An investigation of genetic and reproductive differences between Faroe Plateau and Faroe Bank cod (*Gadus morhua* L.)

A thesis presented by

Petra Elisabeth Petersen

For the degree of

Doctor of Philosophy

UNIVERSITY OF STIRLING

November 2014

Institute of Aquaculture, University of Stirling, Scotland

Acknowledgements

I would like to take this opportunity to thank my supervisors, especially Dr. John B. Taggart for his skilled supervision and support throughout numerous challenges these past years and for overwhelming hospitality on my visits to Scotland. I am also indebted to Dr. David J. Penman and Dr. Øystein Patursson for generous advice and support when it was most needed and to Dr. Geir Dahle for generosity and hospitality on my visits to Bergen.

I wish to acknowledge Dr. Hannes Gislason for giving me the opportunity to conduct research that I am truly passionate about. Special thanks must also go to Dr. Michaël Bekaert for outstanding assistance with the bioinformatic analyses, Dr. Øivind Andersen for his valuable collaboration and input concerning hemoglobin and transferrin screenings and Dr. Petur Steingrund and Eyðfinn Magnussen for their valuable cooperation relating to the field data.

Acknowledgements are also due to the many institutions that lent their expertise and facilities to this study. These include the Faroe Marine Research Institute, the Institute of Marine Research in Bergen, Edinburgh Genomics (University of Edinburgh), the Faculty of Science and Technology (University of the Faroe Islands), the Genetic Biobank of the Faroe Islands, the Research Park iNOVA and Nofima, the Norwegian Institute of Food, Fisheries and Aquaculture Research.

This work has only been possible due to generous and flexible financial support from the Faroese Research Council and Fiskaaling, the Aquaculture Research Station of the Faroes.

To all my colleagues at Fiskaaling, your help, advice and company has been much appreciated. I am particularly grateful to Lisbeth Joensen and Hans Jákup Jacobsen for their assistance with the breeding experiments. I would also like to acknowledge Heðin Abrahamsen, always ready to give a hand with computer problems, and Dr. Knud Simonsen for sharing his great knowledge on the oceanographic features around the Faroe Islands.

To my friends, who I have neglected greatly, thank you for still being there, no matter what. I look forward to spend more time with you.

I would like to express my most sincere gratitude towards my family who has supported me in endless many ways. To my parents, I am grateful for your absolute support and faith in me. To my brother, I hope you know that you have a special place in my heart. To my parents in law, thanks for finding so many ways of being supportive.

A special thanks to You, Morten, I could not have done this without your constant encouragement. I would like to dedicate this thesis to our wonderful children (my greatest achievements during the course of this Ph.D. study).

Abstract

The Atlantic cod (*Gadus morhua* L.) fishery is of great economic importance to the Faroese economy. There are two separately managed cod stocks around the Faroe Islands, the Faroe Plateau and the Faroe Bank cod. Both have experienced dramatic decreases in size and informed management decisions are vital for both stock viability and exploitation. The stocks are geographically isolated by an 800 m deep channel and water temperatures are on average 1 – 2 ºC higher on the Faroe Bank than on the Faroe Plateau. There are clear phenotypic differences between the stocks; in particular, the markedly higher growth rate for the Faroe Bank cod has caught public and scientific attention. There is continuing debate regarding the relative importance of genetics and environmental contributions to the contrasting phenotypes. Analyses of reproductive parameters (field data and experimental captive spawnings) as well as analyses of microsatellite and single nucleotide polymorphism (SNP) markers were undertaken to better resolve the issue.

Field data as well as data from experimental captive spawnings provided evidence of reproductive differences between Faroe Plateau and Faroe Bank cod. Peak spawning occurred earlier on the Faroe Plateau than on the Faroe Bank and this difference in timing of spawning was maintained in captivity. In particular, differences in sizes of eggs (average diameters of 1.40 and 1.30 mm for Faroe Plateau and Faroe Bank cod eggs, respectively) and indirect evidence of greater volumes spawned by the Faroe Bank females suggested stock differences with respect to egg size – egg number trade-off. It was hypothesised that the strategy adopted by cod on the Faroe Bank, with a higher number of smaller eggs, evolved in response to a more hostile environment (bare seabed and higher exposure to predators) experienced by early life stages in this area.

Experimental captive spawnings with Faroe Bank cod showed a large interfamily skew in survival rates of cod eggs and fry. Egg size was identified as a useful indicator of survival rates in the egg stage, but egg survival rates could not be used to predict viability in later developmental stages, thus highlighting the importance of employing some sort of genetic monitoring of cod fry to

ensure sufficient family representation in the progeny. While no tank effect was evident concerning fry survival, a significant tank effect was identified concerning body sizes of fry.

Microsatellite data were analysed using large sample sizes of Faroe Plateau and Faroe Bank cod with the Faroe Plateau divided into two locations, Faroe Plateau North-East and Faroe Plateau West (cod from each of the two were known to belong to separate spawning grounds). Two Norwegian coastal cod samples were included as outlier populations. While no genetic differentiation was detected between the two Faroe Plateau locations, these analyses revealed a detectable, albeit relatively modest, degree of genetic differentiation between cod from the Faroe Plateau and the Faroe Bank (F_{ST} = 0.0014 and 0.0018; D_{Jost_EST} = 0.0027 and 0.0048; $P < 0.0001$ and $P < 0.001$ for the Faroe Plateau North-East – Faroe Bank and the Faroe Plateau West – Faroe Bank comparisons). These values were several times smaller than those between Faroese and Norwegian coastal cod (pairwise F_{ST} and D_{Jost_EST} values in the range of 0.0061 – 0.0137 and 0.0158 – 0.0386, respectively). Despite recent reductions in census population sizes for Faroe Plateau and, particularly, Faroe Bank cod, genetic diversity estimates were comparable to the ones observed for Norwegian coastal cod and there was no evidence of significant genetic bottlenecks. Lastly, data for one of the markers (*Gmo*132) indicated genotype-dependent vertical distribution of cod (as investigated for Faroe Plateau North-East cod).

Contrary to some previously published studies, analysis of SNPs of two candidate genes for adaptive divergence, the hemoglobin gene *Hb-β*$_1$ and the transferrin gene *Tf1*, failed to detect differentiation between samples of Faroe Plateau and Faroe Bank cod analysed in this thesis.

Of 3533 novel SNPs simultaneously discovered and genotyped by restriction-site associated DNA (RAD) sequencing, 58 showed evidence of genetic differentiation between Faroe Plateau North-East and Faroe Bank cod ($P < 0.05$). No single locus was fixed for different alleles between Faroe Plateau and Faroe Bank cod. A set of eight informative SNPs (F_{ST} values between Faroe Plateau and Faroe Bank samples > 0.25; $P < 0.0005$) were selected for validation in larger samples, that included cod from both Faroe Plateau areas and the Faroe Bank as well as

Norwegian coastal and White Sea cod. Six out of the eight loci amplified successfully with a PCR-based method and there was 100 % concordance between genotypes of individuals screened by both techniques. Due to ascertainment bias, the SNPs should only be applied with caution in a broader geographical context. Nonetheless, these SNPs did confirm the genetic substructure suggested for Faroese cod by microsatellite analyses. While no genetic differentiation was evident between the two Faroe Plateau locations, significant genetic differentiation was evident between Faroe Plateau and Faroe Bank cod at five of the SNPs (F_{ST} values in the range of 0.0383 – 0.1914). This panel of five SNPs could confidently be used to trace groups of Faroe Plateau and Faroe Bank cod to their population of origin.

In conclusion, multiple lines of evidence demonstrate that Faroe Plateau and Faroe Bank cod are truly two genetically distinct populations. While the findings contribute to a broader understanding of the biology and the genetics of Faroe Plateau and Faroe Bank cod, the novel SNPs developed may provide a valuable resource for potential future demands of i.e. genetic stock identification methods.

Table of contents

Page

List of tables

List of figures

Glossary of common and Latin names of species

Adriatic sturgeon	*Acipenser naccarii*
Atlantic cod	*Gadus morhua* (Linnaeus, 1758)
Atlantic halibut	*Hippoglossus hippoglossus*
Atlantic herring	*Clupea harengus*
Atlantic salmon	*Salmo salar*
Bacteria causing vibriosis	*Vibrio anguillarum*
Blood parasite	*Trypanoplasma borreli*
Blue whiting	*Micromesistius poutassau*
Brine shrimp	*Artemia*
Brown trout	*Salmo trutta* L.
Capelin	*Mallotus villosus*
Cave fish	*Astyanax maxicanus*
Channel catfish	*Ictalurus punctatus*
Coelacanth	*Latimeria chalumnae*
Common carp	*Cyprinus carpio* L. (Linnaeus, 1758)
Common sole	*Solea solea*
Deepwater redfish	*Sebastes mentella* (Travin, 1951)
Eastern Fence Lizard	*Sceloporus undulatus*
Eel	*Anguilla anguilla*
European clam	*Ruditapes decussatus*
European flounder	*Platichthys flesus* L.
European hake	*Merluccius merluccius*
Greenland cod	*Gadus ogac*
Greenland halibut	*Reinhardtius hippoglossoides*
Haddock	*Melanogrammus aeglefinus*
Japanese medaka	*Oryzias latipes*

Lumpfish	*Cyclopterus lumpus*
Marine copepod	*Calanus finmarchicus*
Nile Tilapia	*Oreochromis niloticus* L.
Norway pout	*Trisopterus esmarkii*
Pacific oyster	*Crassostrea gigas*
Persian sturgeon	*Acipenser persicus*
Platyfish	*Xiphphorus maculatus*
Puffer fish	*Tetraodon nigroviridis*
Rainbow trout	*Oncorhynchus mykiss*
Rotifers	*Brachionus* spp.
Roundworm	*Caenorhabditis elegans*
Russian sturgeon	*Acipenser gueldenstaedtii*
Salmon louse	*Lepeophtheirus salmonis*
Siberian sturgeon	*Acipenser baerii*
Sockeye salmon	*Oncorhynchus nerka*
Spotted gar	*Lepisosteus oculatus*
Stickleback	*Gasterosteus aculeatus*
Tiger puffer fish (fugu)	*Takifugu rubripes*
Trinadian guppy	*Poecilia reticulata*
Tusk	*Brosme brosme*
Walleye pollock	*Theragra chalcogramma*
Whiting	*Merlangius merlangus*
Zebrafish	*Danio rerio*

Glossary of abbreviations and acronyms

ABI	Applied Biosystems
ANOVA	Analysis of variance
AR	Allelic richness
bp	Base pairs
BGF	Used in this study for Borgundfjord.
***c*.**	About, short for “circa” (Latin)
ºC	Degrees Celcius
cDNA	Complementary deoxyribonucleic acid
cm	Centimetre
cM	CentiMorgan
dl	Decilitre
DNA	Deoxyribonucleic acid
dph	Days post-hatch
e.g.	For example, short for “exempli gratia” in Latin
EST	Expressed sequence tag
***et al*.**	And others, short for “et alii” (Latin)
EU	European Union
FAO	Food and Agriculture Organization of the United Nations
FAP	Family assignment program
FB	Used in this study to describe Faroe Bank cod.
F_{MSY}	The fishing mortality at maximum sustainable yield
FPNE	Used in this study for Faroe Plateau North-East
FPW	Used in this study Faroe Plateau West
g	Gram (or “gravity” in the contenxt of centrifugation)
GNP	Gross national product
H_e	Expected heterozygosity

H_o	Observed heterozygosity
HWE	Hardy-Weinberg equilibrium
IAM	Infinite allele model
i.e.	In other words, short for “id est” (Latin)
IPN	Infectious Pancreatic Necrosis
k	Number of alleles (microsatellites)
K	Fulton´s condition factor / number of clusters detected in STRUCTURE analysis
KASP™	Kompetitive allele specific PCR
kg	Kilogram
km	kilometre
l	Litre
LGM	Last glacial maximum
Ltd.	Limited
m	Metre
mg	Milligram
MHC	Major histocompatibility complex
Min	Minute
ml	Millilitre
mm	Millimetre
mM	Millimolar
MPA	Marine protected area
mRNA	Messenger ribonucleic acid
mtDNA	Mitochondrial DNA
N	Number
NCBI	National Center for Biotechnology Information
NCC	Norwegian coastal cod
NEAC	North-East Arctic cod
ng	Nanogram

NGS	Next generation sequencing
nm	Nanometre
nM	Nanomolar
NWB	Used in this study for North-West of Bømlo
P	Probability
PAR	Private allelic richness
PC	Personal computer
PCA	Principal component analysis
PCR	Polymerase Chain Reaction
Per se	By itself (Latin)
PIC	Polymorphism Information content
PIT	Passive integrated transponder
ppm	Parts per million
r	Range in allele size (microsatellites)
r^2	Coefficient of determination (correlation analysis)
RAD	Restriction-site associated DNA
RADseq	Restriction-site associated DNA sequencing
RCF	Relative centrifugal force
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
SD	Standard deviation
SMM	Stepwise mutation model
SNP	Single nucleotide polymorphism
SSR	Simple sequence repeat
STR	Short tandem repeat
t	Metric tonne
t	*t*-test value
TPM	Two-phase mutation model

µg	Microgram
UK	United Kingdom
µl	Microlitre
µm	Micrometre
USA	United Nations of America
UV	Ultra violet
V	Volt
VB	Used in this study for Verrabotn
WWF	World Wildlife Fund
WS	Used in this study for White Sea

Chapter 1. General introduction

1.1. Biology and distribution of the Atlantic cod

The Atlantic cod (*Gadus morhua*, Linnaeus 1758), belonging to the family Gadidae of the order Gadiformes, is a benthopelagic marine species which inhabits continental shelves and banks throughout the North Atlantic Ocean (Figure 1.1) (Cohen *et al*. 1990; Drinkwater 2005). It is mostly found within continental shelf areas from 150 – 200 m of depth, but can occupy a range of depths from the shoreline down to depths greater than 600 m. Cod can reach about 20 years of age and will usually grow up to a maximum of 1 m (fork length), although cod of 2 m length have been observed (Cohen *et al*. 1990).

As one of the most abundant species in the North Atlantic Ocean (Drinkwater 2005), the cod is an important fisheries resource throughout its distribution (Serchuk *et al*. 1994; Kurlansky 1997; Pope *et al*. 2008). Historically, it has had been the most important gadoid species from an economical as well as from a cultural view – there have even been wars fought over the cod (Jensen 1972; Kurlansky 1997; Lear 1998). Being a major predatory species (Macer and Easey 1988; Sparholt 1994; Björnsson *et al*. 2001), the cod plays a fundamental part in ecosystem stability (Hutchinson *et al*. 2001) and is one of the most studied marine species (Brander 1997; Hutchinson *et al*. 2001; Drinkwater 2008). A general decline in fisheries coupled with the success in salmonid aquaculture, has sparked interest in diversification of aquaculture species, in particularly of gadoids (Figure 1.2). Although the production of cultured cod has not reached the levels anticipated earlier this century (Rosenlund and Skretting 2006), Atlantic cod still has the potential to become an important aquaculture species in the future (Rosendal *et al*. 2013) and the importance of this species was recently demonstrated by the sequencing of its genome (Star *et al*. 2011).

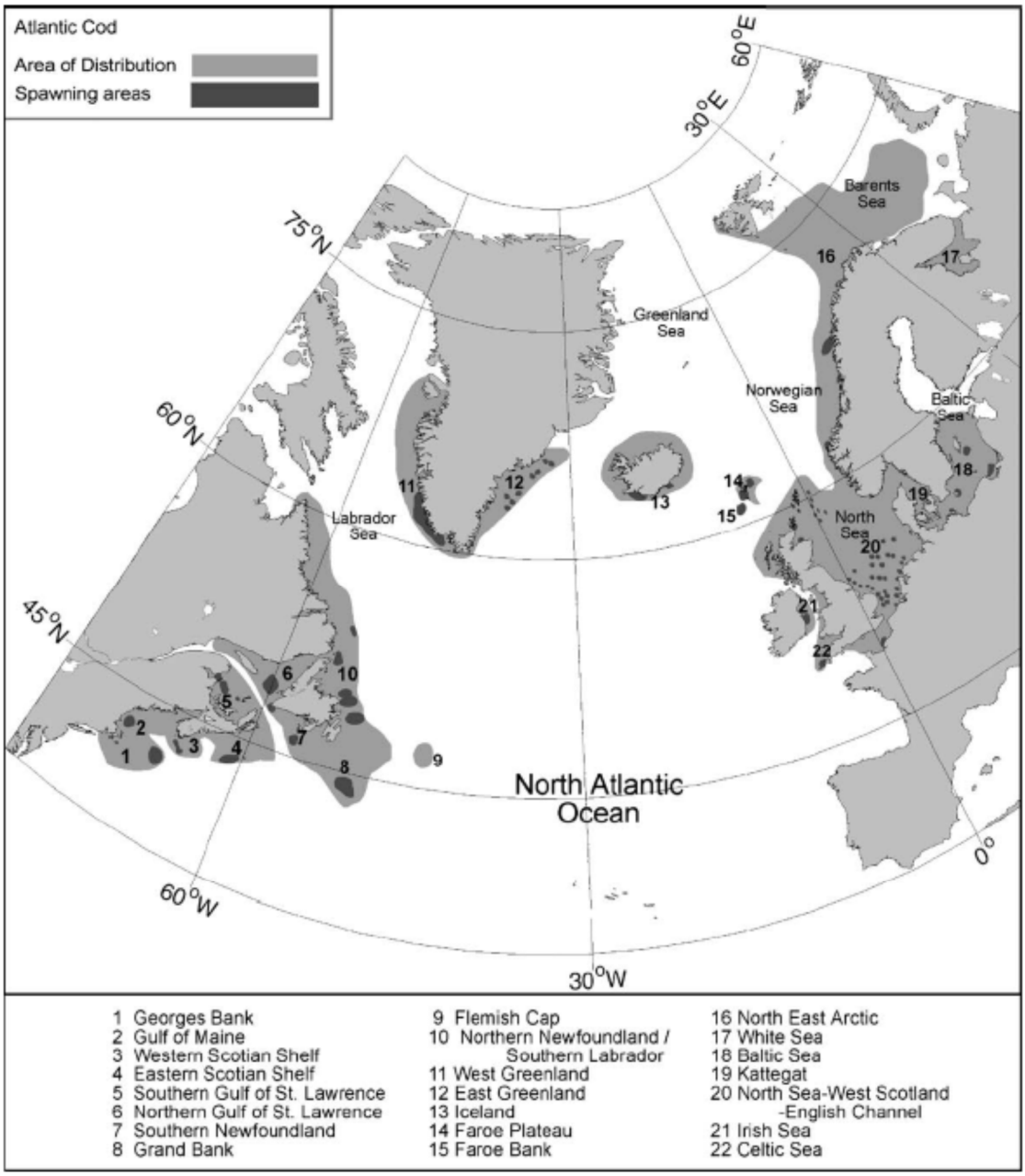

Figure 1.1. The present distribution of Atlantic cod (Drinkwater 2005; reproduced with permission from the ICES Journal of Marine Science).

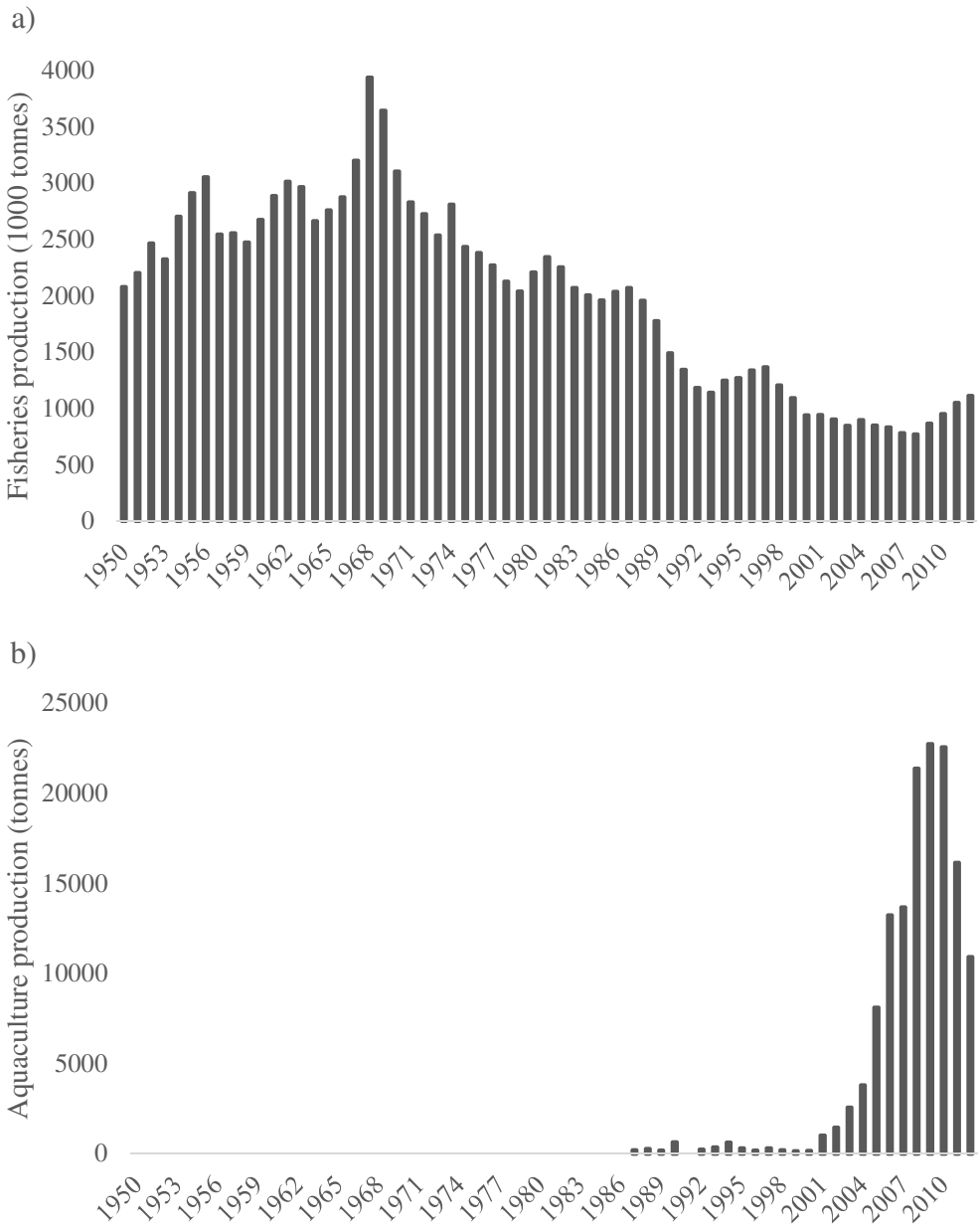

Figure 1.2. Worldwide capture (a) and aquaculture production (b) of Atlantic cod (www.FAO.org, accessed 19 May 2014).

1.1.1. Reproduction

Like most teleost fish, the Atlantic cod falls within the *r*-strategist group, characterised by a very high reproductive potential, a high growth rate and high mortality rates, at least during the early life stages (Codling *et al*. 2005). The cod matures at an early age, usually when it is 2 - 3 years old. During a single spawning season, the female typically releases *c*. 500,000 eggs per kg (Cohen *et al*. 1990). The cod is a determinate multiple batch spawner (Kjesbu 1989; Kjesbu *et al*. 1990),

determinate meaning that the potential fecundity, i.e. the standing stock of advanced or vitellogenic oocytes, is determined prior to the onset of the spawning period (Kjesbu *et al*. 1991; Gordo *et al*. 2008). Individual females may release up to 20 discrete egg batches (Kjesbu *et al*. 2006) over a spawning season that can last from 4 – 15 weeks (Kjesbu 1989; Kjesbu *et al*. 1991; Marteinsdóttir and Björnsson 1999; Kjesbu *et al*. 2006). Although exceptions exist (Gagne and O´Boyle 1984), spawning for most stocks takes place between January and June (Cohen *et al*. 1990) and timing of spawning is important in order to facilitate synchrony between larval development and optimal environmental conditions (Hjort 1914; Cushing 1975; Cury and Roy 1989; Mertz and Myers 1994). The eggs are small, usually in the range of 1.1 – 1.7 mm diameter (Knutsen and Tilseth 1985; Markle and Frost 1985), and are released directly into the water column with no parental care provided (Hutchings *et al*. 1999). This strategy of producing very large numbers of eggs that approximate the physiological minimum size for viability has been proposed to be an adaptive response to environments where egg sizes are of no consistent relevance for survival in the early life stages (Hutchings 1997). At the same time, the importance of studying egg sizes as being predictive of the size and energy reserves of young larvae and, thereby, proposed to play a role in larval survival, is widely accepted (Bagenal 1969; Knutsen and Tilseth 1985; Chambers 1993; Miller *et al*. 1995; Chambers and Waiwood 1996; Chambers 1997). Spawning behaviour is still not fully characterised or understood (Hutchings *et al*. 1999). Under natural conditions cod form spawning schools (Rose 1993; Morgan *et al*. 1997) where they release their gametes, which can stay viable for over an hour, and eggs are externally fertilised (Kjørsvik and Lønning 1983; Trippel and Morgan 1994). However, as reviewed by Nordeide and Folstad (2000), the cod has mistakenly been referred to as a promiscuous spawner. Cod appear to possess a conventional lekking mating system where females visit male aggregations to view their extensive courtship behaviours, which involve dancing, fin display and sound production. In a study on spawning cod in captivity, Hutchings *et al*. (1999) observed the dominant males defending territories near the bottom of the tank. There were two different types of male-female interaction: i) a “ventral mount”, where the male would use his pelvic fins to position himself beneath the female, so the urogenital openings were opposite one another and ii) the female

swimming to the bottom where she remained motionless while one or more males slowly circled around her. Moreover, they observed a positive relationship between male size and male dominance, which was also reflected in an increasing fertilisation success for increasing male size and (or) dominance, although the underlying genetic data for this were scarce (27 eggs from one batch). These authors also reported multiple paternity with unequal contribution for each of the males involved, an observation documented in subsequent studies (Rakitin *et al*. 2001; Bekkevold *et al*. 2002; Hansen *et al*. 2004; Herlin *et al*. 2008). Assortative mating has also been observed in cod in the sense that male reproductive success seems to depend on the magnitude of the male – female size difference, where a small positive male – female difference or zero difference resulted in the highest number of offspring sired, while low paternity success was found for males much larger or much smaller than the females (Bekkevold *et al*. 2002; Herlin *et al*. 2008).

Eggs, larvae and early juveniles are pelagic for approximately two and a half months after which time they settle as metamorphosed juveniles (Cohen *et al*. 1990; Brander 2005). Habitat complexity is important for the survival of early life stages (Anderson *et al*. 2007) and, typically, juveniles have been observed to settle in complex seabed habitats (cobble/rock, macro-algae or other structures) that act to provide protection against predators (for example Gotceitas *et al*. 1995; Borg *et al*. 1997). Larval sizes at hatch may vary from 3 to 7 mm, sizes at metamorphosis may vary from 12 – 15 mm and sizes at settlement from 4 to 10 cm (Brander 2005). For cod eggs, there is a negative relationship between time to hatch and water temperatures, for example at 5 ºC the average predicted time to 50 % hatch is 18 days (Geffen *et al*. 2006). However, other factors, such as population-specific adaptation, parental effects as well as egg size and quality may cause variations in this relationship (Geffen *et al*. 2006, 2012). Cod are voracious and omnivorous and their diet changes as they grow. Larvae mainly feed on plankton, juveniles on invertebrates, in particular small crustaceans, whereas older individuals consume more fish, including conspecifics (Cohen 1990).

1.1.2. Migration patterns

Similar to many other commercially important fish species (Harden-Jones 1968; Secor 2005), some cod stocks travel long distances between their respective nursery, feeding and spawning areas (Godø 1984; Joensen *et al*. 2005). Generally, the larvae will drift with the currents from the spawning area to the nursery area and, when large enough, the juvenile cod migrate to the adult feeding areas, closing the maturation cycle (Cohen *et al*. 1990; Joensen *et al*. 2005). Migration patterns are not similar for all cod stocks (Robichaud and Rose 2004). For Southern Gulf cod off Canada the feeding season is followed by an offshore overwintering period (McKenzie 1941; Halliday and Pinhorn 1982). Also, the nursery or the spawning areas may be interchangeable with the feeding areas, such as for the Barents sea cod (Ottersen and Sundby 2005) and the Faroe Plateau cod, respectively, or the cod may show more or less sedentary behaviour (Joensen *et al*. 2005). After reviewing tagging experiments on 174 groups of Atlantic cod, carried out over an entire century, Robichaud and Rose (2004) partitioned cod into four different behaviours with respect to migration strategies; sedentary cod, accurate homers, inaccurate homers and dispersers, the sedentary group being the largest group (41 %). Sedentary cod were found year-round within a relatively confined area, a type case being the Faroe Bank cod, whose eggs and larvae are kept on the bank by a strong oceanographic gyre. Migration and dispersal seem to be linked to historical abundance, as lower maximum historical biomasses were found for the sedentary compared to other groups. Accurate homers (such as the Faroe Plateau cod) complete seasonal migrations and are believed to home to a comparatively small area, whereas inaccurate homers return to a much wider area. Groups of cod found to move and spawn in large areas in somewhat random patterns were categorised as dispersers. Migration patterns have been postulated to be linked to oceanographic features, the first two strategies being more plentiful among the North-East Atlantic groups due to relatively stable oceanographic conditions whereas more unpredictable environments in the North-West Atlantic were characterised by having more dispersing groups. Cod are sometimes grouped as either coastal or offshore (shelf) cod, and these two groups were not found to differ from each other with regards to the relative frequencies of the four migratory behaviours (Robichaud and Rose 2004).

1.1.3. Occupation of a range of different environments

Atlantic cod inhabit areas with a diverse range of environmental conditions (Brander 1995; Planque and Frédou 1999) and can be found at temperatures between -1 ºC and 20 ºC (Cohen *et al*. 1990), although mostly between 0 ºC and 12 ºC (Drinkwater 2005). Concerning cod abundance and distribution, temperature is the most studied environmental variable (Brander 2005) and it is believed to be the primary regulator of growth rate (Brander 1994; Imsland and Jónsdóttir 2002), although genetic factors may also be involved (Nævdal *et al*. 1992; Imsland *et al*. 2004). Dutil and Brander (2003) compared stock assessments over the entire distributional range of Atlantic cod and found that the stocks divided into four clusters representing four levels of production. The most productive stocks were characterised by higher bottom salinities and temperatures and were found in the North-East Atlantic (the Celtic Sea, Irish Sea and West Scotland cod) (Dutil and Brander 2003) where the oceanographic environment is more stable than in the North-West Atlantic (Robichaud and Rose 2004). These findings may seem contradictory to the findings of Robichaud and Rose (2004), where the sedentary groups of low historical maximum biomasses were found in the North-East Atlantic. However, these high historical biomasses were a result of range expansions, not increased densities.

It appears that cod have the potential to adapt to unstable environments by various means. The strategy of an overwintering period for the Southern Gulf cod is thought to be a means of avoiding temperature levels below a threshold limit (Swain *et al*. 1998). A different strategy to cope with low temperatures is the build-up of plasma antifreeze glycoproteins (Harden-Jones and Scholes 1974) as demonstrated in the Gilbert Bay cod, Labrador (Ruzzante *et al*. 2000a). Furthermore, cod stocks from different regions respond differently to changing temperatures. Whereas in the cold Northern regions recruitment has been found to be enhanced by increasing temperatures the opposite appears to be true in Southern warmer waters with no relationship found in waters of intermediate temperatures (Ottersen 1996; Planque and Frédou 1999).

1.2. The exploitation and management of fish stocks

1.2.1. Cod fisheries

Cod fisheries started around the 10th century (Gallagher 2003) and played a vital role in the development of European economies and societies during late-medieval and early modern times (Holm *et al*. 1996). In the Faroe Islands, the cod fisheries were of particular importance for both economic and social development, and were instrumental in the conversion of a collective farming society to a nation of seamen. Here, the fisheries could sustain far more people than the traditional farming, which meant a steady population increase since the end of the 18th century (Joensen 1996).

In the whole North Atlantic region, technological improvements led to an increased fishing pressure on all commercially important fish stocks, including most cod stocks, from the 19th century onwards (Marteinsdóttir *et al*. 2005) and, consequently, many marine fish species experienced dramatic declines in population sizes. An investigation of data from over 230 marine fish populations revealed a median reduction of 83 % in breeding population sizes compared to historical levels and this is most likely an underestimation, since historical levels were based on the earliest data collected which may not reflect the true historical levels. Seventy populations were investigated within the Gadidae, for example Atlantic cod and haddock (*Melanogrammus aeglefinus*), and more than half of these populations declined by 80 % or more (Hutchings and Reynolds 2004). Perhaps the best-known example of a population decrease comes from cod in Canadian waters. For the northern stock complex, once one of the world´s most abundant cod stocks, numbers of cod decreased an estimated 99.9 % from the early nineteen sixties to the early nineteen nineties (Hutchings 2004).

Fishing gear is selective, often targeting the largest individuals, and when some of the within-species phenotypic variation is genetic in origin, as is often seen for growth rate, length- and age-at-maturation and fecundity, it may have the potential to cause evolutionary changes (Law 2000; van Wijk *et al*. 2013). There are numerous examples of detrimental changes in growth and maturation characters of heavily exploited fish stocks, such as a decrease in

length-at-maturation for North Sea cod (Oosthuizen and Daan 1974; Rowell 1993) and a decrease in age-at-maturation for North-East Arctic cod (Borisov 1978; Jørgensen 1990). In addition, overexploitation of marine fishes may cause changes in species distribution (Csirke 1980; Winters and Wheeler 1985; Swain and Wade 1993; Marshall and Frank 1994) as well as a reduction in size and number of spawning areas, as has been seen for cod and herring (*Clupea harengus*) (Frank *et al*. 1994; Stephenson 2002).

1.2.2. Fisheries management

A high proportion of the world´s fish stocks need rebuilding, perhaps as high as 63 % (Murawski 2010). Stock recovery is a long process, as data from 56 populations showed: even when fishing mortalities were reduced, if populations had experienced a decline of 60 % or more over the last 15 years, little or no recovery was seen 15 years later (Hutchings and Reynolds 2004). The most successful stock recoveries have been observed when immediate and substantial reductions in fishing mortality have been implemented, as opposed to incremental reductions over a long period of time (Murawski 2010). For example, following decades of overfishing, some believed that Georges Bank haddock (*Melanogrammus aeglefinus*) was near to a collapse in the early nineteen nineties and following a lawsuit by an environmental organisation, the fishery management body was forced to eliminate overfishing of haddock and other groundfish species in the New England area. The restrictions implemented included year-round closed areas, increases in trawl mesh size and considerable reductions in fishing mortality. From 1980 to 1993, the fishing mortality had been 0.35 $year^{-1}$ and it decreased to *c*. 0.17 $year^{-1}$, or approximately 30 % of the F_{MSY} (fishing mortality that produces the maximum sustainable yield) from 1994 onwards. The response was dramatic, whereas the spawning stock biomass decreased from over 67,400 t in 1980 to 14,600 t in 1993, it increased again to over 115,000 t in 2003-2004 (Brodziak *et al*. 2008). Most often, however, there is a tendency to increase the fishing effort at the smallest of signs of recovery (Hutchings and Reynolds 2004; Murawski 2010). For instance, the Canadian government closed the fishery for the northern cod off Newfoundland in the early nineties, but reopened it in 1999,

even though the size of the stock had remained historically low, only to close it again in 2003 (Hutchings and Reynolds 2004; Olsen *et al*. 2004).

Stock recovery depends on a variety of factors, apart from fishing mortality, for example life history strategies, species interactions, genetic / evolutionary causes and habitat modification. Repeated bottom trawling has many negative impacts on biodiversity and marine ecosystems. For instance, it may significantly alter bottom morphology, resulting in less refugia for bottom-dwelling marine species. This may cause slower recovery rates for a species like cod compared to pelagic species, such as herring (Hutchings and Reynolds 2004; Pusceddu *et al*. 2014). Concerning the collapse of the northern cod stock complex off Newfoundland, the collapse of its main food, the capelin (*Mallotus villosus*), over a similar period, has often been overlooked. A recent study suggests that the resulting poor diet is the limiting factor for productivity of this cod stock and that a recovery of the cod stock depends on the recovery of capelin in the area (Mullowney and Rose 2014). The abrubt biomass decline of the capelin stock in 1991 was followed by i) a protracted and delayed spawning period; ii) a decrease in size and age at maturity and iii) a decrease in somatic condition. The stock has not yet recovered. Through food availability, capelin population dynamics appears to be best explained by physical ecosystem processes, such as seasonal sea ice dynamics (Buren *et al*. 2014).

Stock assessments may sometimes be complicated by the fact that administrative areas do not necessarily reflect the habitats of individual populations. Hence, what is traditionally managed as a single stock may consist of several smaller sub stocks with limited interbreeding; this has been observed for several cod stocks (Smedbol and Stephenson 2001; Sterner 2007). When a group of distinct sub stocks are being managed as one large panmictic population the result will likely be an overestimation of growth and harvest potential, which can in turn lead to over-exploitation and collapse of fisheries followed by unexpectedly slow recovery rates (Sterner 2007). While population substructuring is not easily detected by traditional methodologies, it can be studied by the use of appropriate molecular genetics tools (Smedbol and Stephenson 2001; Sterner 2007). Genetic studies may aid traditional fisheries management in a number of ways, for example by

providing knowledge on population connectivity, the spatial and temporal scale of population differentiation, effective breeding population sizes, fishery induced changes and adaptation to local habitats (Hauser and Carvalho 2008). An outline of molecular tools that can be used to study fish stocks is provided below.

1.3. Molecular genetic tools used to study fish stocks

Molecular tools have been increasingly applied to fisheries management during the past six decades (Waples *et al.* 2008), and have markedly increased and changed our understanding of marine ecology (Hauser and Carvalho 2008). The basis of these tools is the ability to detect and quantify, either directly or indirectly, inherited mutations in the DNA of organisms that can be passed from generation to generation. Mutation, coupled with selection, genetic drift and migration generate genetic variation among individuals, populations, species and higher order taxonomic groups (Liu and Cordes 2004). The major discoveries made from molecular genetic studies of marine fish populations can be grouped into three categories: i) the presence of extensive population structure within marine species, despite their wide distribution and great dispersal abilities, ii) indications of extensive adaptive variation among populations that might make them more resilient to environmental changes and iii) relatively small effective population sizes, two to six orders of magnitude smaller than census population sizes (Hauser and Carvalho 2008).

Genetic markers can be classified into three major groups: protein variants (i.e. allozymes, hemoglobin and transferrin; these were initially studied as proteins in gels, but can now be assessed as DNA markers), DNA sequence polymorphisms and DNA repeat variation (Schlötterer 2004). The properties and application of the main genetic marker types that have been used to study cod populations are discussed below.

1.3.1. Allozymes

Allozymes are variant forms of an enzyme that are coded by different alleles at the same locus. In some instances, these variants can be distinguished electrophoretically, to provide co-dominant

genetic markers (for example Falconer and Mackay 1996a; Magoulas 1998). This methodology, whereby soluble protein products can be separated by differential mobility through a solid matrix (for example starch, polyacrylamide or cellulose acetate) (Dunham 2004), with subsequent detection of enzymatic activity by specific histochemical stains was established in the 1960s (Magoulas 1998). It provided, for the first time, a practical means to indirectly identify and quantify genetic variation at multiple loci in a wide range of organisms (Schlötterer 2004). For the following two decades protein electrophoresis was extensively applied to studies of molecular variation (Magoulas 1998), one of the major application areas being population studies (Liu and Cordes 2004). Despite its widespread use, the technology does have its drawbacks. The different protein variants originate from amino acid substitutions causing changes in the electric charge of the proteins and since only 5 of the 20 amino acids are charged only about 25 % of the amino acid differences can be detected (Falconer and Mackay 1996a). Furthermore, due to the redundancy within the genetic code many other base substitutions (mutations) are not detectable by this means (Schlötterer 2004). Consequently, only a limited number of isozyme loci tend to exhibit detectable variants (Magoulas 1998). The need for fresh or freshly frozen tissue samples limits its use in some cases, as does the size of tissue sample required, making it difficult to apply to small organisms or immature stages. Hence, the search for more informative markers continued (Liu and Cordes 2004).

1.3.2. Hemoglobin

One of the earliest electrophoretic studies on fish was that of Sick (1961), who described different variants of the oxygen-carrying protein hemoglobin in whiting (*Merlangius merlangus*) and Atlantic cod (Ward 2000). Cod hemoglobins separated into two major components in agar gel electrophoresis, HbI and HbII. Furthermore, the HbI component was observed to separate into three different variants, HbI-1/1, HbI-1/2 and HbI-2/2. The discovery of this two – allele system opened a new era of hemoglobin studies on cod (Frydenberg *et al*. 1965; Sick 1965a; Sick 1965b; Mørk *et al*. 1983; Mørk *et al*. 1984; Fyhn *et al*. 1994; Fyhn *et al*. 1995; Petersen and Steffensen 2003). Allele distributions were found to be strongly influenced by water temperatures, with the

HbI-1 allele dominating in warmer regions (North Sea, southern part of the Norwegian coast and the western part of the Baltic Sea) and the HbI-2 allele mainly found in colder waters (Greenland, Iceland, Canada, northern Norway and the northern part of the Baltic Sea) (Frydenberg *et al.* 1965; Sick 1965a; Sick 1965b; Petersen and Steffensen 2003). Different temperature effects on the hemoglobin – oxygen dissociation curves for the three hemoglobin types provided a possible physiological explanation for the heterogeneous allele distributions. Oxygen affinity for HbI-2/2 individuals was greatest at low temperatures (< 10 ºC), while for HbI-1/1 individuals it was best at high temperatures (> 14 ºC) (Karpov and Novikov 1981, Brix *et al.* 1998; Brix *et al.* 2004). Oxygen affinity values of heterozygous individuals were intermediate between the two homozygous values (Karpov and Novikov 1981). Moreover, of the three hemoglobin forms, individuals with the HbI-2/2 genotype displayed the highest growth rate and competitive performance (Nævdal *et al.* 1992; Salvanes and Hart 2000; Imsland *et al.* 2004). Imsland and Jónsdóttir (2002) hypothesised that the relationship between hemoglobin polymorphisms, oxygen binding capacity and growth could be one where higher oxygen binding capacity resulted in lower metabolic costs, thus leaving more energy available for growth.

It was only recently, however, that researchers managed to reveal the genetic basis of the observed Atlantic cod hemoglobin polymorphisms. Hemoglobins in most fish consist of four globin chains, two α and two ß chains (de Souza and Bonilla-Rodriguez 2007). Nine distinct hemoglobin genes have been discovered for Atlantic cod, four belonging to the *α*- and five to the *ß*-hemoglobin gene family. All nine genes are expressed simultaneously in adult fish, but expression levels vary among different life stages. Adult fish predominantly express the *α1 – α2* and *ß1 – ß4* genes, whereas the *α4* and *ß5* genes are the main genes expressed in early life stages. The role of the *α3* gene is yet to be uncovered (Borza *et al.* 2009). Of importance, researchers also managed to create a link to previously analysed hemoglobin protein variants. By comparing genotypes and phenotypes of 35 juvenile fish caught off Bergen, Andersen *et al.* (2009) established an unambiguous association between polymorphisms in the *ß1* gene and traditional hemoglobin electrophoretic phenotypes. Nonsynonymous substitutions at two linked positions (pos55 and

pos62) in the *ß1* gene resulted in the two non-recombinant alleles Met55 – Lys62 and Val55 – Ala62, representing the previously identified HbI-1 and HbI-2 protein alleles. By different model building methods, the replacements of Met by Val at pos55 and of Lys by Ala at pos62 were found to change the hemoglobin quaternary structure and electrostatic feature, resulting in a better access for oxygen molecules. Alluringly, the same molecular mechanism for facilitating oxygen binding was found to have evolved independently in avian species adapted to high-altitude, low oxygen habitats (Andersen *et al*. 2009). A second study revealed a similar association between *ß1* gene polymorphisms and protein phenotypes, although this second study especially highlighted the importance of the substitution at position 62, which is also found in other fish species, such as rainbow trout (*Oncorhynchus mykiss*) and eel (*Anguilla anguilla*). In addition, other polymorphisms were present within the two major allelic classes, providing a plausible explanation for former protein studies not always being able to establish a clear correlation between hemoglobin types and oxygen and temperature sensitivity as well as growth (Borza *et al*. 2009).

1.3.3. Transferrin

Transferrin has also been used as an informative genetic marker in early electrophoretic studies of cod population structure, although not to the extent that hemoglobins were used (Møller 1968; Jamieson 1967; Jamieson and Jones 1967). Transferrin belongs to the *ß*-globulin family and consists of two homologous domains, forming a cleft for binding iron (Denovan-Wright *et al*. 1996). The iron-binding function is important in the immune response to pathogens by limiting the availability of free iron (Ong *et al*. 2006; Ganz 2009) and transferrin is also believed to play an iron-independent role in the immune system (Macedo *et al*. 2004). Accordingly, a metastudy of 25 vertebrate species found evidence for positive selection acting on transferrin, although, curiously, this was limited to the salmonid species included (Atlantic cod data were not included in this study). For the most part, selected sites were on the outside of the molecule, in relation to areas believed to be targeted by transferrin-binding bacterial proteins (Ford *et al*. 1999). Furthermore, a recent study showed a significant association between transferrin polymorphisms

and resistance of common carp (*Cyprinus carpio* L.) to a certain blood parasite (*Trypanoplasma borreli*), although they did not describe the exact mechanism by which this resistance was conferred (Jurecka *et al*. 2009).

As mentioned, transferrin has not been much studied in Atlantic cod, but a partial transferrin cDNA sequence was isolated from cod two decades ago and cod were found to express transferrin in a higher number of tissue types than other species (Denovan-Wright *et al*. 1996). It was only recently, however, that the full sequence of this cod transferrin gene (*Tf1*) was described together with a second transferrin gene (*Tf2*), a melanotransferrin gene (*MTf*) and a monolobal transferrin gene (*Omp*). Alignment analysis of a Northeast versus a North-West Atlantic version of cod *Tf1* revealed a ratio of nonsynonymous to synonymous substitutions of >> 1 (indicative of positive selection) with several nonsynonymous substitutions involving surface residue changes. The results were further substantiated by a population genetic study of six of the identified SNPs showing highly divergent allele frequencies for 14 North Atlantic cod populations with levels of differentiation significantly higher than those found in other studies using neutral genetic markers. Thus, it appears that this locus is influenced by positive selection for advantageous alleles, possibly caused by evolutionary battles between cod and their pathogens for transferrin-bound iron (Andersen *et al*. 2011).

1.3.4. Nuclear DNA restriction fragment length polymorphism (RFLP)

Whereas the early studies discussed above dealt with protein variations, the discovery of restriction endonuclease enzymes in the 1960s made it possible to study genetic variation at the DNA level, i.e. DNA sequence polymorphisms (Schlötterer 2004). Restriction endonuclease enzymes are used to cut the DNA at restriction sites and genomic changes at or between these sites cause different banding patterns when run through an agarose gel and visualised by fluorescent staining or following Southern blotting / hybridisation of labelled complementary probes. This method provided co-dominant genetic markers that were easy to score and the first DNA based genetic maps were constructed from RFLPs (Dunham 2004; Schlötterer 2004).

A study by Pogson *et al*. (1995) demonstrated very well the advantages of investigating DNA variation directly, as opposed to investigating variation at allozyme loci. Six populations of Atlantic cod that had previously been investigated at 11 allozyme loci (Mørk *et al*. 1985) were investigated at 17 RFLP loci and, in contrast to findings from the allozyme loci, significant differentiation was detected among all populations with the RFLP loci.

1.3.5. Pantophysin (*Pan*I)

One of the most informative outcomes of RFLP analysis in cod stock studies was the identification and characterisation of the polymorphic pantophysin (*Pan*I) locus during the mid-nineties (Pogson *et al*. 1995). The locus was first believed to be part of the synaptophysin gene (*Syp* I) (Fevolden and Pogson 1997) but later revised to reside in the pantophysin I gene (Pogson *et al*. 2001), encoding an integral membrane protein expressed in cytoplasmic transport vesicles (Fernández-Chacon and Südhof 1999). The locus has been linked to growth properties of cod populations (Fevolden and Pogson 1995; Jónsdóttir *et al*. 2002; Case *et al*. 2006), as well as to temperature, salinity and depth (Case *et al*. 2005; Pampoulie *et al*. 2006, 2008b). The *Pan*I locus in cod and other gadoids appears to be influenced by positive selection (Pogson and Mesa 2004), although this could also be through selection acting on nearby loci (Fevolden and Pogson 1997; Case *et al*. 2006).

Of the two alleles (*Pan*I^{A} and *Pan*I^{B}) identified at the *Pan*I locus, *Pan*I^{A} has been associated with larger size at age in wild cod (Jónsdóttir *et al*. 2002). Sarvas and Fevolden (2005) investigated *Pan*I allele frequencies in north-east Arctic and Norwegian coastal cod sampled over nine years and found a latitudinal cline among post-juvenile individuals, with the *Pan*I^{B} allele dominating in the Barents Sea with frequencies ≥ 0.87 and the *Pan*I^{A} allele dominating the southernmost coastal ares with frequencies ≥ 0.73. This correlated well with a higher growth rate demonstrated for Norwegian coastal cod compared to North-East Arctic cod (Van der Meeren *et al*. 1994; Svåsand *et al*. 1996). Under controlled laboratory conditions, Johnston and Andersen (2008) revealed significantly higher numbers of fast muscle fibres in *Pan*I^{BB} than in *Pan*I^{AB} genotypes of North-East Arctic cod.

1.3.6. Mitochondrial DNA (mtDNA)

Mitochondrial DNA (mtDNA) is a circular, covalently closed, double-stranded DNA molecule. Due to its abundance (many thousands of copies per cell) and relative ease of extraction / purification, mtDNA was an initial target for systematic screening for DNA polymorphisms. The molecule is generally maternally inherited as a single "allele" / haplotype – thus the effective population sizes of alleles is ¼ that of nuclear DNA and, since the replication system lacks repair mechanisms, the mutation rate of mtDNA is *c*. one order of magnitude higher than that of nuclear DNA. This combination of attributes has led to extensive studies of mtDNA sequence variation across all organisms, with particular emphasis in the field of population genetics (Avise *et al*. 1987; Dunham 2004; Liu and Cordes 2004). Three types of polymorphisms have been detected: length polymorphisms, restriction-site polymorphisms and heteroplasmy, which can occur due to paternal leakage, the rare inheritance of paternal mtDNA (Dunham 2004).

mtDNA variation has been much studied in cod, for example Dahle (1991) investigated mtDNA variation in Norwegian cod by RFLP analysis and found significant divergence between Arctic and Coastal Norwegian cod. Carr and Marshall (1991a,b) demonstrated the usefulness of the polymerase chain reaction (PCR) to study fish populations when they sequenced a portion of the mitochondrial cytochrome *b* gene and identified several polymorphisms that showed genetic differentiation among Norwegian and Newfoundland cod stocks (Carr and Marshall 1991a) as well as between Greenland cod (*Gadus ogac*) and other Northwestern Atlantic cod (Carr and Marshall 1991b). A few years later, the cod mitochondrial genome was fully sequenced and found to be 16,696 bp in length (Johansen and Bakke 1996). mtDNA variation in cod appears to be mainly influenced by neutral processes and thus is considered to be well suited for studies of population differentiation (Árnason and Pálsson 1996; Árnason *et al*. 1998). Árnason (2004) investigated mtDNA variation in cod from the entire distribution range and found the Newfoundland as well as the Baltic samples to be the most divergent ones whereas there was zero genetic differentiation between cod from the Faroe Plateau and cod from Iceland and Norway, respectively. No Faroe Bank cod were included in this study. Another study included North and

South samples of Faroe Plateau cod and detected no significant differentiation between the areas (Sigurgíslason and Árnason 2003).

1.3.7. Microsatellites

Microsatellites, also called short tandem repeats (STRs) or simple sequence repeats (SSRs), are relatively short stretches of tandemly repeated DNA units, with 1 – 6 bp motifs (Chistiakov *et al*. 2006; Fan and Chu 2007). The length of repeats at a locus can vary, and given that the entire repeat region of a locus is generally less than 300 bp in length, this size polymorphism can be readily analysed by a combination of PCR technology and electrophoretic size separation of alleles (Schlötterer 2004). Microsatellite loci can be highly polymorphic with mutation rate estimates in the range of 10^{-6} – 10^{-2} per locus per generation, compared to a mutation rate of 10^{-9} per locus per generation for nonrepetitive DNA (Fan and Chu 2007). This hypervariability, together with their codominant nature and straightforward PCR-based detection, has resulted in microsatellites becoming the predominant genetic marker system exploited in a wide range of life science studies over the past 25 years (Schlötterer 2004; Chistiakov *et al*. 2006), including studies of population structure in cod and other marine fish (O'Leary *et al*. 2007; Hauser and Seeb 2008). Concerning studies of cod, the discriminatory power of highly variable microsatellite markers was highlighted in a review of studies by O'Leary *et al*. (2007), in which it was observed that whereas many of the previously applied marker systems did not detect significant population differentiation, all microsatellite studies reported did so.

1.3.7.1. Mutational mechanisms

It is 40 years since the first discovery of microsatellites, but the precise mutational mechanisms underlying the observed polymorphisms are still uncertain. However, the main mutational mechanism appears to be strand-slippage replication, during which a nascent strand is created with a different number of repeats compared to the template strand. This happens very often in the genome, but is controlled by the mismatch repair system, thus, the microsatellite mutation rate is decided by both the rate of DNA slippage in the genome as well as the efficiency of the repair system. Two main models have been used to describe microsatellite mutations; the infinite

alleles model (IAM) and the stepwise mutation model (SMM). The former states that a mutation can involve any number of repeats units and the result is always a new allelic state, whereas in the latter model, alleles can mutate up or down by one or a few repeat units, in agreement with the strand-slippage replication mechanism. Although the SMM is thought to incorporate real mutation processes better than the IAM (Fan and Chu 2007), the IAM is still widely used as the default model, since calculations based on the SMM might be highly sensitive to model violations. More realistic (and complex) variations of the SMM have been developed (Selkoe and Toonen 2006), such as the two-phase mutation model (TPM) which is based on the SMM, but also allows larger mutations to occur (Di Rienzo *et al*. 1994).

1.3.7.2. Technical challenges

The challenge in using microsatellites lies mainly in their potential susceptibility to genotyping errors, which may result in altered allele frequencies, Hardy-Weinberg deviations, inbreeding overestimates, difficulties in estimating population size and the inability to differentiate between individuals and analyse parentage (Taberlet *et al*. 1996, 1999; Miller *et al*. 2002; Xu *et al*. 2002; Creel *et al*. 2003; Paetkau 2003; Hoffman and Amos 2005). In addition to genotyping errors originating from human handling or equipment errors, there are a number of DNA-based errors, which can be divided into three forms, size homoplasy, allelic dropout and null alleles (Soulsbury *et al*. 2007).

In a similar manner as described above, strand-slippage can also occur during the PCR reaction, resulting in stutter products that differ from the original template by multiples of the repeat unit length (Shinde *et al*. 2003). This may potentially make the interpretation of results difficult (Selkoe and Toonen 2006).

Allelic dropout is the stochastic failure of one of the two alleles at a heterozygous locus to amplify (Taberlet *et al*. 1996). There are different mechanisms leading to allelic dropout, such as low quality or low quantity of DNA which may result in the preferential amplification of the short allele, causing long allele dropout. Also, one allele may be favoured because of its propensity to

denature (because of low GC content) (Pompanon *et al*. 2005). Some loci and some pairwise combinations of alleles are more prone to allelic dropout (Soulsbury *et al*. 2007).

The most commonly reported source of genotyping error in microsatellite studies is the occurrence of null alleles where some alleles fail to amplify due to unidentified polymorphisms at primer binding sites (Callen *et al*. 1993; O´Reilly and Wright 1995; Paetkau and Strobeck 1995). Often, such null alleles go undiscovered, but they may be identified by unexpected heterozygote deficiency (O´Reilly and Wright 1995) or by specific pedigree analyses.

1.3.8. Single nucleotide polymorphisms (SNPs)

SNPs are point mutations that give rise to different alleles that contain alternative bases at a given nucleotide position within a locus (Liu 2007). Such polymorphisms have been characterised since the beginning of DNA sequencing in 1977, but it took two decades before a greatly increased sequencing capacity allowed large numbers of SNPs to be identified and appropriate gene chip technology to be developed, making it feasible to analyse large sets of SNP data (Liu and Cordes 2004). SNPs can reveal hidden polymorphisms not identified by other marker systems and, therefore, provide the ultimate genetic resolution between individuals (Liu 2007). Candidate polymorphisms underlying important traits are thought to be abundant in SNPs (Jalving *et al*. 2004) and these markers represent valuable tools for studying the genetic basis of multifactorial diseases (Liu 2007; Chelala *et al*. 2008), for the construction of high density genetic maps and for genome-wide association studies (Chelala *et al*. 2008). Importantly, SNPs are readily adaptable to automation. Unlike microsatellites or allozymes, where relatively few methods of screening exist, there are a variety of technologies to choose from when genotyping SNPs (Liu 2007). Relatively affordable options exist, such as quantitative PCR, but large-scale analysis of SNPs still depends on expensive, cutting-edge equipment (Liu and Cordes 2004; Liu 2007; Kumar *et al*. 2012).

1.3.8.1. SNPs useful for delineating population structure

Genome-wide SNP discoveries are relatively new for Atlantic cod, for example Moen *et al*. (2008) validated 318 SNP loci in North-East Arctic cod (NEAC) and Norwegian coastal cod (NCC) cod and Hubert *et al*. (2010) validated 1620 SNP loci in North-West Atlantic cod populations. Both used expressed sequence tags (ESTs) to develop the SNPs. The usefulness of these resources for delineating population structure are exemplified by Heath *et al*. (2014). Whereas microsatellite studies had provided unclear results regarding the population structure of cod in UK waters, these authors used 96 of the developed SNP loci to define the geographic limits of three population units of cod in UK waters.

1.3.8.2. Genomic islands of divergence

Bradbury *et al*. (2013) investigated 1405 SNPs in 466 cod spanning the whole distribution range and found elevated divergence at 5.2 % of the SNPs, suggesting directional selection at one third of the linkage groups. Hemmer-Hansen *et al*. (2013) identified a genomic region, > 20 cM, which showed particularly high levels of differentiation between pairs of migratory and stationary cod populations. Lastly, Karlsen *et al*. (2013) also investigated migratory and stationary cod and found peaks of particularly high F_{ST} values at 25 of the 6647 genomic scaffolds of the cod genome. They estimated the average F_{ST} value of the neutral parts of the genome to be 0.0778, although other genome-wide studies have detected lower average neutral F_{ST} values, for example Hemmer-Hansen *et al*. (2013) found an average neutral F_{ST} value of 0.0024. Moreover, these studies support the existence of "genomic islands of adaptive divergence", i.e. genomic regions yielding greater differentiation estimates than expected under neutral assumptions, in the cod genome, a phenomenon also observed in other species (Nosil *et al*. 2009).

1.3.8.3. Restriction-site associated DNA sequencing (RADseq)

Of the studies mentioned above (Section 1.3.8.2), Karlsen *et al*. (2013) sequenced pooled population samples, a cost-effective approach to investigate genomic divergence of populations by next-generation sequencing. Another approach to circumvent the high costs associated with genome-wide discovery and genotyping of SNPs is restriction-site associated DNA sequencing

(RADseq). This method provides high-resolution population genomic data at randomly sampled sites of the genome, decided by the restriction enzymes used, and can even be applied on species with no genetic data available (Davey and Blaxter 2011). This method was for instance used to identify high levels of genomic divergence between closely located freshwater and oceanic populations of threespine stickleback (*Gasterosteus aculeatus*). Again, some genomic regions showed particularly high levels of divergence and they were found to contain several genes coding phenotypes previously suggested be under selection in freshwater populations (Hohenlohe *et al*. 2010).

1.3.9. Whole genome sequencing efforts

Aquaculture is largely responsible for genome mapping efforts in non-model fish, where whole-genome maps are expected to aid in the understanding of genes underlying economically important traits, such as growth performance and disease resistance (Alcivar-Warren *et al*. 1997). Interest in genome mapping of important aquaculture species began more than a decade ago with the First Aquaculture Species Genome Mapping Workshop where five groups of aquaculture species were identified as targets: salmonids, tilapia, catfish, shrimp and oyster (Alcivar-Warren *et al*. 1997). Since then, a number of genetic linkage maps, largely based on microsatellite loci, have been available for a number of aquaculture species, such as channel catfish (*Ictalurus punctatus*), rainbow trout, Atlantic salmon (*Salmo salar*) and Nile tilapia (*Oreochromis niloticus* L.) (Danzmann and Gharbi 2007). However, the ultimate genetic map is the whole genome sequence and, recently, whole genomic sequences have become available for a number of fish species, including the Atlantic cod (Star *et al*. 2011). Other fish species with their genome sequenced are the blind cave fish (*Astyanax maxicanus*), the Coelacanth (*Latimeria chalumnae*), the tiger puffer fish (*Takifugu rubripes*) (fugu), the Japanese medaka (*Oryzias latipes*), the platyfish (*Xiphphorus maculatus*), the spotted gar (*Lepisosteus oculatus*), the stickleback, the puffer fish (*Tetraodon nigroviridis*), the Nile Tilapia and the zebrafish (*Danio rerio*), all available in Ensembl (www.ensembl.org, accessed 23 March 2014; Flicek *et al*. 2014), and the rainbow trout (Berthelot *et al*. 2014).

1.4. Cod farming and the genetic management of aquaculture broodstocks

1.4.1. Aquaculture

The first evidence of aquaculture production dates back to ancient China and the Far East, but only in recent decades has aquaculture rapidly expanded. This expansion was brought about by a worldwide decline of ocean fisheries (Naylor *et al*. 2000) as well as an increasing support from international institutions and governing agencies (Gjedrem 1997). In contrast to breeding programmes for traditional terrestrial agricultural species, national and international research organisations still support most breeding programmes for aquatic species (Hershberger 2006; Rosendal *et al*. 2013). One noticeable exception is the industry-operated selective breeding programme for Atlantic salmon in Norway (Hershberger 2006).

1.4.2. Cod farming

The artificial rearing of cod started over a century ago in Norway, when Georg O. Sars successfully hatched cod eggs collected from the sea, as well as eggs that were artificially fertilised in the laboratory. In the following years, inspired by Sars, USA and Norway initiated further experiments, releasing cod larvae into the wild, with the aim of restocking local populations (Svåsand *et al*. 2000)**.** Restocking attempts were also widely used for salmonid species in the USA as well as some 90 species of fish and invertebrates in Japan, with much debate over the success and consequences of such practices (Ward 2006). The restocking attempts, using cod larvae, came to a halt in the nineteen seventies (Svåsand *et al*. 2000), although a later study, conducted in the nineteen nineties, released genetically tagged cod larvae into a Norwegian landlocked fjord for research purposes. This study showed that the historical large-scale releases of cod larvae almost certainly had minimal effect on cod recruitment (Kristiansen *et al*. 1997).

From the nineteen seventies onwards, restocking attempts in Norway were done using larger juveniles (> 10 cm). This was followed, in the nineteen nineties, by USA, Denmark, Sweden and the Faroe Islands funded studies, eventually evolving into what we know as cod farming today (Svåsand *et al*. 2000). The first attempts to develop cod farming concentrated on rearing small

wild caught cod and on producing juveniles in natural enclosures, fed with natural zooplankton. These methods provided an unstable supply of juveniles and around 1990 modern hatchery methods were developed where larvae were fed with cultivated plankton such as rotifers (*Brachionus* spp.) and *Artemia* (Rosenlund and Halldórsson 2007).

1.4.3. Challenges in cod farming

There are many challenges in cod farming and one major obstacle for development and growth of the cod farming industry is that, unlike the salmon farming industry, cod farming is constantly competing with the wild cod fisheries, making it vulnerable to fluctuations in the world economics and fish prices (Rosendal *et al*. 2013). Hence, the commercialisation of cod farming in the nineties soon slowed down for economic reasons, one of the main reasons for this being the increase of the natural cod stocks (Rosenlund and Halldórsson 2007). Around the millennium cod farming was in favour again owing to a decline in wild stocks and good prices for cultured cod (Hamre 2006; Rosenlund and Halldórsson 2007) but after the recent global financial crisis, that commenced in USA and Europe in 2008, the cod farming sector has again been hit hard (Rosendal *et al*. 2013).

Although selective breeding has produced very promising results for aquaculture species, such as an average genetic gain in growth rate of 12.5 % per generation, it is estimated that less than 10 % of aquaculture production is based on genetically improved stocks (Gjedrem *et al*. 2012). One of the prerequisites for successful selective breeding is a closed life cycle of the species in question, which has been achieved for cod (Gjedrem 2000; Rosenlund and Skretting 2006). However, some producers still use eggs from wild-caught broodstock because they tend to be of a better quality (better fertilisation and hatching rates) than those produced by farmed broodstock individuals (Roy *et al*. 2007; Lanes *et al*. 2012). Hence, if cod farming is to reach its forecasted potential, broodstock performance needs to become more reliable (Kjesbu *et al*. 2006; Rosenlund and Skretting 2006). Moreover, in the hatchery phase there is a need for more reliable juvenile production with better survival rates and fewer cases of deformities. In the ongrowing phase more cost-efficient feeds produced from sustainable sources need to be developed and more expertise

is needed regarding health management (Rosenlund and Skretting 2006). Early maturation is a major problem, since cod generally mature before they reach the desired market size (Johansen *et al*. 2000; Rosenlund and Skretting 2006; Björnsson *et al*. 2007). Sexual maturation in fish is generally accompanied by a reduction in somatic growth and condition factor and can lead to a significant decrease (up to 30 %) in carcass weight (Braaten 1984; Pedersen and Jobling 1989; Kjesbu *et al*. 2006). Apart from being very unfortunate from an economical point of view, as several months' worth of feed will be wasted, spawning in sea cages may cause ecological concerns, as fertilised eggs may be released into the surrounding environment (Karlsen *et al*. 2006).

1.4.3.1. Juvenile production

Juvenile production has moved from extensive and semi-extensive methods, where the larvae were fed on natural zooplankton, to intensive methods, where larvae are being fed on cultured rotifers and *Artemia*. The latter method is advantageous, since live feed can be produced in sufficient quantities more reliably and, if implemented together with light manipulation of the broodstock, juveniles can be produced year-round (Hamre 2006). Unfortunately, the intensive production of Atlantic cod can also lead to very high and unpredictable mortality rates at the early life stages (Gunnarsli *et al*. 2009). Live feed is associated with the risk of introducing pathogenic microbes into the production system (Ringø and Birkbeck 1999; Makridis *et al*. 2000). Even though this risk can be reduced by treating the water with filtration, ozone and / or UV (Holström and Kjelleberg 1999; Skjermo and Vadstein 1999; Olafsen 2001), the high organic load associated with intensive culture systems can create good growth conditions for different bacteria. Consequenctly, live-feed associated mortalities have been shown to be a major obstacle in the production of juvenile cod (Hansen and Olafsen 1999; Skjermo and Vadstein, 1999). Also, the lower growth rates and higher frequencies of deformities in intensive compared to extensive systems present an urgent economical as well as ethical problem (Hamre 2006). Improved rearing protocols have reduced, to some extent, the frequencies of deformities, but there is a need for more knowledge regarding the first-feeding period of cod (Kjesbu *et al*. 2006) and on fish larval

nutrition in general (Hamre *et al*. 2013). One interesting area of research is that of micro particulate feeds which are believed to have the potential to replace live feeds entirely, but more research is needed in this area (Kjesbu *et al*. 2006; Puvanendran *et al*. 2006; Hamre *et al*. 2013).

1.4.3.2. Photoperiod manipulation

The minimum market size of cod (3 kg; Rosenlund and Skretting 2006) can be achieved when the fish is between 2 and 3 years of age, at the earliest (Björnsson *et al*. 2007). Unfortunately, most cultured cod have been observed to reach sexual maturity by the age of two years (Johansen *et al*. 2000) and males may even reach sexual maturity at the age of one and, therefore experience more than one reproductive cycle before they reach harvest size (Trippel *et al*. 2008). When cod mature prior to reaching market size, it causes great economic losses for farmers, since it may take the fish up to six months to regain condition and weight (Davie 2005; Kjesbu *et al*. 2006). Sexual maturation in marine teleosts is under the regulation of exogenous environmental signals and endogenous neuroendocrine mechanisms. Environmental factors include photoperiod, temperature, nutrition and pheromones, with photoperiod considered the main dictator of the onset of reproduction (Bromage *et al*. 2001). The challenge of early maturation is not unique to cod, but exists in the production of other fish species as well, and different methods exist to delay maturation, such as ploidy manipulation and direct or indirect sex reversal (Pandian and Sheela 1995; Felip *et al*. 2001; Davie 2005). These methods are only in the very early stages of development for cod (Peruzzi *et al*. 2007), and the most realistic solution in the short term is the use of artificial photoperiod manipulation (Davie 2005). Artificial light regimes were initially developed in other fish species to advance and delay spawning in order to maintain an extended or even year-round production of seed (Bromage 1995; Bromage *et al*. 2001). In closed systems, it has proven possible to suppress maturation of cod at 2 and 3 years of age, by exposing the fish to continuous light from the summer solstice when the cod is 1+ years old. It is difficult to achieve such results in open cages, with the overlying natural photoperiod (Davie 2005). Atlantic cod appear to have a different biological response to artificial light than for example salmon, a difference that also extends to the behaviour. The artificial light attraction that has been seen for

salmon in light-controlled sea cages, appears to be absent in Atlantic cod (Skulstad *et al*. 2013 a,b). Nonetheless, Kolbeinshavn *et al*. (2012) reported successfully delaying maturity in Faroe Bank cod up to 4 kg in open cage systems.

1.4.3.3. Health management

As the salmon farming did in its early phases, the cod farming sector is now struggling with disease control (Rosendal *et al*. 2013). A variety of infections and diseases have been identified in cultured gadoid species (Bricknell *et al*. 2006; Kjesbu *et al*. 2006), such as atypical furunculosis (Magnadóttir *et al*. 2002; Lund *et al*. 2008) and vibriosis (Bricknell *et al*. 2006; Kjesbu *et al*. 2006). Based on experience from salmonids, more diseases can be expected as production intensifies (Kjesbu *et al*. 2006), although at the same time modern production techniques and improved husbandry strategies have reduced the risk of disease outbreaks associated with the intensive production of gadoid species (Bricknell *et al*. 2006).

Of all diseases identified to date vibriosis causes the highest mortality rates in farmed cod (Kjesbu *et al*. 2006). Vibriosis is a bacterial disease caused by the bacterium *Vibrio anguillarum*, which has also caused problems in salmonids (Bricknell *et al*. 2006). However, the salmonid pathogens are mainly of different serotypes than those found in cod (Larsen *et al*. 1994), so vaccines developed for salmonids are not necessarily effective in protecting the cod against infection (Schrøder *et al*. 2006). New vaccines have been developed for cod, but these are still not as efficient as those produced for salmonids (Kjesbu *et al*. 2006; Schrøder *et al*. 2006). A part of the explanation for this was recently discovered, with the sequencing of the cod genome. It appears that the immune system in cod is unique in many ways compared to other vertebrate species, cod do not possess some of the genes that are vital for the function of the major histocompatibility complex (MHC) II pathway, but have acquired other compensatory mechanisms. These findings should provide the basis for better disease control in Atlantic cod, for example through the development of more effective vaccines (Star *et al*. 2011).

Fish respond differently to the constant infection pressure in their surrounding environment and some of this variation in response likely reflects genetic variation, meaning that it should be possible to improve disease resistance through artificial selection (Gjedrem 1997). When it comes to vibriosis resistance in Atlantic cod sufficient additive genetic variation has been found for selective breeding to be considered a viable control measure and vibriosis resistance is, currently, one of the traits being selected for in Atlantic cod programmes (Fjalestad *et al*. 2006; Kettunen and Fjalestad 2006; Kettunen *et al*. 2007; Kettunen and Fjalestad 2007). In aquaculture species in general, the rapid accumulation of genetic and genomic data is expected to result in the identification of marker loci and genes that are associated with many important traits, including disease resistance (Quinn *et al*. 2012).

1.4.4. Genetic management of aquaculture breeding programmes

Genetic information has long been exploited in the terrestrial livestock industry (Gjedrem 1997) and a combination of genetic technologies and breeding practices has led to increases in production up to six-fold in some species (Havenstein *et al*. 2003). However, due to several production bottlenecks, such as issues with closing the life cycles and with the domestication of aquatic species (Jerry *et al*. 2001), the aquaculture industry has only relatively recently embraced the use of genetic technologies (De-Santis and Jerry 2007; Gjedrem 1997). Although aquaculture currently lags behind terrestrial animal production in this respect, aquaculture production has the potential to equal or surpass terrestrial farm animal gains, since fish have higher fecundities, and stocks inherently possess relatively high levels of genetic variation. Furthermore, most aquatic species have external fertilisation, which provides the possibility for artificial control (Gjedrem 2005).

Aquaculture species have been found to respond as well, or better, to genetic selection than terrestrial animals. Over a 20 year period the time needed to rear salmon from 50 g to 3.5 kg decreased from 2 years to 10 – 12 months (Gjedrem 1997). Investment in aquaculture breeding programmes can be highly profitable, as seen from estimates of benefit to cost ratios of between 8.5 and 60 for a tilapia breeding programme (Ponzoni *et al*. 2007) and between 22 and 420 for

genetic improvement programmes in common carp (Ponzoni *et al*. 2008). As mentioned above, to gain the most from a breeding programme, the entire life cycle of the species needs to be controlled, so that broodstock can be selected from farmed fish (Gjedrem 2000; Kjesbu *et al*. 2006). Therefore, thorough biological investigations of the biology of the species in question must be undertaken before initiating a breeding programme. Furthermore, breeding goals and the presence and magnitude of genetic and phenotypic variation associated with traits of interest need to be established (Gjedrem 2000).

