

RESEARCH ARTICLE

Open Access



# Analysis of coastal cod (*Gadus morhua* L.) sampled on spawning sites reveals a genetic gradient throughout Norway's coastline

Geir Dahle<sup>1,3\*</sup> , María Quintela<sup>1</sup>, Torild Johansen<sup>2</sup>, Jon-Ivar Westgaard<sup>2</sup>, François Besnier<sup>1</sup>, Asgeir Aglen<sup>1</sup>, Knut E. Jørstad<sup>1</sup> and Kevin A. Glover<sup>1,3</sup>

## Abstract

**Background:** Atlantic cod (*Gadus morhua* L.) has formed the basis of many economically significant fisheries in the North Atlantic, and is one of the best studied marine fishes, but a legacy of overexploitation has depleted populations and collapsed fisheries in several regions. Previous studies have identified considerable population genetic structure for Atlantic cod. However, within Norway, which is the country with the largest remaining catch in the Atlantic, the population genetic structure of coastal cod (NCC) along the entire coastline has not yet been investigated. We sampled > 4000 cod from 55 spawning sites. All fish were genotyped with 6 microsatellite markers and Pan I (Dataset 1). A sub-set of the samples (1295 fish from 17 locations) were also genotyped with an additional 9 microsatellites (Dataset 2). Otoliths were read in order to exclude North East Arctic Cod (NEAC) from the analyses, as and where appropriate.

**Results:** We found no difference in genetic diversity, measured as number of alleles, allelic richness, heterozygosity nor effective population sizes, in the north-south gradient. In both data sets, weak but significant population genetic structure was revealed (Dataset 1: global  $F_{ST} = 0.008$ ,  $P < 0.0001$ . Dataset 2: global  $F_{ST} = 0.004$ ,  $P < 0.0001$ ). While no clear genetic groups were identified, genetic differentiation increased among geographically-distinct samples. Although the locus Gmo132 was identified as a candidate for positive selection, possibly through linkage with a genomic region under selection, overall trends remained when this locus was excluded from the analyses. The most common allele in loci Gmo132 and Gmo34 showed a marked frequency change in the north-south gradient, increasing towards the frequency observed in NEAC in the north.

**Conclusion:** We conclude that Norwegian coastal cod displays significant population genetic structure throughout its entire range, that follows a trend of isolation by distance. Furthermore, we suggest that a gradient of genetic introgression between NEAC and NCC contributes to the observed population genetic structure. The current management regime for coastal cod in Norway, dividing it into two stocks at 62°N, represents a simplification of the level of genetic connectivity among coastal cod in Norway, and needs revision.

**Keywords:** Fishery, Otolith, Population, Fish, Outlier, NCC

\* Correspondence: [geir.dahle@imr.no](mailto:geir.dahle@imr.no)

<sup>1</sup>Institute of Marine Research (IMR), Postbox 1870, N-5817 Bergen, Norway

<sup>3</sup>Department of Biology, University of Bergen, Bergen, Norway

Full list of author information is available at the end of the article



## Background

The Atlantic cod (*Gadus morhua* L.) is a demersal fish found in the northern waters across the Atlantic. Due to its large size and historical high abundance, cod has formed the basis of some of the most commercially important fisheries in the Atlantic for centuries. However, a legacy of over-exploitation, potentially exacerbated by climate-driven changes in distribution, has left several cod populations depleted [1]. In turn, this has also resulted in the decline of many of the commercially significant fisheries [2, 3]. The best-known example of this is the total collapse of the northern cod fishery off Newfoundland which fell from ~ 800,000 tons around 1970 to < 1000 tons by 1992 [4].

Cod is one of the best studied marine fishes, and as for many marine species where it has been investigated, statistically significant spatial population genetic structure has been observed (reviewed by [5–7]). The first studies of population genetic structure of cod were conducted in the 1960's using haemoglobin (e.g. [8–11]) and transferrin [12]. Shortly after, population genetic studies were performed using allozymes [13–15], mtDNA [16, 17], Pantophysin (Pan I) [18–21] and microsatellites [7, 22–28]. More recently, single nucleotide polymorphisms (SNPs) [29–34] and full-genome sequencing [35, 36] have been used to investigate the evolutionary relationships among cod populations. While not all studies on cod have revealed statistically significant population genetic structure [15, 16, 37], the majority have, and collectively, they reveal a species displaying population genetic differentiation throughout its native range.

Studies have revealed genetic differences between cod sampled on the eastern and western sides of the Atlantic [7, 30, 38], and between cod sampled in different ecosystems: in the Baltic vs. North Sea [25], North Sea vs. Norwegian coast [24], Canadian coastal vs. oceanic [38, 39], coastal and oceanic in Gulf of Maine [40], and Greenland coastal vs. offshore with link to the Icelandic offshore cod [41, 42]. Also, genetic differences have been observed between cod sampled in different areas within countries, including the North Sea [23], within UK [7], Iceland [43], North America [38, 40], and Norway (North East Arctic Cod (NEAC) vs. Norwegian Coastal Cod (NCC)) [27, 31, 32, 44]. Finally, even on small spatial scales, such as among neighbouring fjords in Norway, genetic differences have been observed [24, 26]. These genetic studies therefore demonstrate that the species consists of multiple populations displaying varying levels of connectivity.

In addition to spatial genetic structure, genetic differences have been observed among cod “ecotypes” displaying stationary and long-distance migratory behaviours. For example, large genetic differences have been reported between NCC and NEAC, which display stationary and migratory behaviours respectively [29, 31, 32]. The same has also been observed between migratory North Sea cod

(NSC) and coastal cod in southern Norway [45], between migratory and stationary in Iceland [31, 46], and Canada [47]. Furthermore, the population genomic approaches [48] utilised in many of the most recent studies have also revealed large genomic inversions, notably in linkage groups 1, 2 and 7 that are assumed to be responsible for much of this strong divergence between ecotypes [29, 32]. However, genomic islands of divergence, probably also caused by genomic inversions, have also been reported between low-salinity adapted cod in the Baltic, and the North Seas [49].

Despite the considerable number of population genetic studies on cod, there are still a number of remaining questions. This is the case for Norway, which is the country with the largest remaining commercial catch of cod in the Atlantic. In Norway, coastal cod are currently divided into two management stocks, defined as NCC north of 62°, and coastal cod south of 62° (from hereon we refer to both of these components as NCC for simplicity). While the NEAC stock is currently at a high level [50, 51], NCC is depleted and a rebuilding plan has been recommended by ICES [51, 52]. Importantly, while studies of NCC sampled in neighbouring fjords have been conducted in limited geographic areas [24, 26], population genetic structure has not been investigated in detail along the entire Norwegian coastline spanning.

In 2002, a project was initiated to map the population genetic structure of NCC along the entire Norwegian coast. This included sampling cod from 55 spawning locations spanning the Russian to Swedish borders with Norway (Fig. 1). At the same time, biological data and brood-stock fish were sampled [53], some of which have formed the basis of common-garden experiments [54, 55]. Here, we present the results of the genetic analysis based upon data from microsatellite loci and the Pan I locus.

## Methods

### Sampling

In the period from 2002 to 2007, 4422 cod were collected from 55 locations along the coast of Norway (Fig. 1). Eleven of those sites were sampled more than once during the six-year period. Sampling was conducted during the spawning season (late winter and early spring) by taking samples from the catch of local fishermen. All of these fish were collected at known spawning sites. Most of the cod were sampled using gill nets, although for some individuals around the Lofoten islands, demersal trawls were also used. Biometry data such as length, weight and sex of all individuals were recorded, and the gonads were visually inspected to determine the stage of maturation. Fin clips from all the individuals were taken and stored in 96% ethanol prior to DNA extraction.

Although genetic analyses have also been used for identification of NEAC and NCC in Norwegian fisheries



**Fig. 1** Location of the sampling sites along the Norwegian coastline (Figure produced specifically for this manuscript)

[56, 57], in routine Norwegian stock assessment north of 62°, cod are assigned to NEAC and NCC by otolith category. Otoliths from all the fish sampled in this study were analysed to determine age, age at first time of spawning and the number of spawning periods, as well as to identify individuals as NEAC or NCC. Otolith type 1 and 2 is assigned NCC whereas 4 and 5 are assigned to NEAC, as described for the first time by Rollefsen [58] and modified by Berg & Albert [59]. Assignment otolith category is in strong agreement with results from genetic markers [27, 59], and particularly Pan I. At this marker, NCC show a high frequency of the Pan I AA genotype and NEAC show a high frequency of the Pan I BB genotype [26]. In the present study, all individuals with otoliths belonging to types 4 and 5, were removed from the majority of the statistical analyses in order to exclude any NEAC from the samples, which could influence estimates of NCC population genetic structure (76 such individuals were removed in total).

#### Genotyping

DNA extraction was performed using the Qiagen DNeasyH96 Blood & Tissue Kit; each of which contained two or more negative controls. All 4346 individuals were

genotyped at six microsatellite loci: Gmo2, Gmo3, Gmo34, Gmo35, Gmo132 and Tch11 [60–62], together with the Pantophysin locus Pan I [18]. In addition, a subset of 1295 individuals from 17 of the sites were genotyped for an extra set of nine microsatellites (bringing the total microsatellite loci up to 15): GmoC18, GmoC20 [63]; GmoG13, GmoG18 [64]; GmoG25, GmoG40, GmoG43, GmoG45 [64], and Tch22 [62]. Thus, the present study includes two overlapping data sets: one with all 4346 individuals genotyped for six microsatellites and Pan I (55 locations, hereon referred to as Dataset 1), and the other with a sub-set of individuals 1295 individuals genotyped for 15 microsatellites and Pan I (17 locations, hereon referred to as Dataset 2). The PCR conditions are available from the authors upon request. PCR products were analysed on an ABI3130XL sequencer (Applied Biosystems), whereas Pan I was genotyped on 2.5% MetaPhore gels. Microsatellite alleles were scored using GeneMapper v4.0 (Applied Biosystems).

#### Statistical analysis

Statistical analyses were conducted separately for microsatellites and Pan I. The total number of alleles, and allelic richness, were both calculated with MSA [65],

whereas observed ( $H_O$ ) and unbiased expected (uHe) heterozygosity, inbreeding coefficient ( $F_{IS}$ ) and deviations from the expected Hardy-Weinberg distribution, were computed with GenAlEx [66]. Possible linkage (LD) between all locus pairs per population was tested using the program GENEPOP on the web [67] with significance based on the Markov chain method with 10,000 dememorizations, 20 batches and 5000 iterations per batch. Effective population size ( $N_e$ ) per sample was computed using LDNE [68], implementing the threshold values of lowest allele frequency of 0.05 and 0.01. Where applicable, signification was corrected by multiple comparisons by sequential Bonferroni correction [69] implemented in the calculator developed by Justin Gaetano (2013).

To test if loci deviated from neutrality, outlier analyses were conducted with LOSITAN [70] under a stepwise model and the following settings: 1000000 simulations, 99.5% confidence interval, forced mean  $F_{ST}$  and with a 0.01 false discovery rate. Genetic differentiation among sampling sites was tested using the Analysis of Molecular Variance (AMOVA) as well by pairwise  $F_{ST}$ . Both analyses were implemented in ARLEQUIN v.3.5.1.2 [71] and significance was calculated after 10,000 permutations. Likewise, hierarchical AMOVA was conducted by pooling populations according to geographic areas depicted in Table 1.

Several parameters such as the number of alleles, allelic richness, allele frequency,  $H_O$ , uHe and pairwise  $F_{ST}$  were tested for trends in the geographic north-south gradient using the non-parametric Kendall measure of rank correlation [72], which measures the similarity of the orderings of the data when ranked by north-south gradient or by the value of the variable tested [73], and implemented in the R Package 'Kendall' [74].

Population genetic structure and connectivity were investigated using two approaches. First, through BARRIER 2.2 software [75], which aims to reveal genetic barriers between populations using Monmonier's [76] algorithm. The significance for this analyses was tested by bootstrapping 1000 matrices computed with Nei's DA genetic distance [77]. In addition, STRUCTURE v. 2.3.4 [78] was used to identify genetic groups under a model assuming admixture and correlated allele frequencies without population information. STRUCTURE analyses were parallelised with the program ParallelStructure [79] to speed up computation time. After multiple runs of both programs, and probably due to the nature of the genetic structure observed (see results), the software failed to find clear barriers (data not presented for BARRIER). Therefore, in order to complete some of the population genetic analyses (i.e., AMOVA as described above), we subjectively chose six geographic regions for some of the hierarchical analyses, which are not advocated as Management Units.

