Norwegian University
of Life Sciences

Master's Thesis 2023 60 ECTS
Faculty of Biosciences

Development of embryonic cell culture conditions in Atlantic salmon

Maren Ileby Wessel
M.Sc. Biotechnology (Genetics)

Aknowledgements

First and foremost, I would like to thank my main supervisor Victor Boyarchuk (researcher, NMBU) and co-supervisor Prabin Sharma Humagain (PhD candidate, NMBU) for their invaluable guidance, feedback, and support throughout this project. You've trusted me to be independent, but always made time for me when I needed help, which has contributed profoundly to my learning process.

I am also grateful for the rest of the CIGENE lab, who generously shared their expertise and assistance when needed. A special thanks to Guro Katrine Sandvik (associate professor, NMBU) and Simen Rød Sandve (professor, NMBU) for a great initiative on the writing group. Would also like to thank the other participants for truly helpful and honest feedback.

Finally, I would also like to thank my family and loved ones for being of great support throughout these past months. You´ve encouraged me and been a constant source for motivation, and I'm so very grateful to all of you.

Maren Ileby Wessel

Abstrakt

Potensialet til bruk av induserte pluripotente stamceller har fått mye oppmerksomhet siden deres første produksjon. Disse cellene har en fenotype som ligner embryonale stamceller og kan derfor gi opphav til celler i alle kroppens vev. En vellykket regenerering fra differensiert til en pluripotent celle forbeholder kunnskap om hvilke gener som er aktive i de udifferensierte cellene. Med kjennskap til disse er neste steg å indusere uttrykket av de respektive genene i målcellen. Når cellene har blitt regenerert, er det endelige målet å etablere en udødelig cellelinje. Dette avhenger av et optimalisert medium som fremmer cellevekst samtidig som det hemmer celle-differensiering.

I denne oppgaven har RNAseq-data blitt brukt til å se på genuttrykk i embryoniske stamceller hos laks, og identifisere gener involvert i å opprettholde en pluripotent fenotype, samt gener assosiert med differensiering. Disse markørene ble testet ved hjelp av qPCR og Western blotting på forskjellige cellestadier under tidlig utvikling. *Nanog* ble identifisert som en god markør for en pluripotent celle, mens *Apoa1* og *K2c8* var gode markører for differensiering. *Roaa* ble valgt som kontrollgen, og ble brukt for å normalisere qPCR-dataene. Western blot-analysen av Oct4- og Sox17-antistoffene ga ikke konkluderende resultater, og videre analyse er nødvendig for å validere disse genene som markører for henholdsvis pluripotent og differensiert fenotype.

Forskjellige sammensetninger av medier for dyrking av embryonale stamceller ble også testet i denne studien, og det ble gjort forsøk på å utvikle disse forholdene. Av de fem forskjellige mediene som ble testet, overlevde celler bare i fire av dem. Profileringsnivået var imidlertid for lavt, og brønnene ble aldri konfluente. Ingen av mediene oppfylte sitt formål og ville trenge videre optimalisering.

Abstract

The potential use of induced pluripotent stem cells has garnered significant attention since their first production. These cells have a phenotype similar to embryonic stem cells and can therefore give rise to cells in all of the body's tissues. A successful regeneration from differentiated to a pluripotent cell depends on knowledge of which genes are active in the undifferentiated cells. With information of these genes, the next step is to induce the expression of the respective genes in the target cell. Once the cells have been regenerated, the goal is to establish an immortal cell line. This depends on an optimised medium that promotes cell growth while inhibiting differentiation.

This study examined gene expression in embryonic stem cells of salmon using RNAseq data to identify genes involved in maintaining a pluripotent phenotype and those associated with differentiation. These markers were tested using qPCR and Western blotting at different cell stages during early development. *Nanog* was identified as a good marker for the pluripotent phenotype, while *Apoa1* and *K2c8* were good markers for differentiation. *Roaa* was chosen as control gene and used to normalise the qPCR data. The Western blot analysis of Oct4 and Sox17 antibodies did not yield conclusive results, and further analysis is required to validate these genes as markers of pluripotency and differentiation, respectively.

Different compositions of media for the cultivation of embryonic cells were also tested in this study, and an attempt was made to develop these conditions. Of the five different media tested, cells only survived in four of them. However, the profiling level was too low, and the wells were never confluent. None of the media fulfilled their purpose and would need further optimisation.

Keywords: Embryonic stem cells, pluripotency, differentiation, Atlantic salmon, cell culture media, qPCR, Western blotting.

TABLE OF CONTENTS

1 INTRODUCTION AND LITERATURE REVIEW

1.1 Norwegian Salmon farming

1.1.1 Historical perspective and importance

With a long coastline, Norway has great prerequisites conducting aquaculture and fishing. The first salmon farm was introduced outside the east coast of Norway in 1970. Within the same decade, Atlantic salmon (*Salmo salar*, hereafter "salmon") passed the production of rainbow trout (*Oncorhynchus mykiss*), formerly dominating the national seafood industry. Ever since, the Norwegian aquaculture has increased drastically, and became in 2015, the world leading producer of farmed salmon (Norwegian seafood council & Seafood Norway, 2021). The aquaculture industry is of great commercial importance as it provides a large export income. In 2021, Norway passed 113.5 billion Norwegian kroners (NOK) in fish export. ⅔ of the total revenue was made up of salmon (Statistisk sentralbyrå, 2022), also visualised in Figure 1.1 below.

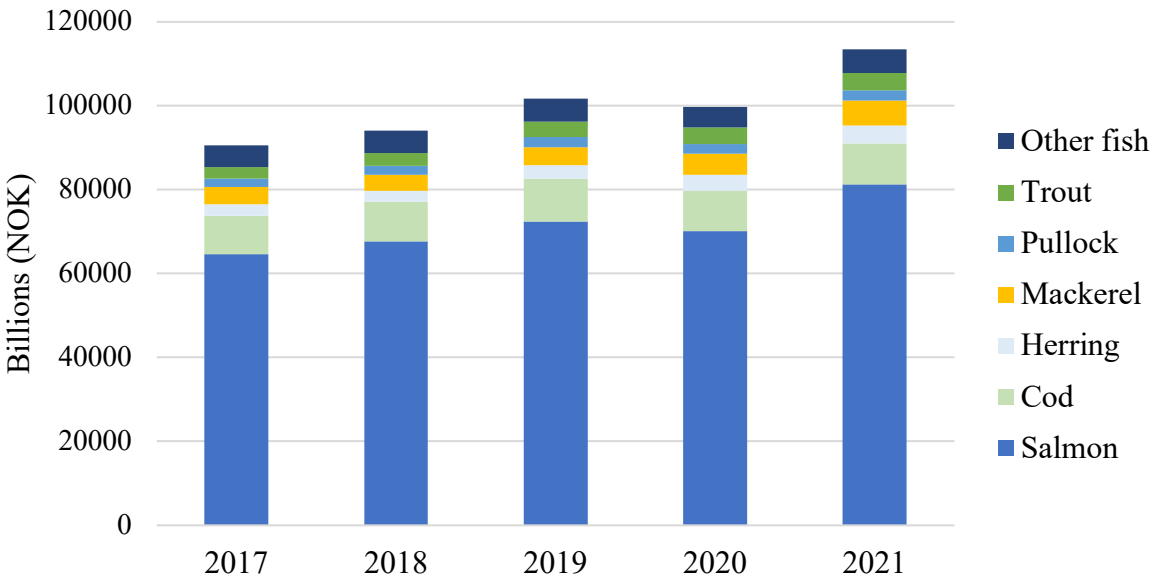


Figure 1.1 | Summary of export value of multiple fish species in Norway over the past five years. From 2017 to 2021, Atlantic salmon made up over 60 % of the total fish export in Norway (Statistisk sentralbyrå, 2022).

A rapid population growth leads to a proportionally increasing food demand, pressuring expansion of the aquatic industry. As an important source of essential nutrients, fisheries are predicted crucial for future sustainable production and global supply of animal protein (Costello et al., 2020). Multiple methods to induce growth in salmon have been approached using biotechnology, one of them through genetic engineering. AquAdvantage salmon is an example of a genetically modified fish with introduced genes of other fish species, which resulted in an increased growth rate. This was accomplished by introducing complementary DNA (cDNA) of growth hormone (GH) isolated from chinook salmon (*Oncorhynchus tshawhytscha*) and regulating promotor and terminal sequences from ocean pout (*Macrozoarces americanus*) antifreeze protein gene. With an extra copy of a GH, the transgenic fish was able to produce GH constantly throughout the year. Growing all year round, and not just during the spring and summer, it could reach the market weight (4-6 kg) a year ahead of the wild-type salmon (Yaskowiak et al., 2006). This fish has later been approved for commercial use by the American Food and Drug Administration (FDA, 2022).

Another approach to increase the total biomass and make the salmon industry more efficient, was through selective breeding programmes. In Norway, these were initiated in the 70's, where wild species were chosen from Norwegian rivers based on growth rate. The programmes soon expanded to include additional commercially favourable traits, such as age at sexual maturation, resistance to common diseases, fat richness and filet colour (Thodesen & Gjedrem, 2006). After 50 years, the domesticated conspecifics have altered both age and size in relation to maturation. These are fitness-related traits that, if introduced to the wild populations, would cause long-term negative effects on the environment (Bolstad et al., 2017).

1.1.2 Challenges and advances of salmon farming

Most of the salmon is farmed in cages in the open sea, and escaped individuals cause a major threat to the wild population. Despite the farmed salmon reportedly have a lower capacity, some are able to spawn with wild salmon species, causing genetic variation in these populations (Havforskningsinstituttet, 2022). Threats include loss of biodiversity as a result of reduced survivability of the offspring and transmission of disease (Kayaci et al., 2015). Efforts to minimise the ecological impacts include genetic engineering of salmon, resulting in sterile fish, thereby creating biological barriers between the farmed and wild populations. The

first approach introduced, and the most commonly used method today is through triploidisation. This is achieved by pressure shock during the second cleavage after fertilisation. This results in an organism with three homologue chromosomes, and consequently sterility (Piferrer et al., 2009). Sterile fish have also been obtained using CRISPR-Cas9 to knock out *dead end* (*dnd*) in F0 generation salmon – a factor essential in vertebrates for primordial germ cell development. This fish, however, is not yet approved for commercial production (Wargelius et al., 2016).

A second major challenge in the salmon farming industry is the fish parasite salmon louse (*Lepeophtheirus salmonis*). The louse threatens both for the wild and farmed species, resulting in reduced growth, delayed puberty, and in worst case, death. As the lice prefers salt water, it is more abundant in farmed salmon, challenging the animal welfare and production, and are the cause of significant economic losses (Havforskningsinstituttet, 2022). To battle the parasite, both chemical, mechanical, and thermal strategies are used. The lice, however, have evolved resistance to the more common chemotherapeutants. Consequently, the non-chemical treatments have dominated the farming industry the last few years. With its high ability to adapt, we won't yet know whether the parasite will evolve e.g., heat resistance or improved ability to attach to the host tissue to circumvent the non-chemical treatment, and how this adaptions will affect the wild populations (Coates et al., 2021). To develop alternative methods to control the parasite, biotechnology is thought to play a central role. With knowledge of its omics and genetic annotations through next-generation technologies, one can see how the lice responds and adapts to the treatments, as well as assessing how susceptible the parasite is to treatment (Chavez-Mardones & Gallardo-Escarate, 2015).

Multiple characteristics of Atlantic salmon makes the species an interesting model for biological research. For one, it spawns in fresh water and migrates to the sea to feed, requiring a unique adaptational transformation (Folmar & Dickhoff, 1980). Moreover, the whole genome of the common ancestor of salmonoids underwent a duplication approximately 80 million years ago. Assembly of high quality of this genome makes salmon an attractive model species for research regarding post-whole genome duplication events, such as rediploidisation. High-throughput sequencing of the Atlantic salmon genome have also opened up for genome characterisation and gene editing in the species (Lien et al., 2016). Biotechnology can play an important role in many different aspects of the aquatic industry and genetic research. By creating effective barriers between the farmed and wild salmon

populations, genetic diversity may be preserved (Fedoroff et al., 2022). To reach the phenotype of interest, such as disease resistance, or development of effective vaccines, one can reduce the use of antibiotics (Vinitnantharat et al., 1999). To advance technology like this, establishment of cell cultures may play a crucial role. Cell cultures enables us to isolate cells from its initial environment, and perpetuate them, as they reproduce indefinitely. One can study the behaviour of the cell type of interest and alter the variables according to the experimental design (Arango et al., 2013).

1.2 Properties of a stem cell

A cell type of especial scientific interest is the stem cell. Stem cells can self-renew indefinitely, meaning they can produce daughter cells identical to themselves. They are unique in their ability to differentiate into specialised cells in the body. This differential potential, called potency, ranges from totipotent ("entirely") to unipotent ("one"), indicating how many different cell types they have the potential to mature into. Stem cells are found both in early development and in adult organisms but hold different properties. The adult stem cells are either multi- or unipotent, meaning they are more lineage-restricted and can mature into a few or a single cell type, respectively (Alison et al., 2002). One example of a multipotent, adult cell is the hematopoietic stem cell. They develop in the bone marrow and are able to differentiate into various types of blood cells, including platelets, and white- and red blood cells. After isolation from a healthy person, these cells can be reintroduced to an individual of a blood-related disorder. The cells then produce new, healthy blood cells. This method was introduced in 1963, and successfully cured a leukaemia patient (Mathe et al., 1963).

The stem cells isolated from early development have a greater potency compared to the adult stem cells. The first cell after fertilisation of an egg is totipotent and is the mother cell of the whole organism. After a few cell divisions, the embryo reaches blastula stage. The blastula consists of a trophoblast surrounding the inner cell mass (ICM) (Figure 1.2). The ICM contains pluripotent embryonic stem (ES) cells. These cells can differentiate into all the three germ layers that comprise and organism: mesoderm, endoderm, and ectoderm (Donovan & Gearhart, 2001).

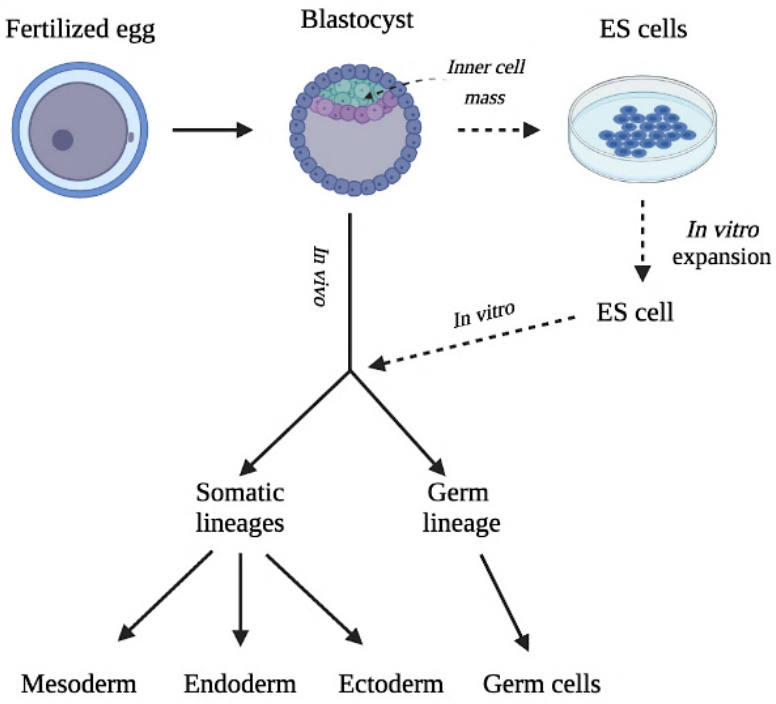


Figure 1.2 | ES cells are isolated from the ICM of a blastocyst. After fertilisation of an egg cell, the cells start dividing, forming a blastocyst, comprising the ICM. It is here the ES cells are isolated from. These cells can be used in culture and further expansion of cell populations. When injected into the blastocyst of the organism, the cells can resume their differentiation into different cell lineages.

With its potential to incorporate various cell lines, ES cells provide a unique opportunity for developmental research of differentiation from pluripotent stem cell to a specific cell type. Four decades ago, pluripotent cells were isolated from a mouse embryo, establishing the very first ES cell culture (Evans & Kaufman, 1981; Martin, 1981). From humans, the first ES cell line was derived in 1998, introducing a new field of regenerative medicine. This area of therapeutics aims to replace damaged tissues or organs, in hope to restore normal function. (Thomson et al., 1998). These cell lines are used in a broad range of research, and have accelerated the field of developmental biology and embryology (Amit et al., 2000).

When establishing and working with stem cell lines, it's important to monitor its stemness, so we know that the cells in the culture have not differentiated. This can be done with multiple approaches, including looking for genes that are highly expressed in this type of cells and down-regulated in differentiated ones, i.e., marker genes for pluripotency (Niwa, 2007).

1.2.1 Common stem cell marker genes

When characterising pluripotency, one would ideally induce the stem cells to differentiate and see whether they enable to form cell types of different tissues from the three germ layers. In humans, the teratoma assay is one such method and has been frequently used. The stem cells are injected under the skin of an immune-compromised mouse and after a few months, a benign tumour has preferably been growing. When isolated, the tumour is analysed and should contain cell types of all three germ layers, confirming a pluripotent phenotype (Wesselschmidt, 2011). This method, however, is controversial from an ethical point of view, and today alternative methods *in vitro* can work as substitutes. One of them is identifying pluripotency markers using RNAseq, qPCR or other qualitative methods (White et al., 2011). Even though many share common pluripotency maintaining genes, these markers are not necessarily the same for all species and are identified looking at gene expression data of the given organism. The markers are expressed at higher concentration in ES cells and at lower, if not at all, in differentiated ones (Park & Patel, 2010).

During early development, transcription factors are essential in determining the cell fate. They work as molecular switches and induce or repress gene expression of a target gene by binding directly to the DNA, altering the epigenome (Tsankov et al., 2015). A combination of different transcription factors preserves the pluripotent phenotype by repression of differentiation. In mammalian stem cells, the POU transcription factor *octamer-binding transcription factor 3/4* (*Oct3/4*, or *Pou5f1*) plays a key role (Niwa et al., 2000). Its non-mammalian homolog, *Pou2*, has been identified in medaka (*Oryzias latipes*) and share similar expression patterns with the murine *Oct4* (Liu et al., 2015). Its activity has also been reported in rohu carp (*Labeo rohita*) (Mohapatra et al., 2014) and Chinese sturgeon (*Acipenser sinensis*) (Ye et al., 2012). The gene is also used as an ES cell marker in Atlantic cod (*Gadus morhua*) (Holen et al., 2010), suggesting common features in teleost fish species. Its expression pattern is well studied in zebrafish (*Danio rerio*), where the homologue is called *Pou5f3*. The protein was confirmed in the same study to be in a synergetic relationship with SRY-box transcription factor B1 (Sox1B) (Kobayashi et al., 2018). The infamous Sox genes are known to partner up with *Oct4* also in murine ES cells. Oct4 form a synergetic partnership with a central transcription factor of the SOX family, called Sox2. The complex targets various genes, including *fibroblast growth factor 4* (*FGF-4*), which is believed to enhance self-renewal and support proliferation (Ambrosetti et al., 1997).

Another component of the Oct 3/4-Sox2 cooperation, is Krüppel-like factor 4 (Klf4). Together the three transcriptions factors are essential for somatic cell reprogramming of induced pluripotent stem cells, as well as co-occupying the *Nanog* promotor (Wei et al., 2009). *Nanog* transcribes a homeodomain protein that is quickly downregulated when ES cells differentiate *in vivo*. A deletion of the gene has shown to result in loss of pluripotency in ES cells of mice, and induce cell differentiation (Mitsui et al., 2003). Moreover, the Nanog protein is proven to target promoters of a total of 353 genes in human ES cells, also occupied by Oct4 and Sox2. These target genes encode important key transcription factors involved in development, and are targeted either by regulating their expression directly or signalling pathways related to their expression levels (Boyer et al., 2005).

Like some genes specify in maintaining pluripotent phenotype, others play essential roles in cell fate towards differentiation. Among these, the GATA factors contribute to induce endoderm formation, and are used as markers for early development in many vertebrates. The GATA family of transcription factors are able to bind the DNA and alter the chromatin accessibility (Heslop et al., 2021). Of this family, especially *Gata4* and *Gata6* are of interest, as their expression is shown to increase during differentiation (Fujikura et al., 2002). When their expression is forced, ES cells have shown to differentiate into extraembryonic endoderm (Mitsui et al., 2003). These genes have also been reported to have increased expression in *Nanog*-knocked out cells, indicating a transcriptional repression before differentiation. *Gata3* is another family member shown important for differentiation. During zebrafish development, among other vertebrates, *Gata3* expression increases in early differentiation of neuronal cells, making it a suitable marker for ectoderm tissue (Neave et al., 1995).

Other markers used for differentiation is *Forkhead box protein A2* (*FoxA2)*, *Sox17* and *Neuronal cell adhesion molecule 1 (Ncam1)*. The two formers are associated with endoderm, with *FoxA2* being a target of direct activation by the *Sox17* protein expression. The Sox17 protein is known to be essential for endodermal differentiation (Sinner et al., 2004). *Ncam1* on the other hand, is used as a mesodermal marker, and is known to play parts in adhesions between cells, as well as between the cell and matrix proteins. Activities include neurite development (i.e., axons, dendrites) and cell migration (Doherty & Walsh, 1996).

In addition to being important for verification of stemness in culture, knowledge of marker genes for pluripotency have contributed to valuable tools such as dedifferentiation of adult cells, also called induced pluripotent stem cells (iPSCs). These cells have had their differentiated phenotype reversed, resulting in cells with characteristics of an embryonic stem cell (Takahashi & Yamanaka, 2006) .

1.2.2 Induced pluripotent stem cells

In 2006, the Japanese scientist Shinya Yamanaka, successfully generated iPSCs from mouse embryonic fibroblast cells. The reprogramming from differentiated to pluripotent phenotype was achieved with use of retroviral transduction, successfully inducing gene expression of four transcription factors: *Oct4*, *Sox2*, *c-Myc*, and *Klf4*. The paper showed that *Nanog*, though infamously known for maintaining pluripotency in ES cells, was not needed when generating iPSCs. *Klf4* and *c-Myc*, however, were considered essential (Ibid.). Within a year, the derivation of human iPSCs (hiPSCs) was established using the same technology (Takahashi et al., 2007; Yu et al., 2007). Yu et al. however, used a different cocktail of genes, comprising *Oct4*, *Sox2*, *Nanog*, and *Lin28* (2007).

A main motivation for generating iPSCs was its ability to circumvent the ethical issues regarding stem cell research on embryos. These cells can instead easily be derived from available adipose or epithelial tissue (Brind'Amour, 2009). Stem cell lines are, however, no longer limited only to mammals. Even though applications such as regenerative medicine are not as relevant to non-mammalian species, ES technology in unconventional model species aid research in cellular, molecular, and developmental biology. Cultures of iPSCs have for example been established in thirteen-lined ground squirrel (*Ictidomys tridecemlineatus*), and is a valuable model when studying mammalian hibernation and their metabolic depression (Ou et al., 2019). iPSCs technology has also been suggested to contribute to the conservation of biological diversity, by recovery of endangered species, as done in three avian species in Japan (Katayama et al., 2022). Other applications include drug discovery and disease modelling, as well as the study of embryonic development and tissue regeneration. In regard to the latter, zebrafish is a popular model species, as it is known for its regenerative properties (Knapp & Tanaka, 2012). Model organisms, such as zebrafish and medaka have also contributed to valuable genetic information related to our understanding of human genetics and therapeutic strategies (Chowdhury et al., 2022).

1.3 Fish stem cell technology

The medaka and zebrafish are the main model species for fish genetics. iPSC cultures have been established from adult zebrafish fibroblasts (Peng et al., 2019), but not yet for medaka. In terms of ES cell technologies, both species serve great advantages for vertebrate research with their short generation time, and production of large embryos quantity which develop rapidly. Fertilisation occurs externally, producing large transparent embryos, convenient for phenotypic observation of development under the microscope. With these properties, both species have been, and continuous to be, important for gene function analysis relevant to mammals (Alvarez et al., 2007).

ES cell cultures in non-model species have, among others, been established in Atlantic cod (Holen et al., 2010), sea perch (*Lateolabrax japonicus*) (Song-Lin Chen et al., 2003), sea bass (Parameswaran et al., 2007), and red sea bream (*Chrysophrys major*) (S.-L. Chen et al., 2003). A prerequisite to establish such a culture is the knowledge of the transcription factors active in a pluripotent stem cell. These genes are well studied in model organisms but are not necessarily conserved in all species. Many genes are associated with pluripotency, and different reprogramming cocktails can efficiently induce pluripotency (Yu et al., 2007). Generating the iPSC culture in non-model species, however, follows the same basic principles.

Genes used as markers in fish stem cells include multiple homologs to the mammalian species, including *Nanog* and *Oct4* in medaka (Yi et al., 2009), and the *Oct4* homolog *Pou2* in Atlantic cod (Holen et al., 2010). Other pluripotency genes common in mammals have been identified in fish, such as *Klf4, Sox2,* and *Transcription factor 3 (Tcf3)*. Their function in fish pluripotency, however, is not well characterised (Sanchez-Sanchez et al., 2011).

