UNIVERSITY OF NORDLAND

MASTER THESIS

Population Genetics of Spawning and Juvenile cod (*Gadus morhua*)

Silje Svendsen

Abstract

The Atlantic cod (*Gadus morhua* L.) consists presumably of several units, but the main ones are the migrating North East Atlantic Cod (NEAC) and the stationary Norwegian Costal Cod (NCC). These are assumed to be two different populations, but due to their intermingling at common spawning grounds in northern Norway, discrimination and management of the two putative populations is challenging. After 80 years of research on the population structure of cod in this area, the results are still inconclusive. Two possible hypotheses might explain the population connectivity in *Gadus morhua*, the "divergent selection hypothesis" which assumes interbreeding and de novo directional selection on each year class, and the "historical isolation hypothesis" which assumes a historical period of allopatry. To test these hypotheses a total of 480 individuals, 144 putative NEAC, 211 putative NCC and 125 cod larva, were genotyped at seven microsatellites and analyzed for allelic- richness and frequencies, Hardy-Weinberg equilibrium (HWE) and population differentiation (F_{ST}). The microsatellites were tested for neutrality to natural selection to detect outlier loci. The outlier test found only two of the seven loci to be neutral, while two were under positive selection and three under balancing selection. The tests were done including all seven microsatellites, and additional testing of deviation from HWE and F_{ST} was done including solely the neutral microsatellites. The pair-wise F_{ST} estimates found a high amount of the sampling locations to be significantly different from each other and the tests for HWE found deviation at both neutral and non-neutral loci.

The most interesting finding in this study was the detection of deviation from HWE within the larvae sample from Vesterålen, while all other sampling locations were seemingly homogenous samples. The deviation in the Vesterålen samples was detected at both neutral and non-neutral loci, and also at the multilocus F_{IS} estimations. These findings might be interpreted to support the "historical isolation hypothesis".

Table of content

Introduction

The cod have been, and still is of great economical importance in Norway. In 2011 the landed catch of cod in Norway was 340 099 tons (live weight) with a landing value of 3,9 billion NOK (Anonymous 2012a). The capture of cod makes out 15 % of the landed catches in Norway and 24 % of the total landing value. This makes cod fisheries the most valuable fishery in Norway (Anonymous 2012a). On a global scale the total capture of cod registered by the Food and Agriculture Organization of the United Nations (FAO) was in 2009, 865 224 tons (Anonymous 2012b).

The Atlantic cod

The coast of Norway presumably consists of several cod units from Stad (62°N) to the Russian border, with North East Arctic cod (NEAC) and Norwegian costal cod (NCC) as the main ones (Johansen et al. 2009a).The NCC is found from the kelp belt to depths of 500 m. They spawn in fjords, and spend their entire life along the Norwegian coast. But they also intermingle with the NEAC in the Vesterålen/Lofoten area. Juvenile cod settle at shallower waters, 0 – 20 m (Anonymous 2012c). According to the Institute of Marine Research the NEAC is the largest cod population in the world and is found mostly along the bottom, but in the Barents Sea it can stay in the open water masses during parts of the year. They spawn of the coast of the Vesterålen/Lofoten islands in February to April. Both eggs and larvae drift with the northwards Norwegian coastal current into the Barents Sea and the juveniles settles in late fall. Most of the population is found in the Barents Sea at the warm part of the polar front (Anonymous 2012c) (Fig. 1). NEAC and NCC differ in life history. NEAC migrate over long distances from their feeding ground in the Barents Sea to their spawning grounds in Lofoten, while NCC migrate only locally (Berg and Albert 2003). It has been suggested that some of NEAC remain in fjords for a prolonged time in their immature phase, depending on abiotic factors such as wind, current, mixing of water masses etc. (Westgaard and Fevolden 2007). From February to May, NEAC and NCC are present simultaneously at the same spawning grounds off the coast of Northern Norway. They might not intermingle randomly and might not interbreed (Nordeide 1998; Nordeide and Folstad 2000). Cod probably do not mate promiscuous, but have a conventional lekking mating system, which could prevent interbreeding between NEAC and NCC (Nordeide and Folstad 2000).
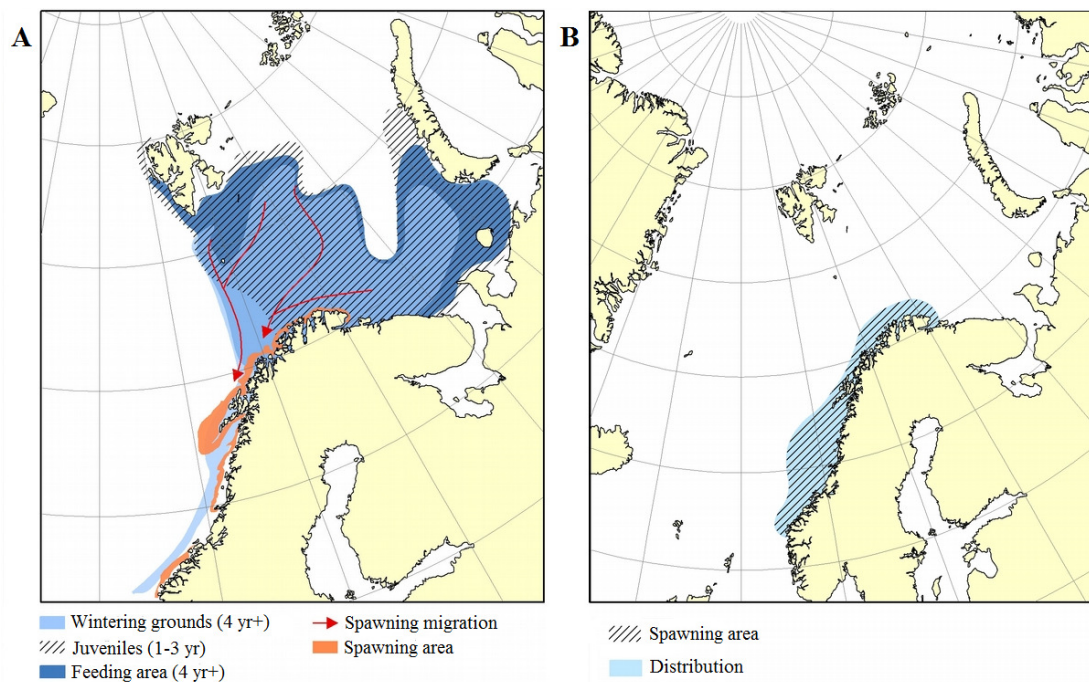
Fig. 1. Distribution of NEAC (A) and NCC (B). Maps adapted from www.imr.no (Anonymous 2012c).

Fishery management

The main goal of fishery management is to maintain a high biological diversity and at the same time get a high dividend from the fishery. North of Stad (62°N) NEAC and NCC intermingle, especially from February to May when NEAC migrate south to spawn (Fig. 1). This gives catches with a mixture of NEAC and NCC, hence managing them as two different populations is difficult. The recruitment and spawning stock biomass (SSB) for NCC was in 2011 reported to be at the lowest ever observed (Anonymous 2012c). The International Council for the Exploration of the Sea (ICES) have advised no catch for this population and a recovery plan have been made, which entails that the quotas for the following years on are continuously evaluated by Institute of Marine Research to ensure a safe management for NCC (Anonymous 2010d). For the NEAC the situation is completely different. The recruitment is good, the size of the SSB is the highest observed and the mortality is low. Based on these data the recommended quota for the NEAC in 2012 is set for 751 000 tons (Anonymous 2010d; Anonymous 2012e).

Population genetics

Studying population genetics of marine species such as cod is challenging, since there are no obvious barriers in the ocean, fish migrate, and they may intermingle. The most common

approach for detecting population structure is to sample from two (or more) putative populations and examine them for genetic, morphological, meristic, phenotypic etc. traits (Waples 1998). For 80 years scientists have tried determining the population structure of NEAC and NCC by this approach. In population genetics it is preferred to use neutral loci, as demography and evolutionary history will have affected these similarly across the genome, while loci under selection, or linked to such genes responds to contrasting environments (Luikart et al. 2003). Differences between putative populations at loci neutral to natural selection could indicate population divergence (Beebee and Rowe 2008). Nordeide et al. (2011) reviewed 54 papers that have studied the relationship between NEAC and NCC and conclude that we still, after all these years of research on this species, do not know if NEAC and NNC make out one large population were the genetic differences are due to gene flow with geographical distance as the limiting factor or several non-interbreeding groups. Throughout the decades of research a lot of different markers have been used to determine their relationship, both phenotypic and genetic markers (Nordeide et al. 2011) (Table 1).

Table 1: Phenotypic and genetic markers used to discriminate NNC and NEAC with reference to some of the work done with these (Nordeide et al. 2011).

	Marker	Reference
Phenotypic	Otoliths	(Rollefsen 1933)
	Migration	(Hylen 1964)
	Blood type	(Møller 1968)
	Number of vertebras	(Løken and Pedersen 1996)
	Growth	(Otterlei et al. 1999)
	Sexual maturation	(Berg and Albert 2003)
	Body shape (K-factor)	(Johansen et al. 2009b)
Genetic	Hemoglobin HbI	(Frydenberg et al. 1965)
	Allozymes	(Mork et al. 1981)
	Minisatellites	(Dahle 1994)
	RFLP	(Pogson et al. 1995)
	mtDNA cytochrome *b*	(Árnason and Pálsson 1996)
	PanI/SypI	(Fevolden and Pogson 1997)
	Microsatellites	(Westgaard and Fevolden 2007)
	SNPs	(Moen et al. 2008)

Population connectivity based on phenotypic markers

Gunnar Rollefsen started the research on the relationship between NEAC and NCC in 1933, by studying their otoliths. He found them to have different shapes and distance between the growth zones from the cods first two years (Rollefsen 1933). However, research has suggested that this subjective measurement of species determination is not reliable. Offspring of NEAC and NCC that grows up in a similar environment show no difference in the shape of the otoliths (Johansen et al. 2009a). Apart from the otoliths there are other phenotypic traits that separate the two populations. The NEAC has a long lean body probably due to their migration patterns, while the NCC are more corpulent (Johansen et al. 2009b). Møller et al. (1968) found a correlation between hemoglobin, transferrin, blood and otolith types when studying cod collected along the entire Norwegian coast and the Barents Sea. They found that when the samples were treated as one large population it was not in Hardy-Weinberg Equilibrium (HWE), but when separating them into NEAC and NCC according to otolith types they were. They suggested that despite sharing spawning grounds at the same time, NEAC and NCC were two genetically separated populations (Møller 1968). Differences between NEAC and NCC has been suggested to be found at the age of maturity, where NEAC reaches maturity close to a year later than NCC (Berg and Albert 2003). Also, NCC has a lower mean vertebrae number than NEAC (Løken and Pedersen 1996).