## Results

Dataset 1 included samples from 55 locations (66 samples when temporal samples were included), made up of 25–129 individuals each, that were genotyped for six microsatellites and Pan I. Across the six microsatellites, the total number of alleles observed per sample ranged from 59 to 85, and allelic richness ranged from 9.0–11.1 (Table 1). Allelic richness displayed a trend in the north-south gradient, increasing towards the south ( $\tau = 0.296$ ,  $P = 0.00045$ ) (Fig. 2a). In contrast, no trend in the number of alleles was observed on the north-south gradient ( $\tau = 0.153$ ,  $P = 0.074$ ) (Fig. 2a). Observed ( $H_O$ ) and unbiased heterozygosity (uHe), ranged from 0.570–0.731 and 0.610–0.758 respectively (Table 1, Additional file 1: Fig. S1a). These parameters also displayed statistically significant (albeit very weak) trends in the north-south gradient ( $\tau = 0.363$ ,  $P = 1.705 \text{ e-}05$  for  $H_O$ , and  $\tau = 0.570$ ,  $P < 2.22 \text{ e-}16$  for uHe).

Dataset 2 included a sub-set of 18 samples from 17 locations from Dataset 1, made up of 33–96 individuals each, that were genotyped with an extra suite of nine microsatellites. Across the microsatellites, a total of 158–237 alleles were observed per sample (Table 2). Neither the number of alleles (158–237), nor allelic richness (11–13.5), displayed a trend in the north-south gradient ( $\tau = 0.243$ ,  $P = 0.1721$ ; and  $\tau = 0.0543$ ,  $P = 0.7888$ , respectively) (Fig. 2b). Two of the sites (Tysfjord\_2003 and more importantly, Finnøy\_2007) showed a lower number of recorded alleles than expected, most probably due to their low sampling sizes ( $N = 42$  and  $N = 33$ , respectively). Neither  $H_O$  (0.632–0.734) nor uHe (0.669–0.769) showed any trend in the north-south gradient ( $\tau = 0.281$ ,  $P = 0.11164$  and  $\tau = 0.21$ ,  $P = 0.23997$ , respectively) (Additional file 1: Figure S1b).

In the six microsatellites common to both datasets, the total number of alleles observed per locus showed extremely similar values, despite that the number of individuals genotyped displayed a 3-fold difference between the two data sets (Table 3). In five of the six microsatellite markers used in Dataset 1, global  $F_{ST}$  per locus significantly differed from zero (Table 3). In eight of the fifteen microsatellite markers used in Dataset 2, global  $F_{ST}$  per locus was significantly different from zero. The locus Gmo132 clearly displayed a much higher global  $F_{ST}$  than any of other loci in both data sets, and was reported to be under directional selection by LOSITAN ( $P = 1.0$ ).

Global  $F_{ST}$  over all microsatellites was low but statistically significant in both datasets (Dataset 1,  $F_{ST} = 0.0075$ ,  $P < 0.0001$ , and Dataset 2,  $F_{ST} = 0.0042$ ,  $P < 0.0001$ ). Global  $F_{ST}$  did not change when removing the individuals showing the genotype Pan I BB, which is the dominating genotype in NEAC and observed in very low frequency in NCC. This meant excluding 104 and 39 individuals

**Table 1** Summary statistics for Dataset 1 (66 samples genotyped at six microsatellites): geographic region the sampling site belongs to, number of sample, sample name, total sampling size (i.e. NEAC and NCC), sampling size for NCC (i.e. individuals with otoliths 1 and 2); % of NEAC in the sample (i.e. individuals with otoliths 4 and 5); number of alleles; AR, allelic richness (based on a 25 diploid individuals); observed ( $H_o$ ) and unbiased expected heterozygosity ( $H_e$ ), inbreeding coefficient ( $F_{IS}$ ), number of deviations from Hardy-Weinberg equilibrium (HWE) and from Linkage Disequilibrium (LD) at  $\alpha = 0.05$

Region	No	Sample	N (total)	N (NCC)	% NEAC	N alleles	Ar	$H_o$	$H_e$	$F_{IS}$	No dev HWE	No dev LD	Ne	
													0.05	0.01
Finnmark	1	Magerøysundet_2002	70	69	1.4	69	9	0.570	0.610	0.038	1	1	254.6	415.9
Finnmark	2	Smørfjord_2002	63	59	6.3	71	9.5	0.602	0.671	0.069	3	1	687.1	∞
Finnmark	3	Smørfjord_2003	78	74	5.1	70	9.2	0.617	0.644	0.028	2	0	∞	494.6
Finnmark	4	Repparfjord_2003	59	58	1.7	67	9.3	0.638	0.694	0.053	0	1	318.3	∞
Troms	5	Balsfjord_2002	94	94	0.0	78	9.3	0.638	0.681	0.042	2	0	1476.5	∞
Troms	6	Balsfjord_2003	81	81	0.0	73	9.1	0.656	0.662	-0.009	0	1	149.8	1080.3
Troms	7	Tranøybotn_2004	102	102	0.0	83	9.8	0.691	0.703	0.007	2	1	∞	∞
Troms	8	Tranøybotn_2005	74	74	0.0	74	9.7	0.631	0.683	0.050	3	0	∞	1037.4
Troms	9	Gratangen_2004	53	53	0.0	67	9	0.657	0.667	0.015	0	0	269.9	285.1
Troms	10	Gratangen_2005	61	61	0.0	72	9.5	0.634	0.666	0.016	2	0	∞	1203
Troms	11	Gratangen_2006	62	62	0.0	75	9.8	0.621	0.674	0.073	0	2	101.2	347.9
Lofoten	12	Eidsfjord_2003	68	68	0.0	77	10.2	0.645	0.694	0.072	2	0	∞	∞
Lofoten	13	Eidsfjord_2004	77	75	2.6	74	9.6	0.689	0.700	0.014	0	1	852.3	∞
Lofoten	14	Valberg_2006	56	43	23.2	67	9.7	0.624	0.680	0.051	1	0	∞	∞
Lofoten	15	Henningsværstraumen_2006	39	39	0.0	70	10.3	0.688	0.732	0.036	0	1	589	∞
Lofoten	16	Austnesfjord_2002	57	45	21.1	65	9.3	0.685	0.706	0.040	0	1	214.6	∞
Lofoten	17	Bresja_2006	40	40	0.0	64	9.7	0.675	0.702	0.017	0	1	260.7	4814.9
Lofoten	18	Kanstadfjord_2006	78	75	3.8	75	9.7	0.636	0.695	0.083	1	1	325.3	∞
Lofoten	19	Fiskerfjord_Tjeldsund_2006	47	47	0.0	68	9.6	0.660	0.688	0.022	1	0	126.1	131.5
Lofoten	20	Stefjord_2002	52	52	0.0	72	9.9	0.638	0.706	0.090	1	3	430.8	426.7
Lofoten	21	Tysfjord_2003	36	36	0.0	62	9.5	0.671	0.707	0.028	0	1	287.7	∞
Lofoten	22	Baldkjosen_Nordfolda_2004	57	56	1.8	73	10.1	0.708	0.712	-0.017	1	0	∞	∞
Helgeland	23	Hopen_2005	46	44	4.3	71	10.2	0.633	0.689	0.095	1	0	248.5	∞
Helgeland	24	Saltenfjord_Hopen_2004	63	63	0.0	73	9.8	0.616	0.690	0.113	1	1	∞	∞
Helgeland	25	Skjerstadfjord_Valnesfjord_2005	89	88	1.1	79	9.8	0.661	0.711	0.068	1	1	508.7	∞
Helgeland	26	Skjerstadfjord_Valnesfjord_2006	47	47	0.0	74	10.2	0.691	0.711	0.021	1	1	464.5	∞
Helgeland	27	Glomfjord_Ørnes_2004	46	46	0.0	68	9.8	0.623	0.683	0.091	2	0	125.1	154.6
Helgeland	28	LilleSjona_2004	78	78	0.0	76	9.8	0.673	0.685	0.003	1	1	∞	3959.9
Helgeland	29	LilleSjona_2005	83	83	0.0	76	9.7	0.596	0.691	0.127	2	0	∞	∞
Helgeland	30	Vega_2002	38	38	0.0	65	9.7	0.645	0.710	0.103	1	0	∞	∞
Helgeland	31	Vega_2003	83	81	2.4	75	9.6	0.652	0.684	0.053	2	1	1406.6	2103.4
Helgeland	32	Langesundet_2004	64	64	0.0	77	10.4	0.661	0.708	0.061	0	0	∞	∞
Helgeland	33	Vikna_2003	79	79	0.0	79	9.9	0.700	0.711	0.033	1	2	34,313.7	7039.7
Helgeland	34	Verrabotn_2005	72	72	0.0	75	10	0.620	0.697	0.094	1	2	1373.1	2636.2
Helgeland	35	Verrabotn_2006	60	60	0.0	69	9.3	0.639	0.695	0.071	2	3	∞	2992.9
Møre	36	Frøya_2004	63	63	0.0	76	10.3	0.720	0.736	0.027	0	0	∞	∞
Møre	37	Hitra_Laksovik_2004	85	85	0.0	83	10.3	0.694	0.724	0.032	1	0	693	∞
Møre	38	Smøla_2003	79	78	1.3	78	10.1	0.722	0.730	0.015	1	1	∞	∞
Møre	39	Vinjefjord_2004	99	99	0.0	85	10.1	0.694	0.723	0.032	0	1	727.1	4552.7
Møre	40	Batnesfjord_2004	56	56	0.0	76	10.3	0.670	0.696	0.026	1	1	∞	5604.4
Møre	41	Buagrunn_2004	48	35	27.1	61	9.4	0.700	0.718	-0.004	1	0	∞	∞

**Table 1** Summary statistics for Dataset 1 (66 samples genotyped at six microsatellites): geographic region the sampling site belongs to, number of sample, sample name, total sampling size (i.e. NEAC and NCC), sampling size for NCC (i.e. individuals with otoliths 1 and 2); % of NEAC in the sample (i.e. individuals with otoliths 4 and 5); number of alleles; AR, allelic richness (based on a 25 diploid individuals); observed ( $H_o$ ) and unbiased expected heterozygosity ( $H_e$ ), inbreeding coefficient ( $F_{IS}$ ), number of deviations from Hardy-Weinberg equilibrium (HWE) and from Linkage Disequilibrium (LD) at  $\alpha = 0.05$  (Continued)

