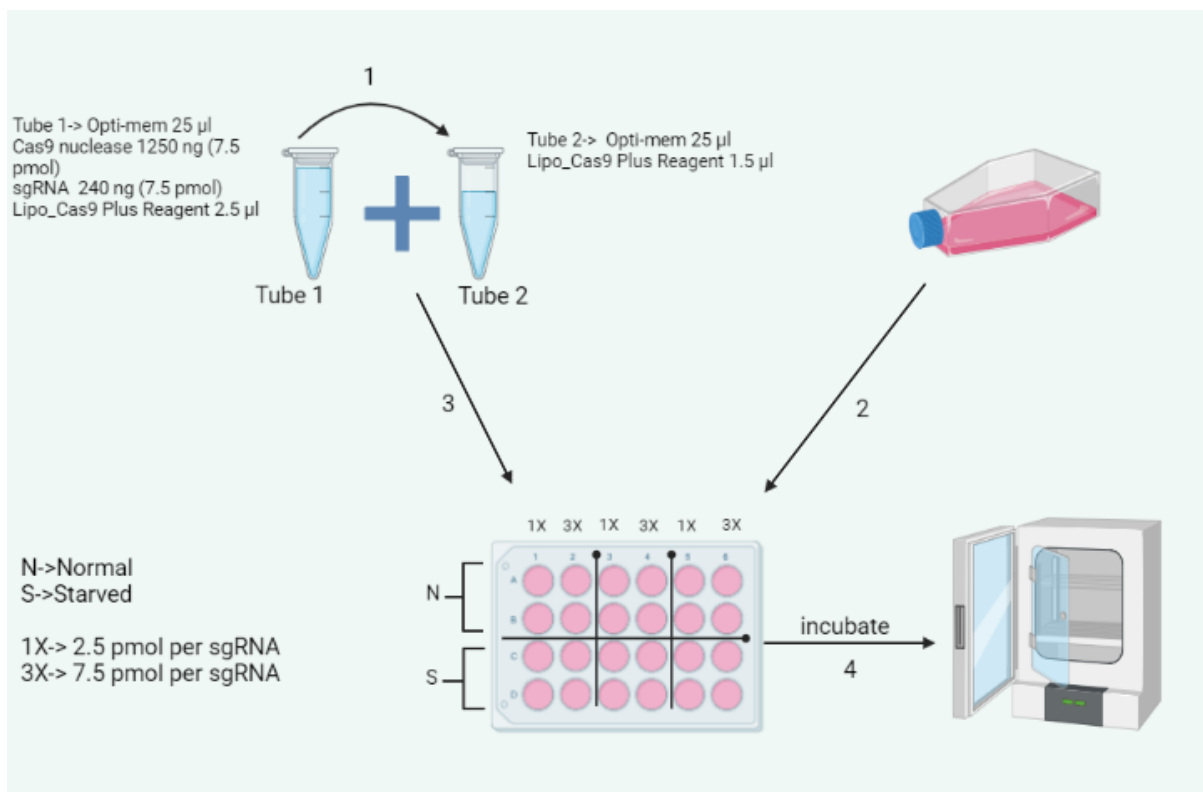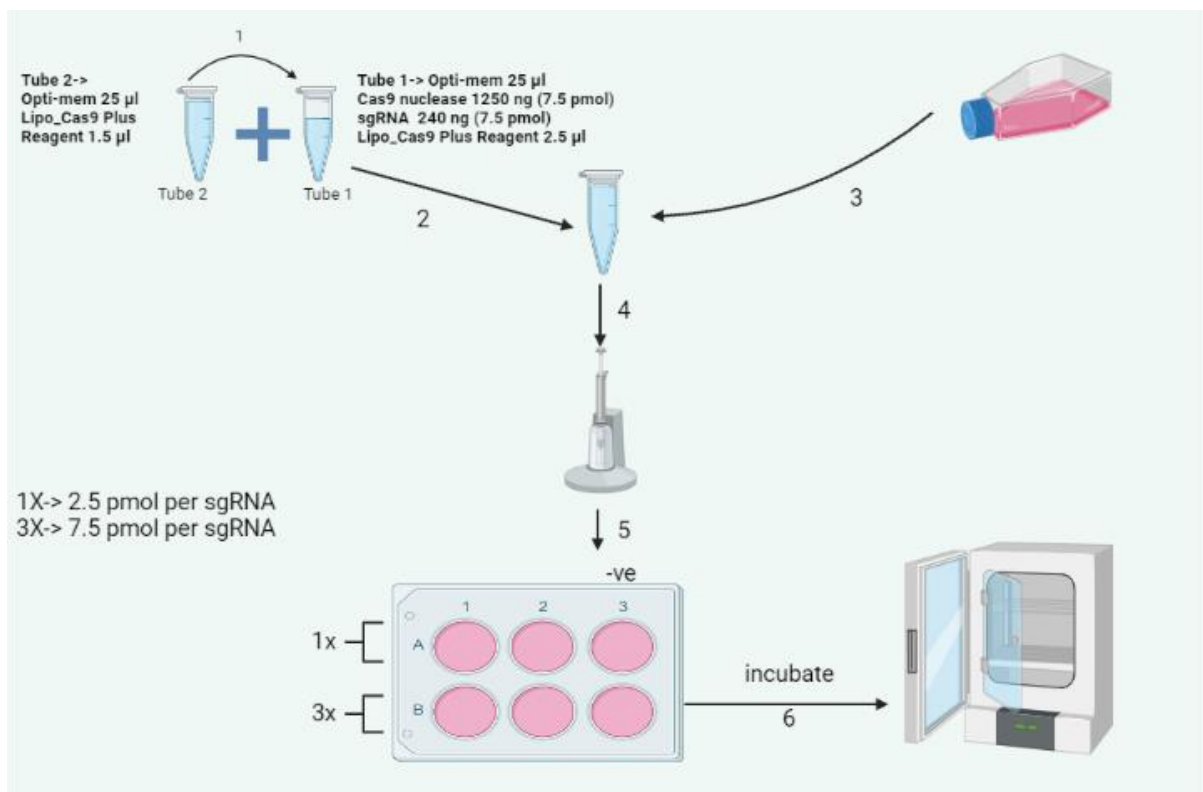Figure 7. Cell electroporation with RNP complexes. Tubes 1 and 2 contain the RNP complex and the transfection solution, respectively. Addition of the two gives the RNP complex-transfection solution, or the transfection complex. This is followed by the electroporation of the cells with the transfection complex and their respective addition to the wells and subsequently incubated at 20 °C for 48 hours.

3.7 Analysis

All sequencing results were analysed using the CLC Main Workbench (version 6.9.2) software for quality controls, alignment, and others. In addition, the products from both the lipofection and electroporation methods were amplified via PCR and then analysed using the cleavage assay with the GeneArt Genomic Cleavage Assay Kit (Thermo Scientific) and the Inference of CRISPR Edits (ICE) (Synthego) for sequenced products, to ascertain the editing efficiency.