Pluripotency is mostly validated through gene expression analysis where expression of marker genes for pluripotency varies between pluripotent stem cells and differentiated ones. To normalise the gene expression in different samples, one often uses housekeeping genes as endogenous controls. Essential genes are commonly used – that is genes that are vital for cell survival. These genes are usually expressed at constant levels across tissues of an organism, and ideally under different conditions (Thellin et al., 1999).

Glyceraldehyde-3P-dehydrogenase (*GAPDH*) is commonly used as a reference in mammals, as it is a critical enzyme in glycolysis. However, reports have shown that its mRNA expression is variable under different physiological conditions, making the gene inappropriate as a reference gene for many experimental studies (Piechaczyk et al., 1984). When regulating experimental settings, choosing the wrong reference gene could thus alter the findings. No universal gene have yet been found ideal to use in all developmental stages and experimental conditions, and the control gene thus varies in between tissue samples (Kozera & Rapacz, 2013).

In salmon, transcription levels of the most common control genes used have been evaluated in eight different adult tissues and in fish undergoing smoltification. The genes tested encoded GAPDH, β-actin, 18S rRNA, S20 ribosomal protein (Rps20), and two paralogs of elongation factor 1A ($EF1A_A$, $EF1A_B$). Findings included different expression values within tissues, as well as between tissues and samples taken of fish going through smoltification. The paper concluded with $EF1A_A$ and $EF1A_B$ being the best candidates as reference genes (Olsvik et al., 2005).

In addition to the endogenous expression, the cell relies on extracellular cues from the environment (Wan et al., 2021). The specific conditions and factors required for successful iPSC generation may need to be optimised for each species.

1.4 Establishment of a cell culture

Cell cultures are isolated cells growing in artificial environments, meant to simulate their natural ones, and are an essential technology in the study of biological processes. The cells are derived from one of two sources: a cell line already established, or they are isolated directly from the source, also referred to as primary cell cultures. Due to their similarity of cells *in vivo*, the latter serve as an ideal model for studying the physiology and biochemistry of cells. The cells of the primary culture, however, have a shorter lifespan, and can only be maintained *in vitro* for a short period of time. Cell lines, on the other hand, have acquired the ability proliferate indefinitely, creating an unlimited supply of research material. As such, cell lines have sped up the research of many biological processes. However, these cells usually differ from the cells in the tissue *in vivo* and are mainly used as models for general research (Oyeleye et al., 2016).

Both marine and freshwater fish cell lines have been established from a broad range of tissues. Many of them were initiated to detect and isolate pathogens related to the species, and to look at the cellular response to infection. Today, established ES cell cultures are mainly used in genetic and molecular analysis of development in vertebrates (Pandey, 2013). As stem cells share different properties, effective optimisation of stem cell cultures necessitates a comprehensive understanding of the *in vivo* microenvironment of the cell type of interest.

1.4.1 Stem cell niche

Given the distinctive characteristics of stem cells, including their ability to balance self-renewal and differentiation, optimising cell culture conditions requires careful consideration of various factors. Different types of stem cells, such as ESCs and adult stem cells, may require different cell culture compositions to obtain their properties (Chou et al., 2008). *In vivo*, stem cells are located in microenvironments, so-called stem cell niches. In adult stem cell niches, these areas promote differentiation when tissue is injured. When quiet, cues for maintenance of stemness are initiated, prohibiting differentiation. Among others, these specialised microenvironments consist of signalling factors, cell-cell contacts, matrix proteins, and niche supporting cells (Figure 1.3). This system is, in combination with neural and systemic stimulation, regulating the homeostasis of the environment, including oxygen level and hormone influx (Peerani & Zandstra, 2010).

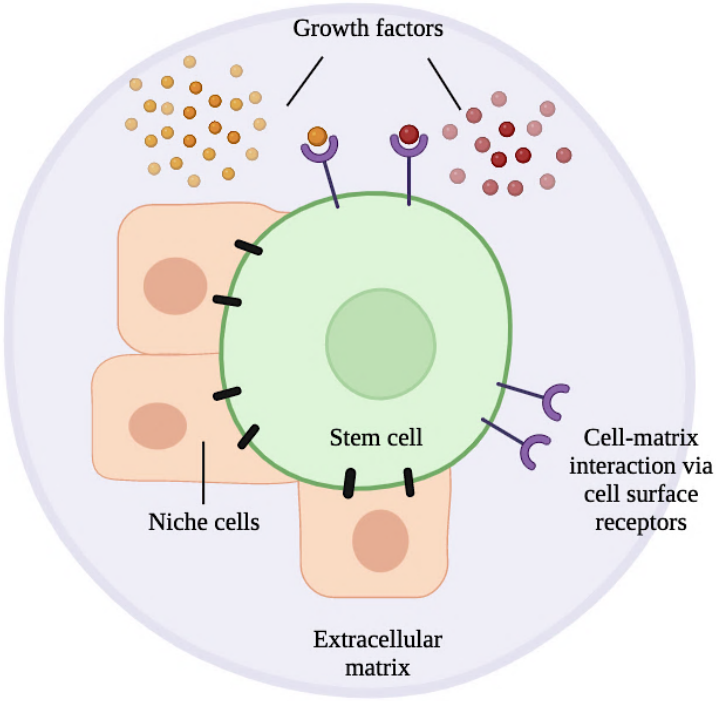


Figure 1.3 | Stem cell niche overview. The stem cells are in a microenvironment and rely on signals from different growth factors. Stem cells have also found to attach to extracellular matrix via cell surface receptors and niche cells, where they can receive signals through interactions (Clark & Pazdernik, 2016).

The importance of the niche has been demonstrated through experiments where ESCs have been isolated from one mouse and injected into the blastocyst of another. These cells resumed their stem cell behaviour (Fuchs et al., 2004). When injected under the skin however, the cells form teratomas – benign tumours comprising a mixture of tissues, such as hair, muscle, and bone. The formation of a teratoma shows pluripotent characteristics, as the tumour comprise all three germ layers. The induced formation of a teratoma is a commonly used assay for pluripotency *in vivo* (Wesselschmidt, 2011).

Common for all cell lines established is that they have been isolated from the organism and are grown in artificial environments with supplements thought to maintain their properties. Not only are basic incubators such as adequate pH and temperature important, but also nutrients, growth factors and antibiotic and -fungal agents (Chandra et al., 2022). Dulbecco's modified eagle medium (DMEM) or Leibovitz L-15 medium (L-15), supplemented with foetal bovine serum (FBS) comes close as a basic culture medium in many vertebrates, including fish species (Pandey, 2013). FBS is commonly used as a growth supplement in cell cultures. As a natural product, its complete composition is not known, and may vary from

batch to batch, and even in between manufacturers. However, the serum is thought to contain high levels of factors inducing growth, and have low levels of growth inhibiting factors (Zheng et al., 2006). FBS have been reported to be important for supporting cell proliferation and metabolism. However, the serum have also shown to induce differentiation, as seen in human mesenchymal stem cells (Shahdadfar et al., 2005). As with other components of a cell culture media, the serum concentration must be optimised for different types of cells and aims. For fish cell culture media, 10 % of serum is typically applied (Batish et al., 2022), but both 7.5 % and 15 % have been used in induced pluripotent stem cells cultures of zebrafish and embryonic stem cell cultures of sea perch, respectively (Song-Lin Chen et al., 2003; Peng et al., 2019).

Multiple media are optimised to exhibit optimal growth in the cell line of interest. Regarding fish ES technology, many common features are observed in their experimental setup. Looking at ES medaka cells isolated in 1996, the cells were cultured in feeder-free conditions and in a rich medium containing basic fibroblast growth factor (bFGF/FGF-2) and embryo extract from fish (Hong et al., 1996). The medaka ESC culture have been used as foundation for later ESC cultures in fish, including Atlantic cod (Holen et al., 2010), red sea bream (S.-L. Chen et al., 2003), and Nile tilapia (*Oreochromis niloticus*) (Fan et al., 2017), all using media comprising similar or equivalent components in terms of nutrients, inhibitory- and growth factors, serum, and fish embryo extract. The latter is thought to support growth by providing unidentified factors with mitogenic effects, i.e., factors inducing mitosis. This is confirmed in multiple species, including rainbow trout, medaka, and zebrafish (Hong et al., 1996). Growth factors, such as bFGF, are in many cultures a critical component of the media as they are proteins that stimulate cell growth. The FGF-family of factors are secreted in tissues and act as signalling molecules in important cellular mechanisms, such as proliferation, differentiation, metabolism, and survival. They have proven important in multiple tissues, including skin, muscle, blood, adipose, and bone, among others. The factor has shown important for self-renewal and is commonly used in human ES cell cultures (Dvorak et al., 2006).

In addition to nutrients, serum and specific growth factors, niche design can include materials such as the extracellular matrix (ECM). Cells interact with the ECM through integrins or other cell surface receptors. Cell-ECM interactions provide cell attachment, guides cell migration, acts as a reservoir for growth factors, and is used as a scaffold in tissue generation (Chen et al., 2007). Even though the ECM is thought important when inducing lineage-specific differentiation, the choice of ECM coating can, however, also restrict the differentiation efficiency, i.e., contribute to maintain a pluripotent phenotype (Ahmed & Ffrench-Constant, 2016).

1.4.2 Cell culture coatings

The first human embryonic stem cell line was cultured on mouse embryonic fibroblast feeder layers and showed to differentiate when they were cultured without a feed layer (Thomson et al., 1998). Feeder cells are made up of a layer of non-dividing cells that promote another cells' proliferation by releasing extracellular secretions, such as growth factors, to the culture medium (Llames et al., 2015). However, co-culturing of cells in this manner has limitations, including an increased risk of pathogen infection and batch-to-batch variation (Yang et al., 2012). As an alternative, artificial coating of ECM proteins is used. ECM is a fibrillar network of proteins surrounding the cells and support important cell functions such as cell adhesion and migrations, as well as proliferation and differentiation. This matrix is used to simulate the tissue environment *in vivo* and include proteins like collagen and fibronectin. Seeding cells on ECM proteins has proven to increase the proliferation rate but give variable results between different types of cells (Gospodarowicz et al., 1980). The ECM is believed to have significant implications in tissue engineering and regenerative medicine, as it can be utilised to manipulate cell fate (Wan et al., 2021).

A protein commonly used for coating is fibronectin. Fibronectin is a glycoprotein of the ECM that binds to integrins in the cell membrane, and are involved in cell growth, adhesion, and differentiation. It has also been demonstrated that the protein increases the proliferation of human lung carcinoma cells and bronchial epithelial cells. One of its observed effects was found to be the induction of protein expression of *c-Myc* (Han & Roman, 2006), an oncogene known to regulate proliferation (Miller et al., 2012).

Gelatin, derived from the ECM protein collagen, is another commonly used protein for coating in cell culture. The arginine-glutamine-aspartic acid (RGD) peptide sequence, which is present in the protein, enables gelatin to exhibit adhesive properties through binding to cell receptors. Through this interaction, gelatin can create an environment that resembles the ECM surrounding the cell and support processes such as cell adhesion, proliferation, and nutrient exchange. Because of this, gelatin is considered a viable option for use in tissue engineering applications (Mushtaq et al., 2022).

Coating techniques have been shown to play a critical role in maintaining pluripotency of stem cells, as well as in strategies aimed at inducing lineage-specific differentiation.
A study conducted in Atlantic cod ES cells demonstrated that cells maintained their undifferentiated state in small wells coated with gelatin and poly D-vinyl, whereas fibronectin- and laminin-coated wells showed a higher frequency of differentiation (Holen et al., 2010). This may, however, not be the case for salmon ES cells.

Ultimately, different stem cell types have distinct properties and require different experimental designs to be effectively cultured and expanded. This optimisation process involves testing different media formulations, growth factors, and culture substrates to maximise cell viability and proliferation. Additionally, to validate the stemness of the culture, gene expression analysis can be performed to assess the expression of key marker genes for pluripotency and differentiation. This analysis can also provide insight into the differentiation potential of the stem cells and help optimise the differentiation protocols.

1.5 Thesis aim

This project will focus on identifying and characterising experimental conditions required for the maintenance of Atlantic salmon embryonic cell pluripotency. This involves testing of cell culture conditions that supress their differentiation while promoting the proliferation of salmon embryonic stem cells. The second aim of the project is to identify the pluripotency and differentiation markers, as well as suitable reference genes using a RNAseq database. The main methods to be used are summarised in Figure 1.4.

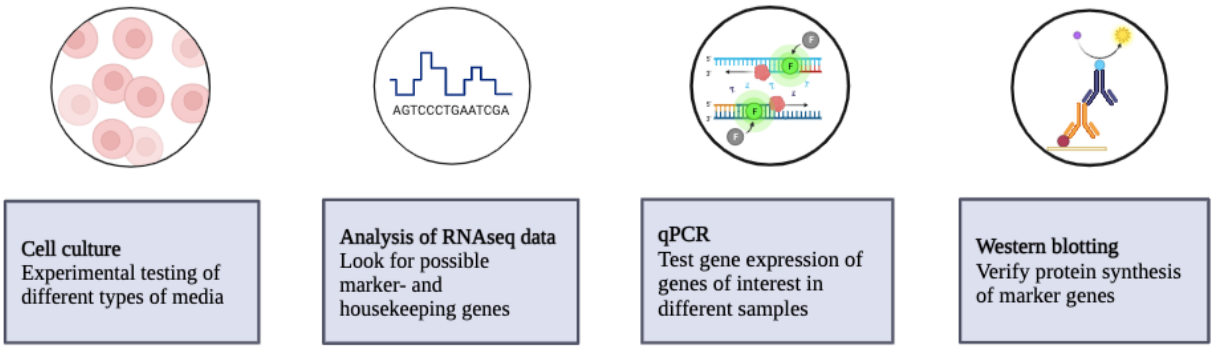


Figure 1.4 | Summary of the main methods used in lab work. One of the main goals is to establish a cell culture for salmon ES cells. Alongside cell lab work, RNAseq data will be analysed in hope to find potential marker genes for pluripotency and differentiation, as well as a suitable endogenous control gene. Genes of interest are tested using qPCR and Western blot assays, to confirm transcription and protein synthesis.

2 MATERIALS AND METHODS

2.1 Atlantic salmon ESC isolation and culturing

Unfertilised egg batches were ordered from MOWI ASA, department Tveitevåg, Norway. To look at how the cells respond to different compositions of media, five different media are here evaluated. These are mainly based on media already seen in other fish ES or iPS cell cultures.

2.1.1 Fertilisation of eggs

Upon arrival, a monolayer of eggs cells was distributed in pre-cooled, washed jars. To each jar, 300-500 µL of milt were added dropwise, together with some water. After careful mixing with a spoon, the jar was covered with aluminium foil and left to sit for 4-5 minutes, to give the sperm time to enter the egg. Next, the eggs were washed twice with water. Water was then added to cover about 0.5 cm over the cells. Then, the jars were covered with aluminium foil, and stored at 4 or 8 ºC.

After fertilisation, the embryo will comprise undifferentiated pluripotent stem cells. These will be used for culturing in different compositions of media. Based on Gorodilovs research findings, one could follow the embryo development and collect samples at blastulation, gastrulation, and somitogenesis stage (1996).

Both cells incubated at 4 ºC and 8 ºC were used for culturing and sample collection. Cells used for culturing were collected at day 8-12 at 4 ºC, and day 4-5 at 8 ºC. Cells collected for samples were collected as stated in Table 2.1 below.

Table 2.1 | Days of collecting samples from fertilised eggs incubated at 4 ºC and 8 ºC. As development is tightly connected to temperature (Gorodilov, 1996), samples were collected at different times as to where the cells were stored. Samples were collected for late blastulation, early/mid gastrulation and early somitogenesis. The latter were only collected from jars incubated at 8 ºC.

Sample type	**4 ºC**	**8 ºC**
Late blastulation	Day 12-14	Day 5-7
Early/mid gastrulation	Day 16-19	Day 8-10
Early somitogenesis	-	Day 13-15

2.1.2 Isolation of ESCs for culturing and sample collection

The fertilised eggs used for culturing were collected at blastulation stage and placed in a petri dish with distilled water on ice. After a brief wash in 70 % ethanol, the eggs were transferred onto a petri dish with phosphate-buffered saline (PBS). In PBS, fine forceps were then used to break the outer chorion. The cells were so rinsed for fat vesicles and collected using a pipette. They were then seeded in either of the media in the tables below (Table 2.2-2.6). In a 96-well plate, three different well surfaces were tested. 1/3rd of the cells were cultured on 2 µg/ml fibronectin-coated wells, 1/3rd in 0.2 % gelatin-coated wells (Sigma chemical G-7041), and the rest were used as control on non-coated wells. To coat the wells, 50 µl gelatin-solution and 25 µl of fibronectin was used. The plate was then lightly rocket to ensure the fluid would cover all the well area. After it was incubated for at least one hour at room temperature, the excessed fluid was removed before media was added. For each well, only a single embryo was cultured. The cells were cultured at 15 °C without CO_2.

As primary cultures are prone to infection (Vierck et al., 2000), both 1x and 2x Penicillium/Streptomycin were tested, as well as the antibiotic Gentamycin (50 µg/ml). As no significant difference was observed, 1 x Penicillium/Streptomycin was used.

Cell collected as samples were collected at their respective cell stage (Table 2.1) and placed on ice in a petri dish with distilled water. The eggs were then transferred to a petri dish with PBS, where fine forceps were used to break the outer chorion. After brief rinsing for fat vesicles, the cells were collected using a pipette. Ten cells were collected for each sample and spun down in a mini centrifuge. The PBS was discarded, leaving only a pellet of cells. Samples were stored at -80 °C or used right away for either sample prep for Western blotting or RNA extraction for qPCR.

The 2i medium (Table 2.4) is initially provided with two differentiation inhibitors (i.e., the name 2i): glycogen synthase kinase 3β (GSK3β) and MEK1/2 (Romito & Cobellis, 2016). These factors, however, have shown to kill the salmon ES cells if both are supplemented in the media, and were here excluded (main project, unpublished).

Table 2.2 | Leibovitz's L-15 (L-15) medium. The table includes the materials and the final concentrations used in L-15 medium.

Product no.	Distributor	Materials	Final concentration
31415029	ThermoSci	Leibovitz's L-15 Medium	
35050061	ThermoSci	GlutaMAX supplement	1x
F7524	Sigma	Foetal Bovine Serum	2 %
P4333	Sigma	Penicillium/Streptomycin	1 x
A2942	Sigma	Amphotericin B	0.5 x

Table 2.3 | Dulbecco's Modified Eagle (DMEM) medium. The table includes the materials and the final concentrations used in DMEM medium.

Product no.	Distributor	Materials	Final concentration
11965092	ThermoSci	Dulbecco's Modified Eagle Medium	
H3375	Sigma	HEPES	20 mM
P4333	Sigma	Penicillium/Streptomycin	1 x
G8540	Sigma	L-Glutamine	2 mM
F7524	Sigma	Foetal Bovine Serum	2 %
S5261	Sigma	Sodium Selenite	2 mM
11360039	ThermoSci	Sodium Pyruvate	1 mM
F3685	Sigma	Basic Fibroblast Growth Factor	5 ng/mL
M3148	Sigma	2-Mercapoethanol	50 μM
M7145	Sigma	Non-Essential Amino Acids	1 mM
		Embryo extract	1 %, 100 embryos/mL
A2942	Sigma	Amphotericin B	0.5 x

Table 2.4 | 2i medium. The table includes the materials and the final concentrations used in 2i medium.

Product no.	Distributor	Materials	Final concentration
SF016	Sigma	ESGRO 2i Medium w/o GSK3β and MEK1/2	
P4333	Sigma	Penicillium/Streptomycin	1 x
A2942	Sigma	Amphotericin B	0.5 x

Table 2.5 | Embryo extract (EE) medium. The table includes the materials and the final concentrations used in EE medium.

Product no.	Distributor	Materials	Final concentration
31415029	ThermoSci	Leibovitz's L-15 Medium	
G8540	Sigma	L-Glutamine	2 mM
P4333	Sigma	Penicillium/Streptomycin	1 x
C5080	Sigma	$CaCl_2 \cdot 2H_2O$	0.8 mM
		Embryo extract	1 %, 100 embryos/mL
F7524	Sigma	Foetal Bovine Serum	2 %
A2942	Sigma	Amphotericin B	0.5 x

Table 2.6 | bFGF medium. The table includes the materials and the final concentrations used in bFGF medium.

Product no.	Distributor	Materials	Final concentration
31415029	ThermoSci	Leibovitz's L-15 Medium	
35050061	ThermoSci	GlutaMAX supplement	1x
P4333	Sigma	Penicillium/Streptomycin	1 x
F7524	Sigma	Foetal Bovine Serum	2 %
F3685	Sigma	Basic Fibroblast Growth Factor	5 ng/mL
A2942	Sigma	Amphotericin B	3.1 x

2.1.3 Preparation of embryo extract

For the DMEM- and EE medium (Table 2.3, 2.5), embryo extract would need to be prepared. To make this, about 100 cells were collected at mid-blastula stage. They were then smashed using a KIMBLE Dounce tissue grinder set (Cat. No.: D9938; Sigma). To crush the eggs more easily, about 0.5 mL PBS were added for every 50 eggs. Next, the mix was further homogenised using a TissueRuptor II (Cat. No.: 9002756; Qiagen). After three cycles of freezing in -80 °C freezer and thawing in water bath (37 °C), the cells were centrifuged for 30 minutes at 15000 rpm at 4°C. The supernatant was then collected and filtered through a 0.2 µm filter, and the pellet discarded. The extract was stored at -20 °C.

2.1.4 Change of media

Media was changed every three to four days. As we don't know the cells sensitivity to mechanic stress, different approaches to change the media were tested. The first method was the gentler one, where the media was exchanged without touching the cells. Half of the media were drawn and replaced by new media using a pipette. This method also allowed for the cells to stay in their blastodisc as within the developing egg cell.

The second approach was to transfer the media with cells into Eppendorf tubes. These were then centrifuged for 5 minutes at 100 x *g*. Meanwhile, the wells were rinsed with PBS to remove waste products and cell debris. After centrifugation, half of the supernatant were removed and replaced with new media. The suspension was then transferred to the same well as before.

2.2 Analysis of RNAseq data

The RNAseq data were partly provided from the main project (unpublished), and partly from AQUA-FAANG. Samples were collected from 18 stages of early development of salmon, ranging from one-cell stage to early somitogenesis, each with three replicates. The data from RNAseq was used to make figures in RStudio and assess possible marker and control genes.

2.2.1 Identification of reference genes and differentiation markers

Methods to extract the candidates from the RNAseq data were inspired by a code available at GitHub, presenting a workflow to sort the best denominator in sequencing data. This script is an extension of 'sleuth'- a tool for analysis of data from RNAseq-experiments - and is called 'sleuth-additive log ratio transformation' (sleuth-ALR) (McGee et al., 2019). The script uses compositional normalisation of the data obtained from RNAseq so that the information acquired is not relative, but absolute in terms of number of transcripts per cell. By calculating the coefficient of variation (cov), the script is able to select candidate reference gene based on variation in the dataset (McGee, 2019). Cov is a value that compares the standard deviation to the mean of the dataset, and is commonly used to look at variation in phenotypic traits (Albert & Zhang, 2010).

Using RStudio, a new table was made with the mean TPM values of the replicates at all the cell stages. Cov was so calculated by dividing the standard deviation on the mean. Next, the data was arranged after ascending cov, with an additional criterion of at least 5 TPM mean. This was done to exclude any genes of low or no expression data (Supplemental code 7.1). Cov gives a relative variation of the expression levels in the samples, where higher values show greater dispersion. When looking for control genes, this value should ideally be as low as possible (Olsvik et al., 2005).

To select for differentiation markers, genes were chosen based on their expression rate at late blastulation and early gastrulation. A code was run to look for genes with at least four times greater TPM value at gastrulation compared to blastulation. The value also had to be bigger than four. The data frame was then sorted based on the highest value in the late blastulation column, and the top 10 genes were selected (Supplemental code 7.2).

2.3 Gene expression analysis using qPCR

Analysis of the mRNA transcript provides an overview of the expressed genes in the cells. To detect the RNA present, qPCR is used. The workflow (Figure 2.1) starts by collecting the cells at blastulation and gastrulation stage. RNA is then extracted, and used for cDNA synthesis, using a reverse transcriptase. The cDNA is so used for qPCR analysis, where the relative gene expression data is measured.