1.4.4.1. Molecular genetic tools in aquaculture

The application of genetics in aquaculture can be divided into 3 main areas: i) monitoring; ii) modification / improvement and iii) genetic mining. Genetic monitoring produces information on the diversity and similarity among wild and captive stocks and provides tools for identifying biological units, such as stocks, families and individuals. Genetic improvement involves the design of improvement programmes based on knowledge of phenotypes, genotypes and potentially pedigrees of the broodstock. Apart from selective breeding, which is the typical method of choice, genetic improvement can also be achieved through crossbreeding and hybridisation, chromosomal manipulation, gynogenesis, control of sex determination, cloning and the production of transgenics (Lester 2002; Hershberger 2006). Accumulating knowledge on DNA, RNA, proteins and other biochemical constituents of the cell have resulted in the third category, genetic mining, which provides researchers with tools to search for information on control, function and gene composition of biological systems. For instance, it is possible, by the use of microarray techniques, to gain information on the activation of thousands of individual genes in an individual (Hershberger 2006).

1.4.4.2. Genetic management to prevent the accumulation of inbreeding

Genetic gain is dependent on genetic variation in the broodstock (Gjedrem 1997), selection intensity and heritability of the trait of interest (Borrell *et al*. 2004). However, the high fecundity of aquatic animals can potentially cause difficulties, as even large-scale hatcheries may utilise a small number of breeders, resulting in a rapid accumulation of inbreeding and have unintended

selection consequences (Eknath and Doyle 1990; Gjedrem 2005). Inbreeding can be detected as an increase in genetic homozygosity and, in addition to a loss of genetic variation, inbreeding may result in reduced performance, termed inbreeding depression (Sonesson *et al*. 2005).

In order to avoid reduced genetic variation in the offspring, appropriate management is needed in terms of maintaining a sufficient number of effective breeders (Estoup *et al*. 1998). A number of early breeding programmes with Nile tilapia failed to improve growth rate, presumably because of too low a level of genetic variation in the base population and subsequent loss in subsequent generations (Hulata *et al*. 1986; Teichert-Coddington and Smitterman 1988; Huang and Liao 1990). Regardless of the size of the aquaculture stock, all closed breeding programmes will inevitably result in an increased relationship among individuals, as alleles are lost through random or selective processes. The solution is to control or manage the inbreeding (Sonesson *et al*. 2005). The loss of genetic variation per generation may be minimised, even in the case of small effective population sizes, by keeping pedigree records on all individuals and using this information to arrange the most suitable matings (Cross *et al*. 1993). When there are no available pedigree data on the broodstock, one way to avoid inbreeding is the use of relatedness coefficients to assist in the selection of suitable breeding pairs (Norris *et al*. 2000). For Atlantic cod, one would think that there was little risk of loss of genetic variation, as most cod hatcheries produce eggs by mass spawnings with 100 or more potential parents (Herlin *et al*. 2007). However, family representation may be highly skewed as for example found in a study of a commercial spawning tank containing 99 potential parents. One single parental pair contributed to 25 % of the fry sampled on a single day and over 90 % of the progeny originated from nine parents (Herlin *et al*. 2008). Using 15 microsatellite loci, a study of three farmed and four wild populations of Atlantic salmon demonstrated lower genetic variability, in terms of allelic diversity, in the farmed compared to the wild salmon. Although it could not be stated for certain, the reason for this could be that the farmed salmon originated from small founder populations resulting in the loss of rarer alleles. Overall heterozygosity was not different between farmed and wild salmon, (though heterozygosity alone is not a good indicator of change when comparing multi-allelic

microsatellite data, number and / or effective number of alleles is a more informative measure). That said, expected heterozygosities were significantly different within the farmed populations. Since all farmed populations had been selectively bred to improve commercially important traits, the differences in expected heterozygosities might indicate different levels of selection pressures with higher selection pressures resulting in greater losses of genetic variability (Norris *et al*. 1999). To conclude, in order to avoid low numbers of contributing parents, mating of close relatives and a resultant loss of genetic variation, some kind of genetic management is fundamental. This is especially true for aquaculture populations which often encounter severe inbreeding rates in the domestication process, when they go from large to small populations (Sonesson *et al*. 2005) and are reared in markedly different environments.

1.4.4.3. Selection for favourable traits

In the process of artificial selection, individuals with desirable parameters for the trait(s) under selection are preferred over others to produce the next generation. Selected traits are generally quantitative, meaning that they are under the control of a large number of genes, and the result of selection is an increase in frequencies of beneficial alleles which may be observed as a change in the population mean for the trait in question (Gjedrem and Thodesen 2005). Which traits one chooses to improve through a breeding programme varies from species to species, but aquaculture breeding programmes usually include one or more of these traits: growth rate, disease resistance, quality traits and age at sexual maturation (Gjedrem 2005). Genetic gain estimates for growth rate in aquatic species typically lie in the range of 10 – 20 % per generation of selection, which are outstandingly high figures compared to terrestrial livestock values (Gjedrem and Baranski 2009). Heritability is a valuable parameter for predicting the response to selection. In its narrow sense, heritability is defined as the ratio of the additive genetic variance to the total phenotypic variance. Heritability estimates can vary between 0 and 1, with high values indicating that relatively large selection responses can be obtained whereas low values mean that environmental effects play a larger role in the observed variation of phenotypes. It should be remembered that heritability estimates are not static, but are only valid for the data used for estimation and under

the particular circumstances in which measurements were taken. If the conditions change or the population responds to selection, the heritability may change (Gjedrem and Olesen 2005). In spite of this, heritability estimates can generally be grouped according to the characters in question (Falconer and Mackay 1996b). Life-history traits generally show low heritability values as genetic variation is more rapidly lost in traits associated with fitness (Mousseau and Roff 1987; Falconer and Mackay 1996b). Mousseau and Roff (1987) reviewed over 1000 estimates for wild outbred animal species and found mean heritability values of 0.26 and 0.46 for life-history traits and morphological traits, respectively. Also, regarding morphological traits, heritability estimates were found to be significantly lower in ectotherms than in endotherms, which may be explained in part by the strong correlation between body size and life history in ectotherms (Mousseau and Roff 1987). Although relatively low, heritability values between 0.2 and 0.3 are still considered high enough to elicit substantial selection responses over a few generations (Law 2000). Possible mechanisms for maintaining such comparatively high genetic variation at life-history traits include mutation, heterozygote advantage, frequency-dependent selection and genotype-by-environment interaction (Freeman and Herron 2007).

1.4.4.4. Domestication and unintended selection

Domestication is a genetic process in which animals adapt to their man-provided environments (Gjedrem and Thodesen 2005). In the intensive cultivation of fish uncontrolled genetic changes are very likely to occur, even in the absence of any selective breeding programme, and may result in serious depletion of genetic variation (Eknath and Doyle 1990; Doyle *et al*. 1995). In a study of brown trout (*Salmo trutta* L.), start-fed with either a low or a high ration diet, the pattern of family mortality differed significantly between the two treatments, indicating a significant genotype × environment interaction at the critical period of first-feeding (Glover *et al*. 2004). As food availability is likely to be very different in a hatchery versus the natural environent (Waples 1999), these results suggest that individuals that have been bred in a captive environment may not perform as well in a natural environment. Accordingly, a number of early studies on salmonids demonstrated a decreased performance for domestic strains compared to wild strains

(Greene 1952; Vincent 1960; Flick and Webster 1964; Moyle 1969; Flick and Webster 1976; Reisenbichler and McIntyre 1977; Fraser 1981; Keller and Plosila 1981; Chilcote *et al*. 1986; Leider *et al*. 1990).

1.4.4.5. Parentage analysis

Parentage analysis exploits the fact that microsatellites (and other types of nuclear markers, e.g. allozymes and SNPs) segregate in a Mendelian fashion, so that in disomic species each offspring inherits one paternal allele and one maternal allele at each nuclear locus (Castro *et al*. 2004; Chistiakov *et al*. 2006). From a panel of several to tens of microsatellite loci, a unique genotype profile can be constructed for each individual tested such that the probability that any individual could be wrongly assigned to a specific family (or parent) can be ascertained, and is likely to be acceptable low (Chistiakov *et al*. 2006).

Parentage analysis offers an alternative to the management of families in separate units until they are large enough to tag and although genotyping may be relatively expensive, it remains advantageous compared to otherwise large and high-cost facilities required for separate family rearing. Furthermore, common rearing from the very early life stages mimicks large-scale production conditions (Estoup *et al*. 1998), which may make it easier to detect real genetic differences, resulting in an increased selection accuracy (Delghandi *et al*. 2003). The discovery of microsatellite markers set off a burst in techniques for resolving parentage (Jones and Ardren 2003) and various programmes exist for analysing parentage, such as PAPA (Duchesne *et al*. 2002), CERVUS (Marshall *et al*. 1998) and FAP (Taggart 2007). Two of the major techniques for reconstructing parentage are exclusion and likelihood-based techniques. Exclusion uses the Mendelian model of inheritance to exclude all but one parent pair, while likelihood-based techniques assign offspring to non-excluded parents by the use of likelihood scores. At first glance exclusion seems like the ideal solution, but strict exclusion rarely achieves 100 % assignment due to sensitivity to genotyping errors (such as null alleles, large allele dropout and failure to resolve closely migrating allelic bands), and the larger the datasets the more prone they are to experience these difficulties (Jones and Ardren 2003). Likelihood approaches generally

result in more assignments, but the mathematical model behind the assignments and the error level set should be carefully considered when interpreting the results (Herlin *et al*. 2007).

1.5. Trans-North Atlantic population structure of marine fish species

The number of population genetic studies are scarce for most of the managed marine fish species in the North Atlantic. Atlantic cod is by far the most studied species, with Atlantic herring in a strong second place. In addition, there is a publication bias towards studies that aim at delineating population structure at relatively small scales (Reiss *et al*. 2009). Hence, there are many gaps in our knowledge regarding Trans-North Atlantic population structure of marine fish and the various parameters that act to shape it. The current picture of the population structure of Atlantic cod and other species in the North Atlantic Ocean is discussed below together with challenges associated with delineating population structure in these high gene flow species.

1.5.1. Population genetic structure of high gene flow marine species – complicating factors

It may be difficult to identify significant genetic differentiation in marine species that are typically characterised by high levels of gene flow (Waples 1998) and non-representative sampling and failure to replicate over years may lead to contradictory results (Reiss *et al*. 2009). Different marker systems may also yield different results (Imsland and Jónsdóttir 2003; O'Leary *et al*. 2007; Reiss *et al*. 2009), as some markers may be evolutionary neutral whereas others are under the influence of natural selection (Hemmer-Hansen *et al*. 2007; Larsen *et al*. 2008a; Pampoulie *et al*. 2011). In Atlantic cod a limited number of blood protein loci detected significant differences between neighbouring cod populations (Møller 1968; Jamieson and Otterlind 1971; Jamieson 1975; Cross and Payne 1978; Dahle and Jørstad 1993). Never the less, studies over a large range of the species´ distribution, applying a large number of allozyme loci, detected only low levels of genetic differentiation among populations (Mørk *et al*. 1982, 1985). More recent microsatellite-based studies generally reported more instances of discernable structure (Bentzen *et al*. 1996; Ruzzante *et al*. 1998; Hutchinson *et al*. 2001; Nielsen *et al*. 2001; Jónsdóttir *et al*. 2002; Knutsen *et al*. 2003).

The Pantophysin (*Pan*I) locus is an example of a locus considered to be under strong selective pressure in Atlantic cod (Fevolden and Pogson 1997; Pogson and Fevolden 1998; Pogson *et al*. 2001; Beacham *et al*. 2002; Karlsson and Mørk 2003; Pogson and Mesa, 2004; Case *et al*. 2005) and in walleye pollock (*Theragra chalcogramma*) (Canino and Bentzen 2004; Canino *et al*. 2005). A study of North-East Arctic cod, Norwegian coastal cod and North Sea cod demonstrated the difference in population structure inference derived from *Pan*I data compared to that of seven microsatellite loci (considered to be neutral). For example, microsatellites revealed genetic heterogeneity in all three groups, the *Pan*I locus only in one of the groups, and in contrast to the *Pan*I locus, microsatellites only detected low levels of genetic differentiation between the North-East Arctic cod and the other groups. Being a marker under selection and possessing only two allelic classes, the *Pan*I locus is likely to respond more rapidly to environmental changes, resulting in an earlier detection of divergence between populations compared to microsatellite loci. Whereas the pattern of differentiation detected by microsatellite loci reflected an isolation by distance mechanism, environmentally induced selection appeared to shape *Pan*I allele frequencies (Skarstein *et al*. 2007). It is not advisable to make conclusions on findings from selective markers only; however, as strong selection can cause genetic differences between fish from areas that are not reproductively isolated (O´Leary *et al*. 2007). Neutral markers could be considered together with markers liable to be influenced by environmental or other constraints to provide reliable estimates of population breeding structure and the factors that shape it (Case *et al*. 2005). This would generate valuable information from a conservational perspective, as the long-term survival of a species depends not only on the preservation of (neutral) genetic diversity, but also of genetic resources or adaptive capabilities, i.e. genetic variation that is actually being expressed on a phenotypic level (Hauser and Carvalho 2008).

1.5.2. The present picture of Trans-North Atlantic population structure in marine fish

For Atlantic cod, genetic studies using micro- and minisatellites (Hutchinson *et al*. 2001; Galvin *et al*. 1995; Bentzen *et al*. 1996; O'Leary *et al*. 2007) as well as RFLPs (Pogson *et al*. 1995; Jónsdóttir *et al*. 2003) and covering a large part of the distributional area have found a clear

separation between North-West and North-East Atlantic cod. In addition, these investigations have detected relatively high levels of genetic differentiation among Barents Sea cod, Baltic Sea cod and the remaining North-East Atlantic populations. A number of studies have also revealed finer scaled population structure (for example Ruzzante *et al*. 1996, 2000b; Beacham *et al*. 2002 for the North-West Atlantic and Nielsen *et al*. 2001; Skarstein *et al*. 2007; Case *et al*. 2005 for the North-East Atlantic).

Other studies of marine fish that have assessed population genetic structure on a Pan-North Atlantic basis, using microsatellites, have replicated the major boundaries to gene flow that have been identified for Atlantic cod. For example, Pampoulie *et al*. (2014) identified three major groups in lumpfish (*Cyclopterus lumpus*), i) a North-West Atlantic group (that included West-Greenland), ii) Iceland-Norway and iii) the Baltic Sea. Knutsen *et al*. (2009) investigated genetic population structure in the deep-sea fish tusk (*Brosme brosme*) and identified major breaks to gene flow between the North-East and the North-West Atlantic, but also between the Mid-Atlantic Ridge and the rest and between Rockall and the rest. The study failed to detect substructuring among tusk from East-Greenland, Iceland, Faroe Islands (sample from the Faroe Plateau) and Norway. McPherson *et al*. (2004) found significant genetic differentiation among herring from the North-West Atlantic, Iceland, the Celtic Sea and the Baltic Sea. Lastly, Bekkevold *et al*. (2005) identified significant differentiation of Atlantic herring samples collected from a longitudinal gradient covering the North Sea and the Baltic Sea, presumably governed by environmental conditions such as temperature and salinity, or related factors.

Whereas there appears to be some similarity in marine fish species concerning the patterns of spatial genetic differentiation, the level of genetic differentiation may differ. For example, the level of genetic differentiation in tusk was found to be comparable with another deep-water fish, the Greenland halibut (*Reinhardtius hippoglossoides*), but lower than those observed for cod (Knutsen *et al*. 2009). Species also differ concerning the main mechanisms governing spatial genetic differentiation patterns. For example, an isolation by distance model has been suggested for cod collected over a large part of the distributional range (Pogson *et al*. 1995; O'Leary *et al*.

2007; Skarstein *et al*. 2007) as well as on a regional scale, as assessed by RFLP variation in the North-West Atlantic (Pogson *et al*. 2001), although no evidence was found of isolation by distance along the Norwegian Skagerrak coast (Knutsen *et al*. 2003). Knutsen *et al*. (2007) found genetic differentiation estimates of Greenland halibut to be closer correlated with geographic distances along ocean currents than with linear geographic distances, suggesting a gene flow mediated by free-floating eggs and larvae. In the deep-sea species, tusk, genetic divergence was related to distances between habitable areas, with strong differentiation values found for two areas (the Mid-Atlantic Ridge and Rockall) that were surrounded by depths greater than that occupiable by the species (Knutsen *et al*. 2009). These authors suggested that barriers to gene flow in tusk are created by limited adult migration across deep ocean basins together with local circumstances, such as gyres, that limit site interchange of eggs and larvae, thus showing that gene flow in a single species may be governed by several factors.

1.5.2.1. Local discrepancies with traditional fisheries management areas

As mentioned above (Section 1.2.2), there may be discrepancies between fisheries management zones and the boundaries of the biological populations. There may be many reasons for this: fisheries management zones are for example frequently based on political and administrative boundaries. In addition, it is a complicated task to assess biologically meaningful population entities and several approaches have been suggested, that broadly fall into one of two categories, the ecological or the evolutionary definition. The ecological definition relies on the co-occurrence and interaction of individuals in time and space and fisheries advice are generally based on this, whereas the evolutionary definition is concerned with the genetic structure of populations, reflecting actual reproductive interaction among individuals (Reiss *et al*. 2009).

Reiss *et al*. (2009) reviewed population genetic studies for the 32 managed fish species in the North-East Atlantic, defined biological entities based on evidence of significant and reproducible genetic differentiation and compared these to the current fisheries management units. Inconsistencies were found for six species, Atlantic cod, haddock, blue whiting (*Micromesistius poutassau*), European hake (*Merluccius merluccius*), whiting and herring.

A special scenario is evident for NEAC and NCC that are assessed separately; they comprise two genetically distinct populations that share spawning sites, so that the fisheries target mixed stocks. Fisheries management deals with this by adding an expected amount of 20,000 tons of NCC to the NEAC quota. For NCC cod and Skagerrak cod, however, genetic structure is suggested at smaller scales than what is accounted for by the fisheries management bodies (Reiss *et al*. 2009). This may be hazardous, as both simulations and real-life examples for Atlantic cod have shown that if several subpopulations are managed as one, the growth and harvest potential may easily be overestimated with the risk of extinction of one or several subpopulations (Smedbol and Stephenson 2001; Fu and Fanning 2004; Sterner 2007). The northern cod off Newfoundland that collapsed due to overfishing had been managed as a single population, despite evidence of several subpopulations that differed in their spatial distribution during winter and spring, although with a certain level of individual exchange (Smedbol and Stephenson 2001). Thus, the challenge lies not only in incorporating knowledge of spatial genetic structure into fisheries management practices, for some species management practices may be complicated by subpopulations that interact during part of the year. Atlantic herring is for example prone to such spatial and temporal complexity. Otolith analyses for this species has shown that larvae and adults of Baltic and North Sea spawning components mix at feeding and nursery areas in the Skagerrak (Clausen *et al*. 2007).

By applying a large panel of informative SNP loci (nearly 100) and sampling a wide region over many years during spawning, a recent study was able to delineate population structure of cod in the North Sea and West of the UK, where microsatellite studies had previously revealed somewhat unclear results. Three reproductively distinct spawning populations were identified, as well as the geographic barriers between them, and two of them cohabited in the North Sea, which is managed as a single fishing area, although divided into sub regions. This study also used trawl survey data on the biology of the fish in the area as well as data on fishery landings to develop a fine-scale spatial model that was used to simulate the responses of the two population units to different spatial patterns of fishing mortality. The model revealed a complex response to

harvesting that was partly determined by competition between the two population units, rendering the more localised population unit at risk under some spatial patterns of harvesting. The fact that current management areas are well embedded in both national and international legislation makes them hard to alter. This study, however, suggests that fine-scale regulation of fishing mortality within a current management area may protect the most vulnerable population unit and thus help to conserve overall diversity and population productivity. Hence, this strategy of combining fine-scale genetic investigations with spatial modelling based on biological information and catch data may be a way forward concerning adopting genetic knowledge into current fisheries management practices (Heath *et al*. 2014).

1.6. Two Faroese cod stocks

1.6.1. Two ecosystems

Two distinct stocks of Atlantic cod are recognised around the Faroe Islands, one from the Faroe Plateau, on which the islands themselves are situated, and the other one from the Faroe Bank (Fjallstein and Magnussen 1996; Figure 1.3). In addition, a third cod stock is located on the Faroe-Iceland Ridge, but this stock is considered to belong to the Icelandic cod stock (ICES 2014a), and focus in this thesis is on the Faroe Plateau and the Faroe Bank cod stocks. The two stocks are geographically isolated by an 800 m deep channel and water temperatures are 1 – 2 ºC higher on the Faroe Bank than on the Faroe Plateau (Magnussen 2006). During the spawning season, the temperature is 6 – 7 ºC on the Faroe Plateau (Steingrund *et al*. 2005) and 7.5 – 8.0 ºC on the Faroe Bank (Magnussen 2002, 2007). Considering depths < 200 m, the Faroe Bank area compares to *c*. 15 % of the Faroe Plateau area (Eliasen *et al*. 2005). On both the Faroe Plateau and the Faroe Bank, anticyclonical currents act to keep plankton, eggs and larvae within the respective areas (Hansen 2000), and as a consequence the two areas are generally considered as two separate ecosystems (Steingrund and Gaard 2005; Magnussen 2007).

The Faroe Plateau shelf water has been shown to support a relatively simple ecosystem, in which higher trophic levels, such as the cod, are closely correlated with primary production (Steingrund and Gaard 2005). This causes large fluctuations in stock biomasses, since primary production

shows great inter-annual variations, suggested to be largely determined by a variable exchange of on-shelf and off-shelf waters (Eliasen *et al*. 2005; Steingrund and Gaard 2005; Debes *et al*. 2008).

The Faroe Bank has a high biodiversity compared with surrounding waters and it supports a number of species that are not found on the Faroe Plateau as well as a number of species that have been classified as threatened in other North-East Atlantic regions. Although the sandy bottom on top of the bank may be poor in macrofauna, it is rich in meiofauna. Moreover, the slopes of the bank are rich in various species of sponges that host a high number of filter feeder species and are likely to be important nursery area for various fish species. Partly because of these characteristics and partly due to threatened state of the cod on the Faroe Bank, which is known as the fastest-growing cod in the world, the Faroe Bank has been proposed as a marine protected area (MPA) by the World Wildlife Fund (WWF) (Schmidt *et al*. 2003).

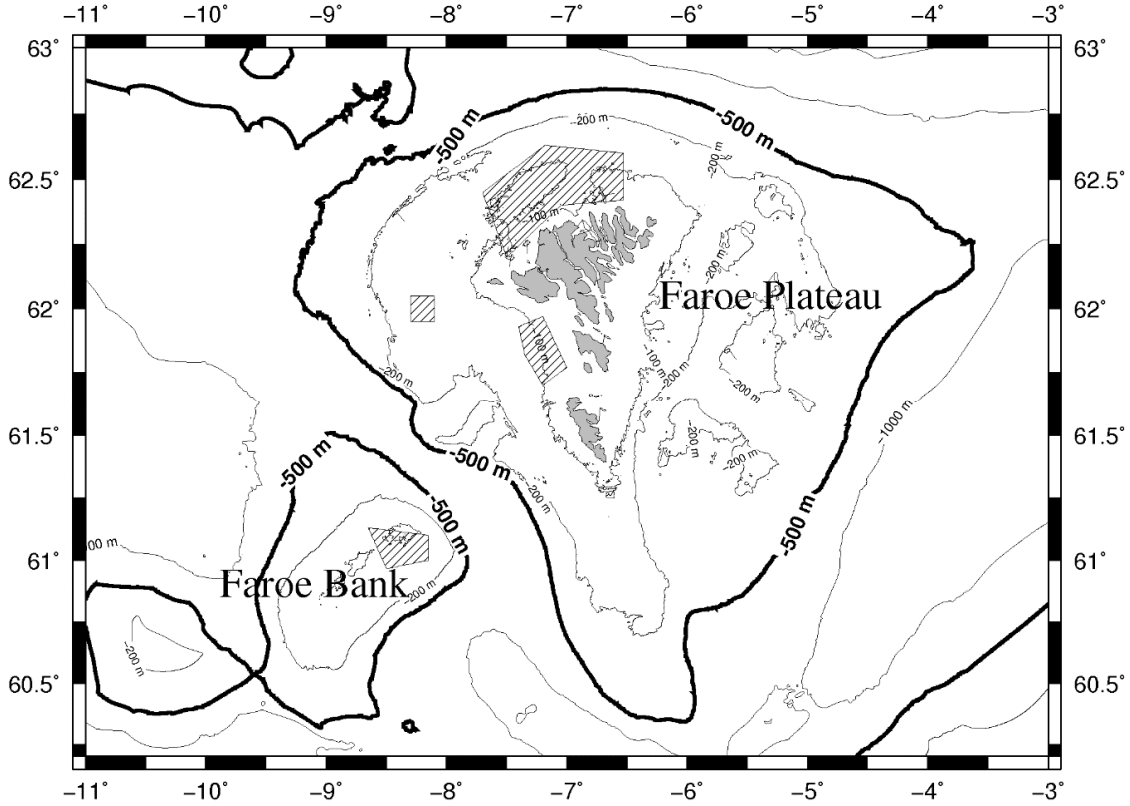

Figure 1.3. Map of the Faroe Plateau and the Faroe Bank area with the main spawning areas hatched. Map constructed by H. Abrahamsen (with assistance by P. Steingrund) and reproduced with permission.

1.6.1.1. Spawning

On the Faroe Plateau peak spawning activity occurs during the second half of March, so the majority of larvae start feeding at the beginning of April, mainly on copepod eggs of the species *Calanus finmarchicus*. At this time, production of zooplankton on the central plateau is still low (Gaard and Steingrund 2001), but the main spawning grounds are located North and West of the islands (Figure 1.3; í Jákupsstovu and Reinert 1994), close to the areas with the highest average egg production of *C. finmarchicus* during early spring (Gaard and Steingrund 2001). These spawning grounds are by far the most important ones, although spawning also takes place all around the plateau at depths between 80 and 180 m. Recent tagging studies (1997 onwards) have shown that cod spawning on the northernmost spawning ground spread more or less over the entire plateau during the rest of the year, whereas cod spawning on the Western spawning ground are more stationary, occupying an area South-West of the islands outside the spawning season. Furthermore, Faroe Plateau cod show spawning site fidelity, i.e. individuals appear to migrate to

the same spawning ground every year (Joensen *et al*. 2005; Steingrund *et al*. 2005; Steingrund 2014). The Faroe Bank stock is believed to be more stationary, occupying the same area regardless of season or stage of life (Joensen *et al*. 2005). Spawning on the Faroe Bank occurs over the period 1 March – 1 May (ICES 2014a).

1.6.2. Stock differences

Differences between cod from the two areas have been detected in post-mortem pH, water content and total protein nitrogen in muscle tissue (Love *et al*. 1974) as well as in fin rays and vertebrae counts, which were proposed to be a consequence of different water temperatures (Schmidt 1930). In addition, the Faroe Bank cod is lighter of colour and more corpulent than cod from the Faroe Plateau and cod from other North Atlantic stocks (Love *et al*. 1974). Of particular interest, however, is the significantly higher growth rate observed for the Faroe Bank cod. For both stocks, the 50 % maturation rate occurs at around three years of age (Steingrund *et al*. 2005). However, on average a three-year-old cod from the Faroe Bank is *c*. 73 cm long and weighs approximately 4.9 kg, whereas a cod of the same age from the Faroe Plateau is, on average, 55 cm long and weighs 1.7 kg (Magnussen 2007, 2011).

Laymen as well as scientists have long speculated over the respective roles of the environment versus the genetics in the observed growth differences between Faroe Plateau and Faroe Bank cod. The environment clearly plays a role, as for example demonstrated by the high zooplankton concentrations that have been recorded on the Faroe Bank compared to the Faroe Plateau and, accordingly, higher food content was also observed in the stomachs of young Faroe Bank cod compared to young Faroe Plateau cod (á Nordi and Poulsen 2000). It has also been proposed that the Faroe Bank cod may have a faster metabolic rate and more efficient food conversion rate, possibly as a result of the 1 – 2 ºC higher water temperatures prevalent on the Faroe Bank (Knutsen and Salvanes 1999).

Spawning temperatures experienced by Faroese cod (7.5 – 8.0 ºC on the Faroe Bank and 6 – 7 ºC on the Faroe Plateau; Section 1.6.1) are broadly in the range of the rearing temperatures (11 and

7 ºC) used in common-garden studies investigating the environmental and genetic components of body shape variation in North-West Atlantic cod populations (Marcil *et al*. 2006a). Of interest, considering the more corpulent phenotype for the Faroe Bank cod, higher rearing temperatures generally resulted in deeper bodied phenotypes of metamorphosed cod, thus indicating plastic responses to rearing temperatures. In addition, there was evidence of genetic divergence at large (< 1000 km) and small-scale (< 100 km) spatial scales, as differences in body shapes were found between populations as well as between spawning-components within a population, when reared under similar environmental conditions (Marcil *et al*. 2006a). Lastly, results for two of the populations (cod from the South-West Scotian Shelf and from the northern-stock complex off Newfoundland) suggested countergradient variation with respect to body shape: in nature, southwest Scotian Shelf cod spawn in winther time whereas northern cod are summer spawners. Since larvae and early juveniles from the two populations grow up at different rearing temperatures (*c*. 2.4 – 3.4 and 5.7 – 7.3 ºC for southwest Scotian Shelf and northern cod, respectively), one could expect differences in body depth between the two with a deeper body for the northern cod. However, no significant differences in body depths were detected between juveniles in the wild, whereas in the common-garden experiments, markedly deeper body shapes were seen for the Scotian Shelf cod compared to the other populations (Marcil *et al*. 2006 a,b). Hence, stabilising selection towards optimal phenotypes appeared to have caused genetic divergence between these populations, despite no apparent phenotypic differences in the wild (Marcil *et al*. 2006a,b).

1.6.2.1. Genetic evidence of population structure

Migration rates between the two Faroese cod stocks and other surrounding stocks are very limited. 24,000 cod were tagged in Faroese waters from 1997 to 2006. Of *c*. 6,000 recaptured cod, one was caught on the Icelandic shelf and one on the Faroe-Icelandic Ridge. During the same period (1997 – 2004), the Icelandic Marine Research Institute tagged 25,572 cod and, of the 3,708 recaptured, 13 were caught on the Faroe-Icelandic Ridge and none on the Faroe Plateau. Additionally, in 2002 168 cod were tagged on the Faroe-Icelandic Ridge and, of ten recaptured

individuals, five were caught in Iceland, three on the Faroe-Icelandic Ridge and none on the Faroe Plateau. The taggings by the Faroe Marine Research Institute alse revealed, that the migration rates between the Faroe Plateau and the Faroe Bank were negligible (ICES 2006), corraborating tagging studies from the first half of the nineteenth century. Based on tagging studies conducted over the period 1909 – 1963, Magnussen (1996) calculated an average migration rate of 0.0029 (95 % confidence interval 0.0010 – 0.0048) from the Faroe Plateau to the Faroe Bank and an average migration rate of 0.0276 (95 % confidence interval 0.0087 – 0.0456) in the opposite direction. However, although migration rates were low, considering the size of the stocks, absolute numbers of migrants were high, with an estimated 45,000 – 220,000 cod travelling from the Faroe Plateau to the Faroe Bank per generation and 13,000 – 74,000 cod travelling from the Faroe Bank to the Faroe Plateau per generation. These high number of migrants were supposedly part of the explanation for the absence of genetic structure, as estimated by electrophoretic studies of blood and muscle enxymes (Magnussen 1996).

However, early electrophoretic studies showed somewhat conflicting evidence regarding the degree of genetic differentiation between Faroe Plateau and Faroe Bank cod. In contrast to the findings by Magnuseen (1996), Jamieson and Jones (1967) found evidence of genetic divergence, using transferrin, and so did Jamieson and Birley (1989), using hemoglobin. However, Jamieson and Thompson (1972), using butyric serum esterase to study population structure of cod from throughout the distribution range, detected no significant differentiation among Faroe Plateau and Faroe Bank cod.

Recent studies, using potentially more informative microsatellite markers, have proposed that the two stocks are genetically distinct (Nielsen *et al*. 2009a). In addition, a comparative growth experiment where cod from both stocks were produced and reared under equal environmental conditions for one year, found a significantly higher condition factor as well as indications of a higher growth for the Faroe Bank cod (Fjallstein and Magnussen 1996). A M.Sc. study on these captive spawning cod also suggested significantly larger egg sizes for Faroe Plateau than for Faroe Bank cod (Magnussen 1993), thus indicating different reproductive strategies for the two.

Lastly, the reared Faroe Plateau and Faroe Bank cod differed with respect to fatty acid profiles of heart tissue (Joensen *et al*. 2000). Hence, these relatively new studies seem to agree on a certain level of genetic differentiation between Faroe Plateau and Faroe Bank cod.

As reviewed above (Section 1.5.2), the major boundaries to gene flow in Atlantic cod appear to be between North-West and North-East Atlantic cod and, furthermore, the North-East Atlantic cod can be divided into Barents Sea cod, Baltic Sea cod and the rest. Hence, only small levels of genetic differentiation might be expected between the Faroese cod and, for example, North Sea, Norwegian coastal or Icelandic cod. In fact, Pampoulie *et al*. (2008a) detected no significant differentiation among Faroe Plateau and East-Icelandic cod, using microsatellite markers and the *Pan*I locus. Another study, using microsatellite markers, detected evidence of genetic discontinuities between the three single populations Eastern Baltic cod, North-East Arctic cod, Faroe Plateau and the remaining samples, which included Faroe Bank cod, a number of North Sea samples, cod from the Belt Sea, cod from the Kattegat / Skagerrak as well as cod from West of Scotland (Nielsen *et al*. 2009a). This latter study, thus, puts the Faroese stocks in a central positions, when it comes to genetic structure of cod in the North Atlantic Ocean, warranting a further investigation of the level and significance of genetic differentiation among Faroe Plateau and Faroe Bank cod, and of the factors that might influence genetic differentiation in this region.

1.6.2.2. The cod fishery in the Faroe Islands

The Faroese society has traditionally been classified as one of the most fishery-dependent societies on earth, with cod being the most important resource, followed by other groundfish species. Consequently, when these fisheries failed in the early nineties, as shown for cod in Figure 1.4, the consequences were catastrophic. Within a few months in 1992 nearly the entire fishing industry went bankrupt, the GNP decreased by 40 % and unemployment increased to 25 %. From the late eighties to the mid nineties the Faroese population declined from nearly 48,000 people to under 44,000 people, mostly an economic migration to Denmark (Hamilton *et al*. 2004).

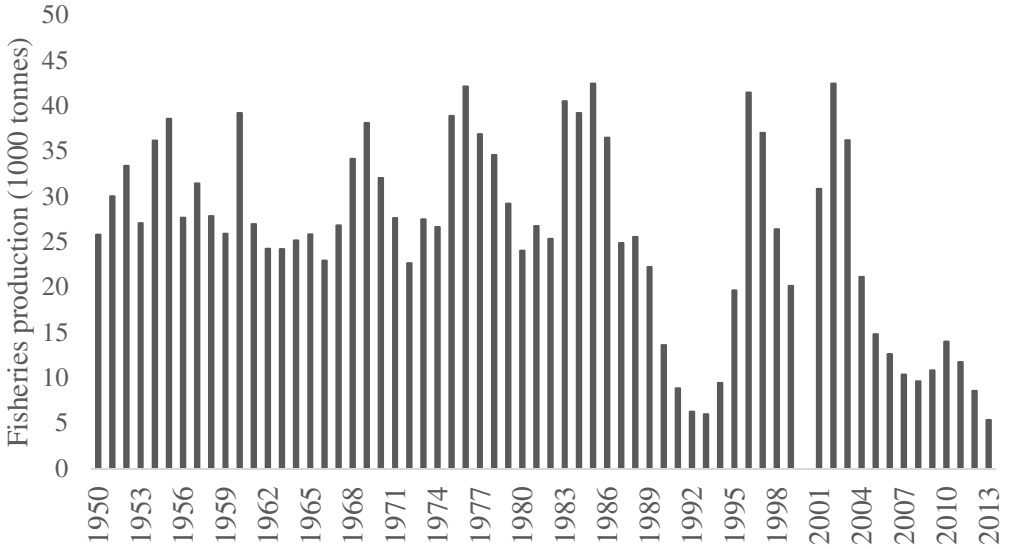

Figure 1.4. Landings of Atlantic cod caught in Faroese waters the last *c*. 60 years. Data for 1950 – 1989 from www.ices.dk (catch data for Atlantic cod caught in ICES subdivisions Vb1 [Faroe Plateau] and Vb2 [Faroe Bank]) and for 1990 – 2013 from www.hagstova.fo, accessed 17 September 2014 (year 2000 data were missing from the database used).

As reflected by catch data (Figure 1.4), the Faroese cod stocks have again reached a "risky" low level. While the latest stock assessments show no anticipation of strong incoming year classes on the Faroe Bank, short-term predictions are slightly positive for Faroe Plateau cod, where the spawning stock biomass is estimated at 25,000 t in 2014 and 2015 and at 27,000 t in 2016 (ICES 2014a). From 1965 cod from the Faroe Plateau and the Faroe Bank were managed as separate units (Magnussen 1996), however, since 1996 vessels have been allowed to fish in both areas during the same trip which has made interpretation of landing data from both areas problematic, although this is a bigger issue for the Faroe Bank stock (ICES 2014a) which is believed to be about 10 % of the size of the Faroe Plateau stock (Magnussen 2006). The Faroe Bank (< 200 m) is closed to trawling. Moreover, due to the severe state of the stock, a near total fishery ban has been in force for this area since 2009. Only a modest jigging fishery has been allowed during summer time (ICES 2014a).

1.6.3. Efforts to breed cod in the Faroe Islands

The Faroese aquaculture industry began in the mid sixties, at first with rainbow trout, but later the Atlantic salmon became the main aquaculture species and it still is. With over 60,000 t

produced per year in 2012 and 2013, over 40 % of the Faroese export value comes from farmed salmon (www.hagstova.fo, accessed 14 October 2014). Like in our neighbouring countries, interest for cod farming in the Faroe Islands started around the year 1990, when the cod fishery was at a low (Figure 1.4) and trials to breed cod were initiated by the Aquaculture Research Station of the Faroes. Juvenile production relied on a semi-intensive method, where larvae were fed a combination of naturally produced zooplankton and cultured *Artemia* (Van der Meeren *et al*. 2005; Paisley *et al*. 2010; Arge 2011; Kolbeinshavn *et al*. 2012). During the periods 1991 to 1994 and 2005 to 2009, totals of 200,000 and 700,000 juveniles were produced, respectively. These trials concentrated on the Faroe Bank cod, due to their particularly high growth rates, as outlined above, but a small number of Faroe Plateau juveniles were also produced (Arge 2011). Research mostly focused on growth differences between the Faroe Plateau and the Faroe Bank cod and on the effect of different first-feeding protocols (Fjallstein and Magnussen 1996; Arge 2011; Kolbeinshavn *et al*. 2012). There were great expectations concerning the Faroe Bank cod and they seemed to deliver. Cod in sea cages reached 4 kg in 28 months (after hatch) and sexual maturation was successfully inhibited up until this point by underwater artificial lightning (Kolbeinshavn *et al*. 2012). In addition to a few trials with landbased recirculation systems, this was the only commercial growth-on of cultured cod in the Faroe Islands (Arge 2011).

1.7. Main aims of the thesis

Despite the current consensus of some level of genetic differentiation between Faroe Plateau and Faroe Bank cod, genetic studies of Faroese cod are sporadic. In some studies, Faroe Plateau and / or Faroe Bank cod samples have only been included for comparative reasons, meaning that sample sizes have generally been small. Hence, an in depth investigation of the potential genetic differentiation between Faroe Plateau and Faroe Bank cod is still warranted, together with an evaluation of factors that may act to shape the genetic variation within and between the two stocks.

The current work grew out of initiatives to breed cod in the Faroe Islands, which primarily focused on breeding Faroe Bank cod in the natural habitat of the Faroe Plateau cod. For monitoring purposes, it was therefore of utmost importance to establish the level of a potential genetic differentiation between the two stocks and to investigate the possibility of designing marker panels that could discriminate between the two. Findings from these investigations were also expected to be of great relevance to the fisheries management sector. Furthermore, since cod farming in general is much hampered by high and unpredictable mortalities in the larval phase, experiments were conducted that addressed reproductive and early life history strategies in the two stocks as well as potential differential family survival rates of young life stages of cultured Faroe Bank cod. This leads to the following overall aims of the thesis:

1. To investigate early life history strategies for cod from the two areas.

2. To investigate if differential family survival rates in the egg stage could be correlated to survival rates in later fry stages and to evaluate potential predictors of survival rates in these two stages.

3. To perform studies of population genetic structure of cod from the Faroe Plateau and the Faroe Bank, using large and temporally replicated samples. Moreover, in order to put the results in context, samples from outlier (non-Faroese populatons) will be included.

4. To develop large numbers of SNP markers and investigate their potential to distinguish between Faroe Plateau and Faroe Bank cod.

Chapter 2. General materials and methods

2.1. Hatchery practices

2.1.1. Collection of wild broodstock

Faroe Bank broodfish were collected from April through August 2008 and the Faroe Plateau broodfish were collected in two rounds, in November 2007 and from April through August 2008. The fish were mainly caught by local fishers using jigging reels. The fish were transported to the Marine Research Centre of Fiskaaling (Aquaculture Research Station of the Faroes) where the Faroe Bank and the Faroe Plateau broodstocks were kept separately in outdoor flow-through tanks. Here, the fish were allowed to adapt to captivity for a minimum of five months prior to further handling. The goal had been to collect *c*. 100 fish from each area, but it proved very difficult to obtain Faroe Bank broodfish; hence, only 72 Faroe Bank broodfish were available for the breeding experiment compared to 146 Faroe Plateau broodfish.

2.1.1.1. Fish anaesthetisation

Prior to any handling, the broodfish were anaesthetised to "handleable" (defined as a loss of equilibrium and loss of reactivity to external stimuli) with a 100 mg l^{-1} concentration of the anaesthetic agent metacaine (MS-222). The fish were only handled for the minimum amount of time necessary and immediately after treatment were put in running seawater to recover.

2.1.1.2. Fish identification and registration of biological details

In January 2009 the fish were weighed and measured (total body length), phenotypic sex was determined by ultrasound using a veterinary ultrasound scanner (model 480, Pie Medical Equipment, Maastricht, the Netherlands: a method first described by Karlsen and Holm, 1994), and fin clips were taken and stored in 96 % ethanol at 4 ºC for DNA analysis. Fish identification was achieved by inserting a passive integrated transponder (PIT) tag (JoJo Automasjon, Sola, Norway) between the first and second dorsal fins, using a needle (sterile tag and needle). Each tag contained a unique code that could be read at any time by a hand-held reader to obtain the identity of the fish in question.

2.1.1.3. Calculation of Fulton´s condition factor (*K*)

Fulton´s condition factor (*K*) has been widely used in fisheries science and studies of fish biology in general as a way to describe the relative "well-being" of the fish from its length – weight relationship (Nash *et al*. 2006). It is calculated as $K = 100{,}000 \times$ body mass (g) divided by [total length (mm)]^3 (Ricker 1975).

2.1.2. Broodstock matings

Spawning periods of Faroe Plateau and Faroe Bank cod spawning in captivity were inferred from previous records of spawning activity of cod held in captivity at the Marine Research Centre 2006 – 2008, where Faroe Plateau cod had been observed to spawn in the period from mid to end of February to mid-April and Faroe Bank cod in the period from mid to end of March to mid-May (Table 2.1). These data were used to decide on the experimental dates used in this study, where there was a limited number of spawning tanks available, so the study had to be accomplished in two sequential phases, with only one stock spawning at a time. Consequently, Faroe Plateau cod were introduced into the spawning tanks and left to spawn between 18 February and 28 March. The tanks were cleaned and sterilised and the Faroe Bank cod were introduced into the tanks and left to spawn between 31 March and 28 April. Generally, cod do not feed during spawning and, in accordance with other similar studies, the fish were not fed while in the spawning tanks (Kjesbu *et al*. 1991,1996).

Table 2.1. Spawning dates observed for captive Faroe Plateau and Faroe Plaeau broodstock held at the Marine Research Centre.

Year	Faroe Plateau	Faroe Bank
2006	2 Mar - 14 Apr	22 Mar - 12 May
2007		10 Apr - 4 May
2008	19 Feb - 13 Apr*	10 Apr - 7 May
2009		15 Apr - 14 May

* Faroe Plateau cod, that had been collected in November 2007 as part of this work.

The majority of spawnings were carried out in circular, light grey, fibreglass reinforced plastic tanks, 1.3 m diameter × 1.5 m high which contained approximately 1.5 m^3 water (N = 30). A few larger silos, originally designed for rearing halibut larvae, were also employed as additional spawning tanks (1.5 m diameter × 4.0 m high; dark grey in colour; N = 5). All spawning tanks were situated indoors and were maintained at ambient photoperiod and temperature. Seawater was filtered (100 μm) and pumped into the tanks at a mean flow rate of 7 l min^{-1}. The effluent water was filtered through a 500 μm mesh net, which retained all eggs. Water temperature was recorded every day. To minimise the occurrence of failed spawning units due to possible male infertility and to stimulate more regular female spawning events, each spawning unit comprised one female and two males. Broodstock were size matched, based on the knowledge that matings with males of similar or slightly larger body lengths than the females are generally the most successful (Bekkevold *et al*. 2002; Herlin *et al*. 2008). There were initially 34 Faroe Plateau and 16 Faroe Bank spawning units, but broodstock mortality rates in the Faroe Plateau cod were high: a total of 37 individuals suffered mortality while in the spawning tanks. The fish were clearly stressed, many were swimming high in the water surface, but veterinary inspection of a large number of both dead and live fish did not provide any conclusive answers and infectious agents were not the primary cause of the mortalities. Histological examination, however, gave some indications that the fish may have suffered gas secretion problems, potentially as a result of poor water quality and / or as a consequence of collection, which may have caused irreversible damage to the swimbladder. However, some females were believed to be uncompromised by such issues, that is females that survived throughout the entire experiment, that were not moved between tanks, and that had available egg diameter data for three or more spawning events (Faroe Plateau: N = 10; Faroe Bank: N = 10). All except one of these ten Faroe Plateau females were caught in 2008.

2.1.3. From eggs to larvae

Eggs were collected daily and one or two dl of the live fertilised egg portions were transferred to incubation in aerated black 15 l cylinders with a cone-shaped bottom, where the incubation

temperature was kept constant at 5 ºC. One third of the water was exchanged daily and dead eggs were siphoned out of the bottom of the cylinders. Upon collection, batch-average egg diameters were assessed from a random sample of 50 eggs, using stereoscopes. Prior to and at the end of incubation the eggs were surface disinfected with glutaraldehyde at 400 ppm for 8 min.

Prior to hatching, 14 and 12 days post-fertilisation for the Faroe Plateau and Faroe Bank egg batches, respectively (calculated for the oldest egg batches), appropriate groups of egg batches were mixed and introduced into 420 l first-feeding tanks which were black on the inside. The earlier intervention for the Faroe Bank egg batches was necessary due an earlier hatching observed for this stock. Appropriate groups of egg batches were defined as batches that were spawned a maximum of two days apart (in order to prevent cannibalism among the larvae). Preferably, these groups should consist of egg batches from at least 10 females, although this criterion had to be relaxed somewhat. Based on these criteria, six groups of egg batches were collected, three from each stock, where each group contained eggs fertilised a maximum of two days apart. The number of batches represented in each group ranged from 7 to 10 (Table 2.2).

Table 2.2. Overview of the six groups of egg batches that were first-fed.

Group code	Spawning dates	Number of batches	Mean incubation temperature (ºC)	Time to 50 % hatch (days × ºC)
FP 1	6 - 8 Mar	8	5.5	94
FP 2 [a)]	16 - 18 Mar	10	5.5	100
FP 3 [a)]	18 - 20 Mar	7	5.5	94
FB 1	12 - 13 Apr	8	5.6	84
FB 2	14 - 16 Apr	7	5.7	88
FB 3 [b)]	17 - 18 Apr	8	5.9	89

[a)] These two groups included some of the same egg batches.
[b)] This group was distributed into three replicate first-feeding tanks, with full numbers (i.e. 45,400 eggs) in each.

45,400 eggs per group of batches, equal number of eggs from each female, were introduced into the first-feeding tanks. Numbers of eggs ml^{-1} (N) were calculated as $N = 1{,}222\ D^{-2.71}$, D = egg diameter (mm) (Kjesbu 1989). In the first-feeding tanks, date at hatch was monitored

and day zero was defined as the day when the first (presumably the oldest) larvae hatched. From day four, the larvae were fed rotifers (*Brachionus* spp.).

2.2. Collection of biological material for the population genetic analyses

The biological material used in the various population genetic analyses encompassed tissue samples of Faroe Plateau and Faroe Bank cod, kindly obtained on research surveys conducted by the Faroe Marine Research Institute, and tissue or DNA samples of Norwegian and White Sea cod, kindly provided by the Institute of Marine Research, Bergen. As outlined below, biological details were only available for the Faroese cod samples.

2.2.1. Collection of wild Faroese cod

Sampling of wild Faroese cod took place between 2008 – 2010 as part of surveys conducted by the Faroe Marine Research Institute with the research vessel RV "Magnus Heinason". Sampling was carried out during spawning in springtime, using as fishing gear a 116 feet box trawl with doors of the Steinham type. The length of the bridles was either 60 m (< ~ 140 m depth) or 120 m (in deeper water). Mesh size in the cod-end was 40 mm. Tow duration was one hour and the tow distance approximately three nautical miles, except for one haul on the Faroe Bank where the trawl, due to technical reasons, was in the water for a shorter time (Table 2.3). The fish were sacrificed by slitting the gills, which kills the fish quickly.

2.2.1.1. Registration of biological details

For each haul a random subsample of the total catch was taken and the fish were weighed, measured (total body length) and scored for sex and maturity stage by experienced personnel using a scale modified after Kesteven (1960): 1 = immature, 2 = mature, small gonads, 3 = mature, a little larger gonads, 4 = gonads fairly large, 5 = gonads fully developed, but not ripe, 6 = ripe gonads (running sperm / eggs), and 7 = spent individuals. In addition, otoliths were taken for subsequent age determination at the Faroe Marine Research Institute. The Fulton´s condition factor (*K*) was calculated as outlined above (Section 2.1.1.3).

Genetic analyses were undertaken for all trawl samples with relevant biological information collected (Table 2.3). The number of cod sampled per haul varied, partly due to the amount of cod in each haul and partly due to time available to process the fish (take fin clips and notes) before the next haul was on board. For each individual a fin clip was taken and stored in 96 % ethanol at 4 ºC for DNA analysis. All of these fish were used in the microsatellite analysis (Chapter 5), whereas subsamples were used in the other genetic analyses (Chapters 6 and 7).

Based on previous research as outlined in Section 1.6.1.1, Faroe Plateau cod caught north and east of the islands were pooled into one group (Faroe Plateau North-East or FPNE). These fish spawn on the main spawning ground to the North. Cod caught west of the islands were pooled into one group (Faroe Plateau West or FPW). These fish spawn on the main spawning ground to the West. Mainly due to weather conditions, the FPW area was only sampled for a single year and no cod were caught North-West of the islands, an area which also contains a smaller spawning ground (Joensen *et al*. 2005; Steingrund *et al*. 2005). There was no reason to believe that more than one stock component had been sampled on the Faroe Bank (FB) (Joensen *et al*. 2005); hence, these data were not divided into sub-areas prior to analysis.

Table 2.3. Overview of Faroese cod sampled for genetic analyses.

Area	Year	Lat.	Long.	*N*	Depth (m)	Age (years)	Total body length (mm)
Faroe Plateau North-East	2008	61.91	-6.40	3	94.0	2.7 (±0.6)	445 (±103)
		61.98	-6.26	6	121.5	4.7 (±1.2)	583 (±105)
		62.04	-6.20	4	119.0	4.5 (±1.7)	629 (±236)
		61.62	-6.13	2	209.0	4.5 (±0.7)	686 (±42)
		61.55	-5.93	4	174.0	6.3 (±2.6)	799 (±154)
		61.54	-5.76	2	194.5	5.0 (±0.0)	728 (±30)
		61.28	-5.27	1	231.0	4.0	584
		62.19	-6.28	19	84.5	3.3 (±0.9)	499 (±95)
		62.03	-5.75	1	193.0	4.0	549
		61.80	-5.56	5	161.0	4.6 (±0.5)	676 (±67)
		61.97	-4.92	2	195.0	5.0 (±0.0)	720 (±58)
		62.29	-4.12	2	344.0	4.0 (±1.4)	559 (±148)
		62.38	-4.56	4	337.0	6.0 (±0.8)	714 (±79)
		62.12	-6.16	21	108.0	5.1 (±1.5)	654 (±121)

Area	Year	Lat.	Long.	*N*	Depth (m)	Age (years)	Total body length (mm)
		62.28	-6.05	13	95.5	4.2 (±0.7)	581 (±64)
		62.36	-6.03	19	95.0	3.4 (±0.8)	482 (±73)
		62.39	-5.95	5	103.5	4.0 (±1.2)	542 (±88)
		62.46	-6.10	5	102.0	4.2 (±0.4)	548 (±77)
		62.54	-5.98	11	146.0	4.4 (±1.5)	618 (±124)
		62.60	-5.91	1	181.0	3.0	561
				130			
	2009	61.64	-6.10	2	215.0	4.0 (±0.0)	697 (±25)
		61.39	-5.76	2	211.0	4.5 (±0.7)	692 (±31)
		61.71	-5.40	1	178.0	4.0	596
		61.79	-5.55	3	157.5	4.0 (±1.0)	560 (±104)
		61.79	-5.17	3	221.0	3.3 (±0.6)	604 (±39)
		61.78	-6.35	9	90.5	3.0 (±0.5)	414 (±41)
		61.98	-6.26	25	121.0	4.6 (±1.0)	611 (±86)
		62.03	-6.19	20	120.5	3.6 (±1.1)	549 (±101)
		62.12	-6.16	20	108.0	4.6 (±1.3)	598 (±71)
		62.18	-6.26	11	88.0	3.9 (±1.5)	578 (±149)
		62.28	-4.07	3	364.0	6.3 (±3.1)	648 (±302)
		62.13	-5.30	5	165.0	3.8 (±1.6)	649 (±214)
		62.21	-5.40	6	162.5	5.0 (±0.9)	733 (±85)
		62.04	-5.71	3	184.5	3.7 (±0.6)	624 (±59)
		62.24	-7.61	9	100.0	2.1 (±1.5)	364 (±207)
		62.29	-7.64	8	101.5	3.5 (±1.8)	524 (±180)
		62.45	-7.41	22	97.5	4.4 (±1.9)	641 (±174)
		62.45	-7.07	18	108.5	5.1 (±0.7)	641 (±68)
				170			
	2010	61.97	-6.29	20	118.5	4.1 (±1.3)	597 (±89)
		62.12	-6.13	15	115.5	3.7 (±1.3)	568 (±126)
		62.05	-5.68	13	176.5	2.8 (±1.1)	541 (±126)
		62.29	-6.04	10	97.0	2.1 (±0.3)	405 (±54)
		62.38	-6.03	30	96.0	2.9 (±0.9)	493 (±90)
		62.53	-5.95	15	150.0	4.3 (±1.4)	649 (±118)
		62.54	-6.23	20	114.5	3.5 (±0.7)	577 (±73)
		62.63	-6.57	20	168.5	5.8 (±1.7)	768 (±128)
		62.47	-6.55	25	94.5	3.3 (±1.7)	519 (±141)
		62.44	-7.09	84	106.5	4.4 (±1.3)	634 (±105)
		62.45	-7.39	15	97.5	3.9 (±1.6)	580 (±99)
		62.53	-7.56	21	125.5	4.0 (±2.3)	631 (±143)
				288			
West	2009	61.88	-7.31	77	114.0	4.9 (±1.1)	608 (±71)
		61.79	-7.36	4	126.0	4.3 (±1.7)	607 (±137)
		61.71	-7.39	35	142.0	5.3 (±1.6)	723 (±95)
				116			
Faroe Bank	2008	61.03	-7.98	1	193.5	5.0	668

Area	Year	Lat.	Long.	*N*	Depth (m)	Age (years)	Total body length (mm)	
		61.17	-8.23	2	134.0	4.5 (±0.7)	799 (±32)	
		61.11	-8.37	2	99.5		910 (±167)	a
		61.10	-8.41	18	100.0	4.3 (±1.3)	793 (±101)	
		61.02	-8.27	34	116.5	4.6 (±1.3)	824 (±83)	
		60.86	-8.58	11	143.0	4.7 (±0.9)	834 (±38)	
		60.94	-8.43	14	118.5	4.2 (±1.2)	791 (±54)	
		61.02	-8.40	23	105.5	4.1 (±1.4)	800 (±107)	
		61.02	-8.56	5	105.5	5.0 (±1.4)	844 (±100)	
		61.01	-8.91	5	123.0	4.6 (±1.3)	841 (±92)	
		60.97	-9.10	7	138.5	4.7 (±1.4)	798 (±46)	
		60.89	-9.24	1	140.0	3.0 (±)	725	
		61.11	-8.57	11	108.5	4.8 (±1.2)	830 (±72)	
		61.20	-8.62	4	159.5	5.5 (±1.3)	853 (±60)	
				138				
	2009	61.26	-8.57	2	222.5	5.5 (±2.1)	835 (±126)	
		60.94	-8.42	10	119.5	4.8 (±1.5)	833 (±142)	
		61.02	-8.24	16	120.0	5.3 (±0.8)	865 (±58)	
		61.01	-8.42	6	107.0	4.7 (±1.6)	808 (±135)	
		61.02	-8.57	18	103.5	4.6 (±1.6)	806 (±134)	
		61.01	-8.92	11	124.0	5.2 (±2.3)	857 (±164)	
		60.98	-9.06	11	133.5	5.5 (±0.9)	873 (±50)	
		60.73	-9.11	3	131.5	6.7 (±2.1)	985 (±163)	
		60.78	-8.93	2	117.0	4.0 (±1.4)	802 (±161)	
		60.86	-8.95	4	104.0	3.5 (±2.4)	687 (±188)	
		61.11	-8.59	22	112.5	5.5 (±1.2)	865 (±54)	b
		61.11	-8.44	17	100.0	4.8 (±1.3)	834 (±99)	
				122				
	2010	61.12	-8.42	14	98.5	6.0 (±2.1)	912 (±88)	
		61.02	-8.23	21	120.5	5.7 (±1.2)	883 (±58)	
		61.02	-8.39	34	108.0	6.2 (±1.6)	892 (±82)	c
		61.02	-8.57	24	103.5	5.6 (±0.9)	892 (±62)	d
		61.01	-8.91	8	123.0	5.9 (±1.2)	890 (±85)	
		60.88	-9.19	4	135.0	5.8 (±0.5)	916 (±29)	
				105				

a) Due to technical reasons tow duration was less than the standard one hour and, therefore, these fish were not age determined by the Faroe Marine Research Institute.

b) No information on size, sex, maturity stage and age for nine individuals.

c) No information on size, sex, maturity stage and age for one individuals.

d) No information on size, sex, maturity stage and age for two individuals.

2.2.1.2. Larger fish caught on the Faroe Bank than on the Faroe Plateau

Sex ratios differed among areas and years with particularly few females for the Faroe Bank area. The overall percentages of females for the three areas, FPNE, FPW and FB, were 47, 63 and 38 %. Cod caught on the Faroe Bank were on average larger and of a higher condition than cod caught on the Faroe Plateau (with the Faroe Plateau West cod intermediate between the Faroe Bank and the Faroe Plateau North-East cod) (Table 2.4). One-way analyses of variance (ANOVA; Section 2.5.7) showed that body length and weight, condition and age differed significantly (*P*-values << 0.0001 in all cases) among the three areas (years grouped). This was not suspected to be due to a skewed sampling strategy, but rather a reflection of the larger sized cod on the Faroe Bank (Magnussen 2007).

Table 2.4. Sample sizes (*N*), average size, age (± SD), sex ratio (% F: percentage females) and proportion of actively spawning individuals (% S; maturity stage 6) of Faroese cod from the various areas and years. FPNE: the Faroe Plateau North-East area, FPW: the Faroe Plateau West area and FB: the Faroe Bank.

Area	Year	*N*	Body length (cm)	Body Weight (g)	Fulton *K*	Age (years)	% F	% S
FPNE	2008	130	58.5 (±12.8)	2451 (±1915)	1.052 (±0.105)	4.2 (±1.4)	47	48
	2009	170	58.7 (±14.5)	2571 (±2261)	1.055 (±0.095)	4.2 (±1.5)	52	48
	2010	288	59.4 (±13.2)	2537 (±1835)	1.037 (±0.086)	3.9 (±1.6)	43	47
Over all years		588	59.0 (±13.5)	2528 (±1981)	1.046 (±0.093)	4.1 (±1.5)	47	48
FPW	2009	116	64.2 (±9.6)	3094 (±1560)	1.084 (±0.095)	5.0 (±1.3)	63	70
FB	2008	138	81.4 (±8.4)	5992 (±1798)	1.082 (±0.105)	4.4 (±1.4)	39	72
	2009	122	84.0 (±11.7)	7144 (±2900)	1.137 (±0.090)	5 (±1.5)	45	44
	2010	105	89.4 (±7.2)	8186 (±2192)	1.126 (±0.094)	5.9 (±1.4)	27	60
Over all years		365	84.5 (±9.9)	6995 (±2475)	1.112 (±0.100)	5.0 (±1.5)	38	60

2.2.2. Norwegian coastal and White Sea cod samples

For comparative reasons, outlier (non-Faroese) samples were included in the genetic analysis. These samples were kindly provided by the Norwegian Institute of Marine Research. An overview of these samples is provided in Table 2.5 and the use of these samples is described in more details in the various sections on genetic analyses (Chapters 5 to 7). A map of all sampling locations is provided in Figure 2.1.

Table 2.5. Overview of the non-Faroese samples used in the various genetic analyses. Sampling year, sample type, number of individuals and geographical position (latitude and longitude as decimal degrees) are indicated.

Area	Year	Type	*N*	Lat.	Long.
Borgundfjord	2005	DNA	48*	62.45	6.26
North-West of Bomlø	2006	DNA	48*	59.90	4.98
Borgundfjord	2004	Tissue	26	62.45	6.26
Verrabotn	2005	Tissue	26	63.81	10.63
White Sea	2003	Tissue	26	65.10	39.00

* The DNA plate was damaged during transport, so only a subset of these could be analysed.

Figure 2.1. Sampling locations of cod for population genetic analyses. FB: Faroe Bank. FPW: Faroe Plateau West. FPNE: Faroe Plateau North-East. NWB: North-West of Bomlø, BGF: Borgundfjord. VB: Verrabotn. WS: White Sea. Figure created via Google Maps (Google Inc., Mountain View, CA, USA).

2.3. Genetic analyses

2.3.1. Extraction of DNA

2.3.1.1. Automated DNA extraction on a Maxwell® 16 platform

DNA was extracted from the Faroe Plateau and Faroe Bank cod samples collected in 2008 (Table 2.3) using the Maxwell® Tissue DNA purification kit on a Maxwell® 16 automated platform (Promega corporation, Ramcon, Birkerød, Denmark). The system uses prefilled reagent cartridges and MagneSil® paramagnetic particles that bind nucleic acids and moves them through a series of wash steps, followed by elution of the DNA in a low-salt solution (Staber and Mann 2010; Davis *et al*. 2012). Prior to loading the cartridges onto the instrument, approximately 35 – 50 mg of tissue (fin clip) was added into the buffer in the first well of each cartridge and 250 µl of elution buffer was added to the last well.

2.3.1.2. DNA extraction using the Real Pure Genomic DNA extraction kit

The Maxwell® system was designed for low to moderate throughput extractions; hence, due to the large number of samples in this study (and the associated costs) an alternative method was needed. Hence, the remaining samples were extracted with the Real Pure Genomic DNA extraction kit (Durviz, Valencia, Spain), that has performed well, as tested on a number of fish species at the Institute of Aquaculture molecular biology laboratory, and that could be adopted for 96 well plate extractions (Herlin 2008). Two variants of the manufacturer´s protocol were used, depending on the downstream analyses.

DNA for microsatellite genotyping was extracted in 96 well plates. In each well a small piece of tissue (fin biopsy of *c*. 2 mm in diameter) was transferred to 75 µl of "lysis solution", 3.0 µl of proteinase K (10 mg/ml) was added and the samples were digested to total lysis (4 hours to overnight) at 55 ºC. At room temperature 0.6 µl of RNase (10 mg/ml) was added and the samples incubated for 1 hour at 37 ºC. Precipitation of proteins was achieved by adding 45 µl of the "protein precipitation solution" followed by centrifugation at 4,100 g (RCF) for 20 min. DNA precipitation was achieved by transferring 50 µl of the supernatant (containing the DNA) to 75 µl of pure, cold isopropanol followed by mixing and centrifugation at 4,100 g for 10 min. After

carefully removing the supernatant (pipetting followed by turning the plate upside down on tissue paper), the DNA pellet was washed with 150 µl of cold 70 % ethanol followed by a last centrifugation step at 4,100 g for 5 min. After carefully removing the ethanol (pipetting and turning the plate upside down), the DNA was left to dry. Finally, the DNA was resuspended in 100 µl "hydration buffer". Some of these samples were also used for genotyping of SNPs. All reagents (excl. isopropanol and ethanol) were supplied with the kit.

To obtain larger quantities of high molecular weight genomic DNA, which was required for the RAD sequencing, a scaled up version of the Real Pure protocol described above was used. DNA extraction was accomplished in 1.5 ml tubes as follows: in each tube a small piece of tissue (fin biopsy of *c*. 3 mm in diameter) was transferred to 150 µl of "lysis solution", 10.0 µl of proteinase K (10 mg/µl) was added and the sample was digested to total lysis (3 hours at 55 ºC with periodical mixing). At room temperature 2 µl of RNase (10 mg/ml) was added and the sample incubated for 1 hour at 37 ºC. To precipitate the proteins, 90 µl of the "protein precipitation solution" was added, the sample mixed and placed at 4 ºC for 15 min followed by centrifugation at 13,400 g for 20 min. Wide tips were used to transfer 180 µl of the supernatant (containing the DNA) to 125 µl of pure (cold) isopropanol followed by mixing and centrifugation at 13,400 g for 10 min. After carefully pouring off the supernatant, the sample was centrifuged once more at 13,400 g for 1 min and the remaining supernatant was removed by pipetting. The DNA pellet was washed with 1 ml of cold 75 % ethanol and left overnight. The sample was centrifuged at 13,400 g for 5 min and after carefully pouring off the supernatant, the sample was centrifuged a last time at 13,400 g for 1 min and the remaining supernatant removed by pipetting. The samples were dried on the bench for 30 min (or on a heat block at 55 ºC for 5 min) to ensure that all ethanol had evaporated. The resulting DNA was resuspended in 50 µl 0.25 × TE buffer (1 × TE buffer: 10 mM Tris, 1 mM EDTA, pH 8.0). Some of these samples were also used for genotyping of SNPs.

2.3.2. Quantification of DNA

DNA quantity and purity was assessed by various methods, depending on the downstream application.

2.3.2.1. Quantification of extractions intended for microsatellite genotyping

For extractions to be used for microsatellite genotyping, random checks of the concentrations were taken (N = 71) using either an Implen NanoPhotometer® (Implen, München, Germany), kindly made available by the Genetic Biobank of the Faroe Islands, or an Eppendorf BioPhotometer Plus (Eppendorf, Horsholm, Denmark). Both of these spectrophotometric methods provided the DNA concentration as well as the ratio of the absorbance at 260 to the absorbance at 280 nm, which indicates the purity of the extraction with an expected ratio of ≥ 1.8 for pure DNA. The average concentration of the samples extracted on the Maxwell® 16 platform was 86 (± 58) ng/µl with an average 260/280 ratio of 1.77 (± 0.19) whereas the samples extracted with the Real Pure plate protocol had an average concentration of 61 (± 39) ng/µl with an average 260/280 ratio of 1.67 (± 0.23).

2.3.2.2. Quantification of extractions intended for RAD sequencing

All extractions to be used for RAD genotyping were quantified by a NanoDrop® 1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA) made for samples of small volumes (1 µl). DNA concentration of the extractions intended for RAD sequencing varied widely with an average concentration of 366 (± 319) ng/µl and an average 260/280 ratio of 1.87 (± 0.06).

2.3.2.3. Integrity of extractions assessed by agarose gel electrophoresis

The quality of the extractions intended for the RAD sequencing study was checked by agarose gel electrophoresis as follows: a 0.8 % agarose gel was made by melting 0.32 g of agarose powder in 40 ml 0.5 × TBE buffer (1 × TBE buffer: 90 mM Tris pH 7.6, 90 mM boric acid, 2 mM EDTA) and ethidium bromide added at 0.1 µg/ml. Gels were run at 8 V/cm for 45 min and visualised

under UV illumination (Ingenius Gel Documentation System; Syngene, Cambridge, UK). A lambda Hind III marker (Promega, UK) was used as a reference ladder.

2.3.3. Genotyping of microsatellites

Ten microsatellite loci were used that have been widely applied to study population structure in cod (Knutsen *et al*. 2003; Nielsen *et al*. 2006; O'Leary *et al*. 2007; Wirgin *et al*. 2007): *Gmo*2 and *Gmo*132 (Brooker *et al*. 1994), *Gmo*3, *Gmo*8, *Gmo*19, *Gmo*34, *Gmo*35 and *Gmo*37 (Miller *et al*. 2000) and *Tch11* and *Tch13* (O´Reilly *et al*. 2000). In the initial experiments two pentaplex systems were tested, with the loci *Gmo*8, *Gmo*19, *Gmo*35, *Gmo*37 and *Tch11* co-amplified in the first multiplex and *Gmo*2, *Gmo*3, *Gmo*34, *Gmo*132 and *Tch13* co-amplified in the second multiplex. The performance of some loci was poor, hence, three new multiplex systems were developed, two triplex systems containing *Gmo*8, *Gmo*19 and *Tch13* (Multiplex 1) and *Gmo*35, *Gmo*37 and *Tch11* (Multiplex 2) respectively, and one tetraplex system containing *Gmo*2, *Gmo*3, *Gmo*34 and *Gmo*132 (Multiplex 3) (Table 2.6). Polymerase chain reactions (PCRs) were performed in 9 µl volumes containing between 1 and 100 ng DNA (exact concentrations were not determined – see Section 2.3.2.1), 0.2 µM each forward and reverse primer, of which all forward primers were fluorescently labelled, 1 × Qiagen multiplex PCR buffer (contains a hotstart polymerase) and 0.5 × Qiagen Q-solution (Qiagen, Copenhagen, Denmark). Reaction conditions involved an initial denaturation step of 15 min at 95 ºC, followed by 25 cycles (except 24 for Multiplex 1) of 30 s denaturation at 95 ºC, 3 min annealing at 60 ºC (except 50 ºC for Multiplex 2) and 60 s extension at 72 ºC, and a final extension step of 30 min at 60 ºC.

The amplified PCR products were processed in 96 well plates on an ABI 3130*xL* Genetic Analyzer (Applied Biosystems, Life Technologies, Taastrup, Denmark). Each well contained 2 µl PCR product (Multiplexes 1 and 3 diluted 1:36; Multiplex 2 diluted 1:20), 7.90 µl Hi-Di formamide solution and 0.10 µl GeneScan -500 LIZ® size standard (reagents from Applied Biosystems). GeneMapper version 4.0 (Applied Biosystems) was used to score the genotypes, which were all inspected visually. Positive and negative control samples were included in each

plate and, to further check consistency of results, 14 % of the samples were re-amplified and re-run and allele calls were consistent between runs for all markers.

Table 2.6. Primer and locus information for the microsatellite markers used in the study.

Locus	Repeat Motif	Primer sequence (5´-3´)	Dye	Allele size range (bp)	Accession no.
		Multiplex 1			
*Gmo*8	GACA	F: GCA AAA CGA GAT GCA CAG ACA CC R: TGG GGG AGG CAT CTG TCA TTC A	NED	108-316	AF159238
*Gmo*19	GACA	F: CAC AGT GAA GTG AAC CCA CTG R: GTC TTG CCT GTA AGT CAG CTT G	VIC	118-230	
*Tch*13	GT	F: TTT CCG ATG AGG TCA TGG R: AAT CCA CTG GTG CAG ACC	6-FAM	76-172	AF178503
		Multiplex 2			
*Gmo*35	ACC	F: GGA GGT GCT TTG AAG ATG R: CCT TAT CAT GTA CGT TGT TAA C	6-FAM	114-159	AF159235
*Gmo*37	GACA	F: GGC CAA TGT TTC ATA ACT CT R: CGT GGG ATA CAT GGG TAC CT	6-FAM	195-307	AF159237
*Tch*11	GATA	F: ATC CAT TGG TGT TTC AAC R: TCG AGT TCA GGT GGA CAA	PET	110-230	AF178501
		Multiplex 3			
*Gmo*2	GT	F: CCC TCA GAT TCA AAT GAA GGA R: GTG TGA GAT GAC TGT GTC G	NED	103-147	
*Gmo*3	GACA	F: AGG CAC GCA GGT GGA CAG GAA C R: GCA GCA CGA GAG AGC TAT TCC TC	6-FAM	159-207	AF159233
*Gmo*34	GACA	F: TCC ACA GAA GGT CTC CTA A R: GGT TGG ACC TCA TGG TGA A	VIC	74-118	AF159234
*Gmo*132	GT	F: GGA ACC CAT TGG ATT CAG GC R: CGA AAG GAC GAG CCA ATA AC	PET	105-221	AF159238

2.3.4. RAD library preparation and sequencing

The RAD library preparation protocol followed closely the methodology originally described by Baird *et al*. (2008) and later comprehensively detailed (Etter *et al*. 2011). The RAD-specific P1 and P2 paired-end adapters and library amplification PCR primers were supplied by the GenePool, University of Edinburgh, their sequence composition being given in Baxter *et al*. (2011). The library construction work was undertaken at the Institute of Aquaculture molecular

biology laboratory, where many previous RAD libraries had been constructed. The available budget ran to a single lane of Illumina HiSeq NGS sequencing. For this it was decided to screen 40 barcoded samples in the RAD analysis; striking a balance between the expected number of RAD loci detectable (using *Sbf*I restriction enzyme with an 8 base recognition sequence – CC_TGCA^GG – and estimated from similar libraries produced for other fish species) and the required sequencing depth for reliable SNP calling (at least 15 × coverage per locus per sample).

