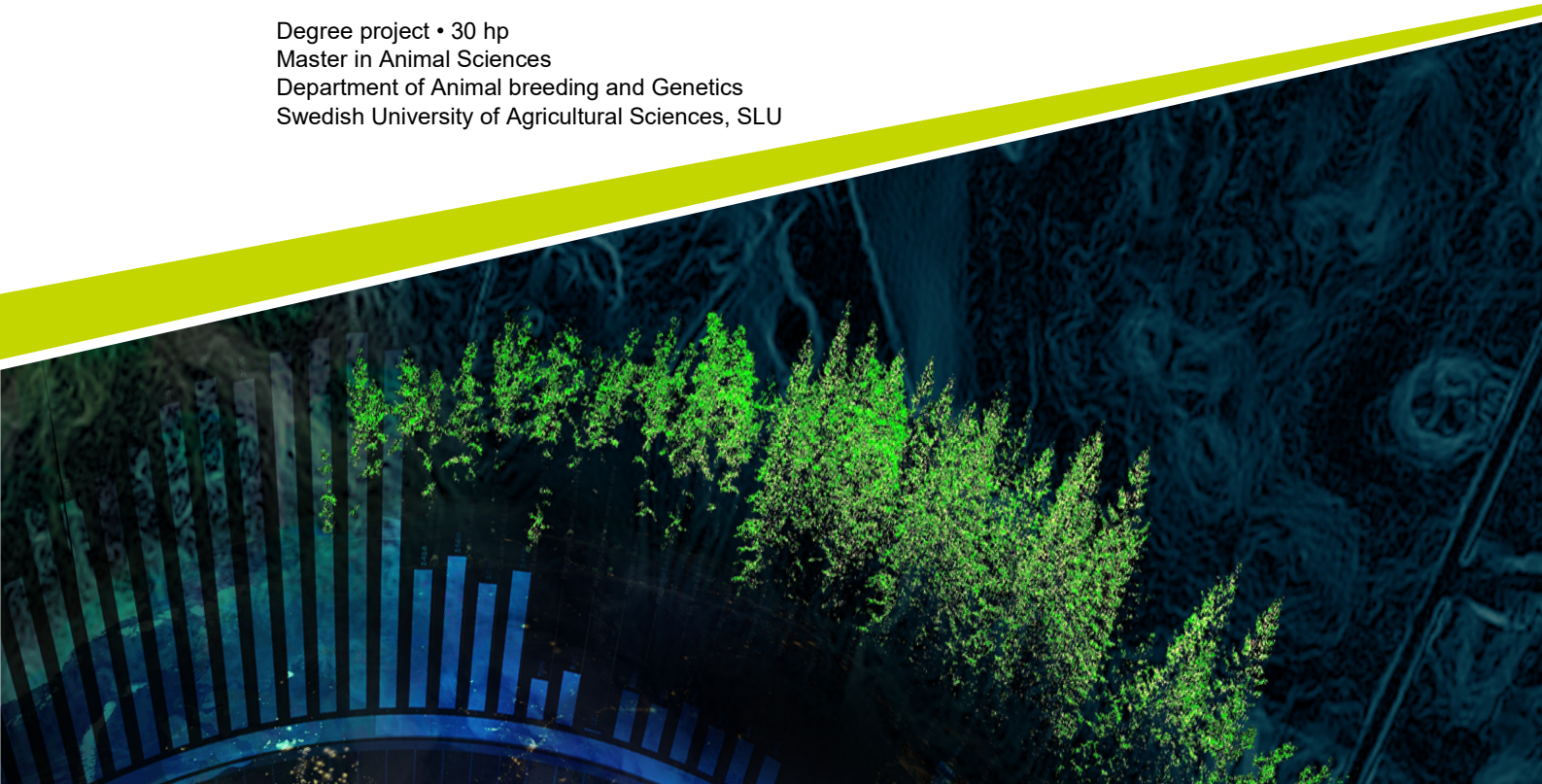
Genetic diversity status and associations with sperm quality characteristics in farmed Arctic charr (*Salvelinus alpinus* L.) across Sweden

Anam Anjum

Degree project • 30 hp
Master in Animal Sciences
Department of Animal breeding and Genetics
Swedish University of Agricultural Sciences, SLU



Genetic diversity status and associations with sperm quality characteristics in farmed Arctic charr (*Salvelinus alpinus* L.) across Sweden

Anam Anjum

Supervisor: **Christos Palaiokostas, Swedish university of Agriculture Sciences, Department of Animal breeding and genetics, Applied genetics.**

Assistant supervisor: Henrik Jeuthe, Swedish university of Agriculture Sciences, Research coordinator of Aquaculture Centre North.

Examiner: Dr. Markus Langeland, Swedish university of Agriculture Sciences, Department of Animal nutrition and management, Aquaculture.

Credits: 30 hp
Level: Second cycle
Course title: Degree project in Animal sciences - Master project
Course code: EX0870
Programme/education: Master in Animal sciences
Course coordinating dept: Department of Animal genetics

Place of publication: Uppsala, Sweden
Year of publication: 2020
Cover picture:

Keywords: Arctic charr, dd-RADseq, genetic diversity, GWAS, reproductive success

Swedish University of Agricultural Sciences
Faculty of Animal genetics
Department of Animal breeding and genetics

Archiving and publishing

Approved students' theses at SLU are published electronically. As a student, you have the copyright to your own work and need to approve the electronic publishing. When you have approved, metadata and full text of your thesis will be visible and searchable online. When the document is uploaded it is archived as a digital file.

[x] YES, I hereby give permission to publish the present thesis in accordance with the SLU agreement regarding the transfer of the right to publish a work. https://www.slu.se/en/subweb/library/publish-and-analyse/register-and-publish/agreement-for-publishing/

[ ] NO, I do not give permission to publish the present work. The work will still be archived and its metadata and abstract will be visible and searchable.

Abstract

The national breeding programme of Arctic charr (*Salvelinus alpinus* L.) has been of fundamental importance for the expansion of Arctic charr farming across Sweden. Selection has been ongoing since 1980s documenting approximately 10% improvement per generation regarding growth related traits. Currently, poor reproductive success (30-70%) as compared to other farmed salmonids is a major challenge for Arctic charr farming in Sweden. Both environmental and genetic factors could affect the reproductive performance in Arctic charr.

The reduction of sequencing cost with the advent of next generation sequencing (NGS) provides opportunities for aquaculture genetics and breeding to investigate in-depth complex traits. Double digestion restriction associated DNA (ddRAD) sequencing is an efficient, robust, flexible, and cost-effective technology for identifying genome wide genetic markers. In the current study, ddRAD sequencing was applied on farmed Arctic charr males from four different farms across Sweden.

A total of 5,159 single nucleotide polymorphisms (SNPs) were used for investigating the status of the genetic diversity and for associations with sperm quality characteristics (concentration, viable cells). The genetic diversity analysis indicated for low genetic differentiation among the sampled populations. Though, the level of genetic diversity within populations (heterozygosity estimates) was higher compared to prior ddRAD studies on farmed fish where no pedigree records were kept. The association analysis indicated for a significantly associated SNP (SNP_402577.127) with sperm cell viability located on chromosome number 31. This novel SNP could be a valuable tool for directing future selection practices and assisting to identify the male broodstocks with the highest sperm quality characteristics.

Keywords: Arctic charr, ddRAD sequencing, Next generation sequencing, Genome wide association study.

Contents

List of tables

List of figures

Abbreviations

FA	Fatty acid
ddRAD	Double digest restriction association DNA
NGS	Next generation sequencing
GBS	Genotyping by sequencing
GS	Genomic selection
RE	Restriction enzyme

1. Introduction

Arctic charr (*Salvelinus alpinus* L.*)* is a cold-water salmonid, native to the Scandinavian peninsula mountain region in a range of 56°-69° N. In addition, Arctic charr populations are also found in large deep lakes of southern Sweden (Johnson, 1980). Compared to the Atlantic region of the Scandinavian peninsula where Arctic charr populations are anadromous the Swedish populations spent their entire lifetime in freshwater (L.-O. Eriksson et al., 1993).

Artic charr can grow in cold water even at temperatures close to 0 °C (Brännäs & Wiklund, 1992). Opposite to other salmonids like Atlantic salmon (*Salmo salar*) and brown trout (*Salmo trutta*) Arctic charr rearing can be efficient even at temperatures below 10^0C (Elliott & Elliott, 2010). In comparison, Atlantic salmon grows well at 13°C (Hevrøy et al., 2013; Thyholdt, 2014) (range 12-18 °C, Elliott and Elliott, 2010) and brown trout at 14 °C (range 12-16 °C) (FAO, 2020).

1.1. Origins of the Swedish Arctic charr breeding programme

In the early 1900s Arctic charr farming in captivity begun in Sweden aiming to produce fry and juvenile fish on a small scale in natural earthern ponds. During the 1950s after the development of intensive tank-based rearing technology a large scale initiative of Arctic charr rearing took place in the country. Most importantly during the 1980s, Arctic charr production for the table market (farming market) was intensified following technological advancements related to the feed requirements and optimal rearing conditions (L. O. Eriksson et al., 2010). Overall, since the 1980s the Arctic char farming industry expanded from a production of about 500 to 800 tones annual to

2300 tons in 2010 (Jeuthe., 2012). In 2017, the production volume of Arctic charr was estimated around 1,310 metric tons (Statista, 2019).

1.2. Strain for selective breeding

In 1982, three populations of Arctic charr were selected aiming to develop a strain with the highest potential for intensive farming. The above populations included a pelagic plankton feeding morph, a benthic invertebrate feeding morph and a predatory morph originating from the Swedish Lakes Torrön, Tinnsjön and Hornavan respectively. During 1986 the native strain of Arctic charr (predatory morph) from Lake Hornavan was selected due to its high growth performance and later onset of sexual maturation characteristics (L.-O. Eriksson et al., 1993; L. O. Eriksson et al., 2010; Nilsson, 1993). The Aquaculture Centre North in Kälarne was the core facility of the breeding programme (Nilsson et al., 2010).

The first generation of selection consisted of 96 full-sib families that were obtained via factorial mating between 36 sires and 32 dams. Growth rate, age at first sexual maturation, resistance to fungal infection and flesh colouration where the focal traits for improvement. Passive integrated Transponder (PIT) tags were used for animal identification and pedigree construction (Nilsson et al., 2010). Typically, the breeding program over the years has been based on 47 – 125 full-sib families (Table 1; Nilsson et al., 2010) and the estimated improvement in terms of weight gain in each generation was approximately 11% (Carlberg et al., 2018).

Table 1. *Number of sires and dams used in each generation of the Arctic charr breeding programme and resulting full-sib families*

Generation	Year class	No. of sires	No. of dams	Full sib family groups	References
1st	1985	36	32	95	Nilsson et al, 2010
2nd	1991	25	51	51	Nilsson et al, 2010
3rd	1996	22	47	47	Nilsson et al, 2010
4th	2000	33	48	48	Nilsson et al, 2010
5th	2005	40	49	49	Nilsson et al, 2010
6th	2009	88	125	125	Nilsson et al, 2010
7th	2013	34	49	49	Carlberg et al. 2018
8th	2017	40	55	55	Personal communication

Overall, after almost 40 years since the Arctic charr breeding program was initiated in Sweden remarkable improvements regarding growth, feed utilization, late maturation, colouration of flesh and significant reductions in terms of production costs were obtained (Brännäs et al., 2011; L.-O. Eriksson et al., 1993; Nilsson et al., 2010). Nowadays, the production cycle required to reach the market size (600-900 g carcass weight) is estimated to be approximately 1.5 years, while the overall production costs are reduced by approximately 40% (Table 2) (L.-O. Eriksson et al., 1993; Nilsson et al., 2010). The selected strain of Arctic charr was named as "Arctic superior" and currently it is highly appreciated by the farmers.

Table 2. *Comparison of Lake Hornavan Arctic charr (1985) with selectively bred "Arctic superior strain" (2010)*

Lake Hornavan Arctic charr (1985)	**Arctic superior-strain (2010)**
Farming cycle 3-4 years	Farming cycle 1.5 -2 years
Relatively slow growth, poor feed utilization	Fast growth, efficient feed utilization
Early maturation (70 -100% before 500 g)	Late maturation (< 5% before 800 g)
Variable fillet colouration and quality	Improve flesh colouration and quality.
Production cost 50 -60 SEK (4.3-5.6 Euro)	Production cost 25-35 SEK (3.3 Euro)

1.3. Challenges

A main challenge in the Arctic charr breeding program is the low reproductive success which is quite far from a satisfactory level. In the study of Jeuthe (2012), the observed gamete quality was very low with high frequency of opaque eggs and watery sperm observed in various broodfish from the national Swedish breeding program. Furthermore, the spawning period in the selectively bred strain of Arctic charr is often prolonged and in many occasions poorly synchronized between dams and sires, leading to increased stress levels for the animals due to repeated handling.

Overall, the successful hatching rate for Arctic charr in Sweden has been ranging between 30 to 70% (L. O. Eriksson et al., 2010; Jeuthe, 2012; Pickova & Brännäs, 2006) which is considered low compared to other farmed salmonids. According to Nilsson et al., (2010) 30-50% of all fertilized eggs survive to the eyed stage at the breeding station in Kälarne. On the contrary in other salmonids farmed across Scandinavia like Atlantic salmon, brown trout and rainbow trout (*Oncorhynchus mykiss*) the egg survival rate is usually above 90% (Jeuthe, 2012).

In addition, due to the poor egg survival a large number of broodfish are needed in order to accomplish the egg requirements for the aquaculture industry resulting in high economic costs (Jeuthe, 2012).

1.4. Biological and environmental factors affecting reproductive success

1.4.1. Temperature

Arctic charr can tolerate different temperatures depending on its life stage. A critical limit in terms of incubation temperature for Arctic charr eggs is 8 °C, while in trout and salmon the corresponding limit is 13 °C and 16 °C, respectively (Elliott & Elliott, 2010; Jeuthe, 2012). More than 80 % mortality was observed when Arctic charr eggs were incubated at 11^0C (Jungwirth & Winkler, 1984). Furthermore, egg mortality at temperatures between 12-13 °C approaches 100% (Jungwirth & Winkler, 1984). Apart from the high mortality during egg incubation suboptimal temperatures could also cause deformities. Increased rates of jaw and vertebral column deformities were observed in various fish like halibut (*Hippoglossus hippoglossus*) (Bolla & Holmefjord, 1988), goldfish (*Carassius auratus*) (Wiegand et al., 1989), Atlantic salmon (Eriksen et al., 2006) and cod (*Gadus morhua)* (Fitzsimmons & Perutz, 2006) when their eggs where incubated at suboptimal temperatures. However, these type of malformations have not yet been observed in the selected strain of Arctic charr (Jeuthe, 2012).

During the parr and later life stages of Arctic charr a water temperature of 12 °C is favourable both for optimal growth and for welfare reasons(Jeuthe, 2012; Larsson & Berglund, 1998). A

temperature range between 19-23 °C during 7 consecutive days resulted in approximately 50% mortality in alevins, parr and fry stages (Elliott & Klemetsen, 2002; Lyytikäinen et al., 1997).