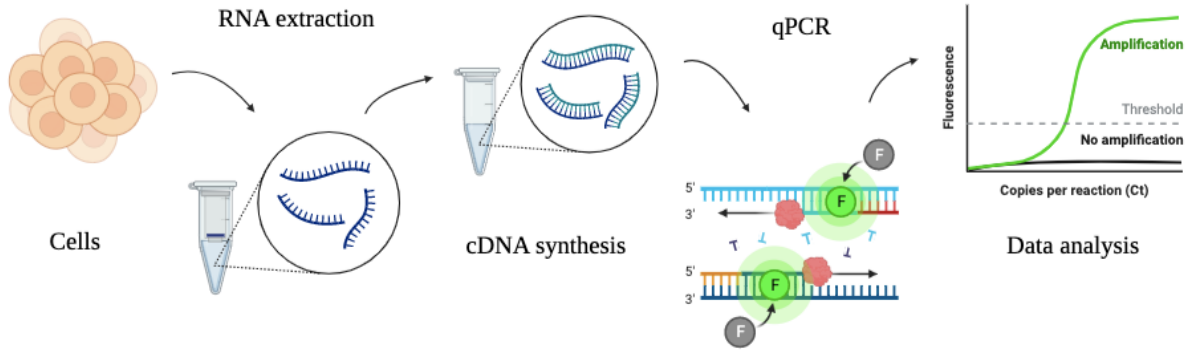


Figure 2.1 | The workflow of gene expression analysis using qPCR. The cells are collected at blastulation and gastrulation stage before their RNA is extracted. After cDNA synthesis, the samples are used for qPCR analysis. The output data analysis will then reveal the relative gene expression.

2.3.1 RNA isolation, quality control and cDNA synthesis

Samples were collected at late blastulation and early-to-mid gastrulation stage and stored in a -80 °C freezer. RNA was extracted using RNeasy Mini Kit according to the manufactures' protocol (Cat. No.: 74104; Qiagen). This kit uses guanidine-isothiocyanate as a lysis agent, disrupting the RNase activity which allows for RNA purification. The sample is loaded onto a RNeasy Mini spin column, where the RNA binds to a silica membrane, and the contaminants are washed off. A few adjustments were made to the protocol in the later steps. When eluting the RNA stuck in the membrane, prewarmed water (55-60 °C) was added. The column was then left at room temperature for 10 minutes, before continuing to centrifugation in the last step. For the RNA to better perform in following analyses, NanoDrop (ND8000.2.3.3) is used for quality control.

Before performing qPCR, the isolated RNA needs to be converted to cDNA. This was done using the iScript™ cDNA Synthesis Kit (Cat no.: 1708890; Bio-Rad). The kit includes reverse transcriptase, an enzyme that enables to generate DNA from the RNA templates. The kit suggests an RNA concentration of 100 fg-1 μg, where 800 ng was used for all samples. The cDNA sample was stored at -20 °C.

2.3.2 qPCR for gene expression analysis

Before running the final qPCR, primers were ordered, and their efficiency measured. To measure the relative number of gene copies in a sample and test the primer efficiency, a 1:4 dilution of all primers was included in all runs to make the standard curve.

To select the region of interest for relative quantification, a set of primers complementary to the sequence is needed. The primers are throughout the qPCR cycle extended by a DNA polymerase. The same primers are then again used on the extension products, thus resulting in an exponential amplification of DNA sequences.

The primers were designed using NCBIs "Primer-BLAST" tool. By implanting the NCBI refseq accession number, the primers were selected after entering certain parameters. For the PCR product size, the number of amino acids were set to 100-250 base pairs. Primer melting minimal temperature to 59 °C, optimum temperature to 62 °C, maximum temperature to 65 °C, and the melting temperature difference to three degrees. To avoid gDNA contamination, the primer was set to span at least one exon-exon junction if multiple exons were present in the gene. To anneal the exon junction, an exon junction match was set in both 5′ and 3′ end. In the 5′ end, the minimum number of matches was set to seven base pairs. For the 3′ end, the minimum match was set to four base pairs, with a maximum of eight base pairs. Next, a database for the search was selected. This was set to "Refseq mRNA", which is the NCBI's Reference Sequence collection of transcripts. Then the organism was specified by entering "8030", which is the taxonomic ID of *Salmo salar*. To further specify the primers, two base pairs was set to mismatch with unintentional targets within the last five base pairs at the 3′end. In the same section, if six or more mismatches were found within the primer sequence, the targets are to be ignored. Lastly, the maximum amplicon target size was set to 4000 base pairs. After the BLAST search, only primers specific to the target were selected (Table 2.7).

Table 2.7 | Primers used in qPCR application. Distributor: Thermo Scientific.

Gene ID	Genes	Forward primers (5′ → 3′)	Reverse primers (5′→ 3′)
100217346	*Klf4*	AGTCTGCGATTTTCCATGACTGG	AGGAGTAAGCCTGCTGAGAAGA
106590258	*Nanog*	GACCCACACCTCACCAGATTC	CTGGGGAAGATGGGGGCAAT
100196696	*Sox2*	TGTACCCGGCACGACGATTA	TGCAGAGTTCCTACCGTTCC
106569532	*Tcf3a*	AAACTGCTCATTCTGCACCAGG	GTACAATCAACAGGAGCTGGGT
106563125	*Pou5f1*	CCTGTGTCTGGAGAGAGATGTGG	CCGGTGTGGCTGGATAGGT
100329176	*Gata3*	TTTCCCCTACGCATCTCGCT	TCTTCGTTACCTTCGTTCACCTGT
106567728	*Ncam1a*	TCGGCCTGAGACATACGAGA	CTGGATCTTGGGAGCATAGCG
100500789	*Gata4*	TCCATTCGACAGCTCCGTCC	TCTTCGAGATGCAGACAGCCG
106610797	*FoxA2*	CTGAAAGACACACATGCGCC	GAGGTGTAACACTCGGGCT
100195814	*Rps20 (control)*	CGGGCAGTTGCACGTGTAGT	GTGATGCGCAGAGTCTTGGTG
106576970	*Eef1a1a (control)*	AAATATGGCTTGGCTTCCCCAG	AAAGTGACAGTGAGAGAGTCTGCG

Eukaryotic translation elongation factor 1 alpha 1a (*Eef1a1a*) and *Ribosomal protein S20* (*Rps20*) were selected as control primers based on a reference gene evaluation performed in salmon by Olsvik et al. (2005).In addition, primers were ordered based on RNAseq data. Possible differentiation genes are showed in Table 2.8, while Table 2.9 comprise suggested endogenous controls.

Table 2.8 | Primers for the possible differentiation markers, based on RNAseq data. Distributor: Thermo Scientific.

Gene ID	Genes	Forward primers (5′ → 3′)	Reverse primers (5′→ 3′)
100196111	*Apoa1*	CAAACACCCACCAGACCACCA	CCTTCACCTGAGCCATGTACTCC
106605287	*Afp4*	AACCCCCAGAAGTAAAGATCCACC	CGGTCAGGGGCGTGATTTCAG
100195744	*Rbp4*	TGCCCACGGCAGAGTTATCA	CCCAGTGGTCATCGTTTCCAG
106583889	*Hebp2*	CGCATCAAGATCAATCGTCTCACG	CCCTGCCAAATAAACCATTGTGC
100195186	*K2c8*	CCCCATCACCGCTGTCCAAG	CTGTTCCAGGAAGCGTACCTTATC

Table 2.9 | Primers for the possible control genes, based on RNAseq data. Distributor: Thermo Scientific.

Gene ID	Genes	Forward primers (5´ → 3´)	Reverse primers (5´→ 3´)
106565140	*Rhoa*	GATGGCGCATGTGGTAAGACG	GTGCCAACTCAACCTGTTTGC
106603799	*SmD2*	TTTTCGCGCAACGTTTTCAGT	TGGGTTTATTCAGCAAACTCATGT
106605287	*Roaa*	GGGCTACGACTACAGACTCAAG	TGCTTCCCCAGTTCCTCAGATG
100195744	*Carm1*	TCTACAACATGAACCAGGGAGGG	TGTATGGTTCACTACCCCTGTG
106583889	*Sae1*	CAGCGATACTCTCCCCGAATCTA	GACCCCCGCAGCCTCTTTT

2.3.3 Relative quantification using qPCR

Two technical replicates were used for two blastulation and gastrulation samples. qPCR was performed using SsoAdvanced™ Universal SYBR® Green Supermix (Cat.no.: 1725271; Bio-Rad), with reagents summarised in Table 2.10. The mix contains SYBR Green I – a dye with high affinity to the minor groove of double stranded DNA (dsDNA) (Dragan et al., 2012). The SYBR Green I-dsDNA complex emits green light, which is collectively enhanced for each binding of dsDNA. The mix was then run according to settings presented in Table 2.11. The results were visualised using CFX Manager Software (Cat.no.: 1845000; Bio-Rad).

Table 2.10 | qPCR master mix reagents.

Reagent	Final concentration
SsoAdvanced™ Universal SYBR® Green Supermix	1 x
Primers (forward/reverse)	1 μM
cDNA	800 ng
dH_2O (nuclease free)	-

Table 2.11 | Settings for qPCR thermocycler.

Step	No. of cycles	Temperature (C°)	Time
Initial denaturation	1	95	30 sec
Thermocycling	40	95	10 sec
		60	25 sec
Melt curve analysis	1	65-95	0.5 increments every 5 sec

The qPCR output is given in threshold cycle (C_t) values. This value corresponds to the moment in the PCR cycle when the fluorescent signal of the reporter dye surpasses a threshold that has been set arbitrarily (Schmittgen & Livak, 2008). These C_t values are used to calculate the expression ratio between blastulation and gastrulation, using a method known as 'The Pfaffl method' (Pfaffl, 2001). This method is corrected for the efficiency and allows for relative quantification. It requires a control to look at the relationship between two samples. For blastulation, gastrulation was used as control, and vice versa. The expression ratio was calculated using the following formula:

$$\text{Expression ratio} = \frac{(E_{target})^{\Delta Ct_{target}\,(control - sample)}}{(E_{ref})^{\Delta Ct_{ref}\,(control - sample)}}$$

- E_{target} = PCR efficiency of the target gene
- E_{ref} = PCR efficiency of the reference gene
- ΔCt_{target} (control – sample) = The difference in mean C_t values of the target in the control in the test sample.
- ΔCt_{ref} (control – sample) = The difference in mean C_t values of the reference in the control in the test sample.

Figures were made in RStudio.

2.4 Protein detection using Western Blot

Western blot is a technique used to separate a mixture of proteins by molecular weight and verify presence of proteins of interest using specific primary and secondary antibodies (Mahmood & Yang, 2012). The secondary antibody binds to the primary antibody and is conjugated with horseradish peroxidase (HRP). This enzyme is used as a reporter, which in the presence of peroxide catalyses the oxidation of luminol. This reaction produces a chemiluminescent substrate. This signal can be detected with a digital imager, making the HRP conjugates suitable for multiple chemical applications (Bio-Rad, 2022). The workflow is summarised in Figure 2.2.

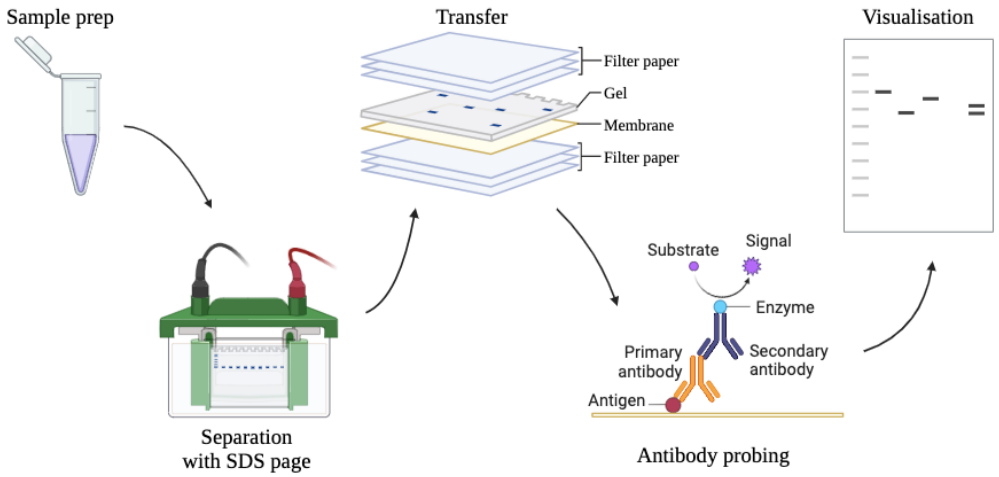


Figure 2.2 | The workflow of Western blotting. The first step involves preparation of the sample. Next, the proteins in the samples are separated by size through SDS-PAGE. After separation, the proteins are transferred to a membrane, before being probed with a primary antibody specific to the target. A secondary antibody then binds to the primary antibody. The signal conjugated to the secondary body is next activated – a signal used for imaging.

All reagents used for Western blot were from Bio-Rad, if not specified otherwise. Western blot was performed using 4-15 % Mini-PROTEAN Precast Gels (Cat.no.: 4568086). These gels consist of polyacrylamide and are designed for electrophoresis. For sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE), the buffers were prepared as stated in the protocol. Samples were collected from late blastulation, early-to-mid gastrulation and somitogenesis stage, and stored at -80 °C. After sample preparation, the protein concentration was measured using Qubit Protein Broad Range Assay Kit (Cat.no.: A50668; Invitrogen). The samples were then diluted to appropriate concentrations before 5-6 µg of total proteins from each sample were loaded on the gel. After 20-30 minutes at constant 200 V the samples had reached the bottom of the cassette, completing the electrophoresis. After the gel was removed from the cassette, transfer was done using a Trans-Blot Turbo system with pre-soaked 0.2 polyvinylidene fluoride (PVDF) membrane transfer packs (Cat.no.: 1704156).

To prepare for immunodetection, the membrane was blocked in Tris Buffered Saline (TBS) with 1% Casein (Cat.no.: 1610782EDU) for 1 hour. After washing with TBS with Tween-20 (TBST) buffer, the membrane was left overnight at 4 °C with diluted primary antibody (Table 2.12). After examining the RNAseq data (Supplementary figure 7.1) and evaluating the available antibodies, Oct4 was selected as a marker for pluripotency and Sox17 was chosen as a marker for differentiation. The membrane was then washed and treated with a secondary antibody (Goat Anti-Rabbit IgG Antibody, HRP-conjugate, cat.no: 12-348; Sigma-Aldrich). After a new series of washing, Clarity Western ECL Substrate kit (Cat.no.: 1705061) was used for detection, followed by imaging using ChemiDoc XRS+.

Table 2.12 | List of primary antibodies available. Of the antibodies available, Oct4 and Sox17 were used for Western blotting, here marked with an underline.

Primary antibody	Distributor	Product number
Oct4	Sigma	AB3209
Sox2	Abcam	AB97959
Klf4	MyBiosource.com	MBS3204079
Sox17	Antibodies-online.com	ABIN2364850
RERE (ATN2)	Antibodies-online.com	ABIN2793839
Sox10	Abcam	AB229331
Pax7	MyBiosource.com	MBS3200912

3 RESULTS

3.1 Growth, coating, and differentiation

The cells grew in 4 out of 5 types of media. Cells suspended in DMEM medium (Table 2.3) died after just a couple days in culture and did not show signs of active cell division. The cells seemed to swell and break, leaving the medium full of cell debris (Figure 3.1). The medium turned pink after transfer of cells – a colour change that indicates an increase in pH, as the medium contains DMEM with the pH indicator phenol red (Welshons et al., 1988). The pH of the medium was 7.4 when initially culturing the cells but increased to 8.5 after three days in culture. After trying to adjust the pH, the same would occur within the next day. This medium was not used for further culturing due to low growth and high death rate.

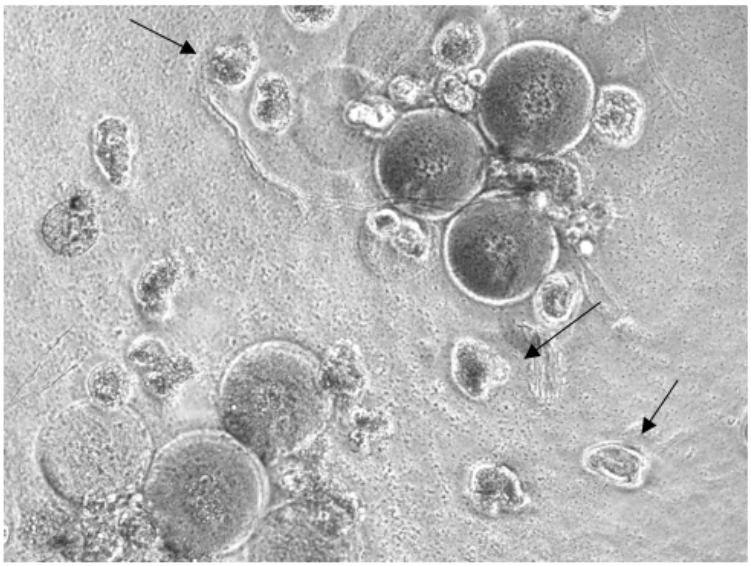


Figure 3.1 | Cells in DMEM medium day one after culturing. Black arrows show cell debris after cell death. The cells swelled right after transfer, resulting in big and circular cells.

In the other media (L-15 medium, 2i medium, EE medium and bFGF medium (Table 2.2, 2.4-2.6)) typical signs of cell division were observed. These were recognised as elongated cells of non-circular structures (Figure 3.2). The cells, however, never grew to confluence. Some of the problems was thought to be their sensitivity in terms of mechanical stress, and different approaches to media exchange were tested. When changing media, the cells were either lifted from the well, or media was removed directly from the well leaving the cells untouched. Even though the cells were resuspended when media was exchanged, the cells formed smaller clusters. When breaking the clusters apart, through pipetting, the cells died off more quickly

when compared to the cultures where the media was exchanged without removing the cells from the well.

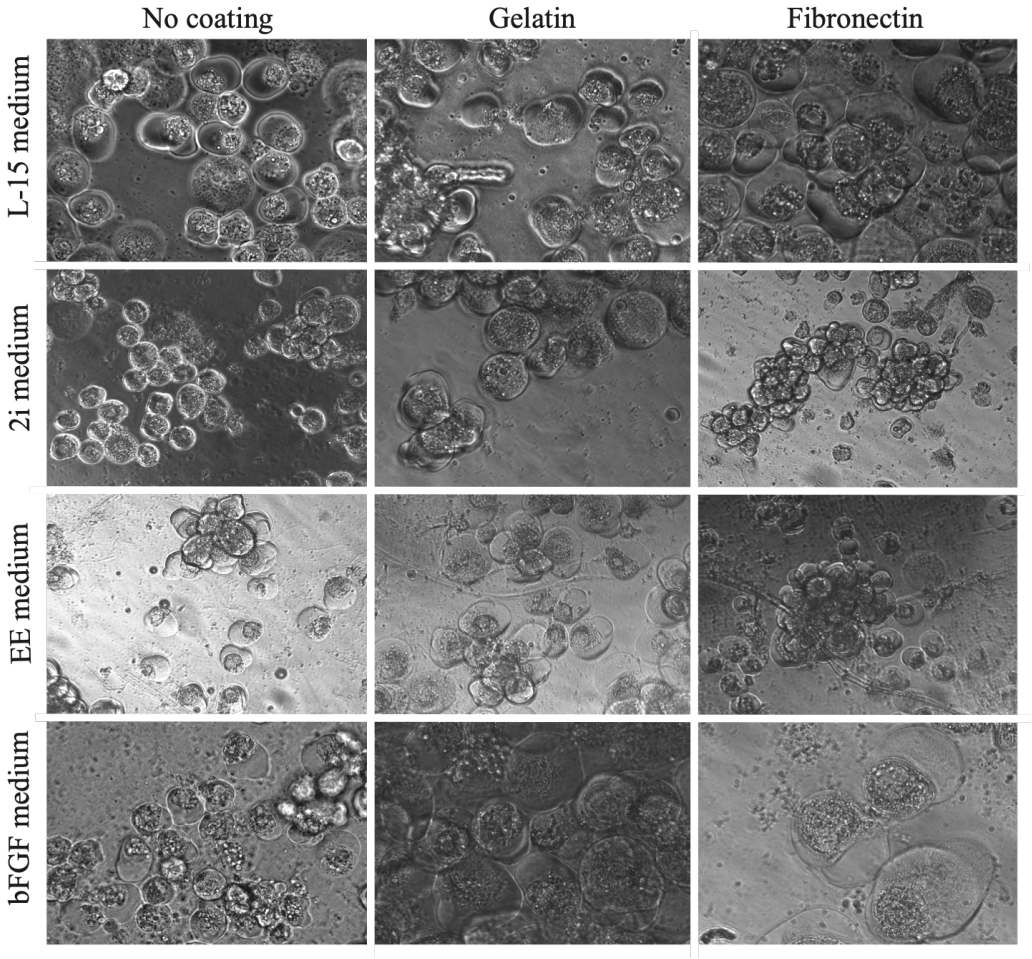


Figure 3.2 | Embryonic stem cells 3 days after transfer to wells with different coating. All cells showed signs of growth in terms of non-circular shapes and elongated structures emerging from the cell membranes. The first column from the left show cells cultured in wells with no coating, the columns in the middle are gelatin-coated, and wells in the right column are coated with fibronectin. Row 1-4 contain L-15, 2i-, EE-, and bFGF medium, respectively.

Despite efforts to reduce the mechanical stress, the cells were dying, and waste products accumulated. Accumulation of cell debris increased the frequency of contamination, causing the cultures only to last for a couple of weeks at most. Some observations were nevertheless done on the fledgling cultures. One observation that was common for all wells, was that masses of cells developed into what resembled embryoid bodies (EBs) (Figure 3.3). These are aggregates of cells and formed in all four media, both with and without coating. They were, however, more frequently observed in EE-, and bFGF medium.

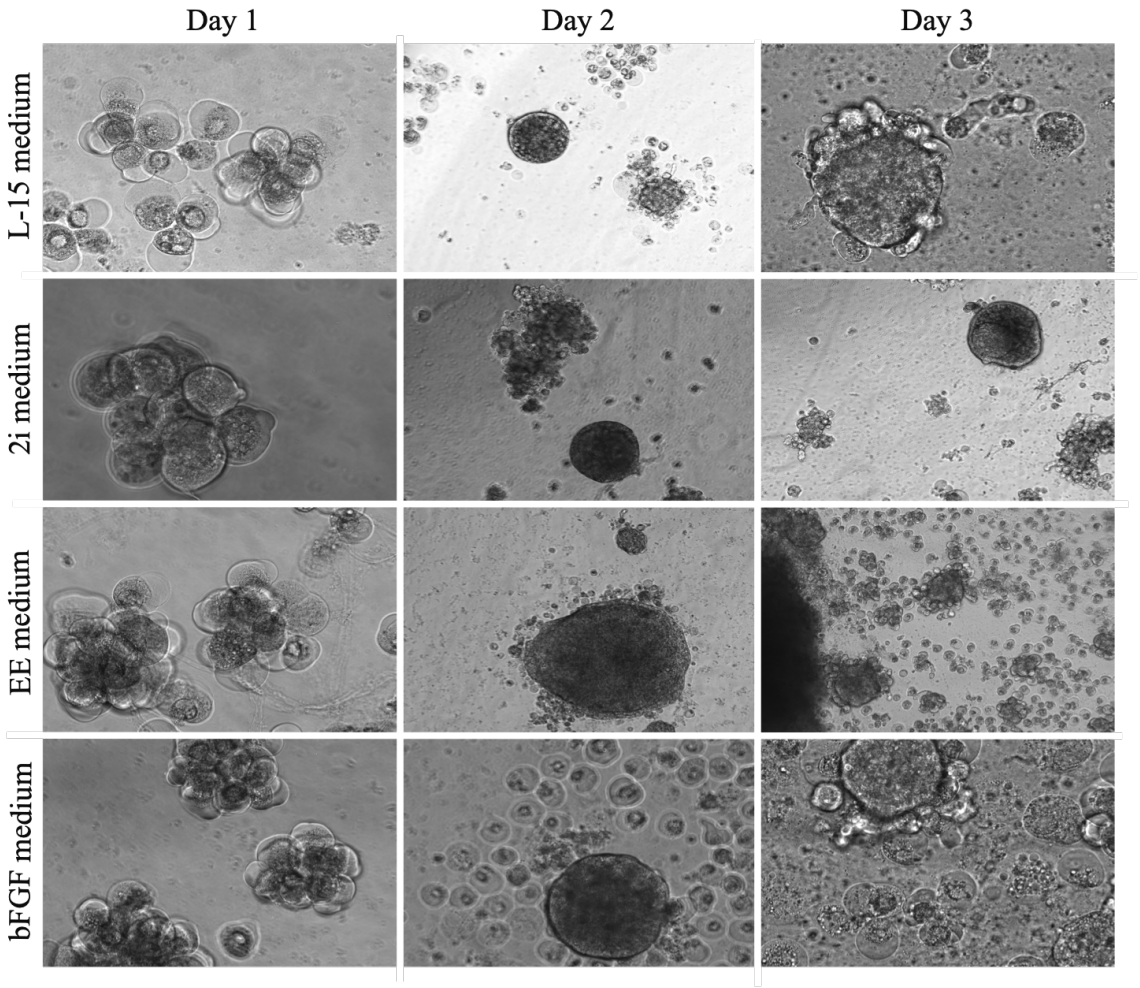


Figure 3.3 | Formation of embryoid bodies (EBs) in cell cultures in all media. The pictures are taken in wells with no coating, but the same behaviour was observed in both gelatin- and fibronectin-coated wells. After one day in culture, the cells aggregated in small clusters. The next day, the clusters formed a lining encircling the cells – a characteristic frequently observed in EBs. EBs were also observed after three days in culture. Row 1-4 contain L-15, 2i-, EE-, and bFGF medium, respectively, and columns show 1-3 days in culture, from left to right.