Chapter 4 Results

4.1 Amplification of the MHC-1-F10 gene in different tissues and TO cell line

As indicated previously from past works (Boison et al., 2019; Hillestad et al., 2020; Hillestad & Moghadam, 2019), two genes have been found and confirmed to be associated with resistance against the PMCV infection. The genes include LOC106588401 and LOC106588402 from the major histocompatibility complex families, class I. Before performing the gene-editing studies, we needed to understand the presence of the different genes in the cell line and different tissues. Therefore, these genes were targeted for amplification through polymerase chain reactions (PCR) and electrophoresis with specially designed primers and the results were analysed. Only the LOC106588402 gene was successfully amplified for different tissues including the heart, liver, head kidney, gill, skin, and spleen, and from the TO cell lines. Below are the results from the tissues and the cell line. Amplification of the MHC-1-F10 gene, as shown in Figures 3a and b, resulted in bands with a size of around the 1300 bp mark, which corresponds to the expected amplicon size. Similar results were obtained from TO cell lines (Fig. 3c) and other tissues (Fig. 3d). Therefore, the LOC106588402 gene was successfully amplified in the cell lines and all the tested tissues. In contrast, no positive PCR results was obtained for LOC106588401 gene, and it was consequently excluded from the subsequent studies.
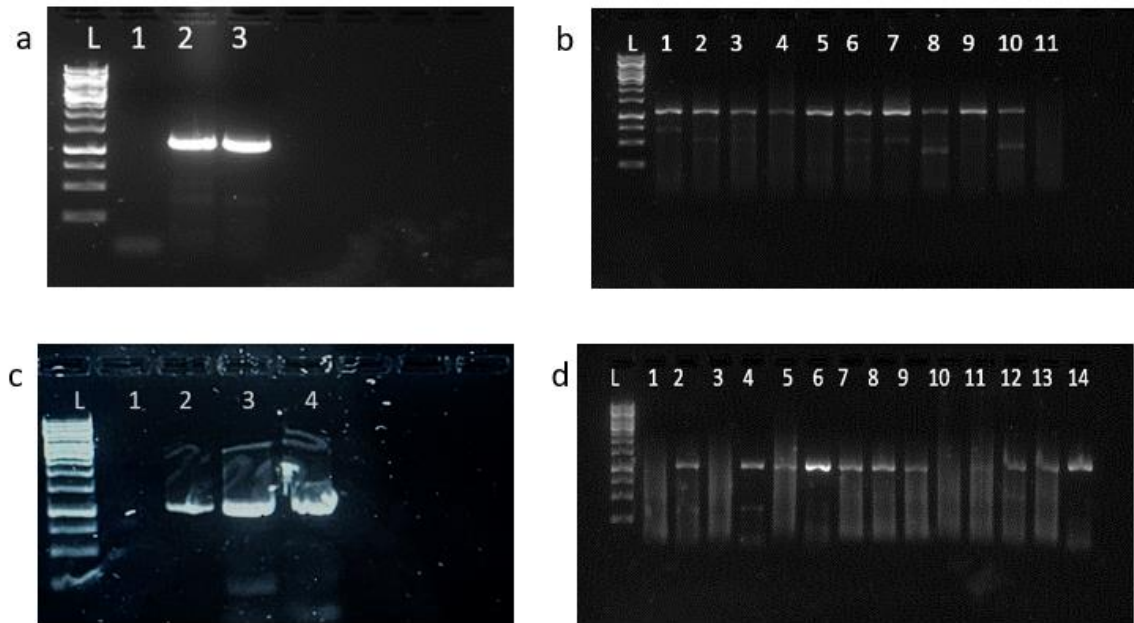
Figure 8. PCR results from the amplification of the MHC-1-F10 gene, where L is the DNA ladder in all cases (a) 1-negative control, 2 and 3- heart samples' bands (b) 1-10-heart samples' bands, 11- negative control (c) 1- negative control, 2-4-TO cell samples' bands(d) 1-4- liver, 5-8- spleen, 9-12- head kidney, and 13-14- skin samples.

4.2 Sequence Diversity of Target Gene in Different Tissues

To gain insigts into the diversity of the targeted MHC I sequences, the PCR results from the tissues mentioned above were extracted and purified and subsequently sequenced using Sanger Sequencing at Eurofins Genomics. We obtained results for 3 or more samples from all tissues except the liver. All the tissues came out with clean chromatograms except for the sequences from the heart samples. This was due to the presence of double peaks at the beginning of the sequences, which spanned to about 319 bp at the beginning of the amplified sequences in the heart samples (Fig.9). Therefore, the heart samples were excluded from the initial alignment that aimed to show the sequence diversity. The remaining sequence with clean chromatograms (from 319 nt to 970 nt of the ref sequence) in all tissues were then aligned with the reference sequence of the MHC-I-F10 gene (LOC106588402; Accession number: NC_059468, region:

10426984…10442999). Figure 10 below shows the alignment results from the tissues with the reference sequences. The results indicated that the sequences obtained from the different tissues were very similar to the reference sequences except for a single nucleotide change at position nt 533 of the reference sequence, where T in the reference was swapped with Gs in the samples (Fig. 10), resulting in aminoacid change from cysteine to tryptophan. The only exceptions were the sequences obtained from the heart, which showed great variability and were only similar to the reference sequences at positions 620 to 920 nt from the alignment (Fig. 11)
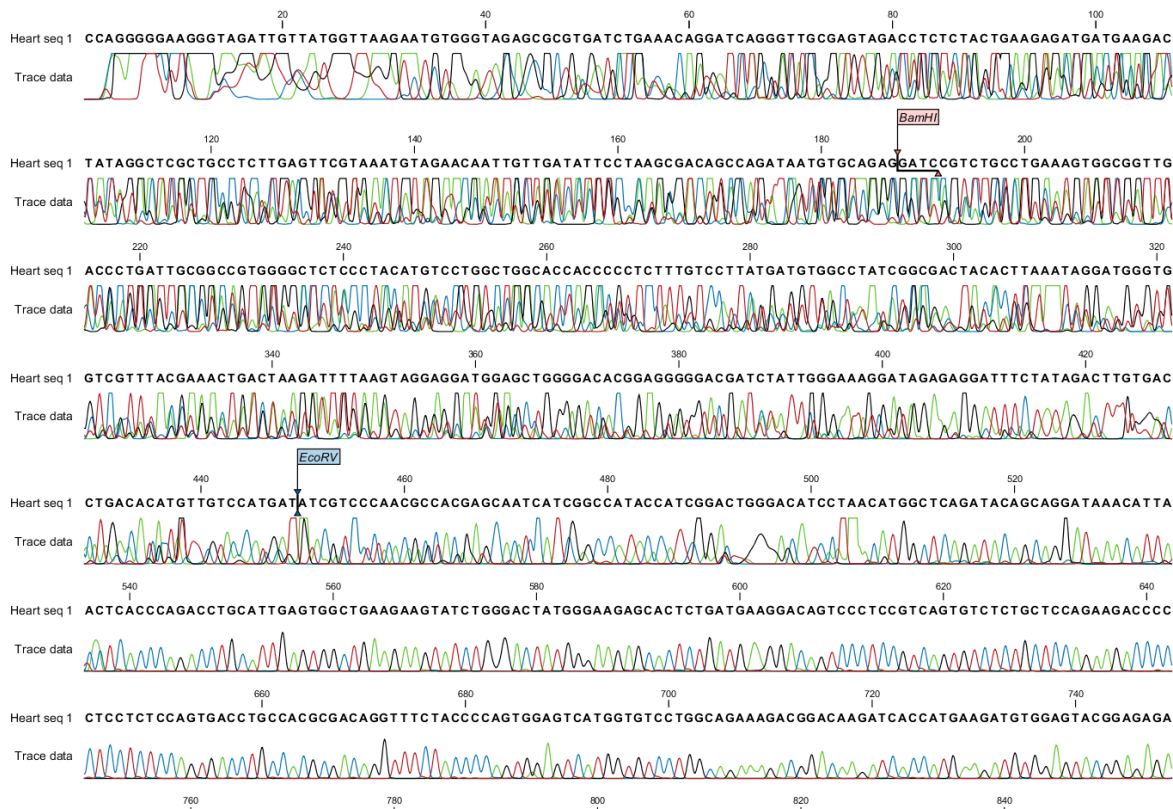
Figure 9. Heart sequence data of MHC-I-F10 gene. This shows the messy results from the beginning of the sequences up to around 500 nt before they become clean.

