



Distinct genetic clustering in the weakly differentiated polar cod, *Boreogadus saida* Lepechin, 1774 from East Siberian Sea to Svalbard

María Quintela¹ · Shripathi Bhat² · Kim Præbel² · Natalia Gordeeva^{3,4} · Gaute W. Seljestad⁵ ·
Tanja Hanebrekke⁶ · Alejandro Mateos-Rivera¹ · Frode Vikebø¹ · Daria Zelenina⁷ ·
Chi-Hing Christina Cheng⁸ · Torild Johansen⁶

Received: 15 October 2020 / Revised: 18 June 2021 / Accepted: 22 June 2021 / Published online: 3 July 2021
© The Author(s) 2021

Abstract

The cold-adapted polar cod *Boreogadus saida*, a key species in Arctic ecosystems, is vulnerable to global warming and ice retreat. In this study, 1257 individuals sampled in 17 locations within the latitudinal range of 75–81°N from Svalbard to East Siberian Sea were genotyped with a dedicated suite of 116 single-nucleotide polymorphic loci (SNP). The overall pattern of isolation by distance (IBD) found was driven by the two easternmost samples (East Siberian Sea and Laptev Sea), whereas no differentiation was registered in the area between the Kara Sea and Svalbard. Eleven SNP under strong linkage disequilibrium, nine of which could be annotated to chromosome 2 in Atlantic cod, defined two genetic groups of distinct size, with the major cluster containing seven-fold larger number of individuals than the minor. No underlying geographic basis was evident, as both clusters were detected throughout all sampling sites in relatively similar proportions (*i.e.* individuals in the minor cluster ranging between 4 and 19% on the location basis). Similarly, females and males were also evenly distributed between clusters and age groups. A differentiation was, however, found regarding size at age: individuals belonging to the major cluster were significantly longer in the second year. This study contributes to increasing the population genetic knowledge of this species and suggests that an appropriate management should be ensured to safeguard its diversity.

Keywords Polar cod · *Boreogadus saida* · Genetic structure · Conservation genetics · Management · Global warming

Introduction

Polar cod, *Boreogadus saida*, Lepechin, 1774 is a small gadid fish whose ability to synthesize antifreeze glycoproteins (Chen et al. 1997) enable survival in the freezing environments of the Arctic Ocean and surroundings seas. In the Atlantic Ocean, it occurs as far south as the Gulf of St Lawrence (Canada), whereas in the Pacific Ocean, it is found as far south as the Sea of Okhotsk and Norton Sound (Western Alaska). This species performs a crucial role in the Arctic food web where it is estimated to funnel 75% of lower trophic energy to upper marine and near-shore predators including birds, seals, beluga whales and eventually polar bears and humans (Ponomarenko 1968; Bradstreet et al. 1986; Crawford and Jorgenson 1996; Hop and Gjørseter 2013). Polar cod is a relatively short-lived species (maximum 7–8 years), reaching maturity at the age of 2–3 years for males and females, respectively (Craig et al. 1982). This species mainly spawns once due to its high post-spawning

✉ María Quintela
maria.quintela.sanchez@hi.no

- ¹ Institute of Marine Research, Nordnes, N-5817 Bergen, Norway
- ² Norwegian College of Fishery Science, The Arctic University of Norway (UiT), N-9019 Tromsø, Norway
- ³ Vavilov Institute of General Genetics, Russian Academy of Sciences, 117971 Moscow, Russia
- ⁴ Shirshov Institute of Oceanology, Russian Academy of Sciences, 119991 Moscow, Russia
- ⁵ Department of Biological Sciences, University of Bergen (UiB), N-5007 Bergen, Norway
- ⁶ Institute of Marine Research, Framcenteret, N-9296 Tromsø, Norway
- ⁷ Russian Federal Institute of Fisheries and Oceanography (VNIRO), 107140 Moscow, Russia
- ⁸ Department of Evolution, Ecology and Behavior, School of Integrative Biology, University of Illinois, Urbana, IL 61801, USA

mortality (Altukhov 1979; Hop et al. 1995; Nahrgang et al. 2016).

B. saida occupies a wide range of habitats encompassing semi-enclosed brackish bays, under landfast and pack ice, as well as throughout the water column from near-shore shallow water to hundreds of kilometres off-shore (Drolet et al. 1991; David et al. 2016; Robertis et al. 2016; Thorsteinson and Love 2016). It is sometimes found in very dense shoals (Crawford and Jorgenson 1996; Melnikov and Chernova 2013) and known to undertake extensive migrations (Ponomarenko 1968; Hop and Gjøsæter 2013). Polar cod aggregates in the areas of the Barents Sea where ice forms during late autumn. Hence, from September onwards, polar cod from the Kara Sea migrates to breed in the Barents Sea passing through all the straits of Novaya Zemlya and around its northern tip to the west (Manteifel 1943). It spawns at near-bottom depths in areas near the ice edge from November to February (Ponomarenko 1968, 2008) producing positively buoyant eggs (Spencer et al. 2020) that drift for two to four months before hatching (Graham and Hop 1995). The exact spawning grounds seem to be difficult to identify due to the ice cover; however, the discontinuous distribution of age 0-group individuals reported in some years suggests two different spawning sites (Eriksen et al. 2015): a western component around Svalbard and an eastern component along the western coast of Novaya Zemlya and near the Timan coast in the south-eastern Barents Sea (Ponomarenko 2000), respectively. A recent study applying a coupled ocean–sea ice and particle tracking model in Svalbard waters reported that particles released in the western fjords were mostly retained in the fjords and suggested that spawning in Svalbard area probably occurs on the southern and eastern sides and later in the south-eastern Barents Sea (Eriksen et al. 2020). However, the revision of Russian historical data (1935–2020) combined with literature published in English reported particularly large abundance of polar cod in the eastern Barents Sea and adjacent waters and revealed crucial knowledge gaps including the uncertainty as to what degree the species is dependent on sea ice, its genetic structure or the distribution of the age 0-group in the NE Barents Sea (Aune 2021).

The first attempt of assessing genetic population structure in this species, using random amplified polymorphic DNA (RAPDs), concluded the existence of a panmictic population in the North Atlantic Ocean (Fevolden et al. 1999). Later on, two lineages of mitochondrial DNA were revealed around Greenland waters, although no strong population differentiation was detected (Pálsson et al. 2009). A more recent study showed that polar cod in Svalbard and NE Greenland was distributed between fjord and oceanic populations, the latter one being a panmictic entity, whereas weak differentiation was observed amongst fjords (Madsen et al. 2016). Polar cod across the Pacific Arctic (Bering, Chukchi and

Beaufort Seas) was found to comprise a single panmictic population with high genetic diversity compared to other gadoids, as indicated across all 13 protein-coding genes in the mitogenome (Wilson et al. 2017). Small but significant differentiation in polar cod along Russian Arctic coast (from the Kara, Laptev and East Siberian seas) was found with seven microsatellite loci (Gordeeva and Mishin 2019). Lack of spatial population structure and isolation by distance in juveniles sampled in the region between the fjords off West Svalbard, the northern Sophia Basin and the Eurasian Basin of the Arctic Ocean was also reported by Maes et al. (2021) using a suite of nine microsatellites. Finally, a circumpolar study (Nelson 2020) conducted with the same set of nine microsatellite markers detected significant differentiation ($F_{ST}=0.01$, $p<0.01$) with a geographically underlying basis across the species' entire range. Four genetic groups were thus identified: (i) Canada East (from Resolute Bay to the Gulf of St. Lawrence), (ii) Canada West (Canadian Beaufort Sea and Amundsen Gulf), (iii) Europe (Greenland Sea, Iceland and the Laptev Sea) and (iv) USA (north Bering, Chukchi and western Beaufort seas). In this case, genetic divergence was thought to be influenced by a combination of physical distance, geophysical structure and oceanography, whereas very little genetic differentiation was detected within the identified groups (Nelson et al. 2020). In addition, a number of genomic resources have been made available in the last years (e.g. Breines et al. 2008; Árnason and Halldórsdóttir 2015; Wilson et al. 2018).

Growing concentrations of greenhouse gases warm the atmosphere and the oceans resulting in sea level rise due to thermal expansion of waters and melting of ice sheets and glaciers (Nicholls and Cazenave 2010). The Arctic amplification of global warming is provoking the retreat in the ice edge of the Arctic–Atlantic at unprecedented pace, thereby negatively affecting this cold-adapted species (Bouchard et al. 2016; Steiner, 2019) with narrow thermal plasticity (Drost et al. 2016; Laurel et al. 2018). Concurrently, boreal species extend their habitats into the northern Barents Sea (Fossheim et al. 2015) and sea ice retreat opens up vast areas for human activities such as petroleum, deep-sea mining and fishing, both adding to the pressures on these pristine Arctic ecosystems (see also Christiansen 2017). As a consequence, the abundance of polar cod, a keystone species of the ice-associated food web, has diminished (Divoky et al. 2015). The lower levels that polar cod stock experienced during the last years together with the fluctuating recruitment cannot be attributed to harvesting and temporally coincide with the rise of sea temperatures in the Barents Sea. A strong mechanistic link between survival in the early stages of the life cycle and changes in ice cover and temperature suggests an imminent collapse of recruitment if the observed ice-reduction and warming continue (Huserbråten et al. 2019). A simulation study indicated that the environmental conditions needed

to account for a successful recruitment to 0-group consist of high ice concentration in summer and low temperatures throughout the pelagic juvenile phase; therefore, the perturbations from the Arctic ocean climate typically found in the northern and eastern Barents Sea appear to be detrimental (Gjøsæter et al. 2020). In contrast, a 30-year data series on the length-at-age of polar cod cohorts (ages 0–4) analysed in relation to sea surface temperature, sea ice concentration, prey biomasses, predator indices and length-at-age the previous year using multiple linear regression suggested that retreating sea ice had positive effects on the growth of polar cod in the Barents Sea as only length-at-age for age 1 showed a positive significant relationship with prey biomass. High length-at-age was significantly associated with low sea ice concentration and high length-at-age the previous year (Dupont et al. 2020).