Furthermore, holding broodstock on suboptimal temperatures especially during the final 6 six months prior to spawning is a factor which can impair the development of gametes (King et al., 2003) and result in poor egg survival in Arctic charr (L. O. Eriksson et al., 2010; Jobling et al., 1998; Pickova & Brännäs, 2006). Additionally, it has also been observed that elevated temperatures could interrupt the oogenesis and oocyte maturation processes (Jeuthe, 2012; King et al., 2003).

1.4.2. Correlation between egg size and egg survival

In some fish species there is a relationship between egg survival and egg size. Smaller egg size has been associated with low egg survival during incubation at elevated temperatures in the case of Atlantic salmon (King et al., 2003). Moreover, there seems to be a positive correlation between survival and egg size in brown trout (*Salmo trutta*)(Bagenal, 1969). On the other hand, fingerling mortality increased with egg size in Chinook salmon (*Oncorhynchus tshawytscha*) (Fowler, 1972). According to the Jónsson and Svavarsson (2000) there is no correlation between egg size and early mortality in Arctic charr. Nevertheless, a recent study revealed that the egg size could be linked with egg survival in Arctic charr (Jeuthe et al., 2013). Moreover, egg survival and egg size both increased with female age up to 6 years, while no further improvements were observed after 6 years of age (Jeuthe et al., 2013).

1.4.3. Dietary fatty acid effect on egg quality

Marine and freshwater diets have significantly different nutrient contents especially in terms of their lipid composition (Kaitaranta & Linko, 1984; Pickova & Brännäs, 2006). The ratio of n-6/ n-3 FA (fatty acid) is quite higher in freshwater based diet (Pickova et al., 2007). Prior studies observed that a low level of Arachidonic acid (20:4 n-6) in the marine water-based diet resulted in low egg survival in farmed Arctic charr (Pickova et al., 2007; Torstensen et al., 2001). In addition, gamete quality in Arctic charr sperm was found to be strongly correlated with FA composition. Moreover, Arctic charr spermatozoa with high fertilization capacity had higher n-3/ n-6 FA ratio and lower levels of short-chained fatty acids (Mansour et al., 2011).

1.4.4. Spawning period

Environmental factors with a major effect on spawning time include the water temperature and the photoperiod (day length). In Sweden the Arctic charr gametogenesis usually starts in spring and accelerates during late summer when the day length starts decreasing. The spawning period starts in late autumn and is associated with a drop in the water temperature. Ovulation could be induced by transferring broodstock to cold water. Photoperiod manipulation could be a valuable management tool to delay spawning until for example the water temperature reaches a suitable range (Jeuthe, 2012).

1.4.5. Inbreeding and genetic variability

Inbreeding increase is inevitable in every animal population system especially in the case of closed populations. Inbreeding depression could amongst other traits affect the reproductive performance of the farmed animals. Since the start of the Arctic charr breeding program in Sweden the

cumulative inbreeding increase was approximately 5% (Personal communication). In prior studies, no correlation was found between the individual female parent inbreeding level and the percentage of egg survival to the eyed stage (Nilsson et al., 2010). Moreover, egg survival of full sib groups and the corresponding amount of inbreeding also had a non-significant association. Finally, the number of eggs per female and egg size exhibited non-significant relationship with inbreeding (Nilsson et al., 2010).

1.5. Types of genetic markers

Microsatellites or simple sequence repeats (SSRs) are repetitive DNA motifs ranging in length from 1-6 or more base pairs. Microsatellites typically span between a few hundred or even thousand bases across an organism's genome. In aquaculture genetics microsatellites are valuable for investigating genetic diversity, parentage assignment, construction of linkage map and QTL mapping (Chistiakov et al., 2006). A wide variety of genetic diversity studies using microsatellites has been done in Atlantic salmon (Slettan et al., 1997) catfish (Zhanjiang Liu et al., 2001, 2003; Tan et al., 1999), common carp (Tanck et al., 2001), tilapias (Cnaani et al., 2002), chinook salmon (Naish & Park, 2002), rainbow trout (Rexroad et al., 2001, 2002a, 2002b) and spotted steeds (*Hemibarbus maculates*) (Zhu et al., 2011).

Amplified fragment length polymorphism (AFLP) and restriction fragment length polymorphism (RFLP) comprise another category of genetic markers that has been widely used. AFLPs have been applied in several aquaculture species in population genetic studies in rainbow trout (Young et al., 1998), channel catfish and blue catfish (Liu et al., 1999; Liu et al., 1998), ayu (*Plecoglossus altivelis)* (Seki et al., 1999), channel catfish (*Ictalurus punctatus*)(Liu et al., 1998; Mickett et al.,

2003), common carp (Z. Wang et al., 2000), oysters (Li & Guo, 2004; Yu & Guo, 2003) and striped mullet (*Mugil cephalus)* (Liu et al., 2009).

Single nucleotide polymorphisms (SNP) are currently the genetic marker of choice in aquaculture genetics. A wide range of technologies have been used for SNP discovery. Some of those technologies were based upon direct polymerase chain reaction (PCR) product sequencing, random shotgun and expressed sequence tags (ESTs) (Liu & Cordes, 2004). Newest technologies for SNP discovery rely upon high throughput sequencing platforms (Abdelrahman et al., 2017; Kumar & Kocour, 2017). According to literature, SNPs have been used amongst others for finding candidate genes of desired traits, for constructing linkage maps, performing parentage assignment, finding quantitative trait loci (QTL) for desired traits, identification of broodstocks and (Oyarzún et al., 2013; Yáñez et al., 2015). The advantage of SNPs over the aforementioned categories of genetic markers (microsatellites, AFLPs, RFLPs) lies in their abundance in an organism's genome and their easiness towards being implemented in the form of large-scale automated genotyping platforms (e.g. SNP arrays). Finally, SNPs are the markers of choice for performing genome wide association studies (GWAS) and genomic selection (GS) in both plants and animals.

1.6. Reduced-representation sequencing platforms and applications in aquaculture

The rapid drop of sequencing costs with the advent of next generation sequencing has provided many opportunities for aquaculture genetics and breeding. Next generation sequencing technologies have been commonly utilized for reference genome assembly studies, for discovering SNPs and for performing GWAS (Metzker, 2010). Reduced-representation sequencing approaches offer a cost-effective alternative to whole genome sequencing allowing the detection of genome-

wide genetic markers (Davey et al., 2011; Robledo et al., 2018). Their main premise involves the usage of restriction enzymes to reduce the genome complexity and thereafter sequence only the fragments resulting from the enzymatic digestion (Davey et al., 2011). Popular methods of reduced-representation sequencing (Table 3) employed in aquaculture species include restriction-site associated DNA sequencing (RAD-seq; Baird et al. 2008), genotyping by sequencing (GBS; Elshire et al., 2011), ddRAD-seq (Peterson et al., 2012), 2b-RAD (Wang et al., 2012), ezRAD (Toonen et al., 2013), quaddRAD (Franchini et al., 2017) and 2RAD/3RAD (Bayona-Vásquez et al., 2019).

Table 3. *Reduced-representation sequencing applications in aquaculture species.*

Study	Species	Aim	Technique
Gonen et al., 2014	*Salmo salar*	Linkage map	RAD
Palaiokostas et al., 2015	*Oreochromis niloticus*	Sex determination QTL	ddRAD
Larson et al., 2016	*Oncorhynchus nerka*	Thermotolerance and growth QTL	RAD
Palaiokostas et al., 2016	*Sparus aurata*	Disease resistance genomic selection	2bRAD
Figueras et al., 2016	*Scophthalmus maximus*	Genomic growth and resistance to patholgens	WGS
Holman et al., 2017	*Salmo salar*	Parentage assignment	RAD
Palaiokostas, et al., 2018a	*Dicentrarchus labrax*	Resistance to viral nervous necrosis	RAD
Palaiokostas, et al., 2018b	*Cyprinus carpio*	Accuracy in juvenile growth rate	RAD
Kyriakis et al., 2019	*Sparus aurata*	Genetic mapping of phenotypic trait	ddRAD
Syaifudin et al., 2019	*Tilapia*	Species specific marker discovery	ddRAD
Moses et al., 2020	*Oreochromis niloticus*	Genetic structure characterization	ddRAD
Huang et al., 2020	*Ctenopharyngodon idellus*	Mapping of growth-related QTLs	2bRAD

1.6.1. RAD sequencing

In RAD-seq genomic DNA from multiple individuals is digested with a restriction enzyme of choice. Adaptors with unique identifiers are usually ligated to the resulting fragments. Thereafter the DNA product is selected for a size most suitable for Illumina sequencing (300-700 bp), amplified through PCR and finally sequenced in an NGS platform of choice producing partial but genome-wide coverage at a fraction of the cost of whole-genome sequencing. In general, the most popular sequencing platforms originate from Illumina. The above sequencing technology involves sequencing one (read 1, single end) or both (reads 1 and 2, paired end) ends of each DNA fragment, typically producing reads between 75-300 bp long. The obtained RAD tags create a reduced representation of the genome, allowing over-sequencing of the nucleotides next to restriction sites and detection of SNPs with high accuracy.

2b-RAD sequencing

2b-RAD is a modified version that is based on the original RAD technique (Wang et al., 2012). 2b-RAD is based on uniform fragmentation through usage of IIB restriction endonucleases. IIB restriction enzymes like BsaXI and A1fI cleave upstream and downstream of genomic DNA recognition motifs. As in RAD-seq during library preparation adapters are ligated after the DNA digestion process and specified barcodes are ligated to the identical sized fragments. Thereafter the template DNA is amplified through PCR. Samples are usually pooled in a single library and sequenced in Illumina platforms. A main advantage of 2b-RAD is that error-prone steps like size selection and the interim purification steps are eliminated therefore simplifying considerably the needed library preparation workflow. Nevertheless, the main caveat of 2b-RAD is that it produces only short reads ranging between 33-36 bp which is less amenable for reference genome

assemblies. Moreover, 2b-RAD is not ideal for the design of SNP assays due to the limited available sequence length that flanks the SNP of interest. This issue is not a big deal for species that have a draft genome available but could become a problem for many aquaculture species where the latter is not available.

ezRAD sequencing

ezRAD sequencing was introduced by Toonen et al., 2013 with its broad concept being similar with the other RAD methods, such as ddRAD and 2bRAD (Peterson et al., 2012; Wang et al., 2012). ezRAD is compatible with a wide range of restriction enzyme or even combination of enzymes. Furthermore, ezRAD does not require purchasing new adaptors for each restriction enzyme. After digestion the downstream sonication step to sheer DNA that is used in the original RAD-seq protocol is eliminated. ezRAD is also flexible for optimizing the number of unique fragments through size selection which is usually performed with agarose gel electrophoresis or SPRI- beads. Library preparation utilizes the Illumina TrueSeq library preparation kit which make it possible for sequencing through any commercial genome core facility. The cost per library preparation is ~$60 (Toonen et al., 2013).

quaddRAD (double digest RAD paired end)

QuaddRAD (ddRAD paired end) protocol is a newer addition in the plethora of reduced-representation sequencing protocols (Franchini et al., 2017). The aim of quaddRAD was to eliminate PCR duplicates by embracing 4-base stretches at the sequencing distal region of each Illumina adapter. Size selection is usually performed on a single pool of hundreds of individuals in the same gel lane in order to increase the number of overlapping loci among individuals and prevent the random between lane size variation. To maximize the accuracy, size selection is

performed through an automated machine (Pippin Prep, Sage sciences). Finally, in the quaddRAD protocol digestion and ligation are performed in a single reaction which reduces hands on time.

Double digest restriction associated DNA (ddRAD) sequencing

ddRAD (digestion with 2 different REs) sequencing is an efficient, robust, flexible and cost-effective platform for identifying genome-wide genetic markers (Peterson et al., 2012; Robledo et al., 2018). In comparison to RAD-seq the step of random shearing and the end repair step of the template DNA are eliminated. ddRAD sequencing allows the construction of highly multiplexed libraries consisting on many individuals due to the lower demands of overall sequencing effort as opposed to RAD-seq (Peterson et al., 2012; Robledo et al., 2018). In ddRAD the library preparation costs are reduced ~$5 per sample while in the corresponding RAD sequencing protocol the costs are ~$25 per sample (Peterson et al., 2012). Furthermore, 100ng or less of starting DNA template is enough for ddRAD library construction due to elimination of high-DNA-loss steps. In addition, the recovered library is obtained as a result of a precise size selection of genomic fragments that excludes the regions flanked either very close or very distant to RE recognition sites. This precise, repeatable size selection has further advantages; firstly, a small fraction of restriction fragments approximately <5% will fall in the target size-selection window which reduces duplicate region sampling (the probability of sampling both direction from the same restriction site is low). Secondly, it results in fragments of more uniform length. These properties make it robust and minimizes unequal read representation especially in pooled sequencing experiments (Craig et al., 2008; Alon et al., 2011; Andolfatto et al., 2011; Peterson et al., 2012; Brown et al., 2016).

2. Aims and objectives

The aim of this study was to investigate the genetic diversity of Arctic charr 'Superior' across farms in Sweden and to gain insight regarding the genetic factors that are associated with sperm quality characteristics (concentration; viable cells). The acquired knowledge will be of value towards a better understanding of genetic factors involved with reproductive success in farmed Arctic charr in Sweden.

Specifically, the main objectives of the dissertation were:

- To analyze the genetic structure and genetic diversity within and among population of Arctic charr.
- To discover genetic polymorphisms associated with sperm quality characteristics in Arctic charr like sperm concentration and viabily.

3. Material and Method

3.1. Sampling and Data collection:

The entire study was in accordance with the Swedish legislation for conducting animal research as described in the Animal Welfare Act 2018:1192 (ethic permit: 5.2.18 09859/2019). Sperm quality metrics were recorder on 150 animals originating from four Arctic charr farms (Farm-A, Farm-B, Farm-C and Farm-D) across Sweden. Sperm concentration and number of viable sperm cells were estimated using the NucleoCounter SP-100 (Figure 1). In addition, tail fin clips were collected from those animals and preserved in 100% ethanol. The above workflow was performed during September – October 2019. A small piece of the preserved fin clip (approximately 3 mm^2) was used for DNA extraction.

Figure 1 *"NucleoCounter SP-100" used for counting sperm quality characteristics.*

3.2. Preparation of genomic DNA

3.2.1. DNA extraction process

- Excess ethanol was evaporated from the fin samples with the usage of heating plate.
- Took 300µl SSTNE buffer, 30µl 10 % SDS and 5µl Proteinase K (10mg/ml) for each tissue sample and incubated them for 3 hours at 55°C.
- The samples were incubated at 70°C for 15 minutes in order to inactivate the proteinase K.
- Cooled the samples at 37 °C. Took 5µl Ribonuclease A (2mg/ml) for each sample and mixed them well. Incubated the samples further for 60 minutes at 37°C with end-over-end mixing.
- Cooled the samples at room temperature. Took 240 µl (5M) NaCl, mixed it well. Left the samples on ice for 10 minutes in order to precipitates the protein. Centrifuge the samples at high speed about 21,000 g for 5 minutes to spin down precipitated proteins.
- Carefully took 400µl upper aqueous phase (supernatant) and put it into a fresh nuclease-free 1.7ml microcentrifuge tube.
- Took an equal volume (400 µl) of isopropanol. Mixed it by 5-6 sharp over end inversions and leave it on ice for 5 minutes.
- DNA precipitate was clearly visible. Centrifuge it at high speed about 14,000g for 2 minutes to make DNA pellet.
- Carefully pour off supernatant, briefly spin and removed all excess isopropanol through pipette.