All DNA samples (previously treated with RNase to remove residual RNA) were quantified by spectrophotometry (Section 2.3.2.2) and quality assessed by agarose gel electrophoresis (Section 2.3.2.3). Those that passed quality control (no observable RNA and comprising predominantly high molecular weight DNA), were selected for use and diluted to a concentration of 50 ng/µl in 5 mM Tris; pH 8.5. A total of 40 Faroe Plateau and Faroe Bank samples were selected, 10 fish from 2009 sampling and 10 fish from 2010 sampling for each stock – each fish to be individually barcoded. Biological information on these fish is provided in Table 7.1 (Section 7.2.2.1). Details of the 40 barcoded samples used in the RAD analysis are given in Table 2.7.

Table 2.7. Details of cod samples and information on RAD library construction.

RAD library	Sample	Barcode
1	FP 2009 14	CGATA
	FP 2009 21	CTAGG
	FP 2009 22	CTGAA
	FP 2009 25	CGCGC
	FP 2009 15	GCTAA
	FP 2009 23	GGCCT
	FP 2009 27	GTCAC
	FP 2009 24	TAGCA
	FP 2009 26	TCCTC
	FP 2009 28	TGTGG
2	FP 2010 43	ATTAG
	FP 2010 16	AGAGT
	FP 2010 41	CGTAT
	FP 2010 25	CTCTT
	FP 2010 31	CCTTG
	FP 2010 49	GACTA
	FP 2010 17	GCGCC
	FP 2010 18	GTTGT
	FP 2010 20	TACGT
	FP 2010 37	TCGAG
3	FB 2009 07	TGACC
	FB 2009 08	TTTTA
	FB 2009 10	ATCGA
	FB 2009 12	CCCCA
	FB 2009 15	CGGCG
	FB 2009 17	GATCG
	FB 2009 13	GCATT
	FB 2009 19	GGTTC
	FB 2009 26	TATAC
	FB 2009 27	TTCCG
4	FB 2010 14	AGCTG
	FB 2010 03	AGTCA
	FB 2010 05	CTTCC
	FB 2010 16	CATGA
	FB 2010 11	CACAG
	FB 2010 02	GAGAT
	FB 2010 29	GCCGG
	FB 2010 04	GTGTG
	FB 2010 21	GTACA
	FB 2010 27	TCTCT

Each sample (0.25 µg DNA) was digested at 37 °C for 30 min with *Sbf*I high fidelity restriction enzyme (New England Biolabs – NEB) using 6U *Sbf*I per µg genomic DNA in 1 × Reaction Buffer 4 (NEB) at a final concentration of 1 µg DNA per 50 µl reaction volume. The reactions (12.5 µl final volume) were then heat inactivated at 65 °C for 20 min. Individual specific P1 adapters, each with a unique five base barcode (Table 2.7), were ligated to the *Sbf*I digested DNA at 22 °C for 30 min by adding 1 µl 100 nM P1 adapter, 0.15 µl 100 mM rATP (Promega), 0.25 µl 10 × Reaction Buffer 2 (NEB), 0.125 µl T4 ligase (NEB; 2 M U/mL) and reaction volumes made up to 15 µl with nuclease free water for each sample. Following heat inactivation at 65 °C for 20 min, the ligation reactions were then combined in appropriate library pools of 10 barcoded samples (Table 2.7). For each library pool 100 µl (*c*. 1.7 µg digested / ligated DNA) was sheared to *c*. 150 – 700 bp size range (Covaris sonicator). The sheared DNA was column purified (PCR MinElute Kit, Qiagen), being eluted in 35 µl EB buffer (Qiagen). Each of the four library samples were then size selected (*c*. 250 – 500 bp) by gel electrophoresis (0.5 × TAE; 1.1 % gel). Gels were run (2 V/cm for 10 min; 8 V/cm for 50 min) in ice-cold buffer – to minimise small fragment diffusion. Each library sample was flanked by a marker lane with fragments 190 and 510 bp in size. The sections of the gel containing the size separated libraries were carefully excised and stored at 4 °C while the remainder of the gel, including flanking DNA size markers was quickly stained with EtBr, viewed under UV, and the appropriate size range flagged by "nicking" the side of each marker lane. The gel was then reassembled and the identified size selected band was excised using a clean scalpel blade (Figure 2.2). In this way, the size-selected DNA was not exposed to EtBr or UV radiation. The remainder of the library construction (i.e. gel purification; end repair, dA overhang addition, P2 paired-end adapter ligation and library amplification) followed the original protocol (Etter *et al*. 2011) exactly. A total of 150 µl of each amplified library (16 PCR cycles) was prepared, column purified, eluted in 35 µl EB buffer and size selected (*c*. 300 – 550 bp) by gel electrophoresis, as described above. Following a final gel elution step into 20 µl EB buffer (MinElute Gel Purification Kit, Qiagen), the libraries were QCed by electrophoresis (agarose gel; Figure 2.3) and accurately quantified by fluorimetry. Equimolar amounts of each library were combined and sequenced on a single lane

of the Illumina HiSeq 2000 platform (100 base, paired-end reads) at the GenePool Genomics Facility, University of Edinburgh.

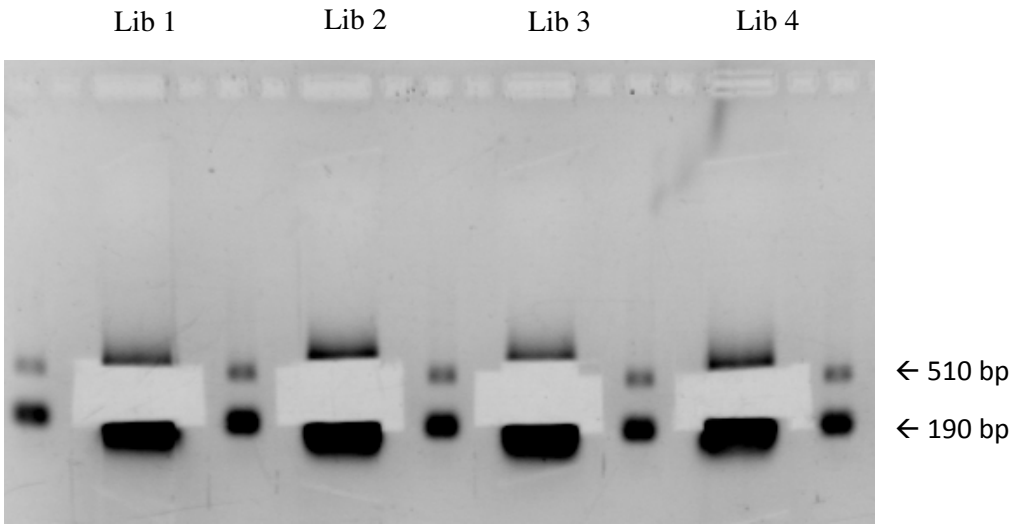

Figure 2.2. Image of RAD size-selection gel following marking of a 190-510 bp region and excision of the genomic DNA gel fragments from all four library preparations.

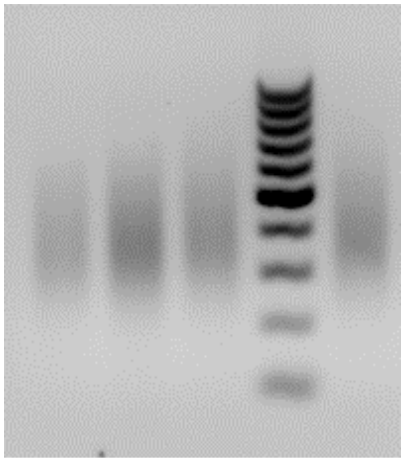

Figure 2.3. Image of the amplified libraries on an agarose gel. Libraries 1 to 4 from left to right with a DNA ladder (100-1000 bp in 100 bp increments) in between library 3 and 4.

2.3.5. Genotyping of single nucleotide polymorphisms (SNPs)

The great majority of single nucleotide polymorphisms (SNPs) analysed in this work were analysed by the Kompetitive Allele Specific PCR genotyping system (KASP™; LGC Genomics, Middlesex, UK). With this method, known SNPs can be genotyped through the competitive

binding of two allele-specific forward primers, that each has a unique tail sequence that corresponds with one of the two universal fluorescence resonant energy transfer (FRET) cassettes. When the relevant primer binds to the DNA template and elongates, the complement of the tail sequence is generated. This will cause the FRET cassette to bind and emit fluorescence, which in the case of a heterozygote is a mixed fluorescent signal.

KASP assays were supplied by LGC Genomics (KBD service – KASP by design), while the SNP screening was undertaken at the Institute of Aquaculture molecular biology laboratory. Each reaction assay (8 µl total volume) comprised *c*. 20 ng genomic DNA, 0.14 µl of supplied SNP-specific KASP primer mix, 4.0 µl KASP assay × 2 mastermix and made up to final volume with nuclease-free water. Assays were performed on a Techne Quantica real-time thermocycler using proprietory analysis software. Cycling conditions were 94 °C for 15 mins (hot-start activation), followed by 10 cycles of touchdown PCR (94 °C for 20 s, 61-55 °C for 60 s - dropping 0.6 °C per cycle); followed by 26 cycles of 94 °C for 20 s and 55 °C for 60 s.

2.4. Calculation of genetic diversity and estimation of population differentiation

The basic techniques and assumptions underlying the population genetic analyses are described here. Further details about population genetic calculations can for example be found in Frankham *et al*. (2007).

2.4.1. Description of genetic diversity of populations

A locus is a segment of DNA and alleles are different forms of the same locus that differ in DNA base sequence (e.g. A_1 and A_2). Genotypes are the combination of alleles at a locus in an individual and they can be homozygous (e.g. A_1A_1) or heterozygous (A_1A_2). A population genetic dataset usually consists of the number of genotypes per investigated loci. If a locus has two different alleles (A_1 and A_2) and p is used to describe the estimated population frequency of allele A_1, it can be calculated from the dataset as follows:

$$p = \frac{(2 \times A_1A_1) + A_1A_2}{2 \times total}$$

Genetic diversity was described in terms of number of alleles, allelic richness (number of alleles standardised to the smallest sample size in the dataset) and observed and expected heterozygosities (Section 2.4.2).

2.4.2. The Hardy-Weinberg equilibrium

The Hardy-Weinberg equilibrium is a central concept in population genetic analyses, which makes a number of assumptions, including random mating, no natural selection, no mutation, no gene flow in or out of the population and an infinitely large population size. When these assumptions are met, the Hardy-Weinberg equilibrium can be used to predict the proportion of diallelic genotypes in the next generation from the initial allele frequencies. Again, using an example with a diallelic locus where p is used to describe the frequency of allele A_1 and q is used to describe the frequency of allele A_2, the expected genotype frequencies of A_1A_1, A_1A_2 and A_2A_2 at Hardy-Weinberg equilibrium will be p^2, $2pq$ and q^2, respectively. Conformance of observed and expected frequencies can be tested with a Chi-square goodness-of-fit test (Section 2.5.6). When there are significant differences between observed and expected frequencies, it may indicate disturbance by some of the forces described above. Observed heterozygosity (H_O) is calculated as the number of heterozygotes at a locus divided by the total number of specimens sampled for that locus. In the case of more than two alleles per locus, expected heterozygosity (H_E) can be calculated as:

$$H_E = 1 - \sum_{i=1}^{No.of\ alleles} p_i^2$$

where p_i describes the frequency of the i^{th} allele. This is also called gene diversity.

2.4.3. Estimation of the level of differentiation among populations

Population subdivision will result in some level of genetic differentiation that is directly related to the inbreeding coefficients within and among populations. Sewall Wright used this principle to describe the hierarchical F statistics, which estimate the proportion of variation (i.e. heterozygosity) found between one level in the hierarchy and a more inclusive level. F_{IS} (also

called the inbreeding coefficient) is the proportion of subpopulation variation that is found among individuals, F_{IT} the proportion of total variation that is found among individuals and F_{ST} the proportion of total variation that is found among subpopulations. Calculation of the three can be achieved as follows:

$$F_{IS} = 1 - \frac{H_I}{H_S}$$
$$F_{IT} = 1 - \frac{H_I}{H_T}$$
$$F_{ST} = 1 - \frac{H_S}{H_T}$$

where H_I is the observed heterozygosity averaged across subpopulations, H_S is the expected heterozygosity averaged across subpopulations and H_T is the expected heterozygosity for the total population.

F_{ST} is also called the fixation index and it describes the probability that two randomly drawn alleles from a subpopulation are identical by descent. F_{ST} ranges from zero (no differentiation among subpopulations) to one (fixation of different alleles in subpopulations).

2.4.3.1. Alternatives to F_{ST}

Values of F_{ST} and analogues, which have traditionally been used to describe the genetic structure of populations, vary with within-population variation, so that for highly variable markers such as microsatellites, even maximum differentiation among populations (no alleles shared) will result in small F_{ST} values (Meirmans and Hedrick 2011, Jost 2008). Hedrick (2005) defined G'_{ST}, which standardised G_{ST} by the maximum value it could obtain given the subpopulation heterozygosity. In a similar manner, a standardised F_{ST} estimate (F'_{ST}) can be obtained (Meirmans 2006), which ranges from zero (populations have similar allele frequencies) to one (populations fixed for different alleles). Whereas these estimates deal with the limitations set on maximum differentiation by heterozygosity levels, Jost (2008) described a second issue, the non-linear relationship between heterozygosity and diversity. He therefore proposed a new statistic, D_{Jost}, which measures "actual differentiation" based on the effective number of alleles rather than on

heterozygosity (Jost 2008). In the present study, both F_{ST} and D_{Jost} were applied to investigate structure based on the microsatellite dataset. For pairwise comparisons, an estimator of D_{Jost} (D_{Jost_EST}) was applied. This estimator reduces the bias originating from the fact that sample frequencies, and not true population frequencies, are used in the calculation of D_{Jost}.

2.4.3.2. Estimation of gene flow

Although D_{Jost} is a useful statistic for estimating population differentiation, it relies heavily on mutation rates, is independent of population sizes and inherently less suited for estimating demographic processes such as migration rates and effective population sizes (Jost 2009; Meirmans and Hedrick 2011). For such purposes, F_{ST} is still a good choice. Hence, for the microsatellite data, overall and pairwise values for the effective number of migrants per generation (Nm) were calculated as $(1 - F'_{ST})\ (4\ F_{ST})^{-1}$ (Meirmans and Hedrick 2011). This formula allows for an estimate of Nm which is unaffected by sub-population heterozygosity and is a modification of the original formula $(1 - F_{ST})\ (4\ F_{ST})^{-1}$. Nm was also inferred from the private allele method (Slatkin 1985; Barton and Slatkin 1986) implemented in GENEPOP version 4.2 (Raymond and Rousset 1995; Rousset 2008; available online at www.genepop.curtin.edu.au).

2.5. Statistical analysis

Most statistical calculations were conducted in Systat (version 11.0, Systat Software, Inc.; www.systat.com) and further details about the techniques can be found in Walpole *et al*. (2012).

2.5.1. Estimation of the mean

The population mean was estimated by the sample mean ($\bar{x}$):

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$$

Each single observation is described by x and n is the total number of observations.

To describe the spread or variability of the data, the mean was presented together with ± 1 standard deviation (SD):

$$\mathrm{SD} = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n-1}$$

2.5.2. Normality assumptions

Prior to performing the statistical tests, the data were checked for conformance to normality and homogeneity of variance assumptions. Data distributions were checked for conformance to the neutral distribution by using the Kolmogorov-Smirnov test and for homogeneity of variance by using the F-test.

2.5.3. Comparison of two sample means

The unpaired *t*-test was used to compare the means of two independent samples assumed to be drawn from populations with normally distributed data. Population variances were unknown and assumed to be unequal. Also, no prior assumptions were made as to which mean might be larger in the case of a significant outcome of the test, i.e. two-tailed critical regions of the test statistic were used. The test statistic, *t*, was calculated as follows:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

s^2 = sample variance = SD^2 and the degrees of freedom (*df*) used when consulting the *t*-distribution table was calculated as:

$$df = \frac{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1 - 1)} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2 - 1)}}$$

2.5.4. Simple linear regression

The simple linear relationship between a variable *Y* and an independent variable *x* was valuated by the model:

$$Y = \beta_0 + \beta_1 x + \mathcal{E}$$

β_0 and β_1 denote the intercept and slope parameters, respectively, and $\mathcal{E}$ describes the random error of the model. The regression coefficients were estimated by the method of least squares as follows:

$b1$ = the slope of the line = $\frac{\sum_{i=1}^{n}(x_i - \overline{x})\,(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$

and b_0 = intercept = $\overline{y} - b_1\,\overline{x}$.

The degree of linear relationship between two variables was assessed by calculating the Pearson product-moment correlation coefficient (*r*) as follows:

$$r = b_1 = \frac{S_{xx}}{\sqrt{S_{xx}S_{yy}}}$$

with $S_{xx} = \sum_{i=1}^{n}(x_i - \overline{x})^2$ and $S_{yy} = \sum_{i=1}^{n}(y_i - \overline{y})^2$

The significance of the regression was tested with analysis of variance (ANOVA; Section 2.5.7).

2.5.5. Multiple linear regression analysis

For *k* independent variables ($x_1, x_2, \ldots, x_k$), the estimated response is obtained by:

$$\hat{y} = b_0 + b_1 x_1 + \cdots + b_k$$

As described above, the regression coefficients were estimated by the method of least squares.

t-tests (Section 2.5.3) were used to test whether variables made a significant contribution to the multiple linear regressions.

2.5.6. Chi-aquare goodness-of-fit test

To test if expected frequencies conformed to those expected based on some theoretical assumptions, the goodness-of-fit test was applied. The test statistic χ^2 was calculated as follows:

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

o_i and e_i describe the observed and expected frequencies, respectively, for the i^{th} cell and k is the total number of cells. Significance was estimated by consulting a table of the chi-square distribution ($df = k - 1$).

2.5.7. Analysis of variance (ANOVA)

A one-way analysis of variance (ANOVA) was used to compare the means of more than two samples, assuming that samples were independent and randomly drawn from normally distributed populations with equal variances. The null hypothesis that samples were drawn from normally distributed populations with equal means and variances was tested by partitioning the total variability (estimated by sum of squares) into its components (variability within and between samples).

Using the same principles as in a one-way ANOVA, a two-way analysis of variance (ANOVA) was used to estimate the effects of two independent variables (as well as their interaction) on a dependent variable.

Chapter 3. A comparison of early life history strategies for Faroe Plateau and Faroe Bank cod

3.1. Introduction

Reproductive parameters are key factors used in defining Atlantic cod stocks / populations and a central concept in the study of early life history strategies is the trade-off between size and number of progeny (Kamler 2005). Such trade-offs are thought to reflect adaptations to various environmental factors and, in teleost species, egg size – egg number trade-offs have been demonstrated among as well as within species (Elgar 1990; Kume 2011). For example, Kume (2011) investigated two migratory forms of threespine stickleback, one form that predominantly breeds in brackish waters and another form that breeds in freshwater, and found the former to produce relatively many but small eggs, while the opposite was true for the latter form.

Environmental conditions as well as temperatures differ between the Faroe Plateau and the Faroe Bank areas (Section 1.6.1) with spawning temperatures 1 – 2 ºC higher on the Faroe Bank than on the Faroe Plateau (Magnussen 2002; Steingrund *et al*. 2005; Magnussen 2007). Two decades ago, a M.Sc. study on Faroe Plateau and Faroe Bank cod found significant stock differences in egg sizes with larger eggs for the Faroe Plateau (1.42 mm in diameter) than the Faroe Bank cod (1.31 mm in diameter) (Magnussen 1993). Cod from the two areas are also known to differ with respect to growth rates. For both stocks, the 50 % maturation rate occurs at around three years of age (Steingrund *et al*. 2005). However, whereas, on average, a three-year-old cod from the Faroe Bank is about 73 cm long and weighs approximately 4.9 kg, a cod of the same age from the Faroe Plateau is, on average, 55 cm long and weighs 1.7 kg (Magnussen 2007, 2011).

In nature, cod on the Faroe Plateau and cod on the Faroe Bank exhibit spawning between February and May with peak spawning time for the Faroe Plateau cod in the second half of March (í Jákupsstovu and Reinert 1994). Compared to the Faroe Plateau cod, there are indications of a more prolonged spawning period for the Faroe Bank cod (Steingrund *et al*. 2005). For the Faroe Plateau cod, a correlation has been demonstrated between timing and location of spawning and

the zooplankton peak (Gaard and Steingrund 2001). Apart from this, not much is known about the reproductive strategies of the respective Faroe Plateau and Faroe Bank cod.

The main objectives of the present study were to investigate whether cod from the two areas differed with respect to early life history strategies, whether the findings of different egg sizes for the stocks could be replicated and whether or not they would translate into different number of eggs spawned, as predicted by the egg size – egg number trade-off theory (Kamler 2005). Comparisons of reproductive parameters (time of spawning, egg and larval sizes as well as female gonadal development as a proxy of fecundity) between Faroe Plateau and Faroe Bank cod are presented and, furthermore, potential predictor variables of egg sizes are explored. Both field data, kindly provided by the Faroe Marine Research Institute, and observations from experimental captive spawnings were used. In addition to being important from a conservation point of view, detailed knowledge of reproductive strategies has wider application to the fishery management sector, for example in refining models which link individual fecundity and spawning-stock biomass to subsequent recruitment to the fishery. Furthermore, detailed knowledge of reproductive potential and strategies of fish stocks is important for the further development of a cod aquaculture industry, where one major challenge is to achieve a stable juvenile production.

3.2. Material and methods

3.2.1. Field study

Data collected by the research vessel RV "Magnus Heinason" 1994 – 2010 were kindly provided by the Faroe Marine Research and used to estimate spawning periods as well as gonadal volumes, as a proxy of fecundity, of cod spawning in the wild. The trawl surveys covered depths from 65 to 700 m, although most stations were between 100 and 300 m. Faroe Plateau spring samples were collected between mid-February and late March at 100 stations during daytime and Faroe Plateau autumn samples were collected in August at 200 stations, taken both day and night (the latter from 1996 onwards). Faroe Bank samples were collected at 29 stations, during daytime in both seasons (mid-March – early April and September). Details about survey gear and trawling

procedures are given in Section 2.2.1 and biologicals details of sampled fish were obtained as described in Section 2.2.1.1.

3.2.2. Captive study

3.2.2.1. Collection of broodstocks

Collection of broodstocks is described in Section 2.1.1. The latest broodstock collections were in August 2008, hence, the fish were allowed to adapt to captivity for at least five months prior to tagging and collection of biological details in January 2009, as described in Sections 2.1.1.1 and 2.1.1.2.

3.2.2.2. Spawning experiment

Broodstock matings are described in Section 2.1.2. Based on previous records of spawning activity of cod held in captivity at the Marine Research Centre, Faroe Plateau cod were introduced into the spawning tanks and left to spawn between 18 February and 28 March and Faroe Bank cod between 31 March and 28 April. Data collection concerning eggs and larvae is further explained in Section 2.1.3. Six groups of egg batches were collected, three from each stock, where each group contained eggs fertilised a maximum of two days apart. Number of batches represented in each group ranged from 7 to 10 batches. Each group was transferred to a first-feeding tank, with the exception of one Faroe Bank group, which was transferred to three replicate first-feeding tanks. At day 14 after hatch random samples of ten larvae per first-feeding tank were taken for length measurement (total body length in mm).

3.2.3. Statistical analyses

3.2.3.1. Estimation of spawning period and gonadal volume of wild-spawning cod

Definition of maturity stages is provided in Section 2.2.1.1. The proportion of spawning and spent females – maturity stages 6 and 7, respectively – were used to estimate spawning time (data range 1994 – 2010). Data were divided into 5-days intervals with between 27 and 3336 (on average 1205) fish per interval being recorded. The difference between average Fulton´s condition factor in spring (maturity stage 5) and autumn (maturity stage 2) was used as a comparative indicator

of gonad development for females (data range 1996 – 2010). For the Faroe Bank spring data there were only a small number of females of length 70 cm and smaller ($N = 14$); hence, for comparative reasons, only large females (i.e. 80 – 110 cm) were included in this analysis. Dr. P. Steingrund, Scientist at the Faroe Marine Research Institute, conducted these analyses.

3.2.3.2. Estimation of reproductive parameters of wild-caught cod spawning in captivity

Female sizes, batch intervals, number of batches and egg and larval sizes were compared between the stocks using *t*-tests. Within each stock, the potential effects of female condition (Fulton´s condition factor, *K*) and experimental day number on batch-average egg diameters were explored by multiple linear regression analyses. Experimental day number one = first day of monitoring in the spawning tank, i.e. 19 February and 1 April for the Faroe Plateau and Faroe Bank cod, respectively. Except for the test regarding larval sizes, tests were only performed on data for the ten females per stock that had spawned three or more batches and were believed to be unaffected by stress or weakened in any other way (Section 2.1.2). This was not practiced when analysing larval sizes, since it was not known which families the sampled larvae belonged to.

3.3. Results

3.3.1. Field study

Data collected over a period of 17 years indicated that, although spawning times at the two locations overlapped, peak spawning occurred *c*. 10 days earlier on the Faroe Plateau than on the Faroe Bank. Peak spawning on the Faroe Bank occurred when the highest proportion of spent cod were recorded on the Faroe Plateau (Figure 3.1).

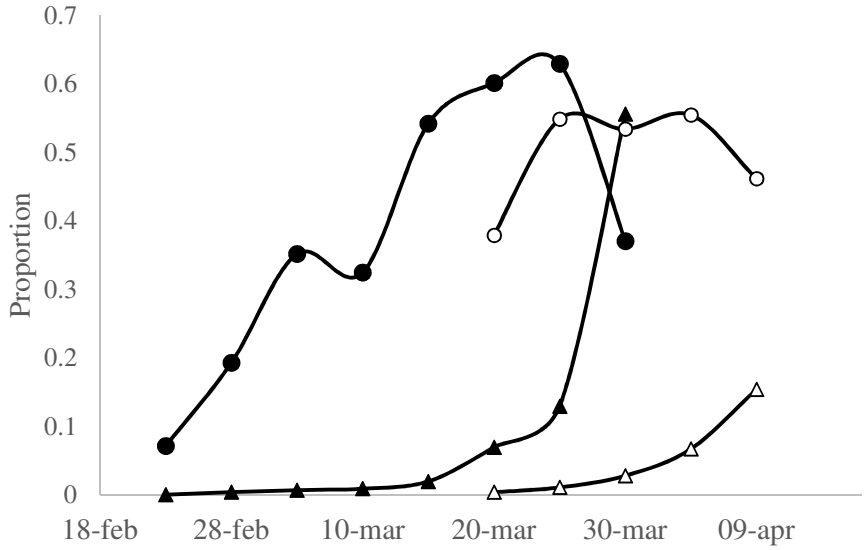

Figure 3.1. Spawning season for cod on the Faroe Plateau and the Faroe Bank 1994 – 2010 as indicated by the proportion of spawning (circles) and spent (triangles) cod. Filled markers for the Faroe Plateau and open for the Faroe Bank. The dates represent midpoints of 5-day intervals.

The decline in average condition factor, *K*, from spring to autumn was 0.116 and 0.144 for the Faroe Plateau and Faroe Bank females, respectively (Table 3.1). This change (approximately 25 % higher for the Faroe Bank) indicated greater gonadal volumes for the Faroe Bank females.

Table 3.1. Average Fulton´s condition factor (*K*) and total body length (± SD) of large (80 – 110 cm) sexually mature females on the Faroe Plateau and the Faroe Bank in spring and autumn 1996 – 2010.

	Faroe Plateau		**Faroe Bank**	
	Spring	Autumn	Spring	Autumn
N	274	900	646	1257
Length (cm)	88.8 (± 7.6 SD)	90.1 (± 8.2 SD)	89.4 (± 10.5 SD)	85.1 (± 13.0 SD)
Condition (*K*)	1.111 (± 0.127 SD)	0.995 (± 0.107 SD)	1.210 (± 0.114 SD)	1.066 (± 0.100 SD)

3.3.2. Captive study

3.3.2.1. Spawning dates

While the breeding experiment was not set up to investigate the spawning periods *per se*, an overview of the batches spawned indicated that the experiment had covered the bulk spawning periods for the two respective stocks, although the Faroe Bank females might have had more

batches left than the Faroe Plateau females when the respective monitoring periods ended (Figure 3.2).

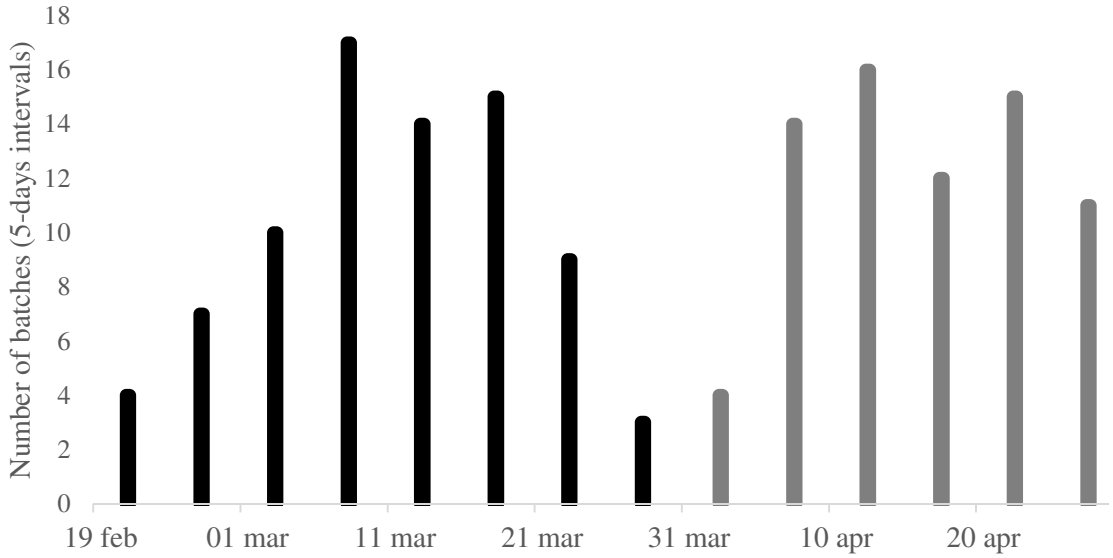

Figure 3.2. Number of batches spawned by the 10 Faroe Plateau (black bars) and 10 Faroe Bank females (grey bars) every five days of the breeding experiment.

3.3.2.2. Egg development times

Mean observed times to hatch, as observed for the various groups of egg batches, ranged from 94 to 100 degree-days (days × ºC) for the Faroe Plateau and 84 to 89 degree-days for the Faroe Bank eggs (Table 2.2). Despite roughly equal incubation temperatures, all three groups of Faroe Bank egg batches used fewer degree-days to hatch compared to the Faroe Plateau replicates. In addition to these observations, time to hatch was also observed for one single batch of Faroe Plateau eggs and three single batches of Faroe Bank eggs, that were not used in any group of batches and could therefore be allowed to hatch in the incubation cylinder (with a constant temperature of 5 ºC). The Faroe Plateau egg batch hatched after 21 days and all three Faroe Bank batches hatched after 14 days. A model by Geffen *et al.* (2006) predicts that cod eggs hatch after 18 days (95 % CI 15 – 21 days) at 5 ºC, so the Faroe Plateau observation was in the scope of what could be expected, although in the upper range. All three Faroe Bank observations, however, were slightly outside the lower 95 % confidence interval suggesting that Faroe Bank cod eggs develop significantly faster than cod eggs from other stocks do.

3.3.2.3. Stock differences with regards to egg and larval sizes

Although the Faroe Bank females were significantly larger than the Faroe Plateau females, they produced significantly smaller eggs. There were, however, no differences in the average number of batches spawned per female, although these data should be seen in the light of the shorter experiment period for the Faroe Bank cod. However, when considering all batch intervals observed, no difference was found between the stocks, and on average, a female released an egg batch every third to fourth day (Table 3.2).

When considering egg sizes, a significant difference was discovered with batch average egg diameters of 1.40 mm (± 0.04 SD) and 1.30 mm (± 0.06 SD) for the Faroe Plateau and the Faroe Bank females, respectively. This difference equated to an approximately 20 % smaller mean volume for a Faroe Bank egg compared to a Faroe Plateau egg and appeared to translate into significant differences in larval sizes two weeks after hatch, when body lengths of Faroe Plateau and Faroe Bank larvae were 5.23 mm (± 0.40 SD) and 4.81 mm (± 0.41 SD), respectively (Table 3.2)

Table 3.2. Data on female sizes, egg batches, egg sizes and larval sizes, mean (± SD). *P*-values show the results from *t*-tests on stock differences.

	Faroe Plateau	**Faroe Bank**	***P***
Females			
N	10	10	
Body length (cm)	71.0 (±4.8 SD)	87.4 (±6.9 SD)	< 0.0001
Eggs			
Batches per female	7.9 (±2.0 SD)	7.2 (±2.0 SD)	*NS*
Batch interval (days)	3.5 (±1.6 SD)	3.3 (±1.4 SD)	*NS*
Egg diameter (mm)	1.40 (±0.04 SD)[a)]	1.30 (±0.06 SD)[b)]	< 0.0001
Larvae			
N	30	50	
Body length (cm)	5.23 (± 0.40 SD)	4.81 (± 0.41 SD)	< 0.0001

a) Calculation based on 50 egg batches that had data available on egg diameters.

b) Calculation based on 56 egg batches that had data available on egg diameters.

3.3.2.4. Effects of experimental day number and female condition on egg sizes

Overall as well as for individual females, there was a trend of decreasing batch-average egg diameters over time in both stocks (Figure 3.3 and Figure 3.4). The decrease in daily mean egg

diameters from start to end of the experiment was 10 % and 11 % for the Faroe Plateau and Faroe Bank females, respectively. Although the water temperature increased while the experiment lasted, it was relatively constant during spawning of the Faroe Plateau females with an average of 6.3 ºC (± 0.2 SD). When the Faroe Bank females spawned, there was an increasing trend in the temperature with an average of 7.2 ºC (± 0.3 SD). Thus, for the Faroe Plateau cod, the temperature did not appear to be involved in the decreasing trend of egg sizes, but temperature could not be precluded from being involved in the seasonal decrease in egg sizes for the Faroe Bank females. The inter-stock differences in egg sizes, however, did not appear to be related to differences in spawning water temperatures. This could be seen by the fact that there was an increase in temperature between the last Faroe Plateau batches and the first Faroe Bank batches, but despite this, the egg sizes of the Faroe Bank batches were of similar sizes or larger than the Faroe Plateau batches in this comparison (Figure 3.3).

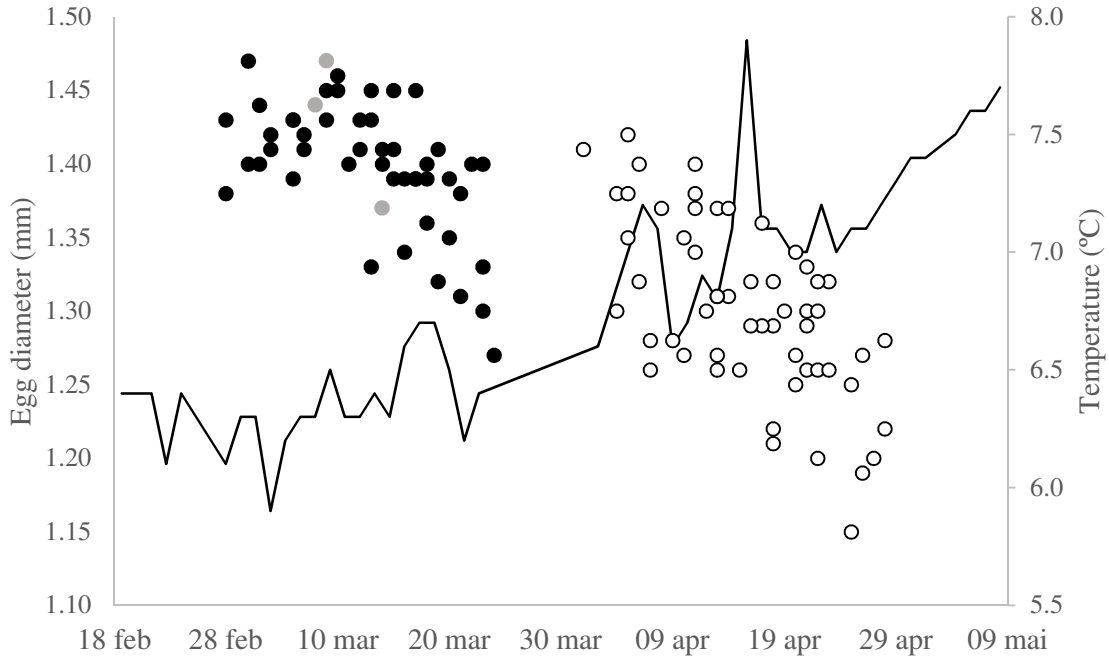

Figure 3.3. Batch-average egg sizes and date of spawning. Black dots show data for the Faroe Plateau females and open circles the Faroe Bank females. The grey dots denote data from a Faroe Plateau female that was caught in 2007. Spawning water temperatures are also shown, as indicated by the line, with temperatures shown on the right axis.

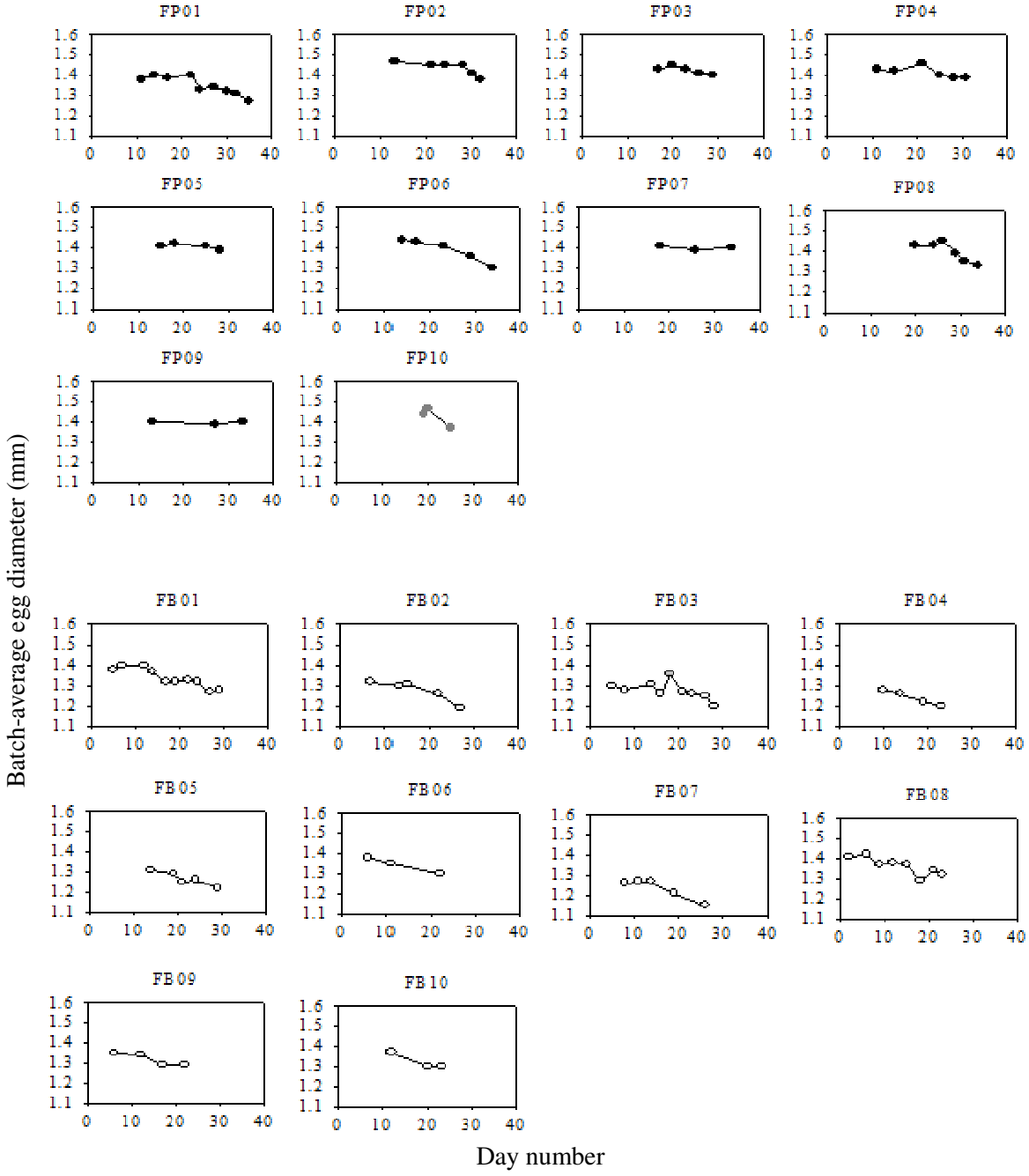

Figure 3.4. Batch-average egg diameters and spawning day number for individual Faroe Plateau (black filled circles) and Faroe Bank (open circles) females. Fish ID in each subfigure. FP10 (shown in grey) was collected in 2007.

In addition to a seasonal effect on egg sizes, condition of the females also appeared to be related with egg sizes, for the Faroe Plateau data as well as for the Faroe Bank data, if an outlier Faroe Bank female with a markedly higher condition factor ($K = 2.19$) was not considered (Figure 3.5).

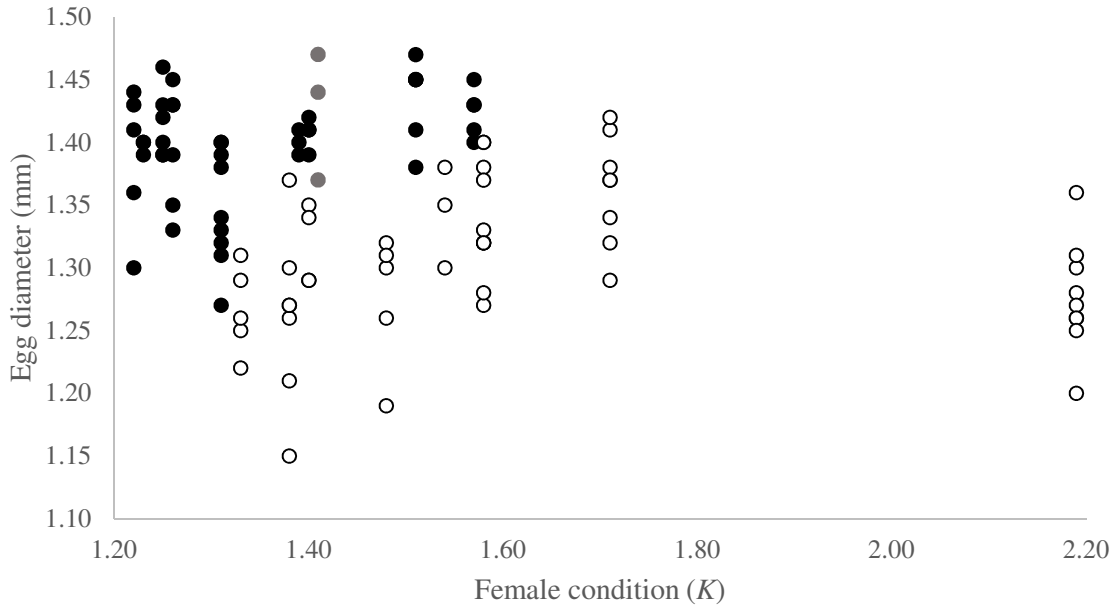

Figure 3.5. Batch-average egg sizes and Fulton´s condition factor (*K*) of the females. Black dots denote Faroe Plateau females and open circles the Faroe Bank females. The grey dots denote data from a Faroe Plateau female that was caught in 2007.

Based on the above observations multiple regression analyses were conducted to investigate the relationship between egg sizes and the potential predictor variables date (experimental day number) and female condition. For both stocks, a significant correlation was found between these two predictor variables and egg size. These models could predict some 40 and 70 % of the variability in egg sizes for the Faroe Plateau and Faroe Bank cod, respectively (Table 3.3).

Table 3.3. Multiple linear regression analyses on the effect of female condition (Fulton´s condition factor, *K*) and experimental day number on batch-average egg sizes (diameters) for the Faroe Plateau and the Faroe Bank data.

	Faroe Plateau (multiple r^2 = 0.430)				**Faroe Bank (Multiple r^2 = 0.730)**			
	Coefficient	SE	*t*	*P*	Coefficient	SE	*t*	*P*
Constant	1.323	0.06	22.067	<0.001	1.087	0.048	22.818	<0.001
Day no.	-0.004	0.001	-0.562	<0.001	-0.005	0.001	-7.293	<0.001
Condition	0.122	0.042	2.906	<0.01	0.204	0.03	6.825	<0.001

3.4. Discussion

This investigation found differences between Faroe Plateau and Faroe Bank cod in a number of reproductive parameters. These include the spawning season with earlier peak spawning observed for the Faroe Plateau cod, the egg development times, which were shorter for the Faroe Bank eggs, egg and larval sizes, which were smaller for the Faroe Bank eggs and the volumes spawned, which were greater for the Faroe Bank females. Furthermore, the egg sizes observed in the current study (average diameters of 1.40 mm and 1.30 mm for the Faroe Plateau and Faroe Bank cod, respectively) were similar to the ones observed in an M.Sc. study of captive spawning cod two decades ago (average diameters of 1.42 mm and 1.31 mm for Faroe Plateau and Faroe Bank cod, respectively) (Magnussen 1993).

3.4.1. Cod spawning in captivity versus cod spawning in the wild

Whereas a part of this work concentrated on cod spawning in their natural surroundings, the majority of the work concentrated on cod spawning in captivity. This means that there is a degree of uncertainty regarding how well the parameters estimated reflect those that are present in wild spawning cod. However, the topic of interest here was the comparison of reproductive parameters between Faroe Plateau and Faroe Bank cod. Such comparisons should be more valid and reflect real stock differences, when conducting laboratory studies, like the captive study conducted here, where the effect of potential different environmental factors are controlled.

3.4.2. Spawning season

Although the differences observed in the captive study should reflect real stock differences, we need to address the fact that, in the captive study, the Faroe Plateau broodstock spawned at temperatures similar to its natural spawning temperatures, whereas the Faroe Bank broodstock spawned at temperatures around 0.5 – 1 ºC lower than its natural spawning temperatures. Kjesbu (1994) found that a temperature decline of this magnitude (1 ºC) during vitellogenesis delayed spawning in cod by up to ten days, leading to the question if this could be the reason for the later peak spawning observed for the captive Faroe Bank cod? Cod initiate vitellogenesis around the time of the autumnal equinox (Kjesbu *et al*. 2010) with photoperiod believed to be the

principal cue regulating spawning time (Bromage *et al*. 2001; Hansen *et al*. 2001; Norberg *et al*. 2004). The fish used in this study were caught relatively early, though, meaning that the Faroe Bank broodstock experienced this decline in temperature well before the onset of vitellogenesis, so the relatively late peak spawning observed for the Faroe Bank cod is unlikely to be due to the temperature difference experienced at the time of capture. Rather, maintaining spawning differential between the two Faroese stocks, while reared under common conditions, strongly suggests a genetic component to this key life history trait.

Despite findings that cod usually spawn earlier in warmer (more southerly) waters (Brander 2005), in nature, Faroe Bank cod, which experience *c*. 1 – 2 ºC warmer water temperatures than the Faroe Plateau cod, are the later of the two stocks to spawn. Thus, temperature differences for the two areas provide further evidence for the hypothesis of genetic factors being involved in the difference in timing of spawning season for the two stocks.

3.4.3. A seasonal trend in egg sizes

The rate of the seasonal decrease in egg sizes was approximately similar for both stocks and was correlated to female condition as well as date (experimental day number). Environmental temperature is a commonly included parameter in investigations of physiological explanations for seasonal trends in egg sizes of marine multiple batch spawners. However, for cod, results are equivocal. Miller *et al*. (1995), investigating cod off the Scotian Shelf, were able to demonstrate a relationship between environmental temperature and egg sizes whereas Kjesbu *et al*. (1996), investigating hatchery-reared Norwegian coastal cod, were not able to find a correlation between temperature and egg sizes. In the present study, a temperature effect could not be precluded as responsible for the variation in the Faroe Bank egg sizes. However, the spawning water temperature was nearly constant during spawning of the Faroe Plateau cod and could not be used to explain the seasonal decrease in those egg sizes. Other studies have proposed that the seasonal decline in egg sizes is due to the depletion of the female´s energy resources, as the spawning season advances, but again, results are mixed. Whereas Chambers and Waiwood (1996) demonstrated a relationship between egg sizes and available energy, indicated by the condition

factor, other studies have not found a similar relationship (Kjesbu *et al*. 1996; Trippel 1998; Ouellet *et al*. 2001). Based on findings in the present study of a relationship between female condition and egg sizes, it may be likely that a seasonal decrease in condition can result in decreasing egg sizes.

3.4.4. Divergent early life history strategies – possibly driven by different environmental conditions

Perhaps the most interesting discovery made here was the finding of diverging stock strategies with respect to the sizes and number of eggs produced. Larger egg sizes were found for the Faroe Plateau cod and smaller egg sizes for the Faroe Bank cod. Furthermore, in agreement with life history theory on the progeny size - progeny number relationship, the field data indicated greater gonadal growth for the Faroe Bank females, as the change in condition between spring and autumn was 25 % higher for Faroe Bank compared to Faroe Plateau females. Per spawning season, with both greater volumes of eggs and smaller egg sizes, each Faroe Bank female can potentially release a considerably higher number of eggs into the water column compared to a Faroe Plateau female. In agreement with smaller eggs spawned by the Faroe Bank females, at 14 days of age, the Faroe Bank larvae were also smaller than the Faroe Plateau larvae of similar ages.

Predation and starvation are widely accepted as the prevailing causes of mortality in young fish larvae (Kamler 2005) and, as such, must be considered potentially potent selective forces in shaping life histories of fishes. This has been demonstrated in Trinidadian guppies (*Poecilia reticulata*) which are live-bearing teleosts with short generation times and, therefore, well suited to study biological changes in response to the environment. Reznick and Endler (1982) found that natural guppy populations that had been exposed to different predation pressures had evolved different life histories. Populations associated with the highest predation pressures were characterised by an earlier age at maturation, devoting relatively more of their body weight to reproduction, reproducing more often and, finally, by producing smaller progeny (Reznick and Endler 1982). This strategy of producing smaller progeny in response to predation has been

termed the "selfish maternal effect", since mothers increase their overall fitness at the cost of survival chances for each individual progeny (Marshall and Uller 2007). Atlantic cod in general must be considered "selfish" in this regard, but based on results here we hypothesise that Faroe Bank cod mothers are even more so compared to their Faroe Plateau colleagues.

To consider the predation risk for early life stages of Atlantic cod one must consider the type of habitat they reside in. Habitat complexity is important for survival and distribution of young Atlantic cod (for example Anderson *et al*. 2007; Ryan *et al*. 2012). Complex seabed patches, composed of e.g. vegetation and rocks, may act as refugia that protect juveniles from predators whereas bare seabed patches expose the juveniles to predators (Ryan *et al*. 2012). Faroe Plateau juveniles inhabitat a seabed that is characterised by a dense vegetation of macroalgae (Steingrund and Gaard 2005; Steingrund *et al*. 2005) and, consequently, rich in refugia, whereas Faroe Bank juveniles inhabitat a bare seabed, consisting mostly of sand and mud with some patches of stone, gravel and rock (Magnussen 2002). Consequently, predation risk is likely to be much higher on early life stages of Faroe Bank cod compared to Faroe Plateau cod.

Based on early life-history theory of marine teleosts, the following hypothesis can be proposed: that the different egg sizes observed for the Faroe Plateau and the Faroe Bank cod have evolved in response to the divergent conditions experienced by settled juveniles in the two areas. This is especially based on the trade-off between number and size of eggs (Elgar 1990 and Kamler 2005, but also on the proposed positive relationship between egg sizes and progeny sizes and viability of the resulting larvae, reviewed in Kamler (2005) and documented for Atlantic cod by Pepin *et al*. (1997). Faroe Plateau juveniles reside in a more sheltered area, compared to their relatives on the Faroe Bank, making the odds of survival for each individual larvae relatively high, so that there is no selective advantage in producing a higher number of small larvae, but rather in producing larger larvae that are better equipped to start an independent life. The strategy of producing larger progeny comes at the cost of the number of progeny produced (for example Elgar 1990). On the Faroe Bank the situation is roughly speaking the reverse, post-settled juveniles are left exposed to predators and stochastic environmental variations, so that here it

would confer a selective advantage to produce a large number of smaller eggs, thus increasing the probability of having at least a few eggs surviving to the adult stage. Moreover, since the Faroe Bank is a highly productive ecosystem (Gaard and Mortensen 1993), competition for food among juveniles on the Faroe Bank may be less fierce than the one their colleagues on the Faroe Plateau experience, making egg, and thus larvae, sizes less important than the number of eggs produced.

Chapter 4. Differential survival among cod families from fertilisation through to metamorphosis

4.1. Introduction

Although the production of cultured cod has not reached the levels anticipated a few years ago, Atlantic cod still has the potential to become an important aquaculture species in the future (Rosendal *et al*. 2013). Pair mating is possible on an experimental scale and is used in breeding programmes, but it is much more expensive (Myhr *et al.* 2012) and commercial scale breeding of cod generally involves mass spawnings in large tanks (Hansen and Puvanendran 2010). This practice has its drawbacks if future broodstock are to be selected from this type of production since the cod mating system is very complex and the effective breeding population per spawning bout may be considerable smaller than the total breeding population (Hutchings *et al.* 1999; Nordeide and Folstad 2000; Herlin *et al.* 2008). Therefore, the risk of accumulating inbreeding over generations is high. Adding to this risk is the fact that intensive production of Atlantic cod still involves high and unpredictable mortality rates in the early life stages (Gunnarsli *et al.* 2009), particularly in the stage from first-feeding through to metamorphosis, where there is a lack of knowledge on optimal rearing protocols (Kjesbu *et al.* 2006).

The cod mating system can be classified as a lekking system (Hutchings *et al.* 1999; Nordeide and Folstad 2000), prerequisites of which include male-male competition and non-random mating by females (Hutchings *et al.* 1999). In addition to sexual traits, characteristics typically tested in this regard include female and male body size and condition (Kjesbu 1989; Kjesbu *et al.* 1991; Hutchings *et al.* 1999; Bekkevold *et al*. 2002; Trippel 2003) as well as male-female size differences (Rakitin *et al.* 2001; Bekkevold *et al.* 2002; Armitage *et al.* 2007).

Irrespective of the precise mechanisms involved, the outcome appears to be a skewed parental contribution to the progeny generation, as seen in both commercial (Wesmajervi *et al.* 2006; Herlin *et al.* 2008) and experimental setups (for example Hutchings *et al.* 1999; Rudolfsen *et al.* 2005). As individual female cod can release up to around 20 egg batches over one to two months (Kjesbu *et al.* 1996), only a proportion of the females can be expected to contribute to each

spawning bout. Accordingly, Herlin *et al.* (2008) investigated paternity in 300 fry (83 dph) originating from eggs collected on a single day from a commercial breeding tank containing 99 parental cod and found that 81 % of the progeny were assigned to a single pair of parents and a total of 26 parental cod were found to have contributed to the assigned progeny. This family skew in reproduction is also seen on a seasonal basis. Hansen *et al.* (2004) analysed egg batches from 15 female and 15 male cod collected over an entire spawning season and found that five females and five males were responsible for 77 and 54 % of the sampled eggs, respectively. Lastly, Bekkevold *et al.* (2002) found a shift in male reproductive dominance, likely explained by sperm-depletion in high-ranking males, thereby allowing second-ranking males to take over. However, the seasonal average reproductive success for males still varied from close to zero progeny sired by some males to nearly 90 % progeny sired by one of the males.

There is a need to understand the factors underlying the variability in survival among egg batches, in order to optimise protocols for cod hatcheries (Hansen and Puvanendran 2010). Egg morphology, biochemical composition and early cleavage patterns amongst others may be useful indicators of egg quality, as seen from a number of studies on cod and other species (Hansen and Puvanendran 2010). How predictive these are of viability in later developmental stages is not always clear though, since most studies have focused on hatching success or survival rates in the early larval stages (Hansen and Puvanendran 2010).

The present study was designed to investigate if differential family survival rates in the egg stage could be correlated to survival rates in later fry stages and to investigate potential predictors of survival rates in these two stages. The experimental setup involved mating wild-caught Faroe Bank cod in tanks with one female and one or two males in each tank. Family survival rates were estimated for two separate periods in early development, the egg stage, here defined as the period from fertilisation to a few days prior to hatch, and the fry stage, here defined as the succeeding period through to metamorphosis. Relationships between survival rates in the two stages and between survival rates and egg sizes, female condition and sizes of metamorphosed fry were explored.

4.2. Materials and methods

Rearing protocols are described in Section 2.1.

4.2.1. Experimental design

The experimental setup involved mating wild-caught Faroe Bank broodfish in spawning tanks. A total of eight females were each mated with one or two males. Mating design is shown in Table 4.1. Their offspring (equal numbers of similar ages from each female; see group FB 3 in Table 2.2) were reared under identical conditions in three replicate tanks until past metamorphosis. On the Faroe Bank, 50 % of the cod mature at an age of three years and body length of *c*. 70 cm (Steingrund *et al*. 2005; Magnussen 2007, 2011) and at a body length of 85 cm, 100 % of the cod have matured (Steingrund *et al*. 2005). The average body length of the females used in this study was 88 cm (± 8 SD), thus all females were considered second-time spawners or more.

4.2.1.1. The egg stage period

At 17 and 18 April, at the respective days of spawning, 2 dl eggs from each of the eight spawning tanks were transferred to incubation. Analyses in this work required that the eggs were introduced into first-feeding tanks prior to hatch, so that egg diameter could be used to calculate the exact number of eggs per volume. Observations at the Marine Research Centre had shown that Faroe Bank eggs could hatch after 14 days at 5 ºC, hence, on 29 April, 45,400 eggs, comprising a similar number (5,675) from each of the eight females, were introduced into each of three first-feeding tanks. At this point, the percentage of eggs surviving incubation was estimated. The eggs hatched on 2 – 3 May. Thus, the egg stage period lasted from fertilisation to 3 – 4 days prior to hatch.

4.2.1.2. The fry stage period

The larvae were fed rotifers (*Brachionus* spp.) initially and weaning from rotifers onto a microparticulate diet started on 9 June. On 4 – 5 August, the cod fry were moved to three larger tanks. Three spawning tanks (Section 2.1.2) were used for this purpose. Sampling was conducted on 19 – 21 August. At this stage, all individuals had undergone metamorphosis. From each of the three replicate tanks, 300 fry were randomly sampled, killed by anaesthetisation with metacaine

(MS-222), measured (total length), preserved in 96 % ethanol and stored at 4 ºC. Thus, the fry stage period lasted for 114 – 116 days.

Table 4.1. Mating design and sizes of pre-spawning parents. F = female; M = male; *K* = Fulton´s condition factor.

ID	Length (cm)	Weight (g)	*K*
F1	87.5*	10300	1.54
M1a	84.0*	9300	1.57
M1b	76.0	5500	1.25
F2	80.5	8900	1.71
M2a	82.0	8600	1.56
M2b	83.0	8500	1.49
F3	77.5	10200	2.19
M3a	79.0	6500	1.32
M3b	81.0	7000	1.32
F4	87.0	7500	1.14
M4a	85.0	8900	1.45
M4b	85.0	8990	1.46
F5	91.5	10600	1.38
M5a	93.0	14400	1.79
M5b	94.0	12500	1.5
F6	97.0	12710	1.39
M6a	94.5	12900	1.53
M6b	95.0	11000	1.28
F7	81.5	7200	1.33
M7	83.0	8900	1.56
F8	100.5	16000	1.58
M8	94.5	11370	1.35
Female averages	88 (± 8 SD)	10426 (± 2871 SD)	1.53 (± 0.32)

*These data were obtained later than the rest, on the day that the fish were introduced into the spawning tanks.

4.2.2. Genotyping

DNA was extracted using the Real Pure Genomic DNA extraction kit (see Section 2.3.1.2 for extraction protocol). Samples were scored at three microsatellite loci *Gmo*8, *Gmo*19 and *Tch*13 (Table 2.6) as described in Section 2.3.3. Full genetic profiles were obtained for all 22 parental cod and 898 of the 900 progeny sampled. One progeny was only genotyped at two of the three loci and one individual was not scored at any loci, presumably due to poor template quality.

4.2.3. Parentage analyses

Progeny were assigned to their respective families by the exclusion-based parental assignment programme FAP (Taggart 2007). By applying one allele mismatch tolerance, all except one of the 899 fry were unambiguously assigned to a single family. The one remaining individual was assigned to a single family by applying two allele mismatch tolerances.

4.2.4. Statistical analyses

In the first part, concerning family survival rates during the egg stage, a multiple linear regression analysis was used to explore the relationships between these and two independent variables, egg diameter and female condition (Fulton´s condition factor, *K*). In the next part, concerning family survival rates in the fry stage, χ^2 goodness-of-fit tests were used to compare observed versus expected numbers of progeny per female, calculated for each of the three replicate tanks. Females were grouped into two groups, a high and a low fry survival group, and a two-way ANOVA was used to investigate the effect of groups and replicate tanks on survival. In the last part, concerning fry sizes, linear regression analyses were used to investigate the relationship between egg diameter and fry sizes, separately for each replicate tank. Finally, fry sizes were compared across the three replicate tanks using a one-way ANOVA. Statistical analyses were conducted in Systat (version 11.0, Systat Software, Inc.; www.systat.com).

4.3. Results

4.3.1. Differential family survival in the egg stage

The mean survival rate in this stage was 69 % (± 20 % SD) and scatter plots suggested a positive association with both egg sizes and female condition factor (Figure 4.1). As shown in Table 4.2, however, multiple linear regression analysis showed that only egg diameter had a significant effect on egg survival rates, accounting for 90 % of the variation in egg survival rates.

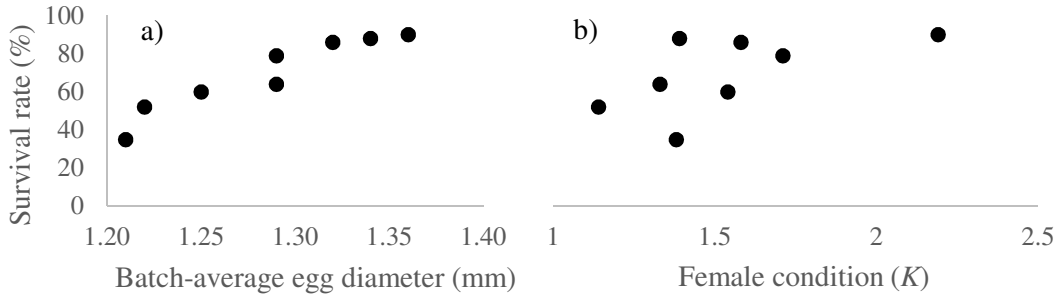

Figure 4.1. Survival rates at the end of the egg stage and the two potential predictor variables a) egg diameter and b) female condition (Fulton´s condition factor, *K*).

Table 4.2. Multiple linear regression analysis of the effect of egg diameter and female condition on survival in the egg stage, multiple $r^2 = 0.904$.

	Coefficient	**Standard Error**	***t***	***P***
Constant	-379.238	75.030	-5.055	<0.01
Egg diameter	351.623	66.388	5.297	<0.01
Female condition	-2.184	11.460	-0.191	*NS*

4.3.2. Differential family survival in the fry stage

In all three replicate tanks, there was a significant difference between observed and expected numbers of progeny originating from the various females (Tank 1: $\chi^2 = 456.9$, $df = 7$, $P < 0.0001$; Tank 2: $\chi^2 = 452.4$, $df = 7$, $P < 0.0001$; Tank 3: $\chi^2 = 296.4$, $df = 7$, $P < 0.0001$). Moreover, the females fell into two very different groups regarding fry survival, with three females (F2, F3 and F6) predominating in all replicate tanks, parenting a total of 91.1 % of the progeny pooled over

replicates. The residual 8.9 % belonged to the five remaining females with one female (F4) only parenting one out of the total of 899 progeny genotyped (Table 4.3).

The females were grouped into two groups, the high survival group (F2, F3 and F6), with each female contributing 30 % (± 7 % SD) of the fry, and the low survival group (F1, F4, F5, F7 and F8), with each of the females contributing 2 % (± 2 % SD) of the fry. In a two-way ANOVA, a significant difference was found between the groups regarding survival, but no difference was detected among the replicate tanks and the interaction term was nonsignificant (Table 4.4).

Table 4.3. Maternal contribution to metamorphosed progeny in the three replicate tanks.

Female	**Replicate**			**Total per female**		
ID	**1**	**2**	**3**	**No.**	**%**	**Cumulative (%)**
F2	126	121	91	338	37.6	37.6
F6	91	92	80	263	29.3	66.9
F3	64	70	84	218	24.2	91.1
F8	10	7	14	31	3.4	94.5
F5	2	4	18	24	2.7	97.2
F1	5	3	10	18	2	99.2
F7	1	2	3	6	0.7	99.9
F4	1	0	0	1	0.1	100
All	300	299	300	899	100	100

Table 4.4. Two-way ANOVA for fry survival across the high and the low survival female groups (see text) and the three replicate tanks.

Source of Variation	***df***	**F**	***P***
Groups	1	205.664	< 0.001
Tanks	2	0.038	*NS*
Group × Tank	2	0.548	*NS*
Error	18		

There was no indication of any significant relationship between fry survival and survival rate in the egg stage, egg diameter or female condition factor.

4.3.3. Sizes of metamorphosed progeny

Although scatter plots indicated positive relationships between egg diameters and fry body lengths (Figure 4.2), regression analysis did not reveal any significant relationship between these variables (Tank 1: $r^2 = 0.683$, $P = NS$; Tank 2: $r^2 = 0.480$, $P = NS$; Tank 3: $r^2 = 0.708$, $P = NS$).

There was wide variability in fry sizes (Figure 4.2) with mean body lengths of 5.08 (± 0.74 SD), 4.75 (± 0.82 SD) and 5.57 cm (± 0.82 SD) for replicate tanks 1, 2 and 3. A one-way ANOVA showed that sizes of fry varied significantly among the three replicate tanks (Table 4.5), thus reflecting a tank effect concerning fry body lengths.

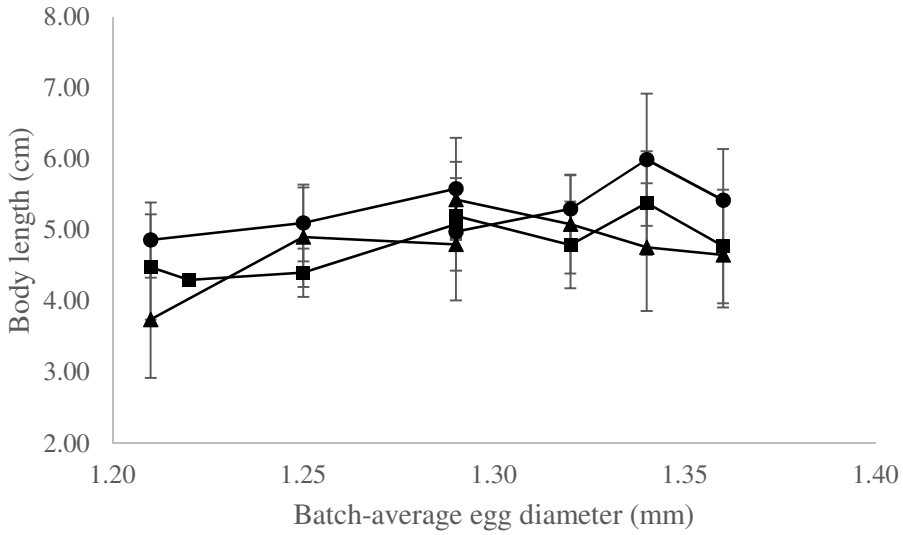

Figure 4.2. Egg diameters and standard body lengths (mean ± SD) of fry from the various females at the end of the fry stage in replicate tanks 1 (squares), 2 (triangles) and 3 (circles).

Table 4.5. One-way ANOVA for fry sizes across the three replicate tanks.

Source of Variation	***df***	**F**	***P***
Tanks	2	80.311	< 0.0001
Error	896		

4.4. Discussion

The results presented here show a large interfamily skew in survival of cod eggs and cod fry, but family survival rates in the egg stage could not simply be used to predict survival rates in the fry

stage. Although there was a strong correlation between egg diameters and survival rates in the egg stage, egg diameter was not predictive of fry survival rates nor of fry body lengths. There was no significant tank effect concerning survival rates in the fry stage, but a significant tank effect was present concerning body sizes of fry.

As pointed out by Hansen and Puvanendran (2010), there are many possible egg quality markers, but it is important that an evaluation of these markers is easy to adapt in daily hatchery work. Seen from this perspective, it is encouraging that 90 % of the variation in egg survival rates could be predicted from egg diameter, an estimate that is easy to obtain. This observation is somewhat intriguing, as other studies, on cod and other species, did not find any association between egg size and hatching success (Ouellet *et al*. 2001). Trippel (1998), however, found a significant effect of egg size on hatching success, but in that study, mean egg size and hatching success were compared among first- and second-time spawners, and all females in the present study were considered to be at least second-time spawners.

Having replicates in the fry stage was a significant strength of the study, since unidentified mortalities in the period from first-feeding to metamorphosis appear to be the rule rather than the exception (Kjesbu *et al.* 2006). Survival rates in this period were comparable among all three replicate tanks, with the same three females dominating. Thus, concerning fry survival, there were no adverse tank effects and the results truly demonstrated a family effect.

Many studies have focused on differential survival among cod families (for example Hutchings *et al.* 1999; Hansen *et al.* 2004; Herlin *et al.* 2008), but a further strength of the present study lies in the fact that survival rates were evaluated for two different periods in early development. Moreover, mixing progeny prior to hatch facilitated the process of transferring a precise number of individuals from each female to the first-feeding tanks and, ultimately, allowed for calculation of precise expected numbers per female.

A key observation was the lack of a relationship between survival rates in the two stages, hence, the practice of evaluating egg quality parameters that may predict hatching success, should be

supplemented with further genetic monitoring of later progeny stages to avoid erosion of genetic variation, which may obstruct any long-term genetic gain in cod breeding.

The fact that female condition factor was not significantly related to survival rate in the egg stage was consistent with a study by Ouellet *et al.* (2001) where no clear association was found between female pre-spawning condition and survival to first hatch. Furthermore, these observations are in agreement with the biology of teleost fishes in general, where female age, but not female size, has been identified as a predictive factor of hatching success. However, in teleost fishes, female size is expected to affect the further viability of the offspring, through its effect on egg size (Kamler 2005). Although results in cod are somewhat ambiguous (Ouellet *et al*. 2001), Kjesbu *et al.* (1991) demonstrated a relationship between female pre-spawning condition and embryo viability. Marteinsdóttir and Steinarsson (1998), investigating eggs from wild cod that were fertilised at sea, found a significant effect of female condition on egg size as well as a significant positive relationship between egg and larval sizes, early larval feeding and growth rates, but no significant association between egg sizes and growth rates at day 20. Similarly, for female condition, there was a significant positive relationship between female condition and early larval feeding, but not with larval growth rates at day 15 and day 20. Hence, the female effect appears to be most apparent in the early life stages and this may explain why the present study found no significant relationship between female condition and sizes and survival rates of metamorphosed cod fry. Furthermore, whereas maternal effects decrease, environmental effects may increase, as the progeny grow, as indicated by a significant tank effect found here with regards to body sizes of fry.

A further explanation for the lack of any significant relationships between female condition factors and larval viability indicators in the present study may lie in the fact that all females were second-time spawners and they were all in good condition, possibly making such effects more difficult to detect.

In conclusion, egg size was found to be a useful indicator of survival rates in the egg stage, but egg survival rates should not be uncritically applied as a proxy of viability in later developmental stages. There was a strong family component concerning survival rates of metamorphosed fry, highlighting the importance of employing some sort of genetic monitoring of later developmental stages of cod fry.

Chapter 5. Genetic differentiation between Faroe Plateau and Faroe Bank cod as investigated by microsatellite markers

5.1. Introduction

Faroe Plateau and Faroe Bank cod differ with respect to a wide array of biochemical, morphometric and growth characteristics (Section 1.6.2) that could reflect the different environments that the fish inhabit and / or genetic differences between fish from the two areas. Magnussen (2007) investigated the potential impact of the Faroe Bank environment on growth performance for 14 common fish species sampled there with published data on growth for these species from a total of 36 different areas. When the fish reached sexual maturity, resident species on the Faroe Bank were on average 36 % larger than fish from other regions, whereas this difference was 6 % for the migratory species that only inhabited the Faroe Bank part of their life (Magnussen 2007). These results clearly demonstrate the positive impact of the Faroe Bank environment on fish growth, probably due to high abundance of food and near optimal temperatures for growth, although the author acknowledged that part of the growth differences might still reflect genetic makeup.