3.1.1 Differentiation in 2i medium

When looking at the differentiation in the types of media, results in 2i medium differ from the three other media. Cells cultured in all media formulations showed neuron-like differentiation but thrived on different types of coating. It's important to note that the observation of neuron-like differentiation is based purely on morphology and that these cells were not tested for neuronal markers.

In 2i medium neuron-like morphology was spotted for a number of cells only three days after culturing in fibronectin-coated wells, as well as in a single well with no-coating (Figure 3.4). This morphology was recognised as elongated cells with pointy arms, attached to the surface. These differentiated cells were most abundant where the cells grouped together in bigger clusters but were also observed in cells isolated from the rest. Even though signs of differentiation were observed in non-coated wells, differentiation was especially prevalent in fibronectin-coated wells. In the latter, a great network of attached neuron-like cells was found underneath the bigger cell masses (Figure 3.4B). Here, the cells were attached to the surface. In gelatin-coated wells, differentiation was not as striking as with fibronectin coating, but some odd structures were noticed that could resemble an early cell maturation (Figure 3.4D).

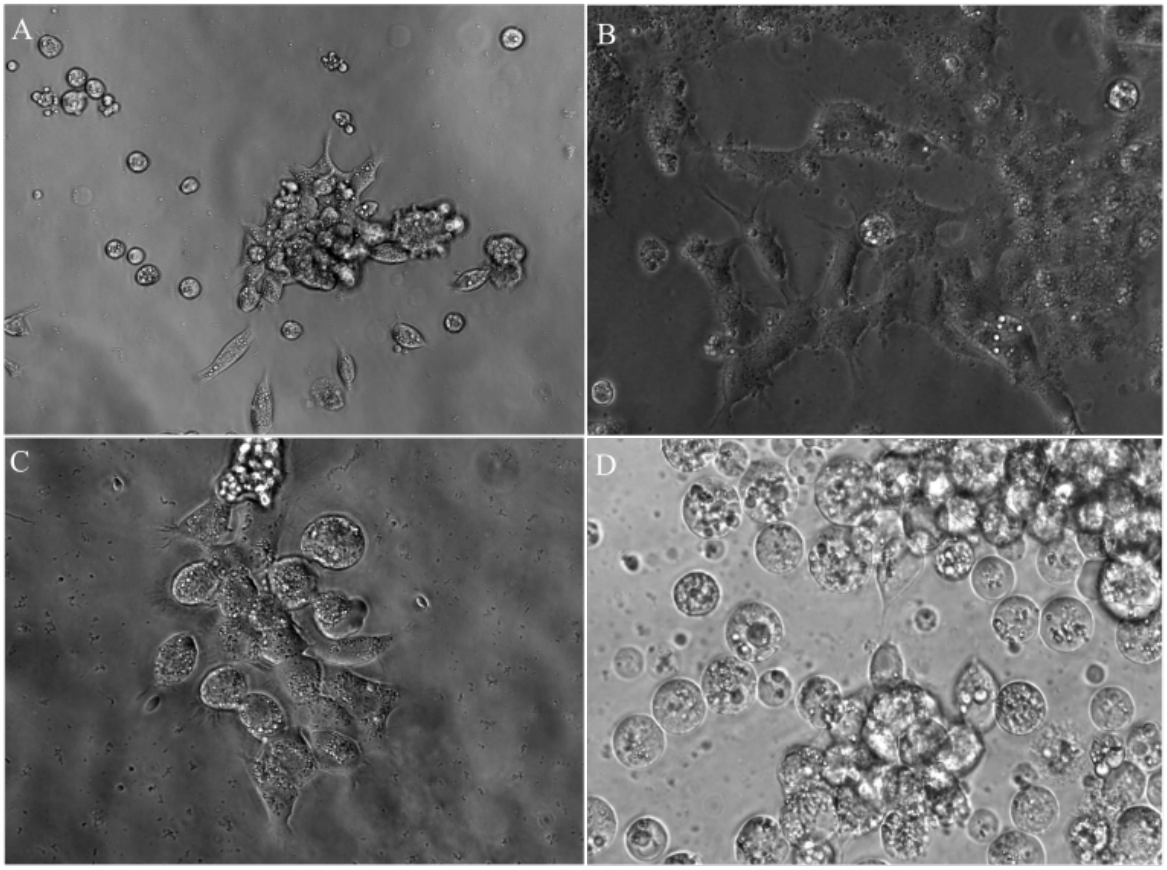


Figure 3.4 | Spontaneous differentiation in 2i medium. Cells in 2i medium started differentiating into neuron-like cells. This was observed 3 days post culturing in fibronectin-coated wells (**A**). After 11 days in culture, more cells had differentiated under the same conditions as in **A** (**B**). This differentiation was also observed after 3 days in wells with no coating (**C**). Figure **D** shows what may be early differentiation after 6 days in culture in a gelatin-coated well.

3.1.2 Differentiation in L-15-, EE-, and bFGF medium

In L-15 based media (L-15-, EE-, and bFGF medium) differentiation was only observed in gelatin-coated wells (Figure 3.5). Cells in these media all showed similar pattern of differentiation, appearing after 8 days in culture. These cells were observed underneath the greater masses of cells and may, however, have been present days earlier. Contouring of attachment aroused suspicion of more cells underneath at this point of time.

The wells with differentiated cells were more abundant where the cells formed bigger aggregates. This was also where attachment was more frequent. Differentiated cells were also observed among cells isolated from the rest. No differentiation was observed in fibronectin- or non-coated wells using the same media.

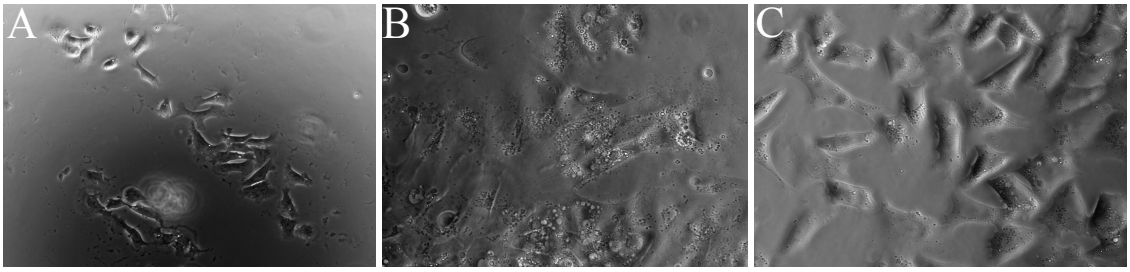


Figure 3.5 | Spontaneous differentiation 8 days after culturing of cells in L-15-, EE-, and bFGF medium on gelatin-coated wells. Differentiated cells are shown in L-15-, EE-, and bFGF medium in figure **A**, **B**, and **C**, respectively. This was especially visible when transferring the bigger cell masses to other wells, exposing the cells attached to the coated surface.

3.1.3 Differentiation in 7.5 % and 15 % FBS media

In embryonic cell line of Japanese flounder (*Paralichthys olivaceus*), proliferation was documented lower in medium containing 7.5 % FBS when compared to medium with 15 % FBS (Chen et al., 2004). To see whether the concentration of FBS could induce the growth rate for the salmon embryonic culture, the concentration of FBS in the L-15-based media was adjusted from 2% to 7.5 and 15%. This was not done for the 2i medium, as it did not initially contain FBS.

After two days in culture in gelatin-coated wells, cultures of all types of media consisted of differentiated cells – all with neuron-like morphology (Figure 3.6). This occurred in both 7.5% and 15% FBS media, and no difference between the two concentrations were observed. In the non- and fibronectin-coated wells no differentiation was observed at this point of time. After three days media was exchanged, and the cells stopped dividing. More and more cells died off, which was the case for cultures in both 7.5 % and 15 % FBS media. No additional differentiation was observed after this point.

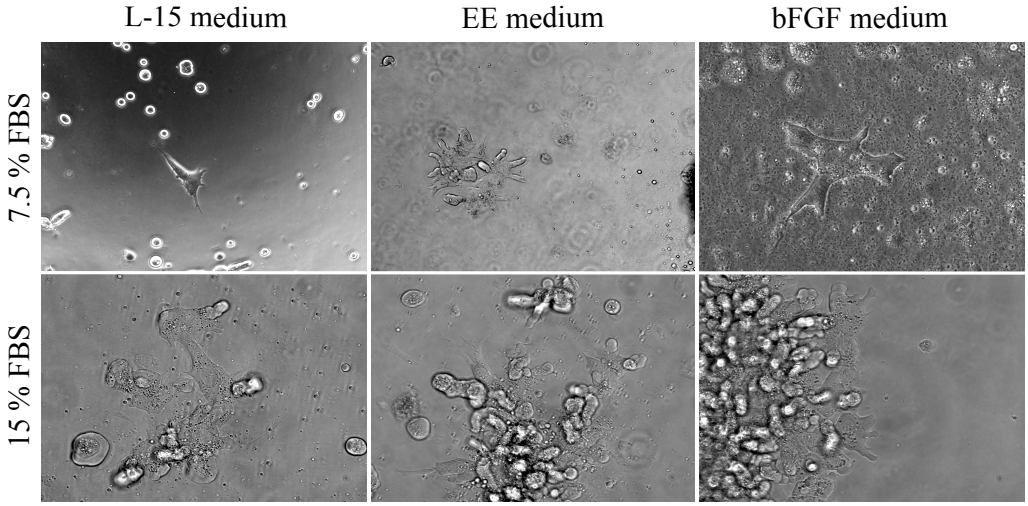


Figure 3.6 | Spontaneous differentiation in L-15-, EE-, and bFGF medium after two days in culture on gelatin-coating. When increasing the concentration of FBS from 2% to 7.5 % and 15 %, the first cells started differentiating already after two days of culturing. This occurred in gelatin-coated wells of all media. Differentiation was not, however, observed in either non- or fibronectin-coated wells at this point of time. Rows show cells in 7.5 % and 15 % FBS, and columns show L-15-, EE-, and bFGF medium, respectively.

3.1.4 Effects of coating

Differences in growth efficiency on different coatings were mainly observed between media containing L-15 (L-15 medium, EE medium, and bFGF medium) and 2i medium. The cells growing in the 2i medium attached more easily to the fibronectin-coated, compared to the gelatin- and non-coated wells. In the L-15-based media, however, the cells attached better in the gelatin-coated wells. This was observed when in all 2, 7.5 and 15 % FBS formulations. When attached, the cells were observed as faded into the background, with a more diffuse cell membrane, and not affected when simply pushing the plate. Common for all of the media was that differentiation primarily was observed where the cells attached to the surface.

3.2 Pluripotency genes expression

3.2.1 Reference genes for qPCR

When using qPCR for relative quantification of gene expression it is important to use endogenous control genes with a stable expression across all samples tested. Both frequently used marker genes and not-so-commonly used genes were assessed as possible controls. In analysis of salmon gene expression *Eef1a1a* and *Rps20* are frequently used (Olsvik et al., 2005). To evaluate these reference genes, Eef1a1a and Rps20 we used RNAseq data to check their expression throughout early salmon development (Figure 3.7)

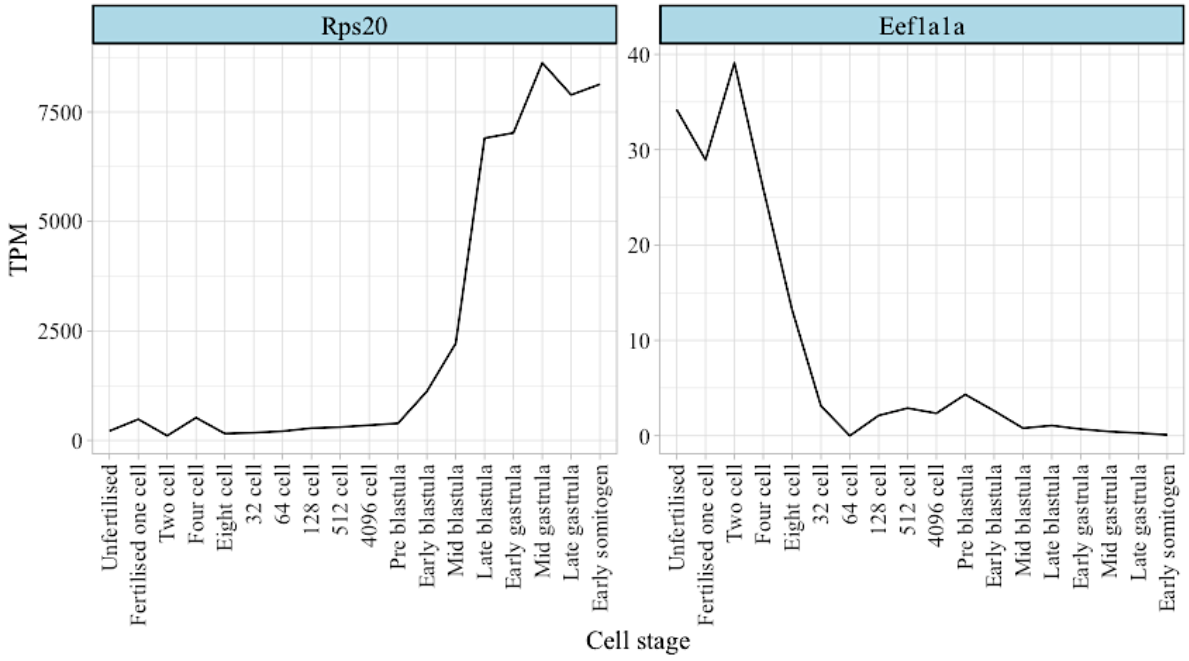


Figure 3.7 | Expression data of *Eef1a1a* and *Rps20* during early development of salmon. The graphs show expression data measured in transcripts per million (TPM). Data is displayed as the mean values of RNAseq data from samples at each cell stage. For *Rps20*, the rate of expression is low before early blastulation, before rapidly increasing, and keeping a high expression rate throughout early somitogen stage. The *Eef1a1a* graph show high transcription rate before eight-cell stage, and an evenly low expression value after. It's worth mentioning that the TPM values of *Eef1a1a* is much lower overall, compared to *Rps20*.

When looking at the pattern of expression in *Eef1a1a*, one can see a relatively high expression rate at early cell stages, that rapidly decreases after eight-cell stage. After this stage, the expression is kept low throughout early somitogen stage. For *Rps20*, the opposite trend is seen. Here, the expression is low at the early stages, and drastically increases at early-to-mid blastulation. *Rps20* is also expressed at much higher rate in general, compared to *Eef1a1a*, as the TPM of *Eef1a1a* ranges from 0-40, and it for *Rps20* ranges from 0-8000 TPM.
As a control gene should hold an even expression rate across different tissues, or in this case, throughout different stages of early development, neither of the genes were further considered. Even though *Eef1a1a* are evenly expressed at the point of time where samples for this experiment was collected, the expression rate is too low, and the gene were thus discarded for this purpose.

To seek out the best suited control genes in early development of salmon, the cov was calculated in the RNAseq dataset. Table 3.1 displays the 5 most prominent genes ranked according to their lowest cov value. The table also provides information on the average TPM value, which indicates the level of gene expression in the cells.

Table 3.1 | List of potential control genes arranged by ascending coefficience of variance. Of the list generated, five genes were selected for testing using qPCR. *Heterogeneous nuclear ribonucleoprotein A/B* (*Roaa*) and *small nuclear ribonucleoprotein Sm D2* (*SmD2*) share a high TPM value, whereas *Transforming protein RhoA* (*Rhoa*), *coactivator-associated arginine methyltransferase 1* (*Carm1*), and *SUMO1 activating enzyme subunit 1* (*Sae1*) all have a TPM value below 90. In terms of cov, *Sae1* sticks out as the candidate with the most variance between the samples.

Gene	TPM (mean)	Cov
Roaa	503.198	0.160
SmD2	471.917	0.161
RhoA	86,144	0.169
Carm1	74.566	0.171
Sae1	82.568	0.728

The RNAseq data was also used to make plots, visualising the five genes' expression rate throughout salmon early development. This is presented in Figure 3.8 below and is based on the mean TPM in the samples collected at each cell stage. The expression patterns of *Heterogeneous nuclear ribonucleoprotein A/B* (*Roaa*) and *Small nuclear ribonucleoprotein Sm D2* (*SmD2*) genes are remarkably similar. Both genes display analogous trends in the levels of their transcripts and respond in a similar manner to up- and downregulation events during development.

When looking at *Transforming protein RhoA* (*Rhoa*), *coactivator-associated arginine methyltransferase 1* (*Carm1*), and *SUMO1 activating enzyme subunit 1*(*Sae1*) they are all expressed at the same TPM range and differ from the two former genes in expression pattern. Common to all is their even expression from unfertilised- to early somitogenesis stage, and none of the genes were excluded from further analysis solely based on expression pattern.

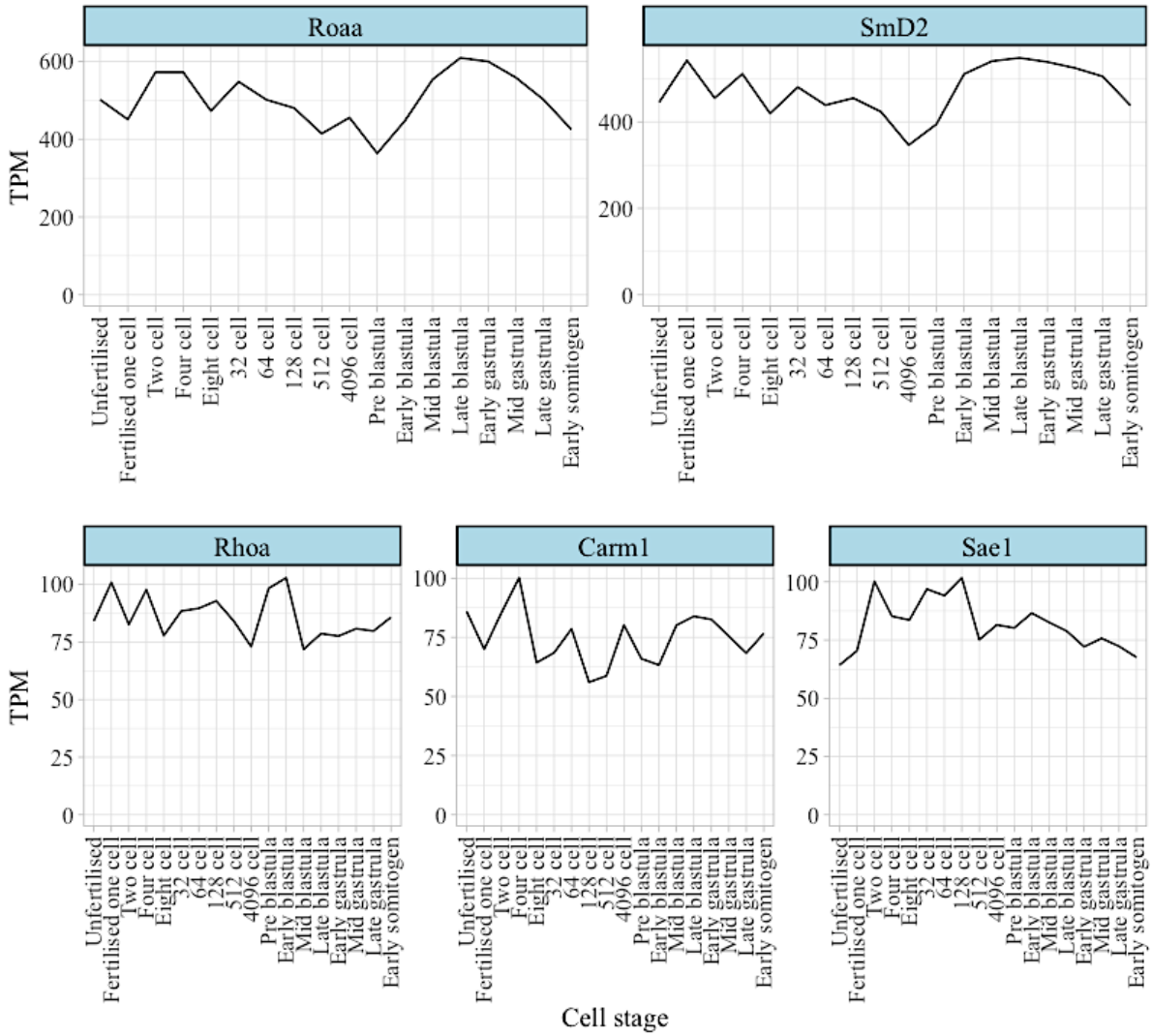


Figure 3.8 | Expression data of possible reference genes during early development of salmon. The line charts visualise the gene expression data in TPM of the genes with lowest coefficience of variance (cov), and a TPM value above 5, generated form the RNAseq data. *Roaa* and *SmD2* show a similar expression pattern with mean TPM values of about 500. *Rhoa*, *Carm1*, and *Sae1* differ in this pattern, but share show a similar expression pattern toward each other, with a much lower mean TPM below 90. Blastula = Blastulation; Gastrula = Gastrulation; Somitogen = Somitogenesis.

After running qPCR, testing the primers of the five potential control genes, only *Roaa*, *SmD2*, and *Rhoa* had primers that passed the wanted efficiency of 90-110 % (Supplementary table 7.1). Because the primer efficiency of *Carm1* and *Sae1* did not fall into this range, they were discarded. To further select a gene, the C_t values generated from the qPCR were evaluated. These values indicate the relative expression of the gene in the sample and should be relatively even for both blastulation and gastrulation samples. The box plot below visualises the expression stability of the genes using C_t values from one run of qPCR (Figure 3.9). The

figure displays both the variation in all the samples in A, and the variation within the cell stage in B. Compared to the other two genes, *Roaa* exhibits a more stable expression, which is evident from the boxplot in Figure 3.9A. When looking at the expression between the two different samples (Figure 3.9B), *Rhoa* stand out at the lesser stable gene of the three genes assessed. Even though the expression is even within the blastulation samples, it clearly separates from the gastrulation samples. As for *SmD2*, both samples look to have variance in gene expression. *Roaa* shares the same pattern of expression as *SmD2* in the blastulation samples but are much more evenly expressed in the gastrulation samples compared to both *SmD2* and *Rhoa*, except for one outlier. The median however is about the same in both sample types.

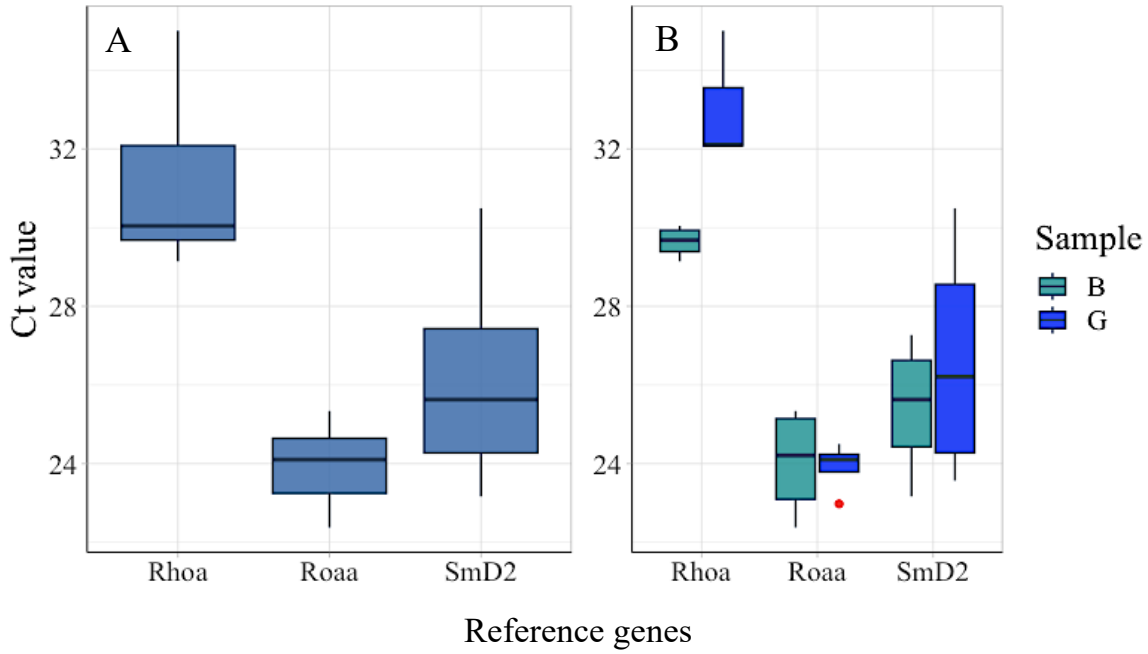


Figure 3.9 | Boxplots representing the expression stability of potential control genes in embryonic development in salmon using C_t values obtained from qPCR. Of the genes with a primer of an adequate efficiency, *Roaa*, *SmD2* and *Rhoa* were considered as endogenous controls. This boxplot presents their C_t values based on two qPCR runs, each with duplicates. On the left (**A**), all samples are considered, whereas the right figure (**B**) displays the variation within either blastulation or gastrulation samples. Boxes represent the lower and higher quantile; Black horisontal line within the box is the median; Whiskers indicate both highest and lowest extreme; Red dot outside the *Roaa* gastrulation box indicate an outlier; B = Blastulation stage; G = Gastrulation stage.