- Took 600 µl (75%) ethanol (the longer the wash the more salt impurities leech out – improving will be 260/230 ratio). Mixed by 10-15 sharp over end inversions. Centrifuge it at 16,000 g for 1-2 minutes and leave it overnight.
- Briefly centrifuge the samples to make the DNA pallet at the bottom. Removed the supernatant and air dried the DNA pallet for 1h, or 60°C heat block for 5 minutes to evaporate all traces of ethanol.
- Re-suspended the DNA pallet in appropriate (30µl) EB Buffer. Incubated for 10 minutes at 55°C and leave overnight to completely dissolve.

3.2.2. Quantification of genomic DNA

Total nucleic acid content and quality (260nm/ 280nm and 260 nm/ 230nm ratios) were determined by spectrometry (Nanodrop 8000; Thermo Scientific) before adjusting nucleic acid concentrations to 50ng/ml with EB buffer. Thereafter DNA concentration was measured with higher accuracy using the Qubit dsDNA Broad Range Assay Kit. The samples were diluted to a concentration of 15ng /ml using EB buffer. The quality of the genomic DNA for ddRAD library construction was assessed by randomly selecting 14 samples for gel electrophoresis.

3.2.3. Gel Electrophoresis:

Preparation of Agarose Gel:

- Measured 0.6 g of agarose. Agarose gels are commonly used in concentrations of 0.7% to 2% depending on the size of bands needed to be separated. For this took 0.6 g of agarose in 40 mL of 1xTAE to make 1.5% gel.

- Mixed agarose powder with 1xTAE in a microwavable flask.
- Microwave for 1-2 min until the agarose completely dissolved. Do not overboil the solution, as some of the buffer will evaporate and thus alter the final percentage of agarose in the gel. Prefer to microwave in pulses and swirling the flask occasionally as the solution heats up.
- Let the agarose solution cooldown about 8-10 mins to 50 °C. On this temperature comfortably keep the hand on the flask.
- Add 4 µl of SYBR Safe (cyanine dye). SYBR Safe binds to the DNA resulting in DNA-dye-complex that absorbs blue light and emits green light.
- Pour the agarose solution into a gel tray with the well comb in place. Pour agarose solution slowly to avoid formation of bubbles which will disrupt the gel. In case of any bubbles produce in the well away with the pipette tip.
- Let sit at room temperature for 20 to 30 mins, until its completely solidified.

Run the Agarose Gel:

- Took 1 µl DNA sample each from randomly selected 14 DNA samples.
- Put 1 µl gel loading dye, purple (6X) and 4 µl water in each tube.
- Load the ladder and DNA samples on gel and run the gel at 90V for 40 minutes.
- The quality of gDNA was approvable (Figure 2)

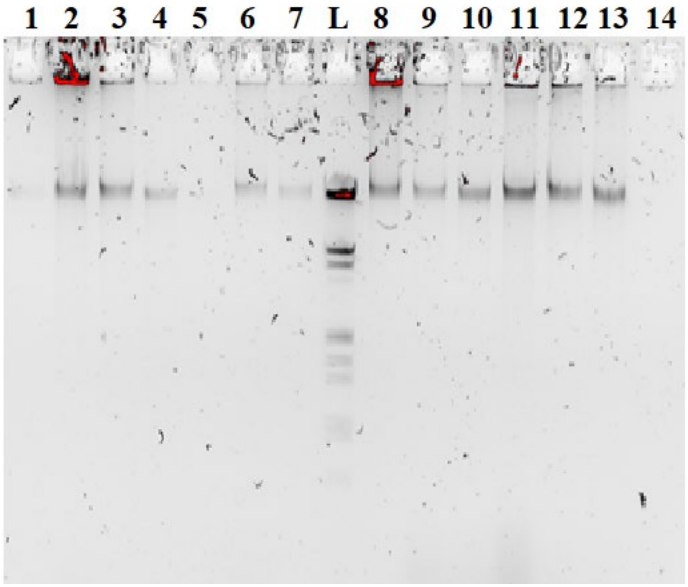


Figure 2 *"DNA" run on Agrose gel to appraise the quality of genomic DNA.*

3.3. ddRAD library preparation and sequencing

We followed the procedure for ddRAD library preparation as described by Peterson et al., 2012 with minor modifications described in Palaiokostas et al., 2015.

3.3.1. Double Digest

Took 1µl (15ng) of genomic DNA (each sample) with *SbfI* (recognizing the CCTGCA|GG motif) and *SphI* (recognizing the GCTAG|C motif) high fidelity restriction enzyme, 10 U of each enzyme per microgram of genomic DNA in 1x CutSmart buffer. The reaction volume made was 6 µl for each sample. Mixed the plate well and incubated samples for 1h at 37 °C for the digestion.

3.3.2. Adapters ligation to DNA fragments

Individual specific combinations of P1 and P2 adapters, each with a unique 5 or 7 bp barcode, ligated to the digested DNA at 22 °C for 10 min by adding 1µl *SbfI* compatible P1 adapter (25nM), 0.7 µl *SphI* compatible P2 adapter (100nM). The reaction volume was 3 µl for each sample.

The reaction was incubated for 2h with 0.12µl 1mM/L rATP, 0.3 µl 1x CutSmart buffer, 0.015µl T4 Ligase (2K ceU/µg (2000 cohesive-end Units per µg)) and 2.565 µl nuclease-free water for each sample. The reaction volume made 3µl for each sample. Following heat inactivation at 65 °C for 20 minutes (Note: Ligase activity was inactivated by the addition of 30 µl PB buffer). Ligation reactions slowly cooled at room temperature approximately over 1h. Combined all sample in a single pool for one sequencing lane. Overall, the total reaction volume was 12µl per sample.

3.3.3. MinElute PCR purification

PCR purification was performed on the pooled library.

- Added 3 volumes of Buffer PB to 1 volume of the pooled DNA template (5400µl PB buffer in total) and mixed it well. Check that the colour of the mixture is yellow (similar to the Buffer PB without the PCR sample). If the colour of the mixture is orange or violet, add 3 µl 3 M sodium acetate, pH 5.0 (Note: it is use for pH adjustment) and mixed it well. The color of the mixture turned to yellow.
- Placed a MinElute column in a provided 2ml collection tube or in a vacuum manifold.
- Applied the pooled library to the MinElute column and leave for 5 minutes to stable the DNA. Centrifuge for 15 seconds to pass the entire sample through the column (Note: 600

µl pooled library applied for each time due to the small space of the tube). Discard flow-through and place the MinElute column back in the same collection tube.

- Added 720 µl Buffer PE to the MinElute column left for 5 minutes then centrifuged for 30sec. Discarded flow-through and place the MinElute column back into same collection tube.
- Opened the column lid (Note: reason to remove excess ethanol) and centrifuged in a 2ml collection tube for 1 minute.
- Placed MinElute column in a clean 1.5ml centrifuge tube and put it on heat box for few seconds with the lid open to evaporate the excess ethanol.
- To elute DNA, added 17 µl Bffer EB (10Mm Tris.Cl, pH 8.5) (Note: before use the buffer put it on the heat box at 55 °C. Its works well on this temperature.) Or water to the center of the MinElute membrane (Ensure that the elution buffer is dispensed directly on to the center of the membrane for complete elution of bound DNA). Column was left for 5 minutes and then centrifuged for 15 seconds.
- To ensure the DNA released from the MinElute membrane added an extra 17 µl Buffer EB on the center of the MinElute membrane and centrifuged it for 1 minute.

Note: all centrifugation steps were carried out at 17,000 g.

3.3.4. Size selection and library formation

Size selection was performed on an agarose gel. Prepared 1.1 % agarose gel (1.875 g of agarose; 150 µl of 1X TAE buffer) with the same method described above and run the DNA on the gel.

Run the gel:

- Prepared twice tubes for each DNA sample.
- Took 3 µl gel loading dye, purple (6X) and 15 µl DNA sample in each tube.
- Run the gel on the ice tray to put the temperature low.
- The gel was run at constant voltage of 30 V for 3 mins, then gradually increase to 45 V, 60 V, 75 V for 3 mins each and at last 90 V for 40 mins until the purple dye front travelled approximately 3cm from the origin.

Excise the DNA fragments

Agarose gel was put on a UV transilluminator (Figure 3). Excised 400 to 600 bp with a clean, sharp scalpel and collected them in the tubes (Figure 4). Checked the weight of excised gel slice before processing through a MinElute column for Gel cleaning. The mean weight was 160 mg.

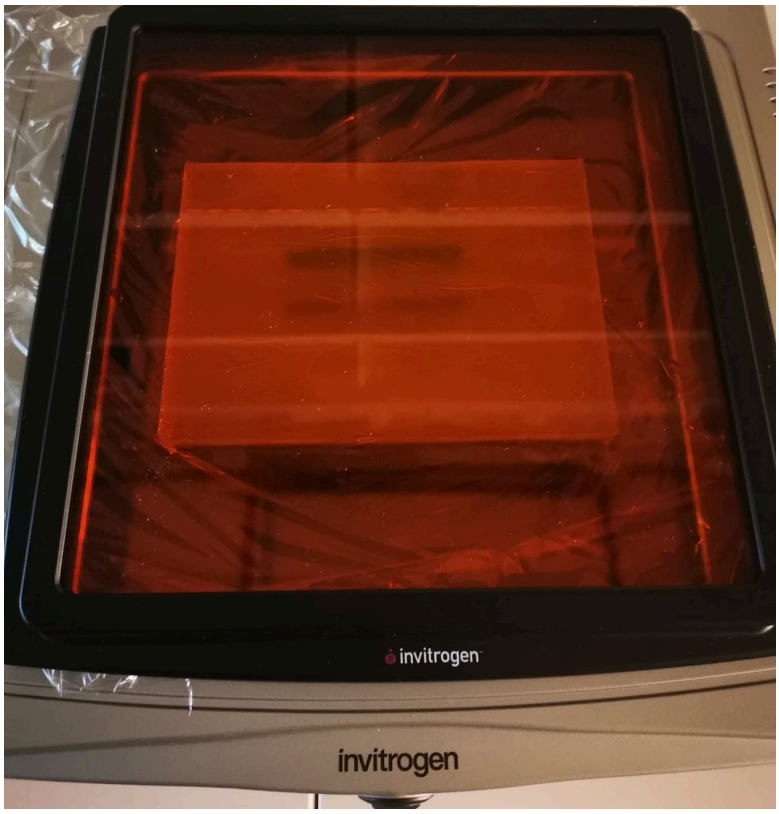


Figure 3 *" DNA fragments" observation through UV Transilluminator.*

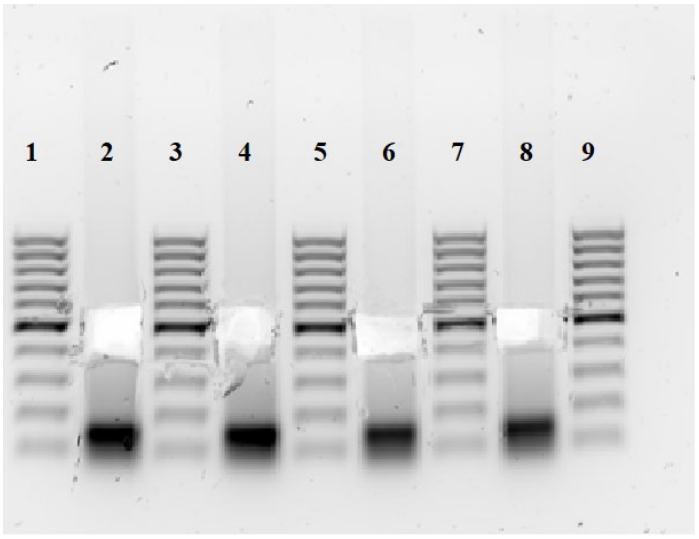


Figure 4 *" Excised Fragments" size selection 400 -600 base pairs.*

MinElute Gel extraction

- Added 3 volumes of Buffer QG to 1 volume of excised gel (160mg x3 = 480 ~ 500µl QG Buffer).
- Incubated the gel slice at 55 °C for 15 mins (or until the gel slice has completely dissolved). During incubation, vortexed the tube every 1-2 mins for proper dissolving.
- After dissolving the gel slice, added 10 µl of 3M sodium acetate to maintain the ph 5.0.
- Added 160 µl (1 gel volume) isopropanol to each gel solution. For proper mixing inverted the tube several times.
- Placed the MinElute column in 2ml collection tube in a suitable rack.
- Applied the gel solution to the MinElute column to bind the DNA to the membrane. Left for 5 minutes then centrifuged for 15 seconds to pass the entire gel solution through the column.
- Discard flow-through and put the MinElute column back into the same collection tube.
- Add 500 µl Buffer QG to the MinElute column and leave for 1 min then centrifuge for 15 sec. Discard flow-through and place the MinElute column back to the same collection tube**.**
- Added 720 µl Buffer PE to the MinElute column, left for 5 minutes then centrifuged for 1 min. Discarded flow-through and placde the MinElute column back into same collection tube.
- Opened the column lid and centrifuged for 1 min.

- Placed each MinElute column in a clean 1.5ml centrifuge tube and put it on heat box for few seconds by open the lid to evaporate the excess ethanol.
- Added 20 µl Bffer EB (10Mm Tris.Cl, pH 8.5) to the center of the MinElute membrane. Centrifuged for 15 seconds.
- Added again 20 µl Buffer EB on the center of the MinElute membrane and centrifuged for 1 minute.
- Collected both samples in a single tube. Note: all centrifugation steps were carried out at 17,000 g.

3.3.5. Quality inspection prior to illumine sequencing

Need to perform few tests for checking the quality of the prepared library.

Amplification test

- Combined 5.5 µl water, 6.25µl Phusion High-Fidelity Master mix, 0.25 µl RAD amplification primer mix (10µM) and 2µl RAD library template. Performed 18 cycles of amplification in a thermal cycler: 30 sec 98°C, 18× (10 sec 98°C, 30 sec 65°C, 30 sec 72°C), 5 min 72°C, hold 4°C. Same concentrations were used for 16 and 14 cycles. Adjusted the thermal cycler: 30 sec 98°C, 16× (10 sec 98°C, 30 sec 65°C, 30 sec 72°C), 5 min 72°C, hold 4°C for 16 cycles and 30 sec 98°C, 14× (10 sec 98°C, 30 sec 65°C, 30 sec 72°C), 5 min 72°C, hold 4°C for 14 cycles. Took the 5µl PCR product (from each 14, 16 and 18 cycles) with 1µl purple (6x) dye. Loaded the sample on 1.5 % agarose gel. Negative control and library template was also load on the gel. Run the gel at 90V for 40 minutes (Figure 5).