Data from a century of tagging studies indeed suggest that genetic divergence is likely to exist between cod from the two areas, as there is very limited movement of adult fish between the Faroe Plateau and the Faroe Bank (Jones 1966; Fjallstein and Magnussen 1996; Magnussen 1996; Joensen *et al*. 2005; ICES 2006). Moreover, both the Faroe Bank and the Faroe Plateau are characterised by having anticyclonical currents that act to retain eggs and larvae within the respective areas (Hansen *et al*. 1986; Hansen 1992; Hansen 2000; Steingrund *et al*. 2005; Larsen *et al*. 2008b). Tagging studies have also revealed the presence of three main spawning sites within the Faroe Plateau, located to the north and west of the islands (Figure 1.3), whereas Faroe Bank cod are believed to be more sedentary. On the Faroe Plateau, although adult cod are believed to show a certain degree of spawning site fidelity (Steingrund *et al*. 2005; Joensen *et al*. 2005), genetic studies undertaken to date, using mtDNA and microsatellites, have failed to detect significant genetic substructure (Sigurgíslason and Árnason 2003; Pampoulie *et al*. 2008a). Thus,

the clockwise circulation system apparently results in an absence of structure with eggs and larvae distributed in a random fashion at the various nursery areas close to land (Pampoulie *et al*. 2008a). The fish caught in tagging experiments have typically been adult fish while there is less knowledge of the movement of juvenile cod (Magnussen 1996). A weakness of tagging studies is that they do not reflect whether immigrant genes are successfully incorporated into the gene pool (Lowe *et al*. 2004a). Genetic studies, on the other hand, can potentially resolve such questions, although gene flow estimates should be interpreted with care, as they might reflect contemporary and / or historical gene flow (Lowe *et al*. 2004b).

Regarding the potential genetic differentiation between Faroe Plateau and Faroe Bank cod, Jamieson and Jones (1967) found evidence of genetic divergence at a transferrin locus, as did Jamieson and Birley (1989), using a hemoglobin marker. However, these results were not replicated by other studies using hemoglobin (Magnussen 1996) and a variety of allozyme markers (Jamieson and Thompson 1972; Magnussen 1996). There is a variety of possible explanations for the discrepancies. For example, a number of studies of cod in Icelandic waters also revealed inconsistencies in hemoglobin allele frequencies, which were later found to be correlated to time of sampling, suggesting the possible presence of moving stock components (Imsland and Jónsdóttir 2003). Furthermore, the denominator for these marker systems is that they are, by definition, linked to genes and potentially under the influence of natural selection, meaning that variation does not necessarily reflect the breeding structure of the stocks, but may for example reflect adaptation to various environmental conditions (Kirk and Freeland 2011). In a more recent study, however, fatty acid profiles in heart tissue of captive bred Faroe Plateau and Faroe Bank cod held in a common environment and with the same diet for > 3.5 years were investigated, and significant stock differences were detected (Joensen *et al*. 2000).

Following the successful identification of two stocks of cod in Faroese waters by analysis of fatty acid profiles in heart tissue, the same method was used to investigate the stock structure of the deep-water redfish, *Sebastes mentella*. Four distinct North-Atlantic stocks were identified. One stock comprised samples from the Faroe Plateau and Norwegian waters, a second stock samples

from the depths South-West of the Faroe Bank, the Wyville Thompson Ridge, South Iceland and the deep waters of the Irminger Sea, a third stock comprised the East Iceland sample and the fourth stock was made up of the more distant oceanic Irminger Sea sample (Joensen and Grahl-Nielsen 2004). The authors (Joensen and Grahl-Nielsen 2004) acknowledged that the method needs further evaluation with respect to the impact of various abiotic and biotic factors, though the identification of genetic divergence between a deep-sea and an oceanic stock component of deep-water redfish in the Irminger Sea was subsequently replicated by allozyme analysis (Daníelsdóttir *et al*. 2008).

In the last two decades, microsatellite markers, which are highly informative (multi-allelic) and considered, mainly, to be selectively neutral in nature, have been the main marker of choice for resolving genetic structuring of natural populations (Schlötterer 2004; Selkoe and Toonen 2006). Knutsen *et al*. (2009) used microsatellites in an investigation of population structure in the deep-sea fish tusk and failed to detect substructuring among tusk from East-Greenland, Iceland, Faroe Islands (sample from the Faroe Plateau) and Norway. Hence, both deep-water redfish as well as tusk appear genetically homogenous across Faroe Plateau and nearby waters (Joensen and Grahl-Nielsen 2004; Knutsen *et al*. 2009).

Microsatellites have also been widely used to study cod populations and have identified significant structure across its distribution range (O'Leary *et al*. 2007) as well as between closely located populations (Nielsen *et al*. 2003). Nielsen *et al*. (2009a) studied population genetic structure of cod samples from the waters west of Scotland and from the Northern North Sea using ten microsatellite loci and included samples from adjacent areas for comparison, including Faroe Plateau and Faroe Bank specimens (69 and 50 individuals, respectively). This study revealed significant genetic differentiation between Faroe Plateau and Faroe Bank cod samples with a pairwise F_{ST} value of 0.0085 ($P < 0.00055$). This estimate was a bit lower, however ($F_{ST} = 0.0077$; $P = 0.00385$), and genetic differentiation was nonsignificant, after adjusting for multiple (78) comparisons, when locus *Gmo*132, believed to be under the influence of hitch-hiking selection (Nielsen *et al*. 2009a), was excluded from the computation. Apart from identifying potential

population structuring, this study also demonstrated the importance of a careful interpretation of microsatellite data, as the assumption of marker neutrality might sometimes be breached. Additionally, loci may behave differently for different samples. In contrast to overall results, pairwise levels of genetic differentiation within the north-eastern Atlantic region were more robust to the inclusion or exclusion of *Gmo*132 (Nielsen *et al*. 2009a).

5.1.1. Objectives

In order to obtain a more detailed picture of the genetic variation present within and between Faroe Plateau and Faroe Bank cod samples, this study applied microsatellite markers to investigate large sample sizes from both areas, collected over three consecutive years to account for potential temporal variation. The main question was that of the potential genetic divergence between the two areas. However, as adult cod on the Faroe Plateau are believed to show a certain degree of spawning site fidelity (Steingrund *et al*. 2005; Joensen *et al*. 2005), this study also tested for a potential divergence between cod from two of the main spawning areas on the Faroe Plateau. To put the data into context, samples from surrounding cod populations were included in the study. Moreover, in order to facilitate comparisons with other studies, microsatellite markers were selected that have been widely used in other studies of cod populations (Section 2.3.3).

In the last twenty years, both Faroese stocks have reached historically low levels on two occasions, and currently there is almost a total moratorium on commercial fishing on the Faroe Bank (Section 1.6.2.2). Large reductions in population sizes may result in a negative impact on the genetic variation present, yet another factor to consider when formulating management strategies. Therefore, analyses were also included that could assess the amount and the temporal stability of the genetic variation present.

On the Faroe Plateau cod migrate to deeper waters (150 m and deeper) as they grow, and cod caught deeper than 150 m are generally in a better condition than cod caught in more shallow waters (Steingrund and Ofstad 2010). However, as opposed to what has been found for some other cod stocks, there is a spatial mismatch between Faroe Plateau cod and their prey. Although

cod on the Faroe Plateau mainly feed on Norway pout and blue whiting, they are generally distributed shallower than the bulk of these prey organisms. The cause of this is largely unknown, although one explanation may be that the cod need time in order to adapt to prey and predator organisms as well as depth and temperatures before moving into deeper waters. Furthermore, the vertical distribution of Faroe Plateau cod has been found to be density-dependent in the sense that year-classes one year older than strong year classes (high biomass) are found in relatively deep waters and vice versa for year classes that are one year older than less strong year classes (Steingrund 2009; Steingrund and Ofstad 2010). Genotype-dependent vertical distribution has previously been suggested for cod (Case *et al*. 2005), hence, to the extent that the data allowed for it, the microsatellite data were also analysed in relation to depth.

5.2. Materials and methods

5.2.1. Sampling of biological material for population genetic analysis

Sampling of biological material for population genetic analysis is described in Section 2.2. The Faroese samples (N = 1069) were collected over three consecutive years, 2008 – 2010. These included the Faroe Bank (FB) area and two Faroe Plateau areas, FPNE (North and East of the Faroe Islands; cod inhabiting this area are believed to spawn at the major spawning ground to the north) and FPW (an area West of the islands; includes a smaller spawning ground). Biological details for these samples (seven in total) is provided in Table 5.1. While FPNE and FB areas were sampled for three consecutive years, only a single sample was obtained for the FPW area (largely due to bad weather, causing changes in the sampling scheme).

In addition to the Faroese samples, Norwegian coastal cod from the Borgundfjord (BGF; N = 48) and an area North-West of Bomlø (NWB; N = 48), collected in 2005 and 2006, were included in the study (though only a subset of these could be analysed, as a number of the samples were damaged during transport). Hence, in total, nine samples were available for the study.

Table 5.1. Sample sizes (*N*), average size (total body length and weight), condition (Fulton´s condition factor, *K*), age (± SD), sex ratio (% F: percentage females) and proportion of actively spawning individuals (% S; maturity stage 6) for the Faroe Plateau North-East (FPNE), Faroe Plateau West (FPW) and Faroe Bank (FB) samples.

Area	Year	*N*	Body length (cm)	Body Weight (g)	Fulton *K*	Age (years)	% F	% S
FPNE	2008	130	58.5 (±12.8)	2451 (±1915)	1.052 (±0.105)	4.2 (±1.4)	47	48
	2009	170	58.7 (±14.5)	2571 (±2261)	1.055 (±0.095)	4.2 (±1.5)	52	48
	2010	288	59.4 (±13.2)	2537 (±1835)	1.037 (±0.086)	3.9 (±1.6)	43	47
FPW	2009	116	64.2 (±9.6)	3094 (±1560)	1.084 (±0.095)	5.0 (±1.3)	63	70
FB	2008	138	81.4 (±8.4)	5992 (±1798)	1.082 (±0.105)	4.4 (±1.4)	39	72
	2009	122	84.0 (±11.7)	7144 (±2900)	1.137 (±0.090)	5.0 (±1.5)	45	44
	2010	105	89.4 (±7.2)	8186 (±2192)	1.126 (±0.094)	5.9 (±1.4)	27	60

5.2.2. DNA extraction

DNA from the Faroese samples, collected in 2008, was extracted using the Maxwell® Tissue DNA purification kit on a Maxwell® 16 automated platform (Promega corporation, Ramcon, Birkeroed, Denmark), as described in Section 2.3.1.1, while the remaining samples were extracted with the Real Pure Genomic DNA extraction kit (Durviz, Valencia, Spain), as described in Section 2.3.1.2.

5.2.3. Microsatellite genotyping

The samples were genotyped at ten microsatellite loci (*Gmo*8, *Gmo*19, *Tch13*, *Gmo*35, *Gmo*37, *Tch11*, *Gmo*2, *Gmo*3, *Gmo*34 and *Gmo*132) as described in Section 2.3.3. The exact same panel of markers has been used by others to study cod population structure (Eiriksson and Árnason 2013). Unfortunately, the plate containing the BGF and NWB DNA was damaged in transport, resulting in too little DNA to genotype all individuals at all loci. Data for some of the missing genotypes were obtained from the Norwegian Institute of Marine Research (allele sizes were

standardised to match those analysed at the Faroese Aquaculture Research Station laboratory). However, this was not possible for four of the loci (*Gmo*8, *Gmo*19, *Gmo*37 and *Tch13*), resulting in many missing BGF and NWB genotypes at these loci. In contrast, for the Faroese cod, full genotypic profiles were available for the vast majority of individuals.

5.2.4. Statistical analyses

5.2.4.1. Genetic diversity and assumption testing

First, basic diversity statistics were calculated; number of alleles (*A*), allelic richness (*AR*; calculated from 1000 resamples based on the smallest sample size), observed (H_O) and expected (H_E) with the diveRsity online application (Keenan *et al*. 2013), and inbreeding coefficients (F_{IS}) with GENEPOP version 4.2 (Raymond and Rousset 1995; Rousset 2008; available online at www.genepop.curtin.edu.au). In addition, PIC (polymorphism information content) was calculated with Microsatellite Toolkit for Excel (Park 2001). This software also has some functions to find potential problems with datasets, like the function of finding matching samples. Three such matches were found in this dataset; by tracing the samples back to laboratory analysis, two of the matches likely originated from mixing or contamination when a few samples were re-extracted and one from mixing or contamination at a PCR-rerun. These individuals were excluded from all analyses.

Deviations from Hardy-Weinberg equilibrium (HWE) as well as linkage disequilibrium were tested with GENEPOP, using the probability test option, with *P*-values estimated from a Markov chain (dememorisation 1000 and 10000 batches of 1000 iterations). When non-conformance to HWE was observed, the data were checked for null alleles, large allele dropout and stuttering in the software Micro-Checker version 2.2.3 (van Oosterhout *et al*. 2004). Moreover, since non-neutral variation has been suggested for a few of the markers (Karlsson and Mørk 2005; Nielsen *et al*. 2006; Skarstein *et al*. 2007; Nielsen *et al*. 2009a), the data was tested for evidence of selection, using the LOSITAN software (Beaumont and Nichols 1996; Antao *et al*. 2008). F_{ST} values as a function of heterozygosity were compared to their neutral distribution generated by coalescent simulations under an island-migration model at mutation-drift equilibrium. For the

analysis, 50,000 simulations were performed on each of the nine samples under both the infinite alleles and the stepwise mutation model with confidence intervals of 99 and 95 %.

5.2.4.2. Genomic location of microsatellites

The microsatellite loci used in this study (Brooker *et al*. 1994; Miller *et al*. 2000; O'Reilly *et al*. 2000) were developed well before the sequencing of the cod genome (Star *et al*. 2011). However, with the cod sequence now available, it is possible to determine the genomic location of the microsatellite loci. This may prove useful in interpreting, for example, potential deviations from neutrality assumptions. Hence, the microsatellite loci were annotated by using the microsatellite sequences available in the NCBI nucleotide database (www.ncbi.nlm.gov) as queries against the cod genome (gadMor1, version 75.1) in Ensembl (www.ensembl.org; Flicek *et al*. 2014). The BLAST-like alignment tool (BLAT) (Kent 2002) was used for the purpose. Two of the markers were not available in NCBI (*Gmo*2 and *Gmo*132). In order to locate these markers by Ensembl, 20 unknown nucleotides ("N´s") flanked by the respective primer sequences, were used as queries in BLASTN searches. The potential biological function of the sequences was inferred from Gene Ontology (GO) annotations (Ashburner *et al*. 2000) in Ensembl, all inferred by electronic annotation.

5.2.4.3. Population structure

Two of the geographic regions (FPNE and FB) had been sampled for three consecutive years; hence, levels of genetic differentiation were first assessed between temporal samples from these geographic regions. Samples were pooled, if no significant differentiation was detected. Then, genetic differentiation among areas was assessed. For microsatellite data, D_{Jost} should provide a better alternative for estimation of population differentiation compared to F_{ST} (Section 2.4.3.1); hence, the degree of genetic differentiation was quantified with D_{Jost} and significance was assessed by 1000 bootstrap replicates to calculate 95 % confidence intervals (CIs). If the lower CI included zero, the null hypothesis of no differentiation was accepted. A nearly unbiased estimator of D_{Jost} ($D_{\text{Jost\_EST}}$) was used for the pairwise comparisons. The diveRsity online application was used for these calculations (Keenan *et al*. 2013). The traditional F_{ST} estimate

(Weir and Cockerham 1984) was also calculated for comparative purposes, also in diveRsity online and with associated 95 % CIs. In addition, the significance of genic (allelic) differentiation was tested by a pseudo-exact probability test, dememorisation 1000 and 1000 batches of 1000 iterations, calculated in GENEPOP.

A Bayesian clustering algorithm in the software STRUCTURE version 2.3.4 (Pritchard *et al*. 2000; Falush *et al*. 2007) was applied in order to determine the number of major distinct genetic clusters (K) from the data. An admixture model with correlated allele frequencies was used and runs consisted of a burn-in length of 250,000 iterations followed by a Monte Carlo Markov Chain of 600,000 iterations. Five replicate runs were conducted for values of K ranging from 1 to 5. To infer the most likely number of clusters, the mean estimated natural logarithm of the posterior probability (P) of the data (D), $\ln P(D)$, was considered as well as an ad hoc criterion ΔK (Evanno *et al*. 2005), which was calculated by using Structure Harvester (Earl and vonHoldt 2012).

5.2.4.4. Phylogenetic analyses

To further investigate the structure and interrelation of populations, evolutionary phylogeny was inferred from genetic distances. The Cavalli-Sforza and Edwards (1967) chord distance D_C and Nei´s genetic distance D_A (Nei *et al*. 1983) which have both been shown to perform well with microsatellite markers (Takezaki and Nei 1996), were used for this purpose. TreeFit version 1.2 (Kalinowski 2009) was used to calculate genetic distances, to generate a treefile and to obtain statistical support for interior branches (estimated from 1000 bootstrap iterations). The trees were visualised by the software TREEVIEW version 1.6.6 (Page 1996).

5.2.4.5. Principal component analysis (PCA)

As a complementary approach to investigate population subdivision, the interrelation of genotypes for all nine samples was also visualised by a principal component analysis of F_{ST} values, conducted with PCA-GEN version 1.2.1 (Goudet 1999). The significance of each principal component was assessed from 1000 randomisations of genotypes.

5.2.4.6. Population assignment of individuals and groups of individuals

Assignment testing in GENECLASS2 (Piry *et al*. 2004) was conducted to obtain statistical certainties of assigning Faroe Plateau and Faroe Bank individuals to their correct population of origin. Individuals were re-allocated to known populations of origin, using the leave-one-out procedure, estimated by the Bayesian method described by Rannala and Mountain (1997) and 10,000 resamplings of individuals, as described in Paetkau *et al*. (2004). Type 1 error was set to 0.05. This was done with all data as well as only single year samples for Faroe Plateau and Faroe Bank cod (the two single year samples for which the highest level of genetic differentiation was suggested). Using these two single year samples as reference samples, assignment of groups of individuals was also evaluated (groups consisting of the first 50 individuals in the remaining year samples).

5.2.4.7. Past and present genetic stability of populations

The Faroe Plateau, and especially the Faroe Bank cod, have experienced severe reductions in census size in recent years. To investigate whether this was reflected in any way in the genetic makeup of the populations, the data was investigated to check for evidence of genetic bottlenecks.

An array of methods are available to test for genetic bottlenecks. One of these is based on the *M*-ratio, which is defined as the mean ratio of the number of alleles, *k*, to the range in allele size, *r*. The rationale behind this method is that in a population that is reduced in size, genetic drift is enhanced and alleles will be lost. However, *k* is reduced faster than *r*, since only the loss of the smallest and largest allele will result in a reduction of *r*. Hence, the *M*-ratio is expected to be smaller in recently reduced populations than in populations at equilibrium. A population with an *M*-ratio of 0.68 or less, calculated over seven or more microsatellite loci, is considered likely to have undergone a recent genetic bottleneck (Garza and Williamson 2001). As well as calculating *M*-ratios, the data were also tested for evidence of recent genetic bottlenecks with the software BOTTLENECK version 1.2.02 (Cornuet and Luikart 1996; Piry *et al*. 1999). For each population and each locus, this software compares the observed heterozygosity to the one expected from the observed number of alleles. Populations showing significant excess of heterozygosity are

considered to have undergone a recent reduction in effective population size. Calculations were based on 1000 replications under the infinite allele model (IAM) and the two-phase model (TPM) with 70 % mutations following the stepwise mutation model and with 30 % variance in size of the multistep mutations. Two statistical tests were performed to assess the significance of potential heterozygosity excess, a sign test (Cornuet and Luikart 1996) and a non-parametric Wilcoxon signed-rank test (Luikart *et al*. 1998a). A third statistical test is available in the software, the standardised differences test (Cornuet and Luikart 1996), but as this test requires a minimum of 20 loci, it was not applied to these data. In addition to the statistical tests applied, the mode-shift option in BOTTLENECK was also chosen. This analysis detects populations that have allele frequency distributions that deviate from the typical L-shape distribution, as such deviations have been found to characterise bottlenecked populations (Luikart *et al*. 1998b). For the FPNE and the FB areas, sampling years were evaluated separately as well as pooled.

5.2.4.8. Estimation of gene flow

Overall and pairwise values for the effective number of migrants per generation (*N*m) were calculated as $(1 - F'_{ST})\ (4\ F_{ST})^{-1}$ (Section 2.4.3.2; Meirmans and Hedrick 2011) and inferred from the private allele method (Slatkin 1985; Barton and Slatkin 1986) implemented in GENEPOP.

5.2.4.9. Significance of the depth that the cod inhabited

The reasons why some cod chose to inhabit deep waters and others do not are not fully understood (Steingrund and Ofstad 2010), therefore the last analysis in this study concerned genetic differentiation between cod caught in deep (≥ 150 m) versus in shallow waters. The FPNE data were used for this, as the FPNE area was the only area with a sufficient number of cod caught in deep waters (Table 2.3). The degree and significance of genetic differentiation was calculated as described above for geographic samples.

5.3. Results

5.3.1. Genetic diversity and assumption testing

Over all samples, locus diversity varied considerably, from *Gmo*3, with a total of 9 alleles, an observed heterozygosity of 0.14 and a PIC value of 0.145, to *Gmo*8, with 53 alleles, an observed heterozygosity of 0.94 and a PIC value of 0.937 (Table 5.2).

Table 5.2. Genetic variation, calculated over all samples, of the ten microsatellite loci. Observed (H_O) and expected (H_E) heterozygosities, total number of alleles (*NA*), polymorphic information content (PIC) and size range of alleles.

Locus	H_o	H_e	*NA*	PIC	Allele size Range (bp)
*Gmo*8	0.94	0.94	53	0.937	108-316
*Gmo*19	0.81	0.92	29	0.915	118-230
Tch13	0.88	0.91	48	0.903	76-172
*Gmo*35	0.84	0.83	15	0.806	114-159
*Gmo*37	0.82	0.85	21	0.830	195-307
Tch11	0.92	0.93	29	0.928	110-230
*Gmo*2	0.70	0.84	21	0.822	103-147
*Gmo*3	0.14	0.15	9	0.145	159-207
*Gmo*34	0.65	0.65	11	0.610	74-118
*Gmo*132	0.91	0.92	39	0.916	105-221

A high proportion (23 / 90) of the total locus × sample combinations deviated significantly from HWE and seven of these remained significant following sequential Bonferroni adjustments (Table 5.3). The data were investigated in Micro-Checker, where the potential occurrence of null alleles at significant frequency were indicated for *Gmo*2 and *Gmo*19 in most samples. The two remaining deviations, not involving these two loci, belonged to two different loci and two different samples / areas. Of 45 tests one case of linkage disequilibrium was found (*Gmo*2 and *Gmo*34), but this was not significant after sequential Bonferroni correction (critical $\alpha = 0.05 / 45 = 0.0011$). *Gmo*2 and *Gmo*19 were excluded from all further analyses. None of the markers showed indications of being influenced by selective forces, as evaluated with outlier tests, applied in LOSITAN. Both the infinite alleles and the stepwise mutation models were examined.

Average expected heterozygosities and allelic richness estimates, as calculated over all loci, were very similar in all samples, when *Gmo*2 and *Gmo*19 were excluded. Average expected heterozygosities ranged from 0.76 (FPNE08, FB09, BGF and NWB) to 0.77 (the remaining samples) and average allelic richness estimates from 11.87 (BGF) to 14.77 (FPNE09).

Table 5.3. Genetic variation of the ten microsatellite loci in all nine samples. Sample size (*N*), number of individuals genotyped (*n*), number of alleles (*A*), allelic richness (*AR*), observed (H_O) and expected (H_E) heterozygosities and inbreeding coefficients (F_{IS}). *AR* was calculated from 1000 resamples based on the smallest sample size. Bold values indicate significant deviation from Hardy-Weinberg equilibrium following sequential Bonferroni correction (first critical $\alpha = 0.05 / 10 \text{ loci} = 0.0050$).

	Area:	**FPNE**	**FPNE**	**FPNE**	**FPW**	**FB**	**FB**	**FB**	**BGF**	**NWB**
	Year:	**2008**	**2009**	**2010**	**2009**	**2008**	**2009**	**2010**	**2005**	**2006**
Locus	***N:***	**130**	**170**	**288**	**116**	**138**	**122**	**105**	**47**	**48**
***Gmo*8**	*n*	126	169	288	116	136	122	105	34	47
	A	37	37	45	39	35	37	38	23	21
	AR	26.20	25.48	26.83	27.27	22.34	25.02	24.99	17.78	17.23
	H_O	0.91	0.95	0.95	0.97	0.96	0.93	0.92	0.88	0.83
	H_E	0.95	0.94	0.94	0.95	0.93	0.94	0.93	0.90	0.83
	F_{IS}	0.038	-0.008	-0.014	-0.022	**-0.030**	0.006	0.017	0.030	0.008
***Gmo*19**	*N*	127	170	287	115	137	122	105	33	46
	A	23	25	25	23	25	25	23	17	21
	AR	18.66	19.11	19.11	19.25	19.54	19.88	18.48	14.92	19.02
	H_O	0.81	0.80	0.80	0.80	0.83	0.77	0.78	0.91	0.87
	H_E	0.91	0.91	0.92	0.92	0.91	0.92	0.92	0.91	0.93
	F_{IS}	**0.112**	0.125	**0.131**	**0.135**	**0.091**	0.162	0.158	0.012	0.079
***Tch*13**	*N*	127	170	288	116	136	120	105	34	47
	A	31	34	42	26	33	30	33	18	22
	AR	19.81	21.61	21.82	19.97	23.26	20.66	21.68	15.66	17.65
	H_O	0.91	0.89	0.85	0.91	0.87	0.84	0.90	0.85	0.98
	H_E	0.89	0.91	0.90	0.91	0.91	0.91	0.91	0.90	0.88
	F_{IS}	-0.014	0.025	0.057	-0.005	0.055	0.079	0.008	0.065	-0.099
***Gmo*35**	*N*	128	170	288	116	137	122	105	47	48
	A	9	12	11	9	10	10	7	9	7
	AR	7.75	8.44	7.66	7.26	8.01	7.98	6.6	8.69	7.00
	H_O	0.88	0.82	0.81	0.90	0.88	0.81	0.85	0.79	0.88
	H_E	0.83	0.83	0.82	0.82	0.83	0.82	0.82	0.79	0.82
	F_{IS}	-0.045	0.006	0.019	-0.091	-0.047	0.015	-0.025	0.018	-0.053
***Gmo*37**	*n*	128	169	288	115	137	122	105	31	48
	A	14	16	16	12	13	14	13	10	11
	AR	10.94	12.14	11.55	11.05	11.43	11.59	10.94	8.94	10.31
	H_O	0.79	0.83	0.85	0.78	0.91	0.79	0.75	0.81	0.85
	H_E	0.83	0.84	0.86	0.82	0.86	0.85	0.84	0.81	0.84
	F_{IS}	0.058	0.015	0.010	0.048	-0.052	0.074	0.106	0.023	-0.009
***Tch*11**	*n*	128	170	288	116	137	122	105	47	48
	A	22	25	25	22	20	20	17	19	17
	AR	18.14	18.86	18.76	18.45	17.51	16.99	15.54	17.25	15.52
	H_O	0.96	0.93	0.93	0.90	0.93	0.91	0.88	0.91	0.96
	H_E	0.93	0.93	0.93	0.93	0.93	0.92	0.92	0.93	0.91
	F_{IS}	-0.033	0.003	0.009	0.043	-0.005	0.019	0.054	0.028	-0.047
***Gmo*2**	*n*	122	165	279	115	134	117	103	46	48

	Area:	**FPNE**	**FPNE**	**FPNE**	**FPW**	**FB**	**FB**	**FB**	**BGF**	**NWB**
	Year:	**2008**	**2009**	**2010**	**2009**	**2008**	**2009**	**2010**	**2005**	**2006**
Locus	***N:***	**130**	**170**	**288**	**116**	**138**	**122**	**105**	**47**	**48**
	A	15	18	14	17	9	10	9	11	11
	AR	10.74	12.38	10.63	11.97	8.26	8.61	8.91	9.86	9.7
	H_O	0.74	0.72	0.61	0.74	0.70	0.72	0.75	0.72	0.79
	H_E	0.82	0.83	0.84	0.85	0.83	0.83	0.84	0.83	0.85
	F_{IS}	0.101	0.138	**0.270**	0.133	0.154	0.137	0.120	0.151	0.079
***Gmo*3**	*n*	128	170	288	116	137	122	105	47	48
	A	6	7	7	5	5	5	7	5	4
	AR	4.23	4.71	3.97	3.97	3.39	3.54	4.85	4.34	3.90
	H_O	0.14	0.15	0.13	0.15	0.08	0.11	0.15	0.21	0.31
	H_E	0.15	0.15	0.15	0.16	0.10	0.10	0.15	0.22	0.28
	F_{IS}	0.056	-0.010	0.134	0.102	0.189	-0.033	-0.044	0.022	-0.116
***Gmo*34**	*n*	128	168	288	115	134	121	105	47	48
	A	8	9	9	8	9	9	8	6	10
	AR	6.80	7.15	6.72	6.11	6.9	6.42	6.90	5.83	9.03
	H_O	0.65	0.65	0.61	0.65	0.71	0.64	0.60	0.70	0.69
	H_E	0.62	0.63	0.63	0.66	0.67	0.66	0.65	0.62	0.67
	F_{IS}	-0.035	-0.033	0.033	0.012	-0.059	0.021	0.077	-0.113	-0.018
***Gmo*132**	*n*	124	169	288	114	135	120	104	47	48
	A	26	28	34	26	26	27	26	18	17
	AR	19.46	19.79	20.59	19.25	19.38	20.25	19.67	16.47	14.88
	H_O	0.91	0.90	0.91	0.94	0.92	0.94	0.89	0.83	0.90
	H_E	0.91	0.92	0.91	0.92	0.92	0.92	0.91	0.87	0.88
	F_{IS}	0.007	**0.030**	0.006	-0.019	0.005	-0.022	0.027	0.061	-0.008
Over	*n*	126.6	169.0	287.0	115.4	136.0	121.0	104.7	41.3	47.6
all	*A*	191	211	228	187	185	187	181	136	141
loci	*AR*	14.27	14.97	14.76	14.46	14.00	14.09	13.86	11.97	12.42
	H_O	0.77	0.77	0.75	0.77	0.78	0.75	0.75	0.76	0.81
	H_E	0.78	0.79	0.79	0.79	0.79	0.79	0.79	0.78	0.79
	F_{IS}	0.023	0.034	0.058	0.028	0.016	0.054	0.058	0.034	-0.011

5.3.2. Genomic location of microsatellites

Annotation of microsatellite loci revealed the fragmented nature of the cod genome, as for two (*Gmo*8 and Gm37) of the eight loci with full query sequences available, the alignment was divided into two different contigs of the genome that had not yet been assembled into scaffolds. All hits were highly significant, though, with *e*-values $\leq 1.60 \times 10^{-87}$.

The query sequences of *Gmo*2 and *Gmo*132 encompassed many missing positions, resulting in high *e*-values (0.011 and 0.027, respectively). However, both sequences aligned to single contigs,

and perfect matches were found for three of the primer sequences. The remaining primer, *Gmo*2 forward primer, differed from the genomic sequence in the two first positions: the forward primer sequence used in this study was (5'-3') CCC-TCA-GAT-TCA-AAT-GAA-GGA, but the corresponding genomic sequence was AGC-TCA-GAT-TCA-AAT-GAA-GGA.

Four of the loci were found in gene regions. Three of them (*Gmo*19, *Gmo*37 and *Gmo*3 – all tetra repeats) were located in introns with the remaining one (*Gmo*2 – a dinucleotide repeat), located at an intron-exon boundary. In addition, two of the loci (*Gmo*8 and *Gmo*34) were located relatively close to genes; the repeat regions of these loci were within 500 and 200 bp down- and upstream of genes, respectively (Table 5.4). All of these genes were classified as "known by projection" meaning that they are homologous to genes that are known in other species (www.ensembl.org). The remaining four loci were found in intergenic regions. Two of them (*Gmo*35 and *Tch*11) aligned to unplaced contigs containing no genes and for the other two (*Tch*13 and *Gmo132*), the nearest genes were > 9,000 and > 18,000 bp away, respectively.

There were no indications that *Tch*11 and *Tch*13, originally developed from walleye pollock (O'Reilly *et al*. 2000), showed less homology than the other loci, when aligned to the cod genome.

Table 5.4. Genomic location of microsatellites based on BLAT hits in Ensembl (www.ensembl.org). Biological function of genes was inferred from electronic annotation by Ensembl. For loci up- and downstream of genes, the basepair (bp) distances are estimated from the repeat region of the microsatellite.

		Gene			Cellular	
Locus	**Genomic region**	**Ensembl ID**	**Symbol**	**Biological process**	**Component**	**Molecular function**
*Gmo*8	< 500 bp 3' of gene	ENSGMOG 00000010129	zgc [a)]	GTP catabolic process, signal transduction, small GTPase mediated signal transduction, protein transport.	Intracellular, membrane.	GTPase activity, GTP binding.
*Gmo*19	Intronic	ENSGMOG 00000012337	CSMD1	Carbohydrate metabolic process.		Hydrolase activity, hydrolysing O-glycosyl compounds.
Tch13	Intergenic					
*Gmo*35	Intergenic					
*Gmo*37	Intronic	ENSGMOG 00000019486	GRK6	Protein phosphorylation, termination of G-protein coupled receptor signaling pathway.		Protein kinase activity, protein serine/threonine kinase activity, protein tyrosine kinase activity, ATP binding.
Tch11	Intergenic					
*Gmo*2	Intronic / exonic	ENSGMOG	ndrg1b			

Locus	Genomic region	Gene		Biological process	Cellular Component	Molecular function
		Ensembl ID	Symbol			
		00000014462				
*Gmo*3	Intronic	ENSGMOG 00000004539	si:ch1073-80j9.1 [b)]			Nucleic acid binding.
*Gmo*34	< 200 bp 5' of gene	ENSGMOG 00000008556	kcnd1	Ion transport, protein homooligomerisation, transmembrane transport.	Membrane.	Ion channel activity, protein binding.
*Gmo*132	Intergenic					

a) Changed to RAB21 in the updated version of the cod genome (gadMor1, version 78.1, released in December 2014).

b) Changed to MYOCD in the updated version of the cod genome (gadMor1, version 78.1, released in December 2014) – with no associated molecular function.

5.3.3. Population structure

The validity of pooling sampling years was evaluated by pairwise D_{Jost_EST} and F_{ST} values as well as by pseudo-exact tests of genic differentiation. None of these analyses detected significant differentiation among years within locations. The genic differentiation test was non-significant in all cases, D_{Jost_EST} and F_{ST} values were zero and their calculated 95 % CIs included zero (boxed portions in Table 5.5). Thus, an absence of short-term temporal instability in FPNE and FB allele frequencies was assumed and temporal samples were pooled for further investigation of area differentiation. These analyses lend some support to the notion that the single samples for the remaining areas, FPW, BGF and NWB, are also representative.

Table 5.5. Matrices of pairwise D_{Jost_EST} (above diagonal) and F_{ST} values (below diagonal) among areas and / or years sampled. 95 % confidence intervals, as calculated from 1000 bootstrap replicates, are indicated below the differentiation estimates. Bold D_{Jost_EST} values indicate significant results of the pseudo-exact test of genic differentiation following sequential Bonferroni adjustment (first critical α = 0.05 / 15 tests = 0.0033). Intra-area values in boxes.

		FPNE			**FB**	
	2008	2009	2010	2008	2009	2010
FPNE						
2008		0.0000 -0.0088-0.0110	0.0000 -0.0071-0.0094	**0.0008** -0.0094-0.0133	**0.0026** -0.0091-0.0180	**0.0014** -0.0098-0.0156
2009	0.0001 -0.0018-0.0025		0.0000 -0.0066-0.0084	0.0005 -0.0079-0.0112	0.0006 -0.0084-0.0129	0.0051 -0.0072-0.0195
2010	-0.0001 -0.0016-0.0018	0.0000 -0.0013-0.0018		**0.0024** -0.0052-0.0131	**0.0037** -0.0048-0.0148	**0.0043** -0.0065-0.0176
FB						
2008	0.0013 -0.0009-0.0039	0.0004 -0.0014-0.0027	0.0014 -0.0004-0.0034		0.0000 -0.0092-0.0113	-0.0001 -0.0108-0.0148
2009	0.0016 -0.0010-0.0048	0.0003 -0.0016-0.0028	0.0016 -0.0002-0.0038	-0.0004 -0.0024-0.0020		-0.0003 -0.0110-0.0139
2010	0.0010 -0.0015-0.0039	0.0014 -0.0011-0.0045	0.0017 -0.0004-0.0043	-0.0003 -0.0028-0.0026	-0.0007 -0.0032-0.0024	

5.3.3.1. Genetic differentiation among areas

Significant genetic differentiation among the five areas was evident by D_{Jost} and the genic differentiation test ($P < 0.001$) although the absolute level of differentiation was low ($D_{Jost} = 0.0426$; 95 % CI 0.04 – 0.07). F_{ST} and pertaining confidence intervals did not reveal any significant differentiation ($F_{ST} = 0.0034$; 95 % CI 0.00 – 0.01). By combining both D_{Jost} and the genic differentiation test, area differentiation was detected at all loci (Table 5.6).

Table 5.6. Overall and locus-specific global differentiation for the areas FPNE (N = 586), FPW (N = 116), FB (N = 364), BGF (N = 47) and NWB (N = 48). D_{Jost}, F_{ST} and 95 % confidence intervals (CI), calculated from 1000 bootstrap replicates, are indicated as well as P-values from pseudo-exact tests of genic differentiation. *NS* indicates no significant differentiation.

Locus	D_{Jost}	**95% CI**	F_{ST}	**95% CI**	P
*Gmo*8	0.2223	0.14-0.31	0.0057	0.01-0.01	**< 0.001**
Tch13	0.0596	0.02-0.11	0.0004	0.00-0.01	*NS*
*Gmo*35	0.0666	0.03-0.11	0.0041	0.00-0.01	**< 0.001**
*Gmo*37	0.0504	0.01-0.10	0.0016	0.00-0.01	*NS*
Tch11	0.1467	0.09-0.21	0.0026	0.00-0.01	**< 0.001**
*Gmo*3	0.0036	0.00-0.01	0.0055	0.00-0.02	**< 0.05**
*Gmo*34	0.0100	0.00-0.02	0.0005	0.00-0.01	**< 0.05**
*Gmo*132	0.1611	0.09-0.23	0.0076	0.01-0.01	**< 0.001**
All	0.0426	0.04-0.07	0.0034	0.00-0.01	**< 0.001**

Pairwise genetic differentiation was also evaluated by the three different statistics D_{Jost_EST}, F_{ST} and pseudo-exact tests of genic differentiation (Table 5.7 and Table 5.8) and, except for the intra-Plateau and intra-Norway comparisons, D_{Jost_EST} values were considerably higher than F_{ST}. There was a strong positive correlation between these two estimates ($r^2 = 0.922$). Levels of differentiation for the intra-Plateau and the intra-Norway comparisons were low and nonsignificant by both D_{Jost_EST} and F_{ST} and the values were of a similar level with D_{Jost_EST} values of 0.0008 and 0.0007 and F_{ST} values of 0.0013 and 0.0012, respectively.

The largest differentiation values were seen in the pairwise comparisons involving Faroese and Norwegian areas and, by all calculations, higher levels of genetic differentiation were detected among Faroese and NWB than among Faroese and BGF cod. All of these differentiation values

were statistically significant, as evaluated by pseudo-exact tests of genic differentiation and the 95 % CIs pertaining to F_{ST}, but the 95 % CIs pertaining to D_{Jost_EST} were only significant for the Faroese – NWB, not the Faroese – BGF comparisons (Table 5.7 and Table 5.8).

Results were less clear regarding differentiation among Faroe Plateau and Faroe Bank cod. Although the D_{Jost_EST} values, 0.0027 for the FPNE – FB and 0.0048 for the FPW – FB comparison, were somewhat higher than the values found for the intra-Plateau and intra-Norway comparisons, the F_{ST} values, 0.0014 for the FPNE – FB and 0.0018 for the FPW – FB comparison were barely higher than these were. However, the 95 % CIs for D_{Jost_EST} did not indicate any significant differentiation for either comparison, but the 95 % CI for F_{ST} was significant for the FPNE – FB comparison. Lastly, pseudo-exact tests of genic differentiation were significant for both Plateau – Bank comparisons (Table 5.7 and Table 5.8).

Table 5.7. Matrices of pairwise D_{Jost_EST} with 95 % confidence intervals, as calculated from 1000 bootstrap replicates, (below diagonal) and geographic distances in km (above diagonal).

Areas	FPNE	FPW	FB	BGF	NWB
FPNE		91	182	655	679
FPW	0.0008 -.0088-0.0110		92	711	700
FB	0.0027 -.0008-0.0071	0.0048 -0.0046-0.0166		779	739
BGF	0.0185 -.0070-0.0473	0.0189 -0.0124-0.0539	0.0158 -0.0117-0.0512		292
NWB	0.0248 0.0098-0.0442	0.0305 0.0087-0.0572	0.0386 0.0208-0.0593	0.0007 -0.0231-0.0352	

Table 5.8. Matrices of pairwise F_{ST} with 95 % confidence intervals, as calculated from 1000 bootstrap replicates, (below diagonal) and *P*-values from pseudo-exact tests of genic differentiation (above diagonal). Bold values indicate significant *P*-values following sequential Bonferroni adjustment (first critical α = 0.05 / 10 tests = 0.005).

Areas	**FPNE**	**FPW**	**FB**	**BGF**	**NWB**
FPNE		*NS*	**<0.0001**	**<0.01**	**<0.0001**
FPW	0.0013 -0.0008-0.0038		**<0.001**	**<0.0001**	**<0.0001**
FB	0.0014 0.0006-0.0024	0.0018 -0.0002-0.0044		**<0.0001**	**<0.0001**
BGF	0.0061 0.0004-0.0129	0.0090 0.0017-0.0178	0.0097 0.0027-0.0177		*NS*
NWB	0.0110 0.0064-0.0169	0.0137 0.0073-0.0217	0.0136 0.0084-0.0203	0.0012 -0.0062-0.0107	

There was no apparent relationship between genetic differentiation estimates and geographic distances (Table 5.7). Geographically, the NWB and BGF areas are approximately equally close to the Faroese areas; still there appears to be a closer genetic connection between BGF and Faroese cod than between NWB and Faroese cod. Likewise, geographic distance alone cannot explain patterns of genetic differentiation seen for Faroese cod, where there appears to be a closer connection between cod from the FPNE area and FB cod than between cod from the FPW area and FB cod.

Altogether, these results indicate significant genetic differentiation between Faroese and Norwegian cod. Furthermore, the results suggest some degree of genetic differentiation between Faroe Plateau and Faroe Bank cod, although not detectable by all of the statistical analyses applied.

5.3.3.2. Bayesian clustering analysis

Bayesian clustering analysis in STRUCTURE failed to replicate the population structure described above. Ln*P*(D) plateaued between $K = 1$ and $K = 2$ and then levelled off, while ΔK was maximised at $K = 2$ (Figure 5.1), thus suggesting two major genotypic clusters in the data.

Proportions of membership of each of the population samples into each of the two inferred clusters were approximately equally distributed between the two (Table 5.9).

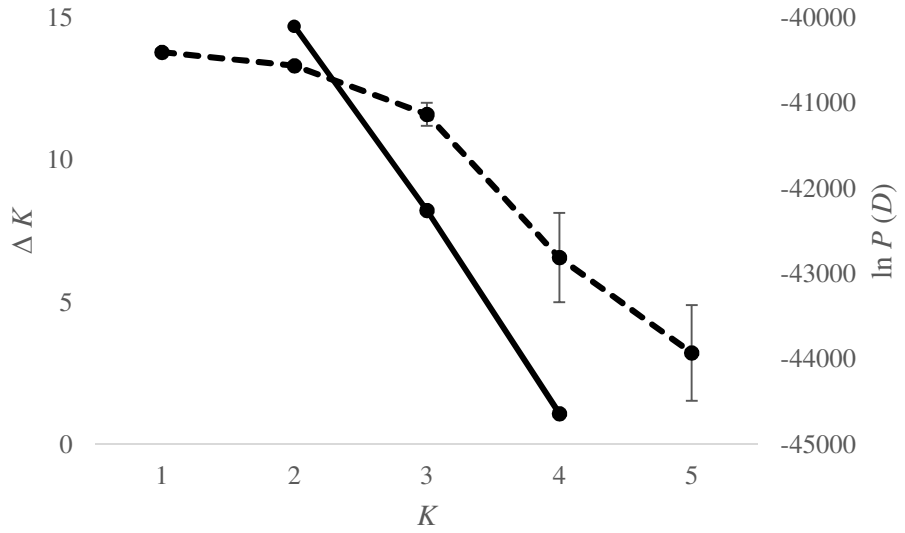

Figure 5.1. Plots used to detect the number of major genotypic clusters (*K*) of the data. Full line and left axis: Δ*K* estimated from Evanno *et al*. (2005). Stipled line and right axis: Average ln*P*(*D*) (±SD) over five runs of each *K*.

Table 5.9. Estimated proportions of membership to each major genotypic cluster, as determined by Bayesian clustering analysis.

Population sample	**Cluster 1**	**Cluster 2**
Faroe Plateau North-East	0.499	0.501
Faroe Plateau West	0.503	0.497
Faroe Bank	0.498	0.502
Borgundfjord	0.500	0.500
North-West of Bomlø	0.493	0.507

5.3.4. Phylogenetic analyses

Evolutionary distances among populations were visualised by a combination of different genetic distances and tree-building algorithms (Figure 5.2). All four trees depicted a similar topology and bootstrap values were high (≥ 0.93) – with one exception, a bootstrap value of 0.39 for the branch separating FPNE and FB cod from the rest in the NJ tree built from Nei´s D_A. All trees grouped FPNE and FB cod as sister nodes and evolutionary distances among the Faroese populations were

small compared to distances within the Norwegian cod. All trees show an evolutionary divergence between FPW cod and the remaining Faroese cod.

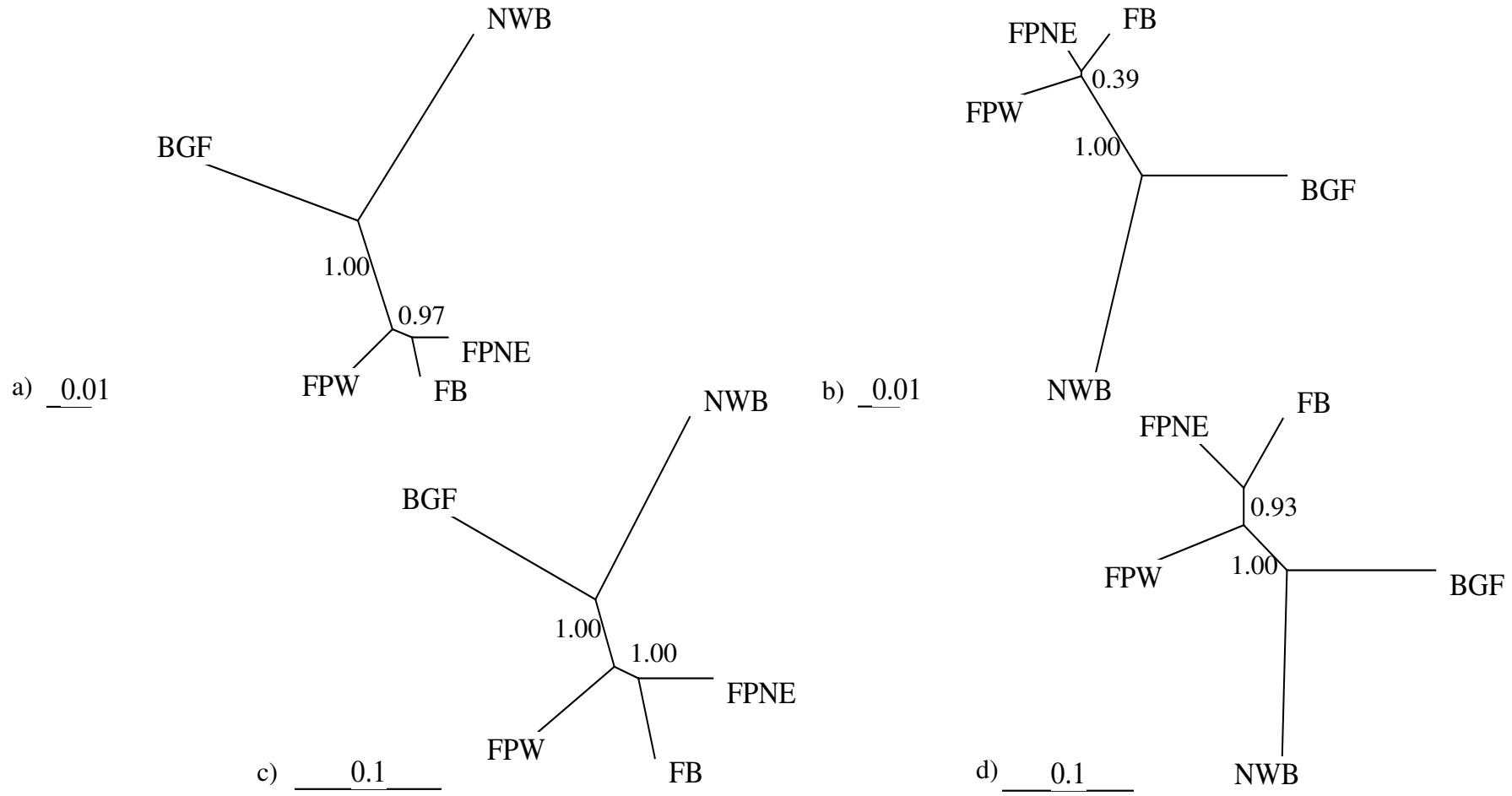

Figure 5.2. Unrooted trees of Cavalli-Sforza, Edward chord distances (D_C) and Nei´s genetic distances (D_A) from eight microsatellite loci created by the neighbour-joining (NJ) and the UPGMA tree-building algorithms. Bootstrap values, obtained after 1000 iterations, are indicated beside the branches. Scaled distances are shown beneath each figure. a) D_A, UPGMA; b) D_A, NJ; c) D_C, UPGMA and d) D_C, NJ.

5.3.5. Principal component analysis

Principal component analysis of the data (years kept separate) showed one significant axis (PC I; P = 0.001), which explained 36.75 % of the total variation. Genotypes were ordered along this axis so that the Norwegian genotypes fell into one cluster and the Faroese genotypes in another cluster (Figure 5.3). These results were in overall agreement with results from above, the largest divergence was seen among Faroese and Norwegian genotypes and all Faroe Plateau samples grouped together and similarly for the Faroe Bank samples. Regarding the FPW cod, the PCA grouped these together with the rest of the Faroe Plateau cod and especially the FPNE sample from 2009 (the FPW cod was also caught in 2009), thus, no statistically significant differentiation in allele frequency composition between the FPW and FPNE cod was detected with this approach.

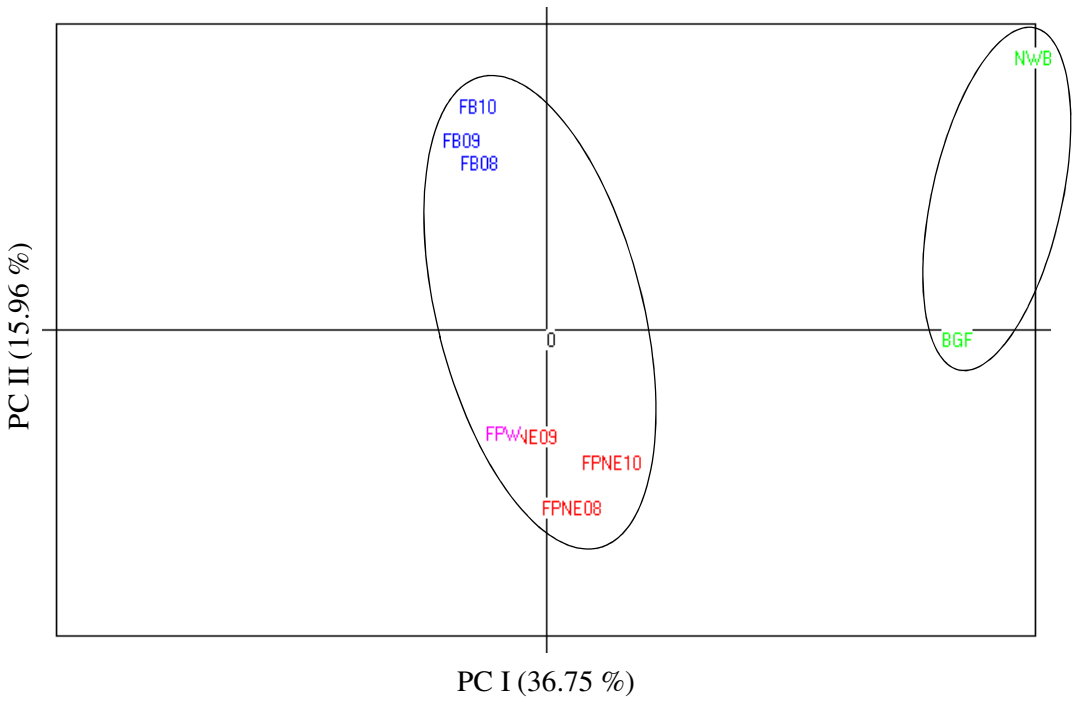

Figure 5.3. Principal component analysis based on microsatellite allele frequencies. The percentage of total variation explained by each of the two principal components is indicated. Faroe Plateau North-East samples in red, Faroe Plateau West sample in pink, Faroe Bank samples in blue and Norwegian samples in green.

5.3.6. Population assignment of individuals and groups of individuals

The results above indicate subtle genetic divergence between Faroe Plateau and Faroe Bank cod. However, the ability of the microsatellite markers to assign individuals to their population of origin was poor. Assignment tests of all Faroe Plateau (based on the results above, FPNE and FPW were grouped for these analyses) and Faroe Bank individuals via GENECLASS2 version 2.0 showed that 63.2 % correct assignments were obtained (quality index = 52.99). Among all pairwise comparisons, analysis above (Table 5.5) showed highest $D_{\text{Jost\_EST}}$ and F_{ST} values between the two 2010 samples. Although pertaining 95 % confidence intervals included zero, the pseudo-exact test of genic differentiation was significant for this comparison. When only these two samples were considered, a few more correct assignments were obtained (64.4 % correct assignments; quality index = 55.12). Hence, the 2010 samples were used as reference samples for the group assignments. The groups consisted of Faroe Plateau and Faroe Bank samples collected in 2008 and 2009 (50 individuals per each of the four groups). These assignments were ineffective, however, as only the Faroe Plateau groups were correctly assigned, while both Faroe Bank groups were wrongly assigned to the Faroe Plateau population.

5.3.7. Inference of past and present genetic stability

M-ratio analysis of the microsatellite data did not provide convincing evidence for recent bottlenecks among the sampled areas, ratios varying between 0.68 and 0.87. Garza and Williamson (2001) considered values of 0.68 and lower to be indicative of potential reduction of population sizes. Similarly, all BOTTLENECK computations were non-significant (Table 5.10).

Table 5.10. *M*-ratios and results from BOTTLENECK for the significance, Wilcoxon and mode-shift tests. *P*-values provided for the Wilcoxon test are from the one-tailed test for heterozygote excess. For FPNE and FB pooled samples are evaluated as well as sampling years. *NS* indicates no significant differentiation.

		Significance test		**Wilcoxon test**		
	***M*-ratios**	**IAM**	**TPM**	**IAM**	**TPM**	**Mode-shift**
FPNE08	0.74	0.320	0.193	0.156	0.770	*NS*
FPNE09	0.79	0.108	0.195	0.098	0.844	*NS*
FPNE10	0.86	0.315	0.173	0.156	0.809	*NS*
FPNE all years	0.87	0.110	0.055	0.098	0.963	*NS*
FPW09	0.81	0.105	0.415	0.098	0.578	*NS*
FB08	0.81	0.109	0.420	0.098	0.629	*NS*
FB09	0.72	0.104	0.190	0.098	0.809	*NS*
FB10	0.77	0.317	0.183	0.125	0.809	*NS*
FB all years	0.81	0.099	0.183	0.098	0.770	*NS*
BGF	0.68	0.323	0.187	0.273	0.809	*NS*
NWB	0.71	0.410	0.189	0.371	0.902	*NS*

5.3.8. Estimation of gene flow

Migration levels varied with the different analytical methods applied. Overall estimates of *N*m were 47 and 95 with the private allele method (Slatkin 1985; Barton and Slatkin 1986) and the $(1 - F'_{ST})\ (4\ F_{ST})^{-1}$ method (Meirmans and Hedrick 2011), respectively. This pattern of lower estimates obtained by the private allele method was also apparent in the pairwise estimates, but both methods showed that apparent gene flow between either of the two Faroese areas and Norway was several times smaller than the gene flow between the two Faroese areas (Table 5.11).

Table 5.11. Pairwise *N*m estimates. Values estimated by the private allele method (Slatkin 1985; Barton and Slatkin 1986) above the diagonal and values estimated from a combination of F_{ST} and F'_{ST} (Meirmans and Hedrick 2011) below the diagonal.

	Faroe Plateau	**Faroe Bank**	**Norway**
Faroe Plateau		44	11
Faroe Bank	624		14
Norway	28	28	

5.3.9. Outlier behaviour at the *Gmo*132 locus in relation to depth

When evaluating differentiation between FPNE cod caught at different depths (above or below 150 meters deep) significant genetic differentiation was detected by D_{Jost} (D_{Jost} = 0.0245; 95 % CI 0.01 – 0.03) and the genic differentiation test ($P < 0.001$). Again, F_{ST} failed to detect any significant differentiation (F_{ST} = 0.0045; 95 % CI 0.00 – 0.01). In contrast to the area differentiation analysis, where differentiation was evident over all loci (Table 5.6), depth differentiation was only significant for three loci (*Gmo*8, *Gmo*34 and *Gmo*132), but one in particular, *Gmo*132, stood out as an outlier locus with a D_{Jost} value of 0.2283 (95 % CI 0.12 – 0.34; Table 5.12).

Visual inspection of allele frequency distributions at *Gmo*132 among all areas as well as the two FPNE depth samples revealed one allele (117 bp) that might be correlated with the outlier behaviour of this locus. This allele was found at much higher frequencies than the surrounding alleles, especially in the FPNE cod caught in deep compared to in shallow waters (Figure 5.4). No such comparison was made for Faroe Bank cod, where only seven fish were caught in deep waters. The 117 allele was found at a relatively low frequency (2.79 %) in the Faroe Bank cod.

Table 5.12. Overall and locus-specific global differentiation at the depth contrast, based on all FPNE cod divided into deep (≥ 150 m; N = 100) and shallow waters (< 150 m; N = 486). D_{Jost}, F_{ST} and 95 % confidence intervals (CI), calculated from 1000 bootstrap replicates are indicated as well as results from pseudo-exact tests of genic differentiation (P). *NS* indicates no significant differentiation.

Locus	D_{Jost}	**95% CI**	F_{ST}	**95% CI**	P
*Gmo*8	0.0894	0.02-0.18	0.0026	0.00-0.01	*NS*
Tch13	0.0242	0.00-0.06	-0.0004	0.00-0.01	*NS*
*Gmo*35	0.0164	-0.01-0.05	0.0005	0.00-0.01	*NS*
*Gmo*37	0.0078	-0.01-0.04	-0.0016	0.00-0.01	*NS*
Tch11	0.0163	-0.02-0.06	-0.0018	0.00-0.00	*NS*
*Gmo*3	0.0006	0.00-0.00	0.0002	0.00-0.02	*NS*
*Gmo*34	0.0247	0.01-0.05	0.0125	0.00-0.03	**< 0.05**
*Gmo*132	0.2283	0.12-0.34	0.0219	0.01-0.04	**< 0.001**
All	0.0245	0.01-0.03	0.0045	0.00-0.01	**< 0.001**

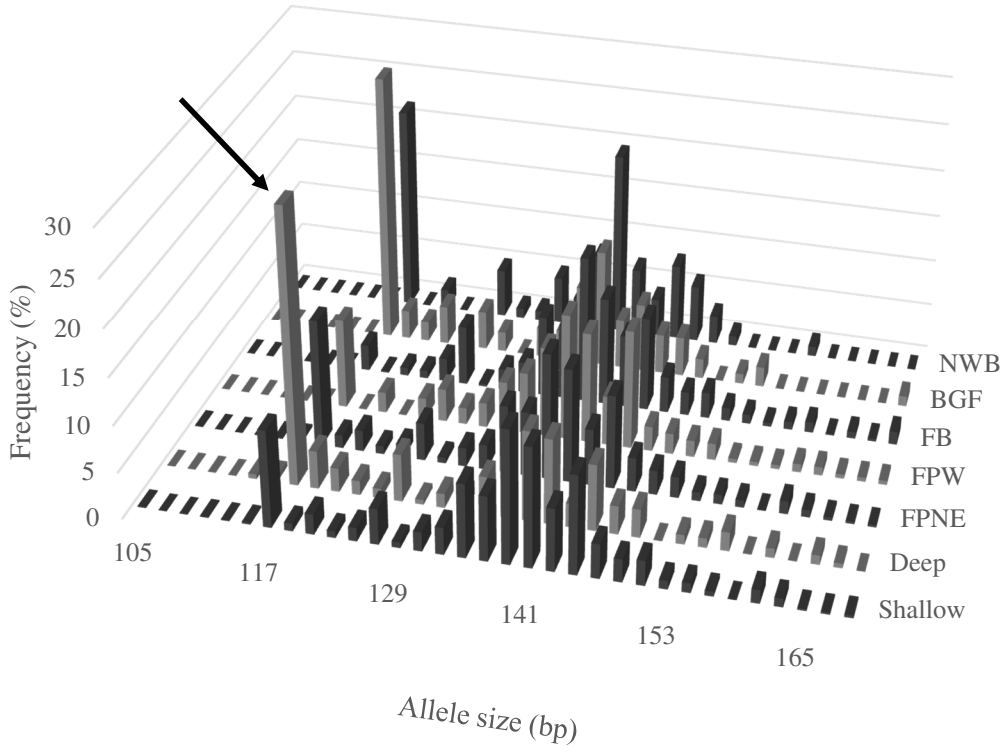

Figure 5.4. Allele frequencies of *Gmo*132 for all areas investigated as well as for the FPNE area divided into deep and shallow waters. The last allele frequency class includes all alleles 171 bp and above. Arrow points to the small allele (117 bp) for the FPNE Deep sample.

5.4. Discussion.

The main objective of this work was to assess generated microsatellite data for evidence of potential population genetic structuring between Faroe Plateau and Faroe Bank cod. Overall, the various analyses herein did suggest that some level of genetic differentiation may exist between Faroe Plateau and Faroe Bank cod, though it is relatively subtle, in that not all tests applied were significant. Concerning structure within the Faroe Plateau cod, all phylogenetic analyses grouped the FPNE and FB areas as sister nodes with the FPW sample placed closer to the Norwegian samples, but the remaining analyses did not detect any significant structure within the Faroe Plateau. As discussed below, this may be the result of sampling bias for this location.

The recent reductions in census population sizes recorded for Faroe Plateau and, particularly, Faroe Bank cod has apparently not resulted in lower genetic diversity compared to another nearby cod population. Average expected heterozygosities as well as allelic richness estimates were comparable among all samples and there was no evidence of genetic bottlenecks.

Lastly, results herein may challenge the assumption of microsatellite neutrality, since approximately half of the loci investigated were either in or close to predicted gene regions of the cod genome and, moreover, data analysis revealed evidence of non-neutrality at the *Gmo*132 locus, although this was only evident in relation to capture depth of the FPNE cod. Potential consequences of this are discussed below.

5.4.1. The assumed neutrality of microsatellite markers

5.4.1.1. Non-uniform distribution throughout the genome

Microsatellite markers are distributed throughout the genome and in eukaryotes, they are more often found in noncoding than in coding regions. For vertebrates only *c*. 9 – 15 % are found in coding regions (Chistiakov *et al*. 2006). Accordingly, microsatellites have generally been considered selectively neutral (Chistiakov *et al*. 2006). However, there is a growing body of evidence to suggest that at least a proportion of these markers may be functionally important. Long short-motif microsatellites have for example been associated with recombination hotspots

in yeast (Bagshaw *et al*. 2008) and microsatellites appear to be associated with gene promoters more often than expected by chance (Sawaya *et al*. 2013) where repeat number variation may be involved in gene regulation (Vinces *et al*. 2009; Lee and Maheshri 2012; Sawaya *et al*. 2013).

There are various methods available for identifying / isolating microsatellite markers and one of these is based on screening expressed sequence tag (EST) sequences, thus resulting in "genic" microsatellites (Senan *et al*. 2014). However, the microsatellites used in the current study were developed by constructing and screening genomic libraries followed by hybridisation with repeat-specific probes (Brooker *et al*. 1994; Miller *et al*. 2000; O'Reilly *et al*. 2000), a method that should predominantly result in markers from noncoding intergenic regions or type II microsatellites (Chistiakov *et al*. 2006). Hence, to find that six out of these ten microsatellites were located in or close to genes was a rather unexpected finding.

Microsatellite markers of various repeat motifs are not uniformly distributed throughout the genome and both tetra- and dinucleotide repeats are, for example, very rarely found in vertebrate exonic regions (Tóth *et al*. 2000; Li *et al*. 2002; Li *et al*. 2004), likely due to purifying selection from frameshift mutations (Li *et al*. 2002). In this study, five of the six microsatellite loci found in (or in close proximity to) genes were tetranucleotide repeats, but were found outside the exonic regions, whereas the sixth locus, a dinucleotide repeat (*Gmo*2) partly aligned to an exonic region. The present as well as other studies (Karlsson and Mørk 2005; Dahle *et al*. 2006), found significant evidence of null alleles at this locus. In the absence of pedigree data, null alleles are assumed due to excess of homozygotes, although such an excess may for example be the result of selection against heterozygotes. However, null alleles may also be the result of unidentified polymorphisms at primer binding sites (Callen *et al*. 1993; O´Reilly and Wright 1995; Paetkau and Strobeck 1995). In concordance, a mismatch was found in the present study between the *Gmo*2 forward primer and the corresponding genomic sequence for cod (www.ensembl.org). The mismatch was located at the 5'-end of the primer, however, and would be less severe for the amplification process compared to mismatches located in the 3'-end of the primers.

5.4.1.2. Previous evidence of non-neutrality

Studies of cod structure have often noted markedly high differentiation estimates at *Gmo*132 compared to other microsatellite loci (Bentzen *et al*. 1996; Ruzzante *et al*. 1998; 2000a,b; Knutsen *et al*. 2003; Karlsson and Mørk 2005). This led Karlsson and Mørk (2005) to investigate temporal stability at this locus and they found that allele frequencies varied significantly among years and cohorts, likely because of natural selection. Other studies have also suggested non-neutrality at *Gmo*132 (Nielsen *et al*. 2006; Skarstein *et al*. 2007; Westgaard and Fevolden 2007; Nielsen *et al*. 2009a), possibly caused by hitch-hiking selection, as sequencing efforts revealed no genes in the genomic region surrounding *Gmo*132 (Nielsen *et al*. 2009a). Non-neutrality has also been indicated for *Gmo*8 (Nielsen *et al*. 2006), *Gmo*37 (Skarstein *et al*. 2007), *Gmo*2 (Westgaard and Fevolden 2007) and *Gmo*34 (Nielsen *et al*. 2006, 2009a; Eiriksson and Árnason 2013). In contrast to *Gmo*132, *Gmo*2 has generally shown low or moderate levels of genetic differentiation among populations (Karlsson and Mørk 2005) and balancing selection has been hypothesised for this locus (Nielsen *et al*. 2009a). This agrees well with the fact that selection appears to act against dinucleotide mutations in coding regions (Chistiakov *et al*. 2006).

There appears to be a certain level of overlap of the microsatellite loci previously suggested to be selectively non-neutral and the loci found herein to be associated with genes, although the locus most often implied to be under natural selection (*Gmo*132) was not found to be located in close proximity to a known gene in the published cod genome. While a microsatellite locus in a genic, or even exonic, region is not synonymous with selection acting on that locus, selection can act on nongenic variants through linkage with adaptive mutations (hitch-hiking effect). In conclusion, although these relatively close relationships between cod microsatellites and genic regions were not expected, the observation could prove useful in interpreting some of the findings of previous population genetic studies on cod.

5.4.1.3. Depth or depth-related environmental conditions may drive selection at *Gmo*132

As discussed above, many studies have detected evidence of selection at the *Gmo*132 locus and although no environmental driver has yet been identified that may be responsible for these

observations (Nielsen *et al*. 2009a), depth has previously been investigated (Eiriksson and Árnason 2013). Results in the present study suggest depth may be a responsible environmental driver for allele frequency distributions at *Gmo*132. Frequency distributions for this locus deviated from the Gaussian-shaped distributions generally associated with microsatellite loci as one small-size allele (*Gmo*132^{117}) was found at relatively high frequencies in most samples and the allele was found at a significantly higher frequency in FPNE cod caught in deep waters compared to FPNE cod caught in shallow waters. When looking at the various geographic areas, this allele was only found at a relatively low frequency on the Faroe Bank (*c*. 3 %), intermediate on the Faroe Plateau (FPNE *c*. 14 %; FPW *c*. 10 %) and at high frequencies in Norwegian fjords (BGF *c*. 29 %; NWB *c*. 22 %).

Vertical distribution of cod in relation to genotypes has been described before (Case *et al*. 2005) and, as noted by those authors, it may not be depth *per se* that causes the deviating allele frequencies, but other related environmental variables. Temperatures on the Faroe Plateau may vary considerably with depth (Case *et al*. 2005) and Steingrund (2009) suggested that the mismatch between the vertical distribution of Faroe Plateau cod and their prey could be caused by the cod needing time to adapt to the changed environment (different prey, predators and temperatures). The *Gmo*132^{117} allele could potentially be a cold-water allele, reflecting temperature-dependent selection. This would agree with the low frequency of this allele on the Faroe Bank where water temperatures are *c*. 1.5 ºC higher than on the Plateau (Magnussen 2006). Interestingly, Dahle (1995), who investigated microsatellite variation in cod from the Faroe Bank, the North Sea, Lofoten and Norwegian fjords, found a small-size *Gmo*132 allele (*Gmo*132^{110}) with particularly high frequencies in the Lofoten (Arctic) and Norwegian fjord samples (*c*. 80 and 30 – 40 %, respectively) and low in the Faroe Bank sample (*c*. 2 %). This could potentially be the same allele as the *Gmo*132^{117} allele described in the present study (with allele frequencies of *c*. 3 % on the Faroe Bank and *c*. 20 – 30 % for the Norwegian fjord samples), since various laboratories may report different allele sizes for the exact same microsatellite alleles. Many factors can potentially contribute to size scoring differences. For example, the nature of the gel

matrix, usage of different fluorescent labels, migration deviations of size standards and even the housing temperature of the genetic analyser machine (Pasqualotto *et al*. 2007). Furthermore, these allele size discrepancies may also occur within a laboratory, if experimental protocols are changed (Presson *et al*. 2006).

Other studies have also found small-size alleles at the *Gmo*132 locus at particularly high frequencies. Bentzen *et al*. (1996) found significantly different allele frequency distributions across populations at the *Gmo*132 locus with two small alleles (*Gmo*132^{111} and *Gmo*132^{117}) dominating in the northern cod stock complex off Newfoundland (frequencies of 24 – 28 and 38 – 44 %, respectively). The respective frequencies of these allele in the Barents Sea and the Flemish Cap samples were 76 and 5 % and 55 and 19 %, possibly indicating that the *Gmo*132^{111} allele identified by Bentzen *et al*. (1996) may be the same as *Gmo*132^{110} allele identified by Dahle 1995) and the *Gmo*132^{117} allele identified in the present study. Nielsen *et al*. (2006) also reported skewed allele frequency distributions for *Gmo*132 with different small-size alleles dominating in different geographical samples. *Gmo*132^{115} was found at a frequency of nearly 70 % in cod from the Barents Sea, whereas the frequencies of this allele in the Baltic Sea, the North Sea and Newfoundland were *c*. 16, 8 and 14 %. In the Newfoundland population, *Gmo*132^{121} dominated (*c*. 42 %) and a small-size allele *Gmo*132^{113} was also found at a relatively high frequency (*c*. 23 %).

In a population genetic study of herring, Larsson *et al*. (2007) found particularly high F_{ST} values at one of the microsatellite loci, reflecting deviating frequencies at two of the alleles that appeared to confer a selective advantage under low salinity conditions. Similarly, findings in the current study could indicate that one specific small-size *Gmo*132 allele confers a selective advantage under certain conditions, although it is unlikely that selection acts directly on this locus, situated in an intergenic region. In terms of repeat numbers, *Gmo*132 is the longest microsatellite locus applied in this study, and long microsatellite alleles have been shown to lose repeat units at a faster rate than they acquire repeat units (Harr and Schlötterer 2000). Hence, the downward bias in *Gmo*132 allele frequencies that has been observed by a number of studies could be a result of

hitch-hiking selection on linked genes, as proposed by others (Nielsen *et al*. 2006), together with a neutral phenomenon that acts to reduce long microsatellite motifs.

5.4.1.4. Consequences for data interpretation

Although many of the microsatellite loci used in this investigation aligned to genic regions, there were some indications that this may not have had any serious implications for the overall interpretations of the results. Firstly, *Gmo*2, the only marker found in an exonic region, was not used in any structure analyses due to evidence of null alleles and another genic locus (*Gmo*19) was also excluded due to evidence of null alleles. Further, outlier analysis in Lositan did not detect evidence of selection on any loci and area differentiation was evident at all, not only the gene-associated, loci. Whereas *Gmo*132 stood out as an outlier locus in the depth contrast, in the geographic comparison this locus showed more moderate differentiation values and was only second highest with respect to the D_{Jost}-value.