Population connectivity based on genetic markers

The frequency of the hemoglobin HbI^1 allele varies along the Norwegian coast, with frequencies of 60 % at the Skagerrak coast and decreasing northward along the coast down to 10 – 15 % in the Lofoten area, and even lower in the Barents Sea (Frydenberg et al. 1965). Fevolden and Pogson (1997) suggested that HbI^1- locus is not a suitable marker for population differentiation, as it might not be neutral to selection. Allozyme markers have been used for population differentiation, some studies show population structuring at non-neutral allozymes, while others have found no subpopulation structuring at allozyme markers (Nordeide et al. 2011). Minisatellites has not shown any population differentiation for cod (Dahle 1994), neither has the studies using the mitochondrial cytochrome *b* DNA (Árnason and Pálsson 1996). Pogson et al. (1995) compared allozymes and restriction fragment length polymorphism (RFLP) and found the RFLPs to show differences between NEAC and NCC whereas the allozymes did not (Pogson et al. 1995). The *Pan*I locus has been a widely used marker in studies of population structure of cod. This highly conserved structure consisting of four transmembrane domains, two intravesicular loops and two cytoplasmic tails allows

mutations to be identified and localized to distinct domains (Pogson 2001). Fevolden and Pogson (1997) found the *Pan*I^A allele to be predominating in NCC, while in NEAC the *Pan*I^B allele dominated (Fevolden and Pogson 1997). The same has been seen at the integral synaptic vesicle membrane protein *Syp*I, which is thought to be a cellular isoform of synaptophysin called pantophysin (*Pan*I). The *Pan*I locus was later shown influenced by natural selection, and therefore non-neutral (Pogson 2001). Later, microsatellites (Box 2) became a popular molecular marker in the search for the genetic population structure for NEAC and NCC (Karlsson and Mork 2005). But microsatellites are not always neutral (Beebee and Rowe 2008). Among the most commonly used microsatellites in studies on cod are GMO34 and GMO132, microsatellite loci that have shown to be non-neutral. GMO132 seems to be under selection, while GMO34 shows linkage disequilibrium with the non-neutral *Pan*I gene. Westgaard and Fevolden (2007) suggest that these markers should not be dismissed but used as a supporting tool for discriminating between NEAC and NCC (Westgaard and Fevolden 2007). The most resent tool for detecting population structure is the use of single nucleotide polymorphisms (SNPs). These SNPs are found in non-coding regions of the genome and show sequence variation at single nucleotides (Beebee and Rowe 2008). Some SNPs have shown high levels of population divergent between NEAC and NCC, suggesting diversifying selection and minor gene flow between the two putative populations. This molecular marker is predicted to be a powerful tool for the future research on this topic (Moen et al. 2008).

The present study

To explain the differences found between NEAC and NCC Nordeide et al. (2011) present two hypotheses, the "divergent selection hypothesis" and the "historical isolation hypothesis". The "divergent selection hypothesis" assumes interbreeding at the common spawning grounds and de novo directional selection on each juvenile year class. Differences between NEAC and NCC may be significant for non-neutral alleles due to selection at their habitats (coast or Barents Sea), whereas neutral alleles would not show any differences due to the interbreeding. The other hypothesis, the "historical isolation hypothesis", assumes that NEAC and NCC have been completely or partially isolated from each other. The differences between NEAC and NCC would affect both neutral and non-neutral alleles, but neutral less than the latter. The differences between the two would be due to genetic drift and founder effects for each group (Nordeide et al. 2011). This study aims to examine if testing of HWE and F_{ST} will support any of these hypotheses, and if there is interbreeding at the Lofoten spawning grounds followed

by selection at an early stage differentiating them. A total of 480 individuals were sampled (285 adults, 70 juveniles and 125 larvae) and genotyped at seven microsatellites. Larvae samples were collected at their yolk sac/pre-flexion stage (Munk et al. 2005), with an age estimated to be between 4 and 30 days post hatching (Auditore et al. 1994). The sampling was done north of the well known spawning areas of the NEAC were the adults were collected, due to the northward drift of the eggs and larvae (Anonymous 2012c). The sampling of NCC was done at two locations south of the main spawning areas of NEAC, Salten and Hordaland, these samples are expected to be less influenced by the NEAC due to the geographical distance, especially the samples from Hordaland. NCC were also collected from a fjord in Troms, an area were NEAC and NCC might intermingle (Westgaard and Fevolden 2007). The larval samples were sequenced for the mitochondrial COI gene in order to distinguish them from similar looking larvae of other species. Due to their early life stage, natural selection has presumably not had time to change their allele frequencies and deviations from the expectations of Hardy-Weinberg equilibrium are expected to be found. The microsatellites GMO34 and GMO132 are expected to be non-neutral as previously described (Westgaard and Fevolden 2007). Differentiation between the two putative populations only at non-neutral loci and not at neutral would be coherent with the "divergent selection hypothesis", while differences at both neutral and non-neutral loci would be coherent with the "historical isolation hypothesis" (Nordeide et al. 2011).

Materials and Methods

Sampling

In this study a total of 480 individuals were studied, 285 adults, 70 juveniles and 125 larvae (Table 2) from various locations (Fig. 2). The sampling of adult cod was done in four different locations in 2010, and all samples were of gills and muscles. A collection of adult cod was done in the Lurefjord area, Hordaland (HO), in the Salten (SA) area and the collection of cod from Troms (TR) was done in Ullsfjorden at various locations (Fig. 2). In addition to being sampled from the same locations as the adults, juveniles from Troms were also collected at Lakselvbukta, Jøvik and Balsfjord. These samples were also assumed to be of the putative NCC population as they were collected outside the spawning season of NEAC and in fjords. The samples from the Lofoten (LO) area were collected at three different locations (Table 2). These samples were assumed to be of the putative NEAC population as they were collected during their spawning season at common spawning sites and west of the Lofoten islands. The larvae samples used were collected in the Vesterålen (VE) area, trawling for larvae was also done in the Salten area, but these samples did not contain any cod larvae.
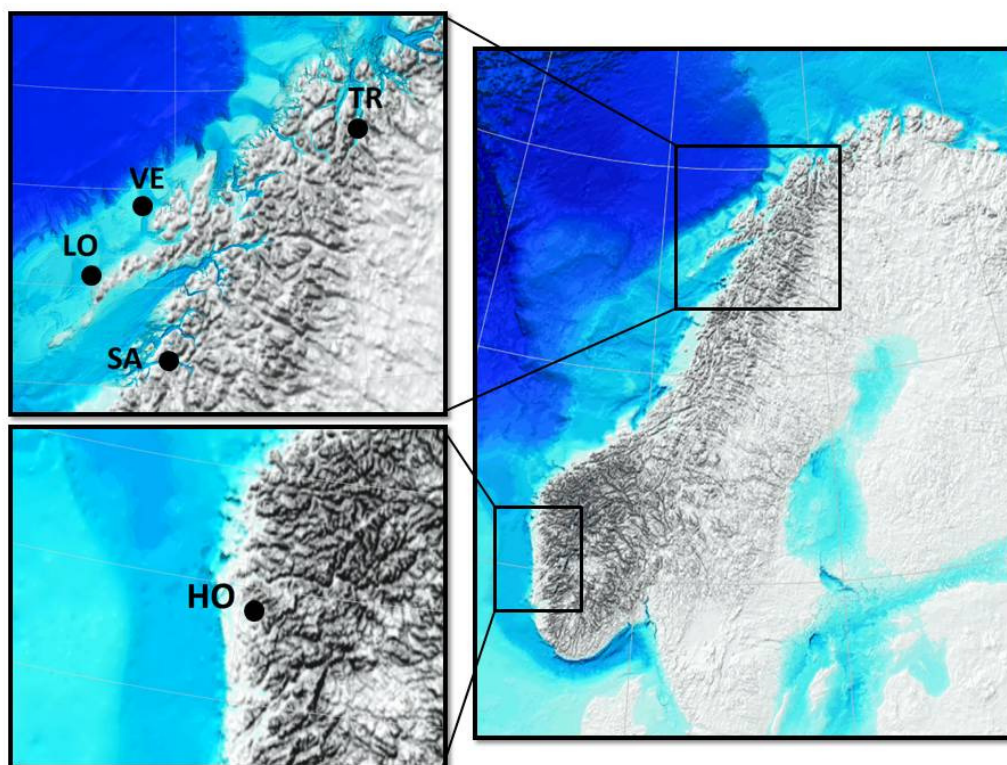

Fig. 2. Map showing sampling locations of cod (TR: Troms, VE: Vesterålen, LO: Lofoten, SA: Salten, HO: Hordaland) (Anonymous 2012g).

Table 2. Sampling site, position, number (N) of cod, date of sampling and life stage

Samples	Location	Position	N	Date	Life stage
HO	Lurefjord	N60°41′ E005°08′	48	2010	Adults
SA	Saltstraumen	N67°13′ E14°36′	45	07.12-12.2010	Adults
LO	Ballstad	N68°40'92" E013°06'68"	48	03.11.2010	Adults
	Laukvik	N68°25'58" E013°56'08"	48	02.22.2010	Adults
	Laukvik	N67°44'79" E012°21'94"	48	03.25.2010	Adults
TR	Eidstranddjupet	N68°53'54" E019°59'57"	24	09.29.2010	Adults
	Eidstranddjupet	N68°53'54" E019°59'57"	24	09.29.2010	Juveniles
	Lyngen Arnøy	N70°02'25" E020°16'38"	24	10.28.2010	Adults
	Lyngen Arnøy	N70°02'25" E020°16'38"	24	10.28.2010	Juveniles
	Lakselvbukt	N69°26'05" E019°38'30"	2	08.24.2010	Juveniles
	Jøvik	N69°36'16" E019°49'05"	5	08.25.2010	Juveniles
	Balsfjord	N69°14'03" E019°22'45"	15	08.27.2010	Juveniles
VE	Vesterålen	N68°55'20" E014°24'00"	125	04.08.2010	Larvae

Sorting of cod larvae

The larvae were sampled with a wp2 plankton net, giving samples containing many different organisms. The samples were stored and cleaned in ethanol. The cod larvae were distinguished from the rest of the sample containing other fish larvae, copepods, small jelly fishes etc. The initial determination of cod was based on morphological traits, mainly the two post anal- and the ventro-caudal melanophores (Munk et al. 2005).

Mitochondrial COI gene for species determination

For all the larvae samples and 48 adults from two adult sampling sites (Lofoten and Troms) the mitochondrial COI sequence (Box 1) was sequenced in order to compare to other gadidae in a phylogenetic tree, as species determination based on morphological traits is subjective and not reliable. The DNA extraction was done using Nexttec[TM] Genomic DNA Isolation Kit (Hilgertshausen, Germany). A 15 µl PCR reaction was set up for each sample with 2 µl DNA, 7, 5 µl Ampli Taq Gold 360, 0, 45 µl forward primer, 0, 45 µl reverse primer (Table 3) and 4, 6 µl Nuclease-free water. The samples were amplified in a Verity Thermo Cycler (Applied Biosystems, Foster City, CA, USA) with denaturation at 95°C for 5 minutes, followed by 45 cycles of denaturation at 95°C for 20 seconds, annealing at 49°C for 30 seconds and extension

at 72°C for 1 minute, and 7 minutes at final extension at 72°C. The PCR product was run on a 1 % agarose gel, and samples with seemingly low concentration were pulled together with an additional PCR amplification of those samples. The PCR products were first cleaned with ExoSAP-IT (USB, Cleveland, OH, USA) and then salts, dyes and nucleotides was cleaned out with Sephadex (Sigma-Aldrich®, Saint Louis, MO, USA). The concentration of PCR product was measured on a Qubit^{TM} flurometer. The needed amount of PCR product was 5 – 30 ng/µl. A Big Dye reaction was done and the product was again cleaned with Sephadex and run on the 3500xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and read in the instrument specific software SEQUENCING ANALYSIS SOFTWARE version 5.4 according to the user bulletin (PN 4401738).