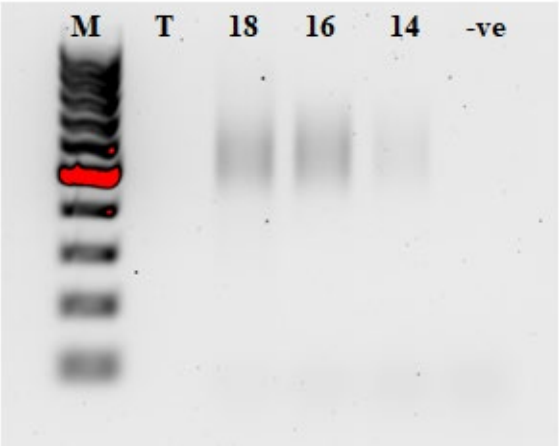


Figure 5 *" Test amplification PCR product"*

- For performing final PCR amplification, prepared the master mix for 48 samples to minimize the effect of PCR duplicates; 264 µl water, 300µl Phusion High-Fidelity Master mix, 12 µl RAD amplification primer mix (10µM) and 24 µl RAD library template and aliquots in 48 wells of PCR plate. Run the PCR reaction for 13 cycles. The thermal cycler was adjusted: 30 sec 98°C, 13× (10 sec 98°C, 30 sec 65°C, 30 sec 72°C), 5 min 72°C, hold 12°C.

- Mixed the plate properly and pooled the whole PCR product in a single tube. Added 18µl buffer BP in the PCR product and mixed it well. Add 3 µl 3 M sodium acetate. Applied the PCR product to the MinElute column for purification. Eluted the DNA with 54µl of EB buffer.

Purification through AMPure magnetic beads

- Used an equal volume of Elute EB (54µl ~ 50µl) of AMPure magnetic beads in the eluted DNA. Mixed it well by pipetting up and down at least 10 times and incubated for 10 minutes. Added 300µl 70% ethanol and waited for 30 seconds (the function of the ethanol is to stable the magnetic beads). The tube was put in the magnetic chamber and waited till all the magnetic beats are attached on the chamber wall. Kept the tube to the magnetic chamber and carefully discarded the clear solution from the tube through pipette without touching the beads. Added again 300µl 70 % ethanol waited for 30 seconds and discarded the ethanol. Waited for a couple of minutes to dry the beads. Eluted the DNA with 30µl elution EB buffer in the magnetic beads and mixed it well by pipetting up and down at least 8-10 times. Briefly centrifuged to consolidate the sample which was placed again on the magnetic camber for 3 mins until the solution became clear. Removed 30µl supernatant and transfered to fresh tube and discarded the beads.

DNA Quality inspection

- Run 1µl sample with 1µl dye and 4µl water on agarose gel at 90V for 40 mins. (Figure 6).

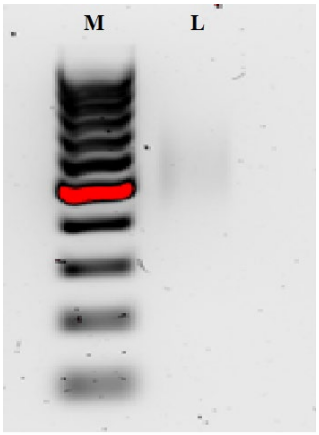


Figure 6 *"Quality checking DNA "*

- Concentration of final PCR product was 6ng /ml measured through Qubit dsDNA high sensitivity Assay kit.

3.4. Illumina sequencing

Illumina sequencing was performed at the National Genomics Infrastructure in Uppsala. Sequencing was performed on a NovaSeq6000 on one lane of a SP flowcell using the 300-cycle sequencing kit for paired-end sequencing. Sequencing was performed by including 5% phiX together with the library due to the low-plexity characteristics of reduced-representation libraries.

3.5. Descriptive analysis and ANOVA

General statistical analysis was performed in order to analyze differences on traits (sperm concentration, number of dead cells and percentage of viable cells) among different populations of Arctic charr. Descriptive analysis and ANOVA (Analysis of Variance) was performed using the R packages pastecs 1.3.21 (Grosjean et al., 2018) and tidyverse 1.3.0. (Grolemund & Wickham, 2016).

3.6. Sequence Data Analysis and SNP detection

Reads of low quality (Phred quality score < 30; Figure 7) and missing the expected restriction sites were discarded. The retained reads were aligned to the publicly available *Salvelinus sp.* genome assembly (Genbank accession number GCF_022910315.2; Christensen et al., 2018) using bowtie2 (Langmead & Salzberg, 2012). Stacks v2.5 (Rochette et al., 2019) was used to identify single nucleotide polymorphisms (SNPs). From each putative ddRAD locus only a single SNP was used for downstream analysis. SNPs with minor allele frequency (MAF) < 0.05 and maximum

heterozygosity > 0.7 across the tested samples were discarded. Finally, only SNPs found in at least 75 % of the samples in each population were retained for downstream analysis.

***Table 4**. Pherd quality scores logarithmically linked to error probabilities.*

Phred quality score	Probability of incorrect base call	Base call accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1000	99.9%
40	1 in 10,000	99.99%
50	1 in 100,000	99.999%
60	1 in 1,000,000	99.9999%

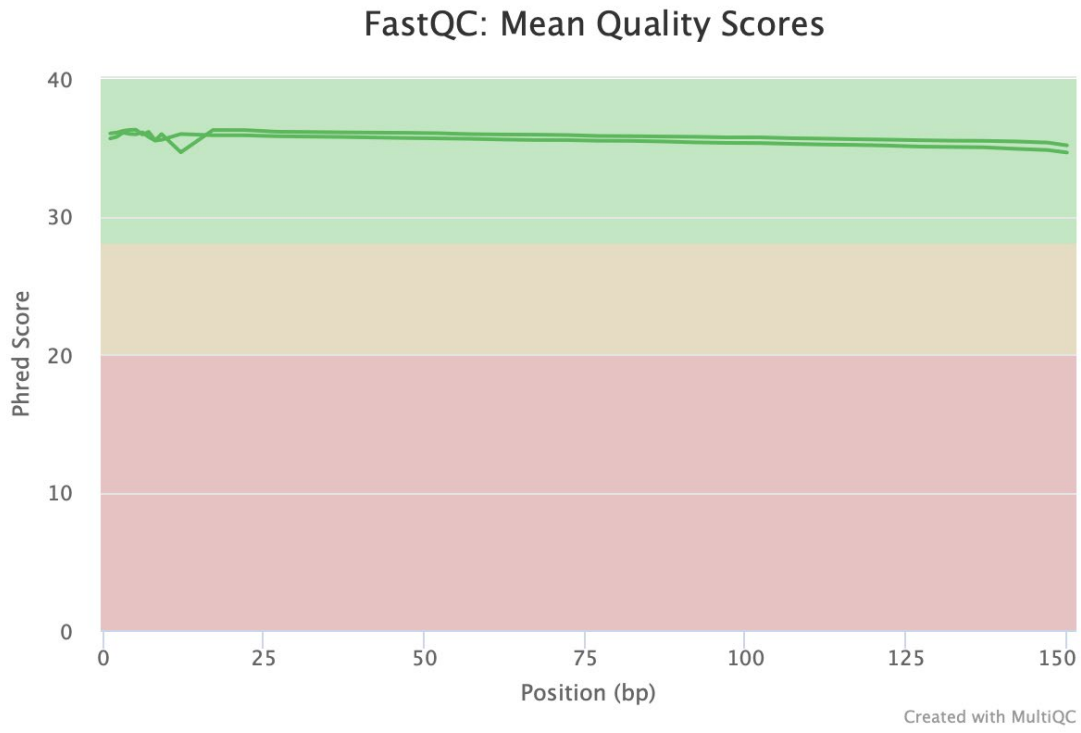


***Figure 7** "Phred Quality Score"*

3.7. Genetic diversity

Estimation of genetic diversity among the different tested populations was performed through the following analysis. Observed heterozygosity (H_o), expected heterozygosity (H_E) and inbreeding coefficient (F_{is}) were estimated by Stacks software version 2.5 (Rochette et al., 2019). Pairwise F_{st} estimates amongst the tested populations were obtained through the R package StAMPP (Pembleton et al., 2013). Finally, principal component analysis (PCA) was performed using the R adegenet 2.0.0 package (Jombart, 2008; Jombart & Collins, 2015). PCA was used to investigate for potential genetic structure within and amongst the tested populations.

3.8. Association analysis

The association analysis was performed using the R packages gaston 1.5.6 (Perdry et al., 2020) and qqman 0.1.4 (Turner, 2017). Association analysis was carried out aiming to identify SNPs associated with either sperm concentration or percentage of viable cells in Arctic charr. The Bonferroni criterion was applied to reduce the number of false positives. According to Bonferroni method the selected statistical significant level (P=0.05) is divided by the total number of genetic markers. Hence, the significance level (P value) for genome wide significance was set as $-\log_{10}(0.05 \text{ /number of genetic markers})$. The Bonferroni method is a highly conservative approach (Palaiokostas & Houston, 2017).

Sperm concentration association analysis

A linear mixed model was applied for investigating associations between the sperm concentration and the genetic markers

$$Y_f = X\alpha + S\beta + \omega + e$$

where $\mathbf{Y_f}$ is the vector of phenotypic records of sperm concentration, **α** is the vector of fixed effect (Farm's location), **β** is the fixed effect of the SNP genotype, **ω** is the random additive genetic effect; X and S are design matrices for α and β, respectively. $\boldsymbol{\omega} \sim \boldsymbol{N}\ \mathbf{(0,}\ \boldsymbol{G}\boldsymbol{\sigma}^2_{\mathbf{a}}\mathbf{)}$ where ***G*** is the genomic relationship matrix, $\boldsymbol{\sigma}^2_{\mathbf{a}}$ is the additive genetic variance and **e** is the vector of residuals.

Cell viability association analysis

A linear mixed model was applied for investigating associations between the percentage of sperm viability and the genetic markers

$$Y_v = X\alpha + S\beta + \omega + e$$

where $\mathbf{Y_v}$ is the vector of phenotypic records of sperm cell viability, **α** is the vector of fixed effect (Farm's location), **β** is the fixed effect of the SNP genotype, **ω** is the random additive genetic effect; X and S are design matrices for α and β, respectively. $\boldsymbol{\omega} \sim \boldsymbol{N}\ \mathbf{(0,}\ \boldsymbol{G}\boldsymbol{\sigma}^2_{\mathbf{a}}\mathbf{)}$ where ***G*** is the genomic relationship matrix, $\boldsymbol{\sigma}^2_{\mathbf{a}}$ is the additive genetic variance and **e** is the vector of residuals.

4. Results

4.1. Statistical analysis of sperm characteristics

Descriptive analysis and ANOVA (Analysis of Variance) was performed using the R packages pastecs 1.3.21 (Grosjean et al., 2018) and tidyverse 1.3.0. (Grolemund & Wickham, 2016) respectively. Sperm concentration, dead cells and percent viability traits were analyzed among the different populations of Arctic charr.

Sperm concentration

Wide variation was observed regarding sperm density among the different farms (Figure 8A). The mean value of sperm density was 3656, 3779, 2911 and 4727 million cells/ml in Farm-A, Farm-B, Farm-C and Farm-D respectively (Table 5). Sperm density was highest in Farm-D while the Farm- had the lowest sperm density. ANOVA indicated for significant variation in sperm density among the tested populations ($P < 0.001$**, Table 6).

Number of dead cell

Highly significant ($P < 0.001$***, Table 6) variation regarding the observed number of non-viable sperm cells was observed amongst the sampled populations. The mean value of dead cell was 350, 314, 285 and 478 million cells/ml in Farm-A, Farm-B, Farm-C and Farm-D respectively (Table 5). The lowest mean of dead cell was observed in Farm-C while, the Farm-D had highest mean (Figure 8B).

Cell viability

The mean value of sperm cell viability percentage was 89.8, 91.4, 90.5 and 89.7 in Farm-A, Farm-B, Farm-C and Farm-D respectively (Table 5, Figure 8C). The obtained results indicated towards a non-significant (P = 0.12, Table 6) effect for the origin of the sampled population.

Table 5. *Descriptive analysis of Sperm density, Dead cell and viability percent among different population of Arctic charr.*

Sperm density		***Mean***	***Median***	***Variance***	***Sd***	***Min***	***Max***	***Quartile (0.25)***	***Quartile (0.75)***
	A	*3656*	*3593*	*2879747*	*1697*	*646*	*8678*	*2297*	*4729*
	B	*3779*	*3088*	*5291851*	*2300*	*1343*	*10169*	*2252*	*4156*
	C	*2911*	*2378*	*5426867*	*2329*	*342*	*8013*	*1468*	*3701*
	D	*4727*	*4956*	*4963634*	*2227*	*1115*	*11461*	*3034*	*6099*
Dead cell		***Mean***	***Median***	***Variance***	***Sd***	***Min***	***Max***	***Quartile (0.25)***	***Quartile (0.75)***
	A	*350*	*357*	*21298*	*146*	*74*	*774*	*238*	*456*
	B	*314*	*254*	*30745*	*175*	*132*	*816*	*201*	*432*
	C	*285*	*175*	*60341*	*245*	*42*	*847*	*129*	*410*
	D	*478*	*481*	*49947*	*223*	*105*	*1037*	*305*	*599*
Percent viability		***Mean***	***Median***	***Variance***	***Sd***	***Min***	***Max***	***Quartile (0.25)***	***Quartile (0.75)***
	A	*89.8*	*90.3*	*11.2*	*3.3*	*67.9*	*95.7*	*89.0*	*91.3*
	B	*91.4*	*91.8*	*1.9*	*1.4*	*88.2*	*92.8*	*90.5*	*92.5*
	C	*90.5*	*90.6*	*2.9*	*1.7*	*87.1*	*93.3*	*89.6*	*91.4*
	D	*89.7*	*89.8*	*2.2*	*1.5*	*86.7*	*94.7*	*89.1*	*90.4*

Table 6. *ANOVA on sperm density, non-viable sperm cell and sperm cell viability among the different populations of Arctic charr.*

		Df	***Sum sq***	***Mean sq***	***F value***	***Pr(>F)***
Sperm density	*Farms*	*3*	*52709356*	*17569785*	*4.3*	*0.006 ***
	Residuals	*146*	*592774100*	*4060097*		
Dead cell	*Farms*	*3*	*737667*	*245889*	*6.9*	*0.0002 ****
	Residuals	*146*	*5183386*	*35503*		
Cell viability	*Farms*	*3*	*38.7*	*12.9*	*1.9*	*0.12*
	Residuals	*146*	*949.9*	*6.5*		