It should be noted that some of the markers used in this study behave differently in other populations. Eiriksson and Árnason (2013) applied the same marker panel to investigate population structure of cod in Icelandic waters and significant genetic differentiation was only evident for *Gmo*34 and *Gmo*132, with outlier analysis detecting evidence for selection at *Gmo*34. Thus, *Gmo*34 behaved very differently than in the present study, where this locus showed the lowest D_{Jost} estimate of all the loci in the geographic comparison. These authors also investigated genetic differentiation for depth contrasts and detected significant differentiation for several loci (*Gmo*19, *Gmo*34, *Gmo*35, *Gmo*37 and *Gmo*132), although outlier analysis did not detect selection at any loci at the depth contrast. The fact that the same loci behave differently concerning depth in two geographic regions (Icelandic waters and the Faroe Plateau) may be a further indication that there are other depth-related conditions that affect allele frequencies at *Gmo*132.

5.4.2. Genetic divergence among cod from the various areas

5.4.2.1. Methods used to estimate the level and significance of genetic differentiation

Pairwise estimates of genetic differentiation were based on the traditional F_{ST} measure as well as an alternative estimate of genetic differentiation, D_{Jost_EST}, the latter metric considered to account better for the high variability and the nonlinear relationship between heterozygosity and diversity at microsatellite loci (Jost 2008). Significance was evaluated by confidence intervals pertaining to these two statistics and by a pseudo-exact test of genic differentiation. As might be expected based on comparisons of G_{ST} and D_{Jost} (Heller and Siegismund 2009; Mermains and Hedrick 2011), there was a strong positive relationship between F_{ST} and D_{Jost_EST} in this study, with higher differentiation estimates obtained from D_{Jost_EST}. The overall D_{Jost} was 0.0426, whereas the overall F_{ST} value was 0.0034. In the majority of the cases, the two statistics provided similar rankings of pairwise differentiation values, though a few discrepancies were evident when differences between values were relatively small. For example, both statistics agreed on a closer connectivity between Faroese and BGF cod than between Faroese and NWB cod, but where D_{Jost_EST} suggested a closer connectivity between Faroe Bank and BGF cod than between Faroe Plateau and BGF cod, the opposite was true for F_{ST}. Furthermore, both statistics agreed in the rankings of the three pairwise comparisons within the Faroese samples, but F_{ST} values were nearly similar, while D_{Jost_EST} values showed higher differentiation values for the Faroe Bank – Faroe Plateau comparisons than for the within-Plateau comparison. This may possibly reflect the better ability of D_{Jost} to detect differentiation when subpopulation heterozygosity is high (Jost 2008), even among closely related populations. Of the corresponding CIs, however, only the 95 % CI of F_{ST} for the FB – FPNE comparison was significant, while pseudo-exact tests of genic differentiation were significant for all Faroe Plateau – Faroe Bank comparisons. In concert, the methods used provide a useful way of interpreting the level and significance of genetic differentiation among populations.

5.4.2.2. Evidence of genetic divergence between Faroe Plateau and Faroe Bank cod

Overall, the various analyses provided weak evidence of potential genetic divergence between Faroe Plateau and Faroe Bank cod. To be able to detect robust patterns of genetic differentiation in marine fish, samples should be of sufficient sizes, replicated over years and adequate number and type of molecular markers should be applied (Waples *et al*. 2008). In this study, sample sizes of Faroese cod were relatively large and replicated over years (except for FPW cod that was only sampled in 2009), and temporal variation was found to be negligible. Suspicion of null alleles affecting allele frequency distributions at *Gmo*2 and *Gmo*19 resulted in the exclusion of these markers, rendering only eight loci for inferring population structure. However, the number of markers used in this study is similar to most population genetic studies of cod. For example, in a number of studies investigating fine-scale population structure of cod, the number of loci used varied between five and seven (Bentzen *et al*. 1996; Ruzzante *et al*. 1996, 2000a,b; Hutchinson *et al*. 2001, Skarstein *et al*. 2007). Furthermore, it appears that marker integrity may be equally important to the number of markers applied (Nielsen *et al*. 2009a).

Grossly estimated, both D_{Jost_EST} and F_{ST} proposed seven-fold higher differentiation between Faroese and Norwegian cod than between Faroe Plateau and Faroe Bank cod. D_{Jost_EST} between Faroe Plateau and Faroe Bank cod was *c*. five times higher than D_{Jost_EST} for within-Faroe Plateau or -Norwegian comparisons, whereas F_{ST} values barely discriminated between these comparisons. Despite the limitations of F_{ST} in microsatellite studies (Jost 2008), the inclusion of F_{ST} makes comparisons with other studies possible, since most microsatellite studies of Atlantic cod still use this statistic (for example Nielsen *et al*. 2009a; Eiriksson and Árnason 2013).

In the present study, pairwise F_{ST} values between Faroese and Norwegian samples ranged from 0.0061 to 0.0137 and the two pairwise F_{ST} values between Faroe Plateau and Faroe Bank cod were 0.0014 and 0.0018. This indicates a genetic divergence between Faroese and Norwegian coastal cod that is of similar magnitude to that found between North Sea and Norwegian coastal cod, where pairwise F_{ST} values calculated at five microsatellite loci ranged from 0.0043 to 0.0133 in several comparisons between sampling sites in the North Sea and a sample from Bergen in

Norway. The smaller level of genetic divergence between Faroe Plateau and Faroe Bank cod was similar in magnitude to some of the within-North Sea comparisons (found to be nonsignificant by the genic differentiation test; Hutchinson *et al*. 2001). However, the level of differentiation between Faroe Plateau and Faroe Bank cod was higher than that found between East Icelandic and Faroe Plateau cod samples, where F_{ST} values ranged from zero to 0.0010, and the majority of these were nonsignificant (Pampoulie *et al*. 2008a).

These comparisons with other studies do not reveal any startling observations; it would for instance seem intuitive that the level of differentiation between Faroese and Norwegian coastal cod was of similar magnitude to the one between North Sea and Norwegian coastal cod. In the present study, of the two F_{ST} values computed between Faroe Plateau and Faroe Bank cod, only the one for the FPNE - FB comparison was significant (F_{ST} = 0.0014). This value is lower than what has previously been reported in pairwise comparisons of Faroe Plateau and Faroe Bank cod at ten (including *Gmo*132) or nine (excluding *Gmo*132) microsatellite loci (F_{ST} = 0.0085 and F_{ST} = 0.0077, respectively), though the comparison excluding *Gmo*132 was nonsignificant following correction for multiple sampling (Nielsen *et al*. 2009a). The study by Nielsen *et al*. (2009a) investigated cod from the North Sea and adjacent areas and, in contrast to the study by Hutchinson *et al*. (2001), Nielsen *et al*. (2009a) found little support for subpopulation structuring in the North Sea region subscribing this to deviating frequency distributions at *Gmo*132. There is no obvious explanation for the large discrepancies regarding genetic differentiation among Faroe Plateau and Faroe Bank cod. It is unlikely that the inclusion or exclusion of *Gmo*132 is the main reason for this, as genetic differentiation, as detected in the present study, including *Gmo*132, was markedly lower than that detected by Nielsen *et al*. (2009a), even when they excluded *Gmo*132. One possible explanation could be the large sample sizes in the present study and the fact that they were sampled over a continuum of years, potentially buffering against chance events caused by sampling effect, although there were no indications of temporal variation among sampling years for neither the Faroe Plateau nor the Faroe Bank samples in this study. Another, or

additional, explanation could be that F_{ST} calculation based on highly variable markers may not be suitable for these kinds of comparisons (Jost 2008).

5.4.2.3. Ambiguous signals regarding the Faroe Plateau West sample

In agreement with previous mtDNA and microsatellite analysis (Sigurgíslason and Árnason 2003; Pampoulie *et al*. 2008a), neither F_{ST}, D_{Jost_EST} nor the genic differentiation test indicated any significant substructure on the Faroe Plateau. Nevertheless, investigations of phylogenetic relationships indicated statistically significant divergence between cod from the Faroe Plateau West and the remaining Faroe Plateau and the Faroe Bank cod. In the current study, Faroe Plateau cod had been assigned to the FPNE and FPW areas based on tagging studies that had found adult cod within each region to show spawning site fidelity (Joensen *et al*. 2005; Steingrund *et al*. 2005). Therefore, genetic divergence between the two Faroe Plateau areas would have a plausible biological explanation though counteracted by the clockwise current circulation homogenising eggs and larvae from different spawning grounds (Pampoulie *et al*. 2008a). A hint as to the ambiguous results for the FPW cod may be found by looking at the PCA (Figure 5.3), which provides a way to visualise the genetic variability of the data without making any strong assumptions (Jombart 2008), in contrast to for example *F* statistics. The PCA grouped the FPW cod together with the remaining FPNE cod, though especially close to the 2009 FPNE cod and the three temporal FPNE samples did not cluster as close together as the FB cod. Thus, the deviating behaviour of the FPW sample could potentially reflect small-scale temporal variation in allele frequencies on the Faroe Plateau and not spatial genetic differentiation.

Following the identification of close genetic connectivity between East-Icelandic and Faroe Plateau cod, Pampoulie *et al*. (2008a) discussed the possibility of a unidirectional transport of eggs and larvae from East Icelandic waters towards the Faroe Plateau, mediated by the East Icelandic Current and the east branch of the Irminger Current. However, the authors acknowledged that this is not supported by 0-group surveys conducted annually since 1974 on the Faroe Plateau. Of the two currents mentioned by Pampoulie *et al*. (2008a), the East Icelandic Current does travel relatively close to the Faroe Islands, but it is at 200 – 400 m depth when it

reaches the Faroe Plateau (Hansen 2000; Larsen *et al*. 2009). This is considerably deeper than the depths that eggs and larvae are normally found in (Brickman *et al*. 2007) and it is outside the front surrounding the Faroe Plateau at *c*. 100 m depth (Hansen 2000; Larsen *et al*. 2009), although mixing with Faroe Plateau Shelf water may occasionally happen (Larsen *et al*. 2008b). Thus, it seems unlikely that any free-floating stages of cod from the East Icelandic Current and the east branch of the Irminger Current would ever end up in the FPW area. Hence, the partly deviating FPW sample cannot be convincingly attributed to immigrant eggs and larvae coming in from East-Icelandic waters.

An alternative explanation of a possible inflow of genes to the Faroe Plateau West area is provided by investigations of pantophysin variation in Faroese and surrounding populations. Nielsen *et al*. (2007) found frequencies of the *Pan*I^{A} allele to be nearly 100 % on the Faroe Plateau (and the Faroe Bank). This was somewhat higher than the *Pan*I^{A} frequency of 88 %, identified by Case *et al*. (2005) for Faroe Plateau cod. As elaborated in Nielsen *et al*. (2007), these differences could be due to temporal fluctuations, but could also be due to hybridisation between Faroe Plateau cod and cod from the Faroe-Iceland Ridge, where the *Pan*I^{A} allele is found at a frequency of 79 % (Case *et al*. 2005). Although the latter could provide a plausible explanation for deviating allele frequencies in the Faroe Plateau West samples, it is not supported by tagging experiments with cod from both areas. Of *c*. 24,000 cod tagged on the Faroe Plateau during 1997 – 2006, *c*. 6,000 cod were recaptured and only one of these was caught on the Faroe-Iceland Ridge. Furthermore, in 2002, 168 individuals were tagged on the Faroe-Iceland Ridge. Of ten recaptured cod, none was caught on the Faroe Plateau, though two specimens had unknown recapture positions (ICES 2006).

In conclusion, the present data were too scarce to conclude either way regarding substructure on the Faroe Plateau. Furthermore, in the case that spawning site fidelity would cause any real substructure, it would likely be on a very small scale, further stressing the need for sampling temporal replicates (Waples 1998).

5.4.3. Genetic resilience towards detrimental effect of reduced population sizes

The level of genetic variability in the Faroese populations was at a similar level to the Norwegian populations. There was no convincing evidence of any bottleneck events for the investigated populations. The apparent genetic resilience towards the large reductions in census population sizes may reflect the fact that these are large marine populations, after all. Moreover, the high fecundity of the Faroe Bank cod, as indicated by the investigation of early life history strategies for Faroe Plateau and Faroe Bank cod (Chapter 3), could potentially help to reduce any detrimental effects of over-exploitation. However, the Faroe Bank population is currently at a historical low (ICES 2014a) and it might still be too early to detect the true consequences of this size reduction.

5.4.4. Factors shaping the current pattern of genetic divergence among populations

As discussed in Section 1.5.2, the level of genetic differentiation among populations may be controlled by various factors. For example, significant differentiation at microsatellite loci has been observed for most of the major spawning populations of cod in the North-West Atlantic, where the divergence between populations appears to be largely determined by oceanographic gyres that retain early life stages (Hauser and Carvalho 2008). Results from the current study, where there is evidence of weak genetic divergence between Faroe Plateau and Faroe Bank cod, are in agreement with these findings, since both the Faroe Plateau and the Faroe Bank ecosystems are characterised by strong gyres.

In the analysis of neutral genetic variation, gene flow will homogenise population structure whereas genetic drift has a diversifying effect. Gene flow in the present study was estimated from the genetic differentiation estimate F_{ST} and the number of migrants exchanged among populations per generation (Nm) were >> 1, indicating that gene flow can potentially have a markedly higher effect on population structure compared to random genetic drift (Lowe *et al*. 2004b). That said, and keeping in mind the limited number of adult cod likely migrating between the areas (Joensen *et al*. 2005; ICES 2006) as well as the gyres on the Faroe Plateau and the Faroe Bank, that act to keep early life stages of cod from drifting away (Hansen 2000), Nm estimates should be

interpreted with caution. Such estimates are based on a mathematical framework that makes many unrealistic assumptions. For example, selection at nearby loci may influence F_{ST} and *N*m estimates, but also background selection, i.e. the continuous selection against deleterious mutations throughout the genome, may substantially increase F_{ST}. Generally, mutation rates are not believed to bias genetic differentiation estimates, but for fast-mutating markers, such as microsatellites, mutation rates may cause biased estimates of *N*m. Furthermore, natural populations are unlikely to conform to the assumptions of similar sizes of populations, that are not spatially structured and that contribute equal number of individuals to the migrant pool (Whitlock and McCauley 1999). At least the last of these assumptions is not true for the Faroese stocks. Although exact stock size is difficult to estimate for the Faroe Bank cod, this stock is estimated to be only *c*. 10 % of the estimated stock size for the Faroe Plateau cod (Section 1.6.2.2).

In addition to the above, *N*m calculations also assume a state of equilibrium between migration and drift, that may for example be breached, if populations have experienced a recent range expansion. In fact, non-equilibrium has been suggested for Atlantic cod in the North Atlantic region, where many areas were uninhabitable during the last glacial maximum (LGM) 25,000 years ago. Following this period, Atlantic cod experienced a rapid range expansion into ice-free areas (Pogson *et al*. 1995; Pampoulie *et al*. 2008a). Despite the lower water level during the LGM (> 100 m; Hansen 2000; Huybrechts 2002), free movement of cod between the Faroe Plateau and the Faroe Bank during this period is unlikely, since they were still separated by the deep Faroe Bank Channel (current depth *c*. 800 m; Section 1.6.1). However, levels of migration may have been different from current levels. Moreover, the central Faroe Bank (at the top shallower than 100 m) was uninhabitable to cod and the distribution of cod on the Faroe Plateau was different from today, where cod usually spend their first 3 to 4 years in shallow waters (< 150 m; Steingrund and Ofstad 2010). Hence, both Faroe Plateau and Faroe Bank cod have experienced a certain change in their distributional range since the LGM. In that case, low F_{ST} values do not necessarily reflect high levels of contemporary gene flow. Time to equilibrium is influenced by both population sizes and migration rates and can take up to 10,000 generations if

population sizes are 1,000 and migration rates are small (< 0.01 %) and more, if population sizes are larger (Whitlock and McCauley 1999). For Faroe Plateau and Faroe Bank cod (average age-at-maturation *c*. 3 years; Section 1.6.2), 10,000 generations would be more than 30,000 years, as cod can spawn for several years; hence, dating further back than the LGM and proposed range expansion. Thus, the populations may not be in migration-drift equilibrium states and the results do not necessarily reflect a high contemporary gene flow.

In line with the above, Pampoulie *et al*. (2008a) suggested that the limited genetic differentiation between Faroe Plateau and East Icelandic cod may be due to a recent common origin of the stocks. Despite current findings of small, but significant, levels of genetic differentiation between Faroe Plateau and Faroe Bank cod, these authors did not detect any significant level of genetic differentiation between East Icelandic and Faroe Plateau cod. As discussed above, these findings are not in agreement with the absence of migration of adult individuals between these locations. However, if genetic differentiation among these populations at microsatellite loci is affected by historical processes, then the smaller level of differentiation between Faroe Plateau and East Icelandic cod may be explained by a relatively more accessible / shallower route (current depth < 500 m; Hansen 2000) compared to the route between the Faroe Plateau and the Faroe Bank.

5.4.5. Conclusion

Two of the three metrics computed from microsatellite data suggested significant genetic differentiation between Faroe Plateau and Faroe Bank cod samples. Although findings in Chapter 3 suggested significant functional differences between Faroe Plateau and Faroe Bank cod (a fuller discussion of overall functional and molecular differences is provided in Chapter 8), the microsatellite data only revealed a relatively modest degree of differentiation between cod from the two areas. The small level of genetic differentiation between Faroe Plateau and Faroe Bank cod (several times smaller than that between Faroese and Norwegian coastal cod) could be a consequence of relatively recent historical processes and contemporary levels of gene flow are suspected to be minimal, in accordance with tagging studies.

Depth or some depth-related variables may be responsible for previously implied non-neutrality at *Gmo*132 and a possible mechanism for the non-random distribution of *Gmo*132 genotypes was suggested.

The marked size reductions in census population sizes for both the Faroe Plateau and the Faroe Bank cod have apparently not resulted in severe reductions of detectable genetic diversity, probably because absolute numbers remain comparatively high. For the Faroe Bank cod, it can be hypothesised that any detrimental effect caused by population decreases may be partly counteracted by the relatively higher fecundity, compared to Faroe Plateau cod, as suggested elsewhere in this work (Chapter 3). However, it could still be too early to detect the full genetic consequences of reduced population size and all measures should be taken to increase the population to sustainable levels. This also applies to the Faroe Plateau population, which may actually warrant more protection than currently believed, considering the lower fecundity compared to the Faroe Bank population.

Chapter 6. Population genetic analyses of Faroe Plateau and Faroe Bank cod by SNP analysis of the previously implied hemoglobin and transferrin genes – a candidate gene approach

All other chapters of this thesis are solely my own work (though samples, biological information as well as data were kindly obtained from the Faroe Marine Research Institute as well as the Marine Research Institute, Bergen, as described in Sections 2.2 and 3.2). The data presented in this chapter, however, in part relate to work (hemoglobin and transferrin screenings) carried out by Øivind Andersen (Nofima, the Norwegian Institute of Food, Fisheries and Aquaculture Research) and colleagues. Of these, the transferrin results have been published (Andersen *et al*. 2011). Faroese cod samples for these screenings consisted of wild Faroe Plateau and Faroe Bank cod, collected as part of the current thesis work in 2008, as described in Section 2.2.1. My part in the collaboration concerned the Faroese samples as well as critical reading of the manuscript for the transferrin study (Andersen *et al*. 2011).

6.1. Introduction

6.1.1. Neutral versus non-neutral markers

In the last few decades, population genetic studies of natural populations have commonly applied neutral genetic markers, such as microsatellites (Morin *et al*. 2004). They are useful markers for analysing aspects of the historical demography and evolution of populations (Morin *et al*. 2004). Microsatellites have for example provided valuable insight into population structure at relatively small geographical scales (i.e. Bentzen *et al*. 1996; Hutchinson *et al*. 2001; Knutsen *et al*. 2003 for Atlantic cod). Variation at neutral marker loci is shaped by migration and genetic drift and is of limited use for the exploration of the adaptive potential of populations. Hence, in recent years, a number of studies have included non-neutral markers, i.e. markers that may be connected to the fitness of the organism and can tell something about the ability of populations to adapt to anticipated changes in the environment (Kirk and Freeland 2011). For example, in a study on population structure in European flounder (*Platichthys flesus* L.), Hemmer-Hansen *et al*. (2007) found that samples grouped according to geographical and historical proximity with regards to

microsatellite variation but according to environmental similarities with regards to variation at a candidate gene for adaptive variation (the heat-shock cognate protein, *Hsc70*). Together with other studies, this study highlighted the importance of including non-neutral markers, as adaptive divergence between populations can potentially be high, despite neutral marker loci showing only subtle genetic differentiation (Hemmer-Hansen *et al*. 2007; Larsen *et al*. 2008a; Pampoulie *et al*. 2011).

6.1.2. Identification of candidate genes for adaptive divergence

The study of candidate genes of adaptive importance is not a new concept, in fact one of the earliest electrophoretic studies of fish populations was that of Sick (1961) who described different variants of the oxygen-carrying protein hemoglobin in whiting and Atlantic cod (Sick 1961; Ward 2000). Such early electrophoretic studies detected variation based on charge differentiation in amino acid side chains and, thus, were only able to detect a small percentage of the genetic variability present. However, when direct analysis of the DNA sequence became a feasible option, the resolution of such studies improved, and even more so in recent years, since technological advances have facilitated the discovery of SNPs, the most common type of sequence variation found (Helyar *et al*. 2011). SNPs can readily be selected so that they are located in or in close proximity to particular gene regions (Helyar *et al*. 2011; Morin *et al*. 2004) and, consequently, there have been a number of cases where researchers have been able to identify a direct link between geno- and phenotypes (Kirk and Freeland 2011).

6.1.3. Fitness-related genes found for Atlantic cod

Three fitness-related genes have been quite extensively described for Atlantic cod: pantophysin (Johnston and Andersen 2008; Pampoulie *et al*. 2008a,b), the hemoglobin gene *Hb-*β_1 (Andersen *et al*. 2009) and the transferrin gene *Tf1* (Andersen *et al*. 2011), described in more detail in Section 1.3. All of these have previously been used to study genetic variation in Faroese cod.

The well-established link between genotype, (protein-) phenotype and environmental temperature (Section 1.3.2) makes the *Hb-*β_1 gene a promising candidate gene for investigating adaptive

variation in the Faroese cod populations. During the spawning season the temperature is 6 – 7 ºC on the Faroe Plateau (Steingrund *et al*. 2005) and 7.5 – 8º C on the Faroe Bank (Magnussen 2002, 2007). Hence, one could expect cod from the two areas to differ with respect to hemoglobin allele frequencies with the Faroe Bank cod harbouring relatively more of the warm-water allele, i.e. HbI-1, as investigated by protein electrophoresis, or the Met55 – Lys62 allele of the *Hb-ß*$_1$ gene (Section 1.3.2). Jamieson and Birley (1989) actually found higher HbI-1 frequencies on the Faroe Bank than on the Faroe Plateau, mean HbI-1 frequencies of 0.19 and 0.06, respectively, though they may have included some Scottish cod in their Faroe Bank dataset, that skewed the HbI-1 frequency upwards for this location (Magnussen 1996). In a more recent study, HbI-1 frequencies of 0.06 and 0.08 were observed for the Faroe Plateau and Faroe Bank cod, respectively, and although these frequencies were not statistically different (Magnussen 1996), the higher frequency of the HbI-1 allele on the Faroe Bank has been proposed to be a reflection of the warmer water temperature in this location (Magnussen 2006).

Concerning transferrin, Jamieson and Jones (1967) detected significant population differentiation between Faroe Plateau and Faroe Bank cod based on frequency analysis of five transferrin types.

Lastly, regarding variation at the pantophysin (*Pan*I) locus, the *Pan*I^A allele, associated with a higher growth rate, has been shown to dominate in both Faroe Plateau and Faroe Bank cod. Case *et al*. (2005) documented a frequency of the *Pan*I^A allele at the Faroe Plateau of 0.88 whereas Nielsen *et al*. (2007) found the frequency of the *Pan*I^A allele to vary from 0.95 – 1.0 in both historical and contemporary Faroe Plateau and Faroe Bank samples. Nielsen *et al*. (2007) suggested that the lower *Pan*I^A allele frequency found by Case *et al*. (2005) might be due to short-term variations in allele frequencies in response to temperature fluctuations or due to hybridisation with individuals from the Faroe – Iceland Ridge where the *Pan*I^B allele has been found at a frequency of 0.79 (Case *et al*. 2005). The latter would also explain the stray *Pan*I^B homozygotes found in the historical Faroese samples (Nielsen *et al*. 2007).

The aim of the present study was to explore if adaptive variation previously suggested by some among Faroe Plateau and Faroe Bank cod, as investigated by protein electrophoresis, could be confirmed by SNP analysis of the actual genes underlying the protein phenotypes. Hence, variation at the *Hb-β*$_1$ and *Tf1* genes was investigated. The *Pan*I locus was not included in this work, since there was no reason to believe that results would deviate from those previously found by using robust PCR-based techniques (Case *et al*. 2005; Nielsen *et al*. 2007). Data for six of the ten transferrin loci that are presented here have been published in Andersen *et al*. (2011).

6.2. Materials and methods

6.2.1. Sampling of biological material

Wild Faroe Plateau and Faroe Bank cod, sampled in 2008, were used in this study. Details concerning samples and sampling methods are described in more detail in Section 2.2. All 2008 specimens (Table 2.3) were included in the hemoglobin analysis (130 from the Faroe Plateau and 138 from the Faroe Bank) and a subset of these were used in the transferrin investigation (50 from each area).

6.2.2. DNA extraction

DNA was extracted with the Maxwell® Tissue DNA purification kit on a Maxwell® 16 automated platform (Section 2.3.1.1).

6.2.3. SNP genotyping

Samples were genotyped for SNPs at the two linked non-synonymous substitutions Met55Val and Lys62Ala of the *Hb-β*$_1$ gene, where the bases a/g in the first position of the Met (atg) or Val (gtg) codon as well as a/c in the second position of the Lys (aag) or Ala (gcg) codon were scored. Polymorphisms at these two positions result in two non-recombinant alleles, Met55 – Lys62 and Val55 – Ala62, which have been found to represent the formerly identified HbI-1 (warm-water) and HbI-2 (cold-water) protein alleles, respectively (Andersen *et al*. 2009; Section 1.3.2).

Transferrin variation was assessed by genotyping ten SNPs, spanning a region over 5,500 bp of the *Tf1* gene. Samples were genotyped by Nofima, the Norwegian Institute of Food, Fisheries and Aquaculture Research, on a MassARRAY system from Sequenom® (San Diego, USA). For further details, see Andersen *et al*. (2009) regarding analysis of the *Hb-ß*$_1$ gene and Andersen *et al*. (2011) regarding analysis of the *Tf1* gene.

6.2.4. Statistical analyses

6.2.4.1. Genetic diversity and assumption testing

GENEPOP version 4.2 (Raymond and Rousset 1995; Rousset 2008; available online at www.genepop.curtin.edu.au) was used to calculate basic diversity statistics and to test for deviations from Hardy-Weinberg equilibrium (HWE), using the probability test option, estimating *P*-values from a Markov chain (dememorisation 1000 and 100 batches of 1000 iterations).

6.2.4.2. Population differentiation

The level of population differentiation was estimated with F_{ST} and the significance of genic (allelic) differentiation was tested with a pseudo-exact probability test (dememorisation 1000 and 100 batches of 1000 iterations), both calculated in GENEPOP.

6.2.4.3. Relating *Hb-ß*$_1$ SNP data to former electrophoretic studies and to individual growth

Hb-ß$_1$ SNP alleles were converted to the formerly applied protein types (Section 1.3.2) and genotype frequencies based on these were calculated and compared to a previous investigation of Faroese cod (Magnussen 1996). A two-way ANOVA was used to investigate the effects of protein type and sex on body length. As cod from the Faroe Plateau and the Faroe Bank have different growth properties, this analysis was achieved for each population separately. Length could only be compared among individuals of the same age, thus the year group with the highest number of individuals for each combination of protein type and sex was chosen for this analysis. For both Faroe Plateau and Faroe Bank cod, this was four-year-old cod.

6.3. Results

6.3.1. Genetic diversity and assumption testing

Allele frequencies and genotypes for SNP loci in the *Hb-β*$_1$ and *Tf1* genes are presented in Table 6.1 and Table 6.2, respectively. At all loci, the same alleles were found to dominate in both Faroe Plateau and Faroe Bank cod. Observed heterozygosities at the *Hb-β*$_1$ loci were higher for the Faroe Plateau (H_O = 0.137) than the Faroe Bank cod (H_O = 0.092) whereas observed heterozygosities at the *Tf1* loci where higher for the Faroe Bank (H_O = 0.152) than the Faroe Plateau cod (H_O = 0.128). Furthermore, two of the *Tf1* loci were monomorphic for both populations and excluded from further analyses. None of the remaining 20 locus × location combinations deviated from HWE.

Table 6.1 Genetic variation of two *Hb-β*$_1$ SNPs in Faroe Plateau and Faroe Bank cod. Minor allele frequencies (MAF), genotype frequencies, observed (H_O) and expected (H_E) heterozygosities of the Met55Val and the Lys62Ala polymorphisms as well as H_O and H_E, calculated over loci.

Locus			**Faroe Plateau**	**Faroe Bank**
Met55Val	MAF		0.070	0.044
	Genotypes:	GG	105	124
		GA	17	12
		AA	0	0
	Ho / He		0.139/0.130	0.088/0.084
Lys62Ala	MAF		0.067	0.047
	Genotypes:	CC	103	124
		CA	16	13
		AA	0	0
	Ho / He		0.134/0.125	0.095/0.090
Over loci	Ho / He		0.137/0.128	0.092/0.087

Table 6.2. Genetic variation of ten *Tf1* SNPs in Faroe Plateau and Faroe Bank cod. Minor allele frequencies (MAF), genotype frequencies, observed (H_O) and expected (H_E) heterozygosities of SNPs as well as H_O and H_E, calculated over loci.

Locus			**Faroe Plateu**	**Faroe Bank**
tf-a	MAF		0.424	0.359
	Genotypes	AA	13	15
		AG	27	20
		GG	6	4
	H_O / H_E		0.589/0.488	0.513/0.460
tf-6	MAF		0.051	0.070
	Genotypes	GG	44	43
		GA	5	7
		AA	0	0
	H_O / H_E		0.102/0.907	0.140/0.130
tf-8	MAF		0.051	0.070
	Genotypes	AA	44	43
		AG	5	7
		GG	0	0
	H_O / H_E		0.102/0.907	0.140/0.130
tf-10	MAF		0.051	0.070
	Genotypes	AA	44	43
		AC	5	7
		CC	0	0
	H_O / H_E		0.102/0.907	0.140/0.130
tf-11	MAF		0.082	0.112
	Genotypes	TT	41	38
		TA	8	11
		AA	0	0
	H_O / H_E		0.163/0.150	0.224/0.199
tf-13	MAF		0.042	0.070
	Genotypes	GG	44	43
		GC	4	7
		CC	0	0
	H_O / H_E		0.083/0.080	0.140/0.130
tf-b	MAF		NA	NA
	Genotypes	GG	50	49
		GA	0	0
		AA	0	0
	H_O / H_E		0.000/0.000	0.000/0.000
tf-c	MAF		0.020	0.041
	Genotypes	CC	47	45
		CG	2	4
		GG	0	0
	H_O / H_E		0.041/0.040	0.082/0.078
tf-d	MAF		NA	NA
	Genotypes	CC	50	50
		CT	0	0
		TT	0	0
	H_O / H_E		0.000/0.000	0.000/0.000
tf-22	MAF		0.051	0.070
	Genotypes	TT	44	43
		TC	5	7
		CC	0	0
	H_O / H_E		0.102/0.097	0.140/0.130
Over all loci	H_O / H_E		0.128/0.115	0.152/0.139

6.3.2. Population structure

Based on variation at the *Hb-β*$_1$ and *Tf1* genes no significant genetic differentiation was found between Faroe Plateau and Faroe Bank cod samples. Levels of genetic differentiation were nonsignificant for all individual loci and genes as well as over all loci (Table 6.3).

Table 6.3. Overall and gene- and locus-specific global differentiation for the Faroe Plateau and Faroe Bank samples. F_{ST} values as well as and results from pseudo-exact tests of genic differentiation (*P*) are shown. *NS* indicates no significant differentiation.

Locus	F_{ST}	*P*
Met55Val	0.0025	*NS*
Lys62Ala	0.0000	*NS*
Ovel all *Hb β1*	0.0012	*NS*
tf-a	-0.0013	*NS*
tf-6	-0.0064	*NS*
tf-8	-0.0064	*NS*
tf-10	-0.0064	*NS*
tf-11	-0.0039	*NS*
tf-13	-0.0021	*NS*
tf-c	-0.0030	*NS*
tf-22	-0.0064	*NS*
Over all *Tf1*	-0.0038	*NS*
Over all loci	-0.0022	*NS*

6.3.3. Relating *Hb-β*$_1$ SNP data to former electrophoretic studies and to individual growth

In concordance with previous studies on hemoglobin variation in Faroe Plateau and Faroe Bank cod (Jamieson and Birley 1989; Magnussen 1996), the great majority of both Faroe Plateau and Faroe Bank cod were homozygous for the Val-Ala haplotype (Table 6.4). To use the former terminology, most individuals were of the protein phenotype HbI-2/2, that has been associated with colder waters. The estimated HbI-1 frequencies were 0.07 and 0.04 for the Faroe Plateau and the Faroe Bank cod.

No Met-Lys homozygotes were observed, but there were a few heterozygotes, where individuals were heterozygous at both SNPs. Since individuals were not sequenced, there is no way of telling

whether these were coupled or repulsive heterozygotes. Although we do not know to what extent genic recombination has occurred, we know that it has occurred, as evident from the two Val-Ala / Val-Lys individuals, one from the Faroe Plateau and one from the Faroe Bank. These observations are in agreement with observations from other cod populations (Andersen *et al*. 2009).

Table 6.4. Number of composite *Hb-β*$_1$ genotypes per population. In this analyses it was not possible to designate a precise genotype for individuals that were heterozygotic at both SNP positions.

Genotype	**Faroe Plateau**	**Faroe Bank**
Val-Ala/Val-Ala	103	123
Met-Lys/Met-Lys	0	0
Met-Lys/Val-Ala or Met-Ala/Val-Lys	15	11
Met-Lys/Met - Ala	0	0
Val-Ala/Val-Lys	1	1
Met-Lys/Val-Lys	0	0
Val-Ala/Met -Ala	0	0

Data on body length for combinations of protein type and sex are presented in Figure 6.1. These data confirm existing literature on different growth patterns for cod from the two areas, with higher growth rates for Faroe Bank cod (Section 1.6.2). Hence, Faroe Plateau and Faroe Bank cod were investigated separately for the potential effects of hemoglobin protein type and sex on body length of the fish. However, for neither Faroe Plateau nor Faroe Bank cod was there any significant effect of hemoglobin protein type, sex or a combination of these on body length, as investigated with two-way ANOVA statistical analyses (Table 6.5), although due to the small sample sizes, the power of the analyses were in all instances less than the desired 80 %.

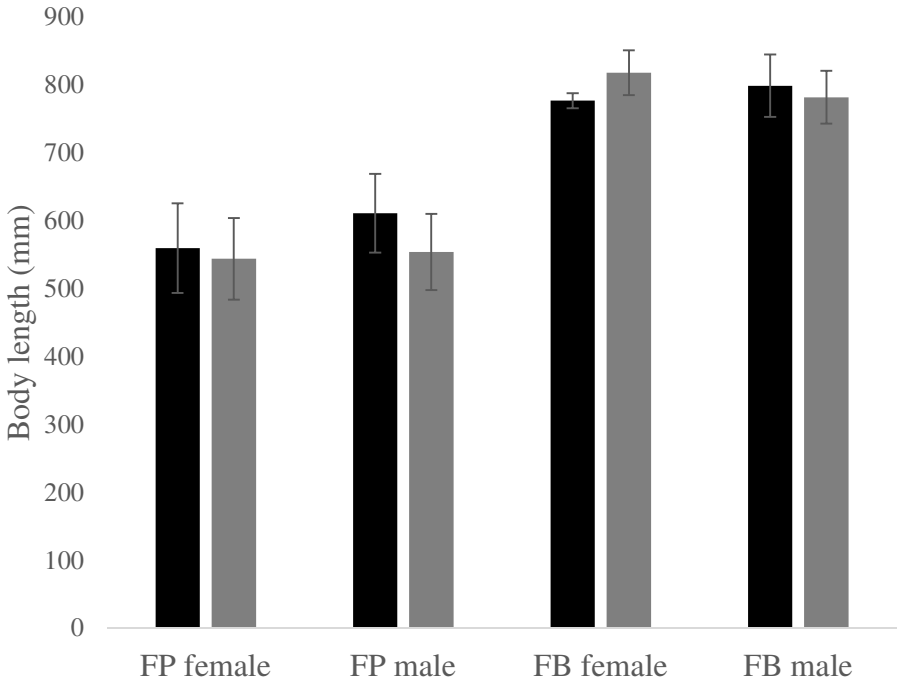

Figure 6.1. Average body lengths (± SD) of four-year-old male and female Faroe Plateau (FP; $N = 53$) and Faroe Bank cod (FB; $N = 55$) cod of different hemoglobin protein types. Black bars represent the HbI-1/2 and grey bars the HbI-2/2 protein type.

Table 6.5. Results of two-way ANOVA with interaction of mean body length of age four Faroe Plateau and Faroe Bank cod by hemoglobin protein type and sex.

	Faroe Plateau			**Faroe Bank**		
Source of variation	*df*	F	*P*	*df*	F	*P*
Protein type	1	2.489	*NS*	1	0.616	*NS*
Sex	1	1.697	*NS*	1	0.224	*NS*
Protein type × sex	1	0.755	*NS*	1	3.483	*NS*
Error	49			51		

6.4. Discussion

In the present study, no significant genetic differentiation was detected between Faroe Plateau and Faroe Bank cod at the hemoglobin gene *Hb-ß*$_1$ or the transferrin gene *Tf1*. This is in agreement with the findings of Magnussen (1996), as investigated by HbI-1 protein allele frequencies. Magnussen (1996) did observe marginally higher frequencies of the warm-water allele on the Faroe Bank (HbI-1 frequencies of 0.06 and 0.08 for the Faroe Plateau and Faroe Bank cod, respectively), whereas the opposite was true in the present study (estimated HbI-1 frequencies of

0.07 and 0.04 for the Faroe Plateau and the Faroe Bank cod, respectively). SNP analysis thus suggested slightly higher frequencies of the warm-water allele on the Faroe Plateau, but the difference was nonsignificant and likely a result of sampling error than anything else.

The transferrin data presented here were part of a larger trans-Atlantic study of cod populations (Andersen *et al*. 2011) where no significant genetic differentiation was detected between Faroe Plateau and Faroe Bank cod. Although four additional loci were included in the present study, one of them with a MAF of 0.02 in the Faroe Plateau cod and two of them monomorphic, this did not affect the results, as no single locus showed any differentiation. The transferrin results presented here as well as in Andersen *et al*. (2011) stand in contrast to the ones by Jamieson and Jones (1967) who investigated serum transferrin types and found different frequencies for Faroe Plateau and Faroe Bank cod. However, as suggested by Magnussen (1996), the results presented by Jamieson and Jones (obtained during spring time) could be biased, as serum sampled during the spawning period may show an elevated amount of electrophoretic bonds (unpublished experiments), disturbing the interpretation of the electrophoretic results (Magnussen 1996).

Information on local adaptation of fish stocks is critical for effective fishery management, as the ability of populations to persist future challenges is related to the adaptive variation present (Hilborn *et al*. 2003; Salmenkova 2011). Findings from the present study, where no divergence was seen between cod from the Faroe Plateau and the Faroe Bank, as investigated for two candidate genes for adaptive variation, do not exclude the potential existence of adaptive divergence between cod from the two areas. The results could merely demonstrate the potential stochastic outcome, when only a few candidate genes are investigated. Furthermore, the long-term survival of fish stocks is not only dependent upon conservation of genetic diversity at the DNA level, but also on preservation of genetic resources, i.e. DNA level diversity that is transformed into phenotypic differences in ecologically important traits (Hauser and Carvalho 2008). Such differences may be present in Faroe Plateau and Faroe Bank cod. Elsewhere in the present thesis, it was shown that Faroe Plateau and Faroe Bank cod differed concerning certain

early life history strategies, that these were likely genetic in origin and might have evolved in response to different environmental conditions in the two areas (Chapter 3).

To conclude, analysis of candidate genetic markers for adaptive divergence did not reveal any divergence between cod from the Faroe Plateau and the Faroe Bank. However, further investigations into this subject are still warranted. Technological advances, such as Next Generation Sequencing (NGS) (Mardis 2008a,b; Imelfort *et al*. 2009) have in recent years made it possible to conduct genome-wide scans in the investigation of adaptive divergence among populations (Bradbury *et al*. 2013, Hemmer-Hansen *et al*. 2013; Karlsen *et al*. 2013). Such techniques could be fruitful for future studies, which should seek to investigate more deeply if there is any adaptive divergence between Faroe Plateau and Faroe Bank cod and to identify potential ecological drivers for this.

Chapter 7. Restriction site associated DNA (RAD) sequencing: identification and analysis of novel SNP loci

7.1.1. Introduction

In the current work, various approaches have been explored to investigate potential genetic divergence between Faroe Plateau and Faroe Bank cod. An investigation of early life history strategies provided indirect evidence of genetic differentiation between the stocks (Chapter 3), microsatellite data revealed evidence of weak genetic differentiation between the stocks (Chapter 5) while an investigation of candidate hemoglobin and transferrin genes failed to detect significant genetic differentiation between the stocks (Chapter 6). In order to investigate in more depth the potential genetic divergence between the two stocks, the relatively recently developed RAD (restriction site associated DNA) technology was considered to be an encouraging approach. RAD technology is recognised as state of the art when it comes to population genomic studies of non-model species (ICES 2012).

RAD technology can be described as a genome "complexity reduction" protocol, as only a pre-defined proportion of the entire genome is investigated, namely short sequences (*c*. 50 – 300 bases) flanking the cut sites of the particular restriction enzyme selected for investigation. Generally, genomic DNA is digested with a rare (8 base) or infrequent (6 base) recognition restriction enzyme to generate a manageable number of DNA fragments. Following ligation of platform-specific NGS adapters the ends of each fragment can be sequenced at relatively high depth for a fraction of the effort required to sequence the entire genome. Barcoded adapters can be used to keep track of individual samples during RAD library construction allowing cost effective library construction and screening of multiple individuals to be undertaken (Baird *et al*. 2008). While the technology enables detection of both repetitive (short tandem repeat) and indel markers, most RAD sequencing projects focus on detection and screening of simple biallelic SNP polymorphisms for practical reasons (most common type of polymorphism and relative ease of bioinformatics processing).

As first demonstrated by Baird *et al*. (2008) who used RAD technology to map two ecologically important traits (lateral plate and pelvic structure phenotypes) in threespine stickleback, RAD sequencing is useful for linking genotype to phenotype, even in non-model species (species with no available reference genome). Recent studies have applied RAD technology to identify the mechanisms of sex determination and major sex determining regions in Atlantic halibut (*Hippoglossus hippoglossus*; Palaiokostas *et al*. 2013a) and Nile Tilapia (Palaiokostas *et al*. 2013b) and to describe a sex-linked SNP marker in the salmon louse (*Leleophtheirus salmonis*; Carmichael *et al*. 2013). Of importance to commercial breeding of Atlantic salmon, RAD technology has also been used for narrowing down the chromosomal region that contains a major IPN (Infectious Pancreatic Necrosis) resistance locus (Houston *et al*. 2012).

RAD sequencing is also a promising approach when it comes to the detection of potential genetic divergence between natural populations as well as identification of introgression between native and farmed populations (Hohenlohe *et al*. 2010, 2011, 2013). Hohenlohe *et al*. (2010) sampled 100 fish from three freshwater and two oceanic populations of threespine stickleback and following RAD sequencing, standard population genetics statistics were calculated using a sliding window approach, allowing the identification of genomic regions among populations that deviated from background values. No regions with deviating F_{ST} values were identified between the two oceanic populations (> 1000 km swim distance), whereas marked divergence was identified at nine genomic regions when oceanic and freshwater populations were compared (despite only 20 – 50 km distance between the freshwater and one of the oceanic populations). The genomic regions involved included the one identified in the study by Baird *et al*. (2008) to control lateral plate phenotype and suggested to be of adaptive importance for freshwater populations.

Fisheries for caviar-producing sturgeons of the genus *Acipenser* are amongst the most overexploited fisheries in the world and genomic resources for assigning fish products to their stock (or farm) of origin, as well as for conservation purposes, are urgently needed. Currently, *c*. 50 % of the caviar on the world market comes from farmed sturgeons. Nonetheless, in many

regions the fisheries continue at levels much higher than the set quotas, resulting in a large trade of mislabelled caviar products. DNA-based methods are already employed for species identification, but these methods (based on sequencing of the mitochondrial *cytochrome b* gene) do not provide the resolution necessary to suggest the stock of origin (Ogden *et al*. 2013). As part of an EU funded project (SturSNiP; https://stursnip.jrc.ec.europa.eu) to develop SNP resources for the management of sturgeon species, Ogden *et al*. (2013) utilised RAD sequencing to discover SNPs in four sturgeon species, the Russian *Acipenser gueldenstaedtii*, Persian *A*. *persicus*, Siberian *A*. *baerii* and Adriatic *A*. *naccarii*. Four population meta-samples as well as eight individual samples for verification were analysed using a single sequencing lane on an Illumina HiSeq platform and thousands of potentially informative markers were identified. Although the functional tetraploidy of these species posed great challenges to the bioinformatics treatment of the sequenced RAD tags, the results were encouraging. In addition to producing a large amount of data available for future studies of sturgeon population structure and product traceability (amongst others), this study found markers that could confidently distinguish between Russian and Persian sturgeon, two species that standard phylogenetic DNA markers have so far been unable to distinguish (Ogden *et al*. 2013).

The main objectives of this study were to investigate the utility of a genome complexity reduction technique to develop SNPs capable of clarifying the potential genetic divergence between Faroe Plateau and Faroe Bank cod. When many thousands of SNP loci are identified from relatively small sample numbers, there is a statistical probability that at least some of the SNPs discovered may not be true SNPs, but artefacts of the sequencing process (ICES 2012). If SNPs discovered by RAD technology are to be used as practical markers in population genetic studies of cod, they need to i) be verified as true SNPs and ii) be assayable by a methodology more cost and time effective compared to RAD sequencing. Furthermore, when loci are identified from small (and potentially unrepresentative) samples, ascertainment bias is also a potential source of error, i.e. allele frequencies detected in samples of limited size may not be representative of true population allele frequencies. In addition, geographic bias may be present, so that SNPs, in this case

identified in Faroese cod, may have markedly different polymorphic content in other cod populations (Bradbury *et al*. 2011; Helyar *et al*. 2011; Seeb *et al*. 2011). Thus, following the main RAD analysis, where SNPs were discovered and assessed for discriminatory potential, further validation screening of a subset of identified SNPs, found to show a significant level of divergence between Faroe Plateau and Faroe Bank cod, was undertaken. These SNPs were genotyped in larger samples (that included outlier populations) using an alternative and more practical approach for aquaculture monitoring and fisheries management (KASP genotyping platform).

7.2. SNP detection in Faroese cod populations using RAD sequencing

7.2.1. Objectives

The aim was to use RAD sequencing methodology to detect and genotype a large number of novel SNP markers from a small panel of Faroe Plateau and Faroe Bank samples.

7.2.2. Materials and methods

7.2.2.1. Sampling of biological material

Sampling procedures are described in Section 2.2. The Faroe Plateau and Faroe Bank samples used in this study were collected in 2009 and 2010 and mature individuals were chosen (based on age and maturity stages; Section 2.2.1.1), equal numbers of each sex. All Faroe Plateau individuals were from the North-East area collected during the breeding season and, thus, should represent the main spawning ground north of the Faroe Islands. Biological information for these samples is provided in Table 7.1.

Table 7.1. Sample sizes (*N*), average size (total body length and weight), condition (Fulton´s condition factor, *K*), age, sex ratio (% F: percentage females), average maturity stage (± SD) and proportion of actively spawning individuals (% S; maturity stage 6) for the Faroe Plateau North-East (FPNE) and Faroe Bank (FB) samples.

Area	Year	*N*	Body length (cm)	Body Weight (g)	Fulton's *K*	Age (years)	% F	Maturity stage	% S
FPNE	2009	10	59.9 (±10.9)	2663 (±1489)	1.121 (±0.071)	4.6 (±1.1)	50	4.7 (±1.6)	40
	2010	10	64.4 (±6.5)	3006 (±930)	1.093 (±0.068)	4.4 (±1.3)	50	5.7 (±0.7)	50
FB	2009	10	88.6 (±6.2)	7850 (±1623)	1.117 (±0.096)	5.4 (±0.8)	50	5.6 (±0.5)	60
	2010	10	91.2 (±6.9)	8717 (±1906)	1.141 (±0.138)	5.6 (±1.3)	50	5.6 (±0.8)	50

7.2.2.2. DNA extraction

DNA was extracted with the Real Pure Genomic DNA extraction kit (Durviz, Valencia, Spain), as described in Section 2.3.1.2. The extraction procedure involved an RNase treatment step to remove residual RNA. The quality and integrity of the DNA was checked, as described in Section 2.3.2.2 and Section 2.3.2.3, and samples were diluted to *c*. 45 ng/µL in 5 mM Tris, pH 8.5.

7.2.2.3. Library construction and sequencing

Library construction is described in detail in Section 2.3.4. Based on previous RAD studies conducted on other fish species, a "rare cutting" restriction enzyme (*Sbf*I – CC_TGCA^GG) was employed to minimise the number of RAD loci generated (*c*. 20,000 – 50,000 loci), allowing a single lane of Illumina HiSeq sequencing (*c*. 150 million reads) to generate a cost-effective, robust dataset (coverage *c*. 30 – 50 ×) for 40 samples.

7.2.2.4. Identification and genotyping of RAD loci

Sequence reads of low quality (score under 30, while the average quality score was 38), missing the restriction site or with ambiguous barcodes were discarded. Retained reads were sorted into loci and genotyped using Stacks software 1.19 (Catchen *et al*. 2011, 2013). The likelihood-based SNP calling algorithm (Hohenlohe *et al*. 2010), implemented in Stacks evaluates each nucleotide

position in every RAD-tag of all individuals, thereby statistically differentiating true SNPs from sequencing errors. The parameters were a minimum stack depth of at least 30, a maximum of 2 mismatches allowed in a locus in an individual and up to 1 mismatch between alleles.

7.2.2.5. Genetic diversity and Hardy-Weinberg equilibrium

Genetic diversity estimates were obtained from GenAlEx version 6.501 (Peakall & Smouse 2006) and GENEPOP version 4.3 (Raymond and Rousset 1995; Rousset 2008). GENEPOP is available as a web application, but for this large dataset, the software was run locally on a PC. Deviations from Hardy-Weinberg equilibrium (HWE) were tested with GENEPOP, using the probability test options and estimating *P*-values from a Markov chain (dememorisation 1000 and 100 batches of 1000 iterations).

7.2.2.6. Identification of outlier loci

The data were tested for evidence of selection in Arlequin version 3.5.1.2 (Excoffier and Lischer 2010). Assuming the finite island model proposed by Beaumont and Nichols (1996), 50,000 coalescent simulations (100 demes per group and expected heterozygosity set from zero to 0.5) were performed on each of the four samples to generate a null distribution of F_{ST} values versus heterozygosity. As this model may potentially result in false positives (Excoffier *et al*. 2009), a conservative approach was taken and only loci above the upper 1 % quantile were considered outlier loci, showing evidence of positive selection.

7.2.2.7. Population structure

Each of the areas investigated (Faroe Plateau and Faroe Bank) had been sampled for two consecutive years; hence, levels of genetic differentiation were first assessed between temporal samples from these geographic regions. Samples were pooled, if no significant differentiation was detected. Then, genetic differentiation among areas was assessed. The degree of genetic differentiation was quantified by F_{ST} and the significance of genic (allelic) differentiation was tested by pseudo-exact probability tests, dememorisation 1000 and 1000 batches of 1000 iterations, both calculated in GENEPOP.

7.2.2.8. Screening for association between SNP loci and phenotypic sex

Potential association between SNP loci and phenotypic sex was explored. Individuals were grouped according to phenotypic sex (Section 2.2.1.1) and the degree of genetic differentiation between the sexes (19 females and 18 males) was quantified by F_{ST} and the significance of genic (allelic) differentiation, tested by pseudo-exact probability tests, dememorisation 1000 and 1000 batches of 1000 iterations, both calculated in GENEPOP. Loci with significant outcomes ($P \leq 0.05$) were evaluated by assignment tests in GENECLASS2 (Piry *et al*. 2004). Individuals were re-allocated to known phenotypic sex, using the leave-one-out procedure, estimated by the Bayesian method described by Rannala and Mountain (1997) and 10,000 resamplings of individuals, as described in Paetkau *et al*. (2004). Type 1 error was set to 0.05.

When relevant, genomic coordinates of SNPs were found by using the respective RAD tag sequences as queries in Ensembl BLAT searches against the cod genome (gadMor1, version 75.1). The Variant Effect Predictor (VEP) Web tool (www.ensembl.org; McLaren *et al*. 2010) was used to determine the effects of the SNPs, i.e. their genomic location and potential consequences, such as synonymous or missense SNP. The potential biological function of genes was inferred from Gene Ontology (GO) annotations (Ashburner *et al*. 2000) in Ensembl, all inferred by electronic annotation.

7.2.2.9. Selection of a set of informative SNP loci for validation in larger sample sizes

Based on pairwise F_{ST} values between the Faroe Plateau and Faroe Bank samples, calculated in Stacks, a set of informative SNP loci (those with the highest corrected F_{ST} values and with Fisher´s $P < 0.0005$) were selected using R / Adegenet (www.R-project.org; Jombart 2008) to evaluate the degree to which they would differentiate between larger samples of Faroe Plateau and Faroe Bank cod.

7.2.3. Results

7.2.3.1. Identification of RAD loci

In total, 311,037,024 raw reads (101 bases long) were produced (155,518,512 paired-end reads). After removing low quality sequences (quality score under 30), ambiguous barcodes and orphaned paired-end reads, 72.2 % of the raw reads were retained (224,608,697 reads). The Stacks package was then used to make the assembly of the sampled loci from each individual: 34,312 unique RAD-tags were retrieved. In order to maximise the number of informative markers and minimise the amount of missing or erroneous data, RAD-tags retrieved in at least 75 % of the samples and harbouring one or two SNPs were used. Out of 13,582 shared RAD-tags (loci), 3,386 diallelic markers were identified, that contained a total of 3533 SNPs.

7.2.3.2. Genetic diversity and Hardy-Weinberg equilibrium

Genetic diversity estimates for the 3533 identified SNPs in the four samples are shown in Table 7.2. Percent polymorphic loci per sample ranged from 44.52 % (Faroe Plateau 2009) to 50.21 % (Faroe Plateau 2010) with 100 % of the loci polymorphic across samples. Many of the loci displayed a low level of polymorphism with minor allele frequencies (MAF) < 0.05. Hence, the Hardy-Weinberg equilibrium was tested for 3075 locus × sample combinations and 110 significant deviations were found, though none of these were significant after sequential Bonferroni correction (first critical $\alpha = 0.05 / 714$ loci $= 0.0001$ for the Faroe Plateau 2009 sample with 714 polymorphic loci [MAF > 0.05]).

As calculated over all SNPs, average number of alleles, effective number of alleles as well as gene diversities (expected heterozygosities) were comparable among the four samples (Table 7.2). Average number of alleles ranged from 1.445 (Faroe Plateau 2009) to 1.502 (Faroe Plateau 2010), average effective number of alleles from 1.155 (Faroe Bank 2010) to 1.162 (Faroe Plateau 2010) and average gene diversities from 0.104 (Faroe Bank 2010) to 0.108 (Faroe Plateau 2010).

Minor allele frequencies across samples ranged from 0.014 to 0.500 (Figure 7.1).

Table 7.2. Genetic variation of the 3533 SNPs in the four samples. Sample size (*N*), mean number of individuals genotyped per locus (*n*), percent polymorphic loci (%P), mean number of alleles (N_a), mean effective number of alleles (N_e), mean observed (H_O) and expected (H_E) heterozygosities and inbreeding coefficient (F_{IS}).

	Faroe Plateau		**Faroe Bank**	
	2009	**2010**	**2009**	**2010**
N	9	9	9	10
n	7.244	8.508	7.857	8.789
%P	44.52	50.21	47.78	49.53
N_a	1.445	1.502	1.478	1.495
N_e	1.160	1.162	1.156	1.155
H_O	0.117	0.121	0.116	0.115
H_E	0.105	0.108	1.105	1.104
F_{IS}	-0.038	-0.059	-0.044	-0.054

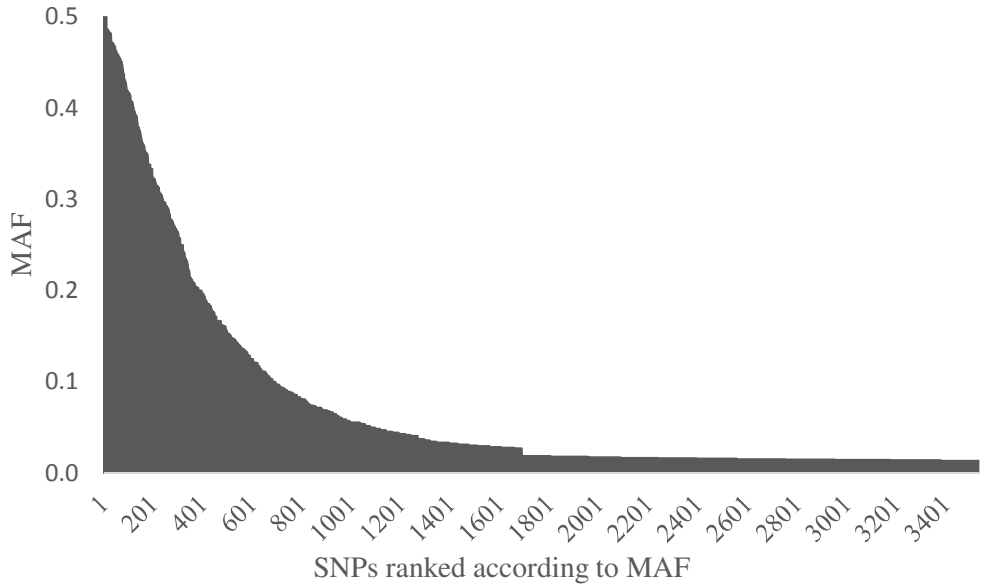

Figure 7.1. Minor allele frequencies (MAF) of SNPs across Faroe Plateau and Faroe Bank samples.

7.2.3.3. Identification of outlier loci

By F_{ST} outlier analysis 37 (1.05 %) of the loci were found to have F_{ST} values above the 1 % quantile (Figure 7.2); hence, more likely to be subject to positive selection.

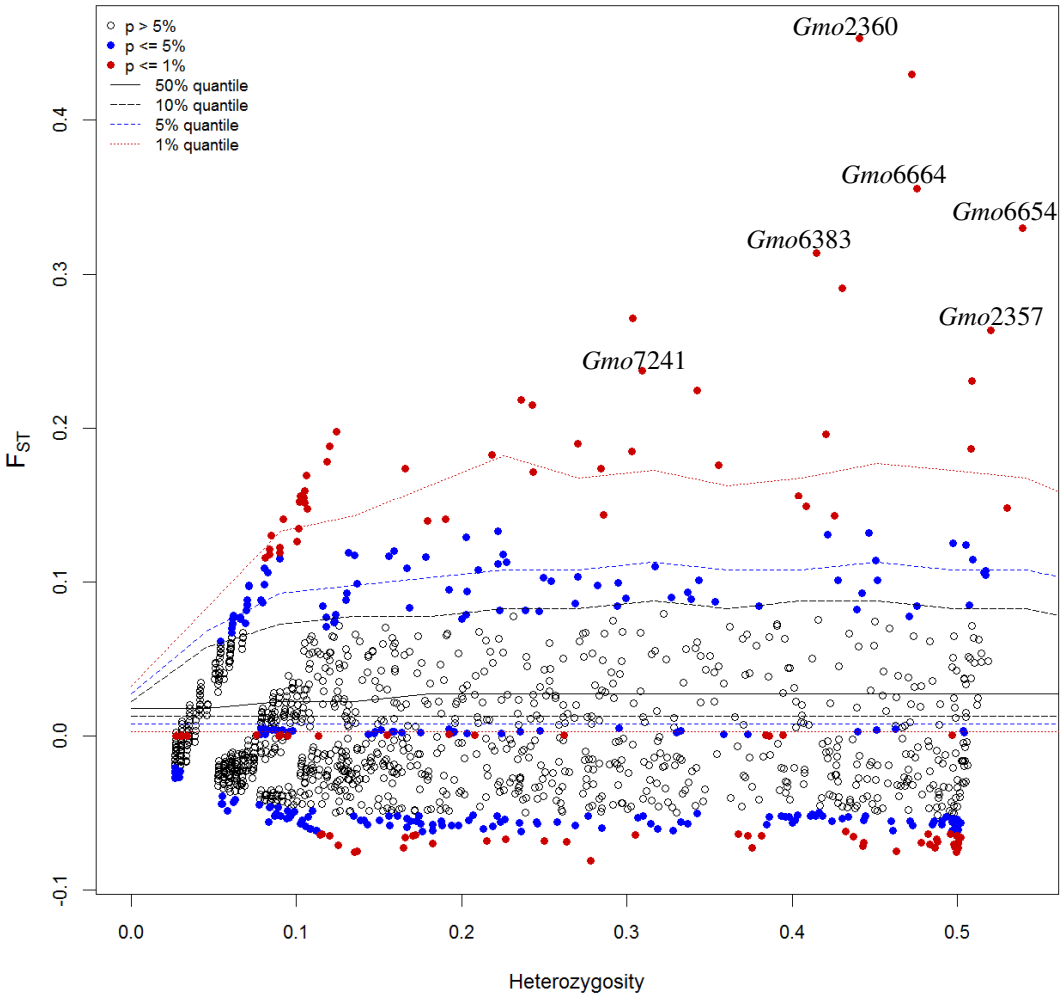

Figure 7.2. F_{ST} outlier analysis for the 3533 SNPs. F_{ST} values as a function of heterozygosity were compared to their neutral distribution under a finite island model. Each dot represents a locus. Loci outside the upper 1 % quantile were considered outlier loci subject to positive selection and loci outside the lower 1 % quantile were considered outlier loci subject to stabilising selection. Names are provided for the six informative SNPs that were validated in a larger dataset (Section 7.3), for functional annotation of these SNPs, please refer to Table 7.11.

7.2.3.4. Population structure

The validity of pooling sampling years was evaluated by pairwise F_{ST} values as well as by pseudo-exact tests of genic differentiation. F_{ST} values were small and the genic differentiation test was non-significant in all cases (boxed portions in Table 7.3). Thus, an absence of short-term temporal instability in Faroe Plateau and Faroe Bank allele frequencies was assumed and temporal samples were pooled for further investigation of area differentiation.

Table 7.3. Matrices of pairwise F_{ST} (below diagonal) among areas and / or years sampled and results of pseudo-exact tests of genic differentiation (above diagonal). Bold value indicates significant result after sequential Bonferroni correction for multiple tests (first critical $\alpha = 0.05 / 6$ tests = 0.008).

	Faroe Plateau		**Faroe Bank**	
	2009	2010	2009	2010
Faroe Plateau				
2009		*NS*	*NS*	**< 0.0001**
2010	0.0003		*NS*	*NS*
Faroe Bank				
2009	0.0096	0.0033		*NS*
2010	0.0101	0.0084	0.0009	

Significant genetic differentiation was found between the Faroe Plateau and the Faroe Bank areas (F_{ST} = 0.0074; $P < 0.0001$). No single diagnostic loci were identified, i.e. no loci were fixed for different alleles between Faroe Plateau and Faroe Bank cod.

As the main object of this study was to investigate the potential of markers to distinguish between Faroe Plateau and Faroe Bank cod, the discriminating power of individual loci was also explored: individual loci showed varying level of differentiation between the two areas (F_{ST} ranging from zero to 0.4986). While the great majority of loci showed no evidence of differentiation, significant genic differentiation was indicated at 58 (1.64 %) of the loci (P values < 0.05). Amongst these 58 loci were 17 of the 37 loci suggested to be subject to positive selection (Section 7.2.3.3). Average F_{ST} values between the two areas were higher, though with a greater spread, as estimated for the 17 putatively selected loci compared to the remaining 41 putatively neutral loci (mean F_{ST} values of 0.2902 [± 0.1271] and 0.1376 [± 0.0322], respectively).

7.2.3.5. Screening for association between SNP loci and phenotypic sex

There were no diagnostic SNPs, i.e. SNPs that were fixed for different alleles between female and male samples. RAD loci were further explored for potential association with phenotypic sex by consulting F_{ST} values and results from pseudo-exact tests of genic differentiation between females and males. Twenty-five SNPs were significant at the 5 % probability level (Table 7.4). Two pairs of SNPs originated from the same RAD tag and yielded redundant information (100 %

correlation between genotypes). Hence, one SNP from each of the two pairs (marked with B's in Table 7.4) were excluded in assignment tests which were conducted with 23 SNP loci. Assignment tests of female and male individuals via GENECLASS2 version 2.0 showed that 89.2 % correct assignments were obtained based on these data (quality index = 57.09).

Table 7.4. Best 25 informative SNP loci for phenotypic sex (as assessed by pseudo-exact tests of genic differentiation between male and female samples). Internal SNP ID, mapping on the cod genome draft (in Ensembl), F_{ST} values between female and male samples and Fisher´s *P*-values. SNPs from the same RAD tags are distinguished by A's and B's.

SNP ID	**Chromosome**	**Position**	F_{ST}	**Fisher´s *P***
*Gmo*995A	GeneScaffold_1287	201367	0.1189	0.0498
*Gmo*995 B	GeneScaffold_1287	201376	0.1189	0.0449
*Gmo*1259	GeneScaffold_1347	26773	0.1254	0.0225
*Gmo*1457	GeneScaffold_1419	14274	0.1201	0.0278
*Gmo*4662	GeneScaffold_2206	100916	0.2112	0.0039
*Gmo*5477	GeneScaffold_2466	20221	0.1400	0.0318
*Gmo*5618	GeneScaffold_2498	31727	0.1252	0.0246
*Gmo*6824	GeneScaffold_2877	584291	0.1378	0.0321
*Gmo*7028	GeneScaffold_2936	79281	0.1243	0.0390
*Gmo*7515	GeneScaffold_308	11831	0.1204	0.0373
*Gmo*9867	GeneScaffold_3761	190588	0.0995	0.0473
*Gmo*10002	GeneScaffold_3789	19654	0.0929	0.0458
*Gmo*10182	GeneScaffold_3838	110695	0.1195	0.0270
*Gmo*10295	GeneScaffold_3883	305683	0.1161	0.0473
*Gmo*11037	GeneScaffold_4164	10925	0.1563	0.0124
*Gmo*12470	GeneScaffold_4590	131866	0.1895	0.0046
*Gmo*12532	GeneScaffold_4599	32255	0.1133	0.0205
*Gmo*12873	GeneScaffold_478	56502	0.1109	0.0289
*Gmo*13428	GeneScaffold_648	92740	0.1067	0.0467
*Gmo*15503	Contig312233	214	0.1298	0.0256
*Gmo*15521	Contig 315598	623	0.2408	0.0011
*Gmo*17077	Contig99732	408	0.1424	0.0261
*Gmo*17530 A	Scaffold_07391	2694	0.1596	0.0209
*Gmo*17530 B	Scaffold_07391	2695	0.1596	0.0193
*Gmo*18260	GeneScaffold_2213	67321	0.1453	0.0261

The data were further explored to see whether similar results could be obtained using fewer SNPs. Consulting Table 7.4, a subset of six SNPs with top high F_{ST} values (> 0.15) were identified (*Gmo*4662, *Gmo*11037, *Gmo*12470, *Gmo*15521, *Gmo*17530 A and *Gmo*17530 B). Consulting

raw allele frequency data, four of these SNPs were found to harbour female-specific alleles, though many females were also homozygotic for the shared allele (Table 7.5). For two of the markers (*Gmo*11037 and *Gmo*15521), however, the genomic sequence of cod (www.ensembl.org), which represents a male specimen (Star *et al*. 2011), was found to display the "female-specific" alleles. Nonetheless, using only these four SNPs slightly enhanced assignment results, as 94.6 % of the individuals were assigned to the correct sex (with a marginally higher quality index of 61.36).

Table 7.5. Female and male genotypes at four SNP loci suggested to be associated with phenotypic sex.

SNP ID			**Female (N = 19)**	**Male (N = 18)**
***Gmo*4662**	Genotypes:	CC	9	14
		CG	9	0
		GG	0	0
				-
***Gmo*11037**	Genotypes:	GG	12	14
		GA	1	0
		AA	3	0
***Gmo*12470**	Genotypes:	TT	11	18
		TC	8	0
		CC	0	0
***Gmo*15521**	Genotypes:	GG	9	14
		GT	9	0
		TT	1	0

Significant BLAT hits (*e*-values $\leq 2.20 \times 10^{-47}$) were found for the RAD-tag sequences harbouring these four SNPs. The length of all query sequences was 96 bp (RAD-tag sequences) and all of these sequences aligned over the entire length of the sequence, with 100 % sequence identity. Two SNPs were located in intergenic regions (*Gmo*4662 and *Gmo*15521), while two were found in coding regions, causing synonymous mutations (*Gmo*11037 and *Gmo*12470). In addition, *Gmo*11037 was classified as an "upstream variant", located *c*. 3000 bp upstream of the gene *dpy30* (Table 7.6). This gene is homologous to the *dpy-30* gene that is involved in dosage

compensation in the worm *Caenorhabditis elagans* (Hsu and Meyer 1994; Petty *et al*. 2011). Dosage compensation refers to regulatory mechanisms that acts to equalise expression levels of genes linked to sex chromosomes between the sexes (Lucchesi *et al*. 2005).

Table 7.6. Annotation of SNP loci based on the Variant Effect Predictor tool and Gene Ontology information from Ensembl. All information on biological function was inferred from electronic annotation in Ensembl. For genomic location of SNPs, please refer to Table 7.4.

SNP ID	Alleles	Ensembl VEP Consequence	Gene Ensembl ID	Symbol	Biological process	Molecular function
*Gmo*4662	G/C	Intergenic				
*Gmo*11037	A/G	Synonymous SNP and < 3200 bp 5' of a second gene:	ENSGMOG00000011826 ENSGMOG00000011824	 dpy30		
*Gmo*12470	C/T	Synonymous SNP	ENSGMOG00000013020	lipea	Metabolic process, cholesterol metabolic process, lipid catabolic process.	Lipase activity, hydrolase activity.
*Gmo*15521	G/T	Intergenic				

7.2.3.6. Selection of a set of informative SNP loci for validation in larger sample sizes

The 3.386 diallelic RAD loci were evaluated for pairwise F_{ST} values between the Faroe Plateau and the Faroe Bank samples and the best eight informative SNP loci (those with the highest corrected F_{ST} values for pairwise comparisons of Faroe Plateau and Faroe Bank population samples and with Fisher's $P < 0.0005$) were identified (Table 7.7). All of these eight SNP loci were among the 37 loci for which outlier analysis suggested positive selection (Section 7.2.3.3). Analyses of the informative SNP panel selected from the RAD sequencing data are provided in Section 7.3.

Table 7.7. Best eight informative SNP loci for discriminating between Faroe Plateau and Faroe Bank cod. Internal ID, mapping on the cod genome draft (in Ensembl), F_{ST} values between Faroe Plateau and Faroe Bank population samples and Fisher´s *P*-values.

SNP ID	Chromosome	Position	Corrected F_{ST}	Fisher´s *P*
*Gmo*2357	GeneScaffold_1671	224530	0.260	8.35×10^{-4}
*Gmo*2360	GeneScaffold_1671	330024	0.370	4.11×10^{-6}
*Gmo*6383	GeneScaffold_2725	76359	0.253	6.27×10^{-6}
*Gmo*6644	GeneScaffold_2818	170306	0.354	2.37×10^{-6}
*Gmo*6654	GeneScaffold_2819	16133	0.258	4.95×10^{-5}
*Gmo*6664	GeneScaffold_2821	233256	0.287	3.15×10^{-4}
*Gmo*7241	GeneScaffold_2996	134658	0.267	4.12×10^{-4}
*Gmo*14283	GeneScaffold_995	105988	0.257	1.36×10^{-5}

7.3. Targeted SNP screening of Faroe Plateau and Faroe Bank cod using discriminatory SNPs identified by RAD sequencing

7.3.1. Objectives

In this study SNPs, identified via RAD sequencing to be among the most informative ones concerning genetic differentiation between Faroe Plateau and Faroe Bank samples, were validated in independent Faroe Plateau and Faroe Bank samples (i.e. samples not used for initial SNP discovery). Analyses were concerned with exploring i) the extent to which SNPs, identified by the RAD sequencing approach, could be analysed by a simpler, more cost-effective PCR-based

method; ii) concordance between genotypes of individuals screened by both technologies (i.e. RAD sequencing and KASP assay); iii) the genomic context of these informative SNPs; iv) the utility of the SNPs for analyses of population structure (geographic outlier samples included); and lastly, v) the utility of the SNPs for assigning individuals and samples of individuals back to their natal origin.

7.3.2. Materials and methods

7.3.2.1. Samples

Sampling of wild Faroese cod is described in Section 2.2.1. In total, 118 Faroese individuals were analysed by the KASP assay. Six of these individuals had been analysed in the RAD sequencing study (Section 7.2) and were included in order to check for concordance in genotypes obtained by the two methods. Excluding these control individuals, the KASP assayed data consisted of 112 "novel" Faroese individuals (FPNE09, $N = 14$; FPNE10, $N = 26$; FPW; $N = 20$; FB09; $N = 25$; FB10; $N = 27$). In addition, Norwegian coastal cod from the Borgundfjord and Verrabotn, collected in 2004 and 2005 (BGF04, $N = 20$; VB05, $N = 20$) and cod from the White Sea, collected in 2003 (WS03; $N = 20$) were included as outlier samples. Biological details for the Faroese samples (excluding the six controls) is provided in Table 7.8.