Figure 10. Alignment of MHC-I-F10 gene sequences from different tissues with the reference sequence. These tissues include the gill, head kidney, liver, skin and the spleen. A single nucleotide polymorphism can be seen at nt 533 with the change from T in the reference sequence to a G in the tissues aligned. There are also some few nucleotide changes, safe, individually.
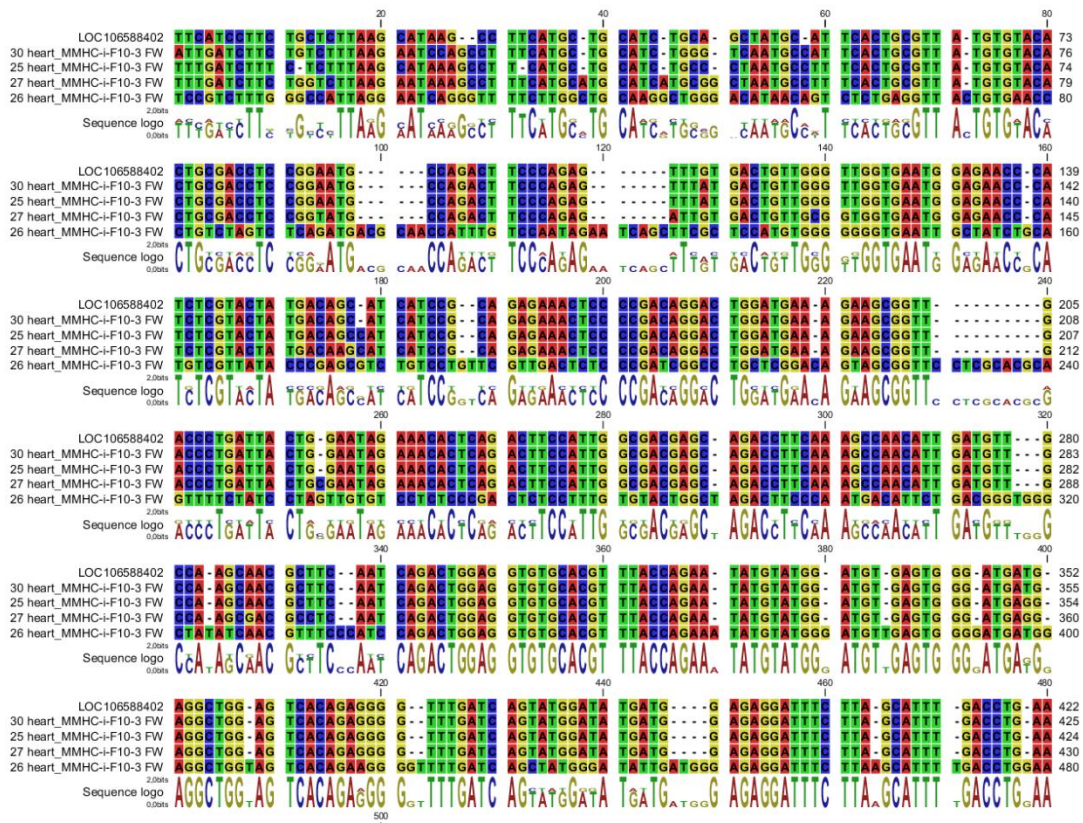
Figure 11. Alignment of MHC-I-F10 gene consensus sequences from the heart samples with the reference sequence. Observations of insertions and deletions from the aligned sequences from the heart samples to the reference sequences.

To understand the reason for the double peaks that occurred at the beginning of the heart MHC-F10 sequences, we designed another reverse primer (GCGTTGGGGCGATACATTTC) and only amplified the first 484 parts of the sequence with the same forward primers used above. The PCR products obtained were then sequenced with both primers. Intriguingly, sequencing with the forward primers resulted in double peaks, while the sequencing results with the reverse primers came clean (Fig. 12). Furthermore, the consensus sequences obtained with the forward primers were aligned to the reference sequence in an attempt to understand the dominant pattern in these sequences. The results showed the presence of multiple deletions and insertions (indels) in the heart sequences relative to the reference sequence (Fig. 13).
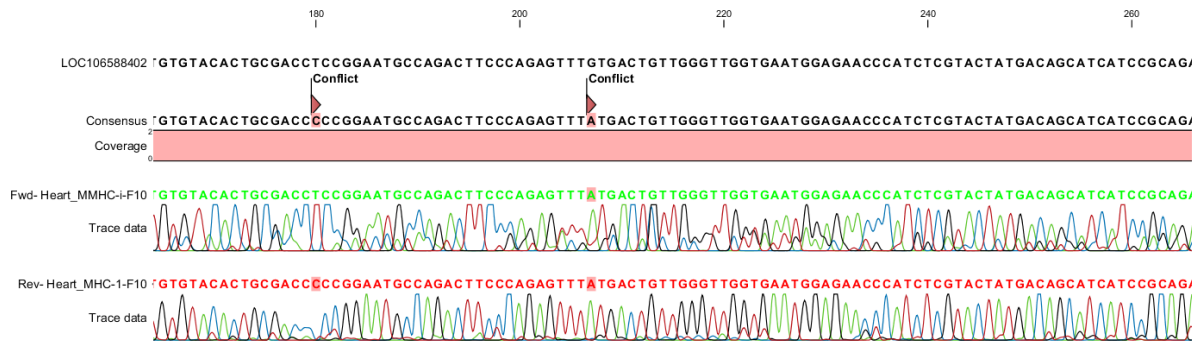
Figure 12. forward and rever sequence results of MHC-I-F10 gene in the heart tissue. Contrasting results from sequencing the heart samples with both the forward and reverse primers.