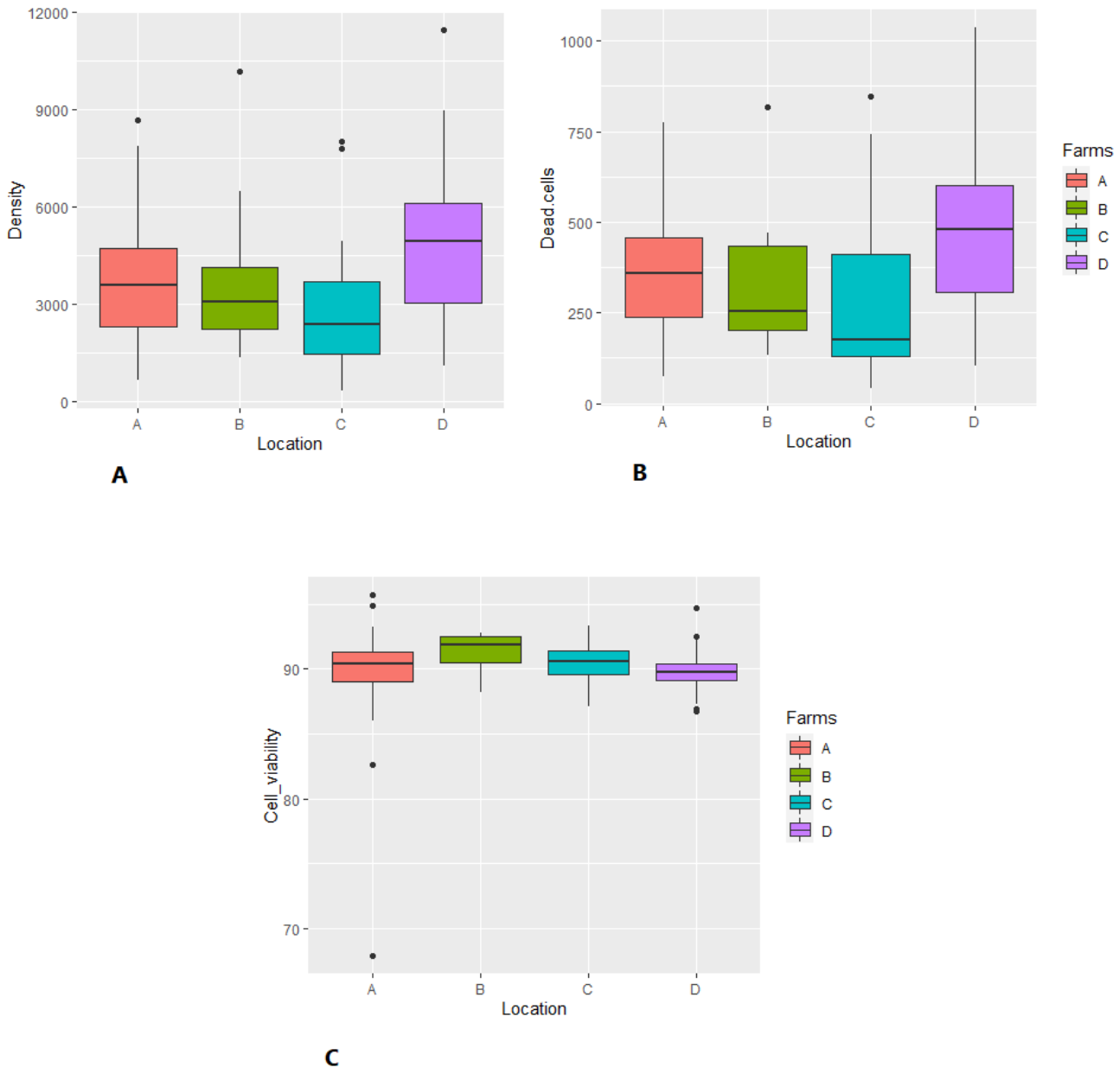


Figure 8 *" Evaluation of Sperm density (A), Non-viable cell (B) and viability percentage (C) among different locations" with outliers, extreme outlier, quartile and interquartile ranges.*

4.2. Sequence output – SNP identification

Approximately 915 million 150 bp long paired-end reads were obtained. During quality control filtering reads were discarded on the basis of low Phred quality score (~ 69 million reads; Figure 7), missing expected enzymatic induced cut site (~ 47 million reads) and missing expected barcode (~ 145 million reads). Overall, 71% of the obtained reads (~ 654 million reads) passed all quality control filtering steps and were retained. Thereafter 29,285 putative loci were identified. In total 5,159 SNPs were found in more than 75% of the genotyped animals and kept for downstream analysis. Finally, 17 animals were removed due to high percentage of missing data for the obtained SNPs (> 30%).

4.3. Distribution of Allele frequency

The distribution of allele frequency (Figure 4.2) was analyzed using the R package adegenet 2.0.0 (Jombart & Collins, 2015). The range of allele frequency was 0.05 - 0.95 (Figure 9).

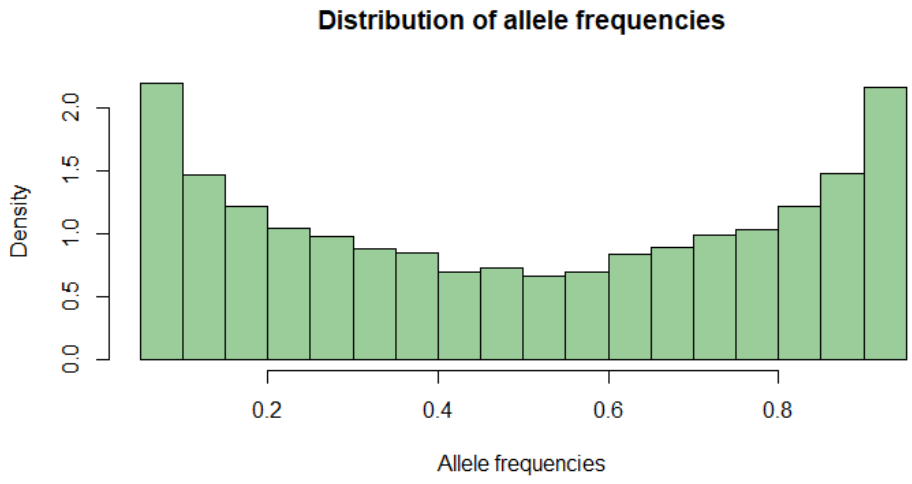


Figure 9 *" Distribution of Allele frequency" in SNPs.*

4.4. Population structure

PCA was used to visualize the individual relationships within and between the sampled populations. The first and second principal components accounted for 4% and 3% of the total variation, respectively. Overall, the data support that the populations are of the same genetic background (Figure 10).

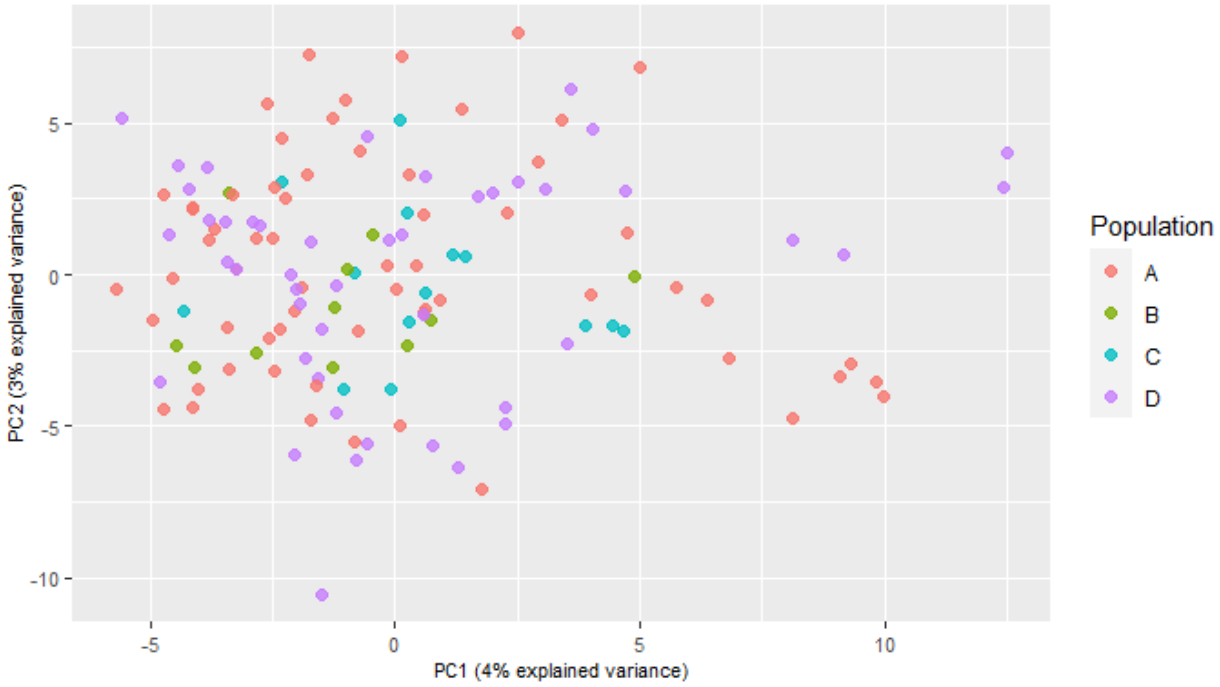


Figure 10 *Principal components analysis (PCA) for 133 individuals based on 5,159 binary SNPs (single-nucleotide polymorphisms). The genetic relationships among individual fish as seen when plotting the first and second principal components (PCA1 and PCA2). Each individual represented by one dot, with its symbol color corresponding to the location origin.*

4.5. Genetic diversity

Observed heterozygosity (H_O), expected heterozygosity (H_E) and inbreeding coefficient (F_{is}) were estimated by Stacks software version v2.5 (Rochette et al., 2019). The overall mean observed

heterozygosity (H_o) in the genotyped Arctic charr samples was 0.321, while the mean expected heterozygosity (H_E) was 0.311. The observed heterozygosity ranged from 0.303 in Farm-B population to 0.369 in Farm -C while the expected heterozygosity ranged from 0.302 in Farm-B population to 0.315 in Farm-D. Regarding the inbreeding coefficient (F_{is}) metric, the lowest value detected was -0.068 in Farm-C while values of 0.051, 0.046 and 0.042 were obtained for Farms A, D, and B respectively (Table 7).

Table 7. *Genetic diversity indices for Arctic charr (Salvelinus alpinus* L.*) from four locations sequenced using ddRADseq. N, number of samples per locus; Ho, observed heterozygosity; H_E, expected heterozygosity; Fis, inbreeding coefficient.*

Location	***N***	***H_o (mean ± SE)***	***H_E (mean ± SE)***	***F_{is} (mean ± SE)***
A	*56.84982*	*0.304 ± 0.002*	*0.314 ± 0.002*	*0.051 ± 0.07*
B	*10.07974*	*0.303 ± 0.003*	*0.302 ± 0.002*	*0.042 ± 0.02*
C	*12.46332*	*0.369 ± 0.003*	*0.312 ± 0.002*	*-0.068 ± 0.02*
D	*42.67428*	*0.307 ± 0.002*	*0.315 ± 0.002*	*0.046 ± 0.05*

4.6. Genetic distance amongst populations

Pairwise F_{st} estimates amongst the tested populations were obtained through the R package StAMPP (Pembleton et al., 2013). The lowest genetic distance was observed among population from Farm A and D (F_{st} =0.002). The genetic distance was approximately equal among population of Farm C & A and Farm C & D (F_{st} = 0.013) while, the highest genetic distance was observed among population of Farm C and B (F_{st}=0.019) (Table 8; Figure 11).

Table 8. *Fst values and corresponding confidence intervals (95%; 1000 bootstraps) between four farmed populations of Arctic charr (Salvelinus alpinus* L.*)*

Population 1	Population 2	Fst	Confidence Interval
A	B	0.004	0.002-0.006
A	C	0.013	0.011-0.015
A	D	0.002	0.002-0.003
B	C	0.019	0.016-0.021
B	D	0.008	0.006-0.010
C	D	0.013	0.011-0.014

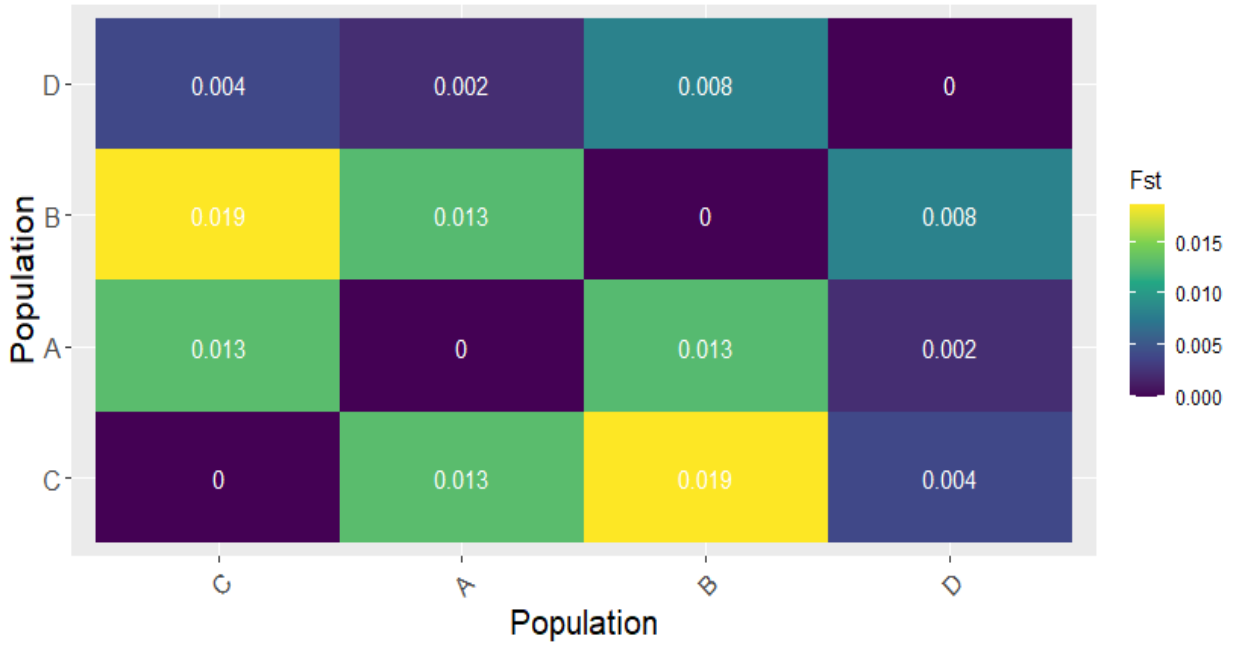


Figure 11 *Pairwise Fst estimates (lower and upper diagonal) among population of Arctic charr belongs to different location.*

4.7. Association analysis for sperm quality characteristics

An association analysis was carried out aiming to identify SNPs associated with either sperm concentration or percentage of viable cells in Arctic charr. Potential inflation of the estimated p-values was assessed through a "quantile-quantile (QQ) plot". No significant genome-wide association was found amongst the investigated genetic variants. The obtained p values of all genetic variants were below the threshold estimated from the Bonferroni criterion. (Figure 12 & 13)