This study aims to characterize the population genetic structure of polar cod in the geographic region ranging from Svalbard to the East Siberian Sea using a panel of specifically developed SNP markers. The deeper genetic knowledge this new panel provides can give useful guidance for conservation and sustainable management of this species.

Materials and methods

Sampling

A total of 1316 individuals was sampled from 17 geographic locations ranging from 75 to 81°N, from East Siberian Sea to Svalbard (Fig. 1). Full biometric data including length, weight, age, stage and sex were available for 596 individuals. Age was determined by counting winter zones from whole otoliths submerged in water and under binoculars as described in Gjøsæter and Ajiad (1994). All samples were collected by pelagic trawl with the exception of ESS, LS and KS (see Table 1), which were sampled using Bongo net. In

all cases, sampling was framed within surveys aiming to assess fish stocks.

DNA isolation, SNP panel selection and genotyping

DNA was extracted from ethanol-preserved fin clips using OMEGA E-Z 96™ Tissue DNA kit (Promega). A panel of single-nucleotide polymorphism (SNP) markers was developed from a restriction site-associated DNA sequencing (RADseq) data set of 270 polar cod individuals from oceanic and fjord populations from the North East Atlantic Ocean. RADseq (Baird 2008; Emerson et al. 2010; Hohenlohe et al. 2010) can identify and score thousands of genetic markers, randomly distributed across the target genome and has become an important tool for ecological population genomics (e.g. Davey et al. 2011; Bergey et al. 2013; Ogden, 2013). The initial dataset used for SNP panel selection consisted of 141,178 SNP markers that fulfilled the following filtering criteria: > 1% major allele frequency (MAF), > 65% presence per population, only a single SNP per sequence tag and removal of suspected paralogues. SNPs were then ranked based on the F_{ST} computed between oceanic and fjord populations, and the 613 SNPs displaying the highest values were selected as a panel for screening two Svalbard samples (63 individuals), which were also part of the initial dataset. Out of those 613 SNPs, 316 were detected in the two Svalbard samples and F_{ST} based on these 316 SNPs was computed between the samples. The top 192 SNPs showing significant differentiation were used to develop the final SNP panel used in all the downstream analysis. A subset of 174 of those SNPs could be distributed on seven multiplex reactions and amplified on a Sequenom MassARRAY iPLEX Platform, as described by Gabriel et al. (2009).

The initial dataset consisted of 1316 individuals genotyped at 174 loci. However, the final number of usable markers was reduced due to (i) non-amplification (9 loci), (ii) > 25% missing genotypes (25 loci) and (iii) monomorphic

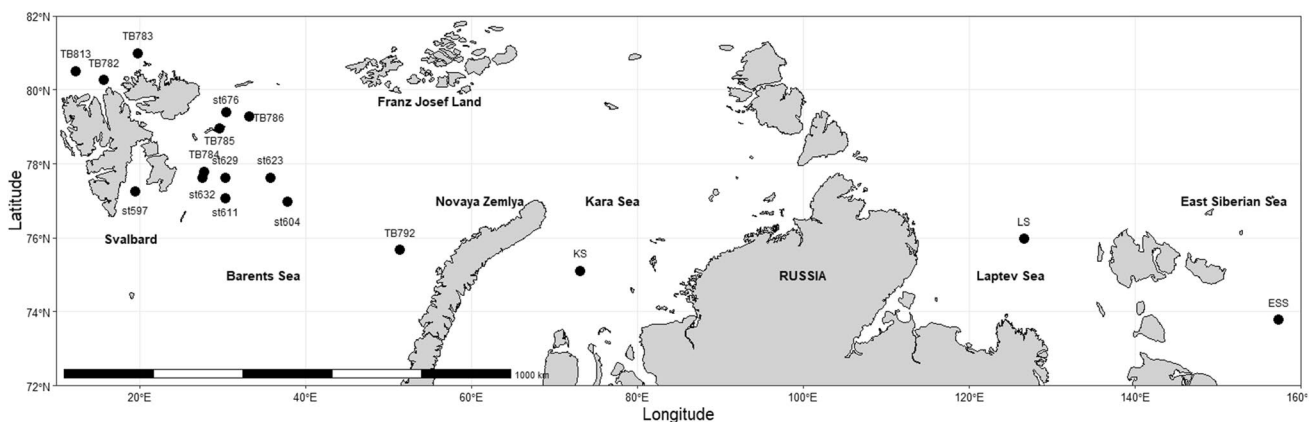


Fig. 1 Map of the sampling sites. Each of the white and black bars in the bottom left scale represents 200 km

Table 1 Summary statistics

Sample	Geographic area	Sampling year	No ind	Latitude	Longitude	H_o	uH_e	F_{IS}
ESS	East Siberian Sea	2017	41	73.77 N	157.30 E	0.138 ± 0.012	0.141 ± 0.012	0.002 ± 0.013
LS	Laptev Sea	2017	59	75.96 N	126.63 E	0.133 ± 0.010	0.130 ± 0.009	-0.029 ± 0.009
KS	Kara Sea	2014, 2017	80	75.10 N	73.19 E	0.134 ± 0.009	0.136 ± 0.009	0.013 ± 0.013
TB792	Barents Sea	2012	78	75.66 N	51.46 E	0.130 ± 0.009	0.135 ± 0.009	0.031 ± 0.016
TB783	Svalbard area	2018	100	80.98 N	19.87 E	0.129 ± 0.010	0.131 ± 0.009	0.030 ± 0.015
TB813	Svalbard area	2019	75	80.49 N	12.33 E	0.121 ± 0.010	0.122 ± 0.009	0.011 ± 0.015
TB782	Svalbard area	2018	99	80.25 N	15.69 E	0.132 ± 0.009	0.130 ± 0.009	-0.022 ± 0.007
st676	Svalbard area	2018	78	79.39 N	30.48 E	0.132 ± 0.009	0.136 ± 0.009	0.018 ± 0.013
TB786	Svalbard area	2018	86	79.27 N	33.28 E	0.127 ± 0.008	0.130 ± 0.009	0.015 ± 0.013
TB785	Svalbard area	2018	96	78.94 N	29.66 E	0.124 ± 0.010	0.128 ± 0.009	0.032 ± 0.012
TB784	Svalbard area	2018	97	77.78 N	27.86 E	0.134 ± 0.008	0.136 ± 0.009	-0.001 ± 0.008
st629	Svalbard area	2018	78	77.62 N	30.35 E	0.136 ± 0.009	0.135 ± 0.009	-0.014 ± 0.010
st623	Svalbard area	2018	48	77.61 N	35.80 E	0.124 ± 0.010	0.126 ± 0.009	0.015 ± 0.015
st632	Svalbard area	2018	46	77.60 N	27.60 E	0.120 ± 0.011	0.119 ± 0.009	0.002 ± 0.016
st597	Svalbard area	2018	94	77.25 N	19.58 E	0.134 ± 0.010	0.137 ± 0.009	0.023 ± 0.014
st611	Svalbard area	2018	50	77.06 N	30.44 E	0.114 ± 0.010	0.112 ± 0.009	-0.005 ± 0.014
st604	Svalbard area	2018	52	76.96 N	37.91 E	0.136 ± 0.010	0.133 ± 0.009	-0.021 ± 0.011

Sample name and geographic area; year of collection; number of individuals (No ind); geographic coordinates; observed heterozygosity, H_o (mean ± SE); unbiased expected heterozygosity, uH_e (mean ± SE) and inbreeding coefficient, F_{IS} (mean ± SE)

markers (13 loci). Thus, 59 individuals showing > 20% missing data were purged, hence leaving 1257 individuals genotyped at 127 polymorphic SNP loci (Table 1).

Population genetic analyses

The genotype frequency per locus and its direction (heterozygote deficit or excess) against the expectations under Hardy–Weinberg equilibrium (HWE) as well as linkage disequilibrium (LD) between pairs of loci were estimated using the programme GENEPOP v7 (Rousset 2008). HWE and LD were examined using the Markov chain parameters: 10,000 steps of dememorization, 1000 batches and 10,000 iterations per batch, and significance was assessed after the post hoc sequential Bonferroni correction (Holm 1979). Observed (H_o) and unbiased expected heterozygosity (uH_e), and also the inbreeding coefficient (F_{IS}), were computed for each sample with GenAEx 6.503 (Peakall and Smouse 2006).

Genetic relationships amongst samples were examined using principal component analysis (PCA) (Pearson 1901; Hotelling 1933; Jolliffe 2002), as implemented in *adeigenet* (Jombart 2008). Population genetic structure was examined using the analysis of molecular variance (AMOVA) and by estimating pairwise F_{ST} (Weir and Cockerham 1984) using ARLEQUIN v.3.5.1.2 (Excoffier et al. 2005) with 10,000 permutations. The Bayesian clustering approach implemented in STRUCTURE v. 2.3.4 (Pritchard et al. 2000) was used to identify genetic groups under a model assuming admixture and correlated allele frequencies without using

population information. The analyses were conducted using the programme ParallelStructure (Besnier and Glover 2013) to reduce the computational time. Ten runs with a burn-in period consisting of 100,000 replications and a run length of 1,000,000 Markov chain Monte Carlo (MCMC) iterations were performed for $K = 1$ to $K = 10$ clusters. To determine the number of genetic groups in the data, STRUCTURE output was analysed using two approaches. Firstly, the ad hoc summary statistic ΔK of Evanno et al. (2005) was calculated. Secondly, StructureSelector (Li and Liu 2018) was used to estimate Puechmaille (2016) statistics (MedMed, MedMean, MaxMed and MaxMean), which have been described as more accurate than the previously used methods to determine the best fit number of clusters, for both even and uneven sampling data. Finally, the ten runs for the selected K s were averaged with CLUMPP v.1.1.1 (Jakobsson and Rosenberg 2007) using the FullSearch algorithm and the G' pairwise matrix similarity statistic and graphically displayed using barplots. Genetic clustering using sampling sites as prior was also examined and visualized using discriminant analysis of principal components (DAPC) (Jombart et al. 2010) in *adeigenet* (Jombart 2008). A number of up to 116 principal components (PCs) were tested to determine the optimal number to avoid overfitting of the data and creating artificially large separation between groups (Jombart and Collins 2015; Miller et al. 2020). The trade-off between power of discrimination and overfitting was measured using the a -score, which is the difference between the proportion of successful reassignment of the analysis (observed

discrimination) and values obtained using random groups (random discrimination), *i.e.* the proportion of successful reassignment corrected for the number of retained PCs.