We also tested whether PMCV infection will affect the sequence diversity by aligning the sequence obtained from infected fish to the other two sequences for the 1300bp PCR product obtained from PMCV infected fish with similar (infected heart 1) or different (infected heart 2) genetic background. The results showed that the double peaks started to disappear from the chromatograms at about nt 450 of the reference sequence in the uninfected fish sequence. For infected fish from both similar and different genetic backgrounds, the sequences continued to have multiple peaks up to about nt 650. Aligning the stretch between nt120 to nt 900 of the reference sequence showed some similarities and differences in the indels detected, but no clear pattern was observed (Fig.13)

Figure 13. Alignment of sequences from infected and uninfected sequences of the heart samples from similar or different genetic backgrounds. This alignment contains sequences from nt 120 to 910 nt from the reference. There are large variabilities exist between the infected and uninfected heart sequences, with uninfected more similar to the reference sequence than the others.

4.3 Editing Efficiency of the Lipofection Method

Two methods were employed to validate the lipofection method's editing efficiency. These include the genomic cleavage assay and the inference of crispr edits (ICE), an online crispr analysis tool developed by Synthego company. For the cleavage assay, lysate from the transfected cells was successfully amplified through PCR reactions with the detection of mismatch and subsequent cleavage carried out through denaturation and application of detection enzyme. The following figure below shows the observed results from the assay. As seen from the results, there were no other bands apart from the single bands seen from the

samples that were not subjected to cleavage. Therefore, the cleavage assay result does not indicate any editing.

Figure 14. Gel electrophoresis of the genomic cleavage assay of CRISPR Cas9 transfected cell line where L is the DNA ladder in all cases. 1-12 CRISPR edited samples' bands. Samples 1-3 are starved cell lines transfected with 1X sgRNA concentration. Samples 4-6 ar are starved cell lines with 3X sgRNA concentration. Samples 7-9 are normal cell lines transfected with 1X sgRNA concentration. Samples 10-12 are normal cell lines transfected with 3X sgRNA concentration.

However, to further validate the result from the cleavage assay, the samples were sequenced with both forward and reverse primers provided in Table 1. The ICE tool was then utilised to analyse the results after Sanger sequencing. The results of the analysis of the CRISPR edit through the ICE tool is displayed in Table 2 below.

Table 2: Numbers of indels created from conditions tested in the transfection method

Editing Method	Transfection			
Conditions	S		N	
	1X	3X	1X	3X
1x Transfection	0	0	0	0
2x Transfection	0	0	0	1
3x Transfection	0	0	1	3
Total		5		

N, normal, S, starved, 1X/3X, one and three times the sgRNA concentrations, 1x, 2x and 3x, are the number of times transfection takes place.

The various conditions tested during the study included the cell condition before transfection (starved or normal), the concentration of the sgRNAs and the transfection times. Contrary to the expectation that starvation of the cells enhances transfection efficiency, no indels were detected in the starved samples compared to five detected in the non-starved or normal samples. Furthermore, more indels were detected in samples with multiple transfections with the RNP complexes.

4.4 Editing Efficiency of the Electroporation Method

Similar to the transfected cells, the editing efficiency for electroporated cells was analysed with genomic cleavage assay and the ICE tool from Synthego. Figure 15 and table 3 show the results from the two methods, respectively. The results from the cleavage assay showed no indication of editing as both the control and the positive samples had the same single bands. However, when using the ICE analysis tools indels were detected in cells electroporated with both 1x (7.5 pmol) and 3x (22.5 pmol) concentration of the sgRNA. Nevertheless, samples with higher concentration (3X) of sgRNAs could be seen with higher numbers of indels than those with lesser concentration (1X).
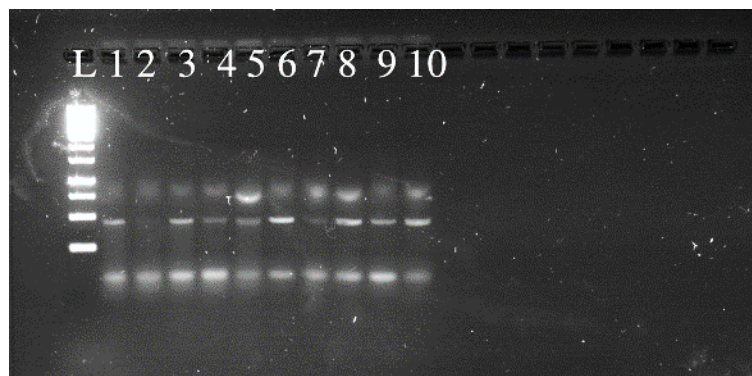
Figure 15. Gel electrophoresis of the genomic cleavage assay of CRISPR Ca9 electroporated cell line where L is the DNA ladder in all cases. 1-5 CRISPR edited samples' bands. 6-10- negative control. Samples 1 and 2 are cell lines electroporated with 1X sgRNA concentration. Samples 3 and 4 are cell lines electroporated with 3X sgRNA concentration. Sample 5, positive control.

Table 3: Numbers of indels created from conditions tested in the electroporation method

Editing Method	Electroporation	
Condition	1X	3X
Indels detected	2	6
Total	8	

1X/3X, one and three times the sgRNA concentrations

In addition, the editing efficiencies obtained by transfection and electroporation were also compared (Fig. 16). The results show that the editing efficiency was generally not high, with the highest editing efficiency being 5%. Nevertheless, relatively higher editing efficiency was detected in electroporated samples. Comparable editing efficiencies were detected when sequencing using both forward and reverse primers, indicating that editing occurred in both strands.
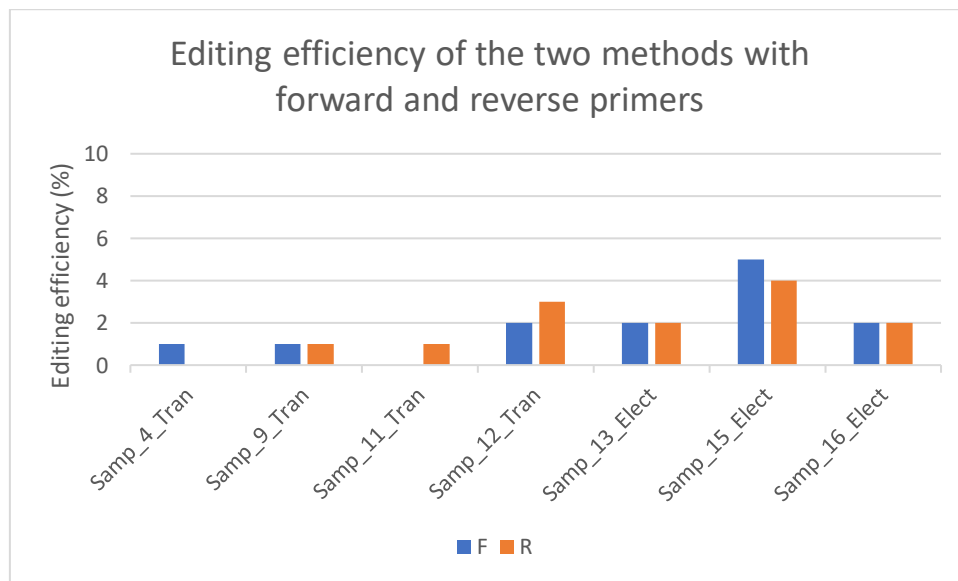
Figure 16. Overall editing efficiency of the two methods with both forward and reverse primers from the ICE tool. Samples 4, 9, 11, and 12 are successful samples edited via the transfection method. Samples 13, 15, and 16 are successful samples edited via electroporation.