As *Roaa* shows highest expression stability, it was chosen as endogenous control when looking for difference in expression using markers for pluripotency and differentiation.

3.2.2 Marker genes for pluripotent phenotype

We expect that pluripotency marker genes are highly expressed at early stages, with a decrease in expression reaching gastrulation. Many marker genes are common among a range of different species, and frequently used to verify a pluripotent phenotype. Of the more common ones are *Pou5f1*, *Nanog*, *Sox2, Klf4* and *Tcf3a*. Based on the RNAseq data presented in Figure 3.10, the gene annotated in salmon genome as *Nanog*, fulfils these criteria of high expression at early cell stage before decreasing post blastulation. *Nanog* expression peaks at mid-blastulation, and significantly decreases at late blastulation-early gastrulation. Despite *Pou5f1* expression being relatively high when entering gastrulation stage, its expression is rapidly decreasing, suggesting its importance in maintaining pluripotency. It could still fit the criteria as a pluripotency marker gene, as the difference in expression between pluripotent and differentiated is significant. Both *Nanog* and *Pou5f1* are thus selected for further analysis.

Candidate genes *Klf4, Tcf3a*, and *Sox2* that were identified based on annotation, are not expressed as expected. Both *Sox2* and *Tcf3a* share a common expression pattern with an increase in expression towards differentiation. *Klf4* peaks at 4096-cell stage, rapidly declines, and then holds its expression rate even throughout early somitogenesis stage. Neither of these genes meet the criteria of a marker gene and are here rejected as pluripotency marker genes.

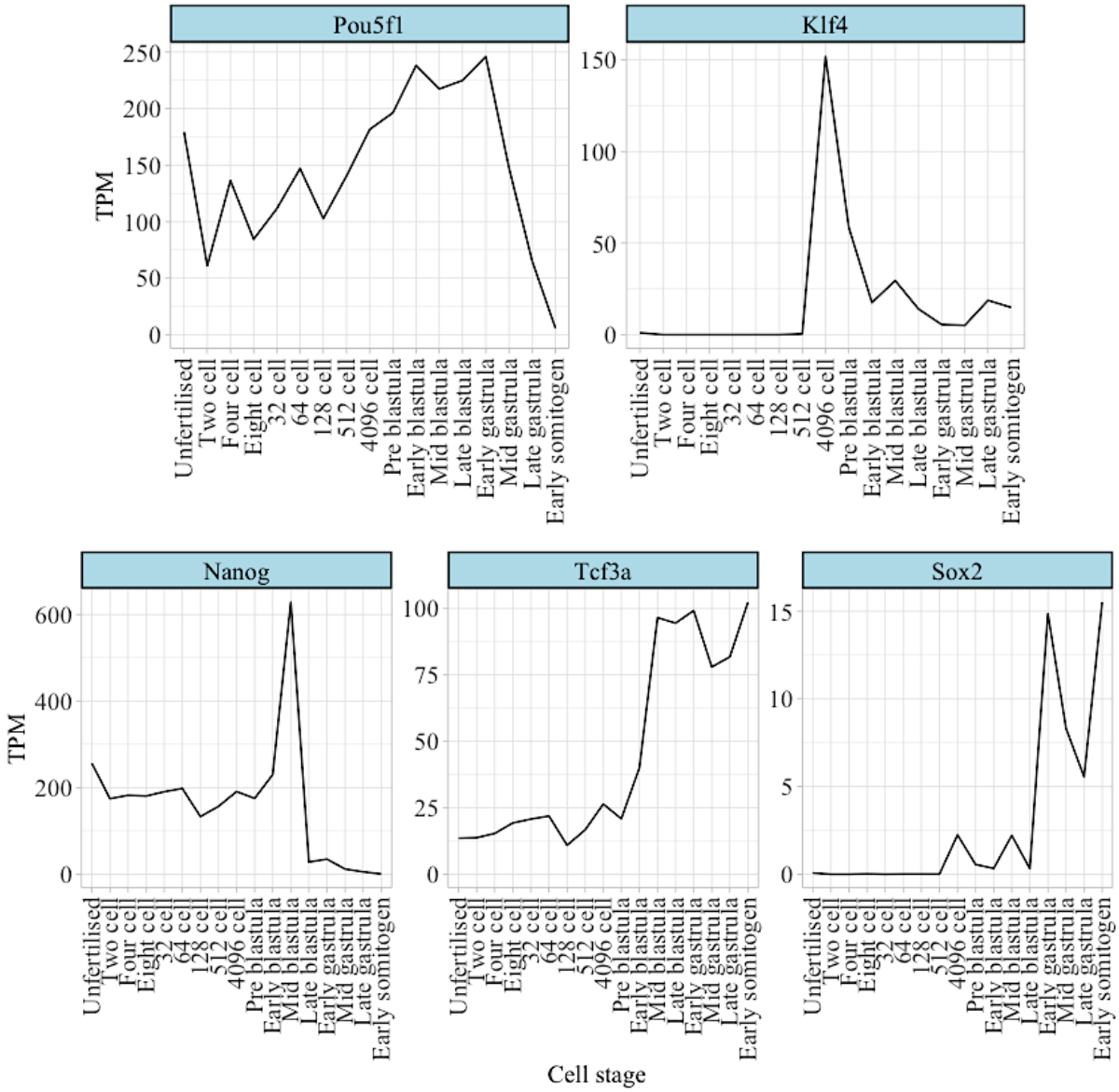


Figure 3.10 | Expression data of salmon in TPM of the most commonly used marker genes for pluripotent phenotype. In theory, *Klf4*, *Oct4*, *Nanog*, *Tcf3a*, and *Sox2* are thought to be highly expressed before gastrulation, where their expression decrease as the cell differentiates. However, according to this RNAseq data obtained from early developing salmon, only *Nanog* are behaving as expected with high transcript values before reaching gastrulation. *Pou5f1* are highly expressed at before mid-gastrulation, before sharply decreasing towards early somitogenesis stage. *Sox2* have an overall low TPM value, whereas *Tcf3a* increase in expression around mid-blastulation stage. *Klf4* have moderate TPM values until 4096-cell stage, where it peaks, before rapidly declining about pre-blastulation stage. Blastula = Blastulation; Gastrula = Gastrulation; Somitogen = Somitogenesis.

3.2.3 Differentiation factors

In search for markers of differentiation one would look for genes that increase in expression when the cell enters gastrulation stage. Figure 3.11 shows the typically used marker genes for differentiation. *Gata4* and *FoxA2* are commonly used as marker genes for endoderm, whereas *Ncam1a* and *Gata3* are used as markers for meso- and ectoderm, respectively. When looking at the figures, *Ncam1a* is not significantly expressed in any of the samples, only with a slight increase at early somitogen stage. As for the other three genes, their expression increases right before and when entering gastrulation, and indication that these genes may serve an important role in development in salmon embryos. Of the four genes, only *Gata3* was used for further analysis, as it had the higher expression rate.

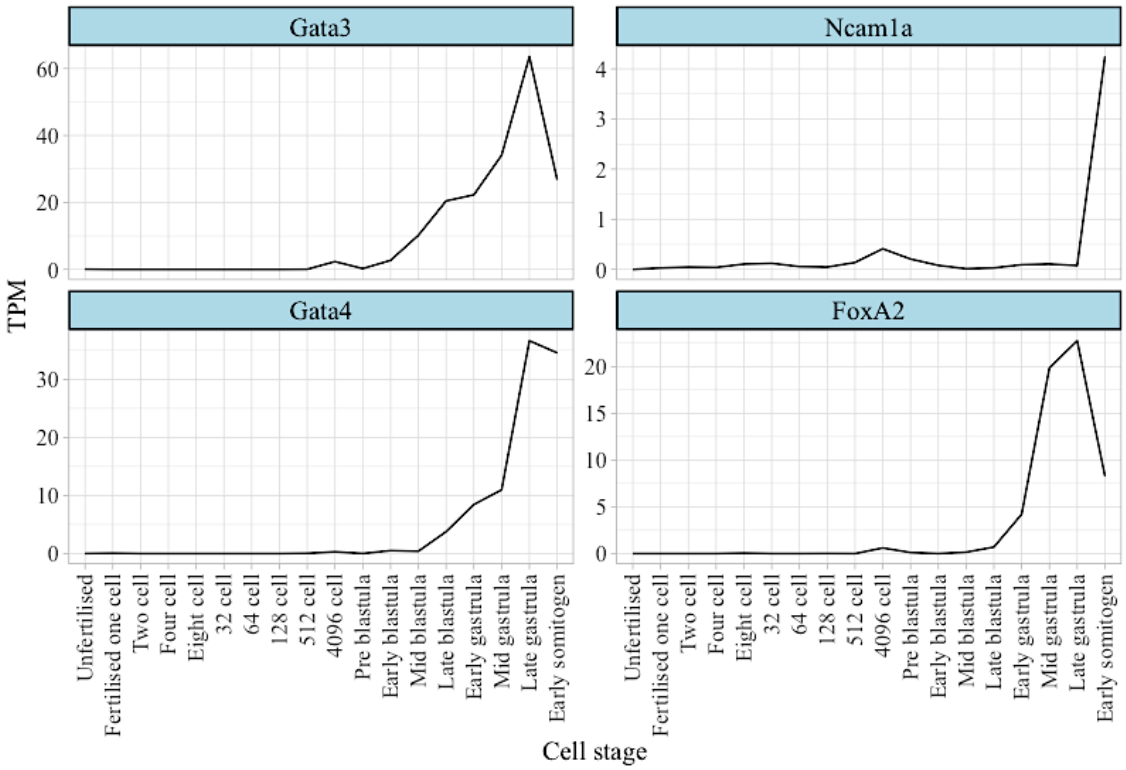


Figure 3.11 | Expression data in TPM of commonly used marker genes for differentiation in salmon embryonic development. *Gata4* and *FoxA2* are in many model organisms used as a marker gene for endoderm. *Ncam1a* and *Gata3* are used as markers for meso- and ectoderm, respectively. *Ncam1a* are in these samples not much expressed at all. The other genes are nevertheless expressed as one could expect with a low expression at undifferentiated stage, and a drastic increase when entering gastrulation. Blastula = Blastulation; Gastrula = Gastrulation; Somitogen = Somitogenesis.

Figure 3.12 presents alternative marker genes for differentiation, selected based on the difference in blastulation and gastrulation samples in RNA sequencing data. Of the top 10 genes generated by the code (Supplemental code 7.2), only five generated primers specific to the target. *Apolipoprotein A1* (*Apoa1*) and *Antifreeze protein type IV* (*Afp4*) were especially highly expressed during differentiation. The other candidates, *Retinol binding protein 4* (*Rpb4*), *Heme binding protein 2* (*Hebp2*), and *Keratin, type II cytoskeletal 8* (*K2c8*) express lower TPM values. As their expression increases drastically when differentiation is initiated, they do, however, all probably serve an important role in development. All genes were selected for further analysis.

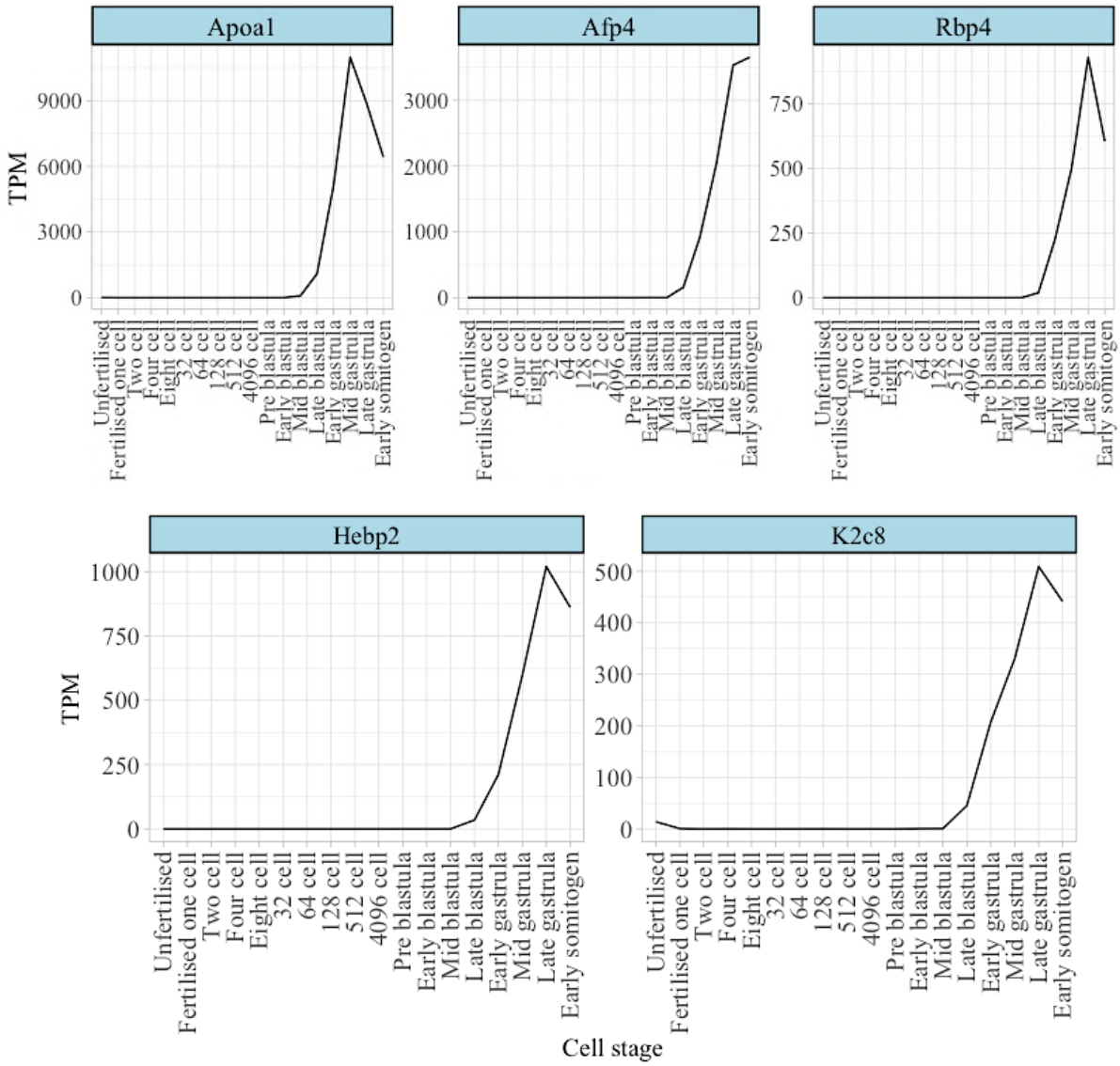


Figure 3.12 | RNAseq data of early developing salmon of the alternative differentiation marker genes, expressed in TPM. All genes share the same pattern of expression, with little-to-no expression before entering early gastrulation. Here, all of them increase in expression significantly. Blastula = Blastulation; Gastrula = Gastrulation; Somitogen = Somitogenesis.

3.3 Gene expression analysis using qPCR and Western blot

To select the candidate genes to use as a reference, as well as markers for pluripotency, and differentiation, in qPCR, both the RNAseq data, primer efficiency and coefficient of determination (R^2) were considered. Common for all the primers selected, was their efficiency between 90-110 %, and a R^2 value close to 1 (Table 3.2). Genes selected as markers for pluripotency had high expression values before entering differentiation, whereas markers selected for differentiation had an increase in expression post gastrulation. For pluripotency, *Nanog* and *Pou5f1* were selected, and for differentiation, *Apoa1* and *K2c8* were chosen.

Some genes were tested using qPCR before acquiring the RNAseq data and their selection were solely based on theory. Others were selected based on RNAseq data. The qPCR results of all genes tested, in terms of primer efficiency and R^2 is visualised in Supplemental table 7.1.

Table 3.2 | Genes selected as markers after testing using qPCR. The table include the genes used for further analysis, selected based on primer efficiency, coefficient of determination (R^2), and their gene expression pattern. *Apoa1* and *K2c8* were selected as marker genes for differentiation, *Nanog* and *Pou5f1* for pluripotency, with *Roaa* as an endogenous control.

Gene	Primer efficiency (%)	R^2
Apoa1	96.05	0.997
Nanog	98.43	0.994
Pou5f1	92.51	0.975
K2c8	90.68	0.997
Roaa (Control)	92.85	1.000

3.3.1 Testing of markers using qPCR

To look at the expression in blastulation and gastrulation for the four marker genes, the ratio between the samples was calculated based on the C_t values obtained from qPCR. Expression of the genes were normalised to *Roaa* as an endogenous control and show expression in blastulation compared to gastrulation (Figure 3.13).

A negative fold-change indicate a down-regulation of the gene in the blastulation sample, compared to gastrulation. This is seen for both *Apoa1* and *K2c8*. These show a negative fold change of about 11- and 4, respectively. Both pluripotency markers, however, increase in expression. *Pou5f1* expression increases by about 2.5-fold, which is a modest increase, compared to *Nanog*s' 25-fold change. This data is also summarised in Table 3.3 below.

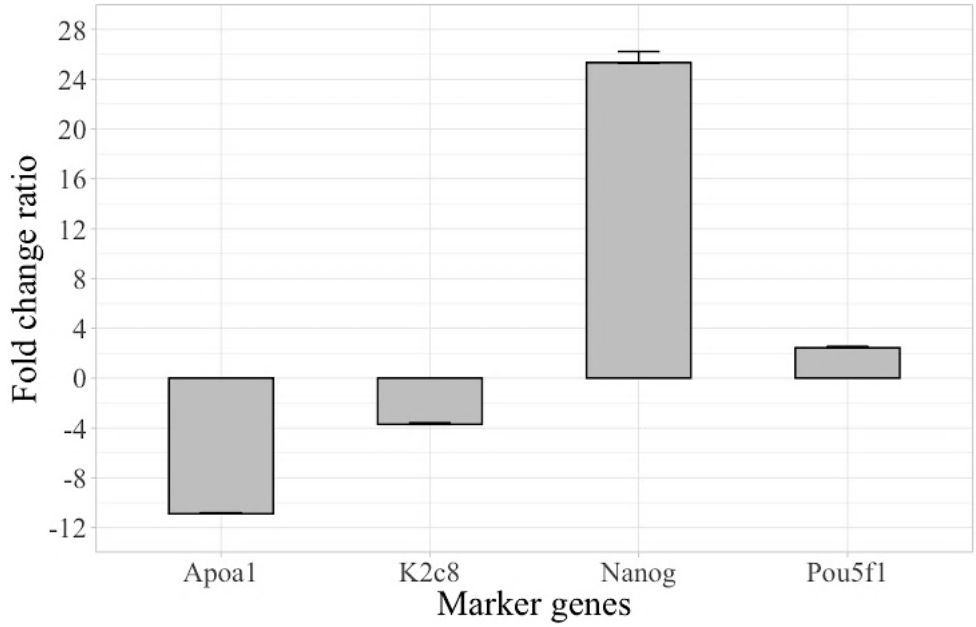


Figure 3.13 | Bar graph of the Pfaffl ratios of the four marker genes relative to the control (*Roaa*) in blastulation compared with gastrulation. Both *Apoa1* and *K2c8* are downregulated in the blastulation sample compared to the gastrulation sample, with about 10- and 4-fold respectively. *Nanog* mRNA expression increase by a 25-fold, whereas *Pou5f1* moderately increase by a 2.5-fold in blastulation compared to gastrulation. Standard deviations are indicated by error bars.

Table 3.3 | The results from qPCR analysis, using Pfaffl method to calculate the fold change in expression of the marker genes with *Roaa* as an internal control. SD = Standard deviation.

	Apoa1	***K2c8***	***Nanog***	***Pou5f1***
Fold change	-10,88	-3,702	25,32	2,445
SD	0,07077	0,1250	0,8953	0,1088

3.3.2 Western blot for protein verification

After evaluating the gene expression pattern of the antibodies available, *Oct4 (Pou5f1)* and *Sox17* were selected for Western blot (Supplemental figure 7.1). Figure 3.14 shows the results from Western blot using the Oct4 antibody. Two bands, at 25 and 20 kDa were detected. The protein is clearly visible in all samples, which are collected at blastulation (B), gastrulation (G) and somitogenesis (S). The bands, however, show a stronger signal in blastulation, compared to in gastrulation and somitogenesis. Salmon Pou5f1 have a molecular weight of 51,827 kDa (UniProt, 2017a). Some protein detection may be observed at about 50 kDa and 37 kDa. These bands are weak compared to the two prominent bands. The molecular size of these two does not match the weight of Pou5f1 protein, indicating that the antibody might be detecting a different protein. This can, however, not be concluded, as the target protein may have undergone alternative splicing.

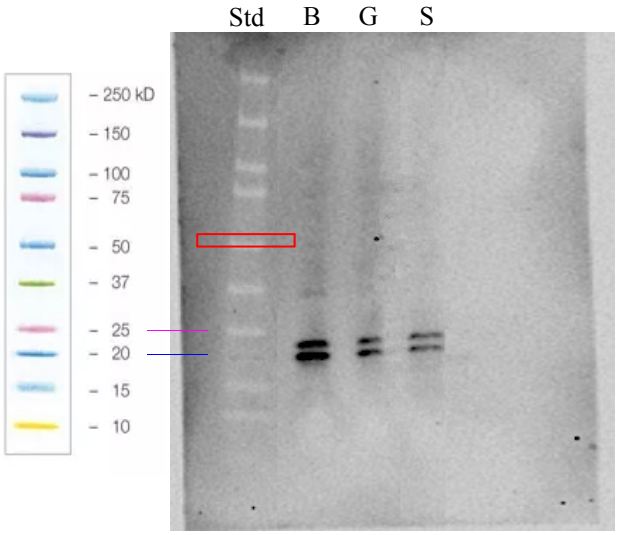


Figure 3.14 | Observation of Oct4 antibody using chemiluminescence detection. In all three wells, two lines are particularly apparent, at about 20 and 25 kDa. This protein is more apparent in the blastulation sample, with a little decrease in both gastrulation and somitogenesis sample. The red box indicates where the expected lines would have been if the antibody was specific to a non-modified salmon Pou5f1 protein. Primary antibody diluted 1:1000; secondary antibody diluted 1:2000. Std = Standard (Precision Plus Protein Kaleidoscope Prestained Protein Standards (Cat no.: 1610375; Bio-Rad)); B = Late blastulation sample; G = Early-to-mid gastrulation sample; S = Early somitogenesis sample.

The salmon Sox17a has a molecular weight of 43.81 kDa and comprise 390 amino acids (UniProt, 2017b). Figure 3.15 shows the results from Western blotting using Sox17 antibody. All the wells contain multiple smeared out bands. There is not much difference between blastulation and gastrulation samples, whereas the expression seems higher in the somitogenesis samples, as the bands appear darker. The area where we would expect a signal (red box), only faint bands are detected in the somitogenesis sample.

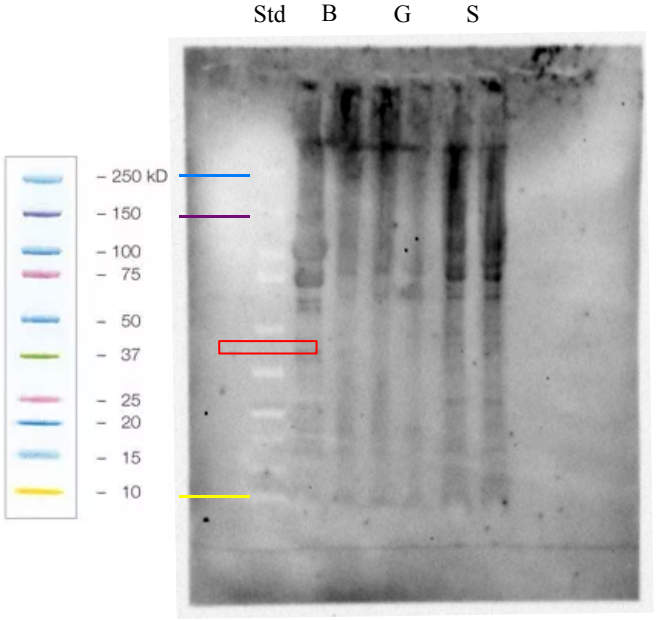


Figure 3.15 | Results from Western blot using Sox17 antibody. The second and third well contains blastulation samples, the fourth and fifth gastrulation sample, whereas the sixth and seventh well contains samples collected from somitogenesis stage. The red box on indicates where the expected lines would have been if the antibody was specific to salmon Sox17a. The lines on the left complement the standard where the lines are faded. Primary antibody diluted 1:3000; secondary antibody diluted 1:2000. Std = Standard (Precision Plus Protein Kaleidoscope Prestained Protein Standards (Cat no.: 1610375; Bio-Rad)); B = Late blastulation sample; G = Early-to-mid gastrulation sample; S = Early somitogenesis sample.

4 DISCUSSION

To establish an iPS cell line in salmon, it is critical to define the specific environment to maintain a pluripotent phenotype *in vitro*. Five different media compositions and two coatings were tested in this research using early developing embryos, known to inhabit pluripotent stem cell features. To validate its characteristics in culture, genes involved in maintenance of pluripotency and differentiation are key to identify. Genes conserved throughout multiple species and alternative genes were tested, based on literature on other stem cells cultures and analysis of RNAseq data, respectively. These markers were then analysed using qPCR and Western blotting.