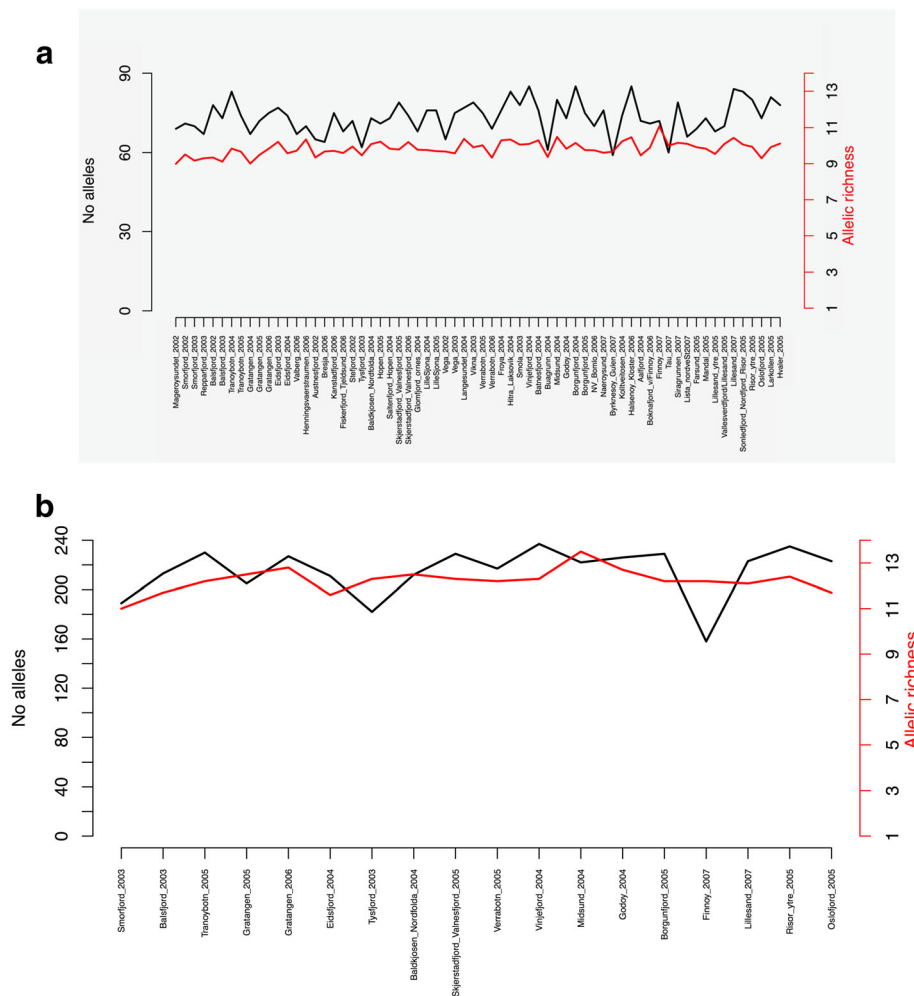
Region	No	Sample	N (total)	N (NCC)	% NEAC	N alleles	Ar	$H_o$	$uH_e$	$F_{IS}$	No dev HWE	No dev LD	Ne	
													0.05	0.01
Møre	42	Midsund_2004	69	69	0.0	80	10.5	0.681	0.723	0.055	0	1	∞	548.8
Møre	43	Godøy_2004	76	65	14.5	73	9.8	0.679	0.705	0.018	3	1	∞	366.1
Møre	44	Borgunfjord_2004	132	129	2.3	85	10.2	0.703	0.735	0.047	1	1	∞	6957.6
Møre	45	Borgunfjord_2005	87	85	2.3	75	9.8	0.665	0.719	0.056	1	0	∞	∞
Vestlandet	46	NV_Bømlo_2006	57	57	0.0	70	9.7	0.722	0.736	-0.006	1	1	186.2	472.6
Vestlandet	47	Nærøysund_2007	92	92	0.0	76	9.6	0.699	0.727	0.028	1	1	∞	659
Vestlandet	48	Byrknesøy-Gulen_2007	27	27	0.0	59	9.7	0.660	0.734	0.119	3	3	45.6	56.1
Vestlandet	49	Kolltveitosen_2004	55	55	0.0	74	10.2	0.721	0.732	-0.001	1	1	∞	∞
Vestlandet	50	Halsenøy_Kloster_2006	93	93	0.0	85	10.5	0.720	0.738	0.009	1	0	378.6	∞
Vestlandet	51	Ålfjord_2004	89	89	0.0	72	9.5	0.700	0.715	0.000	1	1	232.8	258.3
Vestlandet	52	Boknafjord_v/Finnøy_2006	59	59	0.0	71	9.9	0.701	0.723	0.021	1	1	1343.1	1083.3
Vestlandet	53	Finnøy_2007	35	35	0.0	72	11.1	0.710	0.745	0.031	1	1	∞	∞
Vestlandet	54	Tau_2007	25	25	0.0	60	10	0.667	0.709	0.062	1	0	120	269
Sørlandet	55	Siragrunnen_2007	80	80	0.0	79	10.2	0.673	0.723	0.055	1	0	1368.4	∞
Sørlandet	56	Lista_nordvest_2007	36	36	0.0	66	10.1	0.731	0.758	0.053	1	0	255.5	129.5
Sørlandet	57	Farsund_2005	46	46	0.0	69	9.9	0.630	0.710	0.106	1	1	∞	∞
Sørlandet	58	Mandal_2005	67	67	0.0	73	9.8	0.657	0.724	0.061	2	4	∞	∞
Sørlandet	59	Lillesand_ytre_2005	62	62	0.0	68	9.5	0.672	0.712	0.041	0	0	∞	1507.4
Sørlandet	60	Vallesverdfjord/Lillesand_2005	45	45	0.0	70	10.1	0.670	0.727	0.049	1	2	124.1	∞
Sørlandet	61	Lillesand_2007	77	77	0.0	84	10.4	0.716	0.724	-0.005	2	0	961.6	∞
Sørlandet	62	Søndeledfjorden_Nordfjord_Risør_2005	88	88	0.0	83	10.1	0.693	0.727	0.044	2	0	∞	∞
Sørlandet	63	Risør_ytre_2005	78	78	0.0	80	9.9	0.667	0.728	0.070	2	1	∞	∞
Sørlandet	64	Oslofjord_2005	85	85	0.0	73	9.3	0.700	0.713	-0.003	0	2	626.5	3948.3
Sørlandet	65	Larkollen_2005	95	95	0.0	81	9.9	0.688	0.725	0.030	2	0	∞	∞
Sørlandet	66	Hvaler_2005	85	85	0.0	78	10.1	0.686	0.731	0.039	1	1	1229.2	∞

Effective population size (Ne) was calculated using two values of lowest allele frequency (0.05 and 0.01). Negative values of Ne have been replaced into the infinite symbol (∞). Sites have been ordered from north to south and information about the geographic region where they are placed is provided. The total number of deviations from HWE dropped from 73 to 25 after Bonferroni correction, and from 54 to 7 at LD, respectively

from Datasets 1 and 2 respectively (data not shown). After excluding the microsatellite identified to be under positive selection (Gmo132), global  $F_{ST}$  decreased, but it was still statistically significant in both Dataset 1 ( $F_{ST} = 0.00227$ ,  $P < 0.0001$ ), and Dataset 2 ( $F_{ST} = 0.00200$ ,  $P < 0.0001$ ).

The genetic matrix for Dataset 1 revealed that 2.7% of the pairwise  $F_{ST}$  comparisons within the seven geographically-determined regions were significant (which equates to 19% of the total combinations), in contrast with 50.5% of the pairwise comparisons among regions (i.e. 59% of the total, Additional file 2: Table S1). However, when using Bonferroni correction for multiple comparisons, the corrected critical value dropped to

0.00005, which involved a decrease from 2.7 to 0.1% of significant pairwise  $F_{ST}$  within regions, and from 50.5 to 20.5% among regions. Similarly, the genetic matrix for Dataset 2 revealed that 4% of the pairwise  $F_{ST}$  comparisons within the seven geographically-determined regions were significant (i.e. 24% of the total combinations), in contrast with the 55% found among regions (65.6% of the total, Table 4). The corrected critical value dropped to 0.0003 after Bonferroni correction resulting in no significant comparisons within regions, and a drop of 55 to 27% of significant pairwise  $F_{ST}$  among regions. These data suggest isolation by distance, in agreement with the fact that pairwise  $F_{ST}$  values were found to be strongly correlated with the ordering of the samples in the



**Fig. 2** Number of alleles (black line) and allelic richness,  $A_r$  (red line) per sampling site for **(a)** Dataset 1 (6 microsatellites) and **(b)** Dataset 2 (15 microsatellites). Samples are ordered from north to south. In graph **a**,  $A_r$  experienced an increasing N-S trend ( $\tau = 0.296$ ,  $P = 0.00045$ ) but not the number of alleles ( $\tau = 0.153$ ,  $P = 0.074$ ); whereas in graph **b**, neither number of alleles nor  $A_r$  showed any sort of geographic trend ( $\tau = 0.243$ ,  $P = 0.1721$ ; and  $\tau = 0.0543$ ,  $P = 0.7888$ , respectively)

north-south gradient, both for Dataset 1 ( $\tau = 0.899$ ,  $P < 2.22 \times 10^{-16}$ ) and Dataset 2 ( $\tau = 0.895$ ,  $P = 1.79 \times 10^{-6}$ ) (Fig. 3a-b).

Hierarchical AMOVAs were conducted for datasets 1 and 2 using the aforementioned geographic approach to define regions. In both cases, the three levels of division (among regions, among populations within regions and within populations) showed significant values (Table 5). The differentiation observed among regions was higher than the differentiation observed among populations within regions, however, most of the observed genetic variance was hosted within populations (> 99% according to the microsatellites in Datasets 1 and 2).

Hierarchical AMOVAs per locus showed that the locus Gmo132 was the only one significant at all levels of division in both datasets (Table 6). This marker, which was depicted to be under directional selection by LOSITAN

analysis, showed the highest differentiation within populations (average  $F_{ST}$  of 0.034). Locus Gmo34 showed significant  $F_{CT}$  and  $F_{ST}$  (average  $F_{ST} = 0.005$ ) in all analyses. The extra set of nine microsatellites used in Dataset 2 showed one locus also significant at all levels (Gmo45), but also revealing a weak degree of structuring (average  $F_{ST} = 0.006$ ). Interestingly, the most frequent alleles for both Gmo34 and Gmo132 displayed a strong and highly significant trend in the north-south gradient (Fig. 4). No such trend was observed for the most common alleles in Gmo45 (data not presented). Allele 96 from locus Gmo34 ranged from 0.739–0.375 ( $\tau = -0.448$ ,  $P = 1.174 \times 10^{-7}$ ) and allele 115 from locus Gmo132 ranged from 0.681–0.105 ( $\tau = -0.626$ ,  $P = 1.111 \times 10^{-13}$ ) in the direction north to south.

Looking at the results for the Pan I locus, that was analysed for all samples, the three different genotypes

**Table 2** Summary statistics for Dataset 2 (18 samples genotyped at 15 microsatellites): sample size; number of alleles; AR, allelic richness (based on 33 diploid individuals); observed ( $H_o$ ) and unbiased expected heterozygosity ( $uHe$ ), inbreeding coefficient ( $F_{IS}$ ), number of deviations from Hardy-Weinberg equilibrium (HWE) and from Linkage Disequilibrium (LD) at  $\alpha = 0.05$ 

Region	Sample	Sample size	No alleles	Ar	$H_o$	$uHe$	$F_{IS}$	No dev. HWE	No dev. LD	Ne	
										0.05	0.01
Finnmark	Smørfjord_2003	68	189	11.0	0.650	0.728	0.095	6	4	$\infty$	$\infty$
Troms	Balsfjord_2003	77	213	11.7	0.703	0.737	0.029	5	3	$\infty$	$\infty$
Troms	Tranøybotn_2005	89	230	12.2	0.681	0.748	0.082	6	6	$\infty$	$\infty$
Troms	Gratangen_2005	63	205	12.5	0.669	0.696	0.023	3	1	$\infty$	1323.2
Troms	Gratangen_2006	62	227	12.8	0.696	0.752	0.067	3	4	$\infty$	$\infty$
Lofoten	Eidsfjord_2004	77	211	11.6	0.691	0.753	0.072	4	8	$\infty$	$\infty$
Lofoten	Tysfjord_2003	42	182	12.3	0.668	0.719	0.051	3	4	$\infty$	$\infty$
Lofoten	Baldkjosen_Nordfolda_2004	53	212	12.5	0.734	0.763	0.029	4	2	$\infty$	$\infty$
Helgeland	Skjerstadfjord_Valnesfjord_2005	88	229	12.3	0.697	0.764	0.084	4	4	2338.6	3281.5
Helgeland	Verrabotn_2005	73	217	12.2	0.710	0.752	0.043	4	4	$\infty$	$\infty$
Møre	Vinjefjord_2004	96	237	12.3	0.711	0.768	0.066	6	4	$\infty$	$\infty$
Møre	Midsund_2004	67	222	13.5	0.680	0.722	0.060	4	18	451.6	377.1
Møre	Godøy_2004	64	226	12.7	0.722	0.745	0.019	4	4	$\infty$	$\infty$
Møre	Borgunfjord_2005	90	229	12.2	0.713	0.765	0.063	5	2	$\infty$	5147.1
Vestlandet	Finnøy_2007	33	158	12.2	0.632	0.669	0.049	2	1	$\infty$	$\infty$
Sørlandet	Lillesand_2007	80	223	12.1	0.714	0.750	0.044	6	2	$\infty$	$\infty$
Sørlandet	Risør_ytre_2005	86	235	12.4	0.698	0.769	0.091	7	4	$\infty$	$\infty$
Sørlandet	Oslofjord_2005	87	223	11.7	0.705	0.744	0.043	3	4	1702.2	2914.6