From figure 16 above, the highest rate of efficiency was highest with sample 15 when analysed with both forward and reverse primers with the ICE tool, with both 4 and 5% efficiency, respectively. In addition, it is noteworthy that the sample with the highest efficiency was

electroporated with a higher concentration of sgRNA (3X). On the other hand, sample 12 had the highest efficiency among the samples transfected. The sample was also transfected with a higher concentration of sgRNA (3X).

Chapter 5 Discussion

5.1 Amplification and Detection of the MHC-I-F10 gene in different tissues and TO cell line

The genes (LOC106588401 and LOC106588402) used in this study were detected in previous studies focusing on the resistance of Atlantic salmon against the CMS disease in different populations, both in-field and laboratory challenge tests. Before editing the gene, we needed to understand the distribution of these genes in different tissues and TO cell lines because such information is of relevance to the design and execution and practical importance of the study. A PCR procedure was established using primers specifically designed to detect these genes. From the results, only the LOC106588402 (Accession number: NC_059468, region: 10426984…10442999) gene was successfully amplified and detected with the primers in all the tissues and also the TO cell line. It is unclear why we could not successfully amplify the LOC106588401 gene despite both genes were found to have a strong association with the QTL for the resistance against CMS. However, it is noteworthy that the LOC106588401 gene was only detected in the study (Boison et al., 2019), while the other was detected in all three studies (Boison et al., 2019; Hillestad et al., 2020; Hillestad & Moghadam, 2019). It is therefore not clear if the association of LOC106588401 with CMS resistance is real. Differences in virus variants have not been explored which could also contribute to disparate results between studies. Nevertheless, the association between these genes and the QTL related to resistance against CMS should be explored in more detail using selective breeding with marker-assisted selection and/or genomic selection combined with a functional approach to better understand the underlying mechanisms of resistance against PMCV infection.

5.2 Sequence diversity of MHC-I-F10 gene in different tissues

Observation of the sequences after quality control checks and alignment, from various tissues such as the heart, gill, spleen, liver, head kidney, and the skin showed some interesting results.

Firstly, all of the tissue sequences, except those obtained from the heart, were for the most part identical to the reference sequence of the gene (Accession: BT059396.1) except for a common T-G single nucleotide polymorphism (SNP) at nt 533 observed in all the tissues resulting in aminoacid change from cysteine to tryptophan. This SNP should be investigated further to determine whether it is genuine or just an artefact in the sequence landscape. The high sequence similarity is in contrast to the known polymorphic nature of the MHC genes, as they are known to be one of the most polymorphic genes with multiple loci and myriads of alleles at each locus (Marsh et al., 2000). One explanation could be that the target sequence is a conserved (framework) part of the MHC I. Trimming the areas with multiple peaks at the start/end of the sequences may also have contributed to the conformity. It should be added that samples analysed came from a group of fish with similar genetic background which might provide some explanation. Further, it is noteworthy that all heart samples came with double peaks or messy sequences in the first 450-600 bp. At this stage, it is unclear what could have caused this and why it occurred only in the heart samples from different individuals. The genetic background of the fish from which the heart samples originated came from different fish with diverse genetic backgrounds.

To add to the complicated picture, an attempt to sequence the heart samples with reverse primers resulted in a clean sequence. It could be that the reverse primers binds in a part of the gene with less variation compared to the forward primers. This possibility was not further investigated in this study and needs to be addressed in future studies. Another interesting finding in the heart is the difference between control and PMCV infected fish with regard to the sequence stretch with multiple peaks (450 in control and 650 in infected). It must be noted, however, that this difference was not thoroughly investigated in this thesis and should therefore be treated as an observation that warrants further investigation.

Furthermore, a careful look at the alignment of the heart sample sequences with the reference revealed the possible presence of some insertion and deletion. The indels are an essential part of the genomic variants in the genome landscape with substantial effects. When they occur in the coding regions, they result in two variants, known as the non-frameshift and frameshift errors in the genome, which might result in the introduction or removal of one or more amino acids while the protein sequence remains unaltered or a change in the reading frame which might result in an entirely different protein or premature termination of the protein, respectively (Lin et al., 2017). It has also been found that the frameshift error caused by indels are more destructive as they could also result in loss of function variants in vital proteins (Lin et al., 2017). Hence, multiple indels in the heart sequences need to be investigated further. One way to further study the sequence variation in the heart is by cloning the different PCR products and sequencing single clones to identify the presence of indels.

5.3 Editing efficiency for the Transfection and Electroporation methods

In this study, two different methods, transfection and electroporation, were used to deliver the ribonucleoprotein into the cells with the subsequent observation of their efficiency. The reason for using different methods was to check which of the two would be optimal for the delivery of ribonucleoprotein into the TO cells. Furthermore, the concentration of the sgRNAs—one-third of each sgRNA (1X) for the first condition and one for each of the sgRNA (3X) for the other- was utilised. As multiple gRNA had been suggested to be much more efficient than using a single gRNA for genome editing, three sgRNAs were used. These concentrations were determined by following the manufacturer´s recommendations. In addition, the last condition was the starvation of some cells before transfection while leaving some set unchanged, known as normal. Cell starvation before transfection has been found to improve transfection efficiency (Wallenstein et al., 2018) in embryonic stem cells and others and was therefore used as a

condition to enhance the transfection rate and efficiency. However, this was not the case in this study, as no indels were detected in starved cells.