4.1 Senescence and differentiation

Senescence is a stage in a cell's life cycle where it ceases to divide but continues to metabolise. The phenotype is associated with certain biomarkers, including stress-sensing proteins and cell-cycle repressors. These are thought to be triggered by either oncogenes, DNA damage or reactive oxygen species (ROS) (Courtois-Cox et al., 2008). Within the context of senescence, a distinction is often made between replicative and premature senescence. The former is triggered by an intrinsic cell division counter mechanism, such shortening of the telomeres, whereas the latter is induced by external factors. Oncogenes like RAS or suboptimal culture conditions are examples of factors that can induce premature senescence (Mathon & Lloyd, 2001). Inadequate culture conditions have been demonstrated to induce senescence in epithelial cells. This was demonstrated by a study executed by Ramirez et al., where comparison was made with cells seeded in chemical conditioned media and on feeder-cells. When cultured in chemical conditioned media on plastic dishes, the cells would divide 20-30 times until entering senescence. Premature cellular arrest was, however, avoided when cultured on feeder-cells (2001). Of the components of the chemical conditioned media, the serum is hypothesised to induce stress by producing ROS with help from amine oxidase, as seen in mural embryos. ROS are known to induce stress by damaging the DNA. When supplementing the media with antioxidants (2-Mercapoethanol or aminoguanidine) cells grew as normal (Parchment et al., 1990).

Results from the present research involved evaluation of FBS on proliferation on the ESCs. An increased concentration of FBS showed, however, a decrease in proliferation, where no sign of cell growth was observed after two days in culture. With increased amount of FBS follows higher levels of amine oxidase, which could explain why the cells stopped dividing faster in higher concentrations.

A limitation of proliferation is a known characteristic of differentiated cells. Differentiation and senescence are, however, two separate processes and can, as here, occur in a culture simultaneously (Norsgaard et al., 1996). Stem cells have shown to differentiate in densities both too high, and too low. The confluency driving differentiation varies between species and cell types (LeBlanc et al., 2022). The differentiation was, in present study, usually observed in relation to clusters of cells, but also isolated from the rest. Another observation worth mentioning was that it almost exclusively occurred where the cells showed better attachment to the plate. This could indicate that attachment and either high or low cell density is a prerequisite for differentiation.

Mechanical signals, which can be either extrinsic (e.g., tension or compression) or intrinsic (e.g., ECM or density), are recognised as regulators of cell differentiation. Through integrins, the cells can detect mechanical signals. Reactions to mechanical stimuli often affect the cytoskeleton to contract, altering the cytoskeleton tension. This change can activate downstream signalling resulting in transcriptional modifications, affecting cell behaviour such as cell fate (Petzold & Gentleman, 2021). For 2i medium, cells attached and differentiated on fibronectin, whereas the same trend was seen on gelatin-coat in cells seeded in L-15-, EE-, and bFGF medium. There could be multiple reasons as to why the cells attach better to one coating than the other when growing in different media. One of them could be the presence of growth factors, which could have influenced their behaviour and attachment. Growth factors are able stimulate expression of integrin receptors that specifically binds to the gelatin or fibronectin and would consequently make the cells more likely to attach to gelatin- or fibronectin coated surfaces, respectively. Among others, transforming growth factor β enhance expression of α_5, β_1, and α_3 in human monocytes (Wahl et al., 1993). The same is seen in neuronal stem cells, where bFGF and epidermal growth factor enhance expression of β1-integrins (Suzuki et al., 2010). β1-integrins are known to be one of two integrins in a dimer forming the receptor for collagen and gelatin (Zaman, 2007). Growth factors were

present in medium containing FBS (L-15-, EE-, and bFGF medium). Cells in this media attach to gelatin, which could indicate that FBS may induce integrins sensitive to gelatin. Growth factors was not supplied in 2i medium. These cells did not attach well in the gelatin-coated wells, indicating a lack of expression of gelatin-adhering integrins. Instead, the medium contains Leukaemia inhibitor factor (LIF), which could affect its ability to attach to fibronectin. Its effects connected to fibronectin and attachment is not, however, well known.

In order to check whether growth factors or LIF affect the adhesion, it may be useful to investigate the expression of integrin receptors or other signalling proteins in the cells. Further insights could also be provided through testing the effects of different growth factors or changes in the culture conditions.

The fact that the cells do not attach to the non-coated surfaces in both media suggests that there are not supplemented or produced enough adhesion-promoting cues to support cell attachment. Ultimately, factors such as the pH, osmolarity, or temperature could also influence the cell attachment in the different media on different coatings (Kim, 2021; Rico et al., 2010; Takagi et al., 2000).

4.1.1 Embryoid body formation and differentiation

EBs are small, circular, 3-dimentional structures comprised of multiple cells. These structures are thought to mimic some of the morphogenic interactions normally taking place within the embryo *in vivo*. EBs are formed at early development, in which differentiation often follows (Martin & Evans, 1975). In embryonic stem cell cultures EBs are routinely induced as a first step in differentiation protocols for further downstream applications (Pettinato et al., 2014). Protocols for EB formation usually include media depleted of growth factors and inducing agents, such as bFGF and LIF (Schuldiner et al., 2000).

EB structures were formed after two days in all media, indicating that the media compositions were not optimal for maintenance of pluripotent phenotype. EE medium was the only medium containing embryo extract, whereas bFGF medium was the only one containing bFGF. The 2i medium contains LIF – an important factor for maintenance of totipotent ES mouse cell lines (Williams et al., 1988). All these components are thought to be mitogenic (Fan et al., 2017; Furue et al., 2005). If EBs were formed as a result of mitogenic depletion, the L-15 medium

should contain more EBs compared to the other media, as no extra supplements are added in this medium. This was, however, not the case, as the EBs were more frequently observed in EE- and bFGF medium.

Previous studies on human ESCs have stated that embryoid body formation can occur as a result of limited serum exposure. However, in this particular research, a comparison was made between cells cultured in medium containing 15% FBS and cells cultured in medium containing 15% serum replacement solution. The latter comprise amino acids, vitamins, and proteins, such as insulin, and several tracer elements (Kubo et al., 2004). This may have an effect on the results. Alternatively, this could indicate that a 2 % FBS concentration is not enough to prevent EB formation but does not explain why some cells aggregate more in certain media.

In terms of FGF signalling, the pathway has shown important for formation of primitive endoderm (PE) in mammals (Figure 4.1) (Chen et al., 2000). Cells segregate into epiblast cells and PE in the ICM, both still with pluripotent properties. Derivates of the PE is generally referred to as extra-embryonic endoderm and constitute structures such as the yolk sac that help nutrient flow to the embryo (Hermitte & Chazaud, 2014). After the cells have aggregated *in vitro*, spontaneous formation of PE occurs, surrounding the inner epiblast-like cells, forming the embryonic body. This formation is thought to rely on FGF signalling through a pathway known to be induced by bFGF (Bratt-Leal et al., 2009; Ding et al., 2010). Despite this being a study performed in mammals, it could explain why formation of EBs are more abundant in bFGF medium compared to in L-15-, and 2i medium.

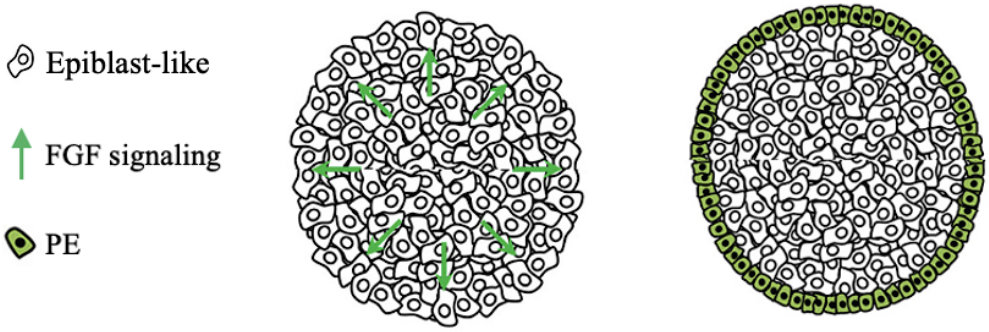


Figure 4.1 | FGF signalling during embryoid body formation. Green arrows simulate FGF signalling, whereas the green cells indicate primitive endoderm (PE) surrounding epiblast-like cells. The outer layer of PE formation is thought to be dependent on FGF factors during EB formation. The epiblast cells remain undifferentiated on the inside (Brickman & Serup, 2017) .

The content of the embryo extract is not fully known, and it is thus hard to say what factors could contribute to additional EB formation in the EE medium. There is evidence to suggest different effects of proliferation in terms of the stage of the embryos at the time of extraction (Canaider et al., 2014). However, the more studied Chick embryo extract comprise multiple identified growth factors, including bFGF (Joseph-Silverstein et al., 1989). Even though this factor is not identified in salmon-based embryo extract, it could be that EB formation is connected to bFGF concentration also for these cells, as hypothesised for the cells cultured in the bFGF medium.

4.1.2 Bursting cells in the DMEM media

Osmosis is the process of the movement of water molecules across a selectively permeable membrane from an area of high concentration of solutes to an area of lower concentration. These solutes, also called osmolytes, can be salts, sugars, or other molecules. In cell culture, if the composition of the media is not correct, osmosis can potentially cause cells to swell and burst. This can happen if the concentration of solutes in the media is higher than the concentration of osmolytes inside the cells (Lopez & Hall, 2023). Therefore, it is crucial to maintain proper osmotic balance in cell culture media to ensure cell viability and growth. It is probable that the cells in DMEM medium were subjected to osmotic stress. Of the components, glucose from DMEM, sodium pyruvate, and glutamine could affect the osmotic balance. If not optimised, the cells will struggle to maintain the osmotic balance, leading to cell death.

In the current study, it was also documented a change in pH in the same medium. The cell metabolism is complex, and it's hard to pinpoint what exactly could have caused the basic transition. Multiple waste products have been reported as toxic or cell growth inhibiting. Of these, the most important ones are lactate and ammonia. The latter is reported toxic in low concentrations but will depend on the cell line and the culture conditions. Ammonia accumulation leads to an increase in pH in the surrounding medium and can result from two possible mechanisms: amino acid metabolism or glutamine decomposition. The primary source of ammonia is through amino acid metabolism, particularly as a result of glutamine metabolism (Schneider et al., 1996).

The extent of glutamine decomposition is influenced by several factors, including the temperature and chemical composition of the media. Specifically, the rate of this reaction tends to increase under conditions of higher temperature and more basic pH, as well as when the culture is supplemented with phosphate, according to various reports (Ibid.).
It is not clear whether the change in pH was a result of amino acid metabolism or glutamine decomposition. The media is, nevertheless, complemented with reagents facilitating both, such as non-essential amino acid, sodium pyruvate, glutamine, and DMEM (high glucose).

4.2 Marker genes for Atlantic salmon early development

4.2.1 Reference genes

A reference gene should ideally be evenly expressed throughout both physiological differences and in-between types of tissues and developmental stages. However, as no universal gene is yet known, endogenous controls are chosen based on the experimental design (Kozera & Rapacz, 2013). Nor *Eef1a1a* or *Rps20* was found to serve its purpose as control genes for early salmon development. *Eef1a1a* is expressed at higher rate before 32-cell stage and decreases in expression between 32-cell stage and early somitogen stage. Not only is the gene unevenly expressed throughout early development but is likely too low in expression to be ideal as an endogenous control gene. *Rps20* have a total of higher expression values but is also unevenly expressed throughout development. Before pre-blastulation, *Rps20* is low in expression before rapidly increasing in TPM value after. Both genes may be used as control genes in adult tissues, but not for this experimental design.

Roaa, *SmD2*, *Rhoa*, *Sae1*, and *Carm1* were selected based on low variation in expression within the samples. *SmD2* is associated with the spliceosome – a large complex of RNAs and proteins that catalyse the removal of introns after transcription. These complexes, also called small nuclear ribonucleoproteins (snRNPs), are divided into subunits, named U1, U2, U4, U5, and U6. The core proteins of these units are called Sm proteins. SmD2 is a protein like this, and falls into the category of common proteins, as they are present in all known spliceosomal snRNP core complexes (Raker et al., 1996). *Roaa* is also a gene associated with RNA metabolism. Heterogenous nuclear ribonucleoproteins (hnRNPs) is a large family of proteins recognised for its involvement in DNA repair mechanisms, cell signalling, and regulation of gene expression, operating at both transcriptional and translational levels (Carpenter et al., 2006). Of the other potential control genes, both *Sae1* and *Carm1* are also associated with

post-translational modifications. *Sae1* encodes a protein called SUMO1 activating enzyme subunit 1. Together with Sae2 protein, Sae1 forms a heterodimeric complex called SUMO-activating enzyme E1, supporting SUMOylation *in vitro*. SUMOylation is a post-translational modification of proteins where small ubiquitin-like modifier (SUMO) 1 proteins are added to lysine residues. This modification regulate its structure and intracellular localisation (Zhao et al., 2020). *Carm1* encodes a multi-functional protein called coactivator-associated arginine methyltransferase 1, Carm1. The protein is active within post-translational modification of mRNA, transcriptional activation, cancer, autophagy, and early development. Related to the latter, the transferase has, in mammalian embryos, been reported to play an important role in determination of cell fate, as it is central in keeping the stem cells in the inner cell mass pluripotent through methylation. Overexpression of the gene in blastomeres induce an inner cell mass phenotype, with upregulation in both *Nanog* and *Sox2* (Torres-Padilla et al., 2007). The last gene, *Rhoa*, encodes a GTPase important for polymerisation of actin. It thus plays a central role in control of the cell shape regulation, as well as its polarity, and motility (Wheeler & Ridley, 2004). The protein is also involved in regulation of the cell cycle, for example by repressing p21 - a well-known Cyclin-dependent kinase (CDK) inhibitor (Coleman et al., 2004). Down-regulation of p21 is often associated with cancer, as the cell proliferation is increased (Gartel & Radhakrishnan, 2005).

Common for all of the genes considered as reference genes, is their association with important roles of cell metabolism. Because of their central functions, they are probably not only transcribed at high rates during early development and may serve as candidates for reference genes also in adult tissues. Even though results from the RNAseq data and variation in C_t values suggests *Roaa* as a good candidate as endogenous control gene, it should be carefully assessed when used. The C_t value evaluation was only based on two runs of qPCR, which is not sufficient for drawing a conclusive result. To further improve the results of the qPCR, more than one endogenous control should have been used (Vandesompele et al., 2002).

There may have been other genes that could have been relevant for this experimental design. A recent study examined four different housekeeping genes in developing primordial germ cells of salmon when investigating sex determination genes. The genes that were evaluated included *GAPDH*, *Ubiquitin-conjugating enzyme E2 L3* (*UB2L3*), *eEf1a*, and *β-actin*. The study concluded that *UB2L3* was the most stable gene, and it was used as a control gene in

qPCR analysis of samples collected throughout salmon embryogenesis (Bhat et al., 2023). Thus, *UB2L3* could also have been considered as a reference gene in this study.

4.2.2 Pluripotency genes in RNAseq data and qPCR

The RNAseq data reveals that *Nanog* acts as expected with high expression before differentiation initiates at gastrulation. The same expression difference is seen in the qPCR results, where the gene is highly upregulated during blastulation stages. This confirms that *Nanog* serves an important role in maintaining pluripotency also in salmon and works well as a pluripotent biomarker. *Pou5f1*, however, is highly expressed also entering gastrulation, but rapidly decrease in expression towards somitogenesis, according to the RNAseq data. This pattern could explain why the gene shows such a moderate fold-change in expression between the blastulation and gastrulation in the qPCR results. Alternatively, *Pou5f1-like* could be used as a marker for pluripotency. The gene show, according to RNAseq data (Figure 4.2), a more expression pattern that would be useful in a marker for pluripotency with a larger amount of mRNA expressed, and an earlier decrease in expression when entering gastrulation.

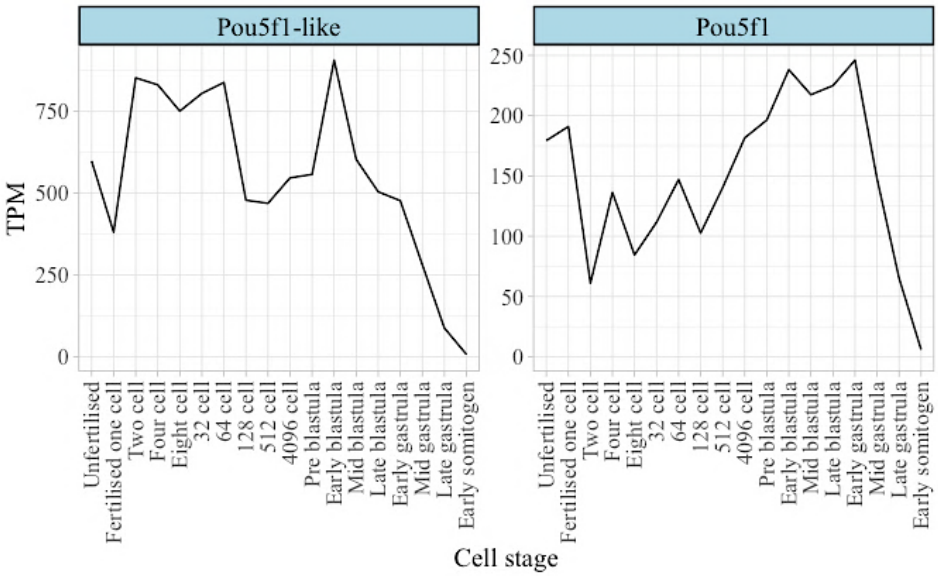


Figure 4.2 | mRNA expression measured in TPM in *Pou5f1* and *Pou5f1-like* during early development of salmon. *Pou5f1-like* is expressed at a higher rate compared to *Pou5f1* and do also decrease in expression before earlier. Blastula = Blastulation; Gastrula = Gastrulation; Somitogen = Somitogenesis.

Why the other pluripotency markers are not expressed as expected in salmon embryos is not clear. One explanation could be evolutionary differences. Humans and fish separated hundreds of millions of years ago in terms of evolution and have resulted in significant divergence in the genome. For instance, the expression pattern of *Sox2* varies between mammals and zebrafish. While the gene is not linked to the preservation of pluripotency in zebrafish, it reportedly plays a crucial part in neural differentiation (Gou et al., 2018). It is possible that salmon *Sox2* also has a similar function.

As a result of WGD, the salmon genome has multiple copies of all chromosomes. This have had a profound effect on the genome, gene expression, protein function and ultimately the phenotype. Multiple copies of a gene may result in genes acquiring new functions as a product of independent evolvement, also called gene retention (McGrath et al., 2014). Gene retention is often associated with the genomic dosage balance, which is the concept of genes needing to be expressed at the right level to ensure proper function. In a diploid organism, the expression of the two alleles of the gene is balanced through the two copies. In a polyploid organism, such as salmon, this process of balancing becomes more complex due to multiple copies of the gene. Imbalance can lead to reduced fitness, which is why many organisms have evolved mechanisms to maintain the balance through epigenetic modifications or gene retention (Wilson & Liberles, 2022).

The differences in expression patterns of the transcription factors thought to regulate pluripotency between the two species may reflect the differences in the evolutionary pressures that have shaped the retention and divergence of duplicated genes after WGD. It could be a result of the differences in environments, developmental processes, or selective pressures that may have led to different function for *Sox2*, *Klf4*, and *Tcf3a* and different patterns of gene expression.

4.2.3 *Apoa1* and *K2c8* as differentiation markers

Of the differentiation factors, *Apoa1* and *K2c8* were tested using qPCR. *Apoa1* encodes the main apolipoproteins of plasma high-density lipoproteins (HDLs). HDLs facilitate the transportation of cholesterol from different tissues to the liver (Cochran et al., 2021). In a study conducted by Nagae et al. the methylation levels on the *Apoa1* gene locus were examined during *in vitro* differentiation of human embryonic stem cells towards the hepatic lineage. After 13 days of development, a decrease in methylation was observed in the gene locus. The study also compared methylation levels in iPSCs and their parental cell. They confirmed a hypomethylation state in the parental cell but found that methylation was fully restored upon reprogramming (Nagae et al., 2011). Methylation is an epigenomic mechanism tightly associated with gene expression. Heavily methylation of a locus causes silence of the target gene. The mechanism mediate which genes are expressed in each cell type and is central upon differentiation of a stem cell (Moore et al., 2013). These results suggest similarity in gene expression patterns between the human and salmon *Apoa1* gene.

K2c8, also known as *Cytokeratin 8* or *Keratin 8*, belongs to a group of genes that encode intermediate filaments, and plays a role in supporting the cytoskeleton and cell motility of epithelial cells (Fillies et al., 2006). In both embryonic and mature zebrafish, *K2c8* is expressed in simple epithelia and its expression is closely associated with the regeneration of the caudal fin. This expression is exclusively observed in epidermal cells, making *K2c8* an appropriate epidermal marker during regeneration. The gene is also found highly expressed during related processes, such as development (Martorana et al., 2002). The equivalent protein is also identified in rainbow trout, where it shows similar expression patterns (Markl & Franke, 1988).

Apoa1 has been found to be involved in the differentiation process of humans, while *K2c8* has been shown to play a role in zebrafish differentiation. The positive results of the present study further suggest that the use of these markers in salmon differentiation studies could yield important insights into the mechanisms of differentiation and regeneration in fish species.

4.2.4 Protein verification using Western blot

Pou5f1 comprise 467 amino acids, and hold a molecular weight of 51.89 kilodaltons (kDa) (Bioinformatics.org, 2023). Another protein that may be sensitive to the antibody is Pou5f1-like. However, it consist of 469 amino acids and weighs 56,270 kDa (UniProt, 2022). According to the manufacturer, the antibody shows reactivity in mouse and human, but no data is shared for salmon (Sigma-Aldrich). To look at how the antibody epitope aligns with the salmon Pou5f1 and Pou5f1-like transcript, a BLAST search was performed using the NCBI BLAST tool. The results reveal a 100 % match of six amino acids in both genes. This region, which is the six first amino acids of the epitope sequence, seemed to be conserved in other POU domain factors as well (Table 4.1). The molecular weights of the proteins listed in Table 4.1 (ranging from 37.4-70.4 kDa), however, do not correspond to the bands observed on the WB membrane.

Table 4.1 | Results obtained from BLAST tool, aligning the epitope sequence of Oct4 antibody against different POU domain factors in salmon. Red letter indicates mismatch with the query.

Protein	RQKGKRSSSDYAQRED (query)	Identities
Pou5f1	RQKGKR	100 %
Pou5f1-like	RQKGKR	100 %
Pou3f2-like	RQKEKR	83 %
Pou3f1	RQKEKR	83 %
Pou3f1-A-like	RQKEKR	83 %
Pou3f3-B	RQKEKR	83 %
Pou3f2	RQKEKR	83 %
Pou4f3	RQKQKR	83 %
Pou4f1	RQKQKR	83 %
Pou4f2	RQKQKR	83 %
Pou3f2a	RQKEKR	83 %
Pou2f2-like	RQKEKR	83 %
Pou2f1 isoform X1	RQKEKRINPPSSS	62 %
Pou2f1-like isoform X1	RQKEKRINPPSSS	62 %

A perquisite to a well performing antibody is its specific binding of the protein of interest. Although the antibody may have a stronger affinity towards Pou5f1 and Pou5f1-like proteins, compared to the other POU-domain factors listed in the table, it may also have the ability to bind to other proteins. This would mean a cross-reactivity of the antibody, leading to a non-specific binding.

An alternate explanation could be that the two bands show two proteins resulting from alternative splicing. This is a process by which different exons of the gene are spliced together, resulting in different isoforms of the protein (Brickman & Serup, 2017). If exons are spliced out, the result would be production of smaller proteins and could explain why bands appear at a lower molecular weight in the membrane than the size of the original protein.

There is no documentation of splicing variants of *Pou5f1* in salmon, but two isoforms have been identified in humans. With different NH_2-termini these variants are thought to have different functional properties and are present at low amounts also in adult tissues (Takeda et al., 1992). They are also expressed during early development (Cauffman et al., 2006). Isomers of the zebrafish homolog *Pou2* have also been identified. These are also thought to have distinct functional capacities but are both thought to serve as regulatory proteins during embryogenesis. The sizes were confirmed to be 63 and 57 kDa (Takeda et al., 1994), which is close the size of the original protein of salmon Pou5f1.

Solely relying on the WB are not sufficient to determine which of the POU domain factors are translated in salmon cells and whether the antibody targets these proteins or not. It could also be that another protein, or its splice variants, are targeted by the antibody.

The epitope of the Sox17 antibody is 50 bp long, giving higher specificity compared to the Oct4 antibody, which is only 16 bp. It is based on *Sox17* gene from human and is reportedly reactive in both human and mice when used for Western blotting (antibodies-online.com). However, cross-reactivity have also been shown in monkey, rat (*Rattus rattus*) and zebrafish.

Why multiple bands appear on the gel could be that the antibody does not bind specifically to the protein of interest. It could bind off-target, in this case to more than one protein, leading to multiple appearances on the gel. To check the specificity between the epitope and the antigen, the epitope sequence was aligned with salmon Sox17a, using the BLAST tool from NCBI. 47 out of 50 amino acids matched (Figure 4.3). Moreover, the "+" sign indicate a similar amino acid – even though not the exact same, resulting in a positive score of 98 %.