Box 1. Mitochondrial DNA.

The mitochondrial genome is confined to the mitochondria which is located in the cell cytoplasm. In a diploid cell the mitochondrial genome (mtDNA) can represent 1 – 10 % of the total DNA contents. The animal mtDNA range in size from 14 kbp to 42 kbp and are circular double-stranded molecules containing two rRNA genes, 22 tRNAs and 13 protein-encoding genes (Beebee and Rowe 2008). In cod the mtDNA is 16,696 bp in length (Johansen and Bakke 1996), and shares the organization of the mitochondrial gene content with most fishes and mammals. But cod mtDNA also contains two unusual nonoding sequence elements. In the control region there is a heteroplasmic 40-bp tandem repeat, and between the tRNA^{Thr} and tRNA^{Pro} a 74 – 102-bp long spacer region (Johansen et al. 2009b).

The nucleotide substitution rate is much higher in mtDNA than nuclear DNA, giving it higher interspecific genetic variation. This combined with non-recombinant maternal inheritance makes the mtDNA a powerful tool in species recognition (Beebee and Rowe 2008). Compared to the African clawed toad (*Xenopus laevis*) the most conserved mtDNA proteins in cod are the three mitochondrially encoded cytochrome oxidase subunits (COI, II and III) (Johansen et al. 1990). These subunits plays different roles, but are combined responsible for the catalytic function of the holoenzyme (Cantatore and Saccone 1987).

Table 3. COI forward and reverse primer and annealing temperature used in the PCR to examine the mitochondrial COI gene.

COI	Sequence	T_{ann}
F	TCGACTAATCATAAAGAYATYGGCAC	45°C
R	ACTTCAGGGTGACCGAAGAATCAGAA	

The sequences were opened and trimmed in FINCHTV version 1.4.0 (Geospiza, Seattle, WA, USA) and then blasted in the National Center for Biotechnology Information (NCBI) (Altschul et al. 1997). The sequences were aligned in BIOEDIT version 7.1.3.0 (Hall 1999) and collapsed into haplotypes in Fabox (Villesen 2007). Orthologues were identified at the internet site FishBase (Froese and Pauly 2012), and obtained in NCBI via MEGA version 5.03 (Tamura et al. 2011). All the Gadidae COI sequences available were included. A phylogenetic Maximum Likelihood tree was made in MEGA 5.03 with 1000 bootstrapping and the Kimura 2-parameter model. The tree file (newik) was converted into a nexus file at phylogeny.fr (Dereeper et al. 2008) and opened in TREEVIEW 1.6.6 (Page 1996) where all sequences obtained from NCBI (except *Gadus morhua*) were selected as outgroups and the tree was rooted with these. One larval individual was not sequence successfully for the COI gene, and therefore this individual is not included in the tree. But the microsatellite fragments were amplified successfully and this individual was not removed from the rest of the study because it did not deviate or stand out in any of the microsatellite tests.

Microsatellites

Seven microsatellites (Box 2) loci were amplified by Polymerase Chain Reaction (PCR). GMO03, GMO19, GMO34, GMO35, GMO36 (Miller et al. 2000), GMO132 (Brooker et al. 1994) and PGMO58 (Jakobsdóttir et al. 2006). Each primer had fluorescent labels (Table 4) where PET is red, VIC is green, NED is yellow and 6-FAM is blue fluorescent dye. A 10 µl PCR reaction contained for each sample; 1 µl diluted DNA (Table 4), 2, 5 µl Ampli Taq Gold 360, 0, 4 µl forward primer, 0, 4 µl reverse primer and 1, 42 µl Nuclease-free water. The PCR was done on a Verity Thermo Cycler (Applied Biosystems, Foster City, CA, USA) with denaturation at 95°C for 1 minute, followed by primer specific annealing temperature (Table 4) for 20 seconds and 72°C for 25 seconds. This was repeated five times. Followed by five similar cycles where the denaturation lasted for 30 seconds instead of 1 minute. Then the primer specific number of cycles (minus the previous 10 cycles) was done with denaturation

10

at 95°C for 20 seconds, the specific annealing temperature for each primer (Table 4) for 20 seconds and extension at 72°C for 25 seconds, and finally 20 minutes at final extension at 72°C. A few samples were difficult to amplify in the PCR or difficult to read after sequencing, these were redone with 3 times the primer volume and amplified at two degrees less in the PCR.

Table 4. Primers for PCR amplification of cod microsatellites, fluorescent label, annealing temperature (T_{ann}), PCR cycles and the DNA concentration used.

Primer	Sequence	Label	T_{ann}	Cycles	DNA dilutions
GMO19		PET	55 °C	38	1:50
F	CAC AGT GAA GTG AAC CCA CTG				
R	GTC TTG CCT GAT AGT CAG CTT G				
GMO34		VIC	55 °C	38	1:50
F	TCC ACA GAA GGT CTC CTA A				
R	GGT TGG ACC TCA TGG TGA A				
GMO36		NED	55 °C	38	1:50
F	GGT GAT GGA GGC TCT AGT				
R	ACC GCA TSC CCT TTT CA				
PGMO58		6-FAM	55 °C	38	1:100
F	CAG CAG ATT GAT GGG TTT AGC				
R	GGA AAC CCT AAG AAC GAG				
GMO35		VIC	53 °C	40	1:50
F	GGA GGT GCT TTG AAG ATG				
R	CCT TAT CAT GTA CGT TGT TAA C				
GMO132		6-FAM	52 °C	38	1:50
F	GGA ACC CAT TGG ATT CAG GC				
R	CGA AAG GAC GAG CCA ATA AC				
GMO03		NED	48 °C	40	1:50
F	AGG CAC GCA GGT GGA CAG GAA C				
R	GCA GCA CGA GAG AGC TAT TCC TC				

Fragment visualization of the PCR products was done on the 3500xl capillary sequencer (Applied Biosystems, Foster City, CA, USA), and alleles were scored as homozygotes or

heterozygotes using the instrument specific software GENEMAPPER®SOFTWARE version 4.1 according to the software installation and administration guide (PN 4403614).

Box 2. Microsatellites.

In population genetics one would ideally use molecular markers that are cheap and easy to develop and use, highly polymorphic and neutral to natural selection. Few markers obtain all these qualities, but among of the most widely used markers are microsatellites (Beebee and Rowe 2008). Also in the studies of the relationship between NEAC and NCC these markers are widely used (Nordeide et al. 2011). Microsatellites are tandem repeats of 1 – 6 nucleotides which are mostly found in non-coding regions. The loci are typically between 5 and 40 repeats in length. To amplify a useful amount of DNA, the polymerase chain reaction (PCR) is employed. The microsatellites are species-specific, and therefore less prone to cross-contamination from other species. When studying present day demography or connectivity, microsatellites are very useful due to their high-mutation rate which gives a high allelic diversity (Selkoe and Toonen 2006).

Statistical tests

F-statistics was used for the F_{ST} and F_{IS} estimates (Box 3). The F_{ST} was estimated in GENETIX 4.05.2 (Belkhir et al. 1996-2004) with 1000 permutations. The same amount of permutations was used to estimate the F_{IS} values, locus per locus. Expected and observed heterozygosity, and the allelic frequencies which gives the number of alleles at a locus were also estimated in GENETIX 4.05.2 (Belkhir et al. 1996-2004).The factorial correspondence analysis (FAC) which visualize the genetic similarities between populations based on the allelic frequencies was done in GENETIX (Belkhir K 1996-2004), while the graphics was for this analysis and the allelic frequencies were done in EXCEL (Microsoft 2007). In STRUCTURE 2.3.3 (Pritchard et al. 2010) a bar plot was made to estimate membership coefficient for each individual in each K (=2) this was done with a admixture model where each individual is deemed to have drawn some fraction of its genome from each of the K populations, and the LOCPRIOR model. The LOCPRIOR model uses the sampling locations to assist with the clustering (Pritchard et al. 2010). The analysis was done with a burn-in of 5 000 000 followed by 5 000 000 MCMC repetitions. Both the FAC and the STRUCTURE bar plot clusters individuals and estimates the memberships to a population and are recommended to be used mainly as a guide (Pritchard et al. 2010). The Allelic Richness which estimates the

mean number of alleles per locus and compensates for the sampling sizes by using rarefaction was obtained in FSTAT 2.9.3.2 (Goudet 1995).

The outlier test was done in ARLEQUIN (Excoffier 2005) by detecting loci under selection with 50 000 simulations and 100 demes, a test which uses coalescent simulations to get the locus-specific p-values from F-statistics on the observed heterozygosity. To produce the graphics for the outlier test R version 2.15.1 was used. An outlier locus is a genomic location which shows extremely divergent patterns or behavior compared to the rest of the genome. Microsatellites are assumed to be neutral, but outlier (non-neutral markers) are common across data sets and may occurs due to various reasons such as strong natural selection in wild populations, selective sweeps that may cause linkage disequilibrium or even due to genotyping errors, like null alleles (Luikart et al. 2003).

Box 3. *F*-statistics.

The expected heterozygosity (H_E) is the expected proportion of heterozygous assuming the samples are under Hardy-Weinberg equilibrium (HWE) while the observed heterozygosity (H_O) is the mean proportion of individuals that are heterozygous across a locus or a set of loci. By using *F*-statistics one can measure the degree of inbreeding (F_{IS}) which shows the homozygous excess relative to the Hardy-Weinberg expectations. This is done by the estimators:

$$F_{IS} = \frac{(H_S - H_I)}{H_S}$$

where I = individual and S = subpopulation. The F_{IS}-values ranges from negative values which shows heterozygote excess, 0 which shows no inbreeding, to 1 which shows full inbreeding (Beebee and Rowe 2008).

The fixation index, F_{ST}, measures the degree of inbreeding of subpopulations relative to the total population. This is done by the estimators:

$$F_{ST} = \frac{(H_T - H_S)}{H_T}$$

where T = total population and S = subpopulation. The F_{ST}-values may range from 0 or negative values showing no population structure, to 1 which shows fully separate populations (Beebee and Rowe 2008).