Loci departing from neutrality and therefore reflecting possible selective responses or linkage disequilibrium with genes under divergent selection (Lewontin and Krakauer 1973) were statistically identified using two outlier analytic approaches: BayeScan (Foll and Gaggiotti 2008) and LOSITAN (Antao et al. 2008). BayeScan was run by setting sample size to 10,000 and the thinning interval to 50 as suggested by Foll et al. (2008). The loci with a posterior probability above 0.99 were retained as outliers, corresponding to a Bayes Factor (BF) > 2, *i.e.* “decisive selection” (Foll and Gaggiotti 2006). In LOSITAN, a neutral distribution of F_{ST} with 100,000 iterations was simulated, with a forced mean F_{ST} at a significance level of 0.05 under an infinite allele model.

Allele frequency shifts at outlier loci are expected to be driven by selective responses towards strong ecological gradients leading to local adaptation, either due to directly associated genes or through hitchhiking (linkage) with associated genes (Gagnaire 2015). Major allele frequencies (MAF) per sample were displayed as appropriate.

Finally, the relationship between genetic distance, *i.e.* $F_{ST}/(1-F_{ST})$, and geographic distance was examined using Mantel (1967) test to investigate if the correlation conformed the expectations of “Isolation by Distance” (IBD), a pattern of increasing genetic differentiation with geographic distance as a result of restricted gene flow and drift (Wright 1943; Slatkin 1993; Rousset 1997). Putative patterns IBD were investigated not only using the full set of SNPs but also with loci decomposed into candidates to selection (*i.e.* SNPs flagged by any of the outlier analyses) and neutrals (*i.e.* SNPs that failed to be identified by any of the outlier detection approaches), both for the total set of samples as well as after excluding the farthest ones (East Siberian Sea, ESS and Laptev Sea, LS). The matrix of geographic distance was calculated using the R-package *marmap* (Pante and Simon-Bouhet 2013) by computing the shortest pairwise distances by water, *i.e.* avoiding land masses. Two-tailed Mantel (1967) tests were conducted using PASSaGE v2 (Rosenberg and Anderson 2011) and significance was assessed via 10,000 permutations.

Results

Genetic variation and overall population structure

The SNP loci and their corresponding flanking regions together with the raw data are available in Online Resource 1 (Tables S1 and S2, respectively). A significant deficit of heterozygotes was detected for 11 of the 127 loci, which

were thus discarded for further analyses. Table 1 shows the summary statistics across the 116 remaining loci for each sample. The observed (H_o) and unbiased expected heterozygosity (uH_e) ranged from 0.114 to 0.138 and 0.112 to 0.141, respectively (Table 1). In both cases, the highest values were reported for sample ESS, taken in the East Siberian Sea.

Principal components analyses (PCA) revealed two main groups (Fig. 2), which were confirmed by the first level of division in STRUCTURE (Fig. 3a). Each genotype was assigned to the two inferred clusters based on the individual proportion of membership (q_i) using a threshold of ancestry of $q > 0.8$, which allowed the identification of a major cluster hosting 1098 individuals and a minor cluster hosting 151, thus leaving out eight admixed individuals. The number of individuals belonging to the minor cluster ranged from 2 (ESS, st611) to 17 (TB784) per sample. These low numbers hampered to conduct any further geographically explicit analysis due to the lack of statistical power. The ratio between females and males of adult fish was around 1 in both clusters.

The subsequent analyses were conducted separately for the total set of individuals as well as for the major cluster. However, given the consistency between both patterns, and unless otherwise stated, only the results corresponding to the full dataset will be presented here, whereas the results for the major cluster can be found in the Online Resource 2. The AMOVA revealed an overall significant genetic structure ($F_{ST} = 0.0045$, $p < 0.0001$), with 99.5% of the variation hosted within populations. The pairwise F_{ST} comparisons showed that the easternmost sample (ESS) was significantly different from all other ones (F_{ST} ranging from 0.013 to 0.038, see also Table S4 for major allele frequency per sample) followed by the Laptev Sea, LS (range from 0 to 0.007), which differed from six of the samples (although four of them would lose significance after Bonferroni correction) (Table 2). No differentiation was observed amongst the samples collected between the Kara Sea and Svalbard.

STRUCTURE analysis after the removal of the individuals from the minor cluster revealed $K = 2$ as the most likely number of clusters (Fig. 3b). The ancestry to cluster on a sampling site basis confirmed the results from the pairwise F_{ST} . These patterns were further corroborated by the discriminant analysis of principal components using the full set of individuals (Fig. S1 in Online Resource 2), where the 36 PCs retained allowed to identify a first axis, which accounted for 43.8% of the variation and isolated sample ESS. Axis 2, responsible for 11.4% of the variation, did not fully separate sample LS, which overlapped to a great extent with the remaining.

LOSITAN identified seven SNPs deviating from neutrality, two of which were confirmed by BayeScan (Table 3). Patterns of IBD were recorded for the full set of

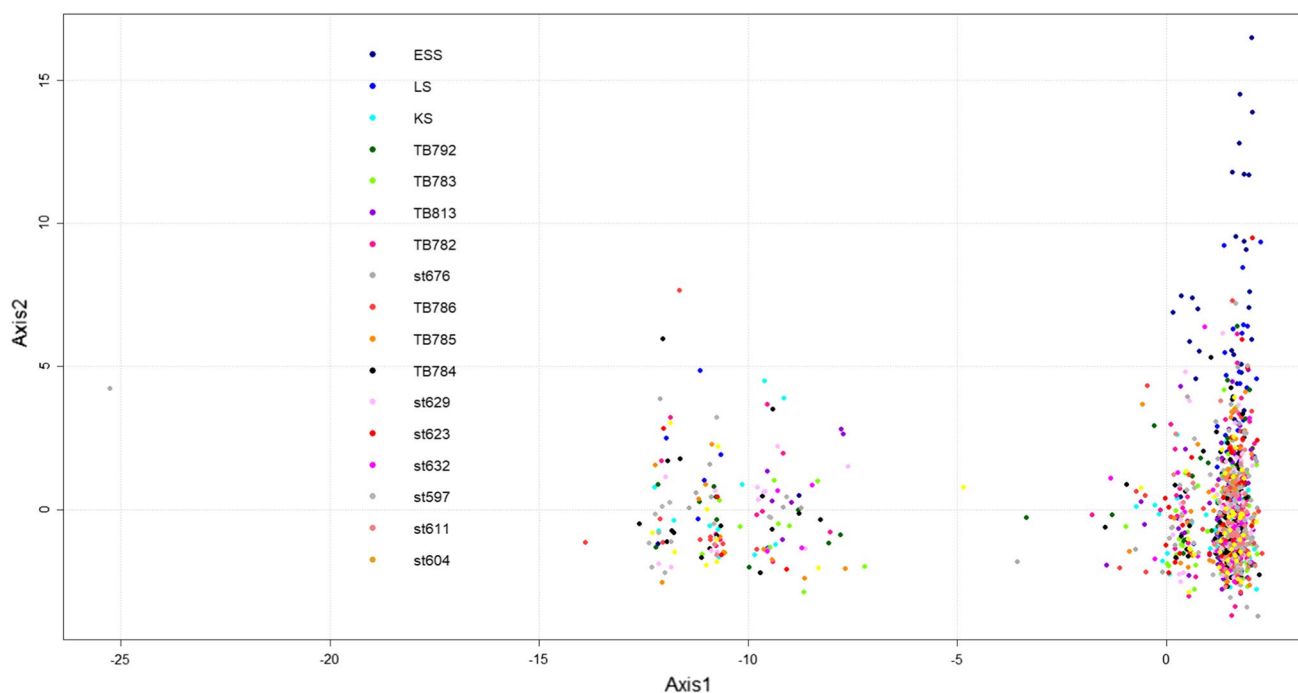


Fig. 2 Principal component analysis (PCA) based on 116 SNP markers. The 1257 individuals are plotted according to their coordinates on the biplot of principal component 1 (PC1) versus PC2 and form two different clusters. Colours depict different sampling locations