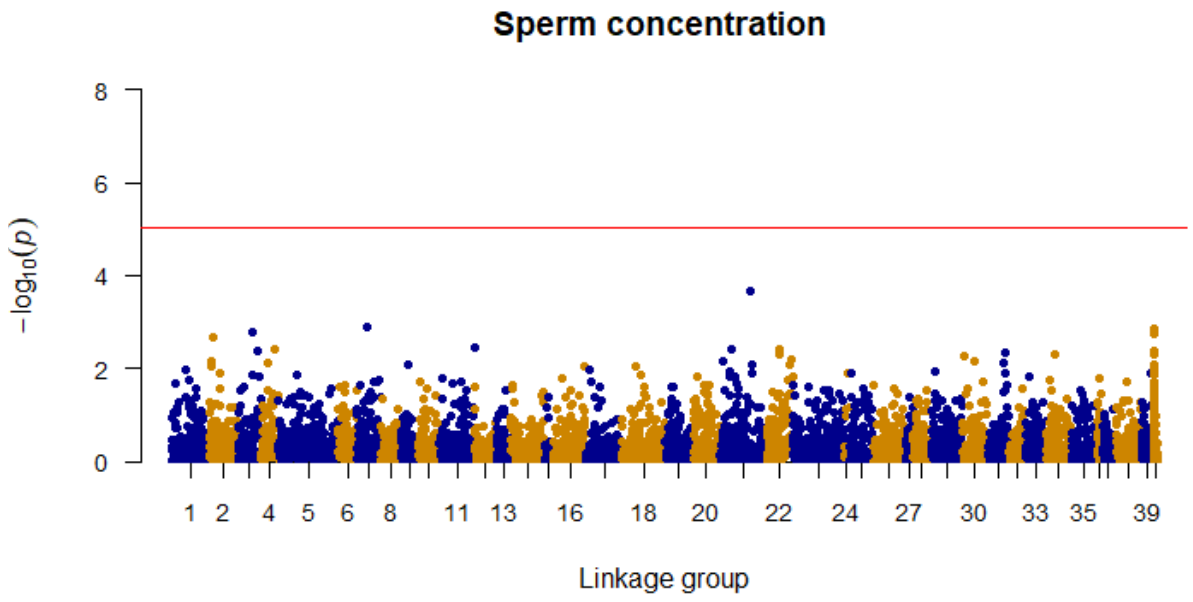


Figure 12 *Manhattan plot highlighting SNPs associated with sperm concentration in Arctic charr (Salvelinus alpinus* L.*). The horizontal red line represents the genome wide significant threshold. The chromosomal position of the SNPs are on the x axis while the corresponding p-value on the y axis.*

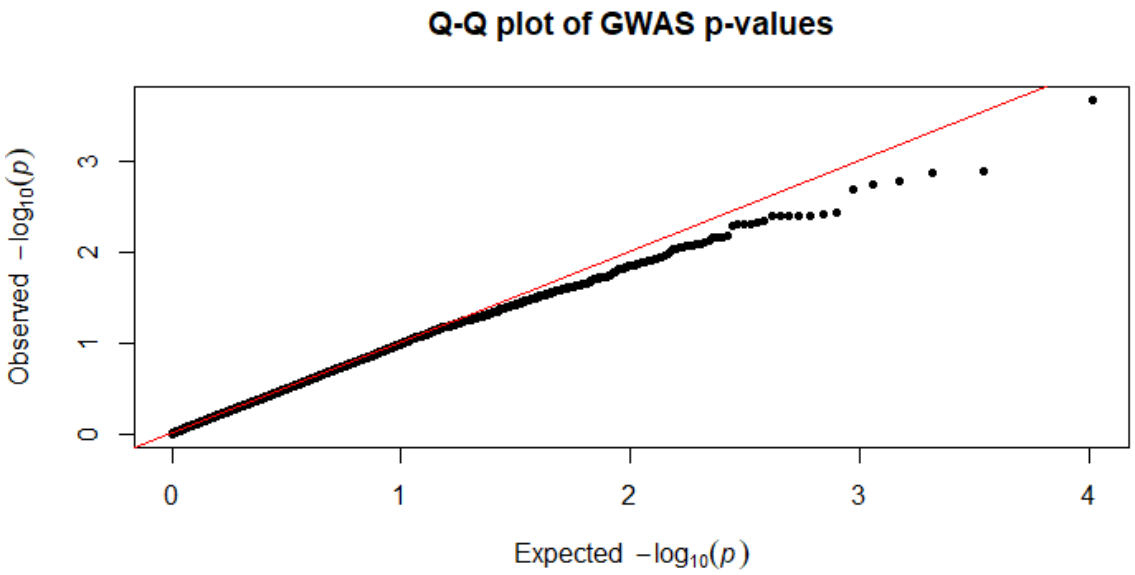


Figure 13 *QQ plot of the p values in GWAS analysis using the MLM model for sperm concentration. The observed negative base 10 logarithm of the p values of all SNPs are on y axis while the expected observed negative base 10 logarithm of the p values under the assumption of no association on the x axis.*

The figure (14) summarizes the association analysis results across percentage of viable cells. A single SNP (SNP_402577.127) located on chromosome number 31was significantly associated with the sperm cell viable percentage. Quantile-quantile plot (Figure 15) also supported the Manhattan plot result and showed a strongly associated SNP that deviate from the diagonal at the upper right end of the plot.

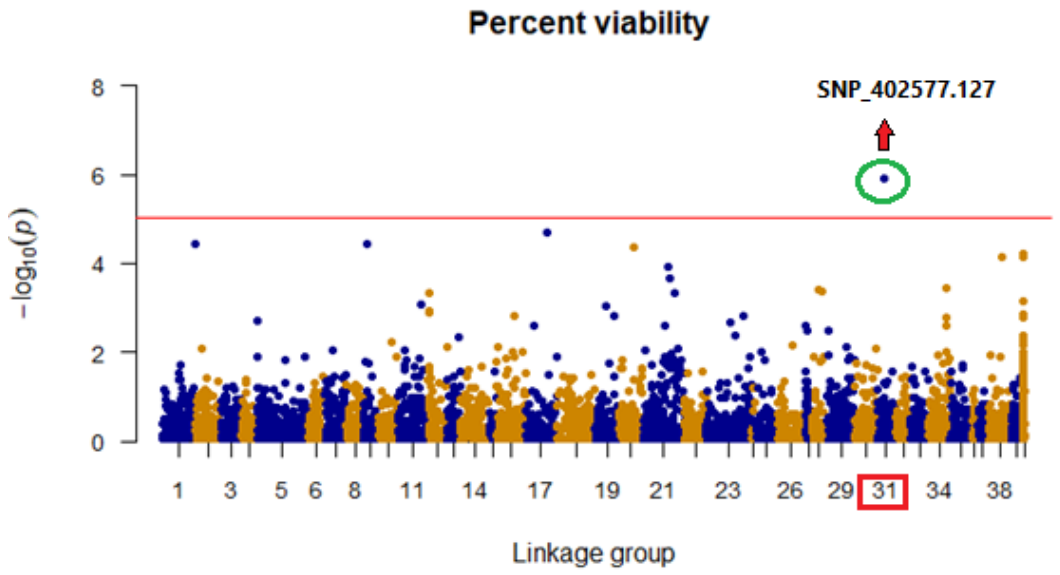


Figure 14 *Manhattan plot highlighting SNPs associated with percent viability in Arctic charr (Salvelinus alpinus* L.*). The horizontal red line represents the genome wide significant threshold. The chromosomal position of the SNPs are on the x axis while the association statistics on the y axis.*

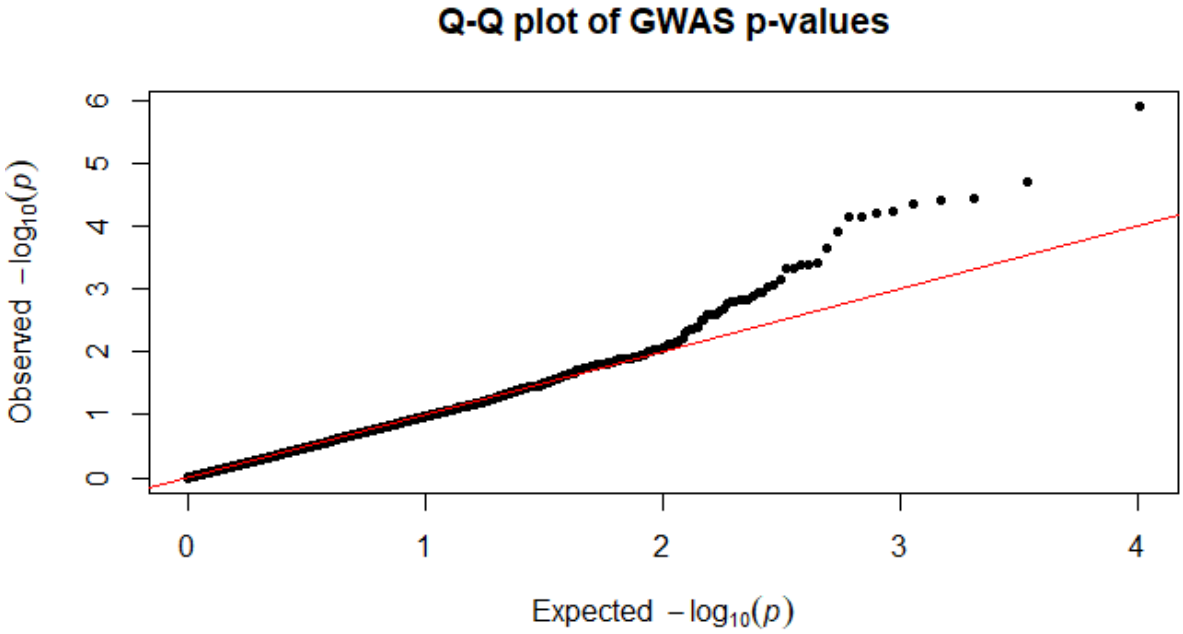


Figure 15 *QQ plot of the p values in GWAS analysis using the MLM model for percent viability. The observed negative base 10 logarithm of the p values of all SNPs are on y axis while the expected observed negative base 10 logarithm of the p values under the assumption of no association on the x axis.*

5. Discussion

In the current study we applied next generation sequencing technologies aiming to gain insights regarding the genetic diversity and the underlying genetic factors that are associated with sperm quality characteristics in farmed Arctic charr. Reproductive success is strikingly lower in farmed fish compared to their wild counterparts (Farquharson et al., 2018). Nevertheless, limited research has been conducted aiming to identify the causative genetic factors influencing reproductive related traits in any farmed fish species. In the review of Gjedrem and Rye (2016), only one study was reported studying genetic improvement in fertility related traits in farmed fish. The most dramatic drop regarding reproductive success amongst salmonids in captivity is observed in the case of Arctic charr. Most interestingly, no prior study attempted to investigate for underlying genetic components associated with sperm quality characteristics in Arctic charr.

To the best of our knowledge no other study investigated the status of the genetic diversity of farmed Arctic charr populations in Sweden using next generation sequencing technologies. Selective breeding enhanced with genomic information can contribute towards a sustainable aquaculture production through identification of the animals with the highest genetic potential and most efficient management of inbreeding accumulation that can be detrimental for amongst others reproductive related traits (Sae-Lim et al., 2017; Sonesson, 2007; Sonesson et al., 2012). ddRAD-seq is one of the most commonly utilized members of the reduced-representation family platforms combining simplicity and cost efficiency (Peterson et al., 2012). Over the last years, ddRAD-seq has been successfully utilized in a plethora of aquaculture studies like in the case of gilthead sea bream (Kyriakis et al., 2019) and Nile tilapia (Kajungiro et al., 2019; Moses et al., 2020).

Genetic diversity

Investigating the status of genetic diversity in farmed animal populations can provide valuable insights that could guide future management decisions. The efficient management of inbreeding accumulation is of the utmost importance for the long-term sustainability of any breeding program. Increased levels of inbreeding can negatively affect reproductive traits as was demonstrated in several studies in livestock (Bjelland et al., 2013; Martikainen et al., 2017; Mc Parland et al., 2007; Pryce et al., 2014).

PCA analysis indicated that the tested animals from all four farms composed a uniform group which was expected as all samples originated from the national Swedish breeding program of Arctic charr. Commonly used genetic diversity metrics were estimated in our study across all sampled populations aiming to gain insight regarding the status of the genetic pool of Arctic charr 'Superior'. In general, the observed heterozygosity level was similar across the studied populations. The population from Farm-C appeared to have the highest values regarding the estimated genetic diversity metrics, indicating a slight excess of heterozygosity. Overall the estimated level of heterozygosity in our study was lower compared to prior studies of Arctic charr from Lake Geneva H_o = 0.67 (Savary et al., 2017). Nevertheless, in the aforementioned study different types of genetic markers were used. As such, a direct comparison between the studies is not possible. However, when compared our results with other ddRADseq studies in other farmed fish species the estimated heterozygosity values were of similar magnitude or higher (Saenz-Agudelo et al., 2015; Leitwein et al., 2016; Moses et al., 2020). Interestingly, when compared our results with cases of farmed fish originating from hatcheries where no pedigree records were kept (Moses et al 2020) the heterozygosity values obtained in our study were almost double. In general,

selective breeding programs can minimize inbreeding accumulation when compared to practices where no prior information is utilized during the performed matings. The above is more pronounced in the case of farmed fish due to their high fecundity meaning that inbreeding can increase rapidly in the absence of pedigree records.

Naturally both environmental and genetic factors could affect fertility-related traits like sperm quality characteristics. Regarding environmental factors water rearing temperature plays a crucial role during the reproduction process. The conducted ANOVA indicated that the rearing location played a significant role regarding sperm concentration and number of non-viable cells. Since, no genetic structure was identified amongst the studied populations the aforementioned difference on sperm characteristics accounted to the rearing location is most likely the effect of the different rearing water temperatures the animals experienced in the different farms.

Association analysis for sperm quality parameters

In our study we investigated for associations between the identified SNP markers and either sperm density or sperm cell viability. No genome-wide significant associations were found in the case of sperm concentration indicating that the trait could be polygenic. In the case of polygenic traits, a large number of genes with small effect each is responsible for controlling the observed phenotype. Therefore in that scenario the most widespread method of choice is to apply genomic selection models in order to estimate genomic breeding values and thereafter perform selection on the highest ranking animals. Nevertheless, we need to acknowledge the fact that in our study a small number of individuals was used (n = 133) limiting our ability to detect small or even QTLs of moderate effect associated with sperm density. As such, follow-up studies with larger sample sizes would of value for deciphering the genetic architecture of the trait. Interestingly, when we

investigated for associations with sperm cell viability a single nucleotide polymorphism (SNP_402577.127) on chromosome number 31 surpassed the genome-wide significant threshold. The above SNP could be of value for directing future selection practices and therefore assisting in identifying animals with high sperm quality characteristics. Nevertheless, once more we need to acknowledge the small number of animals we used in our study. It would be of paramount importance to investigate where the association between the SNP marker and the sperm cell viability we identified in our study persists across larger datasets. Furthermore, it would be of significant importance to investigate the effect of selecting animals with high quality sperm characteristics (regarding e.g. sperm concentration and cell viability) on the number of successfully hatched eggs.

6. Conclusions

Overall, in our study, we discovered 5,159 SNPs from farmed Arctic charr males across different farms in Sweden. A wide range of analysis was performed aiming to gain insights about the genetic diversity of Arctic charr populations among different locations. Moreover, association analysis was performed through a linear mixed model in order to investigate for SNPs that could be utilized as predictors for identifying animals with highest potential regarding sperm quality characteristics. The conducted analysis revealed a significant SNP associated with sperm cell viability. This novel SNP may assist selection of male broodstock based on specific genetic signatures and can therefore improve the efficiency of Arctic charr farming. The results of the present study are a baseline for genetic monitoring of the reproductive success in Arctic charr farming. Future studies are needed to quantify the effect of selecting Arctic charr males based on their genomic profile on egg survival.