Results

Species determination

The phylogenetic maximum likelihood (ML) tree shows all the COI haplotypes (represented as number in Fig. 4) clusters together with the obtained *Gadus morhua* sequence. This clustering, and the high bootstrap values shows good support for the larvae being cod. The ML tree is made as a phylogram and drawn to scale with branch lengths measured in the number of substitutions per site. Bootstrap values higher than 70 are shown (Fig. 4).
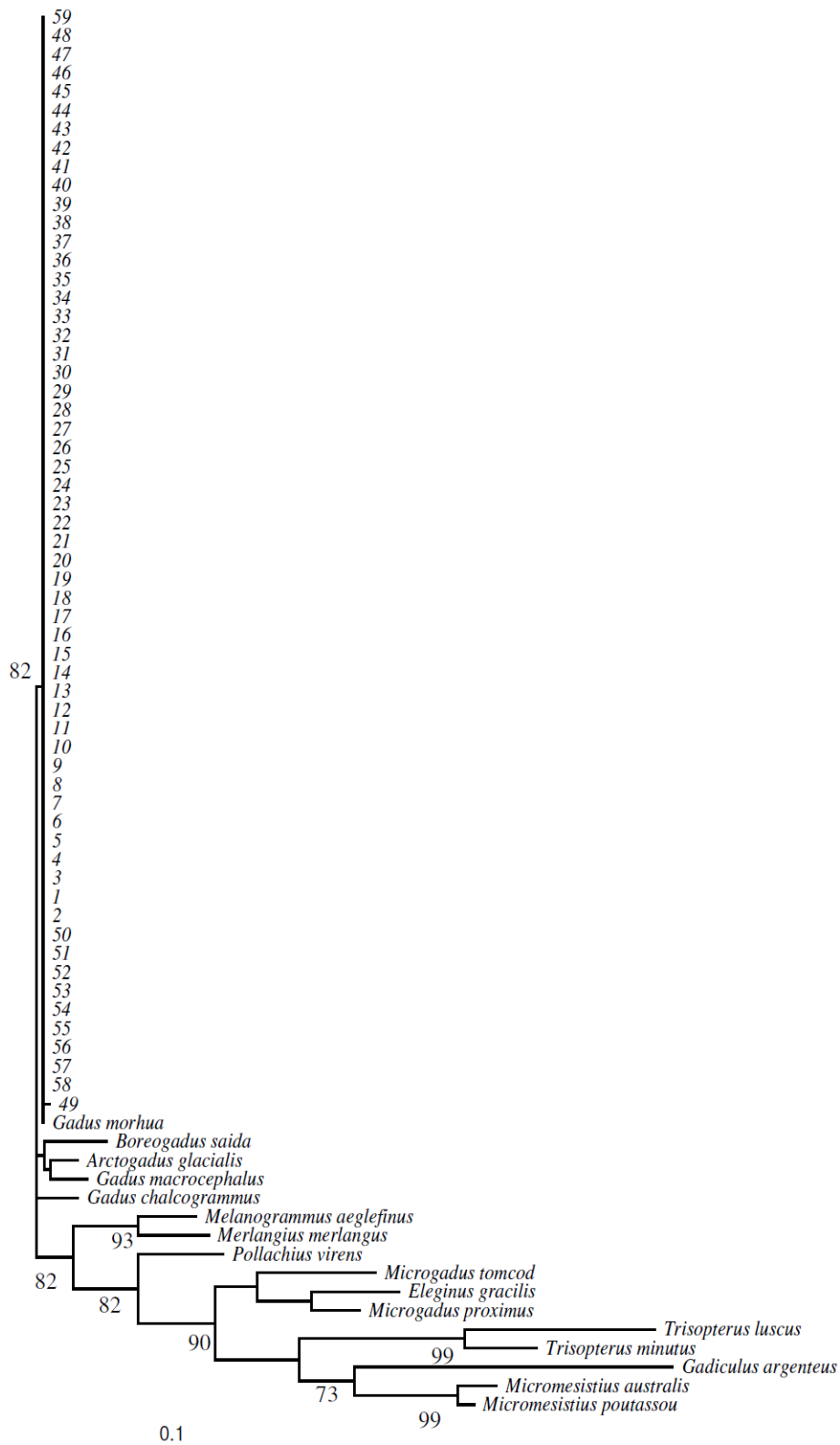
Fig. 3. Maximum likelihood tree showing all haplotypes (numbers at terminal nodes) of the COI gene from the larvae and the selected adults, compared with all other gadidaes orthologs sequences obtained from NCBI (Altschul et al. 1997). All bootstrap values higher than 70 is shown at the appropriate divergence points.

Microsatellites

Allelic richness

The mean allelic richness across all loci does not vary much between sampling sites. The Hordaland samples have the highest allelic richness with 9.786 and the Salten samples have the lowest with 7.828. The variation is much higher at single loci where the values vary from 2.0 at the GMO36 locus in the Salten samples to 20.587 at the GMO19 locus in the Troms samples (Table 5).

Deviation from Hardy-Weinberg equilibrium

A significant difference ($P < 0.05$) between expected and observed heterozygosity across all loci was found for all samples except for the Lofoten samples (F_{IS}) (Table 5). Considering single loci, many deviate from Hardy-Weinberg equilibrium (HWE). GMO34 is the only locus that does not deviate from HWE at any samples, while all samples deviate at the GMO19 locus. The Salten sampling location has four out of seven loci which deviate from HWE, and the Vesterålen sample site deviate at five loci (Table 5). After removing all loci under selection (Fig. 6) only two loci remained, GMO36 and PGMO58, loci which might be neutral to selection (Fig. 6). The F_{IS} estimates using solely these two loci show different results than the F_{IS} estimates including all seven loci. Only the Vesterålen samples (larvae) deviate from HWE based solely the two neutral loci (Table 5).

Table 5. Summary of genetic variation at seven microsatellites loci and five locations (Table 2). Expected (H_E) and observed (H_O) heterozygosity and F_{IS}-values for each locus and per sample site. Deviation from HWE in bold (P < 0.05). Allelic richness (A_r) per loci, mean for all loci and mean at all locations, and the number of alleles per locus (Na).

Location Locus		Troms (N=118)	Hordaland (N=48)	Lofoten (N=144)	Salten (N=45)	Vesterålen (N=125)	Mean A_r	Na
GMO03	A_r	4.602	2.896	4.455	3.956	5.703	4.322	10
	H_E	0.1821	0.2051	0.1201	0.2230	0.2083		
	H_O	0.1864	0.1875	0.1181	0.1556	0.1760		
	F_{IS}	-0.01941	0.09615	0.02055	**0.31250**	**0.15880**		
GMO19	A_r	20.587	18.466	19.089	13.976	19.374	18.298	26
	H_E	0.9235	0.9169	0.9167	0.8737	0.9096		
	H_O	0.7712	0.8125	0.7817	0.5682	0.7360		
	F_{IS}	**0.16905**	**0.12422**	**0.15076**	**0.35974**	**0.19469**		
GMO34	A_r	20.587	8.687	4.843	6.910	4.557	6.118	10
	H_E	0.2617	0.6708	0.1700	0.5956	0.2074		
	H_O	0.2373	0.6458	0.1597	0.5556	0.2000		
	F_{IS}	0.09752	0.04771	0.06403	0.07834	0.03950		
GMO35	A_r	7.992	7.896	8.050	8.000	8.020	7.993	16
	H_E	0.8253	0.7977	0.8287	0.8129	0.8282		
	H_O	0.8220	0.7708	0.7778	0.6279	0.6240		
	F_{IS}	0.00826	0.04424	**0.06487**	**0.23867**	**0.25032**		
GMO36	A_r	3.725	3.998	3.726	2.000	4.269	3.544	8
	H_E	0.5548	0.5586	0.4657	0.2311	0.4670		
	H_O	0.5932	0.4583	0.5069	0.1778	0.0720		
	F_{IS}	-0.06502	0.18966	-0.08502	0.24138	**0.84698**		
GMO132	A_r	9.839	20.559	9.940	14.000	10.064	12.880	28
	H_E	0.4846	0.9206	0.4550	0.6720	0.4351		
	H_O	0.4746	0.7917	0.4653	0.4884	0.4080		
	F_{IS}	0.02485	**0.15033**	-0.01920	**0.28409**	0.06636		
PGMO58	A_r	5.810	6.000	5.993	5.954	6.131	5.978	8
	H_E	0.5346	0.6439	0.6319	0.6889	0.5191		
	H_O	0.5593	0.6875	0.5324	0.7556	0.4240		
	F_{IS}	-0.04196	-0.05726	**-0.18359**	-0.08563	**0.18709**		
Mean A_r		8.307	9.786	8.014	7.828	8.303	8.448	
Mean H_E		0.5381	0.6734	0.4984	0.5853	0.5107		
Mean H_O		0.5206	0.6220	0.4916	0.4756	0.3771		
Multilocus F_{IS}								
All loci		**0.03679**	**0.08670**	0.01703	**0.19857**	**0.26520**		
Neutral loci		-0.05371	0.05760	-0.13759	-0.00326	**0.50005**		

Genetic differentiation

Across all seven microsatellites the pair-wise F_{ST} estimates show only the Vesterålen and Lofoten samples to not differentiate from each other. After removing the five non-neutral loci (Fig. 6) and estimating the F_{ST} based solely on the neutral loci (GMO36 and PGMO58) the Vesterålen and Lofoten samples still do not differentiate, and neither do the Troms and Hordaland samples, while all other sampling locations are significantly different from each other (Table 6).

Table 6. F_{ST} estimates from all seven loci (above diagonal) and from the two neutral loci (below diagonal) significant values in bold (P < 0.05).

	Troms	Hordaland	Lofoten	Salten	Vesterålen
Troms		**0.05637**	**0.00734**	**0.05922**	**0.00916**
Hordaland	0.0047		**0.07678**	**0.04631**	**0.07229**
Lofoten	**0.02664**	**0.03725**		**0.04570**	-0.00073
Salten	**0.13746**	**0.11327**	**0.07036**		**0.04025**
Vesterålen	**0.02836**	**0.04016**	-0.00203	**0.06569**	