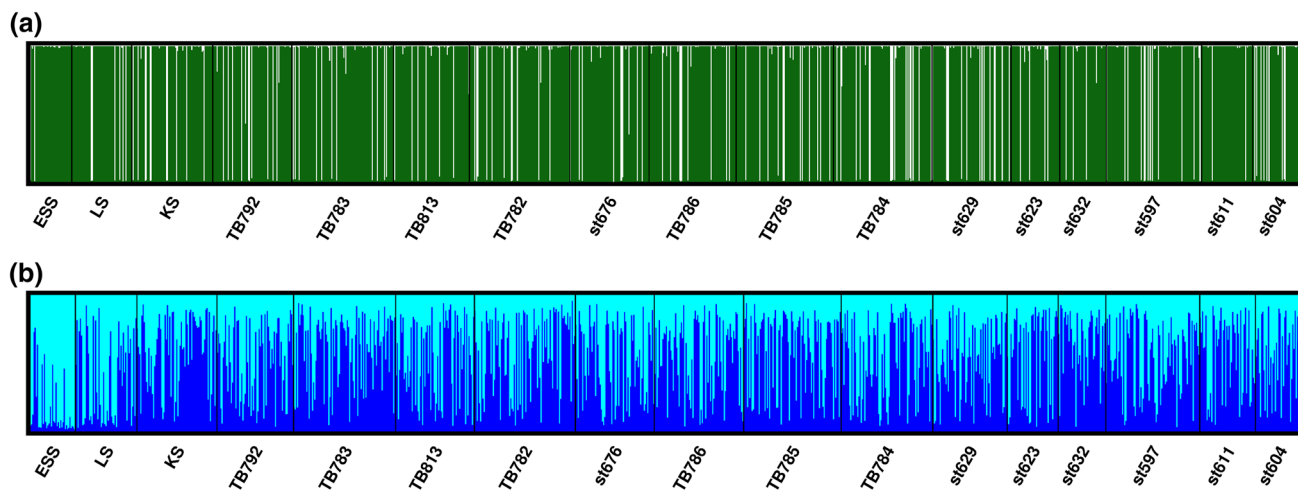


Fig. 3 Barplot representing the proportion of individuals' ancestry to cluster at $K=2$ as inferred from Bayesian clustering in STRUCTURE for **a** the full set of individuals and **b** the 1098 individuals belonging to the major cluster

samples (and individuals) irrespective of the type of markers used (neutrals, candidate to selection or altogether) albeit the strength of the correlation was weaker at neutral loci (Table 4). The significance of IBD was due to the easternmost samples (ESS, LS), as patterns faded away when these samples were excluded from the analyses (see Fig. S3 in Online Resource 2 for illustration).

Study of the genetic clusters identified by principal component analysis

The strongest genetic signal in the dataset was revealed by PCA (Fig. 2) and STRUCTURE (Fig. 3a) as two significantly distinct groups of individuals ($F_{ST}=0.139$, $p<0.0001$) were found in every sampling site, *i.e.* the proportion of individuals belonging to the minor cluster ranged between 4 and

Table 2 Heatmap of pairwise F_{ST} for collections including all the individuals (below diagonal) and for those containing only individuals belonging to the “major” cluster (above diagonal)

	ESS	LS	KS	TB792	TB783	TB813	TB782	st676	TB786	TB785	TB784	st629	st623	st632	st597	st611	st604
ESS	*	0.0163	0.0377	0.0209	0.0404	0.0305	0.0372	0.0294	0.0333	0.0259	0.0375	0.0314	0.0192	0.0333	0.0321	0.0305	0.0345
LS	0.0134	*	0.0033	0.0000	0.0094	0.0041	0.0088	0.0012	0.0045	0.0000	0.0040	0.0000	0.0000	0.0020	0.0054	0.0006	0.0065
KS	0.0340	0.0022	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0017	0.0000	0.0009	0.0000	0.0000	0.0000
TB792	0.0198	0.0000	0.0000	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TB783	0.0377	0.0072	0.0000	0.0000	*	0.0000	0.0000	0.0012	0.0000	0.0000	0.0008	0.0000	0.0000	0.0000	0.0015	0.0002	0.0001
TB813	0.0274	0.0027	0.0000	0.0000	0.0000	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TB782	0.0327	0.0064	0.0000	0.0000	0.0000	0.0000	*	0.0000	0.0010	0.0000	0.0010	0.0000	0.0000	0.0000	0.0010	0.0000	0.0024
st676	0.0255	0.0010	0.0000	0.0000	0.0000	0.0000	0.0000	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TB786	0.0298	0.0024	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	*	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000	0.0000
TB785	0.0265	0.0001	0.0000	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000	*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003
TB784	0.0337	0.0036	0.0000	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	*	0.0000	0.0000	0.0002	0.0000	0.0000	0.0004
st629	0.0319	0.0015	0.0010	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	*	0.0000	0.0000	0.0000	0.0000	0.0000
st623	0.0188	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	*	0.0000	0.0000	0.0000	0.0000
st632	0.0292	0.0014	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0007	*	0.0011	0.0000	0.0000
st597	0.0299	0.0047	0.0000	0.0000	0.0012	0.0000	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	*	0.0000	0.0014
st611	0.0278	0.0004	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0025	0.0000	0.0009	0.0006	*	0.0000
st604	0.0321	0.0052	0.0000	0.0000	0.0000	0.0000	0.0007	0.0000	0.0006	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	*

Boldface type depicts F_{ST} values significantly different from zero after 10,000 permutations following Bonferroni correction for multiple testing. Greener colours indicate lowest differentiation (F_{ST} closer to zero) whereas yellow towards red colours represent larger differentiation

Table 3 Results of outlier detection procedures

Locus	LOSITAN_P (Simul F_{ST} < sample F_{ST})	BayeScan_ log ₁₀ (PO)	<i>Gadus morhua</i> genome assembly
OMLE01074069	<i>l</i>	1000	Chr. 6
OMLE01009325	<i>l</i>	0.77	–
OMLE01104816	<i>l</i>	0.36	Chr. 16
OMLE01016577	<i>l</i>	– 1.04	–
OMLE01007365	<i>l</i>	– 0.65	Chr. 15
OMLE01121548	<i>l</i>	3.22	Chr. 11
OMLE01054446	<i>l</i>	– 0.63	Chr. 20

Values highlighted in italic depict candidate loci to positive selection identified by LOSITAN and BayeScan, respectively. Five of the loci could be annotated to different chromosomes (Chr.) in the Atlantic cod genome assembly

19% per site. The individual’s axis position for PC1 was strongly correlated with STRUCTURE ancestry coefficient Q ($r=0.983$, $p<0.0001$). To rule out that the clustering pattern was driven by the SNP cross-amplification of any taxonomically related species that could be morphologically confounded with polar cod, a subset of individuals was bar-coded. The resulting COI fragment (654 bp) confirmed that samples belonged to *B. saida* (see Online Resource 3 for a detailed description).

The F_{ST} per locus computed between clusters reached statistical significance in 15 out of 116 SNPs (13%); ten of

Table 4 Results of Mantel tests for isolation by distance (IBD)

	Loci	Matrices	Mantel’s r	p value
17 samples	All loci	All $F_{ST} \times$ GeoDist	0.785	0.0001
	Neutrals	$N-F_{ST} \times$ GeoDist	0.666	0.0017
	Outliers	$Sel-F_{ST} \times$ GeoDist	0.840	0.0001
15 samples	All loci	All $F_{ST} \times$ GeoDist	– 0.036	0.725
	Neutrals	$N-F_{ST} \times$ GeoDist	– 0.001	0.991
	Outliers	$Sel-F_{ST} \times$ GeoDist	– 0.025	0.853

Tests have been performed for the full set of samples ($N=17$) and after excluding the furthestmost sites (East Siberian Sea, ESS and Laptev Sea, LS). Three sets of loci have been tested: the total 116 SNPs, the set of neutral markers (*i.e.* those not showing deviation the outlier approaches) and the outliers (*i.e.* those deviating from neutrality under LOSITAN or BayeScan). Boldface type depicts significant correlations

which showed $F_{ST} > 0.6$ (Table 5). Amongst the loci revealing significant structure, the eleven SNPs displaying the largest differentiation in major allele frequency (MAF) between clusters were found to be linked (Fisher’s method, $p < 0.05$), all of them being annotated to chromosome 2 in Atlantic cod, *Gadus morhua* genome assembly with the exception of two for which no match could be found. Furthermore, the allelic patterns showed by nine of those eleven loci (*i.e.* OMLE01011103, OMLE01085950, OMLE01122866, OMLE01075874, OMLE01065648, OMLE01120889, OMLE01059372, OMLE01120901, OMLE01025315)

Table 5 Summary of SNPs showing significant genetic differentiation between clusters: major allele frequency per cluster (MAF), F_{ST} and loadings on the first axis per locus (LD1), respectively

Locus	MAF major cluster	MAF minor cluster	F_{ST} ($p < 0.05$)	LD1	<i>Gadus morhua</i> genome assembly
<i>OMLE01011103</i>	1.000	0.483	0.818	2.570	Chr. 2
<i>OMLE01085950</i>	0.998	0.490	0.803	2.538	–
<i>OMLE01122866</i>	1.000	0.493	0.811	2.497	Chr. 2
<i>OMLE01075874</i>	0.999	0.495	0.841	2.348	Chr. 2
<i>OMLE01065648</i>	1.000	0.500	0.808	2.363	Chr. 2
<i>OMLE01120889</i>	0.982	0.486	0.733	1.652	Chr. 2
<i>OMLE01059372</i>	0.998	0.513	0.786	2.504	Chr. 2
<i>OMLE01120901</i>	0.970	0.500	0.668	0.725	Chr. 2
<i>OMLE01025315</i>	0.991	0.548	0.724	2.147	Chr. 2
<i>OMLE01122584</i>	0.989	0.688	0.603	0.050	–
<i>OMLE01083294</i>	0.735	0.897	0.065	0.025	Chr. 2
OMLE01135302	0.804	0.733	0.013	0.039	Chr. 5
OMLE01075521	0.881	0.821	0.014	0.059	Chr. 14
OMLE01078117	0.945	0.905	0.013	0.132	Chr. 1
OMLE01074820	0.954	0.980	0.006	0.072	Chr. 10

Eight of the SNPs showing the highest F_{ST} values were annotated to chromosome 2 in *Gadus morhua* genome assembly. The eleven SNPs highlighted in boldface type were found to be significantly linked to each other in the full dataset across all samples (Fisher's method, $p < 0.05$), nine of them being annotated to chromosome 2 in cod. The allelic patterns for the nine SNPs in italics correspond to a homozygote in the major cluster and a heterozygote in the minor one

were depicted by a homozygote in the major cluster and conversely to a heterozygote in the minor one.