Effective population size (Ne) was calculated using two values of lowest allele frequency (0.05 and 0.01). Sites have been ordered from north to south and information about the geographic region where they are placed

**Table 3** Summary information for the markers used in this study: linkage group (LG) to which they belong to, whether the linkage group displays a known inversion, global  $F_{ST}$  (and associated  $P$ -value after 10,000 permutations) and number of alleles. n/a stands for "not available"

Locus	LG	Inversions	Global $F_{ST}$ ( $P$ -value)		No alleles	
			Dataset 1	Dataset 2	Dataset 1	Dataset 2
Gmo2	6	No	<b>0.0021 (<math>P = 0.0114</math>)</b>	<b>0.0025 (<math>P = 0.0324</math>)</b>	23	23
Gmo3	18	No	<b>0.0027 (<math>P = 0.0180</math>)</b>	0.0022 ( $P = 0.1191$ )	13	10
Gmo34	1	Yes	<b>0.0047 (<math>P &lt; 0.0001</math>)</b>	<b>0.0028 (<math>P = 0.0308</math>)</b>	10	10
Gmo35	20	No	0.0009 ( $P = 0.1757$ )	<b>0.0030 (<math>P = 0.0200</math>)</b>	16	16
Gmo132	7	Yes	<b>0.0283 (<math>P &lt; 0.0001</math>)</b>	<b>0.0269 (<math>P = 0.0000</math>)</b>	52	41
Tch11	n/a	n/a	<b>0.0020 (<math>P = 0.0003</math>)</b>	0.0015 ( $P = 0.0788$ )	29	27
GmoC18	n/a	n/a		0.0002 ( $P = 0.4351$ )		19
GmoC20	n/a	n/a		0.0000 ( $P = 0.5175$ )		21
GmoG13	n/a	n/a		0.0021 ( $P = 0.1768$ )		18
GmoG18	n/a	n/a		*****		5
GmoG25	n/a	n/a		*****		63
GmoG40	n/a	n/a		0.0035 ( $P = 0.1111$ )		16
GmoG43	n/a	n/a		0.0011 ( $P = 0.1241$ )		30
GmoG45	n/a	n/a		<b>0.0056 (<math>P = 0.0000</math>)</b>		47
Tch22	n/a	n/a		*****		6
Pan I	1	Yes	<b>0.107 (<math>P &lt; 0.0001</math>)</b>		2	

LG is in accordance with nomenclature from Hubert et al. [48]. Microsatellites were blasted against the gadMor2 assembly [101] by Per Erik Jorde (pers. comm)  $P$ -values in boldface type were significantly different from zero



**Table 4** Genetic differentiation between the 18 samples in Dataset 2 (15 microsatellite markers): pairwise  $F_{ST}$  (below diagonal) and  $P$ -value after 10,000 permutations (above diagonal)

	Smø03	Balsf03	Tran05	Grat05	Grat06	Eidsf04	Tysf03	BaldN04	SkjVal05	Verra05	Vinjef04	Mids04	God04	Borg05	Fin07	Lil07	Ris05	Oslo05
Smø03	*	0.0863	<b>0.0267</b>	0.6862	<b>0.0199</b>	<b>0.0310</b>	<b>0.0146</b>	<b>0.0015</b>	<b>0.0032</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0001</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
Balsf03	0.0021	*	0.9300	0.8381	<b>0.0417</b>	0.0535	0.2607	<b>0.0410</b>	<b>0.0104</b>	<b>0.0000</b>	<b>0.0001</b>	<b>0.0000</b>	<b>0.0314</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
Tran05	<b>0.0029</b>	0.0000	*	0.5446	0.0698	0.3481	0.4903	0.1710	0.2027	<b>0.0000</b>	0.0832	<b>0.0002</b>	0.9505	<b>0.0044</b>	<b>0.0196</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
Grat05	0.0000	0.0000	0.0000	*	0.3636	0.2286	0.0911	0.1579	0.3343	<b>0.0022</b>	<b>0.0012</b>	<b>0.0000</b>	0.2597	<b>0.0128</b>	<b>0.0007</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
Grat06	<b>0.0037</b>	<b>0.0028</b>	0.0023	0.0007	*	0.4292	<b>0.0230</b>	<b>0.0121</b>	0.0713	<b>0.0001</b>	<b>0.0016</b>	<b>0.0000</b>	<b>0.0249</b>	<b>0.0058</b>	<b>0.0005</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
Eidsf04	<b>0.0029</b>	<b>0.0022</b>	0.0006	0.0011	0.0005	*	0.2286	0.1716	0.8121	<b>0.0048</b>	<b>0.0243</b>	<b>0.0011</b>	0.2701	<b>0.0056</b>	0.0641	<b>0.0000</b>	<b>0.0001</b>	<b>0.0019</b>
Tysf03	<b>0.0047</b>	<b>0.0013</b>	0.0003	0.0026	<b>0.0046</b>	0.0015	*	0.4561	0.1450	<b>0.0200</b>	0.2457	<b>0.0011</b>	0.6782	0.0699	0.1202	<b>0.0000</b>	<b>0.0004</b>	<b>0.0252</b>
BaldN04	<b>0.0063</b>	<b>0.0030</b>	0.0016	0.0016	<b>0.0044</b>	0.0015	0.0005	*	0.3055	<b>0.0358</b>	0.9137	0.0955	0.2360	0.8199	0.6309	<b>0.0044</b>	<b>0.0128</b>	<b>0.0136</b>
SkjVal05	<b>0.0044</b>	<b>0.0031</b>	<b>0.0012</b>	0.0008	0.0024	0.0000	0.0022	0.0011	*	<b>0.0390</b>	<b>0.0333</b>	<b>0.0008</b>	0.2781	0.1324	0.0941	<b>0.0000</b>	<b>0.0001</b>	<b>0.0008</b>
Verra05	<b>0.0078</b>	<b>0.0086</b>	<b>0.0078</b>	<b>0.0049</b>	<b>0.0068</b>	<b>0.0039</b>	<b>0.0045</b>	<b>0.0033</b>	<b>0.0026</b>	*	<b>0.0004</b>	<b>0.0001</b>	<b>0.0001</b>	<b>0.0047</b>	<b>0.0079</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
Vinjef04	<b>0.0099</b>	<b>0.0063</b>	<b>0.0016</b>	<b>0.0045</b>	<b>0.0046</b>	<b>0.0025</b>	<b>0.0013</b>	<b>0.0000</b>	<b>0.0022</b>	<b>0.0049</b>	*	0.2519	0.5689	0.9032	0.8840	0.1737	0.4764	<b>0.0284</b>
Mids04	<b>0.0154</b>	<b>0.0109</b>	<b>0.0059</b>	<b>0.0079</b>	<b>0.0079</b>	<b>0.0051</b>	<b>0.0070</b>	<b>0.0023</b>	<b>0.0049</b>	<b>0.0076</b>	<b>0.0010</b>	*	0.0610	0.3798	0.6948	0.3487	0.3171	<b>0.0229</b>
God04	<b>0.0060</b>	<b>0.0028</b>	<b>0.0000</b>	<b>0.0010</b>	<b>0.0033</b>	<b>0.0009</b>	<b>0.0000</b>	<b>0.0012</b>	<b>0.0010</b>	<b>0.0064</b>	<b>0.0000</b>	<b>0.0024</b>	*	0.5744	0.8690	0.1331	<b>0.0183</b>	<b>0.0244</b>
Borg05	<b>0.0099</b>	<b>0.0058</b>	<b>0.0034</b>	<b>0.0033</b>	<b>0.0040</b>	<b>0.0034</b>	<b>0.0028</b>	<b>0.0000</b>	<b>0.0015</b>	<b>0.0037</b>	<b>0.0000</b>	<b>0.0007</b>	<b>0.0000</b>	*	0.9163	0.0535	<b>0.0433</b>	<b>0.0017</b>
Fin07	<b>0.0128</b>	<b>0.0096</b>	<b>0.0048</b>	<b>0.0075</b>	<b>0.0084</b>	<b>0.0034</b>	<b>0.0033</b>	<b>0.0000</b>	<b>0.0032</b>	<b>0.0063</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	*	0.9114	0.9212	0.1385
Lil07	<b>0.0175</b>	<b>0.0139</b>	<b>0.0081</b>	<b>0.0109</b>	<b>0.0116</b>	<b>0.0079</b>	<b>0.0082</b>	<b>0.0041</b>	<b>0.0075</b>	<b>0.0110</b>	<b>0.0011</b>	<b>0.0006</b>	<b>0.0014</b>	<b>0.0020</b>	<b>0.0000</b>	*	0.6593	0.0938
Ris05	<b>0.0178</b>	<b>0.0150</b>	<b>0.0080</b>	<b>0.0122</b>	<b>0.0100</b>	<b>0.0061</b>	<b>0.0072</b>	<b>0.0038</b>	<b>0.0069</b>	<b>0.0093</b>	<b>0.0003</b>	<b>0.0009</b>	<b>0.0032</b>	<b>0.0022</b>	<b>0.0000</b>	<b>0.0000</b>	*	<b>0.0034</b>
Oslo05	<b>0.0098</b>	<b>0.0101</b>	<b>0.0060</b>	<b>0.0067</b>	<b>0.0086</b>	<b>0.0038</b>	<b>0.0035</b>	<b>0.0035</b>	<b>0.0042</b>	<b>0.0068</b>	<b>0.0021</b>	<b>0.0027</b>	<b>0.0025</b>	<b>0.0038</b>	<b>0.0021</b>	<b>0.0015</b>	<b>0.0035</b>	*

Likewise, site names have been shortened (see Summary table for full names). Initially, some 59% of the comparisons (90 out of 153) were significantly different from zero but, after Bonferroni correction (corrected critical value = 0.0004), the percentage reduced to 27% (41 out of 153)  $F_{ST}$  with associated  $P$ -values < 0.05 have been shaded in grey to ease the reading

displayed clearly distinct frequencies throughout the NCC samples (Fig. 5). Genotype AA was the most frequent and experienced an increasing trend towards the south, whereas genotype BB, present in low or very low frequencies in all populations, showed a decreasing southwards trend. The frequency of allele Pan I A revealed a negative trend towards the south ( $\tau = 0.23$ ,  $P = 0.0068$ , Fig. 4). Overall genetic structure for Pan I locus was highly significant ( $F_{ST} = 0.107$ ,  $P < 0.0001$ ) and hierarchical AMOVA was also significant at the three levels of grouping (Table 5). Likewise, pairwise  $F_{ST}$  for Pan I showed a highly significant north-south trend ( $\tau = 0.871$ ,  $P < 2.22 \times 10^{-16}$ ).