Population structure

The Factorial Correspondence Analysis (FAC) shows overlap between all the sample locations, but the Hordaland and Lofoten samples cluster the farthest apart from each other, while the other sampling locations cluster between them (Fig. 4). The barplot from STRUCTURE (Pritchard et al. 2010) gives very similar results as the FAC, with the samples from Lofoten and Hordaland being the most divergent from each other, and the other sampling locations being situated between these (Fig. 5).
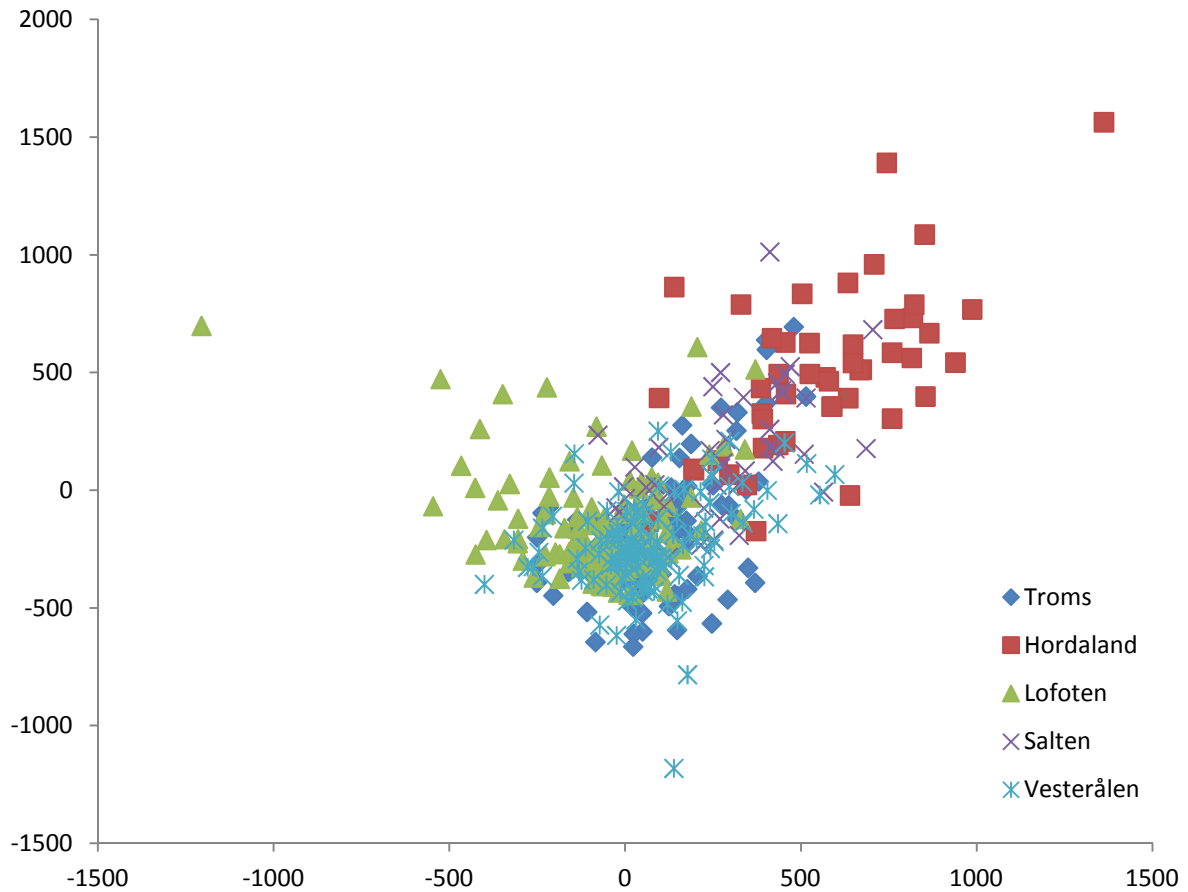
Fig. 4. Factorial Correspondence Analysis (FCA), showing all individuals from each sample site and their distribution
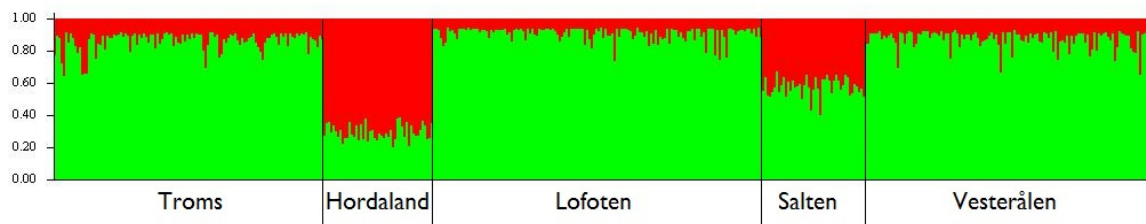

Fig. 5. STRUCTURE bar plot, where each individual is represented as a vertical line and the colors represent individuals estimated membership to the populations.

Testing for outlier loci

The neutrality test detected five outlier loci among the seven microsatellites. GMO34 and GMO132 are outlier loci under positive selection with high F_{ST}–value and low heterozygosity, GMO03, GMO35 and GMO19 are outlier loci under balancing selection with low F_{ST}–values and high heterozygosity, and the remaining two loci, GMO36 and PGMO58 are seemingly neutral markers (Fig. 6).
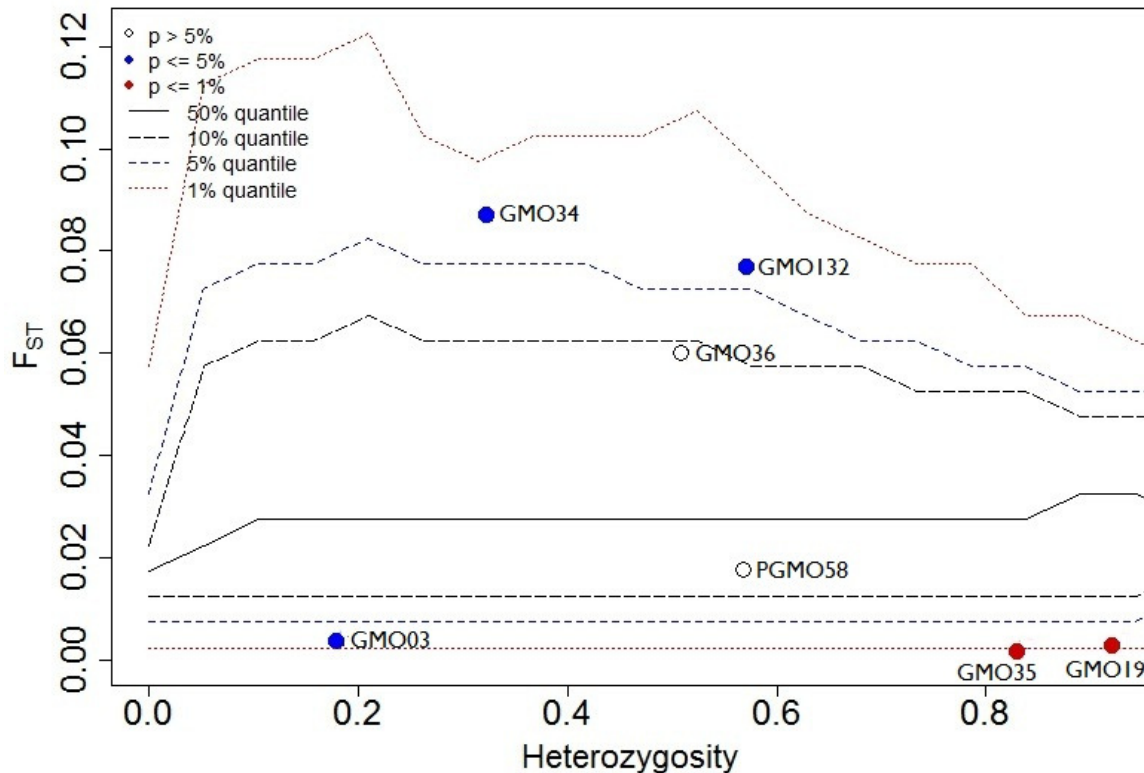
Fig. 6. The outlier test shows two loci (GMO34 and GMO132) to be outliers under positive selection ($P < 0.05$). Two loci (GMO36 and PGMO58) are neutral to selection and three loci are under stabilizing selection GMO03 ($P < 0.05$), GMO35 and GMO19 ($P < 0.01$).

Allelic frequencies

The allelic frequencies for the seven microsatellites vary. The markers which appear to be under positive selection, GMO34 and GMO132, have a relatively high allelic richness, respectively 6.118 and 12.880 (Table 5). The neutral markers have a lower allelic frequencies, both GMO36 and PGMO58 have only 8 alleles each (Fig. 9 and 10), but an allelic richness of 3.544 and 5.978 (Table 5). Three markers, GMO03, GMO35 and GMO19, have relatively high allelic richness of 4.322, 7.993 and 18.298 respectively (Table 5) and allelic frequencies of 10, 16 and 26 (Fig. 11, 12 and 13). The allelic frequencies do not show a clear difference in distribution compared to the sample locations. Except for the GMO132 locus (Fig. 8), that shows that Hordaland and Salten samples seem to have a larger allele sizes than the other sample locations.

Allele frequencies for markers under positive selection

GMO34 and GMO132 have high F_{ST}–value and low heterozygosity (Fig. 6). The GMO34 loci show no differentiation between the putative populations, but there is a difference at the

GMO132 locus. The Hordaland samples, assumed to be NCC have a wider range in allele size than especially the Lofoten samples which are assumed to be NEAC. The pattern of the allele sizes from the Salten samples assumed as NCC is similar to the pattern from the Hordaland samples (Fig. 8)
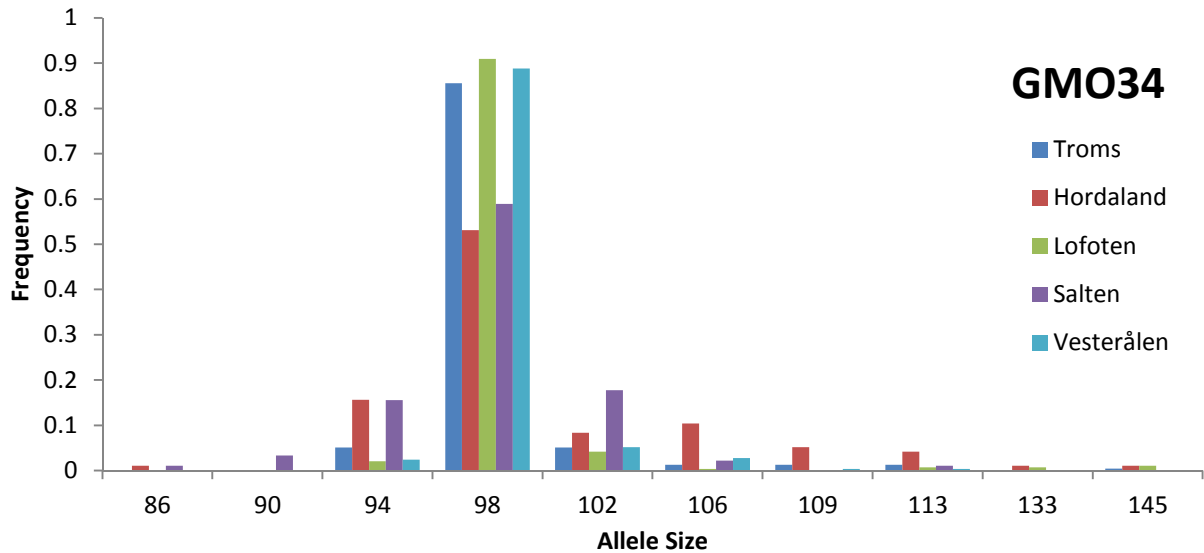