Table 7.8. Sample sizes (*N*), average size (total body length and weight), condition (Fulton´s condition factor, *K*), age, sex ratio (% F: percentage females), maturity stage (± SD) and proportion of actively spawning individuals (% S; maturity stage 6) for the Faroe Plateau North-East (FPNE), Faroe Plateau West (FPW) and Faroe Bank (FB) samples.

Area	Year	*N*	Body length (cm)	Body Weight (g)	Fulton's *K*	Age (years)	% F	Maturity stage	% S
FPNE	2009	14	59.8 (± 9.4)	2340 (± 949)	1.033 (± 0.081)	3.9 (± 0.7)	50	1.9 (± 1.9)	14
	2010	26	58.2 (± 10.8)	2260 (± 1166)	1.032 (± 0.090)	3.5 (± 1.3)	46	3.3 (± 2.1)	15
FPW	2009	20	59.3 (± 8.4)	2427 (± 1061)	1.094 (± 0.102)	4.7 (± 1.2)	55	5.3 (± 1.6)	70
FB	2009	25	83.0 (± 11.3)	6862 (± 2308)	1.141 (± 0.087)	4.9 (± 1.3)	48	4.6 (± 1.6)	36
	2010	27	89.1 (± 7.8)	8329 (± 2526)	1.147 (± 0.077)	6.1 (± 1.8)	30	5.4 (± 0.7)	52

7.3.2.2. DNA extraction

DNA was extracted with the Real Pure Genomic DNA extraction kit (Durviz, Valencia, Spain), as described in Section 2.3.1.2. The extraction procedure involved an RNase treatment step to remove residual RNA. The quality and integrity of the DNA was checked, as described in Section 2.3.2.2 and Section 2.3.2.3, and samples were diluted to 45 ng/µl in 5 mM Tris, pH 8.5.

7.3.2.3. SNP genotyping

Eight SNPs identified from the RAD based analysis as being among the most discriminatory between Faroe Plateau and Faroe Bank samples (Section 7.2.3.6) were selected for genotyping. KASP assay design, based on submitted RAD sequences, was performed by LGC Genomics. SNP genotyping was performed at the Institute of Aquaculture molecular biology laboratory using the Kompetitive Allele Specific PCR (KASP^{TM}) genotyping assay (Section 2.3.5).

7.3.2.4. Statistical analyses

Analyses presented were based on two datasets, the RAD sequencing dataset comprising 40 Faroese samples (FPNE09, N = 10; FPNE10, N = 10; FB09; N = 10; FB10; N = 10; Section 7.2.2.1) and the KASP assayed dataset comprising 112 Faroese samples and 60 outlier samples (see Section 7.3.2.1 above). The RAD sequencing dataset was used for SNP discovery, whereas the KASP assayed dataset was used for SNP validation. Hence, data for six control individuals, analysed by RAD sequencing as well as by the KASP assay, were not included in the KASP assayed dataset, as they might skew the results towards higher genetic divergence between Faroe Plateau and Faroe Bank samples.

7.3.2.4.1. Comparison of results obtained by the RAD sequencing dataset and the KASP assayed dataset

First, the RAD and the KASP data were compared with respect to the power of the SNPs to differentiate among the various samples of Faroese cod. This was evaluated by pairwise levels of genetic differentiation, F_{ST}, by testing the significance of genic (allelic) differentiation by a pseudo-exact probability test, dememorisation 1000 and 100 batches of 1000 iterations, both calculated in GENEPOP. In addition, raw allele frequency data for Faroe Plateau and Faroe Bank areas were presented in order to visually inspect for potential ascertainment bias in the data.

7.3.2.4.2. SNP annotation

Genomic location and effect of SNPs was determined as described in Section 7.2.2.8.

7.3.2.4.3. Analysis of the KASP assayed dataset: Genetic variation, population structure and population assignment

The following analyses were concerned with genetic diversity and the pattern of population structure detected by these “discriminatory” SNPs in the Faroese and the outlier (non-Faroese) samples. Genetic diversity estimates were obtained from GENEPOP (Raymond and Rousset 1995; Rousset 2008). Deviations from Hardy-Weinberg equilibrium (HWE) as well as linkage disequilibrium were also tested with GENEPOP, using the probability test options and estimating P-values from a Markov chain (dememorisation 1000 and 100 batches of 1000 iterations). Overall

and pairwise levels of genetic differentiation were assessed by F_{ST} and the significance of genic differentiation was calculated as described above (Section 7.3.2.4.1). In order to evaluate the relative contribution of each of the loci to the level of differentiation detected among the Faroese locations, loci-specific results were also presented for pairwise comparisons involving Faroese samples.

In addition, the number of major distinct genetic clusters (K) was determined by the Bayesian clustering algorithm in the software STRUCTURE version 2.3.4 (Pritchard *et al*. 2000; Falush *et al*. 2007). An admixture model with correlated allele frequencies was used and runs consisted of a burn-in length of 500,000 iterations followed by a Monte Carlo Markov Chain of 1,000,000 iterations. Five replicate runs were conducted for values of K ranging from 1 to 8. To infer the most likely number of clusters, the mean estimated natural logarithm of the posterior probability (P) of the data (D), $\ln P(D)$, was considered as well as an ad hoc criterion ΔK (Evanno *et al*. 2005), which was calculated by using Structure Harvester (Earl and vonHoldt 2012).

Lastly, assignment testing in GENECLASS2 (Piry *et al*. 2004) was conducted to obtain statistical certainties of assigning Faroe Plateau (FPW and FPNE pooled) and Faroe Bank individuals to their correct population of origin. Individuals were re-allocated to known populations of origin, using the leave-one-out procedure, estimated by the Bayesian method described by Rannala and Mountain (1997) and 10,000 resamplings of individuals, as described in Paetkau *et al*. (2004). Type 1 error was set to 0.05. This was done with the RAD data as well as all data pooled (RAD and KASP). Assignment of groups of individuals was also evaluated in GENECLASS2, using the RAD-sequencing data as reference populations. The samples assigned consisted of the full five Faroese samples as well as these same samples, reduced to the first 10 individuals only.

7.3.3. Results

7.3.3.1. Amplification success and agreement between RAD sequencing and the KASP assay analyses

Two of the eight loci (*Gmo*6644 and *Gmo*14283) did not amplify successfully with the KASP assay. Alignment of the respective two RAD-tag sequences against the cod genome showed

100 % sequence similarities and *e*-values ≤ 1.00×10^{-47}. All of the remaining six loci were successfully amplified in all samples except for *Gmo*6664 in one Faroe Bank individual and *Gmo*2357 in one Faroe Plateau North-East individual. Moreover, in the six individuals that had been analysed by RAD sequencing as well as by the KASP assay, there was 100 % concordance for genotypes scored by both methods, indicating that the KASP assays were identifying the correct loci.

7.3.3.2. Genetic differentiation between Faroe Plateau and Faroe Bank cod – comparison of results obtained by the RAD sequencing dataset and the KASP assay dataset

The current panel of six SNPs was selected from a reduced set of Faroe Plateau and Faroe Bank individuals (FPNE09, N = 10; FPNE10, N = 10; FB09; N = 10; FB10; N = 10), analysed by RAD sequencing, based on their high F_{ST} values and significant outcomes of genic differentiation tests between samples from the two areas. The SNPs were then evaluated in larger datasets by a KASP assay (FPNE09, N = 14; FPNE10, N = 26; FPW; N = 20; FB09; N = 25; FB10; N = 27) to see whether the finding of significant differentiation between Faroe Plateau and Faroe Bank cod was consistent using different individuals of both stocks and larger sample sizes.

The level of genetic differentiation was significantly higher as estimated from the RAD sequencing data compared to the KASP assayed data (Table 7.9 and Table 7.10). However, the KASP assay data did provide evidence of significant differentiation among Faroe Plateau and Faroe Bank cod in four out of six comparisons when temporal samples were kept separately. Hence, these six SNP loci appear to be useful markers for differentiating among Faroe Plateau and Faroe Bank cod. Furthermore, the results suggest no genetic differentiation between sampling years within locations nor between Faroe Plateau North-East and Faroe Plateau West samples (boxed portions in Table 7.10).

Table 7.9. Matrices of pairwise F_{ST} values (below diagonal) among Faroese areas and / or years sampled and results of pseudo-exact tests of genic differentiation (above diagonal), as estimated from the RAD sequencing data. Bold values indicate significant results after sequential Bonferroni correction for multiple tests (first critical α = 0.05 / 6 tests = 0.0083).

		FB		FPNE	
		2009	**2010**	**2009**	**2010**
FB	**2009**		*NS*	**< 0.0001**	**< 0.001**
	2010	0.0284		**< 0.0001**	**< 0.0001**
FPNE	**2009**	0.4146	0.4790		*NS*
	2010	0.2101	0.2277	0.0254	

Table 7.10. Matrices of pairwise F_{ST} values (below diagonal) among all Faroese areas and / or years sampled and results of pseudo-exact tests of genic differentiation (above diagonal), as estimated from the KASP assayed data. Bold values indicate significant results after sequential Bonferroni correction for multiple tests (first critical α = 0.05 / 10 tests = 0.0005).

		FB		FPW	FPNE	
		2009	**2010**	**2009**	**2009**	**2010**
FB	**2009**		*NS*	**< 0.001**	< 0.05	**< 0.0001**
	2010	-0.0062		**< 0.001**	*NS*	**< 0.0001**
FPW	**2009**	0.0768	0.0694		*NS*	*NS*
FPNE	**2009**	0.0447	0.0238	-0.0166		*NS*
	2010	0.1071	0.1219	0.0010	0.0254	

The results above indicate some level of ascertainment bias, as a considerably higher level of genetic differentiation was detected between Faroe Plateau and Faroe Bank cod in the ascertainment samples (RAD data) compared to the validation samples (KASP data). Ascertainment bias is also evident by looking at the raw allele frequency data calculated from the two datasets (Figure 7.3). For all markers, allele frequency differences between Faroe Plateau and Faroe Bank cod were larger in the RAD dataset.

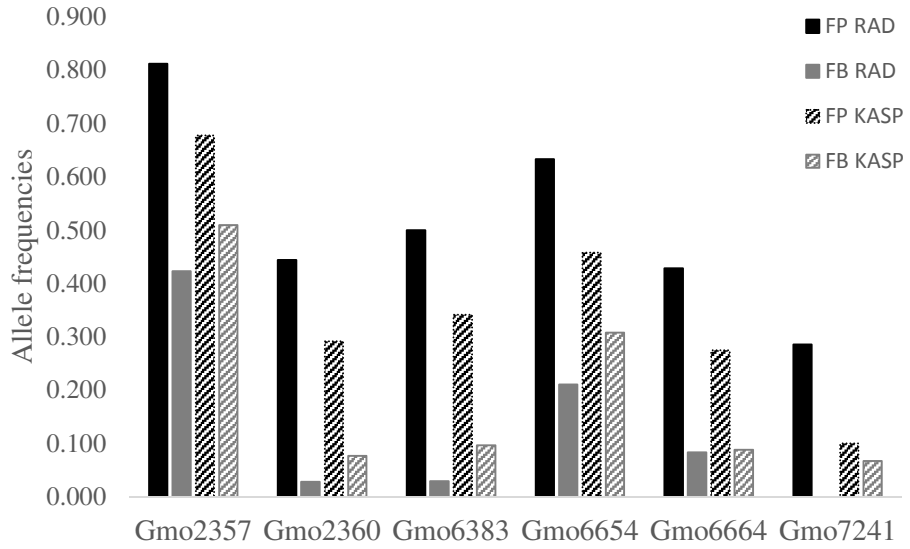

Figure 7.3. Allele frequencies of the six SNPs in samples from the Faroe Plateau and Faroe Bank analysed by RAD sequencing or genotyped by a KASP assay. For the KASP data, Faroe Plateau West and North-East areas were grouped together into a single Faroe Plateau sample.

7.3.3.3. SNP annotation

Significant BLAT hits (*e*-values $\leq 1.20 \times 10^{-76}$) were found for all six RAD-tag sequences investigated. The length of all query sequences was 96 bp (RAD-tag sequences) and all of these sequences aligned over the entire length of the sequence, with 100 % sequence identity. One SNP was located in an intergenic region, one 5' of a gene and four SNPs were located in predicted gene regions. One of these four was assocaited with a splice region, two were intronic and the last one (*Gmo*7241) caused a missense mutation in a gene involved in cyclic nucleotide biosynthetic process and intracellular signal transduction (Table 7.11). *Gmo*2357 and *Gmo*2360 were both found in the same gene scaffold (1671) of the cod genome, although > 100 kb apart (Table 7.7).

Table 7.11. Annotation of SNP loci based on the Variant Effect Predictor tool and Gene Ontology information from Ensembl. All information on biological function was inferred from electronic annotation in Ensembl. For genes with more than one transcript, prediction is provided for each of the transcripts. For genomic location of the SNPs, please refer to Table 7.7. Positive selection was suggested for all six loci (Figure 7.2).

SNP ID	Alleles	Ensembl VEP Consequence	Gene Ensembl ID	Symbol	Biological process	Molecular function
*Gmo*2357	G/C	Intronic	ENSGMOG00000000864	ptprsa	Protein dephosphorylation, de-phosphorylation, olfactory bulb development.	Protein tyrosine phosphatase activity, protein binding, phosphatase activity.
*Gmo*2360	C/T	Intergenic				
*Gmo*6383	C/T	Synonymous SNP, in splice region	ENSGMOG00000018864	ryk	Protein phosphorylation.	Protein kinase activity, protein tyrosine kinase activity, ATP binding.
		Synonymous SNP, in splice region	ENSGMOG00000018864	ryk	Same as other transcript.	Same as other transcript.
*Gmo*6654	A/G	5' of gene	ENSGMOG00000007541	upf1	Embryo development.	Nucleotide binding, DNA binding, ATP binding, hydrolase activity, nucleoside-triphosphatase activity.
*Gmo*6664	A/T	Intronic	ENSGMOG00000001356	si:dkey-46i24.1	Ubiquitin-dependent protein-catabolic process.	Ubiquitin thiolesterase activity.
*Gmo*7241	C/T	Missense SNP	ENSGMOG00000016776	adcy1a	Cyclic nucleotide biosynthetic process, intracellular signal transduction.	Phosphorous-oxygen lyase activity.

7.3.3.4. Analysis of the KASP assayed dataset

7.3.3.4.1. Genetic diversity and assumption testing

Only the KASP assay data were used for the Faroese samples in this section, as the RAD-sequencing data would skew the differentiation between Faroe Plateau and Faroe Bank cod upwards. Genetic diversity estimates for all samples are shown in Table 7.12. All samples were polymorphic at all loci except for the White Sea sample, which was monomorphic for 50 % of the loci (*Gmo*2360, *Gmo*6383 and *Gmo*7241). In addition, *Gmo*6664 displayed a low level of polymorphism in the Faroe Bank 2009 sample with a minor allele frequency (MAF) of 0.04. Hence, the Hardy-Weinberg equilibrium was tested for 44 locus × sample combinations and four significant deviations were found, though three of these were not significant after sequential Bonferroni correction (first critical $\alpha = 0.05 / 6$ loci $= 0.0083$ for samples with 6 polymorphic loci). The only significant deviation was found for *Gmo*6664 in the VB sample ($P < 0.001$), a deviation that was also reflected in the overall Hardy-Weinberg conformance of that sample ($P < 0.05$). A positive overall F_{IS} may indicate non-random mating within a sample, but the overall F_{IS} value of this sample was not markedly high compared to the other samples. Minor allele frequencies ranged from 0.040 to 0.150 and single locus estimates of observed and expected heterozygosities (H_O and H_E) from 0.050 to 0.731 and 0.077 to 0.499, respectively.

Gene diversity (H_E over all loci) was highest in the Faroe Plateau samples (H_E ranging from 0.359 to 0.408; the FPW sample intermediate of the two FPNE samples), followed by the Norwegian (H_E values of 0.316 and 0.336), the Faroe Bank (H_E values of 0.251 and 0.254) and lowest in the White Sea sample ($H_E = 0.172$).

Significant linkage disequilibrium was found in 12 out of 15 tests, nine of these were significant after sequential Bonferroni correction and all loci except *Gmo*2357 were involved. However, only *Gmo*2360 and *Gmo*6383 showed evidence of significant linkage disequilibrium in all samples examined (except the White Sea sample). The high level of non-random association of alleles at

the various markers is likely explained by the fact that the same criteria was used to select all markers.

Table 7.12. Genetic variation of the six SNPs in all eight samples. Sample size (*N*), number of individuals genotyped (*n*), minor allele frequencies (MAF), observed (H_O) and expected (H_E) heterozygosities and inbreeding coefficient (F_{IS}). Bold value indicates significant deviation from Hardy-Weinberg equilibrium following sequential Bonferroni correction (first critical α = 0.05 / 6 loci = 0.0083).

	Area:	**FB**	**FB**	**FPW**	**FPNE**	**FPNE**	**BGF**	**VB**	**WS**
	Year:	**2009**	**2010**	**2009**	**2009**	**2010**	**2004**	**2005**	**2003**
Locus		*N* = 25	*N* = 27	*N* = 20	*N* = 14	*N* = 26	*N* = 20	*N* = 20	*N* = 20
*Gmo*2357	*n*	25	27	20	14	25	20	20	20
	MAF	0.480	0.463	0.300	0.286	0.360	0.275	0.300	0.175
	H_O	0.480	0.481	0.300	0.429	0.480	0.350	0.600	0.250
	H_E	0.499	0.497	0.420	0.408	0.461	0.399	0.420	0.289
	F_{IS}	0.059	0.051	0.309	-0.013	-0.021	0.147	-0.407	0.159
*Gmo*2360	*n*	25	27	20	14	26	20	20	20
	MAF	0.080	0.074	0.300	0.179	0.346	0.150	0.075	-
	H_O	0.160	0.148	0.400	0.214	0.615	0.200	0.150	-
	H_E	0.147	0.137	0.420	0.294	0.453	0.255	0.139	-
	F_{IS}	-0.067	-0.061	0.073	0.305	-0.342	0.240	-0.056	-
*Gmo*6383	*n*	25	27	20	14	26	20	20	20
	MAF	0.120	0.074	0.350	0.214	0.404	0.150	0.250	-
	H_O	0.240	0.148	0.500	0.286	0.731	0.200	0.300	-
	H_E	0.211	0.137	0.455	0.336	0.482	0.255	0.375	-
	F_{IS}	-0.116	-0.610	-0.073	0.186	-0.503	0.240	0.225	-
*Gmo*6654	*n*	25	27	20	14	26	20	20	20
	MAF	0.360	0.259	0.375	0.357	0.577	0.425	0.400	0.350
	H_O	0.240	0.222	0.550	0.429	0.692	0.350	0.400	0.500
	H_E	0.461	0.384	0.469	0.459	0.488	0.489	0.480	0.455
	F_{IS}	0.495	0.437	-0.148	0.103	-0.402	0.307	0.192	-0.073
*Gmo*6664	*n*	25	26	20	14	26	20	20	20
	MAF	0.040	0.135	0.275	0.286	0.269	0.200	0.225	0.175
	H_O	0.080	0.192	0.350	0.288	0.385	0.200	0.050	0.350
	H_E	0.077	0.234	0.399	0.408	0.393	0.320	0.349	0.289
	F_{IS}	-0.021	0.194	0.147	0.333	0.042	0.397	**0.863**	-0.188
*Gmo*7241	*n*	25	27	20	14	26	20	20	20
	MAF	0.060	0.074	0.075	0.143	0.096	0.100	0.150	-
	H_O	0.120	0.148	0.150	0.286	0.192	0.100	0.300	-
	H_E	0.113	0.137	0.139	0.245	0.174	0.180	0.255	-
	F_{IS}	-0.044	-0.061	-0.056	-0.130	-0.087	0.465	-0.152	-
Over	*n*	25.0	26.8	20.0	14.0	25.8	20.0	20.0	20.0
all	%P	100.0	100.0	100.0	100.0	100.0	100.0	100.0	50.0
loci	H_O	0.220	0.223	0.375	0.322	0.516	0.233	0.300	0.184
	H_E	0.251	0.254	0.384	0.359	0.408	0.316	0.336	0.172
	F_{IS}	0.145	0.140	0.048	0.140	-0.246	0.286	0.133	-0.040

7.3.3.4.2. Population structure

Neither the RAD nor the KASP data indicated any genetic differentiation between sampling years within locations (Table 7.9 and Table 7.10); hence, temporal samples were pooled for analyses of population structure. By F_{ST} analysis, significant genetic differentiation was found among the six areas ($F_{ST} = 0.0377$; $P < 0.0001$), with individual loci showing varying levels of differentiation (Table 7.13).

Table 7.13. Overall and locus-specific global differentiation for the areas FB (N = 52), FPW (N = 20), FPNE (N = 40), BGF (N = 20), VB (N = 20) and WS (N = 20). F_{ST} values and results from pseudo-exact tests of genic differentiation (P) are shown. *NS* indicates no significant differentiation.

Locus	F_{ST}	P
*Gmo*2357	0.0412	< 0.05
*Gmo*2360	0.0974	< 0.0001
*Gmo*6383	0.1053	< 0.0001
*Gmo*6654	0.0072	*NS*
*Gmo*6664	0.0240	< 0.05
*Gmo*7241	0.0088	*NS*
All	0.0377	< 0.0001

Pairwise comparisons of F_{ST} showed a clear separation among the Faroe Plateau and the Faroe Bank samples, with F_{ST} values of 0.0809 ($P < 0.0001$) and 0.0853 ($P < 0.0001$) for the FB – FPW and FB – FPNE comparison, respectively. No significant differentiation was evident between the two Faroe Plateau samples. The loci investigated had been selected to show the maximum differentiation possible among Faroe Plateau and Faroe Bank samples, and accordingly, levels of differentiation were high, comparable to the ones detected among the out-group sample (WS) and the Faroese samples (F_{ST} ranging from 0.0671 to 0.1058). All the remaining F_{ST} values (concerned with the Norwegian samples) were nonsignificant (Table 7.14).

Table 7.14. Matrices of pairwise F_{ST} values (below diagonal) and results of pseudo-exact tests of genic differentiation (above diagonal). Bold values indicate significant results after sequential Bonferroni correction for multiple tests (first critical α = 0.05 / 15 tests = 0.0033).

	FB	**FPW**	**FPNE**	**BGF**	**VB**	**WS**
FB		**< 0.0001**	**< 0.0001**	< 0.05	< 0.05	**< 0.01**
FPW	0.0809		*NS*	*NS*	*NS*	**< 0.0001**
FPNE	0.0853	-0.0101		*NS*	*NS*	**< 0.0001**
BGF	0.0280	0.0043	0.0114		*NS*	< 0.05
VB	0.0346	0.0040	0.0112	-0.0207		< 0.01
WS	0.0671	0.1050	0.1058	0.0176	0.0423	

Pairwise F_{ST} estimates among Faroese samples for individual loci are shown in Figure 7.4. For four of the markers (*Gmo*2357, *Gmo*2360, *Gmo*6383 and *Gmo*6664), F_{ST} values were high for the Faroe Plateau – Faroe Bank comparisons and zero for the comparison of sites within the Faroe Plateau region. The SNP with the highest F_{ST} values (0.1914 and 0.1587 for the FB-FPW and the FB-FPNE comparisons, respectively) was *Gmo*6383, which aligned to a splice region of the *ryk* gene involved in protein phosphorylation (Table 7.11). In zebrafish, this gene appears to have a role in egg development and egg activation (Pelegri 2003). For *Gm*6654, F_{ST} was only high for the FB-FPNE comparison, but zero for the FB-FPW, whereas *Gmo*7241 showed zero F_{ST} values in all three comparisons (and nonsignificant outcome of the pseudo-exact test of genic differentiation for all three comparisons).

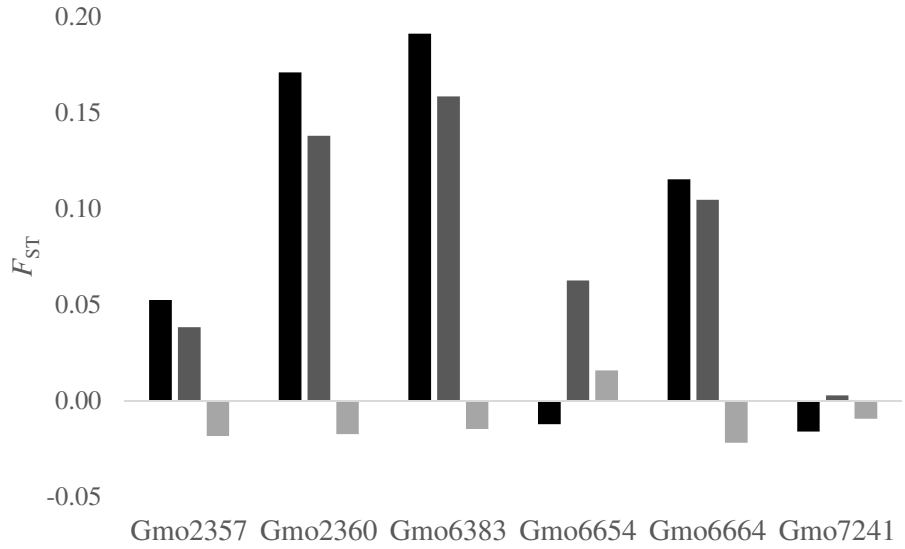

Figure 7.4. Pairwise estimates of genetic differentiation among Faroese locations for the six SNP loci. The bars represent pairwise F_{ST} values for the FB-FPW (black), FB-FPNE (dark grey) and FPW-FPNE (light grey) comparisons.

Population assignment using STRUCTURE grouped the data into two major genotypic clusters, as supported by both the average Ln$P(D)$ as well as ΔK, which were both maximised at $K = 2$ (Figure 7.5). Proportions of membership of White Sea and Faroe Bank cod in Cluster 2 were high and although the remaining samples were more admixed between the two clusters, the Norwegian cod had higher proportions of membership in this cluster, while the Faroe Plateau cod had higher proportions of membership in Cluster 1 (Figure 7.6). Hence, the STRUCTURE analysis agreed with F_{ST} analysis in the way that marked divergence was found between the Faroe Bank sample at one end of the distribution and the two Faroe Plateau samples at the other end, with the Norwegian samples found somewhere in between. However, while F_{ST} analysis indicated significant genetic differentiation between the Faroe Bank and the White Sea samples (Table 7.14); this was not evident with the STRUCTURE analysis.

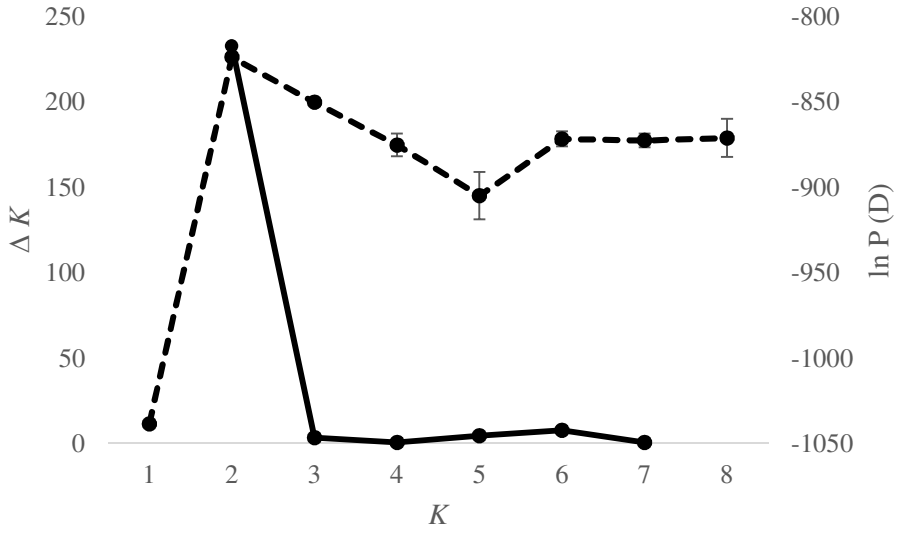

Figure 7.5. Plots used to detect the number of major genotypic clusters (*K*) of the data. Full line and left axis: Δ*K* estimated from Evanno *et al*. (2005). Stipled line and right axis: Average ln*P*(*D*) (±SD) over five runs of each *K*.

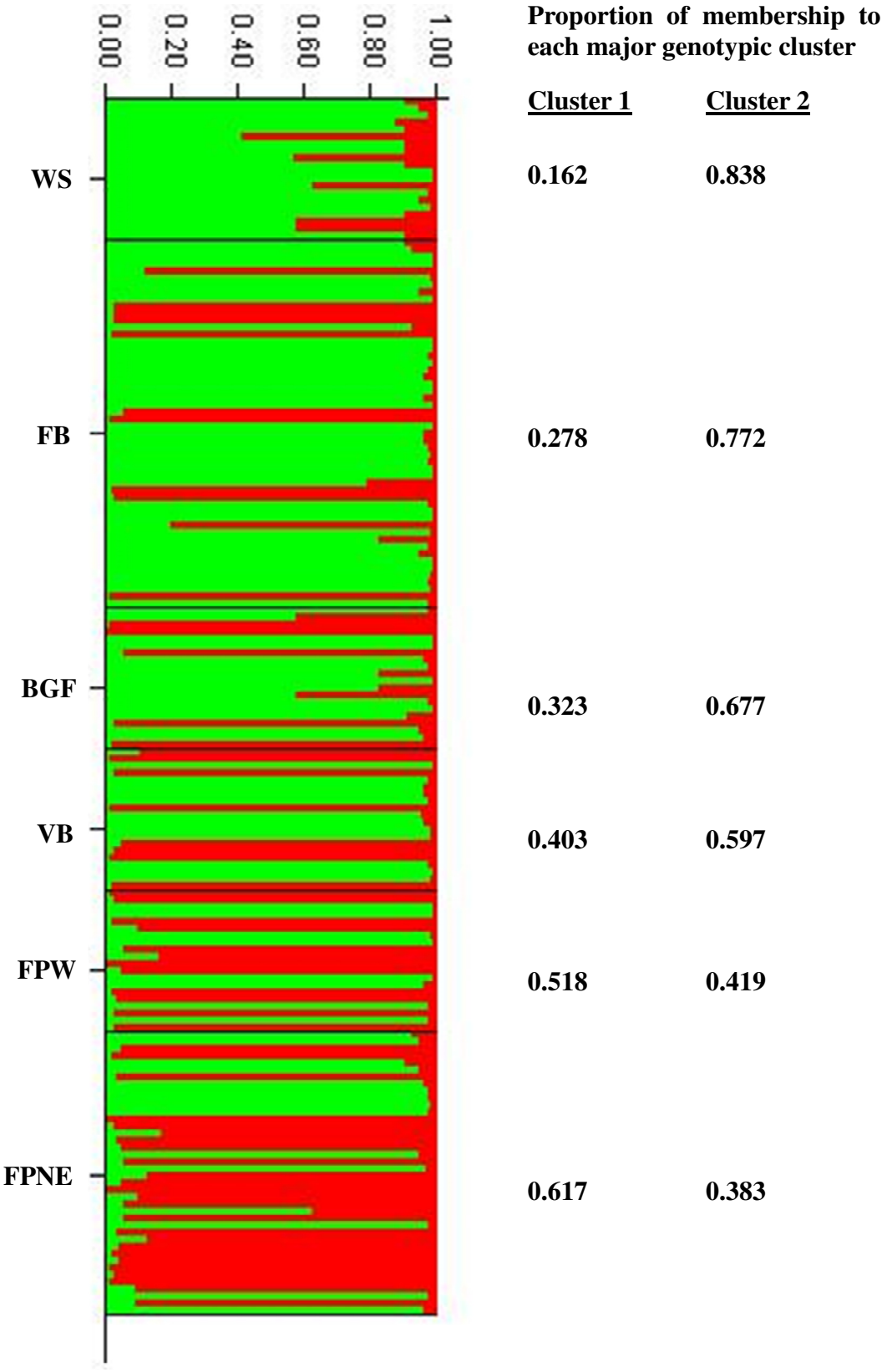

Figure 7.6. To the left, a barplot of the clustering analysis in STRUCTURE, showing a run for a *K* value of 2. Each vertical bar represents an individual and the colours refer to the two population clusters detected. The estimated proportions of membership to each of the two clusters is indicated to the right.

7.3.3.5. Population assignment of individuals and groups of individuals

Although the results above showed a clear separation among Faroe Plateau and Faroe Bank cod at this reduced SNP panel, the ability of the SNPs to assign individuals to their population of origin was limited. Assignment tests of Faroe Plateau and Faroe Bank individuals via GENECLASS2 version 2.0 showed that 75.7 % correct assignments were obtained when individuals were assigned based on the RAD sequencing dataset (quality index = 60.58) and only 67.8 % (quality index = 58.16) when all data were used (the RAD sequencing and the KASP assay data pooled).

Nonetheless, the six SNPs allowed 100 % correct assignments of groups of individuals to their population of origin, when the groups consisted of the five Faroese samples, each collected from one single area and year and, thus, likely to represent catches obtained during commercial fishery. The RAD sequencing data (FB: N = 20; FPNE: N = 18) were used as reference populations. However, when group samples were reduced to ten individuals only, one of the Faroe Plateau samples (FPNE09) was wrongly assigned to the Faroe Bank population (Figure 7.7).

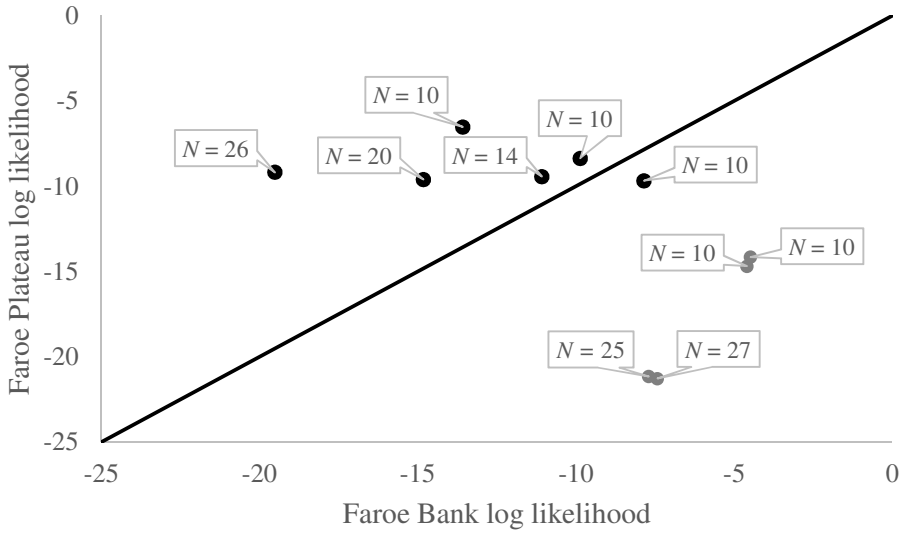

Figure 7.7. Assignment of groups of individuals to populations based on multilocus genotypes at six SNP loci. Assignment plot of log likelihood values of genotypes of groups of Faroe Plateau (black dots) and Faroe Bank (grey dots) individuals to the Faroe Plateau population (above the 45º line) or the Faroe Bank population (below the 45º line). Sample sizes by each group.

Population assignment was also conducted with five loci only (excluding *Gmo*7241, which showed no significant differentiation between Faroe Plateau and Faroe Bank samples; Section 7.3.3.4.2; Figure 7.4). Based on the RAD sequencing data, 75.7 % of the individuals were correctly assigned (quality index = 61.47), while 71.8 % correct assignments were obtained (quality index = 59.65) when all data were used (the RAD sequencing and the KASP assay data pooled). Hence, assignments of individuals to their population of origin were equally / more successful, and with marginally higher quality indices, when assignment tests were based on five loci only.

Similar to analysis with six loci, the five SNPs allowed 100 % correct assignments of groups of individuals to their population of origin, when the groups consisted of the five Faroese samples, each collected from one single area and year and, thus, likely to represent catches obtained during commercial fishery. When the group samples were reduced to ten individuals only, the FPNE09 sample was again wrongly assigned to the Faroe Bank population (Figure 7.8).

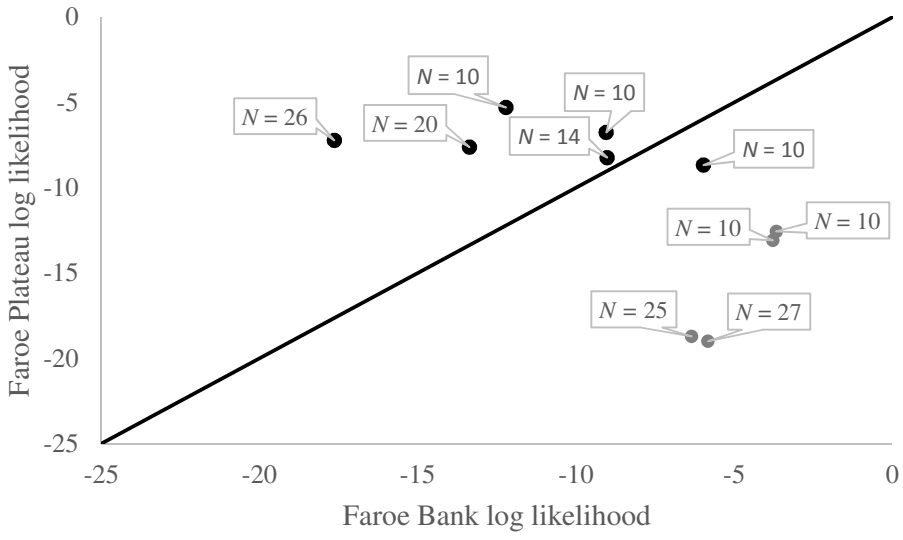

Figure 7.8. Assignment of groups of individuals to populations based on multilocus genotypes at five SNP loci. Assignment plot of log likelihood values of genotypes of groups of Faroe Plateau (black dots) and Faroe Bank (grey dots) individuals to the Faroe Plateau population (above the 45º line) or the Faroe Bank population (below the 45º line). Sample sizes by each group.

In conclusion, these SNPs could be useful for assignment of groups of individuals, i.e. obtained from commercial fishery catches, to their population of origin, if sample sizes are sufficiently large. Moreover, *Gmo*7241 should be excluded from such analysis, as assignments appear to be marginally more successful with five loci only.

7.4. Summary of results obtained in experiment I and II

7.4.1. Experiment I

- 13,582 RAD loci were detected with high confidence.
- 3,386 diallelic markers were identified, carrying a total of 3533 SNP loci..
- No loci were fixed for different alleles between Faroe Plateau and Faroe Bank cod.
- Only a small proportion of the total 3533 SNPs showed evidence of genetic differentiation between Faroe Plateau and Faroe Bank cod (58 SNPs; $P < 0.05$).
- Seventeen of the 37 loci suggested to be subject to positive selection were amongst the 58 informative SNPs.
- Average F_{ST} values between the two areas were higher, as estimated for the 17 putatively non-neutral loci compared to the remaining 41 putatively neutral loci, with mean F_{ST} values of 0.2902 (± 0.1271) and 0.1376 (± 0.0322), respectively.
- Using assignment tests and a subset of five SNP loci (F_{ST} values between phenotypic female and male samples > 0.15; $P < 0.05$), *c*. 95 % of the individuals could be assigned to their respective sex. Two of these SNPs were found in coding regions of the cod genome, though causing synonymous mutations. In addition, one of these was associated with (*c*. 3000 bp upstream of) a gene involved in dosage compensation.
- A set of eight informative SNPs (F_{ST} values between Faroe Plateau and Faroe Bank samples > 0.25; $P < 0.0005$) were chosen for validation in larger samples. Non-neutrality was suggested at all eight loci by outlier analysis.

7.4.2. Experiment II

- Two out of eight loci discovered by RAD sequencing did not amplify successfully with the KASP assay.

- The six individuals that had been analysed by RAD sequencing as well as by the KASP assay, showed 100 % concordance for genotypes scored by the two methods.
- When results from the RAD and the KASP data were compared, a significantly higher level of genetic differentiation between Faroe Plateau and Faroe Bank cod samples was seen with the RAD data, though the KASP data did also show evidence of genetic differentiation between Faroe Plateau and Faroe Bank cod samples.
- Gene diversity was highest in the Faroe Plateau samples, followed by the Norwegian, the Faroe Bank and the White Sea sample, in the order mentioned, and 100 % polymorphic loci were observed for all except the White Sea sample, where only 50 % of the loci were polymorphic.
- Over all loci, significant genetic differentiation was found among the six areas ($F_{ST} = 0.0377$; $P < 0.0001$), although values varied much among the loci.
- Concerning the Faroese samples, there was no evidence of genetic differentiation between sampling years within locations nor between Faroe Plateau North-East and Faroe Plateau West samples. However, a high level of genetic differentiation was seen between Faroe Bank and the Faroe Plateau cod (FB – FPW: $F_{ST} = 0.0809$, $P < 0.0001$; FB – FPNE: $F_{ST} = 0.0853$, $P < 0.0001$), comparable to the level of differentiation detected between the out-group sample (White Sea) and the Faroese samples (F_{ST} ranging from 0.0671 to 0.1058).
- Locus-specific pairwise F_{ST} values for Faroe Plateau – Faroe Bank comparisons were high for four of the loci (*Gmo*2357, *Gmo*2360, *Gmo6383* and *Gmo*6664). For *Gmo*6654, F_{ST} was only high for the FB – FPNE comparison and *Gmo*7241 (the only missense SNP) showed no differentiation.
- Analysis in STRUCTURE suggested two major genotypic clusters, where the White Sea and the Faroe Bank samples had relatively high proportions of membership in Cluster 2. The remaining samples were more admixed between the two clusters, although the Faroe Plateau cod had higher proportions of membership in Cluster 1.
- In assignments of individuals and groups of individuals to their respective population of origin, *Gmo*7241 could be omitted, without any effect on the results. Individual assignments

were successful for *c*. 75 % of the individuals. Assignments of groups of individuals were successful with sample sizes ranging from 14 to 27 individuals, but not when groups were reduced to ten individuals.

7.5. Discussion

By RAD sequencing of barcoded Faroe Plateau and Faroe Bank cod samples in a single sequencing lane of an Illumina HiSeq platform 13,582 RAD loci were detected with high confidence; 3,386 were diallelic, carrying 3533 SNPs. Further analyses of a subset of these in larger sample sizes confirmed that statistically significant genetic differentiation could be detected between Faroe Plateau and Faroe Bank. Though no single locus was fixed for different alleles in the two Faroese populations, five SNPs were identified, that were able to confidently distinguish between samples of Faroe Plateau and Faroe Bank cod (with the proviso that the samples were sufficiently large; > 14).

7.5.1. Ascertainment bias and SNP validation

Although RAD sequencing offers a way to simultaneously discover and genotype SNPs, it is important to validate the SNPs and address the issue of ascertainment bias (ICES 2012). Hence, in this study, following SNP discovery and genotyping by RAD sequencing, an additional experiment was conducted to verify the findings. As would be expected, and outlined in more detail below, a proportion of the SNPs were not assayable by a simpler genotyping platform and there was some level of ascertainment bias because SNPs were discovered (ascertained) in relatively limited sample sizes from a limited geographical area, and selected to maximise differentiation between these two specific samples. Although findings here highlight the importance of addressing potential ascertainment bias, locus by locus, the results also verify the utility of RAD technology for shedding light on potential genetic divergence between natural populations, where other methods may be of more limited use (see for example Ogden *et al*. 2013).

7.5.1.1. SNP validation

Two out of eight candidate SNPs were not successfully amplified by the KASP assay. "Sequencing artefacts" are likely to arise in large-scale SNP discovery (ICES 2012). However, as these two RAD-tag sequences aligned perfectly to the cod genome, the SNPs may simply not have been assayable by the genotyping platform / protocol applied. Primarily for cost reasons, the "KASP by Design" (KBD) service was used for this study, whereby LGC design and provide a single set of primers per SNP. Successful PCR assays are not guaranteed for all such designs. Verification of genotypes obtained by RAD sequencing and the KASP assay was also addressed and genotype agreement was verified in all (six) individuals that were analysed by both methods.

7.5.1.2. Ascertainment bias due to SNP discovery in a limited number of individuals

In the SNP discovery process, a limited number of Faroe Plateau and Faroe Bank individuals were used (40 in total), while a large number of SNPs were identified. Thus, the possibility that correlation between samples and allele frequencies at a particular locus may be due to chance alone, rather than reflect a true stock difference, is a likely outcome for at least a proportion of the identified loci. Furthermore, in the selection process for individuals to include, an effort was made such as to only include individuals with a high probability of belonging to the stock in question (mature and spawning individuals). Hence, it is likely that the SNP discovery process may have introduced ascertainment bias, even when SNPs were analysed in the populations used for SNP discovery. This was confirmed, as ascertainment bias was evident merely by looking at the raw allele frequencies, where all six loci displayed a higher divergence between Faroe Plateau and Faroe Bank samples, as analysed by RAD sequencing compared to the KASP assay of larger population samples. In agreement, substantially higher levels of genetic differentiation estimates between the stocks were observed by the RAD sequencing compared to the KASP assayed data. Despite this, genotyping of SNPs in larger population samples did confirm a significant level of genetic differentiation between the stocks, with overall F_{ST} values >> than those obtained by microsatellite data (Chapter 5). However, one out of the six loci did not display significant genetic

differentiation between Faroe Plateau and Faroe Bank samples (possibly a type I error) and, thus, need not be included in population assignment studies.

7.5.1.3. Ascertainment bias due to SNP discovery from a limited geographical area

An important hypothesis and concern regarding large-scale development of SNP markers for population genetic studies is that SNPs, identified from samples covering a limited geographical area, will skew genetic diversity estimates towards higher diversity in these populations compared to other, geographically more distant, populations (Bradbury *et al*. 2011). This hypothesis was supported by current data, where the lowest gene diversity estimate and the lowest proportion of polymorphic loci were observed for the geographically most distant sample (White Sea). Whereas the remaining samples had 100 % polymorphic loci, the White Sea sample exhibited 50 % polymorphic loci only.

7.5.2. Consideration of ascertainment bias when using these SNPs for downstream analyses

Ascertainment bias can potentially affect all types of downstream analyses (Helyar *et al.* 2011). Although various aspects of ascertainment biases were identified in this study, the ability of a handful of SNPs to identify significant genetic divergence between Faroe Plateau and Faroe Bank cod, as investigated in the RAD sequencing data, was replicated in an independent setup. Moreover, these SNPs could be used for assigning groups of individuals from these two areas back to their natal population.

7.5.2.1. Population assignment testing

There is a growing interest in assigning (migrating) individuals, even animal products, back to their population of origin and the current study provides a valuable resource that could be further explored for such purposes. The handful of SNPs tested in the current study ("best" eight out of a total of 58 loci that were significant at the 5 % probability level) provided limited success concerning individual assignment of Faroese cod; however, the number of SNPs required for accurate assignment of individuals to their natal origin is likely to be considerably larger than this. For example, from 7,000 SNPs, Karlsson *et al*. (2011) found 60 SNPs that could accurately

identify individual salmon as either wild or farmed of origin. The genetic differences detected between the two were considered a result of domestication processes and the study highlighted the value of large genome scans to identify SNPs that can be used to investigate genetic introgression caused by escaped farmed fish. Thus, future studies might want to explore the possibility of assigning individual Faroese cod back to their population of origin using a higher number of the SNPs, discovered in the current study. If cod farming is going to have a comeback in the Faroe Islands, and especially if focus again will be on Faroe Bank cod, these SNPs should provide a valuable resource for monitoring potential genetic introgression from escaped farmed fish and / or eggs.

SNP data provided by the current study may also be a resource of potential use for population assignment in other cod populations, although some considerations are needed before undertaking large-scale investigations of this sort. The success of population assignment tests using SNPs is linked to SNP diversity in the various population samples and may be compromised for many of the populations if the SNPs used have been discovered from a limited geographical context. For example, Bradbury *et al*. (2011) evaluated the affect of ascertainment bias on population assignments of cod from a wide geographical range using a set of SNPs developed from North West Atlantic cod samples. According to theory, diversity estimates (H_E and percent polymorphic loci) were highest in the Western and lowest in the Eastern populations and SNPs with the largest declines in diversity away from the ascertainment region provided the most successful assignment within this region and poor assignment outside the region. Nonetheless, 100 % correct assignments were found for a couple of isolated locations and, when all loci were used, assignment success was surprusingly similar over the whole geographical range (Bradbury *et al*. 2011).

In the current study, there was a considerable decline in diversity from the ascertainment (Faroese) samples to the most distant sample (the White Sea cod). However, no significant differences in diversity estimates were seen between Faroese and Norwegian cod. Hence, although SNPs discovered here may be of limited or no use for assignment testing in White Sea

and similarly distant samples, they could be explored further for assignment use in nearby cod populations.

7.5.2.2. Studies of cod population structure

As discussed above, SNPs with a certain level of geographic bias may still be of use for stock identification purposes. However, careful consideration should be taken when applying SNPs from the current study to studies of cod population structure in a wider geographical context. Rosenblum and Novembre (2007) investigated types of ascertainment bias by comparing frequency and genetic summary statistics at three ascertainment panels with "true population values" for populations of Eastern fence lizard (*Sceloporus undulatus*). Compared to an intentionally diverse panel (constructed from a number of geographical populations) and random panels (individuals chosen at random from the grand total of available individuals), the panel based on one single geographical population showed the highest deviation for all statistics estimated.

The fact that the informative SNP panel showed the highest level of divergence between the Faroe Plateau and the Faroe Bank cod samples was not unanticipated, as the SNPs were selected on basis of their ability to do exactly this. Other aspects of analyses of population structure should be interpreted with caution, however. Results from STRUCTURE indicated most similarity between the White Sea and the Faroe Bank samples. This is likely to be an artefact of the SNP selection process and was not seen by the genic differentiation tests, where comparisons between the White Sea sample and both Norwegian samples were nonsignificant, while comparisons between the White Sea sample and all Faroese samples were significant. The fact that these SNPs failed to detect any divergence between Faroese and Norwegian cod samples (as estimated by genic differentiation tests) should also be interpreted with caution; particularly, since evidence from various microsatellite analyses did suggest significant differentiation between these areas (Chapter 5).

7.5.3. Association between SNP loci and phenotypic sex

Cod may spawn in sea cages, which results in loss of somatic growth and may potentially have a negative impact on the surrounding ecosystem (Section 1.4.3). Consequently, cod aquaculture would benefit from the ability to produce all-female populations (Haugen *et al*. 2012) and recent efforts have been made to understand the sex-determining mechanisms in this species (Johnsen *et al*. 2010; Johnsen and Andersen 2012; Haugen *et al*. 2012, Johnsen *et al*. 2013). Cod are gonochoristic, i.e. individuals develop as and remain either females or males (Devlin and Nagahama 2002; Chiasson *et al*. 2008). Histological examinations of gonadal tissue have shown that sexual differentiation in females can be observed between 14 and 20 mm total body length and somewhat later in males (Haugen *et al*. 2012). A number of genes have been implied in sexual differentiation in cod, as investigated by expression studies, for example *dmrt1* (Johnsen *et al*. 2010), *amh* and paralogs of *sox9a* and *cyp19a* (Haugen *et al*. 2012; Johnsen *et al*. 2013).

Cytogenetic approaches have shown that Atlantic cod possess a diploid number of 46 chromosomes (Ghigliotti *et al*. 2012). No heteromorphic sex chromosome could be visualised, however (Ghigliotti *et al*. 2012). This is often the case for fish, amphibians and reptiles, for which molecular methods provide a useful approach to identify sex determination mechanisms. Recently, RAD technology has been applied to identify sex-specific markers (Carmichael *et al*. 2013; Gamble and Zarkower 2014) and sex-determining genomic regions (Palaiokostas *et al*. 2013a,b) in a variety of species.

Out of the 3533 SNPs evaluated in the current study, 25 SNPs appeared to be statistically associated with phenotypic sex, while four SNPs in particular proved efficient for assigning individual genotypes to their respective sex. All of these four SNPs displayed female-specific alleles, though they were not present in all females. While female-specific alleles at markers associated with sex could potentially suggest females as the heterogametic sex (i.e. Ventura *et al*. 2011), the apparent female-specific alleles reported here could also be an artefact, resulting from the relatively small samples sizes with missing genotypes for some of the individuals (Table 7.5) and / or incorrect scoring of phenotypic sex. The latter is unlikely, however, considering that the

investigated specimens were adult and mature individuals (Table 7.1). Nonetheless, for two of the markers, the genomic sequence of cod (www.ensembl.org), which represents a male specimen (Star *et al*. 2011), was found to display the "female-specific" alleles; hence, these data cannot be used to infer any specific sex-determining mechanism for cod. Furthermore, the relative genomic position of the four SNPs could not be implied from the current annotation of the cod genome, which is yet in its infancy in that it has not been assembled into chromosomes (www.ensembl.org).

The identification of apparently sex-linked SNPs in the current study must be considered to be of a relatively preliminary nature and should be further validated. However, genomic annotation analysis showed association between one of the four informative SNPs (*Gmo*11037) and a gene (*dpy30*) that is homologous to *C. elegans dpy-30,* which is involved in dosage compensation (Hsu and Meyer 1994; Petty *et al*. 2011). Recent studies have also shown early male-specific expression of *dpy-30* homologs in the Pacific Oyster (*Crassostrea gigas*; Dheilly *et al*. 2012) and the European clam (*Ruditapes decussatus*; de Sousa *et al*. 2014), suggesting a role for this gene in sex differentiation.

7.5.4. Significant genetic differentiation between Faroe Plateau and Faroe Bank cod

Despite ascertainment bias, data presented in this chapter should still be suited for inferences of population structure within the ascertainment region and a significant level of genetic differentiation was observed between Faroe Plateau and Faroe Bank cod with no significant genetic differentiation between sites within the Faroe Plateau.

7.5.4.1. Evaluation of potential evidence of adaptive divergence

The current data might suggest some level of adaptive divergence between Faroe Plateau and Faroe Bank cod, when considering the representation of putatively non-neutral loci (17 out of 37) that were among the small proportion of markers (58 out of 3533 SNPs) that indicated statistically significant genetic divergence between cod from the two areas. Average F_{ST} values between the areas were higher as estimated from the putatively non-neutral compared to the putatively neutral

loci. Furthermore, non-neutrality was inferred for all top eight informative markers. Nonetheless, to conclude adaptive divergence, stringent testing is required to connect the three: genotype, phenotype and fitness (Barrett and Hoekstra 2011). Additionally, as outlined above, the data may be subject to ascertainment bias and allele frequencies are likely skewed towards higher differentiation between the areas compared to "real" population frequencies.

Three of the six successfully amplified SNPs in experiment II were found in genic regions of the cod genome. However, as discussed in Section 5.4.1.2, markers in genic regions are not necessarily under any selective pressure, just as markers in non-genic regions can be under the influence of selection acting on nearby genomic locations. SNPs that can cause changes in proteins include missense, nonsense and frameshift SNPs as well as SNPs in splice sites. In addition, SNPs in regulatory regions may alter gene expression levels (Schmitt *et al*. 2010). In fact, one missense SNP was identified, but this was the SNP (*Gmo*7141), that did not show any significant divergence between Faroe Plateau and Faroe Bank cod. *Gmo*6383 showed the highest level of divergence between Faroe Plateau and Faroe Bank cod (F_{ST} values of 0.1914 and 0.1587 for comparisons between the Faroe Bank and the Faroe Plateau West and Faroe Plateau North-East, respectively). This SNP was found to be located in a splice site and could potentially be interesting in terms of local adaptation. Furthermore, the SNP aligned to the *ryk* gene, implied as a maternal factor, involved in zebrafish oogenesis (egg development) and egg activation (Pelegri 2003). Considering the differences in reproductive strategies between Faroe Plateau and Faroe Bank cod (Chapter 3), this gene could be an interesting candidate gene for further investigations of potential adaptive divergence between the two.

7.5.5. Conclusions

The current study found evidence of significant genetic differentiation between Faroe Plateau and Faroe Bank cod and, in agreement with for example Ogden *et al*. (2013), emphasised the utility of RAD technology for resolving population structure where other methods may give inconclusive answers. In addition, SNPs were identified that could confidently trace groups of Faroe Plateau or Faroe Bank cod to their population of origin. However, a number of analyses in

here also highlighted the importance of addressing ascertainment bias, which is inherently associated with large-scale SNP discovery from a limited number of individuals and / or a confined geographical region. Lastly, the data created in this study may be of high value for future stock identification purposes and / or, in the case of a future farming of Faroe Bank cod, for monitoring the potential impact of escaped farmed fish or unintentional egg releases from sea cages.

Chapter 8. General discussion

The primary aim of the present work was to provide an in-depth investigation of potential genetic differences between Faroe Plateau and Faroe Bank cod. Both molecular marker data and aspects of reproductive and early life history strategies provided evidence of significant genetic differences between cod from the two areas. In addition, the genetic data provided evidence of significant genetic differentiation between Faroese and geographically more distant populations. In particular, RAD technology enhanced the investigation and provided the foundation for developing a set of SNP markers that could assign groups of Faroe Plateau and / or Faroe Bank individuals to their population of origin. These results have implications for both the management of the natural populations and for a potential future cod farming industry in the Faroe Islands. Of interest to aquaculture in general, an investigation of progeny from captive spawning Faroe Bank cod revealed a large interfamily skew in survival of both eggs and fry, with no apparent relationship between the two.

8.1. Evidence of genetic and reproductive differences between Faroe Plateau and Faroe Bank cod

8.1.1. Evidence from molecular markers

Low signal to noise ratios may sometimes hamper the detection of potential genetic divergence between marine populations and optimal study designs should include samples of sufficient sizes, temporal replicates and the use of suitable marker systems (Reiss *et al*. 2009). In addition, sampling should be conducted at the time of year when stock integrity is at its maximum, which often means that spawning individuals should be targeted (Hauser and Carvalho 2008). This study largely complied with these recommendations and significant genetic differentiation was identified between Faroe Plateau and Faroe Bank cod.

Non-genetic data from fisheries research indicate that Faroe Plateau cod show some degree of spawning site fidelity (Section 1.6.1.1); hence, samples were obtained from cod that apparently belonged to the Northern spawning ground as well as for cod that belonged to a smaller spawning

ground west of the islands (Faroe Plateau North-East and Faroe Plateau West samples, respectively). Unfortunately, temporal replicates could not be obtained for the Faroe Plateau West location. Hence, sampling effects may explain in part the somewhat ambivalent results for this location, where the microsatellite phylogenies generated in this thesis suggested a closer evolutionary connection between Faroe Plateau North-East and Faroe Bank cod than between cod from the two Faroe Plateau locations. Alternatively, as indicated by principal component analysis based on microsatellite data, these unexpected findings may be explained by some level of temporal variation within the Faroe Plateau samples. However, in agreement with Pampoulie *et al*. (2008a), data on SNPs discovered by RAD sequencing failed to detect any substructure on the Faroe Plateau and, apart from the principal component analysis mentioned, neither microsatellite nor SNP data detected any temporal variation in allele frequencies. Three sampling years are not sufficient to conclude temporal stability of allele frequencies. Nonetheless, these findings corroborate findings by others, which suggest temporally stable allele frequencies in Faroe Plateau and Faroe Bank cod samples. Nielsen *et al*. (2007) investigated *Pan*I and microsatellite variation in four cod populations, including Faroe Plateau cod (samples from 1969 and 2002) and Faroe Bank cod (samples from 1978, 1992 and 2002) and observed no statistically significant temporal variation at microsatellites or the *Pan*I locus, despite a trend of decreasing frequencies of the $Pan\mathrm{I}^{\mathrm{B}}$ allele. In addition, in an analysis of 98 gene-associated SNPs, Nielsen *et al*. (2009b) observed neither directional nor stabilising selection over a period of 24 years for Faroe Bank cod.

The various marker systems applied in the current work revealed varying levels and significance of genetic differentiation between the two areas. Despite early investigations that indicated genetic divergence between Faroe Plateau and Faroe bank cod at hemoglobin (Jamieson and Birley 1989) and transferrin (Jamieson and Jones 1967) genes, analyses of candidate polymorphisms in the hemoglobin gene *Hb-β1* and the transferrin gene *Tf1* present in the current samples showed no significant differentiation between the areas. Concerning hemoglobin results, they are in agreement with the data of Magnussen (1996). Transferrin data used in the present

study have also been published in Andersen *et al*. (2011). While this work suggested an apparent absence of small-scale differentiation at the transferrin gene (*Tf1*), significant adaptive divergence was found between North-West and North-East Atlantic cod populations at this gene (Andersen *et al*. 2011).

Due to the 1 – 2 ºC difference in mean sea temperature between the Faroe Plateau and the Faroe Bank (Magnussen 2006) and the relationship between variation in the *Hb-β1* gene and environmental temperature, this gene might be considered an "obvious" candidate marker to differentiate between cod from the two areas. Hence, failure to detect substructuring using this gene was perhaps unexpected. In this regard, Nielsen *et al*. (2009b) studied two additional hemoglobin genes in cod populations spanning the entire species distribution range and the findings suggested no evidence of local adaptation at these genes. Altogether, results from the current study together with results from others on *Pan*I (Nielsen *et al*. 2007) provided no evidence of genetic divergence between Faroe Plateau and Faroe Bank cod at the three most widely studied fitness-related genes for Atlantic cod (see Section 6.1.3). Based on this, there appears to be no clear evidence of adaptive genetic divergence between cod from the two areas.

The remaining genetic analyses in this body of work concerned microsatellite and SNP markers, the latter identified by RAD sequencing. Analyses of microsatellite data suggested significant, albeit low, genetic differentiation between Faroe Plateau and Faroe Bank cod (Faroe Bank – Faroe Plateau North-East: F_{ST} = 0.0014; Faroe Bank – Faroe Plateau West: F_{ST} = 0.0018), whereas approximately five-fold higher differentiation was computed, based on all 3533 SNP loci (Faroe Bank – Faroe Plateau North-East: F_{ST} = 0.0074). The difference between the levels of differentiation detected by the two marker systems is relatively small, however, if compared to locus-specific F_{ST} values between Faroe Plateau and Faroe Bank cod at the top five informative SNPs, though one of these loci (*Gmo*6654) showed zero differentiation for the Faroe Plateau North-East – Faroe Bank comparison. Excluding this single value, F_{ST} values between Faroe Plateau and Faroe Bank cod samples at these loci range from 0.0383 to 0.1914 and averaged at *c*. 0.1 (Figure 7.4). Accordingly, attempts to assign Faroe Plateau and Faroe Bank individuals and

groups of individuals to their population of origin were unsuccessful using microsatellite markers, but successful for group data using these informative SNP loci.

F_{ST} values calculated from highly variable markers, such as microsatellites, may potentially be downward biased, since the values depend on within-population variation (Section 2.4.3.1). In contrast, in the current work, F_{ST} values calculated from genome-wide SNP data may have been upward biased, as the SNPs used appeared to be markedly influenced by ascertainment bias, i.e. they were skewed towards detecting variation in the (Faroese) populations, from which they were discovered. A relatively new statistic, D_{JOST}, based on the number of effective microsatellite alleles, has been proposed to measure "actual population differentiation" (Jost 2008). Although careful consideration should be taken when comparing markers with different mutational characteristics (Meirmans and Hedrick 2011), it is interesting that this alternative estimate of differentiation suggested intermediate levels of differentiation between Faroe Plateau and Faroe Bank cod samples (Faroe Bank – Faroe Plateau North-East: D_{JOST_EST} = 0.0027; Faroe Bank – Faroe Plateau West: D_{JOST_EST} = 0.0048).

While the informative SNP panel revealed markedly higher levels of differentiation between Faroe Plateau and Faroe Bank cod samples than the total set of 3533 SNPs, this panel of SNP loci failed to detect significant genetic differentiation between Faroese and Norwegian cod samples. Adding to this were the intriguing findings of a relatively close association between White Sea and Faroe Bank cod based on these loci, as suggested by clustering analysis in STRUCTURE. Considering the microsatellite analyses, where both F_{ST} and D_{JOST_EST} agreed on *c.* seven-fold higher genetic differentiation between Faroese and Norwegian cod samples than between Faroe Plateau and Faroe Bank cod, it is obvious that these specifically selected SNPs should only be applied with caution in a broader geographical context.

SNP markers identified by RAD sequencing proved useful in population assignment testing of Faroese cod samples. However, the fact that of the many thousands of markers discovered, none were fully diagnostic, i.e. none were fixed for different alleles between the two areas, adds to the

overall picture of relatively minor genetic divergence between cod from the Faroe Plateau and the Faroe Bank.

8.1.2. Evidence from investigations of early life-history traits

In the current study, the significant genetic differentiation observed between Faroe Plateau and Faroe Bank cod could not be directly linked to functional mutations at specific coding gene loci, although the differentiation detected at some of the SNP markers may be linked to genes under selection (see discussion of *Gmo*6383 below). It is important to realise, however, that even small differentiation at neutral molecular markers can potentially camouflage adaptive differentiation between fish stocks, which may in turn confer resilience to changing conditions and be of importance for the long-term conservation of population complexes and species as a whole (Hilborn *et al*. 2003; Conover *et al*. 2006; Hauser and Carvalho 2008). An illustrative study is that of Hilborn *et al*. (2003) who investigated the sockeye salmon (*Oncorhynchus nerka*) stock complex in Bristol Bay, Alaska, which consists of several hundreds of spawning populations. This stock complex has experienced repeated environmental changes because of the El Niño Southern Oscillation as well as 50 – 70 year climate oscillations that characterise the North Pacific Basin. In response, the populations demonstrated wide variability in spawning habitat, time of spawning, egg size and adult body shape and size, amongst others. Furthermore, though individual populations did experience declines due to unfavourable conditions, the wide spectrum of life history strategies seemingly secured the overall productivity of the stock complex (Hilborn *et al*. 2003).

With increasing challenges from climate change as well as habitat disruption caused by human activity and invasive species, there is a need to understand the complex processes that lead to biocomplexity, or the biological structuring of populations (Hauser and Carvalho 2008), and marine literature on the topic is still relatively scant (Sotka 2012). While common garden experiments, where individuals from different populations / locations are reared under common environmental conditions, are the ultimate way to determine whether traits have a genetic component (Imsland and Jónsdóttir 2003), for many species, including Atlantic cod, they are

logistically difficult to carry out. Nonetheless, a few common garden studies have been used to study Atlantic cod populations. Otterå *et al*. (2006) found peak spawning to differ among groups of Norwegian coastal cod. Hutchings *et al*. (2007) conducted a common garden study of four North-West Atlantic cod populations, for which microsatellite analysis had failed to detect any significant structure, and revealed significant among-population differences in larval growth, survival and their reaction norms. Warm-water populations showed less tolerance to changes in food; cold-water populations showed less tolerance to changes in temperature (Hutchings *et al*. 2007). Important points drawn from these studies are that local adaptation may result in genetically-based differences in life history parameters among populations and those populations may thus respond differently to equal changes in the environment.

Studies with Faroe Plateau and Faroe Bank cod held under equal environmental conditions have revealed stock differences in growth (Fjallstein and Magnussen 1996) and fatty acid composition in heart tissue (Joensen *et al*. 2000).

In the current study, inferences of differences between Faroe Plateau and Faroe Bank cod in reproductive and early life-history traits were based on field data as well as experimental data from a laboratory set-up with controlled environmental conditions. Findings from the experimental data included smaller eggs and larvae as well as faster egg development rates for the Faroe Bank offspring. Field data suggested greater volumes spawned by the Faroe Bank females and an earlier peak spawning for the Faroe Plateau cod, an observation that was also supported by a few years of observations on cod spawning in captivity. Concerning the temporal stability of these observations, the field data included observations over 15 – 17 years, while the captive spawnings were only conducted for a single season. However, the observations on egg sizes were similar to those reported in a M.Sc. study in 1993 (Magnussen 1993) and could therefore suggest temporal stability.

The key finding from these investigations was that of divergent early life-history strategies between cod from the two areas with respect to the sizes and numbers of eggs produced,

presumably as an adaptive response to different predation intensities experienced in the early life stages. It was hypothesised that the strategy adopted by cod on the Faroe Bank, with a higher number of eggs, that were smaller, but developed faster, evolved in response to the more hostile environment (higher exposure to predators) experienced in early life stages in this area. This seems to agree with recent work, which suggests that Faroe Plateau year-class strength is determined during the second summer or winter rather than at the larval stage and cannibalism has been suggested as a controlling factor (Steingrund *et al*. 2010; ICES 2014a). In contrast, cannibalism is almost non-detectable amongst cod on the Faroe Bank, presumably because of the high growth rendering the cod as potential prey organisms for a limited time only (Magnussen 2011).

The differences found between the Faroe Plateau and Faroe Bank cod in egg sizes, egg development rates, larval sizes (at two weeks post-hatch) and spawning season likely had a significant genetic component, as both stocks were held under common conditions. The amount of eggs spawned per female, however, were only indirectly suggested by field data on condition decrease from spring to autumn. Hence, stock differences in this trait could not confidently be explained by genetic causes. However, the findings are in agreement with prevailing theories on early-life history traits, where an evolutionary determined trade-off between sizes and numbers of offspring is a central concept (for example Kamler 2005).

The hypothesis of smaller, but more numerous, eggs produced by the Faroe Bank females as an evolutionary response to a higher predation pressure appears in contrast to findings on sockeye salmon, where size-selective predation in the juvenile stage favours large eggs (West and Larkin 1987; Hilborn *et al*. 2003). However, while salmon have relatively large eggs compared to other teleost species (Hilborn *et al*. 2003), cod eggs approximate the physiological minimum size for viability, a proposed adaptive response to environments where egg sizes are of limited or no importance for survival in the early life stages (Hutchings 1997). Thus, different mechanisms are likely to operate in cod, such that the primary driver for enhanced fitness in hostile / exposed environments would likely be the number of eggs produced per female. Accordingly, although

findings on Faroe Bank progeny in the current work revealed a strong correlation between egg sizes and survival rates in the egg stage, egg sizes were not predictive of fry survival rates nor of fry body lengths (Chapter 4). Moreover, the hypothesis proposed for the different trade-offs in egg sizes and numbers for the Faroe Plateau and Faroe Bank cod are in agreement with observations on Trinidadian guppies, where populations exposed to higher predation pressures allocated more resources to reproduction. These populations reproduced more often, at the cost of progeny size, compared to populations that experienced less intense predation (Reznick and Endler 1982).

Although different exposure to predation during the early life stages was hypothesised as a primary ecological driver for the different early life-history strategies observed in Faroe Plateau and Faroe Bank cod (Chapter 3), alternative or additional explanations cannot be excluded at this point. Interesting research in this regard includes simulation studies of the spatiotemporal distribution of free-floating cod eggs in coastal Newfoundland waters. It appears that predation effects are only minimal and that the number of larvae that will eventually settle, i.e. the recruitment level, is mainly determined by an interplay between advective loss of free-floating life stages and temperatures which likely govern development times and, hence, the time spent as passive particles (Bradbury *et al*. 2000, 2001; Stanley *et al*. 2013). The Faroe Plateau and the Faroe Bank ecosystems are governed by gyres that promote local retention of early life stages of marine species (Section 1.6.1). Of the two areas, more research has been conducted on the Faroe Plateau, where studies have shown a clear link between ecosystem productivity and the physical processes that control the variable exchange rate between on- and off-shelf water systems. Thus, cod production appears to be closely linked to primary production, which is again related to the variable exchange rate between on- and off-shelf water systems (Steingrund and Gaard 2005; Debes *et al*. 2008; Larsen *et al*. 2008b). The spawning areas for cod on the Faroe Plateau are located close to the front between the two water masses; hence, there is a risk of advective loss of early free-floating stages (Gaard and Steingrund 2001). This could potentially influence cod recruitment, similar to that shown for other cod populations (see for example Bradbury *et al*. 2001

for cod in Newfoundland waters and Brickman *et al*. 2007 for Icelandic cod). Less is known about the link between physical processes and ecosystem productivity on the Faroe Bank. The horizontal exchange of on- and off-shelf water has been shown to be less pronounced for banks than for similar sized shelves surrounding islands (Eliasen *et al*. 2005). On the other hand, these authors also showed that the effect of the horizontal exchange on primary production increased as the system decreased in size. At depths shallower than 200 m, the size of the Faroe Bank is only *c*. 15 % of the size of the shelf surrounding the Faroe Islands (Eliasen *et al*. 2005). Thus, although no conclusions can be made at this point, in the case of a higher probability of advective loss of early life stages for Faroe Bank than for Faroe Plateau cod, this might provide an alternative explanation for the apparent evolution of a reproductive strategy for Faroe Bank cod, where smaller, but more numerous eggs, are produced.

8.1.3. Foundation for further investigations of adaptive traits

Despite the increasing awareness of the importance of preserving adaptive genetic variation of marine organisms, it remains a challenge to disentangle the effects of phenotypic plasticity and genetic processes, respectively. It is important to understand the interplay between genetic and environmental processes, especially when considering the threatened state of many marine stocks and the predicted climate change (for example Hauser and Carvalho 2008; Sotka 2012). For example, Drinkwater (2005) predicted a 2 – 3 ºC increase in sea water temperature for both the Faroe Plateau and the Faroe Bank by the year 2100 and based on species-specific limits of temperature, salinity and depths during feeding and spawning, a northwards distributional shift is expected for capelin, herring and cod amongst others (Rose 2005).

The microsatellite locus *Gmo*132 was found to harbour depth-related allele frequencies and could be an interesting candidate marker for future research into selective processes in cod stocks. However, it was also pointed out that the deviating allele frequency distributions frequently encountered at this locus (i.e. Dahle 1995; Bentzen *et al*. 1996; Nielsen *et al*. 2006) could potentially be explained by mutational mechanisms intrinsic to microsatellite sequences (Harr and Schlötterer 2000). Furthermore, no genes have been found in close proximity to *Gmo*132 (the

nearest annotated gene in Ensembl was located > 18,000 bp away; Section 5.3.2), i.e. it is unlikely that selection acts directly on this locus.