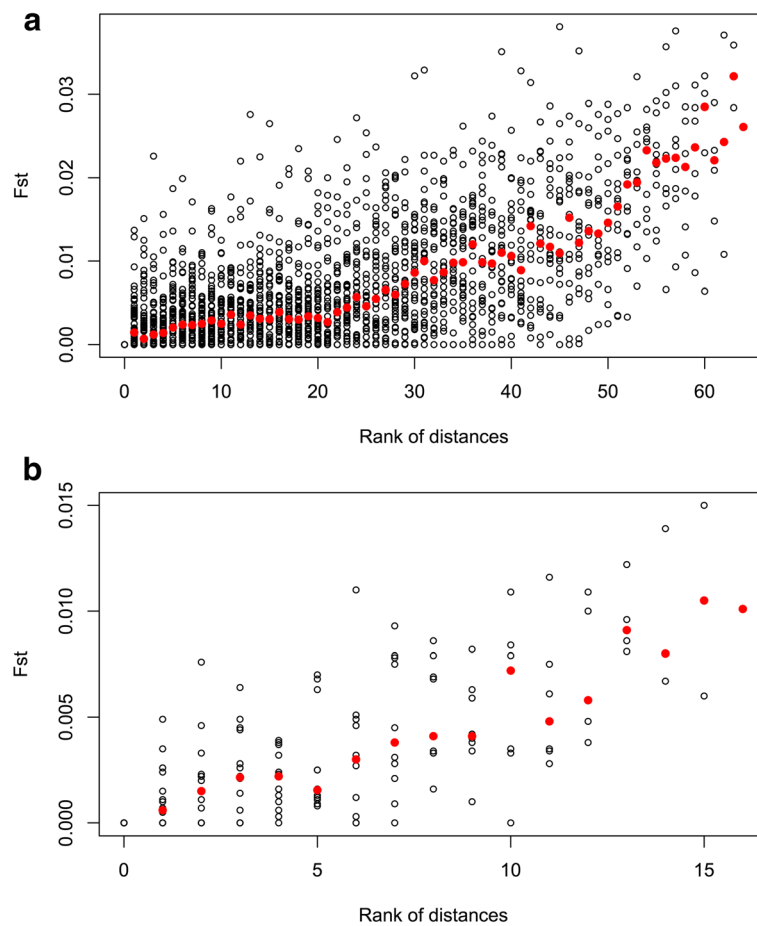
In order to investigate the potential admixture between both types of cod, NEAC was used as an outlier group for NCC in STRUCTURE. These analyses showed no clear clustering of populations in the full dataset genotyped at 6 microsatellites (Fig. 6a) in agreement with BARRIER (results not shown). However, the two outlier loci (Gmo32 and Gmo134) displayed a north-south gradient of admixture (Fig. 6b) whereby NCC showed greater genetic similarity to NEAC in the north than in the south. A similar trend could also be observed in the restricted dataset genotyped at 15 microsatellites (Fig. 6c).

**Discussion**

Norway is the country with the largest remaining commercial catch of Atlantic cod, and this is the first study to investigate population genetic structure of coastal cod (NCC) along the entire Norwegian

coastline. Following the analysis of > 4000 cod sampled from 55 locations, we obtained the following main results: 1. Statistically significant population genetic differentiation was revealed among most of the samples, 2. The observed genetic differentiation followed a pattern of isolation by distance along the north-south gradient, without any clear “breaks”, 3. No distinct change in genetic variation (i.e. allelic diversity or heterozygosity) was observed among the samples in the north-south gradient. Based upon these results, we conclude that NCC displays statistically significant population genetic structure along the Norwegian coastline. Consequently, these results demonstrate that the current management regime, dividing coastal cod in Norway into two management groups, one north and one south of 62°, represents an over-simplification of the true level of population genetic structure, and as such, requires re-evaluation.

As detailed in the introduction, Atlantic cod is one of the best-studied marine fishes, and a large number of population genetic studies have revealed a species displaying extensive population genetic structure (reviewed by [5–7]). While several of the newest studies in this species have utilised genomics methods to identify and characterise population genetic structure (e.g. [29, 32, 38, 41, 45]), our study based upon microsatellites and Pantophysin, has nevertheless made a significant contribution to our understanding of population genetic structure in this species, especially in Norway. First the biological, and then the management implications of our results are discussed below.



**Fig. 3** Pairwise  $F_{ST}$  for: **a)** Dataset 1 and **b)** Dataset 2. Black empty circles depict the pairwise  $F_{ST}$  for pairs of sites within the same rank of distances whereas filled red circles dots depict the median for each class. Both trends are highly significant and reveal increasing levels of differentiation correlating with the rank of distances:  $\tau = 0.899, P \leq 2.22 \text{ e-}16$ ;  $\tau = 0.895, P = 1.79 \text{ e-}06$  and  $\tau = 0.871, P < 2.22 \text{ e-}16$  for **a)** and **b)**, respectively

**Population genetic structure of NCC**

The primary aim of the present study was to investigate population genetic differentiation among coastal cod sampled from the entire Norwegian coastline, which stretches over 2500 km from the south in Oslofjord, to

the north in Porsangerfjord (Fig. 1). In some areas of Norway, in particular the north, NEAC are found using the same spawning areas as NCC [80, 81]. Consequently, some NEAC were inadvertently collected during the sampling process, especially in the north (Table 1). In

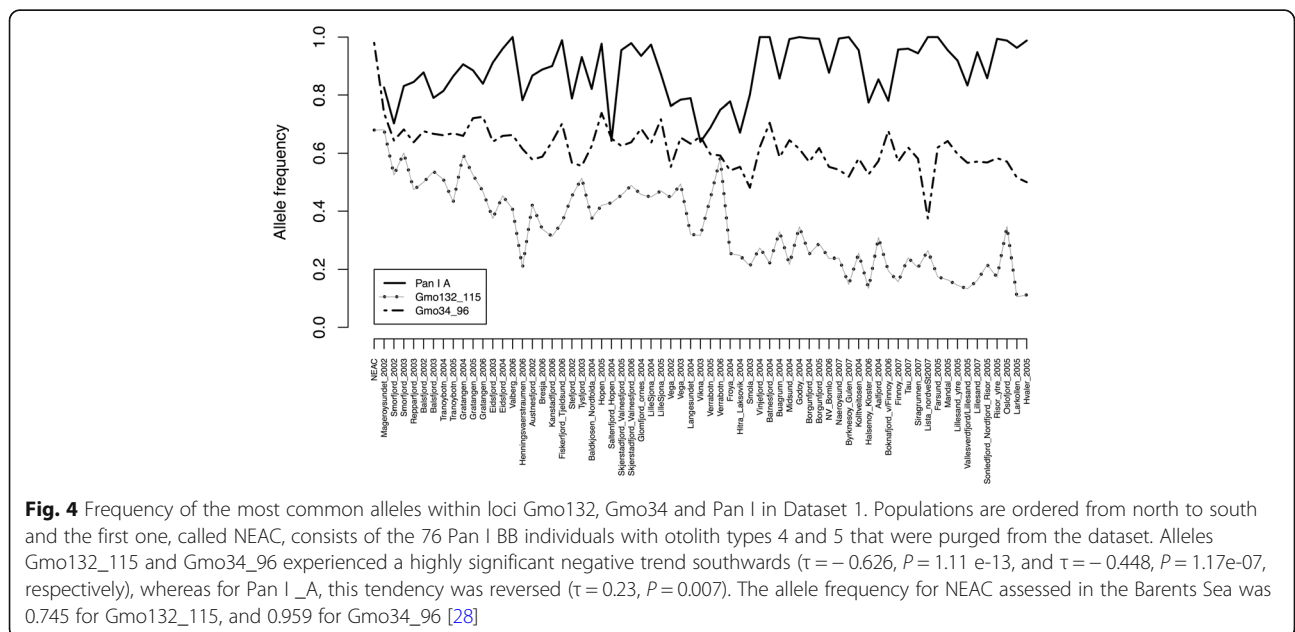
**Table 5** Hierarchical AMOVA based on geographic regions for both sets of microsatellite data and Pan I

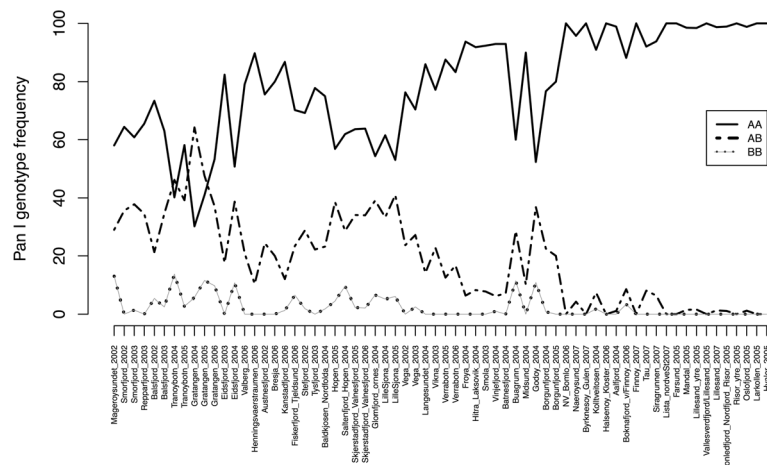
Loci	Source of variation	d.f	S.S.	Variance components (% of variation)	F-statistics
6 microsatellites (Dataset 1)	Among regions	6	118.39	0.014 (0.65%)	$F_{CT} = 0.0065 (P < 0.0001)$
	Among populations within regions	59	155.76	0.004 (0.19%)	$F_{SC} = 0.0019 (P < 0.0001)$
	Within populations	8626	18,279.78	2.119 (99.16%)	$F_{ST} = 0.0084 (P < 0.0001)$
15 microsatellites (Dataset 2)	Among regions	6	67.83	0.017 (0.36%)	$F_{CT} = 0.0036 (P = 0.0001)$
	Among populations within regions	11	60.39	0.006 (0.13%)	$F_{SC} = 0.0013 (P = 0.0081)$
	Within populations	2572	11,769.03	4.626 (99.51%)	$F_{ST} = 0.0049 (P < 0.0001)$
Pan I (Dataset 1)	Among regions	6	66.97	0.009 (8.29%)	$F_{CT} = 0.0829 (P < 0.0001)$
	Among populations within regions	59	33.34	0.004 (3.48%)	$F_{SC} = 0.0379 (P < 0.0001)$
	Within populations	8626	793.87	0.092 (88.24%)	$F_{ST} = 0.1176 (P < 0.0001)$

**Table 6** Hierarchical AMOVA per locus based on geographic regions for both data sets

Data set	Locus	F <sub>CT</sub>	P-value	F <sub>SC</sub>	P-value	F <sub>ST</sub>	P-value
Dataset 1	Gmo2	0.0005	<b>0.0361</b>	0.0017	0.0659	0.0022	<b>0.0116</b>
	Gmo3	0.0008	0.0819	0.0020	0.0734	0.0028	<b>0.0171</b>
	Gmo34	0.0044	<b>0.0000</b>	0.0009	0.1809	0.0053	<b>0.0000</b>
	Gmo35	0.0001	0.4140	0.0008	0.2144	0.0009	0.1757
	<b>Gmo132</b>	0.0279	<b>0.0000</b>	0.0044	<b>0.0000</b>	0.0322	<b>0.0000</b>
	Tch11	0.0009	<b>0.0000</b>	0.0013	0.0517	0.0022	<b>0.0001</b>
Dataset 2	Gmo2	0.0001	0.4766	0.0024	0.0593	0.0026	<b>0.0293</b>
	Gmo3	0.0054	<b>0.0029</b>	0.0000	0.9186	0.0029	0.1193
	Gmo34	0.0057	<b>0.0028</b>	0.0000	0.8920	0.0036	<b>0.0344</b>
	Gmo35	0.0017	0.1859	0.0015	0.2201	0.0033	<b>0.0209</b>
	<b>Gmo132</b>	0.0249	<b>0.0003</b>	0.0054	<b>0.0003</b>	0.0302	<b>0.0000</b>
	Tch11	0.0005	0.2819	0.0011	0.2404	0.0016	0.0779
	GmoC18	0.0006	0.1656	0.0000	0.6458	0.0002	0.4323
	GmoC20	0.0008	0.1404	0.0000	0.7495	0.0001	0.5132
	GmoG13	0.0003	0.4134	0.0019	0.2641	0.0022	0.1730
	GmoG18	*****	*****	*****	*****	*****	*****
	GmoG25	*****	*****	*****	*****	*****	*****
	GmoG40	0.0017	0.1597	0.0020	0.3855	0.0037	0.1067
	GmoG43	0.0006	0.2317	0.0006	0.3822	0.0012	0.1233
	<b>GmoG45</b>	0.0034	<b>0.0027</b>	0.0027	<b>0.0012</b>	0.0061	<b>0.0000</b>
	Tch22	*****	*****	*****	*****	*****	*****