The cleavage assay results from transfection and electroporation methods showed no indication of editing in the cells. This could indicate either no or very low editing. Therefore, the products were further sequenced using the sanger sequencing to make available results that could be used for further analysis of the editing efficiency using the ICE analysis tool provided by Synthego. The results from the ICE analysis using the sanger sequence showed overall low editing efficiency, with more samples from the electroporation method having many indels when the edited and control sequences were contrasted by the tool. The low editing efficiency experienced in this study might have emanated from many factors. The most significant one is the duration of the experiments, which is quite short and dramatically influences the overall results. This left no room to optimise these conditions further to get better results. Furthermore, the TO cell line have not been known to have been used in a genome editing study based on the information available at the time of this study. While Atlantic salmon genomes have been successfully edited using several cell lines (Gratacap et al., 2020; Gratacap et al., 2019), there has been no such study in the TO cell line. Moreover, cell health and viability are other factors that could have impacted the results of this current study. In this regard, it is worth noting that Atlantic salmon have a partially duplicated genome (Lien et al., 2016). Therefore, genome editing in cell lines derived from this species will consequently require more optimised conditions than for mammalian cell lines to ensure maximum efficiency. From our data, there are indications that using high concentrations and multiple transfections results in better editing efficiency—subsequently, this will constitute a good ground for optimisation of the editing protocols. For example, previous studies have used gRNA (crRNA and tracrRNA) with the final concentration of 25 pmol/µl for genome editing in contrast to the maximum of 22.5pmol

used in this study (Gratacap et al., 2020). Using similar or even higher concentrations may therefore lead to better editing.

With the advent of high-tech breeding programs for many of the world's most valuable aquaculture species, the application of genetics and breeding technologies in aquaculture is rising rapidly. Consequently, combating the menace of CMS and other related diseases in Atlantic salmon and other species would enhance the industry's continuous growth and increase its capacity to contribute to food and nutrition security. In addition, most cultured aquatic species provide opportunities for the application of selective breeding as they are still closely genetically related to their wild relatives, and this represents a significant unexplored resource for enhancing the sustainability of farmed seafood (Gratacap et al., 2019). On the other hand, aquaculture species' high fecundity and ease of external fertilization make them ideal candidates for high-resolution genetic studies to understand better and improve complex traits. Therefore, there is great potential in using novel selective breeding and genome-editing technologies like CRISPR/Cas9 to speed up genetic improvement for production traits and functional traits—such as disease, which is a significant obstacle to the industry. Further investigations of the sequence variability in the heart and its association with CMS may provide knowledge that can be used to select disease-resistant fish.

Conclusions and future perspectives

This study has shown the possibility of successful editing of the MHC-I-F10 gene in TO cell line, though with very low efficiency but has laid the foundation for further functional studies for the validation of this important gene. In addition, the TO cell line used in this study is widely used as a model system for immunological studies in Atlantic salmon. Therefore, the capability to perform targeted knockout will enable the evaluation of other candidate genes that might be related to resistance to other diseases. In the future, a satisfactory editing efficiency of the MHC-I-F10 gene could be successfully accomplished given enough time and consideration of other factors that might have impacted the outcome in this study, and might serve as a component of a valuable toolbox for future functional genetics and immunological studies in Atlantic salmon.

In addition, as the influence of the novel CRISPR technology continues to increase across the field of agriculture, medicine and the general life sciences, the Atlantic salmon and aquaculture industry, in general, could benefit immensely from this great innovation in the future. One of these could be the detection, removal, promotion or fixation of functional alleles of interest, segregating within the broodstock at one or many loci in the population. Furthermore, the technology could be deployed to induce introgression-by-editing, subsequently by-passing the inherent problem of linkage drag that accompanies traditional introgression. Finally, the CRISPR technology could help generate and use *de novo* favourable alleles that are unknown by using prior knowledge of the biology of the traits of interest or approaches that utilise genome-wide genetic perturbations to identify candidate genes. All of which could significantly improve the industry's capacity to contribute to the world´s food and nutrition security.

References

Adams, M. J., Lefkowitz, E. J., King, A. M. Q., & Carstens, E. B. (2014). *Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2014)*. 2831–2841. https://doi.org/10.1007/s00705-014-2114-3

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

Amin, A. B., & Trasti, J. (1988). Endomyocarditis in atlantic salmon in Norwegian seafarms. *Bulletin of the European Association of Fish Pathologists*, *8*(4), 70–71.

Barrangou, R., Fremaux, C., Deveau, H., Richardss, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR Provides Against Viruses in Prokaryotes. *Science*, *315*(March), 1709–1712.

Barrett, L. T., Oppedal, F., Robinson, N., & Dempster, T. (2020). Prevention not cure: a review of methods to avoid sea lice infestations in salmon aquaculture. *Reviews in Aquaculture*, *12*(4), 2527–2543. https://doi.org/10.1111/raq.12456

Boison, S., Ding, J., Leder, E., Gjerde, B., Bergtun, P. H., Norris, A., Baranski, M., & Robinson, N. (2019). QTLs Associated with Resistance to Cardiomyopathy Syndrome in Atlantic Salmon. *Journal of Heredity*, *110*(6), 727–737. https://doi.org/10.1093/jhered/esz042

Brouns, JJ, S., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V., & Oost, J. Van Der. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science*, *321*(5891), 960–964.

Capecchi, M. R. (1989). *Altering the Genome Homologous Recombination by From ES Cells to Germ Line Chimera*. *236*(1987).

Cassandro, M. (2020). Animal breeding and climate change, mitigation and adaptation. In *Journal of Animal Breeding and Genetics* (Vol. 137, Issue 2, pp. 121–122). https://doi.org/10.1111/jbg.12469

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

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

Dionne, M., Miller, K. M., Dodson, J. J., & Bernatchez, L. (2009). MHC standing genetic variation and pathogen resistance in wild Atlantic salmon. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1523), 1555–1565. https://doi.org/10.1098/rstb.2009.0011

Dunn, S. E., Li, H., Cardone, G., Nibert, M. L., Ghabrial, S. A., & Baker, T. S. (2013). *Three-dimensional Structure of Victorivirus HvV190S Suggests Coat Proteins in Most Totiviruses Share a Conserved Core*. *9*(3). https://doi.org/10.1371/journal.ppat.1003225

FAO. (2014). *Meat Consumption*. http://www.fao.org/ag/againfo/themes/en/meat/background.html

FAO. (2020). The State of World Fisheries and Aquaculture 2020. Sustainability in action. In *Fao*. https://doi.org/https://doi.org/10.4060/ca9229en

Fevolden, S. E., Røed, K. H., Fjalestad, K. T., & Stien, J. (1999). Poststress levels of

lysozyme and cortisol in adult rainbow trout: Heritabilities and genetic correlations.

Journal of Fish Biology, *54*(4), 900–910. https://doi.org/10.1006/jfbi.1999.0930

Fevoldena, S. E., & Roedb, K. H. (1990). Selection for high and low cortisol stress response

in Atlantic salmon (Sahzo salar) and rainbow. *Atlantic*, *95*(2), 53–65.

Fritsvold, C., Mikalsen, A. B., Poppe, T. T., Taksdal, T., & Sindre, H. (2021).