Fig. 7. The allelic frequencies of the GMO34 loci, with a total of 10 alleles and an allelic richness of 6.118.
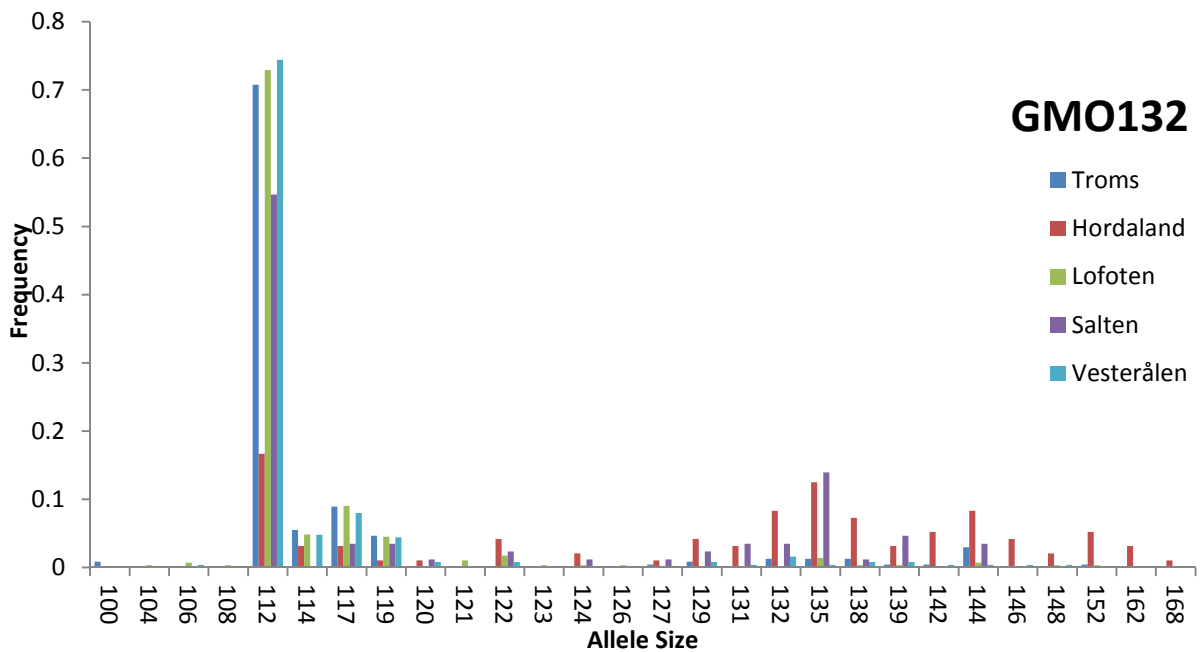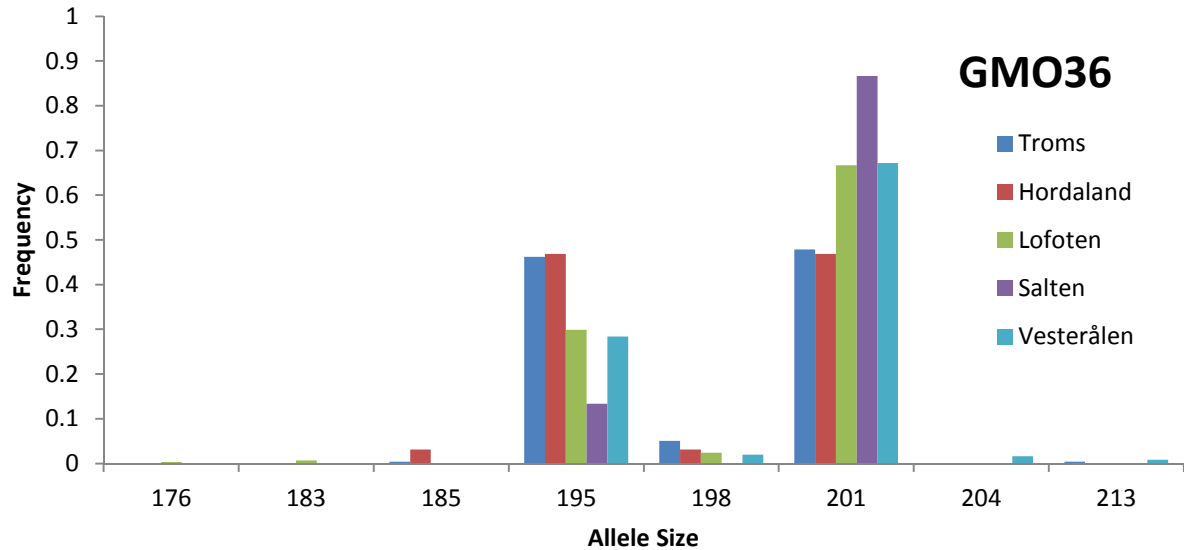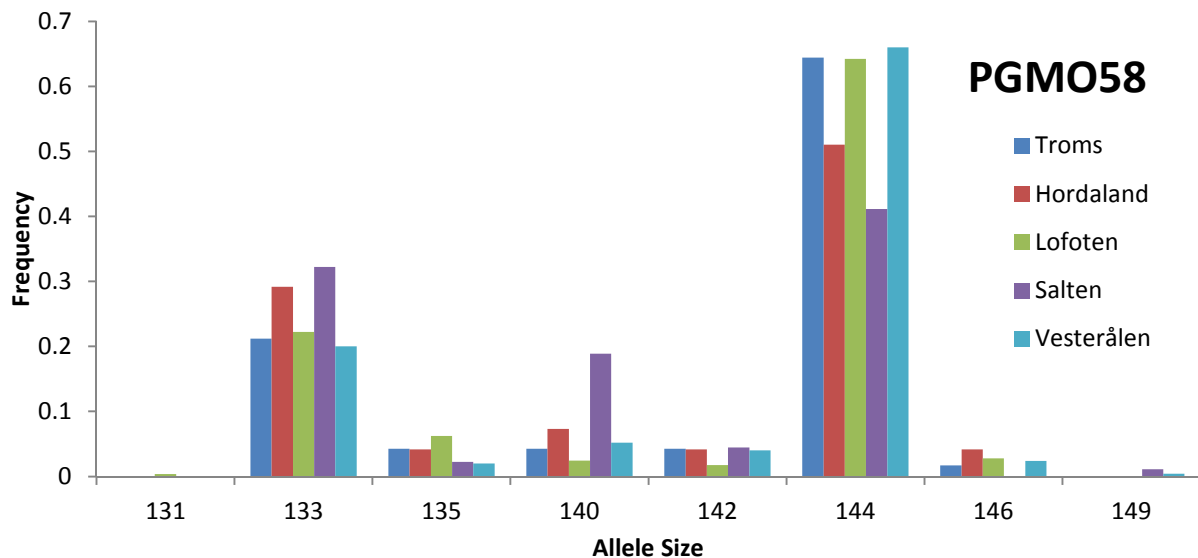

Fig. 8. The allelic frequencies of the GMO132 loci, with total of 28 alleles and an allelic richness of 12.880.

Allelic frequencies for markers neutral to selection

The two loci neutral in the outlier test GMO36 and PGMO58 (Fig. 6) both have 8 different alleles each (Fig. 9 and 10), but vary in allelic richness (Table 5). GMO36 has an allelic richness of 3.544 and PGMO58 an allelic richness of 5.978. The allelic frequencies for these two loci show no differences between the two putative NEAC and NCC samples.
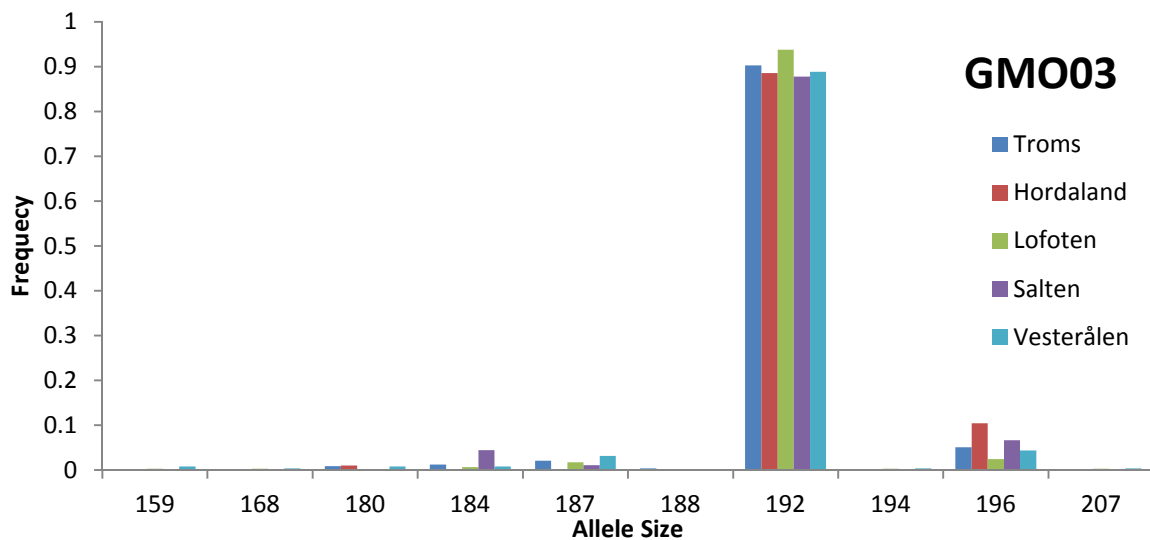

Fig. 9. The allele frequencies of the GMO36 loci, with a total of 8 alleles and an allelic richness of 3.544.

Fig. 10. The allele frequencies of the PGMO58 loci, with a total of 8 alleles and an allelic richness of 5.978.

Allelic frequencies for markers under balancing selection

The allele frequencies for the three loci under balancing selection vary. GMO03 has 10 alleles (Fig. 11) and an allelic richness of 4.332 (Table 5), while GMO35 has 16 alleles (Fig. 12) and an allelic richness of 7.993 (Table 5), and finally GMO19 has 26 alleles (Fig. 13) and an allelic richness of 18.298 (Table 5). The allelic- frequency and richness vary a lot between the three loci, and none of them show any discrimination between the putative NEAC and NCC samples.

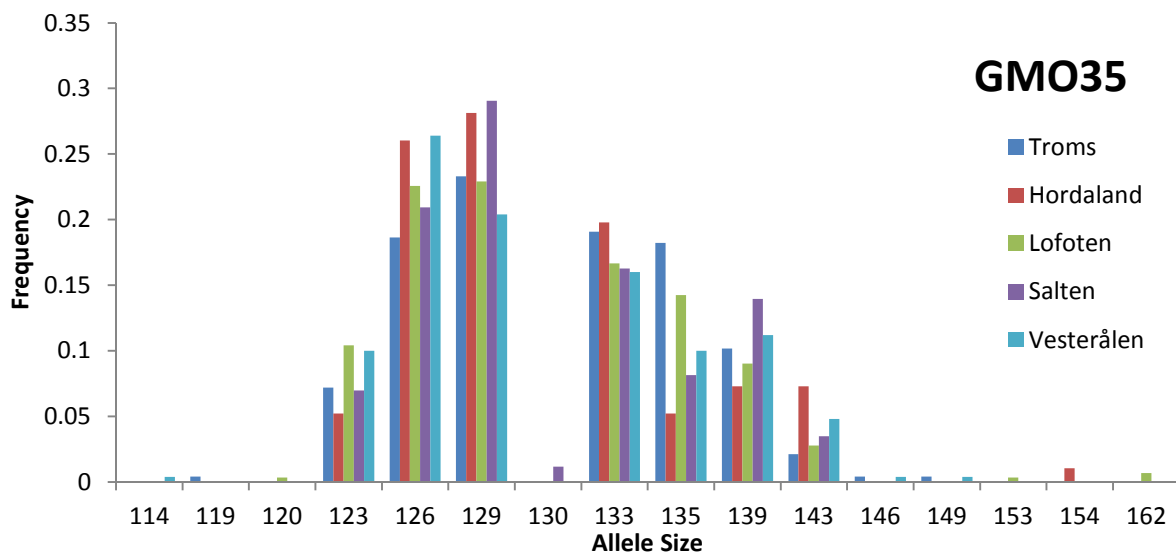**Fig. 11.** The allele frequencies of the GMO03 loci, with a total of 10 alleles and an allelic richness of 4.332.