As biometric information was available for only half of the individuals belonging to the minor cluster, an identical number of individuals ($N = 74$) was randomly sampled from the major cluster to compare length-at-age under even conditions. The distribution of individuals per age was very similar in both cases, but only age classes 1 and 2 had large enough sampling size ($n \geq 25$) to conduct statistical analysis. No differentiation in length was registered when the individuals were one year old ($t = -0.686$, $p = 0.496$); however, individuals from the major cluster reached a larger size at the second year of age ($t = -2.169$, $p = 0.036$), *i.e.* average length of 22.1 cm versus the 16.0 cm in the minor cluster (Fig. 4).

Discussion

The current study, investigating the population genetic structure of polar cod from Svalbard to the East Siberian Sea using a SNP panel, revealed two main results. Firstly, the existence of two significantly distinct genetic groups: a major cluster that contained 87.35% of the individuals and a minor cluster hosting 12%. Geography does not seem to account for the clustering as both genetic groups were represented in every sampling site, *i.e.* individuals belonging to the minor cluster were distributed across all samples in

proportions ranging from 4 to 19%. Secondly, the pattern of isolation by distance found was driven by the two easternmost samples (ESS and LS), whereas a lack of differentiation was reported in the geographic area ranging from the Kara Sea to Svalbard.

Population structure in polar cod

The suite of 116 polymorphic loci revealed a significant pattern of IBD driven by the easternmost Russian samples (East Siberian Sea and Laptev Sea), whilst no spatial population structure was detected in the region between the Kara Sea and Svalbard. This result agrees with the absence of spatial population structure and isolation by distance reported by Maes et al. (2021) using nine microsatellites, which suggests ongoing gene flow in the region between the fjords off West Svalbard, the northern Sophia Basin and the Eurasian Basin of the Arctic Ocean. Likewise, previous studies showed small, but significant, differentiation along the Russian Arctic coast (Gordeeva and Mishin 2019) and lack of genetic structure within the Bering, Chukchi and Beaufort seas of Alaska, despite a consistent isolation by distance pattern (Wilson et al. 2017). The lack of genetic differentiation between the Kara Sea and Svalbard, in agreement with the genetic profile of a panmictic population, typically occurs in large and genetically diverse populations (Frankham 1996). Polar cod is the most widespread and abundant species in the Arctic Ocean, with a stock estimated of almost

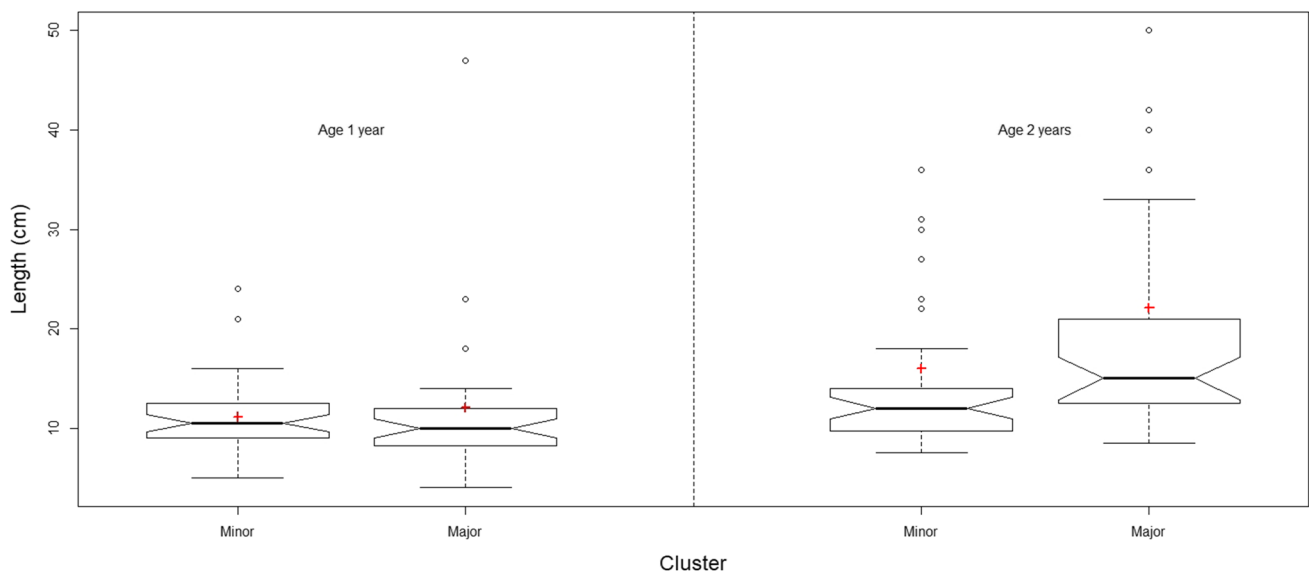


Fig. 4 Box and whiskers plot for length (cm) of the individuals belonging to minor and major cluster at ages 1 and 2 years, respectively. The box represents the interquartile range with the notch showing the median and the red cross depicting the mean. The upper whisker represents the largest value within 1.5 times interquartile

range above 75th percentile, whereas the lower whisker represents the smallest value within 1.5 times interquartile range below 25th percentile. Outliers depict those values > 1.5 and < 3 times the interquartile range beyond either end of the box

2 million tonnes (only in the Barents Sea) in some years (Gjørseter et al. 2020), and has been shown to display large genetic diversity (Wilson et al. 2017). This absence of structure might also result from the significant gene flow due to passive dispersal during early stages and migrations during the adult stage (Ponomarenko 1968; Hop and Gjørseter 2013; Gordeeva and Mishin 2019; Nelson et al. 2020), *i.e.* advection by sea-ice drift has been hypothesized to play an important role for the genetic connectivity of circum-Arctic populations (David et al. 2016). However, the results by Huserbråten et al. (2019) indicate that dispersal features segregate ichthyoplankton drift from the two main spawning grounds of *B. saida* in the Barents Sea, towards the southeast as well as a northward retreat of one of two clearly defined spawning assemblages, possibly in response to warming.

Putative origin of polar cod clustering

The individuals examined in this study displayed a two-clustering pattern irrespective of their geographic explicit locations, as it has been shown in Atlantic cod, *G. morhua* (Berg 2016; Kirubakaran 2016); three-spine stickleback, *Gasterosteus aculeatus* (Jones 2012); Atlantic silverside, *Menidia menidia* (Therkildsen et al. 2019); or lesser sand eel, *Ammodytes tobianus* (Jiménez-Mena et al. 2020). Clustering could involve phenotypically distinct morphotypes as in ballan wrasse, *Labrus bergylta*, where spotty and plain morphs displaying genetic differentiation (Quintela 2016) occur in sympatry (Villegas-Ríos et al. 2013b). The duration and the

timing of the spawning season coincide despite their differences in life-history strategies (Villegas-Ríos et al. 2013a); therefore, body colour has been invoked to play a pivotal role in facilitating sympatric speciation in the absence of impermeable geographic barriers (Casas et al. 2021). In contrast, for other fish species, time of spawning seems to be a crucial evolutionary driver. Genomic data collected from spawning aggregations of Pacific herring (*Clupea pallasii*) across 1600 km of coastline showed that reproductive timing drives population structure in these pelagic fish (Petrou, 2021). On Atlantic herring (*C. harengus*), genomic regions have also been identified in association with spawning season (Martínez Barrio, 2016; Lamichhaney 2017), some of which are located on a supergene maintained by a chromosomal inversion (Pettersson 2019). In both species, spawning phenology was highly correlated with polymorphisms in several genes, in particular SYNE2, which is probably within a chromosomal rearrangement in Pacific herring and known to be associated with spawn timing in Atlantic herring (Lamichhaney et al. 2017; Petrou et al. 2021).

A group of eleven SNPs under strong LD, which were mostly annotated to chromosome 2 (LG02) in Atlantic cod genome assembly, accounted for the clustering pattern in polar cod. This linkage group 2 has been shown to be highly divergent between spring and winter spawners in Atlantic cod within the Gulf of Maine (Barney et al. 2017). Genomic data would be needed to test the hypothesis if the different spawning times for polar cod in Svalbard area and the south-eastern Barents Sea proposed after a particle tracking

model study (Eriksen et al. 2020) align with this clustering pattern. Likewise, rearrangements in LG2, 7 and 12 have been identified within coastal and offshore samples of Atlantic cod on the Norwegian coast (Sodeland 2016; Johansen 2020). In Atlantic cod, cryptic sympatric populations have been described in coastal areas. Thus, in a 14-year period, stable coexistence of two genetically divergent Atlantic cod ecotypes were reported along the Norwegian Skagerrak coast: a “fjord” ecotype that dominates in numbers deep inside fjords, whereas the “North Sea” ecotype was the only type found in offshore North Sea (Knutzen 2018). Likewise, polar cod sampled inside fjords in Svalbard and north-east Greenland has been reported to be different from polar cod sampled over the shelf. The observed genetic variation did not adhere to an isolation by distance pattern and it was speculated that *B. saida* populations inhabiting fjords may have become reproductively isolated from shelf-dwelling *B. saida* during the last post-glacial recolonization (Madsen et al. 2016), whereas the particle tracking study of Eriksen et al. (2020) revealed that particles released in the fjords of western Svalbard were mostly retained in the fjords, thus suggesting that spawning in these areas might lead to isolation. Besides, polar cod is suggested to display several intraspecific forms differing in proportions of body, size and position of fins and in coloration (Chernova 2018). Two ecotypes (coastal and oceanic) of polar cod have also been described in the Russian Arctic, which differ in age of maturation, body shape and colour of larvae and adult fish (Moskalenko 1964). The oceanic one undertakes long migrations, its distribution is associated with ice cover and it reaches the highest latitudes (Moskalenko 1964). Likewise, juveniles sampled in the Russian Arctic seas showed a trend of decreasing body size with increasing latitude (Mishin et al. 2018). Whilst we cannot make the differentiation between such ecotypes in the present study, the average length recorded at the second year of age was larger for the individuals belonging to the major cluster (22.1 vs. 16 cm in the minor cluster). However, sampling sizes urge for a cautious interpretation of this result. Likewise, a differential length in ecotypes was also reported between the “North Sea” and the “fjord” ecotype of Atlantic cod that coexist along the Skagerrak coast, *i.e.* the “North Sea” type grows faster during the summer and is, on average, 2 cm longer in autumn, when they reach 9 and 11 cm, respectively (Knutzen et al. 2018). The settling period in cod occurs within the size range of 40–110 mm, and it was shown that the aforementioned differences in growth rate were found to be small until the juveniles reached ~40 mm and disappeared again in juveniles larger than 110 mm, thus highlighting how this vital rate may change rapidly at ontogenetic transition points leading to different phenotypic trajectories in co-existing ecotypes (Jørgensen et al. 2020). Likewise, migratory NE Arctic cod (NEAC) and stationary Norwegian (NCC) and