Characterization of an outbreak of cardiomyopathy syndrome (CMS) in young Atlantic

salmon, Salmo salar L. In *Journal of Fish Diseases* (Vol. 44, Issue 12, pp. 2067–2082).

https://doi.org/10.1111/jfd.13521

Gjedrem, T., & Gjøen, H. M. (1995). Genetic variation in susceptibility of Atlantic salmon,

Salmo salar L., to furunculosis, BKD and cold water vibriosis. *Aquaculture Research*,

26(2), 129–134. https://doi.org/10.1111/j.1365-2109.1995.tb00892.x

Gjedrem, Trygve, Salte, R., & Gjøen, H. M. (1991). Genetic variation in susceptibility of

Atlantic salmon to furunculosis. *Aquaculture*, *97*(1), 1–6. https://doi.org/10.1016/0044-

8486(91)90274-B

Gjerde, B, Terjesen, B. F., Barr, Y., Lein, I., & Thorland, I. (2004). *Genetic variation for

juvenile growth and survival in Atlantic cod (Gadus morhua)*. *236*, 167–177.

https://doi.org/10.1016/j.aquaculture.2004.03.004

Gjerde, Bjarne. (2006). State of the Art in Selective Breeding of Aquacultured Species. *The

Isreali Journal of Aquaculture-Bamidgeh*, *58*(4), 2006.

Goddard, M. E., & Hayes, B. J. (2009). Mapping genes for complex traits in domestic

animals and their use in breeding programmes. *Nature Publishing Group*, *10*(juNe),

381–391. https://doi.org/10.1038/nrg2575

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

Gratacap, R. L., Regan, T., Dehler, C. E., Martin, S. A. M., Boudinot, P., Collet, B., &
Houston, R. D. (2019). Efficient CRISPR/Cas9 genome editing in a salmonid fish cell
line using a lentivirus delivery system. *BioRxiv*, 1–9. https://doi.org/10.1101/734442

Gratacap, R. L., Wargelius, A., Edvardsen, R. B., & Houston, R. D. (2019). Potential of
Genome Editing to Improve Aquaculture Breeding and Production. *Trends in Genetics*,
35(9), 672–684. https://doi.org/10.1016/j.tig.2019.06.006

Grimholt, U. (2016). MHC and evolution in teleosts. *Biology*, *5*(1).
https://doi.org/10.3390/biology5010006

Grimholt, U., Larsen, S., Nordmo, R., Midtlyng, P., Kjoeglum, S., Storset, A., Saebø, S., &
Stet, R. J. M. (2003). MHC polymorphism and disease resistance in Atlantic salmon
(Salmo salar); facing pathogens with single expressed major histocompatibility class I
and class II loci. *Immunogenetics*, *55*(4), 210–219. https://doi.org/10.1007/s00251-003-
0567-8

Guo, T. C., Ahmed, A., Gamil, A., Koenig, M., & Evensen, Ø. (2015). *Sequence Analysis
and Identification of New Isoform of EP4 Receptors in Different Atlantic Salmon Tissues
(Salmo salar L .) and Its Role in PGE2 Induced Immunomodulation In Vitro*.
https://doi.org/10.1371/journal.pone.0120483

Haugland, Ø., Mikalsen, A. B., Nilsen, P., Lindmo, K., Thu, B. J., Eliassen, T. M., Roos, N.,
Rode, M., & Evensen, Ø. (2011). Cardiomyopathy Syndrome of Atlantic Salmon
(Salmo salar L.) Is Caused by a Double-Stranded RNA Virus of the Totiviridae Family.
Journal of Virology, *85*(11), 5275–5286. https://doi.org/10.1128/jvi.02154-10

Hillestad, B., Kristjánsson, Ó. H., Makvandi-Nejad, S., & Moghadam, H. K. (2020). Genome-Wide Association Study Confirms Previous Findings of Major Loci A ff ecting Resistance to Piscine. *Genes*, *11*, 608.

Hillestad, B., & Moghadam, H. K. (2019). Genome-Wide Association Study of Piscine Myocarditis Virus (PMCV) Resistance in Atlantic Salmon (Salmo salar). *Journal of Heredity*, *110*(6), 720–726. https://doi.org/10.1093/jhered/esz040

Houston, R. D., Haley, C. S., Hamilton, A., Guy, D. R., Gheyas, A. A., Tinch, A. E., Taggart, J. B., Bron, J. E., Starkey, W. G., Mcandrew, B. J., Paley, R. K., Rimmer, G. S. E., Tew, I. J., & Bishop, S. C. (2010). The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic necrosis is largely explained by a major QTL. *Heredity*, 318–327. https://doi.org/10.1038/hdy.2009.171

Jatin M. Vyas, Annemartha G. Van der Veen, H. L. P. (2008). The known unknowns of antigen processing and presentation. *Nature Reviews Immunology*, *8*(8), 607–618. https://doi.org/10.1038/nri2368.The

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). *A Programmable Dual-RNA – Guided*. *337*(August), 816–822.

Karunasagar, I., Pai, R., Malathi, G. R., & Karunasagar, I. (1994). *Mass mortality of Penaeus monodon larvae due to antibiotic-resistant Vibrio harveyi infection*. *128*, 203–209.

Lafferty, K. D., Harvell, C. D., Conrad, J. M., Friedman, C. S., Kent, M. L., Kuris, A. M., Powell, E. N., Rondeau, D., & Saksida, S. M. (2015). *Infectious Diseases Affect Marine Fisheries and Aquaculture Economics*. *September 2014*, 1–26. https://doi.org/10.1146/annurev-marine-010814-015646

Langefors, Ð., Lohm, J., Grahn, M., & Andersen, Ò. (2001). *Association between major*

histocompatibility complex class IIB alleles and resistance to Aeromonas salmonicida in Atlantic salmon. *September 2000*. https://doi.org/10.1098/rspb.2000.1378

Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Matthew, P., Leong, J. S., Minkley, D. R., Zimin, A., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R. A., Schalburg, K. Von, Rondeau, E. B., Genova, A. Di, Samy, J. K. A., & Vik, J. O. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, *533*(7602), 200–205. https://doi.org/10.1038/nature17164

Lin, F. L., Sperle, K., & Sternberg, N. (1985). Recombination in mouse L cells between DNA introduced into cells and homologous chromosomal sequences. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(5), 1391–1395. https://doi.org/10.1073/pnas.82.5.1391

Lin, M., Whitmire, S., Chen, J., Farrel, A., Shi, X., & Guo, J. T. (2017). Effects of short indels on protein structure and function in human genomes. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-09287-x

Løvoll, M., Wiik-Nielsen, J., Grove, S., Wiik-Nielsen, C. R., Kristoffersen, A. B., Faller, R., Poppe, T., Jung, J., Pedamallu, C. S., Nederbragt, A. J., Meyerson, M., Rimstad, E., & Tengs, T. (2010). A novel totivirus and piscine reovirus (PRV) in Atlantic salmon (Salmo salar) with cardiomyopathy syndrome (CMS). *Virology Journal*, *7*, 1–7. https://doi.org/10.1186/1743-422X-7-309

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science*, *339*(6121), 823–826. https://doi.org/10.1126/science.1232033

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

https://doi.org/10.1126/science.1165771

Mccormick, S. D. (2013). 5. Smolt Physiology and Endocrinology. In *Euryhaline Fishes* (Vol. 32). Elsevier Inc. https://doi.org/10.1016/B978-0-12-396951-4.00005-0

Miller, K. M., Winton, J. R., Schulze, A. D., Purcell, M. K., & Ming, T. J. (2004). *Major histocompatibility complex loci are associated with susceptibility of Atlantic salmon to infectious hematopoietic necrosis virus*. 307–316.