P-values in boldface type were significantly different from zero





**Fig. 5** Frequency of the three different genotypes of locus Pan I in Dataset 1. Sites are ordered from north to south

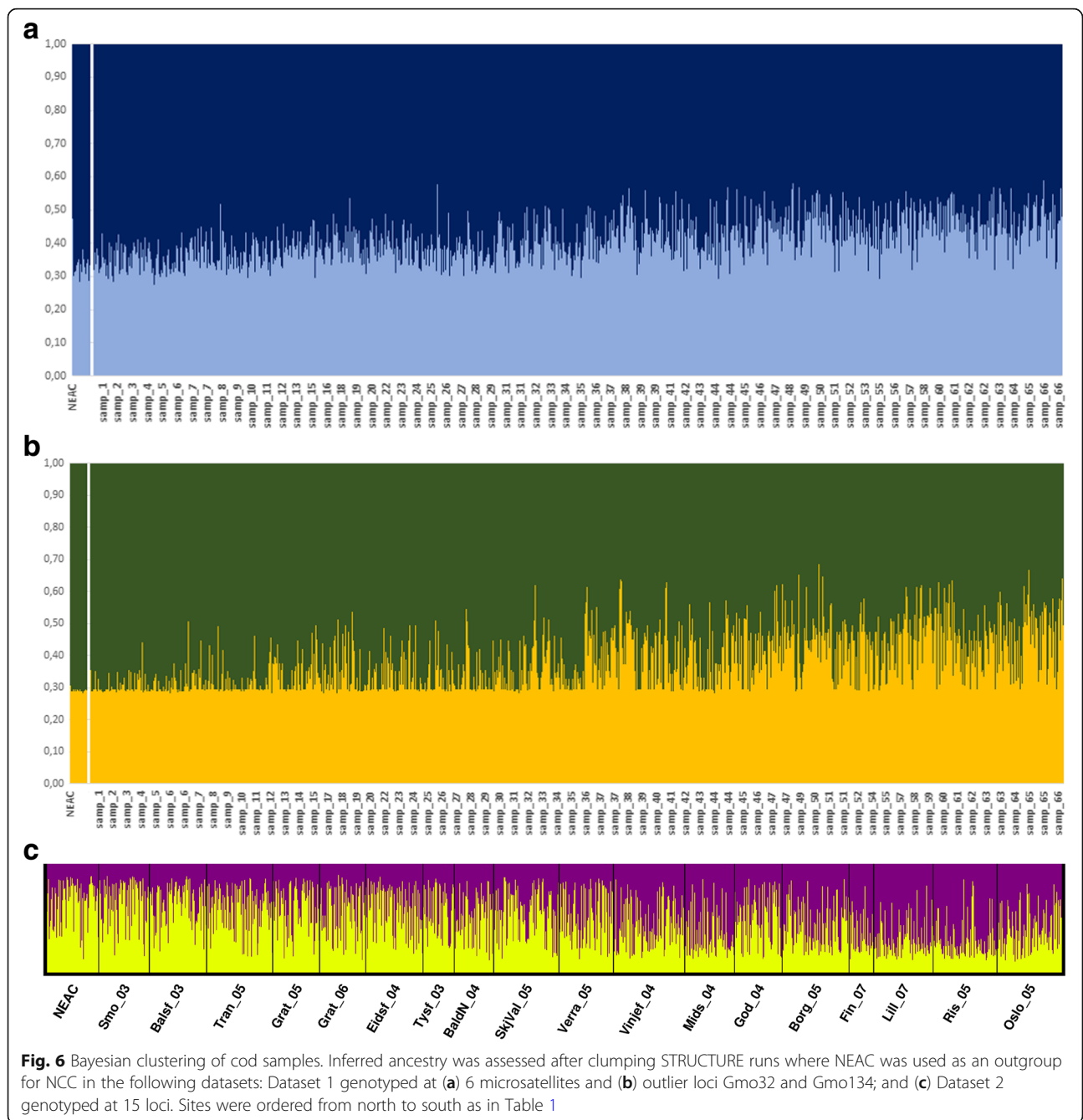
order to eliminate the potential bias caused by NEAC in the present analysis of NCC population genetic structure, we used otolith categories 4 and 5 to exclude any potential NEAC from our biological samples prior to genetic analysis [58].

The Pan I BB genotype is nearly fixed for NEAC [26], and is almost completely diagnostic between NCC and NEAC [18, 26, 27]. In addition, Pan I genotype and otolith type both clearly differentiate between NEAC and NCC [27, 82]. Despite purging otolith category 4 and 5 individuals from the current data set, 104 fish displaying an otolith structure typical for coastal cod (i.e., category 1–2) with the Pan I BB homozygote genotype (i.e. characteristic for NEAC) remained. Importantly however, and in the context of the present study, both including and excluding these 104 BB homozygote individuals did not influence the overall picture of population genetic structure. Therefore, it is concluded that we have effectively excluded NEAC from these analyses, and that the differences reported here primarily reflect genetic differences between NCC populations along the Norwegian coastline. Given that genetic differentiation did not show any clear genetic groupings, nor clear breaks in population connectivity, we suggest that NCC populations belong in a genetic gradient characterised by isolation by distance. This pattern is potentially driven by the generally limited migratory behaviour of NCC as has been seen in tagging experiments ([24, 81, 83], but see [84]), spawning site fidelity [41, 85], and potentially retention of eggs in certain fjord areas [83, 86].

A distinct north-south cline in the frequency of the allele 115 for the microsatellite Gmo132 was observed here (Fig. 4). This pattern may be explained using two different scenarios: 1) Since this genetic marker has been tagged as an outlier in previous genetic analyses [26, 87],

the observed north-south cline at this locus might have been shaped under the influence of environmental factors. In our data however, significant allele frequency differences for this locus also existed among samples within the same region, thus it is unlikely that selection alone has created the observed north-south allelic cline (see [28]). 2) Although NEAC were removed from most of the analyses of the present study, this does not preclude the possibility that genetic introgression and admixture between NEAC and NCC contributes to the observed pattern of population genetic structure in NCC. Supporting this suggestion is the fact that the frequency of the 115 allele in Gmo132, in all of the northern samples investigated here, displayed a frequency around 0.5–0.6, which is very similar to the frequency of this allele in NEAC [ $\sim 0.7$  28]. Therefore, admixture between NEAC and NCC, primarily in the north and following a decreasing gradient of presence towards the south, may contribute to the pattern of population genetic structure observed in NCC. This suggestion is consistent with simulations [88] illustrating that introgression from a genetically distinct source (e.g., NEAC) may generate gradients in allele frequencies along a geographic axis originating at the edge of the contact zone (i.e., north in the current study). These “tails of introgression” may extend to large distances beyond the contact zone itself (i.e., towards the south in the cod case). Importantly, the gradient only appears where dispersal and gene-flow are spatially limited. In our case, the limited gene-flow among NCC populations in Norway results in a gradient in the allele frequency, which starts in the north due to introgression of NEAC.

Population genetic differentiation was observed among samples of NCC throughout all regions of Norway, including among samples from southern



Norway where NEAC was hardly present (as determined either by Pan I genotype or otolith category). In the far southern region, it is unlikely that NEAC has played a direct role in shaping the evolutionary relationships among NCC populations (although it is possible that the tail of potential introgression from NEAC in the north could still extend to the far south and thus play a minor role). However, the interaction between migratory North Sea cod and NCC populations in the south of Norway also remains a potential source of influence on NCC

structure in this region [89]. Clearly, the evolutionary relationship between NEAC and NCC, and North Sea cod and NCC, needs to be resolved in order to evaluate the degree of influence migratory ecotypes have on the population genetic structure of NCC throughout Norway.

Current evidence suggests that the long-distance migratory “ecotype” of Atlantic cod, including NEAC, was derived from the stationary “ecotype” [29], and that this divergence occurred prior to the split between the Northeast and Northwest cod populations [90], which

has been estimated to have occurred approximately 100,000 years ago [91, 92]. As stated above, NEAC and NCC exhibit large genetic and genomic differences between them, and while several potential mechanisms have been proposed [80, 93, 94], the ecological processes leading to and maintaining segregation of these two ecotypes, and the degree of genetic exchange between them in both time and space, remain unresolved. The (~2%) BB homozygotes observed in the present data set, that displayed an otolith structure typical for coastal cod, could fit in one of these scenarios: 1 – “true” NEAC that for some reason did not migrate to and from the Barents sea as is characteristic for NEAC, 2 – hybrid and or admixed individuals between NEAC and NCC that did not migrate to and from the Barents sea, possibly due to only carrying one set of chromosomes with the “inverted supergene” and therefore displaying reduced propensity for long-distance migration, 3 – “true” NCC cod that display the BB genotype as the B allele is observed in NCC although at low frequencies. Quantifying admixture between NEAC and NCC remains a challenge that even recent papers focussing on the genome-wide differences between NEAC and NCC have not completely resolved [29, 32]. Pan I lies within one of the inverted regions of chromosome 1, which together with inversions in linkage groups 2 and 7 are responsible for nearly all of the genomic divergence between NEAC and NCC [29, 32]. As these inversions block recombination in their respective locations on the genome for NEAC, this means that if NEAC and NCC hybridise, their offspring will contain one copy of the inverted parts of linkage groups 1, 2 and 7, and one copy of the ancestral collinear form. We suggest that a genomic analysis of BB homozygotes displaying otolith categories 1–2 may provide an important resource in resolving the issue of hybridisation between NEAC and NCC, and the degree to which this may or may not influence population genetic structure of NCC [95]. Finally, the discrepancy in the allele frequencies in Pan I (linkage group 1) and Gmo132 (linkage group 7) reported between NEAC and NCC along the north-south gradient require further investigation.

### Management implications

Applying the same management strategy to multiple populations or stock components that vary in their abundance and/or resilience to exploitation inevitably results in overfishing and likely collapse of the weaker component [96, 97]. The extent of population genetic structure revealed in the present study, irrespective if it is influenced by a gradient of NEAC admixture or not, and divergent selective forces or not, strongly suggests that the current division of coastal cod in Norway, above and below 62° north, will not be sufficient to ensure

sustainable management of NCC throughout Norway. Our data clearly illustrate population genetic structure within all areas of Norway, and as such, needs to be taken into consideration when renewing management plans. This conclusion is supported by earlier studies of population genetic structure in the south of Norway, where temporally-stable population genetic differentiation has been observed between neighbouring fjords over relatively small distances, effectively demonstrating limited connectivity at least between some fjord systems [24, 98, 99]. Additionally, earlier data from the north of Norway, revealing genetic differences between coastal cod sampled in fjords, supports the main results from the present study [14, 19, 26, 28]. It is nevertheless acknowledged that the lack of any clear “breaks” in population connectivity along the Norwegian coastline means that identifying appropriate management units is challenging.

An additional management contribution from the present analyses is the apparent lack of any clear differences in genetic variation, as reported by numbers of alleles, allelic richness, heterozygosity or the effective population size (Fig. 2-4; Additional file 3), in the north-south gradient. In all areas of Norway, NCC has been documented in decline since 1990s, to which overfishing has probably played a significant role. However, in the southern regions, populations also appear to be influenced by climate-driven recruitment challenges as has been illustrated by the beach-seine survey conducted in this region since 1919 [100]. Despite this, based upon the samples analysed here, there are no indications of severe genetic bottlenecks in any of the samples, and no clear differences in genetic diversity estimates between the north and south.

### Conclusions

Norwegian coastal cod displays statistically significant population genetic structure along the Norwegian coastline, which follows a genetic gradient characterised by isolation by distance. Although not fully resolved, we suggest that genetic introgression and admixture between NEAC and NCC, most in the north and least in the south, together with limited gene-flow among NCC populations, may contribute to the observed population genetic structure of NCC. In turn, we conclude that the current management regime in place, dividing coastal cod in Norway into two management groups, north and south of 62°, represents an over-simplification of the true level of population genetic structure, and as such, requires re-evaluation.