Russian coastal cod have been reported to display both genetic and phenotypic (growth, maturation, counts of vertebrae) differences (Nordeide et al. 2011; Johansen et al. 2020). Thus, for example, BB homozygotes for Pantophysin, Pan I (typical of NEAC), were of significantly longer size than the Pan I AA homozygotes (typical of NCC) (Fevolden et al. 2012).

Chromosomal inversions underlie all four supergenes (*i.e.* sets of genes that are inherited as a single marker and encode complex phenotypes through their joint action) recently identified in Atlantic cod (Matschiner et al. 2021), which have been shown to originate separately between 0.40 and 1.66 million years ago and be linked to migratory lifestyle and environmental adaptations such as to salinity. Such a phenomenon cannot be excluded in the phylogenetically close *B. saida*. Thus, future studies may benefit from expanding the genomic coverage of SNPs to investigate whether structural genomic variants, ecotype specific haplotypes and/or other genomic signatures can shed light on the occurrence of the two discrete clusters observed here.

Management implications

A notable effect of global warming is the poleward shift of the distribution range in marine species (Poloczanska 2016). Over the past 30 years, at least 19 fish species of major economic importance (e.g. cod, herring, saithe, sprat, hake and sole) have experienced changes in distribution in some parts of their range (Baudron 2020). Polar cod is one of the 21 commercially targeted species in the Barents Sea although its economic importance has been relatively low. In the last years, Russia was harvesting this species at very low levels (Popov and Zeller 2018; Wienerroither et al. 2011) and nowadays it is not directly targeted (Aune et al. 2021) although taken as by-catch in the shrimp fishery (Jacques et al. 2019). However, in the light of climate change, the facilitated access to Arctic waters due to the ice retreat might lead to increased exploitation of polar cod, either as a target species or as bycatch in pelagic fisheries and bottom trawling (Christiansen et al. 2014). Additionally, changed physiological constraints or increased competition from other fish species such as the subarctic capelin, *Mallotus villosus* (Hop and Gjørseter 2013), or through predation from boreal species like Atlantic cod and haddock, *Melanogrammus aeglefinus*, expanding their range into the polar cod habitat (Renaud et al. 2012; Christiansen 2016; Andrews et al. 2019) might represent further challenges for the persistence of its populations in the future.

The sustainable management of this species would benefit from the use of comprehensive molecular tools to disentangle the underlying basis of the observed clustering reported in this study. Furthermore, a verification of the existence of the alleged spawning grounds (around Svalbard and along

Novaya Zemlya) is needed, as well as a deeper understanding of the genetic structure of fjord versus oceanic polar cod to clarify if they represent two different management units as it occurs in Atlantic cod.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00300-021-02911-7>.

Acknowledgements We would like to thank Georg Skaret, Jostein Røttingen, Elena Eriksen, the staff of the Polar branch of VNIRO and the crew on board all research vessels in Norway and Russia for collecting the samples. The crew of RV Helmer Hanssen is thanked for providing samples from NE Greenland and Svalbard waters for SNPs development. Georg Skaret and Stine Karlson are acknowledged for performing the age determination of the fish. We are grateful to Bjørghild Breistein, Geir Dahle, Per Erik Jorde and François Besnier for inspiring discussions on an early stage of this manuscript. The following entities are thanked for providing financial support: Equinor through the Grant No. 4590100459; VIGG RAS and IORAS Russian Government basic research programmes for 2019–2021 (# 0112-2019-0001 and # 0149-2018-0008); the Institute of Marine Research through funding from the Norwegian Ministry for Industry and Fisheries (NFD) and the projects “Stock complexes in the Barents Sea” (the Barents Sea program) and “Polar cod” (“Økologiske prosesser”) programme; the UiT–The Arctic University of Norway through the TUNU Programme. The Research Council of Norway is thanked for financial support through the project “Arctic ecosystem impact assessment of oil in ice under climate change” (ACTION - RCN no. 314449/E40). Great thanks are also due to reviewer Sharon Wildes and to the two anonymous referees for their insightful comments to the first version of the manuscript.

Author Contributions TJ, FV, KP, DZ, MQ conceived and designed the research. TJ, NG, DZ conducted the data collection. KP, SB, C-HCC contributed the analytical tools (SNPs). GWS, TH, AMR conducted the laboratory work. MQ performed the statistical analysis and wrote the first draft of the manuscript. All authors contributed, read and approved the final version of the manuscript.

Funding Open access funding provided by Institute Of Marine Research. The funding for the present research has been provided by the following institutions: Equinor through the Grant No. 4590100459; VIGG RAS and IORAS Russian Government basic research programmes for 2019–2021 (# 0112-2019-0001 and # 0149-2018-0008); the Institute of Marine Research through funding from the Norwegian Ministry for Industry and Fisheries (NFD) and the projects “Stock complexes in the Barents Sea” (the Barents Sea program) and “Polar cod” (“Økologiske prosesser”) programme; the UiT–The Arctic University of Norway through the TUNU Programme. The Research Council of Norway is thanked for financial support through the project “Arctic ecosystem impact assessment of oil in ice under climate change” (ACTION - RCN no. 314449/E40).

Declarations

Conflict of interest The authors declare that no competing interest exists, as well as that there is no financial support or relationships that may pose any kind of conflict. Likewise, the authors declare that contributed to the text, agreed with its content and approved it for submission.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long

as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Altukhov KA (1979) About reproduction and development of polar cod *Boreogadus saida* (Lepechin) in the White Sea. *Ichthyology* 19:874–882 (in Russian)
- Andrews AJ, Christiansen JS, Bhat S, Lynghammar A, Westgaard J-I, Pampoulie C, Præbel K (2019) Boreal marine fauna from the Barents Sea disperse to Arctic Northeast Greenland. *Sci Rep* 9:5799. <https://doi.org/10.1038/s41598-019-42097-x>
- Antao T, Lopes A, Lopes B, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a F_{ST} -outlier method. *BMC Bioinform* 9:323. <https://doi.org/10.1186/1471-2105-9-323>
- Árnason E, Halldórsdóttir K (2015) Codweb: Whole-genome sequencing uncovers extensive reticulations fueling adaptation among Atlantic, Arctic, and Pacific gadids. *Sci Adv* 5:034926. <https://doi.org/10.1126/sciadv.aat8788>
- Aune M et al (2021) Distribution and ecology of polar cod (*Boreogadus saida*) in the eastern Barents Sea: a review of historical literature. *Mar Environ Res* 166:105262. <https://doi.org/10.1016/j.marenres.2021.105262>
- Baird NA et al (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376. <https://doi.org/10.1371/journal.pone.0003376>
- Barney BT, Munkholm C, Walt DR, Palumbi SR (2017) Highly localized divergence within supergenes in Atlantic cod (*Gadus morhua*) within the Gulf of Maine. *BMC Genom* 18:271–271. <https://doi.org/10.1186/s12864-017-3660-3>
- Baudron AR et al (2020) Changing fish distributions challenge the effective management of European fisheries. *Ecography* 43:494–505. <https://doi.org/10.1111/ecog.04864>
- Berg PR et al (2016) Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Sci Rep* 6:23246. <https://doi.org/10.1038/srep23246>
- Bergey C, Pozzi L, Disotell T, Burrell A (2013) A new method for genome-wide marker development and genotyping holds great promise for molecular primatology. *Int J Primatol*. <https://doi.org/10.1007/s10764-013-9663-2>
- Besnier F, Glover KA (2013) ParallelStructure: a R package to distribute parallel runs of the population genetics program STRUCTURE on multi-core computers. *PLoS ONE* 8:e70651. <https://doi.org/10.1371/journal.pone.0070651>
- Bouchard C, Mollard S, Suzuki K, Robert D, Fortier L (2016) Contrasting the early life histories of sympatric Arctic gadids *Boreogadus saida* and *Arctogadus glacialis* in the Canadian Beaufort Sea. *Polar Biol* 39:1005–1022. <https://doi.org/10.1007/s00300-014-1617-4>
- Bradstreet MSW, Finley KJ, Sekarak AD, Griffiths WB, Evans CR, Fabjan MF, Stallard HE (1986) Aspects of the biology of Arctic Cod (*Boreogadus saida*) and its importance in arctic marine food chains. Ministère des pêches et des océans, Ottawa, Canada