Nations, U. (2019). World Population Prospects 2019 Highlights. In *United Nations*. https://doi.org/10.3406/tiers.1983.4290

Neefjes, J., Jongsma, M. L. M., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class i and MHC class II antigen presentation. *Nature Reviews Immunology*, *11*(12), 823–836. https://doi.org/10.1038/nri3084

Nelson, J. S (1984). Fishes of the World. 2nd edition. JohnWiley and Sons, Inc., New York.

Parham, P., & Ohta, T. (1996). *PO p u I a ti 0 n B i 0 I ogy of An ti g en Presentation by MHC Class I Molecules*. *272*(April).

Rock, K. L., Reits, E., & Neefjes, J. (2016). Present Yourself! By MHC Class I and MHC Class II Molecules. *Trends in Immunology*, *37*(11), 724–737. https://doi.org/10.1016/j.it.2016.08.010

Rouet, P., Smih, F., & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*, *14*(12), 8096–8106. https://doi.org/10.1128/mcb.14.12.8096-8106.1994

S., P. B. (1963). Introduction to Quantitative Genetics. *Poultry Science*, *42*(2), 547. https://doi.org/10.3382/ps.0420547

Sommerset, I., Walde S., C., Bang Jensen, B., Bornø, G., Haukaas, A., & Brun, E. (red). (2020). *The Health Situation in Norwegian Aquaculture 2019 - Norwegian Veterinary Institute report series nr 5b/2020*. 1–156. https://www.vetinst.no/rapporter-og-publikasjoner/rapporter/2020/fish-health-report-2019

Steven G.E. Marsh, Peter Parham, L. D. B. (2000). *The HLA factsbook*.

Storebakken, T. (2009). Atlantic salmon, Salmo salar. *Nutrient Requirements and Feeding of Finfish for Aquaculture*, *7*(May), 79–102. https://doi.org/10.1079/9780851995199.0079

Su, H., van Eerde, A., Steen, H. S., Heldal, I., Haugslien, S., Ørpetveit, I., Wüstner, S. C., Inami, M., Løvoll, M., Rimstad, E., & Clarke, J. L. (2021a). Establishment of a piscine myocarditis virus (PMCV) challenge model and testing of a plant-produced subunit vaccine candidate against cardiomyopathy syndrome (CMS) in Atlantic salmon Salmo salar. *Aquaculture*, *541*(April). https://doi.org/10.1016/j.aquaculture.2021.736806

Su, H., van Eerde, A., Steen, H. S., Heldal, I., Haugslien, S., Ørpetveit, I., Wüstner, S. C., Inami, M., Løvoll, M., Rimstad, E., & Clarke, J. L. (2021b). Establishment of a piscine myocarditis virus (PMCV) challenge model and testing of a plant-produced subunit vaccine candidate against cardiomyopathy syndrome (CMS) in Atlantic salmon Salmo salar. *Aquaculture*, *541*(December 2020). https://doi.org/10.1016/j.aquaculture.2021.736806

Svendsen, J. C., Nylund, S., Kristoffersen, A. B., Takle, H., Fossberg Buhaug, J., & Jensen, B. B. (2019). Monitoring infection with Piscine myocarditis virus and development of cardiomyopathy syndrome in farmed Atlantic salmon (Salmo salar L.) in Norway. *Journal of Fish Diseases*, *42*(4), 511–518. https://doi.org/10.1111/jfd.12974

Tavakoli, K., Pour-Aboughadareh, A., Kianersi, F., Poczai, P., Etminan, A., & Shooshtari, L. (2021). Applications of CRISPR-Cas9 as an advanced genome editing system in life

sciences. In *BioTech* (Vol. 10, Issue 3). https://doi.org/10.3390/biotech10030014

Timmerhaus, G., Krasnov, A., Nilsen, P., Alarcon, M., Afanasyev, S., Rode, M., Takle, H., & Jørgensen, S. M. (2011). Transcriptome profiling of immune responses to cardiomyopathy syndrome (CMS) in Atlantic salmon. *BMC Genomics*, *12*. https://doi.org/10.1186/1471-2164-12-459

Vassvik, V. (2000). Fish Farming: *A pictorial adventure in the blue pasture* (ISBN 82-995521-1-7)

Verspoor, E., Stradmeyer, L., & Nielsen, J. (2007). The Atlantic Salmon: Genetics, Conservation and Management. In *The Atlantic Salmon: Genetics, Conservation and Management*. https://doi.org/10.1002/9780470995846

Wallenstein, E. J., Barminko, J., Schloss, R. S., & Yarmush, M. L. (2018). *Serum Starvation Improves Transient Transfection Efficiency in Differentiating Embryonic Stem Cells*. *26*(6), 1714–1723. https://doi.org/10.1002/btpr.472.Serum

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

Wetten, M., Aasmundstad, T., Kjøglum, S., & Storset, A. (2007). Genetic analysis of resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar L.). *Aquaculture*, *272*(1–4), 111–117. https://doi.org/10.1016/j.aquaculture.2007.08.046

Wieczorek, M., Abualrous, E. T., Sticht, J., Álvaro-Benito, M., Stolzenberg, S., Noé, F., & Freund, C. (2017). Major histocompatibility complex (MHC) class I and MHC class II proteins: Conformational plasticity in antigen presentation. *Frontiers in Immunology*, *8*(MAR), 1–16. https://doi.org/10.3389/fimmu.2017.00292

Wiik-Nielsen, J., Alarcón, M., Jensen, B. B., Haugland, & Mikalsen, A. B. (2016). Viral co-

infections in farmed Atlantic salmon, Salmo salar L., displaying myocarditis. *Journal of Fish Diseases*, *39*(12), 1495–1507. https://doi.org/10.1111/jfd.12487

Yáñez, J. M., Houston, R. D., & Newman, S. (2014). Genetics and genomics of disease resistance in salmonid species. *Frontiers in Genetics*, *5*(NOV), 1–13. https://doi.org/10.3389/fgene.2014.00415

Zilong Tan, C. K. and W. J. E. (2007). *Health management practices for cage aquaculture in Asia - a key component for sustainability*. https://thefishsite.com/articles/health-management-practices-for-cage-aquaculture-in-asia-a-key-component-for-sustainability