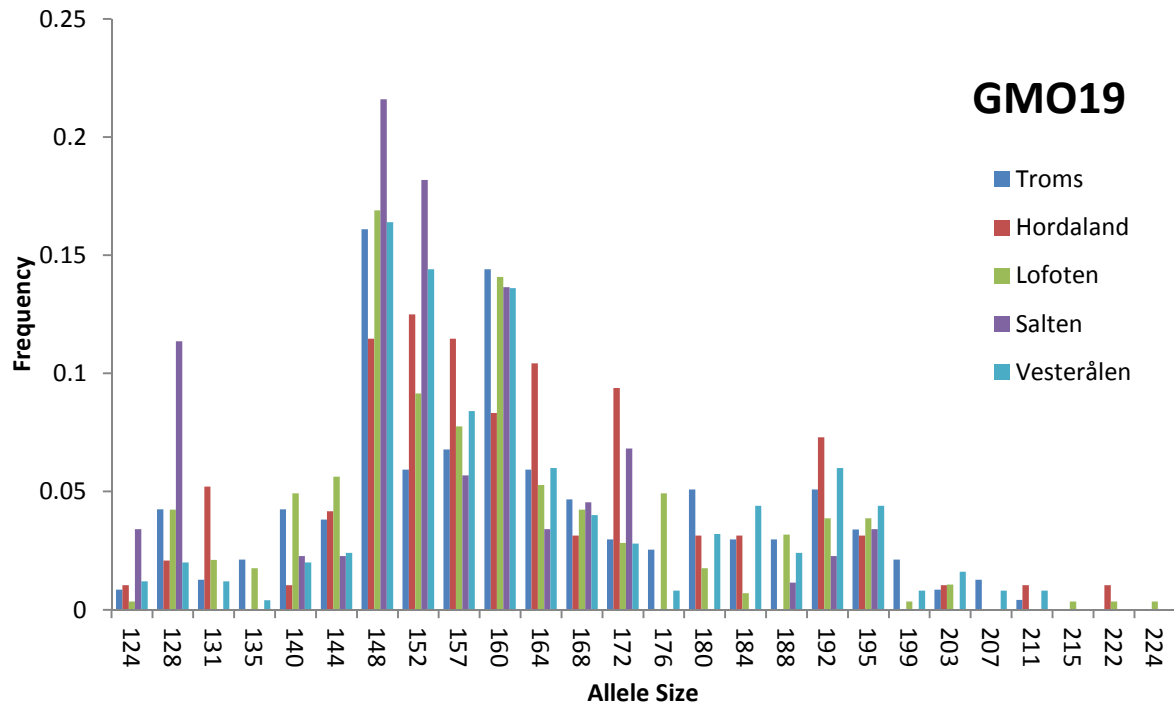**Fig. 12.** The allele frequencies of the GMO35 loci, with a total of 16 alleles and an allelic richness of 7.993.

Fig. 13. The allele frequencies of the GMO19 loci, with a total of 26 alleles and an allelic richness of 18.298.

Discussion

The most interesting finding in this study is the detection of deviation from HWE within the larvae sample from Vesterålen. Deviation was detected at both neutral and non-neutral loci, and also at the multilocus F_{IS} estimations (Table 5). The latter deviation from HWE suggests that there might be a mixture of both NEAC and NCC in the sample. This is plausible due to the sampling location being relatively close to the shore, north of the common spawning area of NEAC were the adult putative NEAC samples were caught (Table 2). This area is influenced by the water masses from the south due to the northward Norwegian coastal current, and from the east of the Lofoten and Vesterålen islands (Vestfjorden) through straits (Mork 1981). Areas known to be spawning grounds of NCC (Anonymous 2012c). The larvae samples might be very informative because these were collected at their yolk sac/pre-flexion stage (Munk et al. 2005). From an egg is fertilized it takes approximately 20 days until it hatches, depending on the temperature (Anonymous 2012f). They reach the stages of the collected samples between 4 and 30 days after hatching (Auditore et al. 1994). Due to their

young age the selective forces which differentiate NEAC and NCC might not have affected them fully yet.

Potential sampling flaws

There are several potential flaws in this study. To avoid intralocus sampling errors, which may cause an upward bias in the F_{ST} estimates, a relatively large number of individuals were sampled. A typical sample size for marine species is 25 individuals from each population (Waples 1998). For the adult and juvenile individuals in this study each sample site was represented with 45 - 48 individuals. In total 144 putative NEAC and 211 putative NCC, and finally 125 cod larva (Table 2). To reduce the possibility of caching a mixture of the two putative populations the NCC was at two locations (Hordaland and Salten) caught south of the main NEAC spawning grounds (Brander 2005). The NCC from Troms was caught in a fjord north of this area (Table 2), an area situated along the NEAC migratory route (Fig. 1) and influenced by the northward Norwegian costal current (Mork 1981). This is an area where NEAC and NCC might intermingle (Westgaard and Fevolden 2007). However, none of the putative NCC samples deviate from HWE at the neutral only multilocus F_{IS} estimates, suggesting these samples might be homogenous for one population (Table 5). The Lofoten sample, assumed to belong to NEAC were caught at well known spawning grounds of the NEAC (Brander 2005), off the west coast off Lofoten in a year with record fisheries on NEAC and low abundance of NCC (Anonymous 2012c). Presumably the NEAC dominated in abundance and based on the location it is likely that this sample is NEAC. In this sample location there is no deviation from HWE at neither of the multilocus F_{IS} estimations (Table 5). The assumptions of population membership of the samples are supported by the FAC (Fig. 4) and the STRUCTURE bar plot (Fig. 5), which both clusters individuals and estimates the memberships to a population. These shows the Hordaland samples to be most divergent from the Lofoten samples, while the remaining samples cluster between these. However these population structuring tools needs to be used with caution and they are recommended to be used only as a guide (Pritchard et al. 2010). The adult samples were collected by commercial fishermen, and with a law decided minimum length of caught cod of 44 cm (Anonymous 2012h), it is likely that they all were adults.

Potential statistical errors

Another plausible flaw in this study is the possibility of type I error. In Table 5, a relatively large amount of statistical tests is preformed with a P-value of 0.05. With this P-value there is a 1 in 20 chance that the variation is due to chance. Hence, rejecting the null hypothesis when

it in fact is true (Sall et al. 2005). To reduce the influence of this possible error, the multilocus F_{IS} estimates (Table 5) is what the main discussion of the results are focused on. For high gene flow species like cod the use of several genetic markers is important in estimating precise F-statistics estimates. Because the estimates vary among loci our ability to detect more precise estimates is enhanced by using several independent loci. This interlocus error does not lead to bias, but limits the precision of the estimates (Waples 1998). In this study a total of seven microsatellites were used (Table 4).

Challenges with microsatellites

Although microsatellites are preferable to many other molecular markers, they also have some drawbacks. The theoretical model of mutation mechanism for microsatellites is the stepwise mutation model (SMM) which is slippage and increasing or decreasing of repeat units (Beebee and Rowe 2008). Non-stepwise mutation processes may also occur, such as point mutation or recombination, but the effects are usually low as the SMM is the dominant force creating new alleles. Another disadvantage of microsatellites is stuttering due to slippage during PCR, which complicates the scoring of alleles (Selkoe and Toonen 2006). However, the most common problem with microsatellites is the occurrence of null-alleles where alleles fail to amplify during PCR due to mutations in the microsatellites flanking region (Beebee and Rowe 2008). These disadvantages with microsatellites may complicate the data analysis and limit their utility (Selkoe and Toonen 2006). Micro-checker is a software commonly used to detect null alleles and stuttering (Van Oosterhout et al. 2004). In this study this was unfortunately not done. However, the allele frequencies (Table 5 and Fig. 7-13) of the microsatellites in this study were coherent with other studies using the same microsatellites (Brooker et al. 1994; Karlsson and Mork 2005; Jorde et al. 2007; Westgaard and Fevolden 2007). So was also the allelic richness (Table 5), which compensates for the sample sizes using rarefaction (Skarstein et al. 2007; Wennevik et al. 2008). The coherence found here might suggests that the data in this study might not have been influenced significantly by the problems microsatellites are prone to.

Outlier test

The outlier test detected five of the seven microsatellites used in this study to be under selection (Fig. 6). Two loci were under positive selection, GMO34 and GMO132, as expected as they previously have been described as such. GMO132 has been shown to be under selective pressure and therefore non-neutral, while GMO34 has shown linkage disequilibrium with the non-neutral *Pan*I gene (Karlsson and Mork 2005; Westgaard and Fevolden 2007).

The positive selection these loci are influenced by may reduce genetic diversity through background selection which eliminates deleterious mutations by purifying selection, or by selective sweeps which entails positive selection in favor of an adaptive mutation (Hanfstingl et al. 1994). Three microsatellites were under balancing selection, GMO03, GMO35 and GMO19 (Fig. 6). Balancing selection is a way to avoid natural selection culling all unfavorable genes, this occurs when natural selection maintains stable allelic frequencies in populations by heterozygote advantage and frequency-dependent selection, called balanced polymorphism (Campbell and Reece 2005). No other work on this topic was found that shows microsatellites under balancing selection. Two microsatellites were shown to be neutral to natural selection, GMO36 and PGMO58 (Fig. 6). These loci do not evolve as a response to natural selection, but by genetic drift and migration (Luikart et al. 2003). In population genetics it is preferred to use neutral loci, as differences found at these loci could indicate population divergence (Beebee and Rowe 2008). The GMO132 locus which is non-neutral might not be suitable for population genetics, but it did suggest differences between the two putative populations based on the allele size, and might therefore be used as a supporting tool for discriminating the populations (Westgaard and Fevolden 2007).

F_{ST} and HWE

The F_{ST} estimates showing the genetic variation between sampling locations shows all the sampling sites to be significantly different from each other, except between the Lofoten and Vesterålen samples when including all seven loci. After removing all the non-neutral loci the Hordaland and Troms also did not show any significant difference from each other (Table 6). This is surprising especially due the geographical distance between the Troms and Hordaland samples, and the possibility of intermingling of NEAC and NCC in the Troms area (Westgaard and Fevolden 2007). But since to the sampling in Troms were done outside the spawning season of NEAC it is possible that the Troms sample is a homogenous collection of NCC. However, a caution needs to be addressed. The use of only two loci in the estimation could limits the precision of the estimates (Waples 1998). The Vesterålen and Lofoten samples show no population structuring between them at neither of the pair-wise F_{ST} estimations (Table 6). This is surprising compared to the F_{IS} estimates (Table 5), were the Vesterålen samples show deviation from HWE suggesting the sample might be influenced by the Wahlund effect, and the Lofoten sample apparently is a homogenous sample.

At the multilocus F_{IS} estimates only the larvae sample from Vesterålen deviate from the expectations of HWE after removing all non-neutral loci. When assessing single locus,

significant deviation from HWE was found at both neutral and non-neutral loci. The neutral PGMO58 locus deviates at both the Lofoten and the larvae sample from Vesterålen, while the other neutral loci, GMO36, deviate only in the Vesterålen sample. The frequency of deviation at non-neutral loci is higher than at the neutral ones (Table 5). This is not surprising as the non-neutral loci are affected by the selective forces at the contrasting environments of the Barents Sea and the coastal areas. These finding might be interpreted to support the "historical isolation hypothesis". This hypothesis assumes no interbreeding between the putative populations despite the intermingling at spawning grounds and the differences between NEAC and NCC might be a result of genetic drift and founder effect caused by a historical period of allopatry (Nordeide et al. 2011).