Bonin (2008) pointed out that despite a boom of genome-scale population genetic studies with ample success in detecting outlier loci (expected to be influenced by selective forces), it generally remained to discover the exact adaptive mutation(s) or gene(s). Bonin (2008) primarily explained this by the fact that most genome scans targeted marker types that were predominantly neutral, such as (most) SNPs or microsatellites, and thus were looking for a needle in a haystack. It follows that if the primary aim of an investigation is to identify potential adaptive mutations, an efficient way would be to look directly at the coding regions, for example by investigating SNPs developed from expressed sequence tags (ESTs) (Bonin 2008). Nielsen *et al*. (2009b), who looked at 98 gene-associated SNPs in Atlantic cod populations spanning the species distribution range, applied this strategy. They found strong statistical support for adaptive selection at eight SNPs. Further, they were able to identify a correlation between temperature and / or salinity (at time of spawning at the respective spawning locations) and variation at several of the outlier loci.

Although model-based tests should usually be applied to investigate potential selection on loci, interlocus comparisons of F_{ST} values from a large number of markers may provide a useful indication, particularly if comparisons are based on populations that inhabit different environments with limited gene flow in between them (Beaumont 2005; Storz 2005). The five SNP loci, verified in the current work to show significant genetic differentiation between Faroe Plateau and Faroe Bank cod (*Gmo*2357, *Gmo*2360, *Gmo*6383, *Gmo*6654 and *Gmo*6664; Section 7.3.3.4.2), were all among the loci for which outlier analysis suggested evidence of positive selection. However, considering the notes by Bonin (2008) and for example Barrett and Hoekstra (2011), that genes should not be termed “adaptive” unless a connection has been established among genotype, phenotype and fitness, only one of these SNPs may warrant further investigation. This is *Gmo*6383, which might potentially have some effect on protein level (as it was associated with a splice site in the cod genome) and, additionally, this marker showed exceptionally high genetic differentiation between Faroe Plateau and Faroe Bank cod. F_{ST} values

were 0.1914 and 0.1587 between the Faroe Bank and the Faroe Plateau West and Faroe Plateau North-East, respectively. As to the relative magnitude of these values, recent genome-scan studies of Atlantic cod could be consulted. In agreement with studies on other species (Nosil *et al*. 2009), genome-scans of Atlantic cod populations support the existence of "genomic islands of adaptive divergence", i.e. genomic regions with elevated levels of divergence compared to background (neutral) levels (Bradbury *et al*. 2013; Hemmer-Hansen *et al*. 2013; Karlsen *et al*. 2013). Whereas Karlsen *et al*. (2013) estimated an average neutral F_{ST} value of 0.0778, a much lower neutral F_{ST} value (0.0024) was indicated by Hemmer-Hansen *et al*. (2013). The F_{ST} values seen for *Gmo*6383 in this study were significantly higher than both of these values. The fact that this SNP was located in the *ryk* gene that appears to play a role in early egg development in zebrafish (Pelegri 2003) is intriguing, seen in light of the evidence of reproductive differences between Faroe Plateau and Faroe Bank cod, where sizes and numbers of eggs appear to be key variables. Hence, this gene could be an interesting candidate gene for further investigations of potential adaptive divergence between cod from the two areas.

8.2. Practical relevance of the results

The main conclusion of the thesis is that the Faroe Plateau and the Faroe Bank cod resolve as two genetically distinct units. This fits well with the wealth of information from tagging studies that show very limited migration among these two and other neighbouring cod populations (ICES 2006). The findings also corroborate findings by Steingrund (2009) who emphasised local processes as the likely cause for the large fluctuations in abundance seen for the Faroe Plateau cod in the nineties. Whereas this information in itself is of value to fisheries management bodies, the findings may also be of more applied use, as discussed below.

8.2.1. RAD technology and population assignment

Overexploitation of many of the world´s fisheries has prompted a consumer demand for eco labelling of fish products and regulations within the European Union (EU) now require certification of origin of all traded fish products. This has prompted a new area of research with focus on the development of genomic resources for traceability purposes. Within the EU funded

project FishPopTrace (https://fishpoptrace.jrc.ec.europa.eu), such resources were developed and validated for Atlantic cod, Atlantic herring, European hake and common sole (*Solea solea*; Nielsen *et al*. 2012). Traceability methods are also suited for monitoring (consequences of) escaped farmed fish (Glover *et al*. 2008, 2009) and data are now in place for a large part of the European salmon production, so that fish can be assigned to their river of origin with a high statistical power (Verspoor *et al*. 2012; ICES 2014b).

Together with research by others (Ogden *et al*. 2013), the present study emphasised RAD technology as a highly relevant approach in this regard. Even if investigations of candidate genes provided no evidence of genetic differentiation between Faroe Plateau and Faroe Bank cod and microsatellite markers revealed subtle differentiation only, RAD technology facilitated the simultaneous discovery and screening of SNP markers that were able to differentiate between groups of Faroe Plateau and Faroe Bank cod. Though these markers did not have the power to resolve population origin at the individual level, assignment of groups may suffice in some circumstances, for example if one were to inspect catches of fish vessels, where samples could consist of a given number of fish for each of a given number of boxes from the hold. This is particularly interesting, as a key concern regarding the management of the Faroe Bank cod is that landing estimates may have been uncertain from 1996 onwards, when fishing vessels were allowed to fish on both the Faroe Plateau and the Faroe Bank during the same trip (ICES 2014a).

In addition to the few SNPs validated in the current study, the extended RAD dataset is a valuable SNP resource, if further markers are needed, for example to meet potential future demands requiring knowledge of population of origin of Faroese cod products.

8.2.2. Implications for a potential future cod aquaculture industry in the Faroe Islands

Short after the start of the current project, initiatives to breed cod in the Faroe Islands ceased, partly due to problems relating to the semi-intensive method of production, but also due to low market prices and no prospect of buyers for the juvenile cod that were produced (Arge 2011). In 2013, the remaining broodstock was culled, hence suspending any further investigation into

aquaculture production of cod. In the meantime, the Atlantic salmon farming industry has increased tremendously in the Faroe Islands; it is now responsible for over 40 % of the total export value, and it takes up most of the protected sea areas, rendering it unlikely that further initiatives to develop cod farming will be conducted in near future (Arge 2011; www.hagstova.fo, accessed 14 October 2014). Nonetheless, potential future initiatives may choose to resume breeding of Faroe Bank cod in the natural environment of the Faroe Plateau cod and it is important to have a tool set ready that can be used to monitor possible genetic introgression between native fish and escapees from fish farms (ICES 2014b).

Escaped farmed fish are generally acknowledged as a threat to native populations as evidenced from extensive research on Atlantic salmon, where the level of introgression appears to be negatively correlated to overall fitness (McGinnity *et al*. 2009; ICES 2014b). The level of genetic introgression depends on a variety of factors, such as number and life-stage of the escaped fish and the level of interbreeding between native and farmed fish. In addition, the genetic background of the escaped fish is important, as it defines the potential fitness of the farmed fish in a natural environment. If maladapted genes, resulting from unintended domestication selection and potentially selective breeding of the farmed fish, enter the wild gene pool, the overall fitness of the native fish may be reduced. Such detrimental effects may be buffered by the relative size of the native population to the escapees, but effects are populations-specific and, at present, difficult to predict (ICES 2014b).

Whereas most fish farm escapes occur because of technical and operational failure of equipment, cod actively search for holes in the net and have a distinct net biting behaviour. In addition, cod are likely to reproduce in sea cages and may escape in the form of fertilised eggs (Section 1.4.3; Jensen *et al*. 2010; Jørstad *et al*. 2008, 2014). Jørstad *et al*. (2008) monitored a Norwegian fjord system, where 1000 genetically tagged cod (specifically bred to be homozygous for a rare allele at an isozyme locus) were allowed to spawn in a sea cage. Eggs from these cod were detected up to 10 km away from the cage and a subsequent nearby plankton survey found that 20 – 25 % of larvae originated from the farmed cod. Jørstad *et al*. (2014) raised genetically tagged cod from

two year-classes in sea-cages under full-scale commercial conditions and discovered otherwise unnoticed episodes of escapes. The escapes spread through the whole fjord system as well as to a neighbouring fjord system and they were found in relatively high proportions at local spawning sites. Moreover, the farmed cod contributed offspring to the fjord system. Though interbreeding between wild and farmed cod could not be precluded, it could not be inferred from these data and the overall conclusion was that a limited number of escape events was unlikely to significantly alter the genetic architecture of the native population.

In the case of potential genetic introgression between Faroe Plateau and Faroe Plateau cod, a further risk to the ones discussed above should be considered, namely that these are two genetically distinct populations that appear to have evolved different reproductive strategies in response to different environments. Again, the consequences of potential introgression cannot be predicted and the Faroe Plateau offspring, which are larger and spawn earlier in the spring, could outcompete the Faroe Bank offspring. On the other hand, Faroe Bank offspring could have an advantage merely due their high numbers. Regardless, potential genetic impact on the native population should always be minimised and monitored, as emphasised in a recent report of the International Council for the Exploration of the Sea (ICES) and part of the tool set is to collect baseline knowledge on the native gene pools, prior to establishment of fish farms (ICES 2014b). Therefore, RAD data from the current work should be of high practical use, if Faroese cod farming is to resume in the future.

8.2.2.1. A good egg from an aquaculture perspective – cues from reproductive strategies of wild populations

The question asked by Brooks *et al*. (1997) about what makes a good egg is still of relevance to the aquaculture industry, whose main objectives include the production of a stable supply of viable eggs (Lubzens *et al*. 2010). This has proven particularly challenging to the cod farming industry due to intrinsically high and unpredictable mortality rates of early life stages of marine fish (Section 1.4.3.1). Although improved rearing protocols have reduced the problem somewhat and research into the development of micro particulate diets for the early life stages could be

promising, more research is needed (Kjesbu *et al*. 2006; Puvanendran *et al*. 2006; Hamre *et al*. 2013). Various indicators of egg quality have been explored in a variety of aquaculture species (egg morphology, biochemical composition and early cleavage patterns), yet there are many unanswered questions concerning the assessment and making of a good and viable egg (Hansen and Puvanendran 2010; Lubzens *et al*. 2010).

The lack of a correlation between survival in egg and fry stages of Faroe Bank cod (Chapter 4) agrees with a review by Lambert *et al*. (2003), who collected information from field and laboratory studies of the most important marine fish stocks of the North Atlantic Ocean. These authors found that the significant relationships seen between maternal, egg and larval characteristics and egg and larval viability rarely persisted into later life stages (Lambert *et al*. 2003). Part of the explanation may lie in the fact that expression levels of maternal genes, deposited in marine eggs, ebb away during early embryonic development (Pelegri 2003). Future insight into factors involved in early development of fish is expected from functional studies of the role of candidate genes and maternally inherited mRNA (Pelegri 2003; Lubzens *et al*. 2010; Drivenes *et al*. 2012).

The present work reports apparently consistent differences between Faroe Plateau and Faroe Bank cod in early life-history parameters and provides the hypothesis that Faroe Bank eggs are produced in higher numbers at a cost of the viability of each single egg (Chapter 3). Further research is required to confirm the hypothesis and to investigate the ecological drivers behind these reproductive differences. Yet, the findings suggest that knowledge of reproductive strategies in the wild may provide valuable information concerning evaluations of which stock(s) to base an aquaculture population on. Due to the remarkably high growth rate, breeding initiatives in the Faroe Islands have concentrated on Faroe Bank cod (1.6.3). However, realising that the critical bottleneck in marine aquaculture is the production of a stable supply of juveniles, it could be hypothesised that potential future farming of cod in the Faroe Islands would benefit from using Faroe Plateau cod as broodstock, since each offspring may potentially have inherently higher chances of survival compared to Faroe Bank cod. Potential future cod farming initiatives might

even gain from using crosses between Faroe Plateau females (larger eggs) and Faroe Bank males (higher growth rate). Conclusions concerning this would be premature though.

There is a general lack of detailed information on teleost oogenesis. *Gmo*6383 was implied here as a candidate locus, involved in early egg development. Moreover, recent studies have identified a large number of candidate genes for reproduction (Hemmer-Hansen *et al*. 2011; Drivenes *et al*. 2012). Drivenes *et al*. (2012) used a cod cDNA microarray to investigate transcriptional profiles of 7,000 genes at various stages during embryogenesis and identified a large number of candidate genes involved in early development, from the oocyte stage to hatched larvae that were ready to start feeding. It would be particularly interesting for future studies to seek to understand the genetic basis of the apparent differences between Faroe Plateau and Faroe Bank cod in reproductive and early life history strategies and these already available SNP resources could provide a promising starting point towards the ultimate goal of linking important phenotypic traits to genotypes. Further analysis of reproductive traits could also involve mapping F2 families from Faroe Plateau × Faroe Bank crosses with RAD sequencing to identify quantitative trait loci.

8.3. Broader implications of the findings

8.3.1. Future research to explore potential stock differences at candidate genes for growth

The broad public interest of potential genetic substructuring of Faroe Plateau and Faroe Bank cod originates mainly from the differences in growth rates between the two. The current study did confirm genetic divergence between cod from the two areas, though no diagnostic "growth genes" were detected. However, in a similar manner as suggested above concerning the genetic basis of early development, available SNP resources could be examined in relation to growth. Interesting candidate SNPs are available from Hemmer-Hansen *et al*. (2011), who described 30 SNPs in 18 candidate genes for growth and reproduction in Atlantic cod. The populations used for SNP discovery covered the species range and included Faroese cod (though not split into Faroe Plateau and Faroe Bank cod) and putative gene function was inferred from literature as well as the National Center for Bioinformatics (NCBI) Entrez Gene database.

8.3.2. Candidate SNP(s) for sex-differentiation in cod

Aquaculture production of cod (and other species as well) is challenged by early maturation (Section 1.4.3) and current research aims at understanding the sex-determining mechanisms in cod (Johnsen *et al*. 2010; Haugen *et al*. 2012; Johnsen and Andersen 2012; Johnsen *et al*. 2013). The current study was not designed specifically to detect sex-linked markers. However, an apparent association was found between SNPs, detected by RAD sequencing, and phenotypic sex. Although these associations should be further validated, one of the identified SNPs (*Gmo*11037) might potentially be an interesting candidate for sex-differentiation in cod. This SNP aligned to a gene (*dpy30*) that is implied in dosage compensation and sex-differentiation in other species (Hsu and Meyer 1994; Petty *et al.* 2011; Dheilly *et al.* 2012; de Sousa *et al.* 2014).

8.3.3. Reproductive parameters and stock assessment methods

Preceding sections have discussed the possible significance of different early life-history strategies for Faroe Plateau and Faroe Bank cod in relation to aquaculture. There is also an increasing awareness concerning the importance of monitoring reproductive characteristics of fishery stocks, as this may allow better estimation of stock reproductive potential (Tomkiewicz *et al*. 2003). Fecundity is an important parameter in this regard (Lambert *et al*. 2003). Thus, findings here of seemingly different relationships between fecundity and individual progeny viability of Faroe Plateau and Faroe Bank cod might become of relevance to fishery scientists, who are currently testing out stock assessment methods for the Faroe Plateau stock, that take the productivity of the ecosystem into account (ICES 2014a).

8.3.4. Monitoring genetic change

Despite marked reductions in census size of both the Faroe Plateau and the Faroe Bank cod, microsatellite analysis in this work detected no evidence of genetic bottlenecks. The smaller Faroe Bank stock should be more prone to loss of genetic variation and it could be intriguing to postulate that loss of genetic variation in this stock may be counteracted by a relatively high level of fecundity. However, as discussed in Chapter 5, it might be too early to detect potential bottlenecks based on genetic data in this work. The latest stock assessments show that the low

levels of both stocks have persisted since sampling for this work was conducted (ICES 2014a); hence, updated estimates of genetic diversity might be warranted and genetic marker data from the current study should provide useful baseline data.

8.4. Conclusions

Work in this thesis presents multiple lines of evidence demonstrating that Faroe Plateau and Faroe Bank cod are two genetically distinct populations. As demonstrated by microsatellite analyses, genetic differences between the two are small compared to genetic differences between Faroese and surrounding cod populations. RAD technology proved a valuable tool for confirming significant genetic differentiation between Faroe Plateau and Faroe Bank cod. Additionally, based on RAD data, a panel of five SNP loci was identified that could be used to assign groups of Faroe Plateau and Faroe Bank cod to their population of origin. Analysis of wild and captive spawning cod and their progeny indicated functional genetic differences between Faroe Plateau and Faroe Bank cod. These differences were seen in the form of divergent early life-history strategies where Faroe Bank cod appear to have maximised their reproductive output at the cost of egg and larval sizes.

Reference List

Alcivar-Warren, A., Dunham, R., Gaffney, P., Kocher, T. and Thorgaard, G. (1997). First aquaculture species genome mapping workshop. *Animal Genetics* **28**, 451-452. doi: 10.1111/j.1365-2052.1997.00202.x.

Andersen, Ø., De Rosa, M. C., Pirolli, D., Tooming-Klunderud, A., Petersen, P. E. and André, C. (2011). Polymorphism, selection and tandem duplication of transferrin genes in Atlantic cod (*Gadus morhua*) – Conserved synteny between fish monolobal and tetrapod bilobal transferrin loci. *BMC Genetics* **12**(51), pp. 1-14. doi: 10.1186/1471-2156-12-51.

Andersen, Ø., Wetten, O., De Rosa, M. C., Andre, C., Carelli Alinovi, C., Colafranceschi, M., Brix, O. and Colosimo, A. (2009). Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations. *Proceedings of the Royal Society of London B* **276**, 833-841. doi: 10.1098/rspb.2008.1529.

Anderson, J. L., Laurel, B. J. and Brown, J. A. (2007). Diel changes in behaviour and habitat use by age-0 Atlantic cod (*Gadus morhua* L.) in the laboratory and field. *Journal of Experimental Marine Biology and Ecology* **351**, 267-275. doi: 10.1016/j.jembe.2007.07.001.

Á Nordi, G. and Poulsen, M. (2000). Føde og fødevalg hos torske- og kulleryngel på Færøplateauet og Færøbanke. B.Sc. Thesis. University of Faroe Islands, 42 pp.

Antao, T., Lopes, A., Lopes, R: J., Beja-Pereira, A. and Luikart, G. (2008). LOSITAN: a workbench to detect molecular adaptation based on a F_{st}-outlier method. *BMC Bioinformatics* **9**, 323, pp. 1-5. doi: 10.1186/1471-2105-9-323.

Arge, R. (2011). Færøyene. In: L. Grønnevet and F. Sigurdsson (Eds.), Nord-Atlantisk Fersk Torsk i Europa – Muligheter og Utfordringer. Project no. 510-033, The Nordic Atlantic Cooperation (NORA). http://www.nora.fo/fileadmin/user_upload/files/13/20121219085535160.pdf. Accessed 19 May 2014. Pp. 72-82.

Armitage, D., Maguire, J., Treasurer, J., and Cross, T. F. (2007). Parental assignment in farmed cod (*Gadus morhua*) progeny in mass spawning using microsatellite DNA loci reveals major disparity in family representation. Genetics in Aquaculture IX, *Aquaculture* **272**(Suppl. 1), S242. doi: 10.1016/j.aquaculture.2007.07.026.

Árnason, E. (2004). Mitochondrial cytochrome *b* DNA variation in the high-fecundity Atlantic cod: trans-Atlantic clines and shallow gene genealogy. *Genetics* **166**, 1871-1885.

Árnason, E. and Pálsson, S. (1996). Mitochondrial cytochrome *b* DNA sequence variation of Atlantic cod, *Gadus morhua*, from Norway. *Molecular Ecology* **5**, 715-724.

Árnason, E., Petersen, P. H. and Pálsson, S. (1998). Mitochondrial cytochrome *b* DNA sequence variation of Atlantic cod, *Gadus morhua*, from the Baltic and the White Seas. *Hereditas* **129**, 37-43.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M. and Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**, 25-29. doi: 10.1038/75556.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. and Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489-522. doi: 10.1146/annurev.es.18.110187.002421.

Bagenal, T. B. (1969). Relationship between egg size and fry survival in brown trout *Salmo trutta*. *Journal of Fish Biology* **1**, 349-353.

Bagshaw, A. T. M., Pitt, J. P. W. and Gemmell, N. J. (2008). High frequency of microsatellites in *S. Cerevisiae* meiotic recombination hotspots. *BMC Genetics* **9**(49), pp. 1-11. doi: 10.1186/1471-2164-9-49.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, A., Selker, E. U., Cresko, W. A. and Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS One* **3**(10), e3376, pp. 1-7. doi: 10.1371/journal.pone.0003376.

Barrett, R. D. H. and Hoekstra, H. E. (2011). Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics* **12**, 767-780. doi: 10.1038/nrg3015.

Barton, N. H. and Slatkin, M. (1986). A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Journal of [* **56**, 409-415. doi: 10.1038/hdy.1986.63.

Baxter, S. W., Davey, J. W., Johnston, J. S., Shelton, A. M., Heckel, D. G., Jiggins, C. D. and Blaxter, M. L. (2011). Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLOS One* **6**(4), e19315, pp. 1-11. doi: 10.1371/journal.pone.0019315.

Beacham, T. D., Brattey, J., Miller, K. M., Le, K. D. and Withler, R. E. (2002). Multiple stock structure of Atlantic cod (*Gadus morhua*) off Newfoundland and Labrador determined from genetic variation. *ICES Journal of Marine Science* **59**, 650-665. doi: 10.1006/jmsc.2002.1253.

Beaumont, M. A. (2005). Adaptation and speciation: what can F_{ST} tell us? *TRENDS in Ecology and Evolution* **20**(8), 435-440. doi: 10.1016/j.tree.2005.05.017.

Beaumont, M. A. and Nichols, A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London B* **263**(1377), 1619-1626. doi: 10.1098/rspb.1996.0237.

Bekkevold, D., André, C., Dahlgren, T. G., Clausen, L. A. W., Thorstensen, E., Mosegaard, H., Carvalho, G. R., Christensen, T. B., Norlinder, E. and Ruzzante, D. R. (2005). Environmental correlates of population differentiation in Atlantic herring. *Evolution* **59**(12), 2656-2668. doi: 10.1111/j.0014-3820.2005.tb00977.x.

Bekkevold, D., Hansen, M. M. and Loeschcke, V. (2002). Male reproductive competition in spawning aggregations of cod (*Gadus morhua*, L.). *Molecular Ecology* **11**, 91-102. doi: 10.1046/j.0962-1083.2001.01424.x.

Bentzen, P., Taggart, C. T., Ruzzante, D. E. and Cook, D. (1996). Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the Northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences* **53**, 2706-2721. doi: 10.1139/f96-238.

Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., Bento, P., Da Silva, C., Labadie, K., Alberti, A., Aury, J.-M., Louis, A., Dehais, P., Bardou, P., Montfort, J., Klopp, C., Cabau, C., Gaspin, C., Thorgaard, G. H., Boussaha, M., Quillet, E., Guyomard, R., Galiana, D., Bobe, J., Volff, J.-N., Genêt, C., Wincker, P., Jaillon, O., Crollius, H. R. and Guiguen, Y. (2014). The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications* **5**, e3657, pp. 1-10. doi: 10.1038/ncomms4657.

Björnsson, B., Steinarsson, A. and Árnason, T. (2007). Growth model for Atlantic cod (Gadus morhua): Effects of temperature and body weight on growth rate. *Canadian Journal of Fisheries and Aquatic Sciences* **59**, 494-502. doi: 10.1016/j.aquaculture.2007.06.026.

Björnsson, B., Steinarsson, A. and Oddgeirsson, M. (2001). Optimal temperature for growth and feed conversion of immature cod (*Gadus morhua* L.). *ICES Journal of Marine Science* **58**, 29-38. doi: 10.1006/jmsc.2000.0986.

Bonin, A. (2008). Population genomics: a new generation of genome scans to bridge the gap with functional genomics. *Molecular ecology* **17**, 3583-3584. doi: 10.1111/j.1365-294X.2008.03854.x.

Borg, Å., Phil, L. and Wennhage, H. (1997). Habitat choice by juvenile cod (*Gadus morhua* L.) on sandy soft bottoms with different vegetation types. *Helgoländer Meeresuntersuchungen* **51**, 197-212.

Borisov, V. M. (1978). The selective effect of fishing on the population structure of species with a long life cycle. *Journal of Ichthyology* **18**, 896-904.

Borrell, Y. J., Álvarez, J., Vásquez, E., Pato, C. F., Tapia, C. M., Sánchez, J. A. and Blanco, G. (2004). Applying microsatellites to the management of farmed turbot stocks (*Scophthalmus maximus* L.) in hatcheries. *Aquaculture* **241**, 133-150. doi: 10.1016/j.aquaculture.2004.08.021.

Borza, T., Stone, C., Gamperl, A. K. and Bowman, S. (2009). Atlantic cod (*Gadus morhua*) hemoglobin genes: multiplicity and polymorphism. *BMC Genetics* **10**(51), pp. 1-14. doi: 10.1186/1471-2156-10-51.

Braaten, B. (1984). Growth of cod in relation to fish size and ration level. In: E. Dahl, D. S. Danielssen, E. Moksnes and P. Solemdal (Eds.), *The Propagation of Cod*. Arendal, Norway: Flødevigen Rapportserie 1, pp. 677-710.

Bradbury, I. R., Hubert, S., Higgins, B., Bowman, S., Borza, T., Paterson, I. G., Snelgrove, P. V. R., Morris, C. J., Gregory, R. S., Hardie, D., Hutchings, J. A. Ruzzante, D. A., Taggart, C. T. and Bentzen, P. (2013). Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. *Evolutionary Applications* **6**(3), 450-461. doi: 10.1111/eva.12026.

Bradbury, I. R., Hubert, S., Higgins, B., Bowman, S., Paterson, I. G., Snelgrove, P. V. R., Morris, C. J., Gregory, R. S., Hardie, D. C., Borza, T. and Bentzen, P. (2011). Evaluating SNP ascertainment bias and its impact on population assignment in Atlantic cod, *Gadus morhua*. *Molecular Ecology resources* **11**(Suppl. 1), 218-225. doi: 10.1111/j.1755-0998.2010.02949.x.

Bradbury, I. R., Snelgrove, P. V. R. and Fraser, S. (2000). Transport and development of eggs and larvae of Atlantic cod, *Gadus morhua*, in relation to spawning time and location in coastal Newfoundland. *Canadian Journal of Fisheries and Aquatic Sciences* **57**(9), 1761-1772. doi: 10.1139/f00-122.

Bradbury, I. R., Snelgrove, P. V. R. and Fraser, S. (2001). The influence of temperature on advective loss of Atlantic cod (*Gadus morhua*) eggs from the inshore environment. *Fisheries Oceanography* **10**(4), 342-352. doi: 10.1046/j.1365-2419.2001.00179.x.

Brander, K. M. (1994). Patterns of distribution, spawning, and growth in North Atlantic cod: the utility of inter-regional comparisons. *ICES Marine Science Symposia* **198**, 406-413.

Brander, K. M. (1995). The effect of temperature on growth of Atlantic cod (*Gadus morhus* L.). *ICES Journal of Marine Science* **52**, 1-10. doi: 10.1016/1054-3139(95)80010-7.

Brander, K. M. (1997). Effects of climate change on cod (*Gadus morhua*) stocks. In: C.M. Wood and D.G. McDonald (Eds.), *Global warming: implications for freshwater and marine fish*. Cambridge, UK: Cambridge Unviersity Press, pp. 255-278.

Brander, K. (2005). Spawning and life history information for North Atlantic cod stocks. ICES cooperative research report no. 274. Copenhagen, Denmark: International Council for the Exploration of the Sea, pp. 1-152.

Brickman, D., Marteinsdóttir, G., Logemann, K. and Harms, I. H. (2007). Drift probabilities for Icelandic cod larvae. *ICES Journal of Marine Science* **64**, 49-59. doi: 10.1093/icesjms/fsl019.

Bricknell, I. R., Bron, J. E. and Bowden, T. J. (2006). Diseases of Gadoid fish in cultivation: a review. *ICES Journal of Marine Science* **63**, 253-266. doi: 10.1016/j.icesjms.2005.10.009.

Brix, O., Forås, E. and Strand, I. (1998). Genetic variation and functional properties of Atlantic cod hemoglobins: introducing a modified tonometric method for studying fragile hemoglobins. *Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology* **119**(2), 575-583. doi: 10.1016/S1095-6433(97)00469-8.

Brix, O., Thorkildsen, S., and Colosimo, A. (2004). Temperature acclimation modulates the oxygen binding properties for the Atlantic cod (*Gadus morhua* L.) genotypes–*HbI*1/1*, *HbI*1/2*, and *HbI*2/2*–by changing the concentrations of their major hemoglobin components (results from growth studies at different temperatures). *Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology* **138**(2), 241-251. doi: 10.1016/j.cbpb.2004.04.004.

Brodziak, J., Cadrin, S. X., Legault, C. M. and Murawski, S. A. (2008). Goals and strategies for rebuilding New England groundfish stocks. *Fisheries Research* **94**, 355-366. doi: 10.1016/j.fishres.2008.03.008.

Bromage, N. (1995). Broodstock management and seed quality-general considerations. In: N. Bromage and R. J. Roberts (Eds.), *Broodstock Management and Egg and Larval Quality*. Oxford, UK: Blackwell, pp. 1-24.

Bromage, N., Porter, M. and Randall, C. (2001). The environmental regulation of farmed finfish with special reference to the role of photoperiod and melanin. *Aquaculture* **197**, 63-98.

Brooker, A. L., Cook, D., Bentzen, P., Wright, J. M. and Doyle, R. W. (1994). Organization of microsatellites differs between mammals and cold-water teleost fishes. *Canadian Journal of Fisheries and Aquatic Sciences* **51**, 1959-1966.

Brooks, S., Tyler, C. R. and Sumpter, J. P. (1997). Egg quality in fish: what makes a good egg? *Reviews in Fish Biology and Fisheries* **7**, 387-416. doi: 10.1023/A:1018400130692.

Buren, A. D., Koen-Alonso, M., Pepin, P., Mowbray, F., Nakashima, B., Stenson, G., Ollerhead, N., Montevecchi, W. A. (2014). Bottom-up regulation of capelin, a keystone forage species. *PLOS One* **9**(2), e87589, pp. 1-11. doi: 10.1371/journal.pone.0087589.

Callen, D. F., Thompson, A. D., Shen, Y., Phillips, H. A., Richards, R. I., Mulley, J. C. and Sutherland, G. R. (1993). Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics* **52**, 922-927.

Canino, M. F. and Bentzen, P. (2004). Evidence for positive selection at the pantophysin (*Pan* I) locus in walleye pollock, *Theragra chalcogramma*. *Molecular Eiology and Evolution* **21**, 1391-1400. doi: 10.1093/molbev/msh141.

Canino, M. F., O´Reilly, P. T., Hauser, L. and Bentzen, P. (2005). Genetic differentiation in walleye pollock (*Theragra chalcogramma*) in response to selection at the pantophysin (*Pan* I) locus. *Canadian Journal of Fisheries and Aquatic Sciences* **62**, 2519-2529. doi: 10.1139/F05-155.

Carmichael, S. N., Bekaert, M., Taggart, J. B., Christie, H. R. L., Bassett, D. I., Bron, J. E., Skuce, P. J., Gharbi, K., Skern-Mauritzen, R. and Sturm, A. (2013). Identification of a sex-linked SNP marker in the salmon louse (*Leleophtheirus salmonis*) using RAD sequencing. *PLOS One* **8**(10), e77832, pp. 1-8. doi: 10.1371/journal.pone.0077832.

Carr, S. M. and Marshall, H. D. (1991a). Detection of intraspecific DNA sequence variation in the mitochondrial cytochrome *b* gene of Atlantic cod (*Gadus morhua*) by the polymerase chain reaction. *Canadian Journal of Fisheries and Aquatic Sciences* **48**(1), 48-52.

Carr, S. M. and Marshall, H. D. (1991b). A direct approach to the measurement of genetic variation in fish populations: applications of the polymerase chain reaction to studies of Atlantic cod, *Gadus morhua* L. *Journal of Fish Biology* **39**(Suppl. A), 101-107. doi: 10.1111/j.1095-8649.1991.tb05073.x.

Case, R. A., Hutchinson, W. F., Hauser, L., Buehler, V., Clemmesen, C., Dahle, G., Kjesbu, O. S., Moksness, E., Otterå, H., Paulsen, H., Svåsand, T., Thorsen, A. and Carvalho, G. R. (2006). Association between growth and *Pan I** genotype within Atlantic cod full-sibling families. *Transactions of the American Fisheries Society* **135**, 241-250. doi: 10.1577/T05-130.1.

Case, R. A. J., Hutchinson,W. F., Hauser, L., Van Oosterhout, C. and Carvalho, G. R. (2005). Macro- and micro-geographic variation in pantophysin (*Pan*I) allele frequencies in NE Atlantic cod *Gadus morhua*. *Marine Ecology Progress Series* **301**, 267-278.

Castro, J., Bouza, C., Presa, P., Pino-Querido, A., Riaza, A., Ferreiro, I., Sánchez, L. and Martínez, P. (2004). Potential sources of error in parentage assessment of turbot (*Scophthalmus maximus*) using microsatellite loci. *Aquaculture* **242**, 119-135. doi: 10.1016/j.aquaculture.2004.09.010.

Catchen, J., Hohenlohe, , P. A., Bassham, S., Amores, A. and Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology* **22**(11), 3124-3140. doi: 10.1111/mec.12354.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. and Postlethwait, J. P. (2011). Stacks: Building and genotyping loci *de novo* from short-read sequences. *Genes, Genomes, Geneticsk* **1**(3), 171-182. doi: 10.1534/g3.111.000240.

Cavalli-Sforza, L. L. and Edwards, W. F. (1967). Phylogenetic analysis: Models and estimation procedures. *American Journal of Human Genetics* **19**, 233-257.

Chambers, R. C. (1993). Phenotypic variability in fish populations and its representation in individual-based models. *Transactions of the American Fisheries Society* **122**, 404-414. doi: 10.1577/1548-8659(1993)122<0404:PVIFPA>2.3.CO;2.

Chambers, R. C. (1997). Environmental influences on egg and propagule sizes in marine fishes. In: R. C. Chambers and E. A. Trippel (Eds.), *Early Life History and Recruitment in Fish Populations*. New York, N. Y.: Chapman and Hall, pp. 63-102.

Chambers, R. C. and Waiwood, K. (1996). Maternal and seasonal differences in egg sizes and spawning characteristics of captive Atlantic cod, *Gadus morhua*. *Canadian Journal of Fisheries and Aquatic Sciences* **53**, 1986-2003.

Chelala, C., Khan, A. and Lemoine, N. R. (2008). SNPnexus: a web database for functional annotation of newly discovered and public domain single nucleotide polymorphisms. *Bioinformatics* **25**, 655-661. doi: 10.1093/bioinformatics/btn653.

Chiasson, M., Benfey, T. J. and Martin-Robichaud, D. J. (2008). Gonadal differentiation in Atlantic cod, *Gadus morhua* L., and Haddock, *Melanogrammus aeglefinus* (L.). *Acta Ichthyologica et Piscatoria* **38**(2), 127-133. doi: 10.3750/AIP2008.38.2.07.

Chilcote, M. W., Leider, S. A. and Loch, J. J. (1986). Differential reproductive success of hatchery and wild summer-run steelhead under natural conditions. *Transactions of the American Fisheries Society* **115**, 726-735.

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

Clausen, L. A. W., Bekkevold, D., Hatfield, E. M. C. and Mosegaard, H. (2007). Application and validation of otolith microstructure as a stock identification method in mixed herring (*Clupea harengus*) stocks in the North Sea and western Baltic. *ICES Journal of Marine Science* **64**(2), 377-385. doi: 10.1093/icesjms/fsl036.

Codling, E. A., Kelly, C. J. and Clarke, M. (2005). Comparison of the effects of exploitation on theoretical long-lived fish species with different life history strategies and the implications for management. ICES Annual Science Conference, (ICES CM 2005/N:24), Aberdeen, Netherlands: International Council for the Exploration of the Sea, 16 pp.

Cohen, D. M., Inada, T. and Scialabba, N. (1990). Gadiform fishes of the world: an annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to date. *FAO Fisheries Synopsis* **10**(125), pp 1-422.

Conover, D. O., Clarke, L. M., Munch, S. B. and Wagner, G. N. (2006). Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology* **69**(Suppl. sc), 21-47. doi: 10.1111/j.1095-8649.2006.01274.x.

Cornuet, J.-M. and Luikart, G. (1996). Detection and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Creel, S., Spong, G., Sands, J.L., Rotella, J., Zeigle, J., Joe, L., Murphy, K.M. and Smith, D. (2003). Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. *Molecular Ecology* **12**, 2003-2009. doi: 10.1046/j.1365-294X.2003.01868.x.

Cross, T. F., Bailey, J., Friars, G. and O'Flynn, F. (1993). Maintenance of genetic variability in reared Atlantic salmon (*Salmo salar*) stocks. In: D. Mills (Ed.), *Salmon in the sea and new enhancement strategies*. Oxford, UK: Fishing News Books, pp. 356-366.

Cross, T. F. and Payne, R. H. (1978). Geographic variation in Atlantic cod, *Gadus morhua*, off eastern North America: a biochemical systematics approach. *Journal of the Fisheries Research Board of Canada* **35**, 117-123.

Csirke, J. (1980). Recruitment in the Peruvian anchovy and its dependence on the adult population. *Rapports et Proces-Verbaux des Reunions - Conseil International pour l'Exploration de la Mer* **177**, 307-313.

Cury, P. and Roy, C. (1989). Optimal environmental window and pelagic fish recruitment success in upwelling areas. *Canadian Journal of Fisheries and Aquatic Sciences* **46**, 670-680.

Cushing, D. H. (1975). Marine ecology and fisheries. Cambridge, UK: Cambridge University Press, 278 pp.

Dahle, G. (1991). Cod, *Gadus morhua* L., populations identified by mitochondrial DNA. *Journal of Fish Biology* **38**, 295-303.

Dahle, G. (1995). Genetic Structure of the North-East-Atlantic cod (*Gadus morhua* L.), an appraisal of different molecular techniques. Ph.D. Thesis, University of Bergen.

Dahle, G. and Jørstad, K. E. (1993). Haemoglobin variation in cod – a reliable marker for Arctic cod (*Gadus morhua* L.). *FisheriesRresearch* **16**, 301-311. doi: 10.1016/0165-7836(93)90143-U.

Dahle, G., Jørstad, K. E., Rusaas, H. E. and Otterå, H. (2006). Genetic characteristics of broodstock collected from four Norwegian coastal cod (*Gadus morhua*) populations. *ICES Journal of Marine Science* **63**, 209-215. doi: 10.1016/j.icesjms.2005.10.015.

Daníelsdóttir, A. K., Gíslason, D., Kristinsson, K., Stefánsson, M., Johansen, T. and Pampoulie, C.

Population structure of deep-sea and oceanic phenotypes of deepwater redfish in the Irminger Sea and Icelandic Continental Slope: are they cryptic species? (2008). *Transactions of the American Fisheries Society* **137**(6), 1723-1740. doi: 10.1577/T07-240.1.

Danzmann, R. G. and Gharbi, K. (2007). Mapping in aquaculture species. In: Z. Liu (Ed.), *Aquaculture Genome Technologies*. 1st edn. IOWA, USA: Blackwell Publishing, pp. 139-167.

Davie, A. (2005). Effects of photoperiod manipulation on growth and reproduction in Atlantic cod (*Gadus morhua* L.) Ph.D. Thesis. University of Stirling, 298 pp.

Davey, J. W. and Blaxter, M. L. (2011). RADSeq: next-generation population genetics. *Briefings in Functional Genomics* **9**(5), 416-423. doi: 10.1093/bfgp/elq031.

Davis, C. P., King, J. L., Budowle, B., Eisenberg, A. J. And Turnbough, M. A. (2012). Extraction platform evaluations: a comparison of Automate Express™, EZ1® Advanced XL, and Maxwell® 16 Bench-top DNA extraction systems. *Legal Medicine* **14**(1), 36-39. doi: 10.1016/j.legalmed.2011.09.005.

Debes, H., Gaard, E. and Hansen, B. (2008). Primary production on the Faroe Shelf: temporal variability and environmental influences. *Journal of Marine Systems* **74**(1-2), 686-697. doi: 10.1016/j.jmarsys.2008.07.004.

Delghandi, M., Mortensen, A. and Westgaard, J.-I. (2003). Simultaneous analysis of six microsatellite markers in Atlantic cod (*Gadus morhua*): a novel multiplex assay system for use in selective breeding studies. *Marine Biotechnology* **5**, 141-148. doi: 10.1007/s10126-002-0086-x.

Denovan-Wright, E. M., Ramsey, N. B., McCormick, C. J., Lazier, C. B. and Wright, J. M. (1996). Nucleotide sequence of transferrin cDNAs and tissue-specific expression of the transferrin gene in Atlantic cod (*Gadus morhua*). *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **113**(2), 269-273.

De-Santis and Jerry, D. R. (2007). Canditate growth genes in finfish-Where should we be looking? *Aquaculture* **272**, 22-38. doi: 10.1016/j.aquaculture.2007.08.036.

De Sousa, J. T., Milan, M., Bargelloni, L., Pauletto, M., Matias, D., Joaquim, S., Matias, A. M., Quillien, V., Leitão, A. and Huvet, A. A microarray-based analysis of gametogenesis in two Portuguese populations of the European clam *Ruditapes decussatus* (2014). *PLOS One* **9**(3), e92202, pp. 1-12. doi: 10.1371/journal.pone.0092202.

de Souza, P. C. and Bonilla-Rodriguez, G. O. (2007). Fish hemoglobins. *Brazilian Journal of Medical and Biological Research* **40**, 769-778.

Devlin, R. H. and Nagahama, Y. (2002). Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* **208**(3), 191-364. doi: 10.1016/S0044-8486(02)00057-1.

Dheilly, N. M., Lelong, C., Huvet, A., Kellner, K., Dubos, M.-P., Riviere, G., Boudry, P. and Favrel, P. (2012). Gametogenesis in the Pacific oyster *Crassostrea gigas*: a microarray-based analysis identifies sex and stage-specific genes. *PLOS One* **7**(5), e36353, pp. 1-15. doi: 10.1371/journal.pone.0036353.

Di Rienzo, A., Peterson, J., Garza, J. C., Valdes, A. M., Slatkin, M. and Freimer, N. B. (1994). Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America* **91**(8), 3166-3170. doi: 10.1073/pnas.91.8.3166.

Doyle, R. W., Herbinger, C., Taggart, C. T. and Lochmann, S. (1995). Use of DNA microsatellite polymorphism to analyze genetic correlations between hatchery and natural fitness. *American Fisheries Society Symposium* **15**, 205-231.

Drinkwater, K. F. (2005). The response of Atlantic cod (*Gadus morhua*) to future climate changes. *ICES Journal of Marine Science* **62**, 1327-1337. doi: 10.1016/j.icesjms.2005.05.015.

Drinkwater, K. (2008). Comparison of the response of Atlantic cod (*Gadus morhua*) in the high latitude regions of the North Atlantic during the warm periods of the 1920s-1960s and the 1990s-2000s. *Deep Sea Research Part II: Topical Studies in Oceanography* **56**(21-22), 2087-2096. doi: 10.1016/j.dsr2.2008.12.001.

Drivenes, Ø., Taranger, G. L. and Edvardsen, R. B. (2012). Gene expression profiling of cod (*Gadus*

morhua) embryogenesis using microarray. *Marine Biotechnology* **14**(2), 167-176. doi: 10.1007/s10126-011-9399-y.

Duchesne, P., Godbout, M.-H. and Bernatchez, L. (2002). PAPA (package for the analysis of parental allocation): a computer program for simulated and real parental allocation. *Molecular Ecology Notes* **2**, 191-193. doi: 10.1046/j.1471-8286.2002.00164.x.

Dunham, R. A. (2004). Biochemical and Molecular Markers. In: R. A. Dunham (Ed.), *Acuaculture and Fisheries Biotechnologies. Genetic Approaches*. Oxfordshire, UK: Cabi Publishing, pp. 85-103.

Dutil, J.-D. and Brander, K. (2003). Comparing productivity of North Atlantic cod (*Gadus morhua*) stocks and limits to growth production. *Fisheries Oceanography* **12**, 502-512. doi: 10.1046/j.1365-2419.2003.00243.x.

Earl, D, A. and vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources* **2**(2), 359-361. doi: 10.1007/s12686-011-9548-7.

Eiríksson, G. M. and Árnason, E. (2013). Spatial and temporal microsatellite variation in spawning Atlantic cod, *Gadus morhua*, around Iceland. *Canadian Journal of Fisheries and Aquatic Sciences* **70**(8), 1151-1158. doi: 10.1139/cjfas-2012-0494.

Eknath, A. E. and Doyle, R. W. (1990). Effective population size and rate of inbreeding in aquaculture of Indian major carps. *Aquaculture* **85**, 293-305.

Elgar, M. A. (1990). Evolutionary compromise between a few large and many small eggs: comparative evidence in teleost fish. *Oikos* **59**, 283-287.

Eliasen, S. K., Gaard, E., Hansen, B. and Larsen, K. M. (2005). A "horizontal Sverdrup mechanism" may control the spring bloom around small oceanic islands and over banks. *Journal of Marine Systems* **56**(3-4), 352-362. doi: 10.1016/j.jmarsys.2005.03.005.

Estoup, A., Gharbi, K., SanCristobal, M., Chevalet, C., Haffray, P. and Guyomard, R. (1998). Parentage assignment using microsatellites in turbot (*Schophthalmus maximus*) and rainbow trout (*Oncorhynchus mykiss*) hatchery populations. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 715-725.

Etter, P. D., Bassham, S., Hohenlohe, P. A., Johnson, E. A. and Cresko, W. A. (2011). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods in Molecular Biology* **772**, 157-178. doi: 10.1007/978-1-61779-228-1_9.

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620. doi: 10.1111/j.1365-294X.2005.02553.x.

Excoffier, L., Hofer, T. and Foll, M. (2009). Detecting loci under selection in a hierarchically structured population, *Heredity* **103**, 285-298. doi: 10.1038/hdy.2009.74.

Excoffier, L. and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology resources* **10**(3), 564-567. doi: 10.1111/j.1755-0998.2010.02847.x.

Falconer, D. S. and MacKay, T. F. C. (1996a). Genetic constitution of a population. In: D. S. Falconer and T. F. C. MacKay (Eds.), *Introduction to Quantitative Genetics*. Essex, UK: Pearson Education Limited, pp. 1-22.

Falconer, D. S. and MacKay, T. F. C. (1996b). Heritability. In: D. S. Falconer and T. F. C. MacKay (Eds.), *Introduction to Quantitative Genetics.* Essex, UK: Pearson Education Limited, pp. 160-183.

Falush, D., Stephens, M. and Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7**(4), 574-578. doi: 10.1111/j.1471-8286.2007.01758.x.

Fan, H. and Chu, J.-Y. (2007). A brief review of short tandem repeat mutation. *Genomics, Proteomics and Bioinformatics* **5:1**, 7-14.

Felip, A., Zanuy, S., Carrillo, M. and Piferrer, F. (2001). Induction of triploidy and gynogenesis in teleost fish with emphasis on marine species. *Genetica* **111**, 175-195.

Fernández-Chacón, R. and Südhof, T. C. (1999). Genetics of synaptic vesicle function: toward the complete

functional anatomy of an organelle. *Annual Review of Physiology* **61**, 753-776. doi: 10.1146/annurev.physiol.61.1.753.

Fevolden, S. E. and Pogson, G. H. (1995). Differences in nuclear DNA RFLPs between Norwegian coastal and the Northeast Arctic population of Atlantic cod. In: H.R. Skjoldahl, C. Hopkins, K. E. Erikstad and H. P. Leinaas (Eds.), *Ecology of Fjords and Coastal Waters, Proceedings of the Mare Nor Symposium on the Ecology of Fjords and Coastal Waters*, Tromsø, Norway, pp. 745-760.

Fevolden, S. E. and Pogson, G. H. (1997). Genetic divergence at the synaptophysin (*Syp* I) locus among Norwegian coastal and north-east Arctic populations of Atlantic cod. *Journal of Fish Biology* **51**, 895-908.

Fjalestad, K. T., Fevolden, S.-E., Jørstad, K. and Olesen, I. (2006). Avl og genetikk – nye artar. In: M. Thomassen, R. Gudding, B. Norberg and L. Jørgensen (Eds.), Havbruksforskning: Fra Merd til Mat. Oslo, Norway: Norges Forskningsråd, pp. 270-285.

Fjallstein, I. and Magnussen, E. (1996). Growth of Atlantic cod (*Gadus morhua* L.) of Faroe Bank strain and Faroe Plateau strain in captivity. *ICES. Mariculture Committee*, ICES CM 1996/F:12, 16 pp.

Flicek, P., Amode, R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Girón, C. G., Gordon, L., Hourlier, T., Hunt, S., Johnson, N., Juetteman, T., Kähäri, A. K., Keenan, S., Kulesha, E., Martin, F. J., Maurel, T., MacLaren, W. M., Murphy, D. N., Nag, R., Overduin, B., Pignatelli, M., Pritchard, B., Pritchard, E., Riat, H. S., Ruffier, M., Sheppard, D., Taylor, K., Thormann, A., Trevanion, S. J., Vullo, A., Wilder, S. P., Wilson, M., Zadissa, A., Aken, B. L., Birney, E., Cunningham, F., Harrow, J., Herrero, J., Hubbard, T. J. P., Kinsella, R., Muffato, M., Parker, A., Spudich, G., Yates, A., Zerbino, D. R. and Searler, S. M. J. (2014). Ensembl 2014. *Nucleic Acids Research* **42**(D1), D749-D755. doi: 10.1093/nar/gkt1196.

Flick, W. A. and Webster, D. A. (1964). Comparative first year survival and production in wild and domestic strains of brook trout, *Salvalinus fontinalis*. *Transactions of the American Fisheries Society* **93**, 58-69.

Flick, W. A. and Webster, D. A. (1976). Production of wild, domestic, and interstrain hybrids of brook trout (*Salvelinus fontinalis*) in natural ponds. *Journal of Fisheries Research Board of Canada* **33**, 1525-1539.

Ford, M. J., Thornton, P. J. and Park, L. K. (1999). Natural selection promotes divergence of transferrin among salmonid species. *Molecular Ecology***,** 8(6), 1055-1061. doi: 10.1046/j.1365-294x.1999.00651.x.

Frank, K. T., Drinkwater, K. F. and Page, F. H. (1994). Possible causes of recent trends and fluctuations in Scotian Shelf/Gulf of Maine cod stocks. *International Commission for the Exploration of the Seas (ICES) Marine Science Symposia* **198**, 110-120.

Frankham, R., Ballou, J. D. and Briscoe, D. A. (2007). *Introduction to Conservation Genetics*. 7th edn. Cambridge, UK: Cambridge University Press, pp. 1-617.

Fraser, J. M. (1981). Comparative survival and growth of planted wild, hybrid, and domestic strains of brook trout (*Salvelinus fontinalis*) in Ontario lakes. *Canadian Journal of Fisheries and Aquatic Sciences* **38**, 1672-1684.

Freeman, S. and Herron, J. C. (2007). Evolution at Multiple Loci: Quantitative Genetics. In: S. Freeman and J. C. Herron (Eds.), *Evolutionary Analysis*, 4th edn. San Francisco, California, Benjamin Cummings pp. 319 - 359.

Frydenberg, O., Møller, D., Nævdal, G. and Sick, K. (1965). Haemoglobin polymorphism in Norwegian cod populations. *Hereditas* **53**, 255-271.

Fu, C. H. and Fanning, L. P. (2004). Spatial considerations in the management of Atlantic cod off Nova Scotia, Canada. *North American Journal of Fisheries Management* **24**(3), 775-784. doi: 10.1577/M03-134.1.

Fyhn, U. E. H., Brix, O., Nævdal, G., and Johansen, T. (1994). New variants of the haemoglobins of Atlantic cod: a tool for discriminating between coastal and Arctic cod populations. *ICES Marine Science Symposia* **198**, 666-670.

Fyhn, U. E. H., Johansen, T., Nævdal, G. and Thorkildsen, S. (1995). Ontogeny of heamoglobin polymorphism in cod. *Journal of Fish Biology* **47**, 56-60.

Gaard, E. and Mortensen, H. (1993). Phyto- and zooplankton communities on the Faroe Bank and their relations to the physical and chemical environment. *ICES CM L* **54**, 1-17.

Gaard, E. and Steingrund, P. (2001). Reproduction of Faroe Plateau cod: spawning grounds, egg advection and larval feeding. *Fródskaparrit* **48**, 87-103.

Gagne, J. A. and O´Boyle, R. N. (1984). The timing of cod spawning on the Scotian Shelf. In: E. Dahl, D. S. Danielssen, E. Moksness and P. Solemdal (Eds.), *The propagation of cod Gadus morhua L. An International Symposium*. Institute of Marine Research, Flødevigen Biological Station, Arendal, Norway, pp. 501-517.

Gallagher G. (2003). The future of cod. *Aquaculture Magazine* **29**, 26-33.

Galvin, P., Sadusky, T., McGregor, D. and Cross, T. (1995). Population genetics of Atlantic cod using amplified single locus minisatellite VNTR analysis. *Journal of Fish Biology* **47**(Suppl. A), 200-208. doi: 10.1111/j.1095-8649.1995.tb06056.x.

Gamble, T. and Zarkower, D. (2014). Identification of sex-specific molecular markers using restriction site-associated DNA sequencing. *Molecular Ecology Resources* **14**, 902-913. doi: 10.1111/1755-0998.12237.

Ganz, T. (2009). Iron in innate immunity: starve the invaders. *Current Opinion in Immunology* **21**(1), 63-67. doi: 10.1016/j.coi.2009.01.011.

Garza, J. C. and Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10**, 305-318. doi: 10.1046/j.1365-294X.2001.01190.x.

Geffen, A. J., Fox, C. J. and Nash, R. D. M. (2006). Temperature-dependent development rates of cod *Gadus morhua* eggs. *Journal of Fish Biology* **69**, 1060-1080. doi: 10.1111/j.1095-8649.2006.01181.x.

Geffen, A. J. and Nash, R. D. M. (2012). Egg development rates for use in egg production methods (EPMs) and beyond. *Fisheries Research* **117-118**, 48-62. doi: 10.1016/j.fishres.2011.01.002.

Ghigliotti, L., Fevolden, S.-E., Cheng, C.-H. C., Babiak, I., Dettai, A. and Pisano, E. (2012). Karyotyping and cytogenetic mapping of Atlantic cod (*Gadus morhua* Linnaeus, 1758). *Animal Genetics* **43**(6), 746-752. doi: 10.1111/j.1365-2052.2012.02343.x.

Gjedrem T. (1997). Selective breeding to improve aquaculture production. *World Aquaculture* **28**, 33-45.

Gjedrem, T. (2000). Genetic improvement of cold-water fish species. *Aquaculture Research* **31**, 25-33. doi: 10.1046/j.1365-2109.2000.00389.x.

Gjedrem, T. (2005). Breeding plan. In: T. Gjedrem (Ed.) *Selection and Breeding Programs in Aquaculture*. Dordrecht, Netherlands: Springer, pp. 251-277.

Gjedrem, T. and Baranski, M. (2009). The Success of Selective Breeding in Aquaculture. In: T. Gjedrem and M. Baranski (Eds.), *Selective Breeding in Aquaculture: An Introduction*, 1st edn. London, UK: Springer, pp. 13-23.

Gjedrem, T. and Olesen, I. (2005). Basic statistical parameters. In: T. Gjedrem (Ed.) *Selection and breeding programs in aquaculture*. Dordrecht, Netherlands: Springer, pp. 45-72.

Gjedrem, T., Robinson, N. and Rye, M. (2012). The importance of selective breeding in aquaculture to meet future demands for protein: a review. *Aquaculture* **350-353**, 117-129. doi: 10.1016/j.aquaculture.2012.04.008.

Gjedrem, T. and Thodesen, J. (2005). Selection. In: T. Gjedrem (Ed.) *Selection and Breeding Programs in Aquaculture*. Dordrecht, Netherlands: Springer, pp. 89-111.

Glover, K. A., Hansen, M. M. and Skaala, Ø. (2009). Identifying the source of farmed escaped Atlantic salmon (*Salmo salar*): Bayesian clustering analysis increases accuracy of assignment. *Aquaculture* **290**(1-2), 37-46. doi: 10.1016/j.aquaculture.2009.01.034.

Glover, K. A., Skilbrei, O. T. and Skaala, Ø. (2008). Genetic assignment identifies farm of origin for Atlantic salmon *Salmo salar* escapees in a Norwegian fjord. *ICES Journal of Marine Science* **65**(6), 912-920. doi: 10.1093/icesjms/fsn056.

Glover, K. A., Taggart, J. B., Skaala, Ø. and Teale, A. J. (2004). A study of inadvertent domestication selection during start-feeding of brown trout families. *Journal of Fish Biology* **64**, 1168-1178. doi:

10.1111/j.0022-1112.2004.00376.x.

Godø, O. R. (1984). Migration, mingling and homing of north-east Arctic cod from two separated spawning grounds. In: O. R. Godø and S. Tilseth (Eds.), *Reproduction and Recruitment of Arctic cod*. Institute of Marine Research Bergen, Norway, pp. 289-302.

Gordo, L. S., Costa, A., Abaunza, P., Lucio, P., Eltink, A. T. G. W. and Figueiredo, I. (2008). Determinate versus indeterminate fecundity in horse mackarel. *Fisheries Research* **89**, 181-185. doi: 10.1016/j.fishres.2007.09.024.

Gotceitas, V., Fraser, S. and Brown, J. A. (1995). Habitat use by juvenile Atlantic cod (*Gadus morhua*) in the presence of an actively foraging and non-foraging predator. *Marine Biology* **123**, 421-430.

Goudet, J. (1999). PCA-GEN for windows. Version 1.2. Institute of Ecology, University of Lausanne, Lausanne, Switzerland.

Greene, C. W. (1952). Results from stocking brook trout of wild and hathcery strains at Stillwater Pond. *Transactions of the American Fisheries Society* **81**, 43-52.

Gunnarsli, K. S., Toften, H. and Mortensen, A. (2009). Effects of nitrogen gas supersaturation on growth and survival in larval cod (*Gadus morhua* L.). *Aquaculture* **288**, 344-348. doi: 10.1016/j.aquaculture.2008.06.029.

Halliday, R. G. and Pinhorn, A. T. (1982). The groundfish resource in the Gulf of St. Lawrence. *Canadian Technical Report of Fisheries and Aquatic Sciences* **1086**, 16 pp.

Hamilton, L. C., Colocousis, C. R. and Johansen, S. T. F. (2004). Migration from resource depletion: the case of the Faroe Islands. *Society and Natural Resources* **17**, 443-453. doi: 10.1080/08941920490430232.

Hamre, K. (2006). Nutrition in cod (*Gadus morhua*) larvae and juveniles. *ICES Journal of Marine Science* **63**, 267-274. doi: 10.1016/j.icesjms.2005.11.011.

Hamre, K., Yúfera, M., Rønnestad, I., Boglione, C., Conceição, L. E. C. and Izquierdo, M. (2013). Fish larval nutrition and feed formulation: knowledge gaps and bottlenecks for advances in larval rearing. *Reviews in Aquaculture* **5**(Suppl. 1), S26-S58. doi: 10.1111/j.1753-5131.2012.01086.x.

Hansen, B. (1992). Residual and tidal currents on the Faroe Plateau. *ICES Hydrography Committee*, ICES CM 1992/C:12, 18 pp.

Hansen, B. (2000). Havid. Tórshavn, Faroe Islands: Føroya Skúlabókagrunnur, pp. 1-232.

Hansen, B., Ellett, D. and Meldrum, D. (1986). Evidence for an anticyclonic circulation on Faroe Bank. ICES CM002FC:15, 15 pp.

Hansen, G. H. and Olafsen, J. A. (1999). Bacterial interactions in early life stages of marine cold water fish. *Microbial Ecology* **38**, 1-26. doi: 10.1007/s002489900158.

Hansen, L. A., Damsgård, B. and Delghandi, M. (2004). Spawning behaviour and reproductive success in cod, *Gadus morhua* L. In: ICES Conference. Book of Abstracts: *Gadoid mariculture: development and future challenges*, p. 41.

Hansen, Ø. J. and Puvanendran, V. (2010). Fertilization success and blastomere morphology as predictors of egg and juvenile quality for domesticated Atlantic cod, *Gadus morhua*, broodstock. *Aquaculture Research* 41, 1791-1798. doi: 10.1111/j.1365-2109.2010.02506.x.

Hansen, T., Karlsen, Ø., Taranger, G. L., Hemre, G.-I., Holm, J. C. and Kjesbu, O. S. (2001). Growth, gonadal development and spawning time of Atlantic cod (*Gadus morhua*) reared under different photoperiods. *Aquaculture* **203**, 51-67.

Harden-Jones, F. R. (1968). Fish migration. London, UK: Edward Arnold Ltd., London, UK, 325 pp.

Harden-Jones, F. R. and Scholes, P. (1974). The effect of low temperature on cod, *Gadus morhua*. *Journal du Conseil - Conseil International pour l´Exploration de la Mer* **35**(3), 258-271. doi: 10.1093/icesjms/35.3.258.

Harr, B. and Schlötterer, C. (2000). Long microsatellite alleles in *Drosophila melanogaster* have a downward mutation bias and short persistence times, which cause their genome-wide underrepresentation. *Genetics* **155**, 1213-1220.

Haugen, T., Almeida, F. F. L., Andersson, E., Bogerd, J., Male, R., Skaar, K. S., Schulz, R. W., Sørhus, E.,

Wijgerde, T. and Taranger, G. L. (2012). Sex differentiation in Atlantic cod (*Gadus morhua* L.): morphological and gene expression studies. *Reproductive Biology and Endocrinology* **10**(47), pp. 1-13. doi: 10.1186/1477-7827-10-47.

Hauser, L. and Carvalho, G. R. (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries* **9**, 333-362.

Hauser, L. and Seeb, J. E. (2008). Advances in molecular technology and their impact on fisheries genetics. *Fish and Fisheries* **9**(4), 473-486. doi: 10.1111/j.1467-2979.2008.00306.x.

Havenstein, G. B., Ferket, P. R. and Qureshi, M. A. (2003). Carcass composition and yield of 1957 versus 2001 broilers when fed representative 1957 and 2001 broiler diets. *Poultry Science* **82**, 1509-1518.

Heath, M. R., Culling, M. A., Crozier, W. W., Fox, C. J., Curney, W. S. C., Hutchinson, W. F., Nielsen, E. E., O'Sullivan, M., Preedy, K. F., Righton, D. A., Speirs, D. C., Taylor, M. I., Wright, P. J. and Carvalho, G. R. (2014). Combination of genetics and spatial modelling highlights the sensitivity of cod (*Gadus morhua*) population diversity in the North Sea to distributions of fishing. *ICES Journal of Marine Science* **71**(4), 794-807. doi: 10.1093/icesjms/fst185.

Hedrick, P. G. (2005). A standardized genetic differentiation measure. *Evolution* **59**(8), 1633-1638. doi: 10.1111/j.0014-3820.2005.tb01814.x.

Heller, R. and Siegismund, R. (2009). Relationship between three measures of genetic differentiation G_{ST}, D_{EST} and G'_{ST}: how wrong have we been. *Molecular Ecology* **18**, 2080-2083. doi: 10.1111/j.1365-294X.2009.04185.x.

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., Cariani, A., Maes, G. E., Diopere, E., Carvalho, G. R. and Nielsen, E. E. (2011). Application of SNPs for population gnetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources* **11**(Suppl. 1), 123-136. doi: 10.1111/j.1755-0998.2010.02943.x.

Hemmer-Hansen, J., Nielsen, E. E., Frydenberg, J. and Loeschcke, V. (2007). Adaptive divergence in a high gene flow environment: *Hsc70* variation in the European flounder (*Platichthys flesus* L.). *Journal of Heredity* **99**, 592-600. doi: 10.1038/sj.hdy.6801055.

Hemmer-Hansen, J., Nielsen, E. E., Meldrup, D. and Mittelholzer, C. (2011). Identification of single nucleotide polymorphisms in candidate genes for growth and reproduction in a nonmodel organism; the Atlantic cod, *Gadus morhua*. *Molecular Ecology Resources* **11**(Supple.1), 71-80. doi: 10.1111/j.1755-0998.2010.02940.x.

Hemmer-Hansen, J., Nielsen, E. E., Therkildsen, N. O., Taylor, M. I., Ogden, R., Geffen, A. J., Bekkevold, D., Helyar, S., Pampoulie, C., Johansen, T., Fishpoptrace Consortium and Carvalho, G. R. (2013). A genomic island linked to ecotype divergence in Atlantic cod. *Molecular Ecology* **22**, 2653-2667. doi: 10.1111/mec.12284.

Herlin, M. C. G. (2008). Genetic management of Atlantic cod (*Gadus morhua* L.) hatchery populations. Ph.D. Thesis, University of Stirling, 223 pp.

Herlin, M., Delghandi, M., Wesmajervi, M., Taggart, J. B., McAndrew, B. J. and Penman, D. J. (2008). Analysis of the parental contribution to a group of fry from a single day of spawning from a commercial Atlantic cod (*Gadus morhua*) breeding tank. *Aquaculture* **274**, 218-224. doi: 10.1016/j.aquaculture.2007.11.034.

Herlin, M., Taggart, J. B., McAndrew, B. J. and Penman, D. J. (2007). Parentage allocation in a complex situation: a large commercial Atlantic cod (*Gadus morhua*) spawning tank *Aquaculture* **272**(Suppl. 1), S195-S203. doi: 10.1016/j.aquaculture.2007.08.018.

Hershberger, W. K. (2006). What can history tell us about the future of aquaculture genetics. *The Israeli Journal of Aquaculture - Bamidgeh* **58**, 223-229.

Hilborn, R., Quinn, T. P., Schindler, D. E. and Rogers, D. E. (2003). Biocomplexity and fisheries sustainability. *Proceedings of the National Academy of Sciences of the United States of America* **100**(11), 6564-6568. doi: 10.1073/pnas.1037274100.

Hjort, J. (1914). Fluctuations in the great fisheries off northern Europe viewed in the light of biological research. Rapports et Procs-Verbaux des Reunions *Journal du Conseil - Conseil International pour l'Exploration de la Mer* **20**, 1-228.

Hoffman, J.I. and Amos, W. (2005). Microsatellite genotyping errors: detection approaches, common

sources and consequences for paternal exclusion. *Molecular Ecology* **14**, 599-612. doi: 10.1111/j.1365-294X.2004.02419.x.

Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W. and Luikart, G. (2011). Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources* **11**(Suppl. 1), 117-122. doi: 10.1111/j.1755-0998.2010.02967.x.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A. and Cresko, W. A. (2010). Population genomics of parallel adaptation in Threespine Stickleback using sequenced RAD tags. *PloS Genetics* **6**, e1000862, pp. 1-23. doi: 10.1371/journal.pgen.1000862.

Hohenlohe, P. A., Day, M. D., Amish, S. J., Miller, M. R., Kamps-Hughes, N., Boyer, M. C., Muhlfeld, C. C., Allendorf, F. W., Johnson, E. A. and Luikart, G. (2013). Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology* **22**, 3002-3013. doi: 10.1111/mec.12239.

Holm, P., Starkey, D. J. and Thór, J. T. (1996). The North Atlantic fisheries, 1100-1976. National perspectives on a common resource. Studia Atlantica 1, Fiskeri- og Søfartsmuseet, Esbjerg, Denmark, pp. 1-209.

Holström, C. and Kjelleberg, S. (1999). Marine *Pseudoalteromonas* species are associated with higher organisms and produce biologically active extracellular agents. *FEMS Microbiology Ecology* **30**, 185-193. doi: 10.1111/j.1574-6941.1999.tb00656.x.

Houston, R. D., Davey, J. W., Bishop, S. C., Lowe, N. R., Mota-Velasco, J. C., Hamilton, A., Guy, D. R., Tinch, A. E., Thomson, M. L., Blaxter, M. L., Gharbi, K., Bron, J. E. and Taggart, J. B. (2012). Characterisation of QTL-linked and genome-wide restriction site-associated DNA (RAD) markers in farmed Atlantic salmon. *BMC Genetics* **13**(244), pp. 1-15. doi: 10.1186/1471-2164-13-244.

Hsu, D. R. and Meyer, B. J. (1994). The *dpy-30* gene encodes an essential component of the *Caenorhabditis elegans* dosage compensation machinery. *Genetics* **137**, 999-1018.

Huang, C.-M. and Liao, I.-C. (1990). Response to mass selection for growth rate in *Oreochromis niloticus*. *Aquaculture* **85**, 199-205. doi: 10.1016/0044-8486(90)90019-J.

Hubert, S., Higgins, B., Borza, B. and Bowman, S. (2010). Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*). *BMC Genetics* **11**(191), pp. 1-14. doi: 10.1186/1471-2164-11-191.

Hulata, G., Wohlfarth, G. W. and Halevy, A. (1986). Mass selection for growth rate in the Nile tilapia (*Oreochromis niloticus*). *Aquaculture* **57**, 177-184. doi: 10.1016/0044-8486(86)90195-X.

Hutchings, J. A. (1997). Life history responses to environmental variability in early life. In: R. C. Chambers and E. A. Trippel (Eds.), *Early Life History and Recruitment in Fish Populations*. London, UK: Chapman and Hall, pp. 139-168.

Hutchings, J. A. (2004). The cod that got away. *Nature* **428**, 899-900. doi: 10.1038/428899a.

Hutchings, J. A., Bishop, T. D. and McGregor-Shaw, C. R. (1999). Spawning behaviour of Atlantic cod (*Gadus morhua*: evidence of mate competition and mate choice in a broadcast spawner. *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 97-104.

Hutchings, J. A. and Reynolds, J. D. (2004). Marine fish population collapses: consequences for recovery and extinction risk. *BioScience* **54**(4), 297-309. doi: 10.1641/0006-3568(2004)054[0297:MFPCCF]2.0.CO;2.

Hutchings, J. A., Swain, D. P., Rowe, S., Eddington, J. D., Puvanendran, V. and Brown, J. A. (2007). Genetic variation in life-history reaction norms in a marine fish. *Proceedings of the Royal Society of London B* **274**, 1693-1699. doi: 10.1098/rspb.2007.0263.

Hutchinson, W. F., Carvalho, G. R. and Rogers S. I. (2001). Marker genetic structuring in localised spawning populations of cod *Gadus morhua* in the North Sea and adjoining waters, as revealed by microsatellites. *Marine Ecology Progress Series* **223**, 251-260.

Huybrechts, P. (2002). Sea-level changes at the LGM from ice-dynamic reconstructions of the Greenland and Antarctic Ice sheets during the glacial cycles. *Quaternary Science Reviews* **21**, 203-231. doi: 10.1016/s0277-3791(01)00082-8.

ICES (2006). Report of the North-Western Working Group (NWWG). ICES CM 2006/ACFM:26.

ICES (2012). Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM). ICES CM 2012/SSGHIE:12.

ICES (2014a). Report of the North-Western Working Group (NWWG). ICES CM 2014/ACOM:07.

ICES (2014b). Report of the Working Group on the Application of Genetics in Fisheries and Mariculture (WGAGFM). ICES CM 2014/SSGHIE:13.

í Jákupsstovu, S. H. and Reinert, J. (1994). Fluctuations in the Faroe Plateau cod stock. *ICES Marine Science Symposia* **198**, 194-211.

Imelfort, M., Duran, C., Batley, J. and Edwards, D. (2009). Discovering genetic polymorphisms in next-generation sequencing data. *Plant Biotechnology Journal* **7**(4), 312-317. doi: 10.1111/j.1467-7652.2009.00406.x.

Imsland, A. K., Foss, A., Nævdal, G., Johansen, T., Folkvord, A., Stefansson, S. O. and Jonassen, T. M. (2004). Variations in growth in haemoglobin genotypes of Atlantic cod. *Fish Physiology and Biochemistry* **30**, 47-55.

Imsland, A. K., and Jónsdóttir, Ó. D. B. (2002). Is there a genetic basis to growth in Atlantic cod. *Fish and fisheries* **3**, 36-52.

Imsland, A. K. and Jónsdóttir, Ó. D. B. (2003). Linking population genetics and growth of Atlantic cod. *Reviews in Fish Biology and Fisheries* **13**, 1-26. doi: 10.1023/A:1026373509576.

Jalving, R., van´t Slot, R. and van Oost, B. A. (2004). Chicken single nucleotide polymorphism identification and selection for genetic mapping. *Poultry Science* **83**, 1925-1931. doi: 10.1093/ps/83.12.1925.

Jamieson, A. (1967). New genotypes in cod at Greenland. *Nature* **215**, 661-662. doi: 10.1038/215661a0.

Jamieson, A., 1975. Enzyme type of Atlantic cod stocks on the North American banks. In: C. L. Markert (Ed.) *Isozyme*s. IV. Genetics and evolution. New York, USA: Academic Press, pp. 491-515.

Jamieson, A. and Birley, A. J. (1989). The demography of haemoglobin polymorphism in the Atlantic cod, *Gadus morhua* L. *Journal of Fish Biology* **35** 193-204.

Jamieson, A. and Jones, B. W. (1967). Two races of cod at Faroe. *Journal of Heredity* **22**, 610-612.

Jamieson, A. and Otterlind, G. (1971). The use of cod blood protein polymorphisms in the Belt Sea, the Sound and the Baltic Sea. ICES Special Meeting on the Biochemical and Serological Identification of Fish Stocks. Rapports et Proces-verbaux des Reunions - Conseil International pour lÉxploration de la mer, Dublin, Ireland, pp. 55-59.

Jamieson, A. and Thompson, D. (1972). Butyric esterase differences in Atlantic cod, *Gadus morhua* L. *ICES*, C.M 1972/F:32, 5 pp.

Jensen, A. C. (1972). The cod; the uncommon history of a common fish and its impact on American life from Viking times to present. Toronto, Canada: Fitzhenry and Whiteside Ltd., 182 pp.