Score	Expect	Method	Identities	Positives	Gaps
99.0 bits(245)	2e-31	Composition-based stats.	47/50(94%)	49/50(98%)	0/50(0%)

```
Query  65   PMNAFMVWAKDERKRLAQQNPDLHNAELSKMLGKSWKALPVSEKRPFVEE
            PMNAFMVWAKDERKRLAQQNPDLHNAELSKMLGKSWKAL ++EKRPFVEE
Sbjct  1    PMNAFMVWAKDERKRLAQQNPDLHNAELSKMLGKSWKALTLAEKRPFVEE
```

Figure 4.3 | BLAST results from Atlantic salmon and the epitope sequence. The query is here the epitope of the antibody, and the subject (Sbjct) the *Sox17a* gene of salmon. 47 out of 50 amino acids align perfectly, where 3 amino acids don't. These are marked either with no symbol, indicating a non-similar substitution, or with a "+" symbol, indicating similarity but no exact match. With two "+" symbols, the positive score is 98 %.

While a good match between an antibody and its intended target may reduce the likelihood of off-target effects, it does not completely eliminate the possibility. A majority of the bands were visible at a higher molecular weight than expected. Multiple bands could indicate different protein isoforms, as a result of post-translational modifications (PTM) or alternative splicing variants (Ramos-Vara, 2005). Alternative splicing would generate a smaller protein and should thus be visible at molecular wights lower than the initial protein weight. This is not likely, as the bands appear at a higher molecular weight than expected. During PTM functional groups such as carbohydrates or phosphate are added to the protein – a process known to have multiple important roles in cell signalling as well as protein functionality (Beltrao et al., 2013). These alterations can affect the protein mobility through the gel, resulting in multiple bands. The bands could also appear due to a shift in molecular weight as a result of the PTM. As far as the available literature indicates, there is no conclusive evidence regarding the specifics of PTM on Sox17 proteins. However, PTMs of the Sox family members such as Sox2 and Sox9 have been frequently reported. Modifications include acetylation, phosphorylation, ubiquitination, methylation, and SUMOylation (Williams et al., 2020). Common methods to detect PTMs is mass spectroscopy or biochemical assay (Meftahi et al., 2021), and could be used to identify variants of Sox17.

When looking at the expression data from RNAseq of *Sox17a* (Figure 4.4), one can see that the expression increases already at mid blastulation, which would explain why bands are appearing in the late blastulation sample of the Western blot. However, one would expect more apparent bands in the gastrulation samples compared to the other two, which was not the case. The expression, however, seemed more apparent with a stronger signal in the somitogenesis samples, which is not what we would have expected when looking at the expression pattern. The reason for this could be attributed to the date of sample collection, which was determined based on Gorodilov's description of early developing salmon embryos. To accurately track the development of fertilised eggs, it is crucial to maintain a stable temperature (Gorodilov, 1996). However, the temperature may not have been consistently monitored, which could have resulted in the collection of underdeveloped embryos if the temperature was lower than estimated.

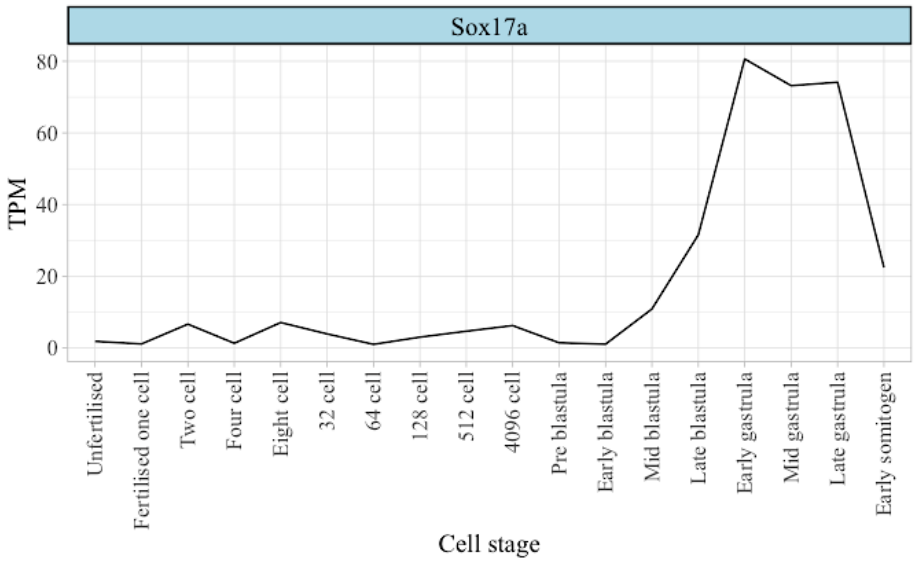


Figure 4.4 | Expression pattern of *Sox17* based on RNAseq data. The expression of the Sox17 protein is low before reaching mid blastulation, where the expression rate increases drastically. With a peak at early gastrulation, the expression sharply decreases when getting to somitogenesis. Blastula = Blastulation; Gastrula = Gastrulation; Somitogen = Somitogenesis.

5 CONCLUSION

This thesis aimed to develop a medium used for salmon embryonic stem cells and investigate potential markers of pluripotency and differentiation within these cells.

The media tested showed a common limitation of low proliferation capacity. Premature senescence was observed within days in culture in four out of five media, while the cells in the fifth culture medium died as a result of swelling and rupturing. Interestingly, differentiation was observed in the media with senescent cells, where L-15-based media induced differentiation on gelatin-coated wells and 2i-medium on fibronectin coats. Furthermore, better cell attachment was observed in these media, indicating a correlation between adhesion and cell differentiation. Further development should focus on optimising the concentrations of proliferating agents as well as supplements of antioxidants, to improve cell division and reduce differentiation.

Nanog exhibited the highest fold-change increase in expression during blastulation compared to gastrulation and was confirmed as a good marker for pluripotent phenotype based on qPCR analysis. *Pou5f1* showed a moderate increase in comparison, and *Pou5f1-like* gene was suggested as an alternative based on RNAseq data. As for the differentiation markers tested, both *Apoa1* and *K2c8* proved to be effective. The qPCR results for and differentiation markers demonstrated a significant difference in expression between the samples. *Roaa* was chosen as an endogenous control and used to normalise the data.

In Western blot analysis, *Pou5f1* was employed as a marker of pluripotency, while *Sox17a* was used as a marker of differentiation. The membrane probed with the Pou5f1 antibody exhibited two bands, indicating alternative splicing of the protein target. Conversely, the signal for Sox17a appeared to be streaked in the wells, potentially resulting from post-translational modifications like glycosylation. Sox17 is not linked to any PTMs, but members of the Sox-family are, suggesting a common feature within the family. We're not able to conclude that either of the antibodies are non-specific based on this information. Additional analysis is suggested to confirm *Pou5f1* and *Sox17a* as genetic markers during early development in salmon.

6 REFERENCES

Ahmed, M., & Ffrench-Constant, C. (2016). Extracellular Matrix Regulation of Stem Cell Behavior. *Curr Stem Cell Rep*, *2*(3), 197-206. https://doi.org/10.1007/s40778-016-0056-2

Albert, A., & Zhang, L. (2010). A novel definition of the multivariate coefficient of variation. *Biometrical Journal*, *52*(5), 667-675.

Alison, M. R., Poulsom, R., Forbes, S., & Wright, N. A. (2002). An introduction to stem cells. *J Pathol*, *197*(4), 419-423. https://doi.org/10.1002/path.1187

Alvarez, M. C., Bejar, J., Chen, S., & Hong, Y. (2007). Fish ES cells and applications to biotechnology. *Mar Biotechnol (NY)*, *9*(2), 117-127. https://doi.org/10.1007/s10126-006-6034-4

Ambrosetti, D. C., Basilico, C., & Dailey, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Molecular and Cellular Biology*, *17*(11), 6321-6329. https://doi.org/doi:10.1128/MCB.17.11.6321

Amit, M., Carpenter, M. K., Inokuma, M. S., Chiu, C.-P., Harris, C. P., Waknitz, M. A., Itskovitz-Eldor, J., & Thomson, J. A. (2000). Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. *Developmental biology*, *227*(2), 271-278.

antibodies-online.com. (2023). *SOX17 antibody (AA 70-120)*. https://www.antibodies-online.com/antibody/2364850/anti-SRY+Sex+Determining+Region+Y-Box+17+SOX17+AA+70-120+antibody/

Arango, M., Quintero-Ronderos, P., Castiblanco, J., & Montoya-Ortíz, G. (2013). Cell culture and cell analysis. *Autoimmunity: from bench to bedside*, 741-754.

Batish, I., Zarei, M., Nitin, N., & Ovissipour, R. (2022). Evaluating the Potential of Marine Invertebrate and Insect Protein Hydrolysates to Reduce Fetal Bovine Serum in Cell Culture Media for Cultivated Fish Production. *Biomolecules*, *12*(11). https://doi.org/10.3390/biom12111697

Beltrao, P., Bork, P., Krogan, N. J., & van Noort, V. (2013). Evolution and functional cross-talk of protein post-translational modifications. *Molecular systems biology*, *9*(1), 714.

Bhat, I. A., Dubiel, M. M., Rodriguez, E., & Jónsson, Z. O. (2023). Insights into Early Ontogenesis of Salmo salar: RNA Extraction, Housekeeping Gene Validation and Transcriptional Expression of Important Primordial Germ Cell and Sex-Determination Genes. *Animals*, *13*(6), 1094.

Bio-Rad. (2022). *Clarity™ Western ECL Substrate Instruction Manual*. https://lab.vanderbilt.edu/wp-content/uploads/sites/21/public_files/083013_BioRad_Clarity_Substrate%20Manual.pdf

Bioinformatics.org. (2023). Protein Molecular Weight. https://www.bioinformatics.org/sms/prot_mw.html

Bolstad, G. H., Hindar, K., Robertsen, G., Jonsson, B., Sægrov, H., Diserud, O. H., Fiske, P., Jensen, A. J., Urdal, K., & Næsje, T. F. (2017). Gene flow from domesticated escapes alters the life history of wild Atlantic salmon. *Nature Ecology & Evolution*, *1*(5), 1-5.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., & Jenner, R. G. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, *122*(6), 947-956.

Bratt-Leal, A. M., Carpenedo, R. L., & McDevitt, T. C. (2009). Engineering the embryoid body microenvironment to direct embryonic stem cell differentiation. *Biotechnology Progress*, *25*(1), 43-51.

Brickman, J. M., & Serup, P. (2017). Properties of embryoid bodies. *Wiley Interdiscip Rev Dev Biol*, *6*(2). https://doi.org/10.1002/wdev.259
Brind'Amour, K. (2009). Ethics and induced pluripotent stem cells. *United State: Embryo Project*.
Canaider, S., Maioli, M., Facchin, F., Bianconi, E., Santaniello, S., Pigliaru, G., Ljungberg, L., Burigana, F., Bianchi, F., & Olivi, E. (2014). Human stem cell exposure to developmental stage zebrafish extracts: a novel strategy for tuning stemness and senescence patterning. *CellR4*, *2*(5), e1226.
Carpenter, B., MacKay, C., Alnabulsi, A., MacKay, M., Telfer, C., Melvin, W. T., & Murray, G. I. (2006). The roles of heterogeneous nuclear ribonucleoproteins in tumour development and progression. *Biochim Biophys Acta*, *1765*(2), 85-100. https://doi.org/10.1016/j.bbcan.2005.10.002
Cauffman, G., Liebaers, I., Van Steirteghem, A., & Van de Velde, H. (2006). POU5F1 isoforms show different expression patterns in human embryonic stem cells and preimplantation embryos. *Stem cells*, *24*(12), 2685-2691.
Chandra, V., Tiwari, A., Pant, K. K., & Bhatt, R. (2022). Animal Cell Culture: Basics and Applications. In *Industrial Microbiology and Biotechnology* (pp. 691-719). Springer.
Chavez-Mardones, J., & Gallardo-Escarate, C. (2015). Next-Generation Transcriptome Profiling of the Salmon Louse Caligus rogercresseyi Exposed to Deltamethrin (AlphaMax): Discovery of Relevant Genes and Sex-Related Differences. *Mar Biotechnol (NY)*, *17*(6), 793-810. https://doi.org/10.1007/s10126-015-9661-9
Chen, S.-L., Ren, G.-C., Sha, Z.-X., & Shi, C.-Y. (2004). Establishment of a continuous embryonic cell line from Japanese flounder Paralichthys olivaceus for virus isolation. *Diseases of aquatic organisms*, *60*(3), 241-246.
Chen, S.-L., Sha, Z.-X., & Ye, H.-Q. (2003). Establishment of a pluripotent embryonic cell line from sea perch (Lateolabrax japonicus) embryos. *Aquaculture*, *218*(1-4), 141-151.
Chen, S.-L., Ye, H.-Q., Sha, Z.-X., & Hong, Y. (2003). Derivation of a pluripotent embryonic cell line from red sea bream blastulas. *Journal of Fish Biology*, *63*(3), 795-805. https://doi.org/https://doi.org/10.1046/j.1095-8649.2003.00192.x
Chen, S. S., Fitzgerald, W., Zimmerberg, J., Kleinman, H. K., & Margolis, L. (2007). Cell-cell and cell-extracellular matrix interactions regulate embryonic stem cell differentiation. *Stem cells*, *25*(3), 553-561. https://doi.org/10.1634/stemcells.2006-0419
Chen, Y., Li, X., Eswarakumar, V. P., Seger, R., & Lonai, P. (2000). Fibroblast growth factor (FGF) signaling through PI 3-kinase and Akt/PKB is required for embryoid body differentiation. *Oncogene*, *19*(33), 3750-3756.
Chou, Y. F., Chen, H. H., Eijpe, M., Yabuuchi, A., Chenoweth, J. G., Tesar, P., Lu, J., McKay, R. D., & Geijsen, N. (2008). The growth factor environment defines distinct pluripotent ground states in novel blastocyst-derived stem cells. *Cell*, *135*(3), 449-461. https://doi.org/10.1016/j.cell.2008.08.035
Chowdhury, K., Lin, S., & Lai, S.-L. (2022). Comparative study in Zebrafish and medaka unravels the mechanisms of tissue regeneration. *Frontiers in Ecology and Evolution*, *10*, 34.
Clark, D. P., & Pazdernik, N. J. (2016). Cloning and Stem Cells. *Biotechnology*, 565-591.
Coates, A., Phillips, B. L., Bui, S., Oppedal, F., Robinson, N. A., & Dempster, T. (2021). Evolution of salmon lice in response to management strategies: A review. *Reviews in Aquaculture*, *13*(3), 1397-1422.
Cochran, B. J., Ong, K.-L., Manandhar, B., & Rye, K.-A. (2021). APOA1: a protein with multiple therapeutic functions. *Current Atherosclerosis Reports*, *23*, 1-10.

Coleman, M. L., Marshall, C. J., & Olson, M. F. (2004). RAS and RHO GTPases in G1-phase cell-cycle regulation. *Nature reviews Molecular cell biology*, *5*(5), 355-366.
Costello, C., Cao, L., Gelcich, S., Cisneros-Mata, M. A., Free, C. M., Froehlich, H. E., Golden, C. D., Ishimura, G., Maier, J., Macadam-Somer, I., Mangin, T., Melnychuk, M. C., Miyahara, M., de Moor, C. L., Naylor, R., Nostbakken, L., Ojea, E., O'Reilly, E., Parma, A. M., . . . Lubchenco, J. (2020). The future of food from the sea. *Nature*, *588*(7836), 95-100. https://doi.org/10.1038/s41586-020-2616-y
Courtois-Cox, S., Jones, S. L., & Cichowski, K. (2008). Many roads lead to oncogene-induced senescence. *Oncogene*, *27*(20), 2801-2809. https://doi.org/10.1038/sj.onc.1210950
Ding, V. M., Ling, L., Natarajan, S., Yap, M. G., Cool, S. M., & Choo, A. B. (2010). FGF-2 modulates Wnt signaling in undifferentiated hESC and iPS cells through activated PI3-K/GSK3β signaling. *Journal of cellular physiology*, *225*(2), 417-428.
Doherty, P., & Walsh, F. S. (1996). CAM-FGF receptor interactions: a model for axonal growth. *Molecular and Cellular Neuroscience*, *8*(2-3), 99-111.
Donovan, P. J., & Gearhart, J. (2001). The end of the beginning for pluripotent stem cells. *Nature*, *414*(6859), 92-97. https://doi.org/10.1038/35102154
Dragan, A. I., Pavlovic, R., McGivney, J. B., Casas-Finet, J. R., Bishop, E. S., Strouse, R. J., Schenerman, M. A., & Geddes, C. D. (2012). SYBR Green I: fluorescence properties and interaction with DNA. *J Fluoresc*, *22*(4), 1189-1199. https://doi.org/10.1007/s10895-012-1059-8
Dvorak, P., Dvorakova, D., & Hampl, A. (2006). Fibroblast growth factor signaling in embryonic and cancer stem cells. *FEBS Lett*, *580*(12), 2869-2874. https://doi.org/10.1016/j.febslet.2006.01.095
Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, *292*(5819), 154-156. https://doi.org/10.1038/292154a0
Fan, Z., Liu, L., Huang, X., Zhao, Y., Zhou, L., Wang, D., & Wei, J. (2017). Establishment and growth responses of Nile tilapia embryonic stem-like cell lines under feeder-free condition. *Dev Growth Differ*, *59*(2), 83-93. https://doi.org/10.1111/dgd.12341
FDA, U. S. F. A. D. A. (2022). *AquAdvantage salmon fact sheet*. https://www.fda.gov/animal-veterinary/aquadvantage-salmon/aquadvantage-salmon-fact-sheet
Fedoroff, N., Benfey, T., Giddings, L. V., Jackson, J., Lichatowich, J., Lovejoy, T., Stanford, J., Thurow, R. F., & Williams, R. N. (2022). Biotechnology can help us save the genetic heritage of salmon and other aquatic species. *Proceedings of the National Academy of Sciences*, *119*(19), e2202184119.
Fillies, T., Werkmeister, R., Packeisen, J., Brandt, B., Morin, P., Weingart, D., Joos, U., & Buerger, H. (2006). Cytokeratin 8/18 expression indicates a poor prognosis in squamous cell carcinomas of the oral cavity. *BMC Cancer*, *6*, 10. https://doi.org/10.1186/1471-2407-6-10
Folmar, L. C., & Dickhoff, W. W. (1980). The parr—Smolt transformation (smoltification) and seawater adaptation in salmonids: A review of selected literature. *Aquaculture*, *21*(1), 1-37.
Fuchs, E., Tumbar, T., & Guasch, G. (2004). Socializing with the neighbors: stem cells and their niche. *Cell*, *116*(6), 769-778.
Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J.-i., & Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Genes & development*, *16*(7), 784-789.
Furue, M., Okamoto, T., Hayashi, Y., Okochi, H., Fujimoto, M., Myoishi, Y., Abe, T., Ohnuma, K., Sato, G. H., & Asashima, M. (2005). Leukemia inhibitory factor as an

anti-apoptotic mitogen for pluripotent mouse embryonic stem cells in a serum-free medium without feeder cells. *In Vitro Cellular & Developmental Biology-Animal*, *41*(1), 19-28.
Gartel, A. L., & Radhakrishnan, S. K. (2005). Lost in transcription: p21 repression, mechanisms, and consequences. *Cancer Res*, *65*(10), 3980-3985. https://doi.org/10.1158/0008-5472.CAN-04-3995
Gorodilov, Y. N. (1996). Description of the early ontogeny of the Atlantic salmon, Salmo salar, with a novel system of interval (state) identification. *Environmental Biology of fishes*, *47*, 109-127.
Gospodarowicz, D., Delgado, D., & Vlodavsky, I. (1980). Permissive effect of the extracellular matrix on cell proliferation in vitro. *Proc Natl Acad Sci U S A*, *77*(7), 4094-4098. https://doi.org/10.1073/pnas.77.7.4094
Gou, Y., Vemaraju, S., Sweet, E. M., Kwon, H. J., & Riley, B. B. (2018). sox2 and sox3 Play unique roles in development of hair cells and neurons in the zebrafish inner ear. *Dev Biol*, *435*(1), 73-83. https://doi.org/10.1016/j.ydbio.2018.01.010
Han, S. W., & Roman, J. (2006). Fibronectin induces cell proliferation and inhibits apoptosis in human bronchial epithelial cells: pro-oncogenic effects mediated by PI3-kinase and NF-kappa B. *Oncogene*, *25*(31), 4341-4349. https://doi.org/10.1038/sj.onc.1209460
Havforskningsinstituttet. (2022). Risikorapport norsk fiskeoppdrett 2022 - risikovurdering. https://www.hi.no/hi/nettrapporter/rapport-fra-havforskningen-2022-12
Hermitte, S., & Chazaud, C. (2014). Primitive endoderm differentiation: from specification to epithelium formation. *Philos Trans R Soc Lond B Biol Sci*, *369*(1657). https://doi.org/10.1098/rstb.2013.0537
Heslop, J. A., Pournasr, B., Liu, J.-T., & Duncan, S. A. (2021). GATA6 defines endoderm fate by controlling chromatin accessibility during differentiation of human-induced pluripotent stem cells. *Cell reports*, *35*(7), 109145.
Holen, E., Kausland, A., & Skjaerven, K. (2010). Embryonic stem cells isolated from Atlantic cod (Gadus morhua) and the developmental expression of a stage-specific transcription factor ac-Pou2. *Fish Physiol Biochem*, *36*(4), 1029-1039. https://doi.org/10.1007/s10695-010-9381-z
Hong, Y., Winkler, C., & Schartl, M. (1996). Pluripotency and differentiation of embryonic stem cell lines from the medakafish (Oryzias latipes). *Mechanisms of development*, *60*(1), 33-44.
Joseph-Silverstein, J., Consigli, S. A., Lyser, K. M., & Ver Pault, C. (1989). Basic fibroblast growth factor in the chick embryo: immunolocalization to striated muscle cells and their precursors. *The Journal of cell biology*, *108*(6), 2459-2466.
Katayama, M., Fukuda, T., Kaneko, T., Nakagawa, Y., Tajima, A., Naito, M., Ohmaki, H., Endo, D., Asano, M., Nagamine, T., Nakaya, Y., Saito, K., Watanabe, Y., Tani, T., Inoue-Murayama, M., Nakajima, N., & Onuma, M. (2022). Induced pluripotent stem cells of endangered avian species. *Commun Biol*, *5*(1), 1049. https://doi.org/10.1038/s42003-022-03964-y
Kayaci, A., Mehmet, C., GÜNER, Y., GÜLEÇ, F., & Mehmet, İ. (2015). Genetic Impact Determination of Farmed Fish on Native Fish by mtDNA Markers. *KSÜ Doğa Bilimleri Dergisi*, *18*(1), 18-25.
Kim, N. (2021). pH variation impacts molecular pathways associated with somatic cell reprogramming and differentiation of pluripotent stem cells. *Reprod Med Biol*, *20*(1), 20-26. https://doi.org/10.1002/rmb2.12346
Knapp, D., & Tanaka, E. M. (2012). Regeneration and reprogramming. *Current opinion in genetics & development*, *22*(5), 485-493.

Kobayashi, K., Khan, A., Ikeda, M., Nakamoto, A., Maekawa, M., & Yamasu, K. (2018). In vitro analysis of the transcriptional regulatory mechanism of zebrafish pou5f3. *Exp Cell Res*, *364*(1), 28-41. https://doi.org/10.1016/j.yexcr.2018.01.023
Kozera, B., & Rapacz, M. (2013). Reference genes in real-time PCR. *Journal of applied genetics*, *54*, 391-406.
Kubo, A., Shinozaki, K., Shannon, J. M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H. J., & Keller, G. (2004). Development of definitive endoderm from embryonic stem cells in culture. *Development*, *131*(7), 1651-1662. https://doi.org/10.1242/dev.01044
LeBlanc, L., Kim, M., Kambhampati, A., Son, A. J., Ramirez, N., & Kim, J. (2022). beta-catenin links cell seeding density to global gene expression during mouse embryonic stem cell differentiation. *iScience*, *25*(1), 103541. https://doi.org/10.1016/j.isci.2021.103541
Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., Hvidsten, T. R., Leong, J. S., Minkley, D. R., & Zimin, A. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, *533*(7602), 200-205.
Liu, R., Li, M., Li, Z., Hong, N., Xu, H., & Hong, Y. (2015). Medaka Oct4 is essential for pluripotency in blastula formation and ES cell derivation. *Stem Cell Reviews and Reports*, *11*(1), 11-23.
Llames, S., Garcia-Perez, E., Meana, A., Larcher, F., & del Rio, M. (2015). Feeder Layer Cell Actions and Applications. *Tissue Eng Part B Rev*, *21*(4), 345-353. https://doi.org/10.1089/ten.TEB.2014.0547
Lopez, M. J., & Hall, C. A. (2023). Physiology, Osmosis. In *StatPearls*. https://www.ncbi.nlm.nih.gov/pubmed/32491541
Mahmood, T., & Yang, P. C. (2012). Western blot: technique, theory, and trouble shooting. *N Am J Med Sci*, *4*(9), 429-434. https://doi.org/10.4103/1947-2714.100998
Markl, J., & Franke, W. W. (1988). Localization of cytokeratins in tissues of the rainbow trout: fundamental differences in expression pattern between fish and higher vertebrates. *Differentiation*, *39*(2), 97-122.
Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A*, *78*(12), 7634-7638. https://doi.org/10.1073/pnas.78.12.7634
Martin, G. R., & Evans, M. J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. *Proceedings of the National Academy of Sciences*, *72*(4), 1441-1445.
Martorana, M. L., Tawk, M., Lapointe, T., Barre, N., Imboden, M., Joulie, C., Geraudie, J., & Vriz, S. (2002). Zebrafish keratin 8 is expressed at high levels in the epidermis of regenerating caudal fin. *International Journal of Developmental Biology*, *45*(2), 449-452.
Mathe, G., Amiel, J. L., Schwarzenberg, L., Cattan, A., & Schneider, M. (1963). Haematopoietic Chimera in Man after Allogenic (Homologous) Bone-Marrow Transplantation. (Control of the Secondary Syndrome. Specific Tolerance Due to the Chimerism). *Br Med J*, *2*(5373), 1633-1635. https://doi.org/10.1136/bmj.2.5373.1633
Mathon, N. F., & Lloyd, A. C. (2001). Cell senescence and cancer. *Nat Rev Cancer*, *1*(3), 203-213. https://doi.org/10.1038/35106045
McGee, W. (2019). *sleuth-CN*. GitHub. https://github.com/warrenmcg/sleuth-CN/blob/master/R/choose_denom.R
McGee, W. A., Pimentel, H., Pachter, L., & Wu, J. Y. (2019). Compositional Data Analysis is necessary for simulating and analyzing RNA-Seq data. *bioRxiv*, 564955.