### Additional files

**Additional file 1: Figure S1.** Ho and uHe per sampling site for (a) Dataset 1 (66 samples genotyped at 6 microsatellite markers) and (b) Dataset 2 (18 samples genotyped at 15 microsatellite markers). Locations

are ordered from north to south along the coastline. In graph a), Ho and uHe experienced significantly increasing N-S trends ( $\tau = 0.363$ ,  $P = 1.705 \times 10^{-5}$ , and  $\tau = 0.570$ ,  $P < 2.22 \times 10^{-16}$ , respectively). Unlikewise, in graph b), neither Ho nor uHe showed any kind of N-S trend ( $\tau = 0.281$ ,  $P = 0.11164$  and  $\tau = 0.21$ ,  $P = 0.23997$ , respectively). (ZIP 1059 kb)

**Additional file 2: Table S1.** Genetic differentiation at Dataset 1 (66 samples genotyped at 6 microsatellite markers): Pairwise  $F_{ST}$  are shown in the lower diagonal whereas  $P$ -values computed after 10,000 permutations are shown in the lower diagonal. (XLSX 77 kb)

**Additional file 3: Raw data.** Genotypes for Dataset 1 (66 samples genotyped at 6 microsatellites) and Dataset 2 (18 samples genotyped at 15 microsatellites). Pan I genotypes, and otolith information is provided for each individual as well. (XLSX 563 kb)

### Abbreviations

AR: Allele Richness;  $F_{CT}$ : F-statistics among groups/regions;  $F_{SC}$ : F-statistics within populations;  $F_{ST}$ : a commonly used measure of population differentiation due to genetic structure;  $H_O$ : observed heterozygosity; LD: Linkage Disequilibrium; NCC: Norwegian Coastal Cod; NEAC: North East Arctic Cod; NSC: North Sea Cod; Pan I: Panthophysin I; PCR: Polymerase Chain Reaction; SNP(s): Single Nucleotide Polymorphism(s); uHe: unbiased expected heterozygosity

### Acknowledgements

We would like to acknowledge the constructive comments by the two anonymous referees, and for the microsatellite linkage group locations supplied to us by Per Erik Jorde (IMR, Norway). This work could not have been made possible without the assistance from several people at the Institute of Marine Research; Ole I Paulsen, Eva Farestveit, Gunnar Bakke, Arve Kristiansen, Per Ågotnes, as well as the crew on board the RV's "Fangst", "GM Dannevig", "Johan Hjort" and "GO Sars". We also thank all the local fishermen along the Norwegian coast that helped us getting samples and providing us with information about spawning grounds. This study was financed by the Norwegian Ministry for Industry and Fisheries (NFD).

### Funding

The study was financed by the Norwegian Ministry for Industry and Fisheries.

### Availability of data and materials

Supporting information is available in the additional files and further information is available from the authors on request.

### Authors' contributions

KEJ conceived the study, GD conducted the microsatellite and panthophysin analyses, and drafted the paper. FB, MQ and JIW conducted the statistical analysis, KEJ, TJ and AA provided background material to the study. KAG led the process of data interpretation and the final version of the manuscript, with scientific contributions from all other authors. All authors read and approved the final manuscript.

### Ethics approval and consent to participate

Sampling (fin clippings) were conducted from dead fish caught in the legal commercial cod fisheries along the Norwegian coast.

### Consent for publication

There is no material in the manuscript (personal details, images or videos) that would require any consent to publish.

### Competing interests

The authors declare that they have no competing interests.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Institute of Marine Research (IMR), Postbox 1870, N-5817 Bergen, Norway.

<sup>2</sup>Institute of Marine Research (IMR), Postbox 6404, N-9019 Tromsø, Norway.

<sup>3</sup>Department of Biology, University of Bergen, Bergen, Norway.

Received: 26 September 2017 Accepted: 18 May 2018

Published online: 09 July 2018

### References

- Mieszowska N, Genner MJ, Hawkins SJ, Sims DW. Effects of climate change and commercial fishing on Atlantic cod *Gadus morhua*. *Adv Mar Biol.* 2009; 56:213–73.
- Myers RA, Hutchings JA, Barrowman NJ. Why do fish stocks collapse? The example of cod in Atlantic Canada. *Ecol Appl.* 1997;7(1):91–106.
- Hutchings JA. Thresholds for impaired species recovery. *Proc R Soc B Biol Sci.* 1809;282:2015.
- Olsen EM, Knutsen H, Gjøsæter J, Jorde PE, Knutsen JA, Stenseth NC. Life-history variation among local populations of Atlantic cod from the Norwegian Skagerrak coast. *J Fish Biol.* 2004;64(6):1725–30.
- Imslund AK, Jónsdóttir ÓDB. Linking population genetics and growth properties of Atlantic cod. *Rev Fish Biol Fish.* 2003;13(1):1–26.
- Nordeide JT, Johansen SD, Jørgensen TE, Karlsen BO, Moum T. Population connectivity among migratory and stationary cod *Gadus morhua* in the Northeast Atlantic—a review of 80 years of study. *Mar Ecol Prog Ser.* 2011; 435:269–83.
- O'Leary DB, Coughlan J, Dillane E, McCarthy TV, Cross TF. Microsatellite variation in cod *Gadus morhua* throughout its geographic range. *J Fish Biol.* 2007;70:310–35.
- Frydenberg OVE, Møller DAG, Nævdal G, Sick K. Haemoglobin polymorphism in Norwegian cod populations. *Hereditas.* 1965;53(1–2): 257–71.
- Møller D. Genetic diversities in spawning cod along the Norwegian coast. *Hereditas.* 1968;60:1–21.
- Sick K. Haemoglobin polymorphism of cod in the Baltic and the Danish Belt Sea. *Hereditas.* 1965;54
- Sick K. Haemoglobin polymorphism of cod in the North Sea and the North Atlantic Ocean. *Hereditas.* 1965;54(1):49–69.
- Møller D. Genetic differences between cod groups in the Lofoten area. *Nature.* 1966;212
- Jørstad KE, Dahl E, Danielssen DS, Moksness E, Solemdal P. Genetic analysis of cod in northern Norway. In: *The Propagation of Cod (Gadus morhua L.)*. Vol. 1. Arendal; 1984. p. 745–60.
- Jørstad KE, Nævdal G. Genetic variation and population structure of cod, *Gadus morhua* L., in some fjords in northern Norway. *J Fish Biol.* 1989;35: 245–52.
- Mork J, Ryman N, Ståhl G, Utter F, Sundnes G. Genetic variation in Atlantic cod (*Gadus morhua*) throughout its range. *Can J Fish Aquat Sci.* 1985;42(10): 1580–7.
- Árnason E, Pálsson S, Arason A. Gene flow and lack of population differentiation in Atlantic cod, *Gadus morhua* L., from Iceland, and comparison of cod from Norway and Newfoundland. *J Fish Biol.* 1992;40(5): 751–70.
- Dahle G. Cod, *Gadus morhua* L., populations identified by mitochondrial DNA. *J Fish Biol.* 1991;38(2):295–303.
- Fevolden SE, Pogson GH. Genetic divergence at the synaptophysin (Synp) locus among Norwegian coastal and north-East Arctic populations of Atlantic cod. *J Fish Biol.* 1997;51(5):895–908.
- Sarvas TH, Fevolden S-E. The scnDNA locus *Pan I* reveals concurrent presence of different populations of Atlantic cod (*Gadus morhua* L.) within a single fjord. *Fish Res.* 2005;76(3):307–16.
- Sarvas TH, Fevolden SE. Pantophysin (*Pan I*) locus divergence between inshore v. Offshore and northern v. Southern populations of Atlantic cod in the north-East Atlantic. *J Fish Biol.* 2005;67(2):444–69.
- Pogson GH, Fevolden S-E. Natural selection and the genetic differentiation of coastal and Arctic populations of the Atlantic cod in northern Norway: a test involving nucleotide sequence variation at the pantophysin (*Pan I*) locus. *Mol Ecol.* 2003;12(1):63–74.
- Fevolden SE, Westgaard JI, Pedersen T, Præbel K. Settling-depth vs. genotype and size vs. genotype correlations at the *Pan I* locus in 0-group Atlantic cod *Gadus morhua*. *Mar Ecol Prog Ser.* 2012;468:267–78.
- Heath MR, Culling MA, Crozier WW, Fox CJ, Gurney WSC, Hutchinson WF, Nielsen EE, O'Sullivan M, Preedy KF, Righton DA, et al. Combination of genetics and spatial modelling highlights the sensitivity of cod (*Gadus morhua*) population diversity in the North Sea to distributions of fishing. *ICES J Mar Sci.* 2014;71(4):794–807.