Bibliography

Abdelrahman, H., ElHady, M., Alcivar-Warren, A., Allen, S., Al-Tobasei, R., Bao, L., Beck, B., Blackburn, H., Bosworth, B., Buchanan, J., Chappell, J., Daniels, W., Dong, S., Dunham, R., Durland, E., Elaswad, A., Gomez-Chiarri, M., Gosh, K., Guo, X., … Zhou, T. (2017). Aquaculture genomics, genetics and breeding in the United States: Current status, challenges, and priorities for future research. *BMC Genomics*, *18*(1), 1–23. https://doi.org/10.1186/s12864-017-3557-1

Alon, S., Vigneault, F., Eminaga, S., Christodoulou, D. C., Seidman, J. G., Church, G. M., & Eisenberg, E. (2011). Barcoding bias in high-throughput multiplex sequencing of miRNA. *Genome Research*, *21*(9), 1506–1511. https://doi.org/10.1101/gr.121715.111

Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T. T., Mast, J., Sunayama-Morita, T., & Stern, D. L. (2011). Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Research*, *21*(4), 610–617. https://doi.org/10.1101/gr.115402.110

Bagenal, T. . (1969). Relationship between egg size and fry survival in brown trout Salmo trutta L. *Journal of Fish Biology*, *1*, 349–353. https://doi.org/doi:10.1111/j.1095-8649.1969.tb03882.x.

Baird, N. a, Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. a, Selker, E. U., Cresko, W. a, & Johnson, E. a. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One*, *3*(10), e3376. https://doi.org/10.1371/journal.pone.0003376

Bayona-Vásquez, N. J., Glenn, T. C., Kieran, T. J., Pierson, T. W., Hoffberg, S. L., Scott, P. A., Bentley, K. E., Finger, J. W., Louha, S., Troendle, N., Diaz-Jaimes, P., Mauricio, R., & Faircloth, B. C. (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ*, *2019*(10), 1–25. https://doi.org/10.7717/peerj.7724

Bjelland, D. W., Weigel, K. A., Vukasinovic, N., & Nkrumah, J. D. (2013). Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *Journal of Dairy Science*, *96*(7), 4697–4706. https://doi.org/10.3168/jds.2012-6435

Bolla, S., & Holmefjord, I. (1988). Effect of temperature and light on development of Atlantic halibut larvae. *Aquaculture*, *74*, 355–358.

Brännäs, E., Nilsson, J., & Eriksson, L. . (2011). *Rödingavel - en summering av det Svenska avelsprogrammet från 1982-2011. In Rapport (Sveriges lantbruksuniversitet, Institutionen för vilt, fisk och miljö) (12093069)*.

Brännäs, E., & Wiklund, B. S. (1992). Low temperature growth potential of Arctic charr and rainbow trout. *Nordic Journal of Freshwater Research*, *67*, 77–81.

Brown, J. K., Taggart, J. B., Bekaert, M., Wehner, S., Palaiokostas, C., Setiawan, A. N., Symonds, J. E., & Penman, D. J. (2016). Mapping the sex determination locus in the hapuku (Polyprion oxygeneios) using ddRAD sequencing. *BMC Genomics*, *17*(1), 1–12. https://doi.org/10.1186/s12864-016-2773-4

Carlberg, H., Nilsson, J., Brännäs, E., & Alanärä, A. (2018). An evaluation of 30 years of selective breeding in the Arctic charr (Salvelinus alpinus L.) and its implications for feeding management. *Aquaculture*, *495*(June 2017), 428–434. https://doi.org/10.1016/j.aquaculture.2018.06.008

Chistiakov, D. A., Hellemans, B., & Volckaert, F. A. M. (2006). Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture*, *255*(1–4), 1–29. https://doi.org/10.1016/j.aquaculture.2005.11.031

Christensen, K. A., Rondeau, E. B., Minkley, D. R., Leong, J. S., Nugent, C. M., Danzmann, R. G., Ferguson, M. M., Stadnik, A., Devlin, R. H., Muzzerall, R., Edwards, M., Davidson, W. S., & Koop, B. F. (2018). The arctic charr (salvelinus alpinus) genome and transcriptome assembly. *PLoS ONE*, *13*(9), 1–30. https://doi.org/10.1371/journal.pone.0204076

Cnaani, A., Ron, M., Lee, B. ., Hulata, G., Kocher, T. ., & Seroussi, E. (2002). Mapping the transferrin gene in tilapia. *Anim. Genet.*, *33*(1), 78–80. https://doi.org/10.1046/j.1365-2052.2002.0742e.x

Craig, D. W., Pearson, J. V., Szelinger, S., Sekar, A., Redman, M., Corneveaux, J. J., Pawlowski, T. L., Laub, T., Nunn, G., Stephan, D. A., Homer, N., & Huentelman, M. J. (2008). Identification of genetic variants using bar-coded multiplexed sequencing. *Nature Methods*, *5*(10), 887–893. https://doi.org/10.1038/nmeth.1251

Davey, J. W., Hohenlohe, P. a, Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L.

(2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews. Genetics*, *12*(7), 499–510. https://doi.org/10.1038/nrg3012

Elliott, J. M., & Elliott, J. A. (2010). Temperature requirements of Atlantic salmon Salmo salar, brown trout Salmo trutta and Arctic charr Salvelinus alpinus: Predicting the effects of climate change. *Journal of Fish Biology*, *77*(8), 1793–1817. https://doi.org/10.1111/j.1095-8649.2010.02762.x

Elliott, J. M., & Klemetsen, A. (2002). The upper critical thermal limits for alevins of Arctic charr from a Norwegian lake north of the Arctic circle. *Journal of Fish Biology*, *60*(5), 1338–1341. https://doi.org/10.1006/jfbi.2002.1934

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), 1–10. https://doi.org/10.1371/journal.pone.0019379

Eriksen, M. S., Bakken, M., Espmark, Å., Braastad, B. O., & Salte, R. (2006). Prespawning stress in farmed Atlantic salmon Salmo salar: Maternal cortisol exposure and hyperthermia during embryonic development affect offspring survival, growth and incidence of malformations. *Journal of Fish Biology*, *69*(1), 114–129. https://doi.org/10.1111/j.1095-8649.2006.01071.x

Eriksson, L.-O., Alanärä, A., Brännäs, E., Nilsson, J., & Kiessling, A. . (1993). Arctic charr farming in Sweden. *Bulletin of the Aquaculture Association Canada*, *93*(1), 18–24.

Eriksson, L. O., Alanärä, A., Nilsson, J., & Brännäs, E. (2010). The Arctic charr story: Development of subarctic freshwater fish farming in Sweden. *Hydrobiologia*, *650*(1), 265–274. https://doi.org/10.1007/s10750-010-0248-1

FAO. (2020). *Fisheries and Aquaculture Department.* http://www.fao.org/fishery/culturedspecies/Salmo_trutta/en.

Farquharson, K. A., Hogg, C. J., & Grueber, C. E. (2018). A meta-analysis of birth-origin effects on reproduction in diverse captive environments. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03500-9

Figueras, A., Robledo, D., Corvelo, A., Hermida, M., Pereiro, P., Rubiolo, J. A., Gómez-Garrido, J., Carreté, L., Bello, X., Gut, M., Gut, I. G., Marcet-Houben, M., Forn-Cuní, G., Galán, B., García, J. L., Abal-Fabeiro, J. L., Pardo, B. G., Taboada, X., Fernández, C., … Martínez, P. (2016). Whole genome sequencing of turbot (Scophthalmus maximus; Pleuronectiformes):

A fish adapted to demersal life. *DNA Research*, *23*(3), 181–192. https://doi.org/10.1093/dnares/dsw007

Fitzsimmons, S., & Perutz, M. (2006). Effects of egg incubation temperature on survival, prevalence and types of malformations in vertebral column of Atlantic Cod (Gadus morhua) larvae. *Bulletin - European Association of Fish Pathologists*, *26*(2), 80–86.

Fowler, L. . (1972). Growth and mortality of fingerling chinook salmon as affected by egg size. *The Progressive Fish-Culturist*, *34*, 66–69. https://doi.org/10.1577/1548-8640(1972)34[66:gamofc]2.0.co;2.

Franchini, P., Monné Parera, D., Kautt, A. F., & Meyer, A. (2017). quaddRAD: a new high-multiplexing and PCR duplicate removal ddRAD protocol produces novel evolutionary insights in a nonradiating cichlid lineage. *Molecular Ecology*, *26*(10), 2783–2795. https://doi.org/10.1111/mec.14077

Gjedrem, T., & Rye, M. (2016). Selection response in fish and shellfish: a review. *Reviews in Aquaculture*, *10*(1), 168–179. https://doi.org/10.1111/raq.12154

Gonen, S., Lowe, N. R., Cezard, T., Gharbi, K., Bishop, S. C., & Houston, R. D. (2014). Linkage maps of the Atlantic salmon (Salmo salar) genome derived from RAD sequencing. *BMC Genomics*, *15*(1), 1–17. https://doi.org/10.1186/1471-2164-15-166

Grolemund, G., & Wickham, H. (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data.* https://r4ds.had.co.nz/index.html

Grosjean, P., Ibanez Frederic, & Etienne, M. (2018). *Package "pastecs": Package for Analysis of Space-Time Ecological Series*. 76. https://cran.r-project.org/web/packages/pastecs/pastecs.pdf

Hevrøy, E. M., Hunskår, C., de Gelder, S., Shimizu, M., Waagbø, R., Breck, O., Takle, H., Sussort, S., & Hansen, T. (2013). GH-IGF system regulation of attenuated muscle growth and lipolysis in Atlantic salmon reared at elevated sea temperatures. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, *183*(2), 243–259. https://doi.org/10.1007/s00360-012-0704-5

Holman, L. E., Garcia de la serrana, D., Onoufriou, A., Hillestad, B., & Johnston, I. A. (2017). A workflow used to design low density SNP panels for parentage assignment and traceability in aquaculture species and its validation in Atlantic salmon. *Aquaculture*, *476*(December 2016), 59–64. https://doi.org/10.1016/j.aquaculture.2017.04.001

Huang, X., Jiang, Y., Zhang, W., Cheng, Y., Wang, Y., Ma, X., Duan, Y., Xia, L., Chen, Y., Wu, N., Shi, M., & Xia, X. Q. (2020). Construction of a high-density genetic map and mapping of growth related QTLs in the grass carp (Ctenopharyngodon idellus). *BMC Genomics*, *21*(1), 313. https://doi.org/10.1186/s12864-020-6730-x

Jeuthe, H. (2012). *Introductory Research Essay Reproductive success of farmed Arctic charr*. *16*.

Jeuthe, H., Brännäs, E., & Nilsson, J. (2013). Effects of egg size, maternal age and temperature on egg, viability of farmed Arctic charr. *Aquaculture*, *408–409*, 70–77. https://doi.org/10.1016/j.aquaculture.2013.05.034

Jobling, M., Tveiten, H., & Hatlen, B. (1998). Cultivation of Arctic charr: An update. *Aquaculture International*, *6*(3), 181–196. https://doi.org/10.1023/A:1009246509657

Johnson, L. (1980). The Arctic charr, Salvelinus alpinus. In Balon, E. K. and W. Junk (eds), Charrs: Salmonid Fishes of the Genus Salvilinus. *The Hague, Netherlands*, 15–98.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Collins, C. (2015). Analysing genome-wide SNP data using adegenet. *R Package*, 1–37. https://doi.org/10.2217/imt-2016-0023

Jónsson, B., & Svavarsson, E. (2000). Connection between egg size and early mortality in Arctic charr, Salvelinus alpinus. *Aquaculture*, *187*(3–4), 315–317.

Jungwirth, M., & Winkler, H. (1984). The temperature dependence of embryonic development of grayling (Thymallus thymallus), Danube salmon (Hucho hucho), Arctic char (Salvelinus alpinus) and brown trout (Salmo trutta fario). *Aquaculture*, *38*(4), 315–327. https://doi.org/10.1016/0044-8486(84)90336-3

Kaitaranta, J. K., & Linko, R. R. (1984). Fatty acids in the roe lipids of common food fishes. *Comparative Biochemistry and Physiology -- Part B: Biochemistry And*, *79*(3), 331–334. https://doi.org/10.1016/0305-0491(84)90384-5

Kajungiro, R. A., Palaiokostas, C., Pinto, F. A. L., Mmochi, A. J., Mtolera, M., Houston, R. D., & de Koning, D. J. (2019). Population Structure and Genetic Diversity of Nile Tilapia (Oreochromis niloticus) Strains Cultured in Tanzania. *Frontiers in Genetics*, *10*(December), 1–12. https://doi.org/10.3389/fgene.2019.01269

King, H. R., Pankhurst, N. W., Watts, M., & Pankhurst, P. M. (2003). Effect of elevated summer temperatures on gonadal steroid production, vitellogenesis and egg quality in female

Atlantic salmon. *Journal of Fish Biology*, *63*(1), 153–167. https://doi.org/10.1046/j.1095-8649.2003.00137.x

Kumar, G., & Kocour, M. (2017). Applications of next-generation sequencing in fisheries research: A review. *Fisheries Research*, *186*, 11–22. https://doi.org/10.1016/j.fishres.2016.07.021

Kyriakis, Di., Kanterakis, A., Manousaki, T., Tsakogiannis, A., Tsagris, M., Tsamardinos, I., Papaharisis, L., Chatziplis, Di., Potamias, G., & Tsigenopoulos, C. S. (2019). Scanning of genetic variants and genetic mapping of phenotypic traits in gilthead seabream through ddRAD sequencing. *Frontiers in Genetics*, *10*(JUN), 1–15. https://doi.org/10.3389/fgene.2019.00675

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Larson, W. A., McKinney, G. J., Limborg, M. T., Everett, M. V., Seeb, L. W., & Seeb, J. E. (2016). Identification of Multiple QTL Hotspots in Sockeye Salmon (Oncorhynchus nerka) Using Genotyping-by-Sequencing and a Dense Linkage Map. *Journal of Heredity*, *107*(2), 122–133. https://doi.org/10.1093/jhered/esv099

Larsson, S., & Berglund, I. (1998). Growth and food consumption of 0+ Arctic charr fed pelleted or natural food at six different temperatures. *Journal of Fish Biology*, *52*(2), 230–242. https://doi.org/10.1006/jfbi.1997.0575

Leitwein, M., Gagnaire, P. A., Desmarais, E., Guendouz, S., Rohmer, M., Berrebi, P., & Guinand, B. (2016). Genome-wide nucleotide diversity of hatchery-reared Atlantic and Mediterranean strains of brown trout Salmo trutta compared to wild Mediterranean populations. *Journal of Fish Biology*, *89*(6), 2717–2734. https://doi.org/10.1111/jfb.13131

Li, L., & Guo, X. (2004). AFLP-based genetic linkage maps of the Pacific oyster Crassostrea gigas Thunberg. *Marine Biotechnology*, *6*(1), 26–36. https://doi.org/10.1007/s10126-003-0001-0

Liu, J. Y., Lun, Z. R., Zhang, J. Bin, & Yang, T. B. (2009). Population genetic structure of striped mullet, Mugil cephalus, along the coast of China, inferred by AFLP fingerprinting. *Biochemical Systematics and Ecology*, *37*(4), 266–274. https://doi.org/10.1016/j.bse.2009.04.010

Liu, Z., Li, P., Kucuktas, H., Nichols, A., Tan, G., Zheng, X., Argue, B. ., Yant, R., & Dunham,

R. (1999). Development of AFLP markers for genetic linkage mapping analysis using channel catfish and blue catfish interspecific hybrids. *Trans. Am. Fish. Soc.*, *128*, 317– 327.