- Breines R, Ursvik A, Nymark M, Johansen SD, Coucheron DH (2008) Complete mitochondrial genome sequences of the Arctic Ocean codfishes *Arctogadus glacialis* and *Boreogadus saida* reveal oriL and tRNA gene duplications. *Polar Biol* 31:1245–1252. <https://doi.org/10.1007/s00300-008-0463-7>
- Casas L, Saenz-Agudelo P, Villegas-Ríos D, Irigoien X, Saborido-Rey F (2021) Genomic landscape of geographically structured colour polymorphism in a temperate marine fish. *Mol Ecol* 30:1281–1296. <https://doi.org/10.1111/mec.15805>
- Chen L, DeVries AL, Cheng C-HC (1997) Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. *Proc Natl Acad Sci* 94:3817. <https://doi.org/10.1073/pnas.94.8.3817>
- Chernova N (2018) Arctic cod in the Russian Arctic: new data, with notes on intraspecific forms. *JAMB* 7:28–37. <https://doi.org/10.15406/jamb.2018.07.00180>
- Christiansen JS (2017) No future for Euro-Arctic ocean fishes? *Mar Ecol Prog Ser* 575:217–227. <https://doi.org/10.3354/meps12192>
- Christiansen JS et al (2016) Novel biodiversity baselines outpace models of fish distribution in Arctic waters. *Sci Nat* 103:8. <https://doi.org/10.1007/s00114-016-1332-9>
- Christiansen JS, Mecklenburg CW, Karamushko OV (2014) Arctic marine fishes and their fisheries in light of global change. *Glob Change Biol* 20:352–359. <https://doi.org/10.1111/gcb.12395>
- Craig PC, Griffiths WB, Haldorson L, McElderry H (1982) Ecological studies of Arctic cod (*Boreogadus saida*) in Beaufort Sea coastal waters, Alaska. *Can J Fish Aquat Sci* 39:395–406. <https://doi.org/10.1139/f82-057>
- Crawford RE, Jorgenson JK (1996) Quantitative studies of Arctic cod, *Boreogadus saida* schools: Important energy stores in the Arctic food web. *Arctic* 49:181–193. <https://doi.org/10.14430/arctic1196>
- Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* 12:499–510. <https://doi.org/10.1038/nrg3012>
- David C, Lange B, Krumpen T, Schaafsma F, van Franeker JA, Flores H (2016) Under-ice distribution of polar cod *Boreogadus saida* in the central Arctic Ocean and their association with sea-ice habitat properties. *Polar Biol* 39:981–994. <https://doi.org/10.1007/s00300-015-1774-0>
- Divoky GJ, Lukacs PM, Druckenmiller ML (2015) Effects of recent decreases in arctic sea ice on an ice-associated marine bird. *Prog Oceanogr* 136:151–161. <https://doi.org/10.1016/j.pocean.2015.05.010>
- Drolet R, Fortier L, Ponton D, Gilbert M (1991) Production of fish larvae and their prey in subarctic southeastern Hudson Bay. *Mar Ecol Prog Ser* 77:105–118
- Drost HE, Lo M, Carmack EC, Farrell AP (2016) Acclimation potential of Arctic cod (*Boreogadus saida*) from the rapidly warming Arctic Ocean. *J Exp Biol* 219:3114. <https://doi.org/10.1242/jeb.140194>
- Dupont N, Durant JM, Langangen Ø, Gjøsæter H, Stige LC (2020) Sea ice, temperature, and prey effects on annual variations in mean lengths of a key Arctic fish, *Boreogadus saida*, in the Barents Sea. *ICES J Mar Sci* 77:1796–1805. <https://doi.org/10.1093/icesjms/fsaa040>
- Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, Holzapfel CM (2010) Resolving postglacial phylogeography using high-throughput sequencing. *Proc Natl Acad Sci USA* 107:16196–16200. <https://doi.org/10.1073/pnas.1006538107>
- Eriksen E, Huserbråten M, Gjøsæter H, Vikebø F, Albretsen J (2020) Polar cod egg and larval drift patterns in the Svalbard archipelago. *Polar Biol* 43:1029–1042. <https://doi.org/10.1007/s00300-019-02549-6>
- Eriksen E, Ingvaldsen RB, Nedreaas K, Prozorkevich D (2015) The effect of recent warming on polar cod and beaked redfish juveniles in the Barents Sea. *Reg Studies Mar Sci* 2:105–112. <https://doi.org/10.1016/j.rsma.2015.09.001>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1:47–50. <https://doi.org/10.1177/117693430500100003>
- Fevolden S-E, Martinez I, Christiansen JS (1999) RAPD and scnDNA analyses of polar cod, *Boreogadus saida* (Pisces, Galidae), in the north Atlantic. *Sarsia* 84:99–103. <https://doi.org/10.1080/00364827.1999.10420437>
- Fevolden SE, Westgaard JI, Pedersen T, Præbel K (2012) 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* 468:267–278. <https://doi.org/10.3354/meps09990>
- Foll M, Beaumont MA, Gaggiotti O (2008) An approximate Bayesian computation approach to overcome biases that arise when using Amplified Fragment Length Polymorphism markers to study population structure. *Genetics* 179:927–939. <https://doi.org/10.1534/genetics.107.084541>
- Foll M, Gaggiotti O (2006) Identifying the environmental factors that determine the genetic structure of populations. *Genetics* 174:875–891. <https://doi.org/10.1534/genetics.106.059451>
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180:977–993. <https://doi.org/10.1534/genetics.108.092221>
- Fosshem M, Primicerio R, Johannesen E, Ingvaldsen RB, Aschan MM, Dolgov AV (2015) Recent warming leads to a rapid borealization of fish communities in the Arctic. *Nat Clim Change* 5:673–677. <https://doi.org/10.1038/nclimate2647>
- Frankham R (1996) Relationship of genetic variation to population size in wildlife. *Conserv Biol* 10:1500–1508. <https://doi.org/10.1046/j.1523-1739.1996.10061500.x>
- Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the Sequenom MassARRAY iPLEX Platform. *Curr Protoc Hum Genet*. <https://doi.org/10.1002/0471142905.hg0212s60>
- Gagnaire P-A et al (2015) Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evol Appl* 8:769–786. <https://doi.org/10.1111/eva.12288>
- Gjøsæter H, Ajiad AM (1994) Growth of polar cod, *Boreogadus saida* (Lepechin), in the Barents Sea. *ICES J Mar Sci* 51:115–120. <https://doi.org/10.1006/jmsc.1994.1011>
- Gjøsæter H, Huserbråten M, Vikebø F, Eriksen E (2020) Key processes regulating the early life history of Barents Sea polar cod. *Polar Biol* 43:1015–1027. <https://doi.org/10.1007/s00300-020-02656-9>
- Gordeeva NV, Mishin AV (2019) Population genetic diversity of Arctic cod (*Boreogadus saida*) of Russian Arctic Seas. *J Ichthyol* 59:246–254. <https://doi.org/10.1134/S0032945219020073>
- Graham M, Hop H (1995) Aspects of reproduction and larval biology of Arctic cod (*Boreogadus saida*). *Arctic* 48:130–135. <https://doi.org/10.14430/arctic1234>
- Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet* 6:e1000862–e1000862. <https://doi.org/10.1371/journal.pgen.1000862>
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* 6:65–70
- Hop H, Gjøsæter H (2013) Polar cod (*Boreogadus saida*) and capelin (*Mallotus villosus*) as key species in marine food webs of the