Concluding remarks

In conclusion, the present study which analyzed 144 putative North East Atlantic cod (NEAC), 211 putative Norwegian costal cod (NCC) and 125 cod larva across seven microsatellites might support the "historical isolation hypothesis" presented by Nordeide et al (2011). The high amount of significant differentiation at the F_{ST} estimations may suggest that there is an even more complex population structure in *Gadus morhua* that what we currently are aware of.

The results of this study indicate that more neutral markers for cod should be developed in order to better understand the complexity of the population structuring in Atlantic cod. Sampling of the two putative should be done in a manner that reduces the possibility of mixed caches of NEAC and NCC. NEAC should be sampled in the Barents Sea, and the NCC along the coast well outside the spawning season of the NEAC. Larvae samples should be collected from areas which are presumably less influenced by NEAC in addition to samples were NEAC larvae is traditionally found. Nordeide et al. (2011) suggests the use of next-generation sequencing techniques on more neutral loci and controlled experiments to compare fitness components for future studies of the population structure in *Gadus morhua*.

Acknowledgements

References

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs."   Retrieved May, 2012.

Anonymous. (2010d). "Book III - The Barents Sea and the Norwegian Sea."   Retrieved August, 2012, from http://www.ices.dk/indexfla.asp.

Anonymous. (2012a). "10.05 Fiske og fiskeoppdrett, Statistisk årbok 2011."   Retrieved April, 2012, from http://www.ssb.no/emner/10/05/aarbok.html.

Anonymous. (2012b). "Fishery statistics."   Retrieved March, 2012, from http://www.fao.org/fishery/statistics/collections/en.

Anonymous. (2012c). "Torsk."   Retrieved August, 2012, from http://www.imr.no/temasider/fisk/torsk/nb-no.

Anonymous. (2012e). "Kvoteråd."   Retrieved August, 2012, from http://www.imr.no/radgivning/kvoterad/kvoterad_2012/ices-rad_pa_arktiske_bestander/nb-no.

Anonymous. (2012f). "Når skreien gyt."   Retrieved November, 2012, from http://www.imr.no/temasider/fisk/torsk/nordaustarktisk_torsk_skrei/nar_skreien_gyt/nb-no.

Anonymous. (2012g). "Mareano - Samler kunnskap om havet."   Retrieved August, 2012, from http://www.mareano.no/.

Anonymous. (2012h). "Minimum sizes for salt water fish."   Retrieved November, 2012, from http://www.fiskeridir.no/english/recreational-fishing/minimum-sizes.

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

Auditore, P. J., G. Lough, et al. (1994). "A review of the Comparative Development of Atlantic Cod (Gadus morhua L.) and Haddock (Melanogrammus aeglefinus L.) Based on an Illustrated Series of Larvae and Juveniles from Georges Bank." NAFO Scientific Council Studies 20: 7-18.

Beebee, T. and G. Rowe (2008). An Introduction to Molecular Ecology, Oxford University Press.

Belkhir, K., P. Borsa, et al. (1996-2004). "GENETIX 4.05, Logiciel Sous Windows TM Pour la Génétique des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier.".

Belkhir K, B. P., Chikhi L, Raufaste N, et al. (1996-2004). "GENETIX 4.05, Logiciel Sous Windows TM Pour la Génétique des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier.".

Berg, E. and O. Albert (2003). "Cod in fjords and coastal waters of North Norway: distribution and variation in length and maturity at age." ICES Journal of Marine Science 60: 787-797.

Brander, K. (2005). Spawning and life history information for North Atlantic cod stocks. ICES Cooperative Research Report. No. 274: 152 pp.

Brooker, A. L., D. Cook, et al. (1994). "Organization of Microsatellites Differ between Mammals and Cold-water Teleost Fishes." Canadian Journal of Fisheries and Aquatic Sciences 51: 1956-1966.

Campbell, N. A. and J. B. Reece (2005). Biology - Seventh Edition, Pearson.

Cantatore, P. and C. Saccone (1987). "Organization, Structure, and Evolution of Mammalian Mitochondrial Genes." International Review of Cytology 108: 149-208.

Dahle, G. (1994). "Genetic studies of families of Arcto-Norwegian and Norwegian costal cod: preliminary random amplified polymorphicdeoxyribonucleic acid studies." ICES Journal of Marine Science Symposia 198: 684-687.

Dereeper, A., V. Guignon, et al. (2008). "Phylogeny.fr: robust phylogenetic analysis for the non-specialist Nucleic Acids Research." Nucleic Acids Research 36(Web Server Issue).

Excoffier, L. G. L., and S. Schneider (2005). "Arlequin ver. 3.0: An integrated software package for population genetics data analysis." Evolutionary Bioinformatics Online 1: 47-50.

Fevolden, S. E. and G. H. Pogson (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.

Froese, R. and D. Pauly. (2012). "FishBase."   Retrieved September, 2012, from http://www.fishbase.org/identification/specieslist.php?famcode=183&areacode=.

Frydenberg, O., D. Møller, et al. (1965). "Haemoglobin polymorphism in Norwegian cod populations." Hereditas 53: 257-271.

Goudet, J. (1995). "FSTAT (Version 2.9.3.2): a computer program to calculate F-statistics." Journal of Heredity 86: 485-486.

Hall, T. A. (1999). "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." 41(95-98).

Hanfstingl, U., A. Berry, et al. (1994). "Haplotypic divergence coupled with lack of diversity at the Arabidopsis thaliana alcohol dehydrogenase locus: roles for both balancing and directional selection? ." Genetics 138: 811-828.

Hylen, A. (1964). "Costal cod and skrei in the Lofoten area." Fiskeridirektoratets skrifter Serie Havundersøkelser 13: 27-42.

Jakobsdóttir, K. B., ó. D. Jörundsdóttir, et al. (2006). "Nine new polymorphic microsatellite loci for the amplification of archived otolith DNA of Atlantic cod, Gadus morhua L." Molecular Ecology Notes 6(2): 337-339.

Johansen, S. and I. Bakke (1996). "The complete mitochondrial DNA sequence of Atlantic cod (Gadus morhua): relevance to taxonomic studies among codfishes." Molecular Marine Biology and Biotechnology 5(3): 203-214.

Johansen, S., P. H. Guddal, et al. (1990). "Organization of the mirochondrial genome of Atlantic cod, *Gadus morhua*." Nucleic Acids Research 18: 411-419.

Johansen, S. D., D. H. Coucheron, et al. (2009b). "Large-scale sequence analyses of Atlantic cod." New Biotechnology 25(5): 263-271.

Johansen, T., E. Berg, et al. (2009a). 1.7 Kysttorsk – et tema med mange variasjoner. Kyst og Havbruk: 40-44.

Jorde, P. E., H. Knutsen, et al. (2007). "Spatial scale of genetic structuring in costal cod *Gadus morhua* and geographic extent of local populations." Marine Ecology Progress Series 343: 229-237.

Karlsson, S. and J. Mork (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.

Luikart, G., P. R. England, et al. (2003). "The Power and Promise of Population Genomics: From Genotyping to Genome Typing." Nature Reviews Genetics 4: 981-994.

Løken, S. and T. Pedersen (1996). "Effects of parental type and temperature on vertebrae number in juvenile cod, *gadus morhua* (L.), in Northern Norway." Sarsia(80): 293-298.

Microsoft (2007). Excel (Part of Microsoft Office Professional Edition).

Miller, Le KD, et al. (2000). "Development of tri- and tetranucleotide repeat microsatellite loci in Atlantic cod (*Gadus morhua*)." Molecular Ecology 9: 238-239.

Moen, T., B. Hayes, et al. (2008). "Identification and characterisation of novel SNP markers in Atlantic cod: Evidence for directional selection." BMC Genetics 9: 18.

Mork, J., R. Giskeødegård, et al. (1981). "LDH gene frequencies in cod samples from two locations on the Norwegian coast." Journal du Conseil - Conseil International pour l'Exploration de la Mer 39: 110-113.

Mork, M. (1981). "Circulation Phenomena and Frontal Dynamics of the Norwegian Coastal Current." Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences 302(1472): 635-647.

Munk, P., J. G. Nielsen, et al. (2005). Eggs and larvae of North Sea fishes. Frederiksberg, Biofolia.

Møller, D. (1968). "Genetic diversity in spawning cod along the Norwegian coast." Hereditas 60: 1-32.

Nordeide and Folstad (2000). "Is cod lekking or a promiscuous group spawner?" Fish and Fisheries 1(1): 90-93.

Nordeide, J. T. (1998). "Coastal cod and north-east Arctic cod - do they mingle at the spawning grounds in Lofoten?" Sarsia 83(5): 373-379.

Nordeide, J. T., S. D. Johansen, et al. (2011). "Population connectivity among migratory and stationary cod *Gadus morhua* in the Northeast Atlantic - A review og 80 years of study." Marine Ecology Progress Series 435: 269-283.

Otterlei, E., G. Nyhammer, et al. (1999). "Temperature- and size-dependent growth of larval and early juvenile Atlantic cod (*Gadus morhua*): a comparative study of Norwegian costal cod and northeast Arctic cod." Canadian Journal of Fisheries and Aquatic Sciences 56: 2099-2111.

Page, R. D. M. (1996). "TREEVIEW: An application to display phylogenetic trees on personal computers." Computer Applications in the Biosciences 12: 357-358.

Pogson, G. H. (2001). "Nucleotide Polymorphism and Natural Selection at the Pantophysin (*Pan* I) Locus in the Atlantic Cod, *Gadus morhua* (L.)." Genetics 157: 317-330.

Pogson, G. H., K. A. Mesa, et al. (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.

Pritchard, J. K., X. Wena, et al. (2010). "Documentation for structure software: Version 2.3." http://pritch.bsd.uchicago.edu/structure.html.

Rollefsen, G. (1933). "The Otoliths of the Cod - preliminary report." Fiskeridirektoratets skrifter Serie Havundersøkelser 4: 1-18.

Sall, J., L. Creighton, et al. (2005). JMP Start Statistics - Third Edition, Curt Hinrichs.

Selkoe, K. A. and R. J. Toonen (2006). "Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers." Ecology Letters 9: 615-629.

Skarstein, T. H., J. I. Westgaard, et al. (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: 271-290.

Tamura, K., D. Peterson, et al. (2011). " MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. ." Molecular Biology and Evolution (submitted).

Van Oosterhout, C., W. F. Hutchinson, et al. (2004). "MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data." Molecular Ecology Notes 4: 535-538.

Villesen, P. (2007). "FaBox: an online toolbox for fasta sequences." Molecular Ecology Notes 7(6): 965–968.

Waples, R. S. (1998). "Separating the Wheat From the Chaff: Patterns of Genetic Differentiation in High Gene Flow Species." Journal of Heredity 89: 438-445.

Wennevik, V., K. Jørstad, et al. (2008). "Mixed stock analysis and the power of different classes of molecular markers in discriminating coastal and oceanic Atlantic cod (<i>Gadus morhua</i> L.) on the Lofoten spawning grounds, Northern Norway." Hydrobiologia 606(1): 7-25.

Westgaard, J.-I. and S. E. Fevolden (2007). "Atlantic cod (*Gadus morhua* L.) in inner and outer costal zones of northern Norway display divergent genetic signature at non-neutral loci." Fisheries Research 85: 306-315.