Liu, Z. J., & Cordes, J. F. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*, *238*(1–4), 1–37. https://doi.org/10.1016/j.aquaculture.2004.05.027

Liu, Z, Li, P., Argue, B. ., & Dunham, R. . (1998). Inheritance of RAPD markers in channel catfish (Ictalurus punctatus), blue catfish (I. furcatus), and their F1, F2 and backcross hybrids. *Animal Genetics*, *29*(1), 58–62. https://doi.org/10.1046/j.1365-2052.1998.00284.x

Liu, Zhanjiang, Karsi, A., Li, P., Cao, D., & Dunham, R. (2003). An AFLP-Based Genetic Linkage Map of Channel Catfish (Ictalurus punctatus) Constructed by Using an Interspecific Hybrid Resource Family. *Genetics*, *165*(2), 687–694.

Liu, Zhanjiang, Li, P., Kocabas, A., Karsi, A., & Ju, Z. (2001). Microsatellite-containing genes from the channel catfish brain: Evidence of trinucleotide repeat expansion in the coding region of nucleotide excision repair gene RAD23B. *Biochemical and Biophysical Research Communications*, *289*(2), 317–324. https://doi.org/10.1006/bbrc.2001.5978

Lyytikäinen, T., Koskela, J., & Rissanen, I. (1997). Thermal resistance and upper lethal temperatures of underyearling Lake Inari Arctic charr. *Journal of Fish Biology*, *51*(3), 515–525. https://doi.org/10.1006/jfbi.1997.0456

Mansour, N., Lahnsteiner, F., McNiven, M. A., Richardson, G. F., & Pelletier, C. S. (2011). Relationship between fertility and fatty acid profile of sperm and eggs in Arctic char, Salvelinus alpinus. *Aquaculture*, *318*(3–4), 371–378. https://doi.org/10.1016/j.aquaculture.2011.05.023

Martikainen, K., Tyrisevä, A. M., Matilainen, K., Pösö, J., & Uimari, P. (2017). Estimation of inbreeding depression on female fertility in the Finnish Ayrshire population. *Journal of Animal Breeding and Genetics*, *134*(5), 383–392. https://doi.org/10.1111/jbg.12285

Mc Parland, S., Kearney, J. F., Rath, M., & Berry, D. P. (2007). Inbreeding effects on milk production, calving performance, fertility, and conformation in Irish Holstein-Friesians. *Journal of Dairy Science*, *90*(9), 4411–4419. https://doi.org/10.3168/jds.2007-0227

Metzker, M. L. (2010). Sequencing technologies the next generation. *Nature Reviews Genetics*, *11*(1), 31–46. https://doi.org/10.1038/nrg2626

Mickett, K., Morton, C., Feng, J., Li, P., Simmons, M., Cao, D., Dunham, R. A., & Liu, Z. (2003). Assessing genetic diversity of domestic populations of channel catfish (Ictalurus

punctatus) in Alabama using AFLP markers. *Aquaculture*, *228*(1–4), 91–105. https://doi.org/10.1016/S0044-8486(03)00311-9

Moses, M., Mtolera, M. S. P., Chauka, L. J., Lopes, F. A., de Koning, D. J., Houston, R. D., & Palaiokostas, C. (2020). Characterizing the genetic structure of introduced Nile tilapia (Oreochromis niloticus) strains in Tanzania using double digest RAD sequencing. *Aquaculture International*, *28*(2), 477–492. https://doi.org/10.1007/s10499-019-00472-5

Naish, K., & Park, L. (2002). Linkage relationships for 35 new microsatellite loci in chinook salmon Oncorhynchus tshawytscha. *Anim. Genet.*, *33*(4), 218– 316. https://doi.org/10.1046/j.1365-2052.2002.t01-4-00886.x

Nilsson, J. (1993). Arctic charr strain crosses: effects on growth and sexual maturity. *Journal of Fish Biology*, *43*, 163–171. https://doi.org/10.1111/j.1095-8649.1993.tb00420.x.

Nilsson, J., Brännäs, E., & Eriksson, L. O. (2010). The Swedish Arctic charr breeding programme. *Hydrobiologia*, *650*(1), 275–282. https://doi.org/10.1007/s10750-010-0232-9

Oyarzún, P. A., Toro, J. E., & Navarro, J. M. (2013). Comparison of the physiological energetics between Mytilus chilensis, Mytilus galloprovincialis and their hybrids, under laboratory conditions. *Aquaculture Research*, *44*(12), 1805–1814. https://doi.org/10.1111/j.1365-2109.2012.03184.x

Palaiokostas, C., Bekaert, M., Khan, M. G. Q., Taggart, J. B., Gharbi, K., McAndrew, B. J., & Penman, D. J. (2015). A novel sex-determining QTL in Nile tilapia (Oreochromis niloticus). *BMC Genomics*, *16*(1), 1–10. https://doi.org/10.1186/s12864-015-1383-x

Palaiokostas, C., Cariou, S., Bestin, A., Bruant, J. S., Haffray, P., Morin, T., Cabon, J., Allal, F., Vandeputte, M., & Houston, R. D. (2018). Genome-wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (Dicentrarchus labrax) using RAD sequencing. *Genetics Selection Evolution*, *50*(1), 1–11. https://doi.org/10.1186/s12711-018-0401-2

Palaiokostas, C., Ferraresso, S., Franch, R., Houston, R. D., & Bargelloni, L. (2016). Genomic prediction of resistance to pasteurellosis in gilthead sea bream (Sparus aurata) using 2B-RAD sequencing. *G3: Genes, Genomes, Genetics*, *6*(11), 3693–3700. https://doi.org/10.1534/g3.116.035220

Palaiokostas, C., & Houston, R. D. (2017). Genome-wide approaches to understanding and improving complex traits in aquaculture species. *CAB Reviews: Perspectives in Agriculture,*

Veterinary Science, Nutrition and Natural Resources, *12*(055), 1–10. https://doi.org/10.1079/PAVSNNR201712055

Palaiokostas, C., Kocour, M., Prchal, M., & Houston, R. D. (2018). Accuracy of genomic evaluations of juvenile growth rate in common carp (Cyprinus carpio) using genotyping by sequencing. *Frontiers in Genetics*, *9*(MAR), 1–9. https://doi.org/10.3389/fgene.2018.00082

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*(5), 946–952. https://doi.org/10.1111/1755-0998.12129

Perdry, H., Dandine-Roulland, C., Bandyopadhyay, D., & Kettne, L. (2020). *Gaston pakage: version 1.5.6 Genetic Data Handling (QC, GRM, LD, PCA) & Linear Mixed Models.* https://doi.org/10.1159/000488519>.License

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0037135

Pickova, J., & Brännäs, E. (2006). Egg quality in Arctic charr (Salvelinus alpinus). *Archiv Fur Tierzucht*, *49*, 86–90.

Pickova, J., Brännäs, E., & Andersson, T. (2007). Importance of fatty acids in broodstock diets with emphasis on Arctic char (Salvelinus alpinus) eggs. *Aquaculture International*, *15*(3–4), 305–311. https://doi.org/10.1007/s10499-007-9084-3

Pryce, J. E., Haile-Mariam, M., Goddard, M. E., & Hayes, B. J. (2014). Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genetics Selection Evolution*, *46*(1), 1–14. https://doi.org/10.1186/s12711-014-0071-7

Rexroad, C., Coleman, R., Hershberger, W., & Killefer, J. (2002a). Eighteen polymorphic microsatellite markers for rainbow trout (Oncorhynchus mykiss). *Anim. Genet*, *33*, 76–78. https://doi.org/10.1046/j.1365-2052.2002.0742d.x

Rexroad, C., Coleman, R. ., Hershberger, W., & Killefer, J. (2002b). Thirty-eight polymorphic microsatellite markers for mapping in rainbow trout. *J. Anim. Sci.*, *80*(2), 541–542. https://doi.org/10.2527/2002.802541x

Rexroad, C., Coleman, R., Martin, A., Hershberger, W., & Killefer, J. (2001). Thirty-five polymorphic microsatellite markers for rainbow trout (Oncorhynchus mykiss). *Anim. Genet.*, *32*, 317–319. https://doi.org/10.1046/j.1365-2052.2001.0730b.x

Robledo, D., Palaiokostas, C., Bargelloni, L., Martínez, P., & Houston, R. (2018). Applications of genotyping by sequencing in aquaculture breeding and genetics. *Reviews in Aquaculture*, *10*(3), 670–682. https://doi.org/10.1111/raq.12193

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754. https://doi.org/10.1111/mec.15253

Sae-Lim, P., Kause, A., Mulder, H. A., & Olesen, I. (2017). Breeding and genetics symposium: Climate change and selective breeding in aquaculture. *Journal of Animal Science*, *95*(4), 1801–1812. https://doi.org/10.2527/jas2016.1066

Saenz-Agudelo, P., DiBattista, J. D., Piatek, M. J., Gaither, M. R., Harrison, H. B., Nanninga, G. B., & Berumen, M. L. (2015). Seascape genetics along environmental gradients in the Arabian Peninsula: Insights from ddRAD sequencing of anemonefishes. *Molecular Ecology*, *24*(24), 6241–6255. https://doi.org/10.1111/mec.13471

Savary, R., Dufresnes, C., Champigneulle, A., Caudron, A., Dubey, S., Perrin, N., & Fumagalli, L. (2017). Stocking activities for the Arctic charr in Lake Geneva: Genetic effects in space and time. *Ecology and Evolution*, *7*(14), 5201–5211. https://doi.org/10.1002/ece3.3073

Seki, S., Agresti, J., Gall, G. A., Taniguchi, N., & May, B. (1999). AFLP analysis of genetic diversity in three populations of Ayu plecoglossus altivelis. *Fisheries Science*, *65*(6), 888–892. https://doi.org/10.2331/fishsci.65.888

Slettan, A., Olsaker, I., & Lie, Ø. (1997). Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. *Heredity*, *78*(6), 620–627. https://doi.org/10.1038/hdy.1997.101

Sonesson, A. K. (2007). Within-family marker-assisted selection for aquaculture species. *Genetics Selection Evolution*, *39*(3), 301–317. https://doi.org/10.1051/gse:2007005

Sonesson, A. K., Woolliams, J. A., & Meuwissen, T. H. E. (2012). Genomic selection requires genomic control of inbreeding. *Genetics Selection Evolution*, *44*(1), 1–10. https://doi.org/10.1186/1297-9686-44-27

Statista. (2019). *Production volume of arctic char in Sweden from 2007 to 2017. Agriculture Fisheries and Aquaculture*. https://www.statista.com/statistics/795740/production-volume-of-arctic-char-in-sweden/

Syaifudin, M., Bekaert, M., Taggart, J. B., Bartie, K. L., Wehner, S., Palaiokostas, C., Khan, M.

G. Q., Selly, S. L. C., Hulata, G., D'Cotta, H., Baroiller, J. F., McAndrew, B. J., & Penman, D. J. (2019). Species-Specific Marker Discovery in Tilapia. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-48339-2

Tan, G., Karsi, A., Li, P., Kim, S., Zheng, X., Kucuktas, H., Argue, B. ., Dunham, R. ., & Liu, Z. (1999). Polymorphic microsatellite markers in Ictalurus punctatus and related catfish species. *Mol. Ecol.*, *8*, 1758–1760. https://doi.org/10.1046/j.1365-294x.1999.00723-3.x

Tanck, M., Palstra, A., van der Weerd, M Leffering, C., van der Poel, J.J Bovenhuis, H., & Komen, J. (2001). Segregation of microsatellite alleles and residual heterozygosity at single loci in homozygous androgenetic common carp (Cyprinus carpio L.). *Genome.*, *44*(5), 743–751. https://doi.org/10.1139/g01-072

Thyholdt, S. B. (2014). the Importance of Temperature in Farmed Salmon Growth: Regional Growth Functions for Norwegian Farmed Salmon. *Aquaculture Economics and Management*, *18*(2), 189–204. https://doi.org/10.1080/13657305.2014.903310

Toonen, R. J., Puritz, J. B., Forsman, Z. H., Whitney, J. L., Fernandez-Silva, I., Andrews, K. R., & Bird, C. E. (2013). ezRAD: A simplified method for genomic genotyping in non-model organisms. *PeerJ*, *2013*(1), 1–15. https://doi.org/10.7717/peerj.203

Torstensen, B. E., Frøyland, L., & Lie, Ø. (2001). *Kapittel 3: Lipider. In Fiskeernæring (Waagbø, R., Espe, M., Hamre, K. & Lie, Ø., eds.), pp. 57-75: Kystnæringen.*

Turner, S. (2017). qqman: Q-Q and Manhattan Plots for GWAS Data. R package version 0.1.4. *Https://CRAN.R-Project.Org/Package=qqman*. https://cran.r-project.org/package=qqman

Wang, S., Meyer, E., Mckay, J. K., & Matz, M. V. (2012). 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nature Methods*, *9*(8), 808–810. https://doi.org/10.1038/nmeth.2023

Wang, Z., Jayasankar, P., & Khoo, S. (2000). AFLP fingerprinting reveals genetic variability in common carp stocks from Indonesia. *Asian Fisheries Science*, *13*, 139–147.

Wiegand, M., Hataley, J., Kitchen, C., & Buchanan, L. (1989). Induction of developmental abnormalities in larval goldfish, Carassius auratus L., under cool incubation conditions. *Journal of Fish Biology*, *35*, 85–95. https://doi.org/10.1111/j.1095-8649.1989.tb03395.x

Yáñez, J. M., Newman, S., & Houston, R. D. (2015). Genomics in aquaculture to better understand species biology and accelerate genetic progress. *Frontiers in Genetics*, *6*(APR), 1–3. https://doi.org/10.3389/fgene.2015.00128

Young, W. P., Wheeler, P. A., Coryell, V. H., Keim, P., & Thorgaard, G. H. (1998). A detailed linkage map of rainbow trout produced using doubled haploids. *Genetics*, *148*(2), 839–850.

Yu, Z., & Guo, X. (2003). Genetic linkage map of the eastern oyster Crassostrea virginica Gmelin. *Biological Bulletin*, *204*(3), 327–338. https://doi.org/10.2307/1543603

Zhu, J. J., Liu, S. L., Ma, K. Y., Li, J. Le, & Lian, Q. P. (2011). Twenty-four polymorphic microsatellite markers from spotted barbell (Hemibarbus maculates). *Conservation Genetics Resources*, *3*(2), 239–241. https://doi.org/10.1007/s12686-010-9331-1

Acknowledgements

This work was performed at the Department of Animal Breeding and Genetic, Swedish University of Agricultural Sciences (SLU). I am grateful to everyone who have helped me and provided convenience for this work during this half a year. I learned a lot from some of you. Here I would like to acknowledge particularly:

Christos Palaiokostas, I would like to thank you for accepting me as your master student. As a supervisor, your patience and guidance were appreciable. All the experiences that you taught me to manipulate the laboratory apparatuses during my experimental work. Thanks for the help you provided to resolve all kind of problems that I faced during statistical analysis. I especially would like to thank you for every effort that you made to improve my dissertation. I always learned something new from each of the conversation that we had. I admire your efforts.

Tytti Vanhala, for informative lab survey.

Cano Merkan, for fixing computer problems.

A would like to thanks my family for all kinds of support and encouragement during this time. A special thank to my cute daughter, Adeen Farid for her nice smiles to elevate my ambitions when I was in blue mood. Special thanks to my beloved husband, Farid Masood for his support and cooperation.