- Arctic and the Barents Sea. *Mar Biol Res* 9:878–894. <https://doi.org/10.1080/17451000.2013.775458>
- Hop H, Trudeau VL, Graham M (1995) Spawning energetics of Arctic cod (*Boreogadus saida*) in relation to seasonal development of the ovary and plasma sex steroid levels. *Can J Fish Aquat Sci* 52:541–550. <https://doi.org/10.1139/f95-055>
- Hottelling H (1933) Analysis of a complex of statistical variables into principal components. *J Educ Psychol* 24:498–520
- Huserbråten MBO, Eriksen E, Gjøsaeter H, Vikebø F (2019) Polar cod in jeopardy under the retreating Arctic sea ice. *Commun Biol* 2:407. <https://doi.org/10.1038/s42003-019-0649-2>
- Jacques N, Herrmann B, Larsen RB, Sistiaga M, Brčić J, Gökçe G, Brinkhof J (2019) Can a large-mesh sieve panel replace or supplement the Nordmøre grid for bycatch mitigation in the north-east Atlantic deep-water shrimp fishery? *Fish Res* 219:105324. <https://doi.org/10.1016/j.fishres.2019.105324>
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Jiménez-Mena B, Le Moan A, Christensen A, van Deurs M, Mosegaard H, Hemmer-Hansen J, Bekkevold D (2020) Weak genetic structure despite strong genomic signal in lesser sandeel in the North Sea. *Evol Appl* 13:376–387. <https://doi.org/10.1111/eva.12875>
- Johansen T et al (2020) Genomic analysis reveals neutral and adaptive patterns that challenge the current management regime for East Atlantic cod *Gadus morhua* L. *Evol Appl* 13:2673–2688. <https://doi.org/10.1111/eva.13070>
- Jolliffe IT (2002) *Principal component analysis*. Springer, New York
- Jombart T (2008) *adegen*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. <https://doi.org/10.1093/bioinformatics/btm129>
- Jombart T, Collins C (2015) A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0.
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11:94. <https://doi.org/10.1186/1471-2156-11-94>
- Jones FC et al (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61. <https://doi.org/10.1038/nature10944>
- Jørgensen KEM, Neuheimer AB, Jorde PE, Knutsen H, Grønkvær P (2020) Settlement processes induce differences in daily growth rates between two co-existing ecotypes of juvenile cod *Gadus morhua*. *Mar Ecol Prog Ser* 650:175–189. <https://doi.org/10.3354/meps13433>
- Kirubakaran TG et al (2016) Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Mol Ecol* 25:2130–2143. <https://doi.org/10.1111/mec.13592>
- Knutsen H et al (2018) Stable coexistence of genetically divergent Atlantic cod ecotypes at multiple spatial scales. *Evol Appl* 11:1527–1539. <https://doi.org/10.1111/eva.12640>
- Lamichhaney S et al (2017) Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean. *Proc Natl Acad Sci* 114:E3452. <https://doi.org/10.1073/pnas.1617728114>
- Laurel BJ, Copeman LA, Spencer M, Iseri P (2018) Comparative effects of temperature on rates of development and survival of eggs and yolk-sac larvae of Arctic cod (*Boreogadus saida*) and walleye pollock (*Gadus chalcogrammus*). *ICES J Mar Sci* 75:2403–2412. <https://doi.org/10.1093/icesjms/fsy042>
- Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74:175
- Li Y-L, Liu J-X (2018) StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol Ecol Resour* 18:176–177. <https://doi.org/10.1111/1755-0998.12719>
- Madsen ML, John Nelson R, Fevolden S-E, Christiansen JS, Præbel K (2016) Population genetic analysis of Euro-Arctic polar cod *Boreogadus saida* suggests fjord and oceanic structuring. *Polar Biol* 39:1357–1357. <https://doi.org/10.1007/s00300-015-1812-y>
- Maes SM, Christiansen H, Mark FC, Lucassen M, Van de Putte A, Volckaert FAM, Flores H (2021) High gene flow in polar cod (*Boreogadus saida*) from West-Svalbard and the Eurasian Basin. *J Fish Biol*. <https://doi.org/10.1111/jfb.14697>
- Manteifel BP (1943) Arctic cod and its fishery, edn. Arkhangelsk Publishing
- Mantel N (1967) The detection of disease of clustering and a generalized regression approach. *Can Res* 27:209–220
- Martínez Barrio A et al (2016) The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. *Elife* 5:e12081. <https://doi.org/10.7554/eLife.12081.001>
- Matschiner M et al. (2021) Origin and fate of supergenes in Atlantic cod. *bioRxiv:2021.2002.2028.433253*. <https://doi.org/10.1101/2021.02.28.433253>
- Melnikov IA, Chernova NV (2013) Characteristics of under-ice swarming of polar cod *Boreogadus saida* (Gadidae) in the Central Arctic Ocean. *J Ichthyol* 53:7–15. <https://doi.org/10.1134/S0032945213010086>
- Miller JM, Cullingham CI, Peery RM (2020) The influence of a priori grouping on inference of genetic clusters: simulation study and literature review of the DAPC method. *Heredity* 125:269–280. <https://doi.org/10.1038/s41437-020-0348-2>
- Mishin AV, Evseenko SA, Bolshakov DV, Bolshakova YY (2018) Ichthyoplankton of Russian Arctic Seas: 1 Polar cod *Boreogadus saida*. *J Ichthyol* 58:710–716. <https://doi.org/10.1134/S0032945218050156>
- Moskalenko BK (1964) Biology of the Arctic cod (*Boreogadus saida*, Lepechin, 1774). *Vopr Ikhtologii* 4:433–443
- Nahrgang J, Dubourg P, Frantzen M, Storch D, Dahlke F, Meador JP (2016) Early life stages of an arctic keystone species (*Boreogadus saida*) show high sensitivity to a water-soluble fraction of crude oil. *Environ Pollut* 218:605–614. <https://doi.org/10.1016/j.envpol.2016.07.044>
- Nelson RJ et al (2020) Circumpolar genetic population structure of polar cod, *Boreogadus saida*. *Polar Biol*. <https://doi.org/10.1007/s00300-020-02660-z>
- Nicholls RJ, Cazenave A (2010) Sea-level rise and its impact on coastal zones. *Science* 328:1517. <https://doi.org/10.1126/science.1185782>
- Nordeide J, Johansen S, Jørgensen T, Karlsen B, Moum T (2011) Population connectivity among migratory and stationary cod *Gadus morhua* in the North-East Atlantic—a review of 80 years of study. *Mar Ecol Prog Ser* 435:269–283. <https://doi.org/10.3354/meps09232>
- Ogden R et al (2013) Sturgeon conservation genomics: SNP discovery and validation using RAD sequencing. *Mol Ecol* 22:3112–3123. <https://doi.org/10.1111/mec.12234>
- Pálsson S, Källman T, Paulsen J, Arnason E (2009) An assessment of mitochondrial variation in Arctic gadoids. *Polar Biol* 32:471–479. <https://doi.org/10.1007/s00300-008-0542-9>
- Pante E, Simon-Bouhet B (2013) marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE* 8:e73051–e73051. <https://doi.org/10.1371/journal.pone.0073051>
- Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol*

- Notes 6:288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Pearson K (1901) On lines and planes of closest fit to systems of points in space. *Phil Mag* 6:559–572
- Petrou EL et al (2021) Functional genetic diversity in an exploited marine species and its relevance to fisheries management. *Proc R Soc B* 288:20202398. <https://doi.org/10.1098/rspb.2020.2398>
- Pettersson ME et al (2019) A chromosome-level assembly of the Atlantic herring genome—detection of a supergene and other signals of selection. *Genome Res*. <https://doi.org/10.1101/gr.253435.119>
- Poloczanska ES et al (2016) Responses of marine organisms to climate change across oceans. *Front Mar Sci*. <https://doi.org/10.3389/fmars.2016.00062>
- Ponomarenko VP (1968) Migration of cod in the Soviet Sector of Arctic. *Trudy PINRO*:505–512
- Ponomarenko VP (2000) Distribution of pre-spawning Arctic cod *Boreogadus saida* near Timan coast of the Barents Sea. VNIRO, Moscow
- Ponomarenko VP (2008) Characteristics of feeding of Arctic cod (*Boreogadus saida*) in the Barents Sea. *Vopr Rybolov* 9:330–343 (in Russian)
- Popov S, Zeller D (2018) Reconstructed Russian fisheries catches in the Barents Sea: 1950–2014. *Front Mar Sci*. <https://doi.org/10.3389/fmars.2018.00266>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Puechmaille SJ (2016) The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Mol Ecol Resour* 16:608–627. <https://doi.org/10.1111/1755-0998.12512>
- Quintela M et al (2016) Is the ballan wrasse (*Labrus bergylta*) two species? Genetic analysis reveals within-species divergence associated with plain and spotted morphotype frequencies. *Integr Zool* 11:162–172. <https://doi.org/10.1111/1749-4877.12186>
- Renaud PE, Berge J, Varpe Ø, Lønne OJ, Nahrang J, Ottesen C, Hallanger I (2012) Is the poleward expansion by Atlantic cod and haddock threatening native polar cod, *Boreogadus saida*? *Polar Biol* 35:401–412. <https://doi.org/10.1007/s00300-011-1085-z>
- Robertis A, Taylor K, Wilson C, Farley E (2016) Abundance and distribution of Arctic cod (*Boreogadus saida*) and other pelagic fishes over the US continental shelf of the Northern Bering and Chukchi Seas. *Deep Sea Res*. <https://doi.org/10.1016/j.dsr2.2016.03.002>
- Rosenberg MS, Anderson CD (2011) PASSaGE: pattern analysis, spatial statistics and geographic exegesis. Version 2. *Methods Ecol Evol* 2:229–232. <https://doi.org/10.1111/j.2041-210X.2010.00081.x>
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145:1219–1228
- Rousset F (2008) GENEPOP'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* 8:103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47:264–279. <https://doi.org/10.1111/j.1558-5646.1993.tb01215.x>
- Sodeland M et al (2016) “Islands of Divergence” in the Atlantic cod genome represent polymorphic chromosomal rearrangements. *Genome Biol Evol* 8:1012–1022. <https://doi.org/10.1093/gbe/evw057>
- Spencer ML, Vestfals CD, Mueter FJ, Laurel BJ (2020) Ontogenetic changes in the buoyancy and salinity tolerance of eggs and larvae of polar cod (*Boreogadus saida*) and other gadids. *Polar Biol*. <https://doi.org/10.1007/s00300-020-02620-7>
- Steiner NS et al (2019) Impacts of the changing ocean-sea ice system on the key forage fish Arctic cod (*Boreogadus saida*) and subsistence fisheries in the Western Canadian Arctic—Evaluating linked Climate, Ecosystem and Economic (CEE) Models. *Front Mar Sci*. <https://doi.org/10.3389/fmars.2019.00179>
- Therkildsen NO, Wilder AP, Conover DO, Munch SB, Baumann H, Palumbi SR (2019) Contrasting genomic shifts underlie parallel phenotypic evolution in response to fishing. *Science* 365:487. <https://doi.org/10.1126/science.aaw7271>
- Thorsteinson LK, Love MS (2016). Alaska Arctic Marine Fish Ecology Catalog. <https://doi.org/10.3133/sir20165038>
- Villegas-Ríos D, Alonso-Fernández A, Domínguez-Petit R, Saborido-Rey F (2013a) Intraspecific variability in reproductive patterns in the temperate hermaphrodite fish, *Labrus bergylta*. *Mar Freshw Res* 64:1156–1168. <https://doi.org/10.1071/MF12362>
- Villegas-Ríos D, Alonso-Fernández A, Fabeiro M, Bañón R, Saborido-Rey F (2013b) Demographic variation between colour patterns in a temperate protogynous hermaphrodite, the ballan wrasse *Labrus bergylta*. *PLoS ONE* 8:e71591. <https://doi.org/10.1371/journal.pone.0071591>
- Weir BS, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370. <https://doi.org/10.2307/2408641>
- Wienerroither R et al. (2011) Atlas of the Barents Sea Fishes vol 1.
- Wilson R, Menning D, Wedemeyer K, Talbot S (2018) A transcriptome resource for the Arctic cod (*Boreogadus saida*). *Mar Genomics*. <https://doi.org/10.1016/j.margen.2018.03.003>
- Wilson RE, Sage GK, Sonsthagen SA, Gravley MC, Menning DM, Talbot SL (2017) Genomics of Arctic Cod. US Dept. of the Interior, Bureau of Ocean Energy Management, Alaska OCS Region, Anchorage, AK
- Wright S (1943) Isolation by distance. *Genetics* 28:114–138

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