McGrath, C. L., Gout, J. F., Johri, P., Doak, T. G., & Lynch, M. (2014). Differential retention and divergent resolution of duplicate genes following whole-genome duplication. *Genome Res*, *24*(10), 1665-1675. https://doi.org/10.1101/gr.173740.114
Meftahi, G. H., Bahari, Z., Zarei Mahmoudabadi, A., Iman, M., & Jangravi, Z. (2021). Applications of western blot technique: From bench to bedside. *Biochem Mol Biol Educ*, *49*(4), 509-517. https://doi.org/10.1002/bmb.21516
Miller, D. M., Thomas, S. D., Islam, A., Muench, D., & Sedoris, K. (2012). c-Myc and cancer metabolism. *Clin Cancer Res*, *18*(20), 5546-5553. https://doi.org/10.1158/1078-0432.CCR-12-0977
Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., & Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, *113*(5), 631-642.
Mohapatra, C., Patra, S. K., Panda, R. P., Mohanta, R., Saha, A., Saha, J. N., Das Mahapatra, K., Jayasankar, P., & Barman, H. K. (2014). Gene structure and identification of minimal promoter of Pou2 expressed in spermatogonial cells of rohu carp, Labeo rohita. *Mol Biol Rep*, *41*(6), 4123-4132. https://doi.org/10.1007/s11033-014-3283-6
Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology*, *38*(1), 23-38.
Mushtaq, F., Raza, Z. A., Batool, S. R., Zahid, M., Onder, O. C., Rafique, A., & Nazeer, M. A. (2022). Preparation, properties, and applications of gelatin-based hydrogels (GHs) in the environmental, technological, and biomedical sectors. *Int J Biol Macromol*, *218*, 601-633. https://doi.org/10.1016/j.ijbiomac.2022.07.168
Nagae, G., Isagawa, T., Shiraki, N., Fujita, T., Yamamoto, S., Tsutsumi, S., Nonaka, A., Yoshiba, S., Matsusaka, K., Midorikawa, Y., Ishikawa, S., Soejima, H., Fukayama, M., Suemori, H., Nakatsuji, N., Kume, S., & Aburatani, H. (2011). Tissue-specific demethylation in CpG-poor promoters during cellular differentiation. *Hum Mol Genet*, *20*(14), 2710-2721. https://doi.org/10.1093/hmg/ddr170
Neave, B., Rodaway, A., Wilson, S. W., Patient, R., & Holder, N. (1995). Expression of zebrafish GATA 3 (gta3) during gastrulation and neurulation suggests a role in the specification of cell fate. *Mechanisms of development*, *51*(2-3), 169-182.
Niwa, H. (2007). How is pluripotency determined and maintained?
Niwa, H., Miyazaki, J.-i., & Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genetics*, *24*(4), 372-376. https://doi.org/10.1038/74199
Norsgaard, H., Clark, B. F., & Rattan, S. I. (1996). Distinction between differentiation and senescence and the absence of increased apoptosis in human keratinocytes undergoing cellular aging in vitro. *Experimental gerontology*, *31*(5), 563-570.
Norwegian seafood council, & Seafood Norway. (2021). *Norsk havbrukshistorie*. Laksefakta. https://laksefakta.no/#
Olsvik, P. A., Lie, K. K., Jordal, A. E., Nilsen, T. O., & Hordvik, I. (2005). Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. *BMC Mol Biol*, *6*, 21. https://doi.org/10.1186/1471-2199-6-21
Ou, J., Rosa, S., Berchowitz, L. E., & Li, W. (2019). Induced pluripotent stem cells as a tool for comparative physiology: lessons from the thirteen-lined ground squirrel. *Journal of Experimental Biology*, *222*(19), jeb196493.
Oyeleye, O., Ogundeji, S., Ola, S., & Omitogun, O. (2016). Basics of animal cell culture: Foundation for modern science. *Academic Journals*.
Pandey, G. (2013). Overview of fish cell lines and their uses. *Int. J. Pharm. Res. Sci*, *2*, 580-590.

Parameswaran, V., Shukla, R., Bhonde, R., & Hameed, A. S. (2007). Development of a pluripotent ES-like cell line from Asian sea bass (Lates calcarifer)--an oviparous stem cell line mimicking viviparous ES cells. *Mar Biotechnol (NY)*, *9*(6), 766-775. https://doi.org/10.1007/s10126-007-9028-y

Parchment, R. E., Lewellyn, A., Swartzendruber, D., & Pierce, G. B. (1990). Serum amine oxidase activity contributes to crisis in mouse embryo cell lines. *Proceedings of the National Academy of Sciences*, *87*(11), 4340-4344.

Park, E., & Patel, A. N. (2010). Changes in the expression pattern of mesenchymal and pluripotent markers in human adipose-derived stem cells. *Cell Biol Int*, *34*(10), 979-984. https://doi.org/10.1042/CBI20100124

Peerani, R., & Zandstra, P. W. (2010). Enabling stem cell therapies through synthetic stem cell–niche engineering. *The Journal of clinical investigation*, *120*(1), 60-70.

Peng, L., Zhou, Y., Xu, W., Jiang, M., Li, H., Long, M., Liu, W., Liu, J., Zhao, X., & Xiao, Y. (2019). Generation of Stable Induced Pluripotent Stem-like Cells from Adult Zebra Fish Fibroblasts. *Int J Biol Sci*, *15*(11), 2340-2349. https://doi.org/10.7150/ijbs.34010

Pettinato, G., Wen, X., & Zhang, N. (2014). Formation of well-defined embryoid bodies from dissociated human induced pluripotent stem cells using microfabricated cell-repellent microwell arrays. *Sci Rep*, *4*, 7402. https://doi.org/10.1038/srep07402

Petzold, J., & Gentleman, E. (2021). Intrinsic mechanical cues and their impact on stem cells and embryogenesis. *Frontiers in Cell and Developmental Biology*, 3112.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*, *29*(9), e45. https://doi.org/10.1093/nar/29.9.e45

Piechaczyk, M., Blanchard, J. M., Marty, L., Dani, C., Panabieres, F., El Sabouty, S., Fort, P., & Jeanteur, P. (1984). Post-transcriptional regulation of glyceraldehyde-3-phosphate-dehydrogenase gene expression in rat tissues. *Nucleic Acids Res*, *12*(18), 6951-6963. https://doi.org/10.1093/nar/12.18.6951

Piferrer, F., Beaumont, A., Falguière, J.-C., Flajšhans, M., Haffray, P., & Colombo, L. (2009). Polyploid fish and shellfish: production, biology and applications to aquaculture for performance improvement and genetic containment. *Aquaculture*, *293*(3-4), 125-156.

Raker, V., Plessel, G., & Lührmann, R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *The EMBO journal*, *15*(9), 2256-2269.

Ramirez, R. D., Morales, C. P., Herbert, B. S., Rohde, J. M., Passons, C., Shay, J. W., & Wright, W. E. (2001). Putative telomere-independent mechanisms of replicative aging reflect inadequate growth conditions. *Genes Dev*, *15*(4), 398-403. https://doi.org/10.1101/gad.859201

Ramos-Vara, J. A. (2005). Technical aspects of immunohistochemistry. *Veterinary pathology*, *42*(4), 405-426.

Rico, F., Chu, C., Abdulreda, M. H., Qin, Y., & Moy, V. T. (2010). Temperature modulation of integrin-mediated cell adhesion. *Biophys J*, *99*(5), 1387-1396. https://doi.org/10.1016/j.bpj.2010.06.037

Romito, A., & Cobellis, G. (2016). Pluripotent Stem Cells: Current Understanding and Future Directions. *Stem Cells Int*, *2016*, 9451492. https://doi.org/10.1155/2016/9451492

Sanchez-Sanchez, A. V., Camp, E., & Mullor, J. L. (2011). Fishing pluripotency mechanisms in vivo. *Int J Biol Sci*, *7*(4), 410-417. https://doi.org/10.7150/ijbs.7.410

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nature protocols*, *3*(6), 1101-1108.

Schneider, M., Marison, I. W., & von Stockar, U. (1996). The importance of ammonia in mammalian cell culture. *J Biotechnol*, *46*(3), 161-185. https://doi.org/10.1016/0168-1656(95)00196-4

Schuldiner, M., Yanuka, O., Itskovitz-Eldor, J., Melton, D. A., & Benvenisty, N. (2000). Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells. *Proc Natl Acad Sci U S A*, *97*(21), 11307-11312. https://doi.org/10.1073/pnas.97.21.11307

Shahdadfar, A., Frønsdal, K., Haug, T., Reinholt, F. P., & Brinchmann, J. E. (2005). In vitro expansion of human mesenchymal stem cells: choice of serum is a determinant of cell proliferation, differentiation, gene expression, and transcriptome stability. *Stem cells*, *23*(9), 1357-1366.

Sigma-Aldrich. (2023). *Anti-Oct-4 Antibody*. https://www.sigmaaldrich.com/NO/en/product/mm/ab3209

Sinner, D., Rankin, S., Lee, M., & Zorn, A. M. (2004). Sox17 and beta-catenin cooperate to regulate the transcription of endodermal genes. *Development*, *131*(13), 3069-3080. https://doi.org/10.1242/dev.01176

Statistisk sentralbyrå. (2022). *Eksport av fisk, etter statistikkvariabel, varegruppe og år*. https://www.ssb.no/statbank/sq/10074188

Suzuki, Y., Yanagisawa, M., Yagi, H., Nakatani, Y., & Robert, K. Y. (2010). Involvement of β1-integrin up-regulation in basic fibroblast growth factor-and epidermal growth factor-induced proliferation of mouse neuroepithelial cells. *Journal of Biological Chemistry*, *285*(24), 18443-18451.

Takagi, M., Hayashi, H., & Yoshida, T. (2000). The effect of osmolarity on metabolism and morphology in adhesion and suspension chinese hamster ovary cells producing tissue plasminogen activator. *Cytotechnology*, *32*(3), 171-179. https://doi.org/10.1023/A:1008171921282

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, *131*(5), 861-872. https://doi.org/10.1016/j.cell.2007.11.019

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, *126*(4), 663-676. https://doi.org/10.1016/j.cell.2006.07.024

Takeda, H., Matsuzaki, T., Oki, T., Miyagawa, T., & Amanuma, H. (1994). A novel POU domain gene, zebrafish pou2: expression and roles of two alternatively spliced twin products in early development. *Genes & development*, *8*(1), 45-59.

Takeda, J., Seino, S., & Bell, G. I. (1992). Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues. *Nucleic acids research*, *20*(17), 4613-4620.

Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A., & Heinen, E. (1999). Housekeeping genes as internal standards: use and limits. *Journal of biotechnology*, *75*(2-3), 291-295.

Thodesen, J., & Gjedrem, T. (2006). Breeding programs on Atlantic salmon in Norway: lessons learned.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., & Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, *282*(5391), 1145-1147. https://doi.org/10.1126/science.282.5391.1145

Torres-Padilla, M. E., Parfitt, D. E., Kouzarides, T., & Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature*, *445*(7124), 214-218. https://doi.org/10.1038/nature05458

Tsankov, A. M., Gu, H., Akopian, V., Ziller, M. J., Donaghey, J., Amit, I., Gnirke, A., & Meissner, A. (2015). Transcription factor binding dynamics during human ES cell differentiation. *Nature*, *518*(7539), 344-349.

UniProt. (2017a). *A0A1S3LM66 · A0A1S3LM66_SALSA* https://www.uniprot.org/uniprotkb/A0A1S3LM66/entry
UniProt. (2017b). *A0A1S3QA46 · A0A1S3QA46_SALSA* https://www.uniprot.org/uniprotkb/A0A1S3QA46/entry
UniProt. (2022). *A0A1S3KZV8 · A0A1S3KZV8_SALSA* https://www.uniprot.org/uniprotkb/A0A1S3KZV8/entry
Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*, *3*(7), RESEARCH0034. https://doi.org/10.1186/gb-2002-3-7-research0034
Vierck, J. L., Byrne, K., Mir, P. S., & Dodson, M. V. (2000). Ten commandments for preventing contamination of primary cell cultures. *Methods in cell science*, *22*, 33-41.
Vinitnantharat, S., Gravningen, K., & Greger, E. (1999). Fish vaccines. *Advances in veterinary medicine*, *41*, 539-550.
Wahl, S. M., Allen, J. B., Weeks, B. S., Wong, H. L., & Klotman, P. E. (1993). Transforming growth factor beta enhances integrin expression and type IV collagenase secretion in human monocytes. *Proceedings of the National Academy of Sciences*, *90*(10), 4577-4581.
Wan, X., Liu, Z., & Li, L. (2021). Manipulation of stem cells fates: the master and multifaceted roles of biophysical cues of biomaterials. *Advanced Functional Materials*, *31*(23), 2010626.
Wargelius, A., Leininger, S., Skaftnesmo, K. O., Kleppe, L., Andersson, E., Taranger, G. L., Schulz, R. W., & Edvardsen, R. B. (2016). Dnd knockout ablates germ cells and demonstrates germ cell independent sex differentiation in Atlantic salmon. *Sci Rep*, *6*, 21284. https://doi.org/10.1038/srep21284
Wei, Z., Yang, Y., Zhang, P., Andrianakos, R., Hasegawa, K., Lyu, J., Chen, X., Bai, G., Liu, C., Pera, M., & Lu, W. (2009). Klf4 interacts directly with Oct4 and Sox2 to promote reprogramming. *Stem cells*, *27*(12), 2969-2978. https://doi.org/10.1002/stem.231
Welshons, W. V., Wolf, M. F., Murphy, C. S., & Jordan, V. C. (1988). Estrogenic activity of phenol red. *Mol Cell Endocrinol*, *57*(3), 169-178. https://doi.org/10.1016/0303-7207(88)90072-x
Wesselschmidt, R. L. (2011). The teratoma assay: an in vivo assessment of pluripotency. In *Human Pluripotent Stem Cells* (pp. 231-241). Springer.
Wheeler, A. P., & Ridley, A. J. (2004). Why three Rho proteins? RhoA, RhoB, RhoC, and cell motility. *Exp Cell Res*, *301*(1), 43-49. https://doi.org/10.1016/j.yexcr.2004.08.012
White, M. G., Al-Turaifi, H. R., Holliman, G. N., Aldibbiat, A., Mahmoud, A., & Shaw, J. A. (2011). Pluripotency-associated stem cell marker expression in proliferative cell cultures derived from adult human pancreas. *J Endocrinol*, *211*(2), 169-176. https://doi.org/10.1530/JOE-11-0123
Williams, C. A. C., Soufi, A., & Pollard, S. M. (2020). Post-translational modification of SOX family proteins: Key biochemical targets in cancer? *Semin Cancer Biol*, *67*(Pt 1), 30-38. https://doi.org/10.1016/j.semcancer.2019.09.009
Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., & Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature*, *336*(6200), 684-687. https://doi.org/10.1038/336684a0
Wilson, A., & Liberles, D. (2022). Dosage balance acts as a time-dependent selective barrier to subfunctionalization.

Yang, K., Lee, J., & Cho, S. W. (2012). Engineering biomaterials for feeder-free maintenance of human pluripotent stem cells. *Int J Stem Cells*, *5*(1), 1-5. https://doi.org/10.15283/ijsc.2012.5.1.1

Yaskowiak, E. S., Shears, M. A., Agarwal-Mawal, A., & Fletcher, G. L. (2006). Characterization and multi-generational stability of the growth hormone transgene (EO-1alpha) responsible for enhanced growth rates in Atlantic Salmon. *Transgenic Res*, *15*(4), 465-480. https://doi.org/10.1007/s11248-006-0020-5

Ye, H., Du, H., Chen, X. H., Cao, H., Liu, T., & Li, C. J. (2012). Identification of a pou2 ortholog in Chinese sturgeon, Acipenser sinensis and its expression patterns in tissues, immature individuals and during embryogenesis. *Fish Physiol Biochem*, *38*(4), 929-942. https://doi.org/10.1007/s10695-011-9579-8

Yi, M., Hong, N., & Hong, Y. (2009). Generation of medaka fish haploid embryonic stem cells. *Science*, *326*(5951), 430-433.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, II, & Thomson, J. A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science*, *318*(5858), 1917-1920. https://doi.org/10.1126/science.1151526

Zaman, M. H. (2007). Understanding the molecular basis for differential binding of integrins to collagen and gelatin. *Biophysical journal*, *92*(2), L17-L19.

Zhao, Q., Ma, Y., Li, Z., Zhang, K., Zheng, M., & Zhang, S. (2020). The Function of SUMOylation and Its Role in the Development of Cancer Cells under Stress Conditions: A Systematic Review. *Stem Cells Int*, *2020*, 8835714. https://doi.org/10.1155/2020/8835714

Zheng, X., Baker, H., Hancock, W. S., Fawaz, F., McCaman, M., & Pungor Jr., E. (2006). Proteomic Analysis for the Assessment of Different Lots of Fetal Bovine Serum as a Raw Material for Cell Culture. Part IV. Application of Proteomics to the Manufacture of Biological Drugs. *Biotechnology Progress*, *22*(5), 1294-1300. https://doi.org/https://doi.org/10.1021/bp060121o

7 APPENDIX

Supplemental code 7.1 | Code used to generate the list of potential endogenous controls. The code selects candidates based on the lowest Cov value. It also selects for genes with a TPM mean over 5, and finally shows the top 5 genes in the list.

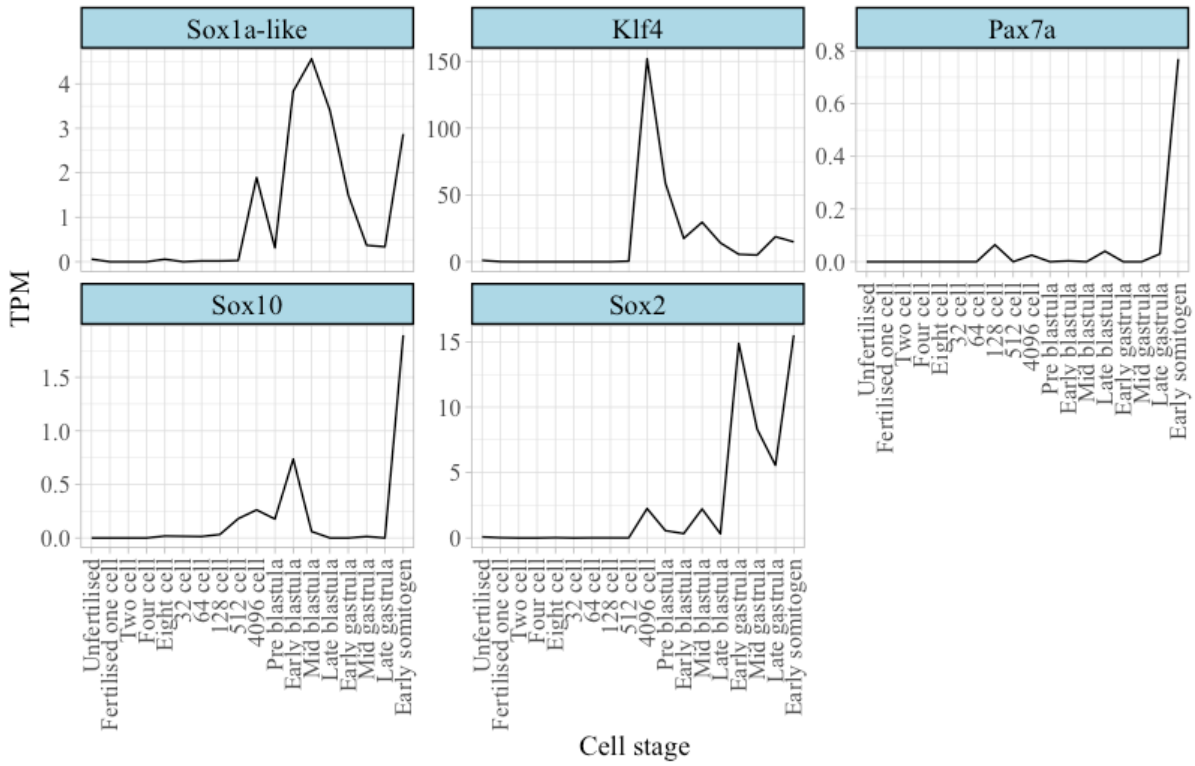
```
obj<-read.table(file = 'RNASeq_data.tsv', sep = '\t', header = TRUE)


allNonZero <- !matrixStats::rowAnys(as.matrix(obj[,-1]), value = 0)


mat <- obj[allNonZero,-1]


mean_vals <- rowMeans(mat, na.rm = T)


cov <- matrixStats::rowSds(as.matrix(mat), na.rm = T) / rowMeans(as.matrix(
mat), na.rm = T)


mean_cov_merge <- data.frame(mean_vals, cov)
c <- subset(mean_cov_merge, mean_vals >= 5 & cov > 0, na.rm = T) %>%
  arrange(cov)


head(c,5)
##                    mean_vals       cov
## ENSSSAG00000015001 503.19835 0.1600350
## ENSSSAG00000067428 471.91709 0.1612816
## ENSSSAG00000080426  86.14416 0.1687418
## ENSSSAG00000120073  74.56560 0.1706358
## ENSSSAG00000057386  82.56819 0.1727762
```

Supplemental code 7.2 | Code used to generate a list of possible markers for differentiation. The code selects objects where the value from the column late blastulation multiplied by four is less than the value in column early gastrulation, and where the value in the latter column is greater than 4. The list is then sorted after the highest values in late blastulation samples. The final output shows the top 10 genes in the list, where the five genes that generated specific primers were chosen.

```
obj_mean <- read.table(file = 'test.tsv', sep = '\t', header = T)


diff_markers <- obj_mean[which(obj_mean$X1_Late_blastulation * 4 < obj_mean
$X10_Early_gastrulation & obj_mean$X10_Early_gastrulation > 4), ]


sorted_indices <- order(diff_markers$X1_Late_blastulation, decreasing = T)


sorted_diff_markers <- diff_markers[sorted_indices, ]


head(row.names(sorted_diff_markers), 10)
##  [1] "ENSSSAG00000097945" "ENSSSAG00000007048" "ENSSSAG00000074245"
##  [4] "ENSSSAG00000045682" "ENSSSAG00000096881" "ENSSSAG00000062385"
##  [7] "ENSSSAG00000049420" "ENSSSAG00000114130" "ENSSSAG00000120869"
## [10] "ENSSSAG00000060970"
```

Supplemental table 7.1 | Genes tested using qPCR. The genes tested were based both on literature and RNAseq data. Based on primer efficiency, coefficient of determination (R^2), and their gene expression pattern, *Apoa1* and *K2c8* were selected as marker genes for differentiation, *Nanog* and *Pou5f1* for pluripotency, with *Roaa* as an endogenous control. The selected genes are here marked in with an underline.

Gene	**Efficiency (%)**	**R^2**
Apoa1	95.1	0.997
Afp4	55.2	0.616
Nanog	98.4	0.994
Pou5f1	92.5	0.975
Klf4	69.4	0.966
Sox2	87.0	1.00
Roaa	92.9	1.00
Rhoa	105	0.988
SmD2	94.6	0.999
Sae1	57.8	0.924
K2c8	90.7	0.997
Rsp20	80.6	0.999
Eef1a1a	159	0.386
Hepb2	-96.8	0.591
Carm1	383	0.904
Krt5	112	0.999



Supplemental figure 7.1 | Gene expression based on RNAseq data of the antibodies that were not selected for Western blotting.

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

Postboks 5003
NO-1432 Ås
Norway