24. Knutsen H, Jorde PE, André C, Stenseth NC. Fine-scaled geographic population structuring in a highly mobile marine species: the Atlantic cod. *Mol Ecol*. 2003;12:385–94.
25. Nielsen EE, Hansen MM, Schmidt C, Meldrup D, Gronkjaer P. Fisheries: population of origin of Atlantic cod. *Nature*. 2001;413(6853):272–2.
26. Skarstein TH, Westgaard JI, Fevolden SE. Comparing microsatellite variation in north-East Atlantic cod (*Gadus morhua* L.) to genetic structuring as revealed by the pantophysin (*Pan I*) locus. *J Fish Biol*. 2007;70:271–90.
27. Wennevik V, Jørstad KE, Dahle G, Fevolden S-E. Mixed stock analysis and the power of different classes of molecular markers in discriminating coastal and oceanic Atlantic cod (*Gadus morhua* L.) on the Lofoten spawning grounds, Northern Norway. In: Davenport J, Burnell GM, Cross T, Emmerson M, McAllen R, Ramsay R, Dordrecht RE, editors. *Challenges to Marine Ecosystems: Proceedings of the 41st European Marine Biology Symposium*. Netherlands: Springer; 2008. p. 7–25.
28. Westgaard J-I, Fevolden S-E. Atlantic cod (*Gadus morhua* L.) in inner and outer coastal zones of northern Norway display divergent genetic signature at non-neutral loci. *Fish Res*. 2007;85(3):306–15.
29. Berg PR, Star B, Pampoulie C, Sodeland M, Barth JMI, Knutsen H, Jakobsen KS, Jentoft S. Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Sci Rep*. 2016;6:23246.
30. Bradbury IR, Hubert S, Higgins B, Borza T, Bowman S, Paterson IG, Snelgrove PVR, Morris CJ, Gregory RS, Hardie DC, et al. Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proc R Soc B Biol Sci*. 2010;277(1701):3725–34.
31. Hemmer-Hansen J, Nielsen EE, Therkildsen NO, Taylor MI, Ogdén R, Geffen AJ, Bekkevold D, Helyar S, Pampoulie C, Johansen T, et al. A genomic island linked to ecotype divergence in Atlantic cod. *Mol Ecol*. 2013;22(10):2653–67.
32. Kirubakaran TG, Grove H, Kent MP, Sandve SR, Baranski M, Nome T, De Rosa MC, Righino B, Johansen T, Otterå H, et al. Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Mol Ecol*. 2016;25(10):2130–43.
33. Moen T, Hayes B, Nilsen F, Delghandi M, Fjalestad KT, Fevolden SE, Berg PR, Lien S. Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. *BMC Genet*. 2008;9
34. Therkildsen NO, Hemmer-Hansen J, Als TD, Swain DP, Morgan MJ, Trippel EA, Palumbi SR, Meldrup D, Nielsen EE. Microevolution in time and space: SNP analysis of historical DNA reveals dynamic signatures of selection in Atlantic cod. *Mol Ecol*. 2013;22(9):2424–40.
35. Halldórsdóttir K, Árnason E. Whole-genome sequencing uncovers cryptic and hybrid species among Atlantic and Pacific cod-fish. In: bioRxiv; 2015.
36. Karlén BO, Klingan K, Emblem Å, Jørgensen TE, Jueterbock A, Furmanek T, Hoarau G, Johansen SD, Nordeide JT, Moum T. Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. *Mol Ecol*. 2013;22(20):5098–111.
37. Árnason E, Pálsson S. Mitochondrial cytochrome b DNA sequence variation of Atlantic cod *Gadus morhua*, from Norway. *Mol Ecol*. 1996;5(6):715–24.
38. Bradbury IR, Hubert S, Higgins B, Bowman S, Borza T, Paterson IG, Snelgrove PVR, Morris CJ, Gregory RS, Hardie D, et al. Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. *Evol Appl*. 2013;6(3):450–61.
39. Ruzzante DE, Wroblewski JS, Taggart CT, Smedbol RK, Cook D, Goddaard SV. Bay-scale population structure in coastal Atlantic cod in Labrador and Newfoundland, Canada. *J Fish Biol*. 2000;56(2):431–47.
40. Kovach AI, Breton TS, Berlinsky DL, Maceda L, Wirgin I. Fine-scale spatial and temporal genetic structure of Atlantic cod off the Atlantic coast of the USA. *Mar Ecol Prog Ser*. 2010;410:177–95.
41. Bonanomi S, Overgaard Therkildsen N, Retzel A, Berg Hedeholm R, Pedersen MW, Meldrup D, Pampoulie C, Hemmer-Hansen J, Grønkvær P, Nielsen EE. Historical DNA documents long-distance natal homing in marine fish. *Mol Ecol*. 2016;25(12):2727–34.
42. Therkildsen NO, Hemmer-Hansen J, Hedeholm RB, Wisz MS, Pampoulie C, Meldrup D, Bonanomi S, Retzel A, Olsen SM, Nielsen EE. Spatiotemporal SNP analysis reveals pronounced biocomplexity at the northern range margin of Atlantic cod *Gadus morhua*. *Evol Appl*. 2013;6(4):690–705.
43. Pampoulie C, Ruzzante DE, Chosson V, Jörundsdóttir TD, Taylor L, Thorsteinsson V, Daniëlsdóttir AK, Marteinsdóttir G. The genetic structure of Atlantic cod (*Gadus morhua*) around Iceland: insight from microsatellites, the *Pan I* locus, and tagging experiments. *Canadian Journal of Fisheries and Aquatic Sciences*. 2006;63(12):2660–74.
44. Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D. Population genomics of marine fishes: identifying adaptive variation in space and time. *Mol Ecol*. 2009;18
45. Sodeland M, Jorde PE, Lien S, Jentoft S, Berg PR, Grove H, Kent MP, Arnyasi M, Olsen EM, Knutsen H. "Islands of divergence" in the Atlantic cod genome represent polymorphic chromosomal rearrangements. *Genome Biol Evol*. 2016;8(4):1012–22.
46. Pampoulie C, Daniëlsdóttir AK, Thorsteinsson V, Hjörleifsson E, Marteinsdóttir G, Ruzzante DE. The composition of adult overwintering and juvenile aggregations of Atlantic cod (*Gadus morhua*) around Iceland using neutral and functional markers: a statistical challenge. *Can J Fish Aquat Sci*. 2012;69(2):307–20.
47. Ruzzante DE, Taggart CT, Cook D. A review of the evidence for genetic structure of cod (*Gadus morhua*) populations in the NW Atlantic and population affinities of larval cod off Newfoundland and the Gulf of St. Lawrence. *Fisheries Research*. 1999;43(1):79–97.
48. Hubert S, Higgins B, Borza T, Bowman S. Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*). *BMC Genomics*. 2010;11(1):191.
49. Berg PR, Jentoft S, Star B, Ring KH, Knutsen H, Lien S, Jakobsen KS, André C. Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). *Genome Biol Evol*. 2015;7(6):1644–63.
50. ICES. Cod (*Gadus morhua*) in subareas I and II (Northeast Arctic). In: *ICES Advice*. Vol. book 3; 2015.
51. ICES. Report of the Arctic fisheries working group (AFWG). In: *ICES CM 2015/ACOM:05*. Hamburg-Germany; 2015.
52. ICES. Cod (*Gadus morhua*) in Subareas I and II (Norwegian coastal waters cod). In: *ICES Advice*; 2015. p. 2015.
53. Dahle G, Jørstad KE, Rusaas HE, Otterå H. Genetic characterisation of broodstock collected from four Norwegian coastal cod (*Gadus morhua*) populations. *ICES J Mar Sci*. 2006;63:209–15.
54. Jørstad KE, Van Der Meer T, Dahle G, Gnalt ALA, Taggart JB, Glover K, Otterå H. Offspring performance of three different cod populations, measured under identical environmental conditions. *Aquaculture*. 2007; 272:5273.
55. Otterå H, Agnalt A-L, Jørstad KE. Differences in spawning time of captive Atlantic cod from four regions of Norway, kept under identical conditions. *ICES J Marine Science*. 2006;63(2):216–23.
56. Dahle G, Johansen T, Westgaard J-I, Aglen A, Glover KA. Genetic management of mixed-stock fisheries "real-time": The case of the largest remaining cod fishery operating in Atlantica in 2007–2017. *Fish Res*. 2018; 205:77–85.
57. Johansen T, Westgaard J-I, Seliussen BB, Nedreaas K, Dahle G, Glover KA, Kvalsund R, Aglen A. Handling editor: Stewart G: "real-time" genetic monitoring of a commercial fishery on the doorstep of an MPA reveals unique insights into the interaction between coastal and migratory forms of the Atlantic cod. *ICES J Mar Sci*. 2017; <https://doi.org/10.1093/icesjms/fsx224>.
58. Rollefson G. The otoliths of cod. In: *Fiskeridirektoratets skrifter, serie Havundersøkelser*, vol. 4; 1933. p. 1–14.
59. Berg E, Albert OT. Cod in fjords and coastal waters of North Norway: distribution and variation in length and maturity at age. *ICES J Mar Sci*. 2003;60(4):787–97.
60. Brooker AL, Cook D, Bentzen P, Wright JM, Doyle RW. Organization of microsatellites differs between mammals and cold-water teleost fishes. *Can J Fish Aquat Sci*. 1994;51(9):1959–66.
61. Miller KM, Le KD, Beacham TD. Development of tri- and tetranucleotide repeat microsatellite loci in Atlantic cod (*Gadus morhua*). *Mol Ecol*. 2000;9(2):238–9.
62. O'Reilly PT, Canino MF, Bailey KM, Bentzen P. Isolation of twenty low stutter di- and tetranucleotide microsatellites for population analyses of walleye Pollock and other gadoids. *J Fish Biol*. 2000;56(5):1074–86.
63. Stenvik J, Wesmajervi MS, Fjalestad KT, Damsgård B, Delghandi M. Development of 25 gene-associated microsatellite markers of Atlantic cod (*Gadus morhua* L.). *Mol Ecol Notes*. 2006;6(4):1105–7.
64. Wesmajervi MS, Tafese T, Stenvik J, Fjalestad KT, Damsgård B, Delghandi M. Eight new microsatellite markers in Atlantic cod (*Gadus morhua* L.) derived from an enriched genomic library. *Mol Ecol Notes*. 2007;7(1):138–40.
65. Dieringer D, Schlötterer C. MICROSATELLITE ANALYSER (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes*. 2003;3(1):167–9.
66. Peakall R, Smouse PE. GenA1Ex 6: genetic analysis in excel. Population genetic software for teaching and research. *Mol Ecol Notes*. 2006;6(1): 288–95.



67. Rousset F. GENEPOP'007: a complete re-implementation of the genepop software for windows and Linux. *Mol Ecol Resour.* 2008;8(1):103–6.
68. Waples RS, Do Chi LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour.* 2008;8(4):753–6.
69. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat.* 1979;6(2):65–70.
70. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: a workbench to detect molecular adaptation based on a  $F_{ST}$ -outlier method. *BMC Bioinformatics.* 2008;9(1):323.
71. Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinformatics Online.* 2005;1:47–50.
72. Kendall M, Gibbons JD. Rank correlation methods. 5th ed; 1976.
73. Valz PD, Thompson ME. Exact inference for Kendall's s and Spearman's rho. *J Comput Graph Stat.* 1994;3:459–72.
74. Team RDC. A language and environment for statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
75. Manni F, Guérard E, Heyer E. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Hum Biol.* 2004;76(2):173–90.
76. Monmonier MS. Maximum-difference barriers: an alternative numerical regionalization method. *Geogr Anal.* 1973;5(3):245–61.
77. Nei M, Tajima F, Tatenyo Y. Accuracy of estimated phylogenetic trees from molecular data II Gene frequency data. *J Mol Evol.* 1983;19(2):153–70.
78. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155(2):945–59.
79. Besnier F, Glover KA. ParallelStructure: a R package to distribute parallel runs of the population genetics program STRUCTURE on multi-Core computers. *PLoS One.* 2013;8(7):e70651.
80. Nordeide JT, Pettersen IH. Haemoglobin frequencies and vertebral numbers of cod (*Gadus morhua* L.) off northern Norway – test of a population structure hypothesis. *ICES J Mar Sci.* 1998;55(1):134–40.
81. Jakobsen T. Coastal cod in northern Norway. *Fish Res.* 1987;5(2):223–34.
82. Berg E, Sarvas T, Harbitz A, Fevolden SE, Salberg AB. Accuracy and precision in stock separation of north-East Arctic and Norwegian coastal cod by otoliths - comparing readings, image analyses and a genetic method. *Mar Freshw Res.* 2005;56(5):753–62.
83. Knutsen H, Olsen EM, Ciannelli L, Espeland SH, Knutsen JA, Simonsen JH, Skreslet S, Stenseth NC. Egg distribution, bottom topography and small-scale cod population structure in a coastal marine system. *Mar Ecol Prog Ser.* 2007;333:249–55.
84. Michalsen K, Johansen T, Subbey S, Beck A. Linking tagging technology and molecular genetics to gain insight in the spatial dynamics of two stocks of cod in Northeast Atlantic waters. *ICES J Mar Sci.* 2014;71(6):1417–32.
85. Skjæraasen JE, Meager JJ, Karlens Ø, Hutchings JA, Fernö A. Extreme spawning-site fidelity in Atlantic cod. *ICES J Mar Sci.* 2011;68(7):1472–7.
86. Ciannelli L, Knutsen H, Olsen EM, Espeland SH, Asplin L, Jelmert A, Knutsen JA, Stenseth NC. Small-scale genetic structure in a marine population in relation to water circulation and egg characteristics. *Ecology.* 2010;91(10):2918–30.
87. Nielsen EE, Hansen MM, Meldrup D. Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Mol Ecol.* 2006;15(11):3219–29.
88. Gagnaire PA, Broquet T, Aurelle D, Viard F, Souissi A, Bonhomme F, Arnaud-Haond S, Bierne N. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evol Appl.* 2015;8(8):769–86.
89. Knutsen H, André C, Jorde PE, Skogen MD, Thuróczy E, Stenseth NC. Transport of North Sea cod larvae into the Skagerrak coastal populations. *Proc Biol Sci.* 2004;271(1546):1337.
90. Berg PR. Genomic divergence in Atlantic cod populations. Oslo: University of Oslo; 2017.
91. Bigg GR, Cunningham CW, Ottersen G, Pogson GH, Wadley MR, Williamson P. Ice-age survival of Atlantic cod: agreement between palaeoecology models and genetics. *Proc R Soc B Biol Sci.* 2008;275(1631):163.
92. Carr SM, Marshall HD. Phylogeographic analysis of complete mtDNA genomes from walleye Pollock (*Gadus chalcogrammus* Pallas, 1811) shows an ancient origin of genetic biodiversity. *Mitochondrial DNA.* 2008;19(6):490–6.
93. Hutchings JA, Fraser DJ. The nature of fisheries- and farming-induced evolution. *Mol Ecol.* 2008;17
94. Nordeide JT, Folstad I. Is cod lekking or a promiscuous group spawner? *Fish Fish.* 2000;1(1):90–3.
95. Nielsen EE, Hemmer-Hansen J, Poulsen NA, Loeschcke V, Moen T, Johansen T, Mittelholzer C, Taranger G-L, Ogden R, Carvalho GR. Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evol Biol.* 2009;9(1):276.
96. Case RAJ, Hutchinson WF, Hauser L, Van Oosterhout C, Carvalho GR. Macro- and micro-geographic variation in pantophysin (Pan I) allele frequencies in NE Atlantic cod *Gadus morhua*. *Mar Ecol Prog Ser.* 2005;301
97. Ruzzante DE, Taggart CT, Cook DA. Nuclear DNA basis for shelf- and bank scale population structure in Northwest Atlantic cod (*Gadus morhua*): Labrador to George Bank. *Mol Ecol.* 1998;7
98. Jorde PE, Knutsen H, Espeland SH, Stenseth NC. Spatial scale of genetic structuring in coastal cod *Gadus morhua* and geographic extent of local populations. *Mar Ecol Prog Ser.* 2007;343:229–37.
99. Knutsen H, Olsen EM, Jorde PE, Espeland SH, André C, Stenseth NC. Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. *Mol Ecol.* 2011;20(4):768–83.
100. Barceló C, Ciannelli L, Olsen EM, Johannessen T, Knutsen H. Eight decades of sampling reveal a contemporary novel fish assemblage in coastal nursery habitats. *Glob Chang Biol.* 2016;22(3):1155–67.
101. Tørresen OK, Star B, Jentoft S, Reinart WB, Grove H, Miller JR, Walenz BP, Knight J, Ekholm JM, Peluso P, et al. An improved genome assembly uncovers prolific tandem repeats in Atlantic cod. *BMC Genomics.* 2017; 18(1):95.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