Jensen, Ø., Dempster, T., Thorstad, E. B., Uglem, I. and Fredheim, A. (2010). Escapes of fishes from Norwegian sea-cage aquaculture: causes, consequences and prevention. *Aquaculture Environment Interactions* **1**, 81-83. doi: 10.3354/aei00008.

Jerry, D. R., Pruwis, I. W. and Piper, L. R. (2001). Opportunities for genetic improvement in crustacean species. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* **14**, 55-58.

Joensen, H. and Grahl-Nielsen, O. (2004). Stock structure of *Sebastes mentella* in the North Atlantic revealed by chemometry of the fatty acid profile in heart tissue. *ICES Journal of Marine Science* **61**, 113-126. doi: 10.1016/j.icesjms.2003.10.002.

Joensen, H., Steingrund, P., Fjallstein, I. and Grahl-Nielsen, O. (2000). Discrimination between two reared stocks of cod (*Gadus morhua*) from the Faroe Islands by chemometry of fatty acid composition in the heart tissue. *Marine Biology* **136**, 573-580.

Joensen, J. P. (1996). The fisheries of the Faroe Islands. An overview. In: P. Holm, D. J. Starkey, J. T. Thór (Eds.), *The North Atlantic Fisheries, 1100-1976*. Studia Atlantica, 1, Fiskeri- og Søfartsmuseet, Esbjerg, Denmark, pp. 27-47.

Joensen, J. S., Steingrund, P., Henriksen, A. and Mouritsen, R. (2005). Migration of cod (*Gadus morhua*): tagging experiments at the Faroes 1952-1965. *Fródskaparrit* **53**, 100-135.

Johansen, S. and Bakke, I. (1996). The complete mitochondrial DNA sequence of Atlantic cod (*Gadus morhua*): relevance to taxonomic studies among codfishes. *Molecular Marine Biology and Biotechnology* **5**, 203-214.

Johansen, T., Nyhammer, G., Otterlei, E. and Nævdal, G. (2000). Sexual maturation in cod (*Gadus morhua* L.) in relation to strain, temperature, sex, size and haemoglobin type. In: B. Norberg, O. S. Kjesbu, G. L. Taranger, E. Andersson and S.O. Stefansson (Eds.), Proceedings of the 6th International Symposium on the Reproductive Physiology of Fish. Institute of Marine Research and University of Bergen, Bergen, Norway, pp. 1-343.

Johnsen, H. and Andersen, Ø. (2012). Sex dimorphic expression of five *dmrt* genes identified in the Atlantic cod genome. The fish-specific *dmrt2b* diverged from *dmrt2a* before the fish whole-genome duplication. *Gene* **505**(2), 221-232. doi: 10.1016/j.gene.2012.06.021.

Johnsen, H., Seppola, M., Torgersen, J. S., Delghandi, M. and Andersen, Ø. (2010). Sexually dimorphic expression of *dmrt1* in immature and mature Atlantic cod (*Gadus morhua* L.). *Comparative Biochemistry and Physiology, Part B* **156**, 197-205. doi: 10.1016/j.cbpb.2010.03.009.

Johnsen, H., Tveiten, H., Torgersen, J. S. and Andersen, Ø. (2013). Divergent and sex-dimorphic expression of the paralogs of the Sox9-Amh-Cyp19a1 regulatory cascade in developing and adult Atlantic cod (*Gadus morhua* L.). *Molecular Reproduction and Development* **80**(5), 358-370. doi: 10.1002/mrd.22170.

Johnston, I. A. and Andersen, Ø. (2008). Number of muscle fibers in adult Atlantic cod varies with temperature during embryonic development and pantophysin (*Pan*I) genotype. *Aquatic Biology* **4**, 167-173. doi: 10.3354/ab00101.

Jombart, T. (2008). *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**(11), 1403-1405. doi: 10.1093/bioinformatics/btn129.

Jones, A. G. and Ardren, W. R. (2003). Methods of parentage analysis in natural populations. *Molecular Ecology* **12**, 2511-2523. doi: 10.1046/j.1365-294X.2003.01928.x.

Jones B. W. (1966). The cod and cod fishery at Faroe. Fisheries and Food. *Fishery Investigations, Series II* **24**. Her Majesty´s Stationary Office (HMSO), London, UK, pp. 1-32.

Jónsdóttir, Ó. B. D., Imsland, A. K., Atladóttir, Ó. Ý. and Daníelsdóttir, A. K. (2003). Nuclear DNA RFLP variation in Atlantic cod in the North Atlantic Ocean. *Fisheries Research* **63**(3), 429-436. doi: 10.1016/S0165-7836(03)00098-5.

Jónsdóttir, Ó. D. B., Imsland, A. K., Daníelsdóttir, A. K. and Marteinsdóttir, G. (2002). Genetic heterogeneity and growth properties of different genotypes of Atlantic cod (*Gadus morhua* L.) at two spawning sites off south Iceland. *Fisheries Research* **55**, 37-47.

Jørgensen, T. (1990). Long-term changes in age at sexual maturation of Northeast Arctic cod (*Gadus morhua* L.). *Journal du Conseil - Conseil International pour l´Exploration de la Mer* **46**, 235-248.

Jørstad, K. E., Otterå, H., van der Meeren, T., Dahle, G., Paulsen, O. I., Bakke, G. and Svåsand, T. (2014). Genetic marking of farmed Atlantic cod (*Gadus morhua* L.) and detection of escapes from a commercial cod farm. *ICES Journal of Marine Science* **91**(3), 574-584. doi: 10.1093/icesjms/fst171.

Jørstad, K. E., van der Meeren, T., Paulsen, O. I., Thomsen, T., Thorsen, A. and Svåsand, T. (2008). "Escapes" of eggs from farmed cod spawning in net pens: recruitment to wild stocks. *Reviews in Fisheries Science* **16**(1-3), 285-295. doi: 10.1080/10641260701678017.

Jost, L. (2008). G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* **17**, 4015-4026. doi: 10.1111/j.1365-294X.2008.03887.x.

Jost, L. (2009). *D* vs. G_{ST}: Response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Molecular Ecology* **18**, 2088-2091. doi: 10.1111/j.1365-294X.2009.04186.x.

Jurecka, P., Wiegertjes, G. F., Rakus, K. Ł., Pilarczyk, A. and Irnazarow, I. (2009). Genetic resistance of carp (*Cyprinus carpio* L.) to *Trypanoplasma borreli*: influence of transferrin polymorphisms. *Veterinary Immunology and Immunopathology* **127**, 19-25. doi: 10.1016/j.vetimm.2008.09.006.

Kalinowski, S. T. (2009). How well do evolutionary trees describe genetic relationships among populations? *Journal of Heredity* **102**, 506-513. doi: 10.1038/hdy.2008.136.

Kamler, E. (2005). Parent–egg–progeny relationships in teleost fishes: an energetics perspective. *Reviews in Fish Biology and Fisheries* **15**, 399-421. doi: 10.1007/s11160-006-0002-y.

Karlsen, B. O., Klingan, K., Emblem, Å., Jørgensen, T. E., Jueterbock, A., Furmanek, T., Hoarau, G., Johansen, S. D., Nordeide, J. and Moum, T. (2013). Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. *Molecular Ecology* **22**, 5098-5111. doi: 10.1111/mec.12454.

Karlsen, Ø. and Holm, J. C. (1994). Ultrasonography, a non-invasive method for sex determination in cod (*Gadus morhua*). *Journal of Fish Biology* **44**(6), 965-971. doi: 10.1111/j.1095-8649.1994.tb01268.x.

Karlsen, Ø., Norberg, B., Kjesbu, O. S. and Taranger, G. L. (2006). Effects of photoperiod and exercise on growth, liver size, and age at puberty in farmed Atlantic cod (*Gadus morhua* L.). *ICES Journal of Marine Science* **63**, 355-364. doi: 10.1016/j.icesjms.2005.10.013.

Karlsson, S., Moen, T., Lien, S., Glover, K. A. and Hindar, K. (2011). Generic genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP-chip. *Molecular Ecology Resources* **11**(Suppl. s1), 247-253. doi: 10.1111/j.1755-0998.2010.02959.x.

Karlsson, S. and Mørk, J. (2003). Selection-induced variation at the pantophysin locus (*Pan*I) in a Norwegian fjord population of cod (*Gadus morhua* L.). *Molecular Ecology* **12**, 3265-3274. doi: 10.1046/j.1365-294X.2003.01993.x.

Karlsson, S. and Mørk, J. (2005). Deviation from Hardy-Weinberg equilibrium, and temporal instability in allele frequencies at microsatellite loci in a local population of Atlantic cod. *ICES Journal of Marine Science* **62**, 1588-1596. doi: 10.1016/j.icesjms.2005.05.009.

Karpov, A. K. and Novikov, G. G. (1981). Hemoglobin alloforms in cod *Gadus morhua* (Gadiformes, Gadidae), their functional characteristics and occurence in populations. *Journal of Ichthyology* **6**, 45-49.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W. and Prodöhl, P. A. (2013). diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* **4**, 782-788. doi: 10.1111/2041-210X.12067.

Keller, W. T. and Plosila, D. S. (1981). Comparison of domestic, hybrid and wild strains of brook trout in a pond fishery. *New York Fish and Game Journal* **28**, 123-137.

Kent, W. J. (2002). BLAT-The BLAST-like alignment tool. *Genome Research* **12**, 656-664. doi: 10.1101/gr.229202.

Kesteven, G. L. (1960). Manual of field methods in fisheries biology. Rome, Italy: FAO, 152 pp.

Kettunen, A. and Fjalestad, K. T. (2006). Resistance to vibriosis in Atlantic cod (*Gadus morhua* L.): first challenge results. *Aquaculture* **258**, 263-269. doi: 10.1016/j.aquaculture.2006.04.042.

Kettunen, A. and Fjalestad, K. T. (2007). Genetic parameters for important traits in the breeding program for Atlantic cod (*Gadus morhua* L.). *Aquaculture* **272**(Suppl. 1), S276.

Kettunen, A., Serenius, T. and Fjalestad, K. T. (2007). Three statistical approaches for genetic analysis of disease resistance to vibriosis in Atlantic cod (*Gadus morhua* L.). *Journal of Animal Science* **85**, 305-313. doi: 10.2527/jas.2006-112.

Kirk, H. and Freeland, J. R. (2011). Applications and implications of neutral *versus* non-neutral markers in molecular ecology. *International Journal of Molecular Sciences* **12**, 3966-3988. doi: 10.3390/ijms12063966.

Kjesbu, O. S. (1989). The spawning activity of cod, *Gadus morhua* L. *Journal of Fish Fiology* **34**, 195-206.

Kjesbu, O. S. (1994). Time of start of spawning in Atlantic cod (*Gadus morhua*) females in relation to vitellogenic oocyte diameter, fish length and condition. *Journal of Fish Biology* **45**, 719-735.

Kjesbu, O. S., Klungsoyr, J., Kryvi, H., Witthames, P. R. and Greer Walker, M. (1991). Fecundity, atresia, and egg size of captive Atlantic cod (*Gadus morhua*) in relation to proximate body composition. *Canadian Journal of Fisheries and Aquatic Sciences* **48**, 2333-2343.

Kjesbu, O. S., Righton, D., Krüger-Johnsen, M., Thorsen, A., Michalsen, K., Fonn, M. and Witthames, P. R. (2010). Thermal dynamics of ovarian maturation in Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences* **67**, 606-625. doi: 10.1139/F10-011.

Kjesbu, O. S., Solemdal, P., Bratland, P. and Fonn, M. (1996). Variation in annual egg production in individual captive Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences* **53**, 610-620.

Kjesbu, O. S., Taranger, G. L. and Trippel, E. A. (2006). Gadoid mariculture: development and future challenges. *ICES Journal of Marine Science* **63**, 187-191. doi: 10.1016/j.icesjms.2005.12.003.

Kjesbu, O. S., Witthames, P. R., Solemdal, P. and Greer Walker, M. (1990). Ovulatory rhythm and a method to determine the stage of spawning in Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences* **47**, 1185-1193.

Kjørsvik, E. and Lønning, S. (1983). Effects of egg quality on normal fertilisation and early development of the cod, *Gadus morhua* L. *Journal of Fish Biology* **23**, 1-12. doi: 10.1111/j.1095-8649.1983.tb02877.x.

Knutsen, G. M. and Tilseth, S. (1985). Growth, development and feeding success of Atlantic cod *Gadus morhua* related to egg size. *Transactions of the American Fisheries Society* **114**, 507-511.

Knutsen, H., Jorde, P. E., Albert, O. T., Hoelzel, A. R. And Stenseth, N. C. (2007). Population genetic structure in the North Atlantic Greenland halibut (*Reinhardtius hippoglossoides*): influenced by ocean currents? *Canadian Journal of Fisheries and Aquatic Sciences* **64**, 857-866. doi: 10.1139/F07-070.

Knutsen, H., Jorde, P. E., André, C. and Stenseth, N. C. (2003). Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. *Molecular Ecology* **12**, 385-394. doi: 10.1046/j.1365-294X.2003.01750.x.

Knutsen, H., Jorde, P. E., Sannæs, H., Hoelzel, A. R., Bergstad, O. A., Stefanni, S., Johansen, T. and Stenseth, N. C. (2009). Bathymetric barriers promoting genetic structure in the deepwater demersal fish tusk (*Brosme brosme*). *Molecular Ecology* **18**, 3151-3162. doi: 10.1111/j.1365-294X.2009.04253.x.

Knutsen, I. and Salvanes, A. G. V. (1999). Temperature-dependent digestion handling time in juvenile cod and possible consequences for prey choice. *Marine Ecology Progress Series* **181**, 61-79. doi: 10.3354/meps181061.

Kolbeinshavn, A. G., Vestergaard, P., Patursson, Ø. and Gislason, H. (2012). Rapid growth of cod in sea cages in the Faroe Islands. *Aquaculture* **358-359**, 224-233. doi: 10.1016/j.aquaculture.2012.06.036.

Kristiansen, T. S., Jørstad, K. E., Otterå, H., Paulsen, O. I. and Svåsand, T. (1997). Estimates of larval survival of cod by releases of genetically marked yolk-sac larvae. *Journal of Fish Biology* **51**(Suppl. A), 264-283. doi: 10.1111/j.1095-8649.1997.tb06103.x.

Kumar, S., Banks, R. and Cloutier, S. (2012). SNP discovery through next-generation sequencing and its applications. *International Journal of Plant Genomics* **831460**, pp. 1-15. doi: 10.1155/2012/831460.

Kume, M. (2011). Clutch and egg sizes of two migratory forms of the Threespine Stickleback *Gasterosteus aculeatus* in Eastern Hokkaido, Japan. *Zoological Studies* **50**(3), 309 – 314.

Kurlansky, M. (1997). Cod: a Biography of the Fish that changed the World. London, UK: Jonathan Random House, 282 pp.

Lambert, Y., Yaragina, N. A., Kraus, G., Marteinsdóttit, G. and Wright, P. J. (2003). Using environmental and biological indices as proxies for egg and larval production of marine fish. *Journal of Northwest Atlantic Fishery Science* **33**, 115-159.

Lanes, C. F. C., Bizuayehu, T. T., Bolla, S., Martins, C., Fernandes, J. M. de O., Bianchini, A., Kiron, V. and Babiak, I. (2012). Biochemical composition and performance of Atlantic cod (*Gadus morhua* L.) eggs and larvae obtained from farmed and wild broodstocks.*Aquaculture* **324-325**, 267-275. doi: 10.1016/j.aquaculture.2011.10.036.

Larsen, J. L., Pedersen, K. and Dalsgaard, I. (1994). *Vibrio anguillarum* serovars associated with vibriosis in fish. *Journal of Fish Diseases* **17**, 259-267. doi: 10.1111/j.1365-2761.1994.tb00221.x.

Larsen, K. M. H., Hansen, B. and Svendsen, H. (2008b). Faroe Shelf water. *Continental Shelf Research* **28**, 1754-1768. doi: 10.1016/j.csr.2008.04.006.

Larsen, K. M. H., Hansen, B. and Svendsen, H. (2009). The Faroe Shelf front: properties and exchange. *Journal of Marine Systems* **78**, 9-17. doi: 10.1016/j.jmarsys.2009.02.003.

Larsen, P. F., Nielsen, E. E., Williams, T. D. and Loeschcke, V. (2008a). Intraspecific variation in expression of candidate genes for osmoregulation, heme biosynthesis and stress resistance suggests local adaptation in European flounder (*Platichthys flesus*). *Journal of Heredity* **101**, 249-259. doi: 10.1038/hdy.2008.54.

Larsson, L. C., Laikre, L., Palm, S., André, C., Carvalho, G. R. and Ryman, N. (2007). Concordance of allozyme and microsatellite differentiation in a marine fish, but evidence of selection at a microsatellite locus. *Molecular Ecology* **16**, 1135-1147. doi: 10.1111/j.1365-294X.2006.03217.x.

Law, R. (2000). Fishing, selection and phenotypic evolution. *ICES Journal of Marine Science* **57**(3), 659-668. doi: 10.1006/jmsc.2000.0731.

Lear, W. H. (1998). History of fisheries in the Northwest Atlantic: the 500-year perspective. *Journal of Northwest Atlantic Fishery Science* **23**, 41-73.

Lee, T.-H. and Maheshri, N. (2012). A regulatory role for repeated deoxy transcription factor binding sites in target gene expression. *Molecular Systems Biology* **8**(576), pp. 1-11. doi: 10.1038/msb.2012.7.

Leider, S. A., Hulett, P. L., Loch, J. J. and Chilcote, M. W. (1990). Electrophoretic comparison of the reproductive success of naturally spawned transplanted and wild steelhead trout through the returning adult stage. *Aquaculture* **88**, 239-252.

Lester, L. J. (2002). What role does genetics play in responsible aquaculture. In: R. R. Stickney and J. P. McVey (Eds.), *Responsible Marine Aquaculture*. New York, USA: CABI Publishing, pp. 233-261.

Li, Y.-C., Korol, A. B., Fahima, T., Beiles, A. and Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational machanisms: a review. *Molecular Ecology* **11**(12), 2453-2465. doi: 10.1046/j.1365-294X.2002.01643.x.

Li, Y.-C., Korol, A. B., Fahima, T. and Nevo, E. (2004). Microsatellites within genes: structure, function and evolution. *Molecular Biology and Evolution* **21**(6), 991-1007. doi: 10.1093/molbev/msh073.

Liu, Z. (2007). Single nucleotide polymorphism (SNP). In: Z. Liu (Ed.) *Aquaculture Genome Technologies*. 1st edn. Iowa, USA: Blackwell Publishing, Ames, pp. 59-72.

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

Love, R. M., Robertson, I., Lavéty, J. and Smith, G. L. (1974). Some biological characteristics of cod (*Gadus morhua* L.) from the Faroe Bank compared with those from other fishing grounds. *Comparative Biochemistry and Physiology* **47**, 149-162.

Lowe, A., Harris, S. and Ashton, P. (2004a). Gene flow and mating system. In: *Ecological Genetics: Design, Analysis and Application*. Oxford, UK: Blackwell Publishing, pp. 106-149.

Lowe, A., Harris, S. and Ashton, P. (2004b). Intraspecific phylogenies and phylogeography. In: *Ecological Genetics: Design, Analysis and Application*. Oxford, UK: Blackwell Publishing, pp. 150-188.

Lubzens, E., Young, G., Bobe, J. and Cerdà, J. (2010). Oogenesis in teleosts: how fish eggs are formed. *General and Comparative Endocrinology* **165**, 367-389. doi: 10.1016/j.ygcen.2009.05.022.

Lucchesi, J. L., Kelly, W. G. and Panning, B. (2005). Chromatin remodeling in dosage compensation. *Annual Review of Genetics* **39**, 615-651. doi: 10.1146/annurev.genet.39.073003.094210.

Luikart, G., Allendorf, F. W., Cornuet, J.-M. and Sherwin, W. B. (1998b). Distortion of allele frequency distributions provides a test for recent population bottleneck. *Journal of Heredity* **89**(3), 238-247. doi: 10.1093/jhered/89.3.238.

Luikart, G., Sherwin, W. B., Steele, B. M. and Allendorf, F. W. (1998a). Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology* **7**(8), 963-974. doi: 10.1046/j.1365-294x.1998.00414.x.

Lund, V., Arnesen, J. A., Mikkelsen, H., Gravningen, K., Brown, L. and Schrøder, M. B. (2008). Atypical

furunculosis vaccines for Atlantic cod (*Gadus morhua*); vaccine efficacy and antibody responses. *Vaccine* **26**, 6791-6799.

Macedo, M. F., De Sousa, M., Ned, R. M., Mascarenhas, C., Andrews, N. C. and Correia-Neves, M. (2004). Transferrin is required in early T-cell differentiation. *Immunology* **112**(4), 543-549. doi: 10.1111/j.1365-2567.2004.01915.x.

Macer, T. and Easey, M. W. (1988). North Sea cod and the English fishery. Laboratory leaflet Number 61. Ministry of Agriculture, Fisheries and Food, Directorate of Fisheries Research, Lowestoft, UK. 23 pp.

Magnadóttir, B., Bambir, S. H., Gudmundsdóttir, B. K., Pilström, L. and Helgason, S. (2002). Atypical *Aeromonas salmonicida* infection in naturally and experimentally infected cod, *Gadus morhua* L. *Journal of Fish Diseases* **25**, 583-597.

Magnussen, E. (1993). Hvorfor vokser torsken hurtigere på Færøbanke end på Færøplateauet? M.Sc. Thesis. University of Copenhagen, 104 pp.

Magnussen, E. (1996). Electrophoretic studies of cod (*Gadus morhua*) from the Faroe Bank and Faroe Plateau compared to results found in other distribution areas. *ICES. Demersal Fish Committee*, CM 1996/G:10, 18 pp.

Magnussen, E. (2002). Demersal fish assemblages of Faroe Bank: species composition, distribution, biomass spectrum and diversity. *Marine Ecology Progress Series* **239**, 211-225.

Magnussen, E. (2006). Genetic activities on fish in the Faroe Islands. In: *NVDrit* 2006: 09. Facutly of Science and Technology, University of the Faroe Islands, pp. 1-6.

Magnussen, E. (2007). Interpopulation comparison of growth patterns of 14 fish species on Faroe Bank: are all species on the bank fast-growing? *Journal of Fish Biology* **71**, 453-475. doi: 10.1111/j.1095-8649.2007.01502.x.

Magnussen, E. (2011). Food and feeding habits of cod (*Gadus morhua*) on the Faroe Bank. *ICES Journal of Marine Science* **68**(9), 1909-1917. doi: 10.1093/icesjms/fsr104.

Magoulas (1998). Application of molecular markers to aquaculture and broodstock management with special emphasis on microsatellite DNA. In: *Proceeding of the Seminar Genetics and Breeding of Mediterranean Aquaculture Species, Network on Technology of Aquaculturein the Mediterranean (TECAM)*, Zaragossa, Spain, 28-30 April 1997, pp. 153-168.

Makridis, P., Fjellheim, A. J., Skjermo, J. and Vadstein, O. (2000). Control of the bacterial flora of *Brachionus plicatilis* and *Artemia franciscana* by incubation in bacterial suspensions. *Aquaculture* **185**, 207-218. doi: 10.1016/S0044-8486(99)00351-8.

Marcil, J., Swain, D. P. and Hutchings, J. A. (2006a). Genetic and environmental components of phenotypic variation in body shape among populations of Atlantic cod (*Gadus morhua* L.). *Biological Journal of the Linnean Society* **88**, 351-365. doi: 10.1111/j.1095-8312.2006.00656.x.

Marcil, J., Swain, D. P. and Hutchings, J. A. (2006b). Countergradient variation in body shape between two populations of Atlantic cod (*Gadus morhua*). *Proceedings of the Royal Society of London B* **273**, 217-223. doi: 10.1098/rspb.2005.3306.

Mardis, E. R. (2008a). Next-generation DNA sequencing methods. *Annual Reviews of Genomics and Human Genetics* **9**, 387-402. doi: 10.1146/annurev.genom.9.081307.164359.

Mardis, E. R. (2008b). The impact of next-generation sequencing technology on genetics. *Trends in Genetics* **24**(3), 133-141. doi: 10.1016/j.tig.2007.12.007.

Markle, D. F. and Frost, L. A. (1985). Comparative morphology, seasonality and a key to planktonic fish eggs from the Nova Scotian shelf. *Canadian Journal of Zoology* **63**, 246-257.

Marshall, C. T. and Frank, K. T. (1994). Geographic responses of groundfish to variation in abundance: methods of detection and interpretation. *Canadian Journal of Fisheries and Aquatic Sciences* **51**, 808-816.

Marshall, D. J. and Uller, T. (2007). When is a maternal effect adaptive? *Oikos* **116**, 1957-1963. doi: 10.1111/j.2007.0030-1299.16203.x.

Marshall, T. C., Slate, J., Kruuk, L. E. B. and Pemberton, J. M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639-655.

doi: 10.1046/j.1365-294x.1998.00374.x.

Marteinsdóttir, G. and Björnsson, H. (1999). Time and duration of spawning of cod in Icelandic waters. ICES CM 1999/Y:34, 22 pp.

Marteinsdóttir, G., Ruzzante, D. and Nielsen, E. E. (2005). History of the North Atlantic cod stocks. ICES CM 2005/AA, 17 pp.

Marteinsdóttir, G. and Steinarsson, A. (1998). Maternal influence on the size and viability of Iceland cod *Gadus morhua* eggs and larvae. *Journal of Fish Biology* **52**, 1241-1258. doi: 10.1111/j.1095-8649.1998.tb00969.x.

McGinnity, P. Jennings, E., deEyto, E., Allott, N., Samuelsson, P., Rogan, G. Whelan, K. and Cross, T. (2009). Impact of naturally spawning captive-bred Atlantic salmon on wild populations: depressed recruitment and increased risk of climate-mediated extinction. *Proceedings of the Royal Society of London B* **276**(1673), 3601-3610. doi: 10.1098/rspb.2009.0799.

McKenzie, R. A. (1941). Cod migrations from tagging results. Fisheries Research Board of Canada, Annual report **18**, pp. 18-19.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P. And Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**(16), 2069-2070. doi: 10.1093/bioinformatics/btq330.

McPherson, A. A., O'Reilly, P. T. and Taggart, C. T. (2004). Genetic differentiation, temporal stability, and the absence of isolation by distance among Atlantic herring populations. *Transactions of the American Fisheries Society* **133**(2), 434-446. doi: 10.1577/02-106.

Meirmans, P. G. (2006). Using the AMOVA framework to estimate a standardized genetic differentiation estimate. *Evolution* **60**(11), 2399-2402. doi: 10.1111/j.0014-3820.2006.tb01874.x.

Meirmans, P. G. and Hedrick, P. W. (2011). Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources* **11**, 5-18. doi: 10.1111/j.1755-0998.2010.02927.x.

Mertz, G. and Myers, R. A. (1994). Match/mismatch predictions of spawning duration versus recruitment variability. *Fisheries Oceanography* **3**, 236-245. doi 10.1111/j.1365-2419.1994.tb00101.x.

Miller, C.R., Joyce, P. and Waits, L.P. (2002). Assessing allelic dropout and genotype reliability using maximum likelihood. *Genetics* **160**, 357-366.

Miller, K. M., Le, K. D. and Beacham, T. D. (2000). Development of tri- and tetranucleotide repeat microsatellite loci in Atlantic cod (*Gadus morhua*). *Molecular Ecology* **9**(2), 238-239. doi: 10.1046/j.1365-294x.2000.00804-2.x.

Miller, T. J., Herra, T. and Leggett, W. C. (1995). An individual-based analysis of the variability of eggs and their newly hatched larvae of Atlantic cod (*Gadus morhua*) on the Scotian Shelf. *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 1083-1093.

Moen, T., Hayes, B., Nilsen, F., Delghandi, M., Fjalestad, K. T., Fevolden, S.-E., Berg, P. R. and Lien, S. (2008). Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. *BMC Genetics* **9**(18), pp. 1-9. doi: 10.1186/1471-2156-9-18.

Møller, D. (1968). Genetic diversity in spawning cod along the Norwegian coast. *Hereditas* **60**, 1-32.

Morgan, M. J., DeBlois, E. M., and Rose, G. A. 1(997). An observation of the reaction of Atlantic cod (*Gadus morhua*) in a spawning shoal to bottom trawling. *Canadian Journal of Fisheries and Aquatic Sciences* **54**(Suppl. 1), 217-223.

Morin, P. A., Luikart, G., Wayne, R. K and the SNP workshop group (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution* **19**(4), 208-216. doi: 10.1016/j.tree.2004.01.009.

Mørk, J., Giskeødegård, R. and Sundnes, G. (1983). Haemoglobin polymorphism in *Gadus morhua*: genotypic differences in maturing age and within-season gonad maturation. *Helgoländer Meeresuntersuchungen* **36**, 313-322.

Mørk, J., Giskeødegård, R. and Sundnes, G. (1984). The haemoglobin polymorphism in Atlantic cod (*Gadus morhua* L.); genotypic differences in somatic growth and in maturing age in natural populations. In E. Dahl, D. S. Danielssen, E. Moksnes and P. Solemdal (Eds.). *The Propagation of Cod Gadus morhua L.* An International Symposium. Institute of Marine Research, Flødevigen

Biological Station, Arendal, Norway. pp. 1-12.

Mørk, J., Reuterwall, C., Ryman, N. and Ståhl, G. (1982). Genetic variation in Atlantic cod (*Gadus morhua* L.): a quantitative estimate from a Norwegian coastal population. *Hereditas* **96**, 55-61.

Mørk, J., Ryman, N., Ståhl, G., Utter, F. and Sundnes, G. (1985). Genetic variation in Atlantic cod (*Gadus morhua*) throughout its range. *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 1580**-**1587.

Mousseau, T. A. and Roff, D. A. (1987). Natural selection and the heritability of fitness components. *Journal of Heredity* **59**, 181-197. doi: 10.1038/hdy.1987.113.

Moyle, P. B. (1969). Comparative behaviour of young brook trout of domestic and wild origin. *Progressive Fish-culturist* **31**, 51-59.

Mullowney, D. R. J. and Rose, G. A. (2014). Is recovery of northern cod limited by poor feeding? The capelin hypothesis revisited. *ICES Journal of Marine Science* **71**(4), 784-793. doi: 10.1093/icesjms/fst188.

Murawski, S. A. (2010). Rebuilding depleted fish stocks: the good, the bad, and, mostly, the ugly. *ICES Journal of Marine Science* **67**(9), 1830-1840. doi: 10.1093/icesjms/fsq125.

Myhr, A. I., Rosendal, G. K. and Olesen, I. (2012). New evelopments in Biotechnology and IPR in Aquaculture – Are They Sustainable? In: Z. A. Muchlisin (Ed.), *Aquaculture Book 1*. Rijeka, Croatia: InTech – Open Access Publisher, pp. 317-342. doi: 10.5772/30830.

Nash, R. D. M., Valencia, A. H. and Geffen, A. J. (2006). The origin of Fulton´s condition factor – setting the record straight. Fisheries **31**(5), 236-238.

Naylor, R. L., Goldburg, R. J., Primavera, J. H., Kautsky, N., Beveridge, M. C. M., Clay, J., Folke, C., Lubchenco, J., Mooney, H. and Troell, M. (2000). Effect of aquaculture on world fish supplies. *Nature*, **405**, 1017-1033. doi: 10.1038/35016500.

Nei, M., Tajima, F. and Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153-170.

Nielsen, E. E., Cariani, A., Mac Aoidh, E., Maes, G. E., Milano, I., Ogden, R., Taylor, M., Hemmer-Hansen, J., Babbucci, M., Bargelloni, L., Bekkevold, D., Diopere, E., Grenfell, L., Helyar, S., Limborg, M. T., Martinsohn, J. T., McEwing, R., Panitz, F., Patarnello, T., Tinti, F., van Houdt, J. K. J., Volckaert, F. A. M., Waples, R. S., FishPopTrace consortium and Carvalho, G. R. (2012). Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature Communications* **3**(851), pp. 1-6. doi: 10.1038/ncomms1845.

Nielsen, E. E., Hansen, M. M. and Meldrup, D. (2006). Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Molecular Ecology* **15**, 3219-3229. doi: 10.1111/j.1365-294X.2006.03025.x.

Nielsen E. E., Hansen, M. M., Ruzzante, D. E., Meldrup, D. and Grønkjær, P. (2003). Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. *Molecular Ecology* **12**, 1497-1508. doi: 10.1046/j.1365-294X.2003.01819.x.

Nielsen, E. E., Hansen, M. M., Schmidt, C., Meldrup, D. and Grønkjær, P. (2001). Fisheries: Population of origin of Atlantic cod. *Nature* **413**, p. 272. doi: 10.1038/35095112.

Nielsen, E. E., Hemmer-Hansen, J., Poulsen, N. A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G.-L., Ogden, R. and Carvalho, G. R. (2009b). Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology* **9**(276), pp. 1-11. doi: 10.1186/1471-2148-9-276.

Nielsen, E. E., MacKenzie, B. R., Magnussen, E. and Meldrup, D. (2007). Historical analysis of *Pan*I in Atlantic cod (*Gadus morhua*): temporal stability of allele frequencies in the southeastern part of the species distribution. *Canadian Journal of Fisheries and Aquatic Sciences* **64**, 1448–1455. doi: 10.1139/F07-104.

Nielsen, E. E., Wright, P. J., Hemmer-Hansen, J., Poulsen, N. A., Monro Gibb, I. and Meldrup, D. (2009a). Microgeographical population structure of cod Gadus morhua in the North Sea and west of Scotland: the role of sampling loci and individuals. *Marine Ecology Progress Series* **376**, 213-225. doi: 10.3354/meps07798.

Norberg, B., Brown, C. L., Halldórsson, A., Stensland, K. and Björnsson, B. T. (2004). Photoperiod regulates the timing of sexual maturation, spawning, sex steroid and thyroid hormone profiles in the Atlantic cod (*Gadus morhua*). *Aquaculture* **229**, 451-467. doi: 10.1016/S0044-8486(03)00393-4.

Nordeide, J. T. and Folstad, I. (2000). Is cod lekking or a premiscuous spawner? *Fish and Fisheries* **1**, 90-93. doi: 10.1046/j.1467-2979.2000.00005.x.

Norris, A. T., Bradley, D. G. and Cunningham, E. P. (1999). Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. *Aquaculture* **180**, 247-264. doi: 10.1016/S0044-8486(99)00212-4.

Norris, A. T., Bradley, D. G. and Cunningham, E. P. (2000). Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture* **182**, 73-83. doi: 10.1016/S0044-8486(99)00247-1.

Nosil, P., Funk, D. J. and Ortiz-Barrientos, D. (2009). Divergent selection and heterogenous genomic divergence. *Molecular Ecology* **18**, 375-402. doi: 10.1111/j.1365-294X.2008.03946.x.

Nævdal, G., Folkvord, A., Otterlei, E. and Thorkildsen, S. (1992). Growth rate related to genotype of 0-group cod at three environmental temperatures. *Sarsia* **77**, 71-73.

Ogden, R., Gharbi, K., Mugue, N., Martinsohn, J., Senn, H., Davey, W., Pourkazemi, M., Mcewing, R., Eland, C., Vidotto, M., Sergeev, A. and Congiu, L. (2013). Sturgeon conservation genomics: SNP discovery and validation using RAD sequencing. *Molecular Ecology* **22**, 3112-3123. doi: 10.1111/mec.12234.

Olafsen, J. A. (2001). Interactions between fish larvae and bacteria in marine aquaculture. *Aquaculture* **200**(1-2), 223-247. doi: 10.1016/S0044-8486(01)00702-5.

O'Leary, D. B., Coughlan, J., Dillane, E., McCarthy, T. V. and Cross, T. F. (2007). Microsatellite variation in cod *Gadus morhua* throughout its geographic range. *Journal of Fish Biology* **70**(Suppl. C), 310-335. doi: 10.1111/j.1095-8649.2007.01451.x.

Olsen, E. M., Heino, M., Lilly, G. R., Morgan, M. J., Brattey, J., Ernande, B. and Dieckmann, U. (2004). Maturation trends indicative of rapid evolution preceded the collapse of northern cod. *Nature* **428**, 932-935. doi: 10.1038/nature02430.

Ong, S. T., Ho, J. Z., Ho. B. and Ding, L. J. (2006). Iron-withholding strategy in innate immunity. *Immunobiology* **211**(4), 295-314. doi: 10.1016/j.imbio.2006.02.004.

Oosthuizen, E. and Daan, N. (1974). Egg fecundity and maturity of North Sea cod, *Gadus morhua*. *Netherlands Journal of Sea Research* **8**, 378-397. doi: 10.1016/0077-7579(74)90006-4.

O'Reilly, P. T., Canino, M. F., Bailey, K. M. and Bentzen, P. (2000). Isolation of twenty low stutterdi- and tetranucleotide microsatellites for population analyses of walleye pollock and other gadoids. *Journal of Fish Biology* **56**, 1074-1086. doi: 10.1006/jfbi.2000.1230.

O´Reilly, P. and Wright, J.M. (1995). The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. *Journal of Fish Biology* **47**(Suppl. A), 29-55. doi: 10.1111/j.1095-8649.1995.tb06042.x.

Otterå, H., Agnalt, A.-L. and Jørstad, K. E. (2006). Differences in spawning time of captive Atlantic cod from four regions of Norway, kept under identical conditions. *ICES Journal of Marine Science* **63**(2), 216-223. doi: 10.1016/j.icesjms.2005.11.004.

Ottersen, G. (1996). Environmental impact on variability in recruitment, larval growth and distribution of Arcto-Norwegian cod. Ph.D. Thesis. University of Bergen, 136 pp.

Ottersen, G. and Sundby, S. (2005). Arcto-Norwegian cod. In: K. Brander (Ed.) *ICES Cooperative Research Report No. 274. Spawning and Life History Information for North Atlantic cod stocks*. Copenhagen, Denmark: International Council for the Exploration of the Sea, pp 11-18.

Ouellet, P., Lambert, Y. and Bérubé, I. (2001). Cod egg characteristics and viability in relation to low temperature and maternal nutritional condition. *ICES Journal of Marine Science* **58**, 672-686. doi: 10.1006/jmsc.2001.1065.

Paetkau, D. (2003). An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology* **12**, 1375-1387. doi: 10.1046/j.1365-294X.2003.01820.x.

Paetkau, D., Slade, R., Burden, M. and Estoup, A. (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**, 55-65. doi: 10.1046/j.1365-294X.2004.02008.x.

Paetkau, D. and Strobeck, C. (1995). The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology* **4**, 519-520. doi: 10.1111/j.1365-294X.1995.tb00248.x.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357-358.

Paisley, L. G., Ariel, E., Lyngstad, T., Jónsson, G., Vennerström, P., Hellström, A. and Østergaard, P. (2010). An overview of aquaculture in the Nordic countries. *Journal of the World Aquaculture Society* **41**(1), 1-17. doi: 10.1111/j.1749-7345.2009.00309.x.

Palaiokostas, C., Bekaert, M., Davie, A., Cowan, M. E., Oral, M., Taggart, J. B., Gharbi, K., McAndrew, B. J., Penman, D. J. and Migaus, H. (2013a). Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. *BMC Genetics* **14**(**566**), pp. 1-12. doi: 10.1186/1471-2164-14-566.

Palaiokostas, C., Bekaert, M., Khan, M. G. Q., Taggart, J. B., Gharbi, K., McAndrew, B. J. and Penman, D. J. (2013b). Mapping and validation of the major sex-determining region in Nile Tilapia (*Oreochromis niloticus* L.). *PLOS One* **8**(7), e68389, pp. 1-9. doi: 10.1371/journal.pone.0068389.

Pampoulie, C., Daníelsdóttir, A. K., Storr-Paulsen, M., Hovgård, H., Hjörleifsson, E. and Steinarsson, B. Æ. (2011). Neutral and nonneutral genetic markers revealed the presence of inshore and offshore stock components of Atlantic cod in Greenland waters. *Transactions of the American Fisheries Society* **140**(2), 307-319. doi: 10.1080/00028487.2011.567850.

Pampoulie, C., Jakobsdóttir, K. B., Marteinsdóttir, G. and Thorsteinsson, V. (2008b). Are vertical behaviour patterns related to the pantophysin locus in the Atlantic cod (*Gadus morhua* L.)? *Behavior Genetics* **38**, 76-81. doi: 10.1007/s10519-007-9175-y.

Pampoulie, C. Jörundsdóttir, T. D., Steinarsson, A., Pétursdóttir, G., Stéfansson, M. Ö. and Daníelsdóttir, A. K. (2006). Genetic comparsion of experimental farmed strains and wild Ocelandic populations of Atlantic cod (*Gadus morhua* L.). *Aquaculture* **261**, 556-564. doi: 10.1016/j.aquaculture.2006.07.044.

Pampoulie, C., Skírnisdóttir, S., Ólafsdóttir, G., Helyar, S., Thorsteinsson, V., Jónsson, S., Fréchet, A., Durif, C., Sherman, S., Lampart-Kaluzniacka, M., Hedeholm, R., Ólafsson, H., Daníelsdóttir, A. and Kasper, J. M. (2014). Genetic structure of the lumpfish *Cyclopterus lumpus* across the North Atlantic. *ICES Journal of Marine Science*. doi. 10.1093/icesjms/fsu071.

Pampoulie, C., Steingrund, P., Stefánsson, M. Ö. and Daníelsdóttir, A. K. (2008a). Genetic divergence among East Icelandic and Faroese populations of Atlantic cod provides evidence for historical imprints at neutral and non-neutral markers. *ICES Journal of Marine Science* **65**(1), 65-71. doi: 10.1093/icesjms/fsm181.

Pandian, T. J. and Sheela, S. G. (1995). Hormonal induction of sex reversal in fish. *Aquaculture* **138**, 1-22.

Park, S. D. E. (2001). Trypanotolerance in West African cattle and the population genetics effects of selection. Ph.D. Thesis, University of Dublin.

Pasqualotto, A. C., Denning, D. W. and Anderson, M. J. (2007). A cautionary tale: lack of consistency in allele sizes between two laboratories for a published multilocus microsatellite typing system. *Journal of Clinical Microbiology* **45**(2), 522-528. doi: 10.1128/JCM.02136-06.

Peakall, R. and Smouse, P. S. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**(1), 288-295. doi: 10.1111/j.1471-8286.2005.01155.x.

Pedersen, T. and Jobling, M. (1989). Growth rates of large, sexually mature cod, *Gadus morhua*, in relation to condition and temperature during an annual cycle. *Aquaculture* **81**, 161-168. doi: 10.1016/0044-8486(89)90242-1.

Pelegri, F. (2003). Maternal factors in zebrafish development. *Developmental Dynamics* **228**, 535-554. doi: 10.1002/dvdy.10390.

Pepin, P., Orr, D. C. and Anderson, J. T. (1997). Time to hatch and larval size in relation to temperature and egg size in Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences*

54(1), 2-10.

Peruzzi, S., Kettunen, A., Primicerio, R. and Kaurić, G. (2007). Thermal shock induction of triploidy in Atlantic cod (*Gadus morhua* L.). *Aquaculture Research* **38**, 926-932. doi: 10.1111/j.1365-2109.2007.01742.x.

Petersen, M. F. and Steffensen, J. F. (2003). Preferred temperature of juvenile Atlantic cod *Gadus morhua* with different haemoglobin genotypes at normoxia and moderate hypoxia. *The Journal of Experimental Biology* **206**, 359-364. doi: 10.1242/jeb.00111.

Petty, E., Laughlin, E. and Csankovszki, G. (2011). Regulation of DCC localization by HTZ-1/H2A.Z and DPY-30 does not correlate with H3K4 methylation levels. *PLOS One* **6**(10), e25973, pp. 1-12. doi: 10.1371/journal.pone.0025973.

Piry, S., Alapetite, A., Cornuet,, J.-M., Paetkau, D., Baudouin, L. and Estoup, A. (2004). GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**(6), 536-539. doi: 10.1093/jhered/esh074.

Piry, S., Luikart, G. and Cornuet, J.-M. (1999). Bottleneck: a computer program for detecting reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**, 248-249. doi: 10.1093/jhered/90.4.502.

Planque, B. and Frédou, T. (1999). Temperature and the recruitment of Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 2069-2077.

Pogson, G. H. and Fevolden, S. E. (1998). DNA heterozygosity and growth rate in the Atlantic cod *Gadus morhua* (L.). *Evolution* **52**, 915-920.

Pogson, G. H. and Mesa, K. A. (2004). Positive Darwinian Selection at the Pantophysin (*Pan* I) locus in marine gadid fishes. *Molecular Biology and Evolution* **21**(1), 65-75. doi: 10.1093/molbev/msg237.

Pogson, G. H., Mesa, K. A. and Boutilier, R. G. (1995). Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. *Genetics* **139**, 375-385.

Pogson, G. H., Taggart, C. T., Mesa, K. A. and Boutilier, R. G. (2001). Isolation by distance in the Atlantic cod, *Gadus morhua*, at large and small geographic scales. *Evolution* **55**(1), 131-146. doi: 10.1111/j.0014-3820.2001.tb01279.x.

Pompanon, F., Bonin, A., Bellemain, E. and Taberlet, P. (2005). Genotyping errors: causes, consequences and solutions. *Genetics* **6**, 847-859. doi: 10.1038/nrg1707.

Ponzoni, R. W., Nguyen, N. H. and Khaw, H. L. (2007). Investment appraisal of genetic improvement programs in Nile tilapia (*Oreochromis niloticus*). *Aquaculture* **269**, 187-199. doi: 10.1016/j.aquaculture.2007.04.054.

Ponzoni, R. W., Nguyen, N. H., Khaw, H. L. and Ninh, N. H. (2008). Accounting for genotype by environment interaction in economic appraisal of genetic improvement programs in common carp *Cyprionus carpio*. *Aquaculture* **285**, 47-55. doi: 10.1016/j.aquaculture.2008.08.012.

Pope, J. G., Falk-Pedersen, J., Jennings, S., Rice, J. C., Gislason, H. and Daan, N. (2008). Honey, I cooled the cods: Modelling the effect of temperature on the structure of Boreal/ Arctic fish ecosystems. *Deep-Sea Research Part II: Topical Studies in Oceanography* **56**, 21-22, 2097-2107. doi: 10.1016/j.dsr2.2008.11.021.

Presson, A. P., Sobel, E., Lange, K. and Papp, J. C. (2006). Merging microsatellite data. *Journal of Computational Biology* **13**(6), 1131-1147. doi: 10.1089/cmb.2006.13.1131.

Pritchard, J. K., Stephens, M. and Donnelly, P. (2000). Inference of structure using multilocus genotype data. *Genetics* **155**, 945-959.

Pusceddu, A., Bianchelli, S., Martín, J., Puig, P., Palanques, A., Masqué, P. And Danovaro, R. (2014). Chronic and intensive bottom trawling impairs deep-sea biodiversity and ecosystem functioning. *Proceedings of the National Academy of Sciences of the United States of America*, pp. 1-6. doi: 10.1073/pnas.1405454111.

Puvanendran, V., Burt, A. L. and Brown, J. A. (2006). Can Atlantic cod (*Gadus morhua*) larvae be weaned faster onto dry feed at higher temperatures? *Aquaculture* **255**, 334-340.

doi: 10.1016/j.aquaculture.2005.11.027.

Quinn, N. L., Gutierrez, A. P., Koop, B. F. and Davidson, W. S. (2012). Genomics and Genome Sequencing: Benefits for finfish Aquaculture. In: Z. A. Muchlisin (Ed.), *Aquaculture Book 1*. Rijeka, Croatia: InTech – Open Access Publisher, pp. 127-154. doi: 10.5772/30316.

Rakitin, A., Ferguson, M. M. and Trippel, E. A. (2001). Male reproductive succes and body size in Atlantic cod *Gadus morhus* L. *Marine Biology* **138**, 1077-1085. doi: 10.1007/s002270100551.

Rannala, B. and Mountain, J. L. (1997). Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94**(17), 9197-9201.

Raymond, M. and Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**(3), 248-249.

Reisenbichler, R. R. and McIntyre, J. D. (1977). Genetic differences in growth and survival of juvenile hatchery and wild steelhead trout. *Journal of the Fisheries Research Board of Canada* **34**, 123-128.

Reiss, H., Hoarau, G., Dickey-Collas, M. and Wolff, W. J. (2009). Genetic population structure of marine fish: mismatch between biological and fisheries management units. *Fish and Fisheries* **10**, 361-395. doi: 10.1111/j.1467-2979.2008.00324.x.

Reznick, D. and Endler, J. A. (1982). The impact of predation on life history evolution in Trinidadian Guppies (*Poecilia reticulata*). *Evolution* **36**(1), 160-177.

Ricker, W. E. (1975). Computation and interpretation of biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada* **191**, pp. 1-382.

Ringø, E. and Birkbeck, T. H. (1999). Intestincal microflora of fish larvae and fry. *Aquaculture Research* **30**, 73-93.

Robichaud, D. and Rose, G. A. (2004). Migratory behaviour and range in Atlantic cod: inference from a century of tagging. *Fish and Fisheries* **5**, 185-214. doi: 10.1111/j.1467-2679.2004.00141.x.

Rose, G. A. (1993). Cod spawning on a migration highway in the north-west Atlantic. *Nature* **366**, 458-461. doi: 10.1038/366458a0.

Rose, G. A. (2005). On distributional responses of North Atlantic fish to climate change. *ICES Journal of Marine Science* **62**, 1360-1374. doi: 10.1016/j.icesjms.2005.05.007.

Rosenblum, E. B. and Novembre, J. (2007). Ascertainment bias in spatially structured populations: a case study in the Eastern fence lizard. *Journal of Heredity* **98**(4), 331-336. doi: 10.1093/jhered/esm031.

Rosendal, G. K., Olesen, I. and Tvedt, M. W. (2013). Evolving legal regimes, market structures and biology affecting access to and protection of aquaculture genetic resources. *Aquaculture* **402-403**, 97-105. doi: 10.1016/j.aquaculture.2013.03.026.

Rosenlund, G. and Halldórsson, Ó. (2007). Cod juvenile production: research and commercial developments. *Aquaculture* **268**, 188-193. doi: 10.1016/j.aquaculture.2007.04.040.

Rosenlund, G. and Skretting, M. (2006). Worldwide status and perspective on gadoid culture. *ICES Journal of Marine Science* **63**, 194-197. doi: 10.1016/j.icesjms.2005.11.012.

Rousset, F. (2008). GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**(1), 103-106. doi: 10.1111/j.1471-8286.2007.01931.x.

Rowell, C. A. (1993). The effects of fishing on timing of maturity in North Sea cod (*Gadus morhua* L.). In: T. K. Stokes, J. M. McGlade and R. Law (Eds.), *The Exploitation of Evolving Resources*. Julich, Germany: Springer-Verlag, pp. 44-61.

Roy, W., Bell, G., Sawanboonchun, J., Davie, A., Franco, J., Fernandes, D., Gnassou, J. and Robertson, D. (2007). SARF014: Cod broodstock nutrition: arachidonic acid and astaxanthin as determinants of egg quality. Final report. Prepared for The Scottish Aquaculture Research Forum. University of Stirling, pp. 1-27.

Rudolfsen, G., Figenschou, L., Folstad, I., Nordeide, J. T. and Søreng, E. (2005). Potential fitness benefits from mate selection in the Atlantic cod (*Gadus morhua*). *Journal of Evolutionary Biology* **18**(1), 172-179. doi: 10.1111/j.1420-9101.2004.00778.x.

Ruzzante, D. E., Taggart, C. T. and Cook, D. (1998). A nuclear DNA basis for shelf- and bank-scale population structure in northwest Atlantic cod (*Gadus morhua*): Labrador to Georges Bank. *Molecular Ecology*, **7**, 1663-1680. doi: 10.1046/j.1365-294x.1998.00497.x.

Ruzzante, D. E., Taggart, C. T., Cook, D. and Goddard, S. (1996). Genetic differentiation between inshore and offshore Atlantic cod (*Gadus morhua*) off Newfoundland: microsatellite DNA variation and antifreeze level. *Canadian Journal of Fisheries and Aquatic Sciences* **53**, 634-645.

Ruzzante, D. E., Taggart, C. T., Lang, S. and Cook, D. (2000b). Mixed-stock analysis of Atlantic cod near the Gulf of St. Lawrence based on microsatellite DNA. *Ecological Applications* **10**(4), 1090-1109.

Ruzzante, D. E., Wroblewski, J. S., Taggart, C. T., Smedbol, R. K., Cook, D. and Goddard, S. (2000a). Bay-scale population structure in coastal Atlantic cod in Labrador and Newfoundland, Canada. *Journal of Applied Ecology* **56**, 431-447. doi: 10.1111/j.1095-8649.2000.tb02116.x.

Ryan, M. R., Killen, S. S., Gregory, R. S. and Snelgrove, P. V. R. (2012). Predators and distance between habitat patches modify gap crossing behaviour of juvenile Atlantic cod (*Gadus morhua*, L. 1758). *Journal of Experimental Marine Biology and Ecology* **422-423**, 81-87. doi: 10.1016/j.jembe.2012.04.017.

Salmenkova, E. A. (2011). New view on the population genetic structure of marine fish. *Russian Journal of Genetics* **47**(11), 1279-1287. doi: 10.1134/S1022795411110159.

Salvanes, A. G. V. and Hart, P. J. (2000). Is individual variation in competitive performance of reared juvenile cod influenced by haemoglobin genotype? *Sarsia* **85**, 265-274. doi: 10.1080/00364827.2000.10414578.

Sarvas, T. H. and Fevolden, S. E. (2005). Pantophysin (*Pan* I) locus divergence between inshore v. offshore and northern v. southern populations of Atlantic cod in the north-east Atlantic. *Journal of Fish Biology* **67**, 444-469. doi: 10.1111/j.0022-1112.2005.00738.x.

Sawaya, S., Bagshaw, A., Buschiazzo, E., Kumar, P., Chowdhury, S., Black, M. E. and Gemmell, N. (2013). Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. *PLOS One* **8**(2), e54710, pp. 1-11. doi: 10.1371/journal.pone.0054710.

Schlötterer, C. (2004). The evolution of molecular markers - just a matter of fashion? *Nature Reviews Genetics* **5**, 63-69.

Schmidt, J. (1930). Racial investigation. X. The Atlantic cod (*Gadus callaris* L.) and local races of the same. *Comptes-Rendus des Travaux du Laboratoire Carlsberg*, 18.

Schmidt, S. F., Tendal, O. S. and Magnussen, E. (2003). The Faroe Bank – a potential MPA. Briefing for the World Wildlife Fund. http://www.ngo.grida.no/wwfneap/Projects/MPAmap.htm, accessed 21 May 2014. Pp. 1-2.

Schmitt, A. O., Aβmus, J., Bortfeldt, R. H. and Brockmann, G. A. (2010). CandiSNPer: a web tool for the identification of candidate SNPs for causal variants. *Bioinformatics* **26**(7), 969-970. doi: 10.1093/bioinformatics/btq068.

Schrøder, M. H., Mikkelsen, H., Børdal, S., Gravningen, K. and Lund, V. (2006). Early vaccination and protection of Atlantic cod (*Gadus morhua* L.) juveniles against vibriosis. *Aquaculture* **254**, 46-53. doi: 10.1016/j.aquaculture.2005.10.015.

Secor, D. H. (2005). Fish migration and the unit stock: three formative debates. A review of ecological and historical issues related to stock connectivity and metapopulations. In: S. X. Cadrin, K. D. Friedland, J. R. Waldman (Eds.), *Stock Identification Methods. Applications in Fishery Science*. London, UK: Elsevier Academic Press, pp. 17-44.

Seeb, J. E., Carvalho, G., Hauser, L., Naish, K., Roberts, S. and Seeb, L. W. (2011). Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources* **11**(Suppl. 1), 1-8. doi: 10.1111/j.1755-0998.2010.02979.x.

Selkoe, K. A. and Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629. doi: 10.1111/j.1461-0248.2006.00889.x.

Senan, S., Kizhakayil, D., Sasikumar, B. and Sheeja, T. E. (2014). Methods for development of microsatellite markers: an overview. *Notulae Scientia Biologicae* **6**(1), 1-13.

Serchuk, F. M., Grosslein, M. D., Lough, R. G., Mountain, D. G. and O´Brien, L. (1994). Fishery and the environmental factors affecting trends and fluctuations in the Georges Bank and Guld of Maine Atlantic cod stocks: an overview. *ICES Marine Science Symposia* **198**, 77-109.

Shinde, D. Y., Lai, F. Z. and Arnheim, N. (2003). Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. *Nucleic Acids Research*, **31**, 974-980. doi: 10.1093/nar/gkg178.

Sick, K. (1961). Haemoglobin polymorphism in fishes. *Nature* **192**, 894-896.

Sick, K. (1965a). Haemoglobin polymorphism of cod in the Baltic and the Danish Belt Sea. *Hereditas* **54**, 19-48.

Sick, K. (1965b). Haemoglobin polymorphism of cod in the North Sea and the North Atlantic Ocean. *Hereditas* **91**, 557-564.

Sigurgíslason, H. and Árnason, E. (2003). Extent of mitochondrial DNA sequence variation in Atlantic cod from the Faroe Islands: a resolution of gene genealogy. *Journal of Heredity* **91**, 557-564. doi: 10.1038/sj.hdy.6800361.

Skarstein, T. H., Westgaard, J.-I. and Fevolden, S. E. (2007). Comparing microsatellite variation in north-east Atlantic cod (*Gadus morhua* L.) to genetic structuring as revealed by the pantophysin (*Pan* I) locus. *Journal of Fish Biology* **70**(Suppl. C), 271-290. doi: 10.1111/j.1095-8649.2007.01456.x.

Skjermo, J. and Vadstein, O. (1999). Techniques for microbial control in the intensive rearing of marine larvae. *Aquaculture* **177**, 333-343. doi: 10.1016/S0044-8486(99)00096-4.

Skulstad, O. F., Karlsen, Ø., Fosseidengen, J. E., Kristiansen, T. S., Taranger, G. L. and Oppedal, F. (2013a). Vertical distribution and sexual maturation iin cage-farming of Atlantic cod (*Gadus morhua* L.) exposed to natural or continuous light. *Aquaculture Research* **44**(6), 903-917. doi: 10.1111/j.1365-2109.2012.03095.x.

Skulstad, O. F., Taylor, J., Davie, A., Migaud, H., Kristiansen, T., Mayer, I., Taranger, G. L., Olsen, R. E. and Karlsen, Ø. (2013b). Effects of light regime on diurnal plasma melatonin levels and vertical distribution in farmed Atlantic cod (*Gadus morhua* L.). *Aquaculture* **414-415**, 280-287. doi: 10.1016/j.aquaculture.2013.08.007.

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics* **16**, 393-430. doi: 10.1146/annurev.es.16.110185.002141.

Smedbol, R. K. and Stephenson, R. (2001). The importance of managing within-species diversity in cod and herring fisheries of the north-western Atlantic. *Journal of Fish Biology* **59**(Suppl. A), 109-128. doi: 10.1111/j.1095-8649.2001.tb01382.x.

Sonesson, A. K., Woolliams, J. A. and Meuwissen, T. H. E. (2005). Kinship, relationship and inbreeding. In: T. Gjedrem (Ed.), *Selection and breeding programs in aquaculture*. Dordrecht, the Netherlands: Springer, pp. 73-111.

Sotka, E. E. (2012). Natural selection, larval dispersal, and the geography of phenotype in the sea. *Integrative and Comparative Biology* **52**(4), 538-545. doi: 10.1093/icb/ics084.

Soulsbury, C. D., Iossa, G., Edwards, K. J., Baker, P. J. and Harris, S. (2007). Allelic dropout from a high-quality DNA source. *Conservation Genetics* **8**, 733-738. doi: 10.1007/s10592-006-9194-x.

Sparholt, H. (1994). Fish species interactions in the Baltic Sea. *Dana* **10**, 131-162.

Staber, C. and Mann, R. (2010). Quick and easy isolation of genomic DNA from *Drosophila* using the Maxwell® 16 instrument. Promega resources. www.promega.com, accessed 24 April 2014.

Stanley, R. R. E., deYoung, B., Snelgrove, P. V. R. and Gregory, R. S. (2013). Factors regulating early life history dispersal of Atlantic cod (*Gadus morhua*) from Coastal Newfoundland. *PLOS One* **8**(9), e75889, pp. 1-15. doi: 10.1371/journal.pone.0075889.

Star, B., Nederbragt, A. J., Jentoft, S., Grimholt, U., Malmstrøm, M., Gregers, T. F., Rounge, T. B., Paulsen, J., Solbakken, M. H., Sharma, A., Wetten, O. F., Lanzén, A., Winer, R., Knight, J., Vogel, J.-H., Aken, B., Andersen, Ø., Lagesen, K., Tooming-Klunderud, A., Edvardsen, R. B., Tina, K. G., Espelund, M., Nepal, C., Previti, C., Karlsen, B. O., Moum, T., Skage, M., Berg, P. R., Gjøen, T., Kuhl, H., Thorsen, J., Malde, K., Reinhardt, R., Du, L., Johansen, S., Searle, S., Lien, S., Nilsen,

F., Jonassen, I., Omholt, S. W., Stenseth, N. C. and Jakobsen, K. (2011). The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**, 207-210. doi: 10.1038/nature10342.

Steingrund, P. (2009). The near-collapse of the Faroe Plateau cod (*Gadus morhua* L.) stock in the 1990s: the effect of food availability on spatial distribution, recruitment, natural production and fishery. Ph.D. Thesis. University of Bergen, 110 pp.

Steingrund, P. (2014). Ferðing hjá toski undir Føroyum (Migration patterns of Faroe Plateau cod). In: *Speki*, television programme, National Faroese Broadcasting Company, 7th May 2014. Can be accessed via www.kvf.fo.

Steingrund, P. and Gaard, E. (2005). Relationship between phytoplankton production and cod production on the Faroe Shelf. *ICES Journal of Marine Science* **62**, 163-176. doi: 10.1016/j.icesjms.2004.08.019.

Steingrund, P., Hansen, B. and Gaard, E. (2005). Cod in Faroese waters. In K. Brander (Ed.), *Spawning and Life History Information for North Atlantic Cod Stocks*, ICES Cooperative Research Report 274, pp. 50-55.

Steingrund, P., Mouritsen, R., Reinert, J., Gaard, E. and Hátún, H. (2010). Total stock size and cannibalism regulate recruitment in cod (*Gadus morhua*) on the Faroe Plateau. *ICES Journal of Marine Science* **67**(1), 111-124. doi: 10.1093/icesjms/fsp240.

Steingrund, P. and Ofstad, L. H. (2010). Density-dependent distribution of Atlantic cod (*Gadus morhua*) into deep waters on the Faroe Plateau. *ICES Journal of Marine Sciences* **67**(1), 102-110. doi: 10.1093/icesjms/fsp230.

Stephenson, R. L. (2002). Stock structure and management structure: an ongoing challenge for ICES. *International Commission for the Exploration of the Seas (ICES) Marine Science Symposia* **215**, 305-314.

Sterner, T. (2007). Unobserved diversity, depletion and irreversibility. The importance of subpopulations for management of cod stocks. *Ecological Economics* **61**, 566-574. doi: 10.1016/j.ecolecon.2006.05.015.

Storz, J. F. (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology* **14**, 671-688. doi: 10.1111/j.1365-294X.2005.02437.x.

Svåsand, T., Jørstad, K. E., Otterå, H. and Kjesbu, O. S. (1996). Differences in growth performance between Arcto-Norwegian and Norwegian coastal cod reared under identical conditions. *Journal of Fish Biology* **49**, 108-119.

Svåsand, T., Kristiansen, T. S., Pedersen, T., Salvanes, A. G. V., Engelsen, R., Nævdal, G. and Nødtvedt, M. (2000). The enhancement of cod stocks. *Fish and Fisheries* **1**, 173-205. doi: 10.1046/j.1467-2979.2000.00017.x.

Swain, D. P., Chouinard, G. A., Morin, R. and Drinkwater, K. F. (1998). Seasonal variation in the habitat associations of Atlantic cod (*Gadus morhua*) and American plaice (*hippoglossoides platessoides*) from the southern gulf of St. Lawrence. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 2548-2561.

Swain, D. P. and Wade, E. J. (1993). Density-dependent geographic distribution of Atlantic cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. *Canadian Journal of Fisheries and Aquatic Sciences* **50**, 725-733.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P. and Bouvet, J. (1996). Reliable genotyping with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194. doi: 10.1093/nar/24.16.3189.

Taberlet, P., Waits, L.P. and Luikart, J. (1999). Noninvasive genetic sampling: look before you leap. *Trends in Ecology and Evolution* **14**, 323-327. doi: 10.1016/S0169-5347(99)01637-7.

Taggart, J. B. (2007). FAP: an exclusion-based parental assignment with enhanced predictive functions. *Molecular Ecology Notes* **7**, 412-415. doi: 10.1111/j.1471-8286.2006.01616.x.

Takezaki, N. and Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**, 389-399.

Teichert-Coddington, D. R. and Smitterman, R. O. (1988). Lack of response by *Tilapia nilotica* to mass

selection for rapid early growth rate. *Transactions of the American Fisheries Society* **117**, 297-300. doi: 10.1577/1548-8659(1988)117<0297:LORBNT>2.3.CO;2.

Tomkiewicz, J., Morgan, M. J., Burnett, J. and Saborido-Rey, F. (2003). Available information for estimating reproductive potential of Northwest Atlantic groundfish stocks. *Journal of Northwest Atlantic Fishery Science* **33**, 1-21.

Tóth, G., Gáspári, Z. and Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* **10**, 967-981. doi: 10.1101/gr.10.7.967.

Trippel, E. A. (1998). Egg size and viability and seasonal offspring production of young Atlantic cod. *Transactions of the American Fisheries Society* **127**(3), 339-359. doi: 10.1577/1548-8659(1998)127<0339:ESAVAS>2.0.CO;2.

Trippel, E. A. (2003). Estimation of male reproductive success of marine fishes. *Journal of Northwest Atlantic Fishery Science* **33**, 81-113.

Trippel, E. A., Benfey, T. J., Neil, S. R. E., Cross, N., Blanchard, M. J. and Powell, F. (2008). Effects of continuous light and triploidy on growth and sexual maturation in Atlantic cod, *Gadus morhua*. *CYBIUM* **32**(2), 136-138.

Trippel, E. A. and Morgan, M. J. (1994). Sperm longevity in Atlantic cod (*Gadus morhua*). *Copeia* **1994**, 1025-1029.

Van der Meeren, T., Jørstad, K. E., Solemdal, P. and Kjesbu, O. S. (1994). Growth and survival of cod larvae (*Gadus morhua* L.): comparative enclosure studies of Northeast Arctic cod and coastal cod from western Norway. *ICES Marine Science Symposia* **198**, 633-645.

Van der Meeren, T., Pedersen, J. P. og Kolbeinshavn, A.-G. (2005). Yngelproduksjon – larvefase. In: H. Otterå, G. L. Taranger and J. Borthen (Eds.), *Oppdrett av Torsk – Næring med Framtid*. Bergen, Norway: Norsk Fiskeoppdrett AS, pp. 85-111.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. and Shipleu, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Resources* **4**(3), 535-538. doi: 10.1111/j.1471-8286.2004.00684.x.

Van Wijk, S. J., Taylor, M. I., Creer, S., Dreyer, C., Rodrigues, F. M., Ramnarine, I. W., van Oosterhout, C. and Carvalho, G. R. (2013). Experimental harvesting of fish populations drives genetically based shifts in body siza and maturation. *Frontiers in Ecology and the Environment* **11**, 181-187. doi: 10.1890/120229.

Ventura, T., Aflalo, E. D., Weil, S., Kashkush, K. and Sagi, A. (2011). Isolation and characterization of a female-specific DNA marker in the giant freshwater prawn *Macrobrachium rosenbergii*. *Journal of Heredity* **107**, 456-461. doi: 10.1038/hdy.2011.32.

Verspoor, E., Consuegra, S., Fridjonsson, O., Hjorleifsdottir, S., Knox, D., Olafsson, K., Tompsett, S., Wennevik, V. and Garciá de Leániz, C. (2012). Regional mtDNA SNP differentiation in European Atlantic salmon (*Salmo salar*): an assessment of potential utility for determination of natal origin. *ICES Journal of Marine Science* **69**(9), 1625-1636. doi: 10.1093/icesjms/fss029.

Vincent, R. E. (1960). Some influences of domestication upon three stocks of brook trout (*Salvelinus fontinalis* Mitchill). *Transactions of the American Fisheries Society* **89**, 35-52.

Vinces, M. D., Legendre, M., Caldara, M., Hagihara, M. and Verstrepen, K. J. (2009). Unstable tandem repeats in promoters confer transcriptional evolvability. *Science* **324**(5931), 1213-1216. doi: 10.1126/science.1170097.

Walpole, R. E., Myers, R. H., Myers, S. L. and Ye, K. (2012). Probability & Statistics for Engineers & Scientists. 9th edn. Boston, USA: Prentice Hall, pp. 1-791.

Waples, R. S. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity* **89**, 438-450. doi: 10.1093/jhered/89.5.438.

Waples, R. S. (1999). Dispelling some myths about hatcheries. *Fisheries* **24**(2), 12-21. doi: 10.1577/1548-8446(1999)024<0012:DSMAH>2.0.CO;2.

Waples, R. S., Dickhoff, W. W., Hauser, L. and Ryman, N. (2008). Six decades of fishery genetics: taking stock. *Fisheries* **33**, 76-79.

Ward, R. D. (2000). Genetics in fisheries management. *Hydrobiologia* **420**, 191-201.

doi: 10.1023/A:1003928327503.

Ward, R. D. (2006). The importance of identifying spatial population structure in restocking and stock enhancement programmes. *Fisheries Research* **80**, 9-18. doi: 10.1016/j.fishres.2006.03.009.

Weir, B. S. and Cockerham, C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Wesmajervi, M. S., Westgaard, J.-I. and Delghandi, M. (2006). Evaluation of a novel pentaplex microsatellite marker system for paternity studies in Atlantic cod (*Gadus morhua*). *Aquaculture Research* **37**, 1195-1201. doi: 10.1111/j.1365-2109.2006.01549.x.

West, C. J. and Larkin, P. A. (1987). Evidence for size-selective mortality of juvenile sockeye salmon (*Oncorhynchus nerka*) in Babine Lake, British Colombia. *Canadian Journal of Fisheries and Aquatic Sciences* **44**(4), 712-721. doi: 10.1139/f87-086.

Westgaard, J.-I. And Fevolden, S.-E. (2007). Atlantic cod (*Gadus morhua* L.) in inner and outer coastal zones of northern Norway display divergent genetic structure at non-neutral loci. *Fisheries Research* **85**(3), 306-315. doi: 10.1016/j.fishres.2007.04.001.

Whitlock, M. C. and McCauley, D. E. (1999). Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. *Journal of Heredity* **82**, 117-125. doi: 10.1038/sj.hdy.6884960.

Winters, G. H. and Wheeler, J. P. (1985). Interaction between stock area, stock abundance, and catchability coefficient. *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 989-998.

Wirgin, I., Kovach, A. I., Maceda, L., Roy, N. K., Waldman, J. and Berlinsky, D. L. (2007). Stock identification of Atlantic cod in U.S. waters using microsatellite and single nucleotide polymorphism DNA analyses. *Transactions of the American Fisheries Society* **136**(2), 375-391. doi: 10.1577/T06-068.1.

Xu, J., Turner, A., Little, J., Bleecker, E. and Meyers, D. (2002). Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Human Genetics* **111**, 573-574. doi: 10.1007/s00439-002-0819-y.

Appendix: communications of research

Peer-reviewed publication

Andersen, Ø., De Rosa, M. C., Pirolli, D., Tooming-Klunderud, A., Petersen, P. E. and André, C. (2011). Polymorphism, selection and tandem duplication of transferrin genes in Atlantic cod (*Gadus morhua*) – Conserved synteny between fish monolobal and tetrapod bilobal transferrin loci. *BMC Genetics* **12**(51), pp. 1-14. doi: 10.1186/1471-2156-12-51.

Communications at conferences

Oral presentations

Petersen, P. E., Taggart, J., Penman, D., Dahle, G. and Gislason, H. Comparison of Atlantic cod from the Faroe Bank and the Faroe Plateau by genetic analyses – initial studies. Cod Farming in Nordic Countries, Reykjavík, Iceland, 30 September – 1 October 2008.

Petersen, P. E., Taggart, J., Penman, D. Dahle, G. and Patursson, Ø. Genetic studies of Atlantic cod (*Gadus morhua*) from the Faroe Bank and the Faroe Plateau. FarGen Summit, Tórshavn, Faroe Islands, 23 – 24 September 2012.

Poster presentation

Petersen, P. E., Taggart, J., Penman, D., Dahle, G. and Gislason, H. Genetic studies of Atlantic cod (*Gadus morhua*) from the Faroe Bank and the Faroe Plateau. Institute of Aquaculture Ph.D. Research Conference, University of Stirling, Stirling, Scotland, 29 October 2008.

Other communications

Public seminars

"Ílegukanningar av landgruns- og bankatoski" ("Genetic studies of Faroe Plateau and Faroe Bank cod"). Presented at two occasions in March 2012, for the public (arranged by the Biologists´ Society) and for the staff working at the Faroe Marine Research Institute.

"Ílegur í sambandi við aling" ("Genetics in aquaculture"). A seminar held on national tv, programme "Speki", 16 April 2014. Can be accessed via www.kvf.fo.

Short articles in "Alitíðindi" (Newsletter of the Aquaculture Research Station of the Faroes)

"Míni, tíni og okkara børn – familjuviðurskifti í alitoski" ("Family relationships in cultured cod"), vol. 1, 2011.

"Transferrin ílega í toski" ("Transferrin genes in cod"), vol. 2, 2011.

"Funnið ílegumarkørar, sum vísa týðandi mun millum landgruns- og bankatosk" ("Identification of genetic markers that can distinguish between Faroe Plateau and Faroe Bank cod"), vol. 2, 2014.