Canadian Science Publishing

ARTICLE

The making of a genetic cline: introgression of oceanic genes into coastal cod populations in the Northeast Atlantic

Per Erik Jorde, Mats B.O. Huserbråten, Bjørghild B. Seliussen, Mari S. Myksvoll, Frode B. Vikebø, Geir Dahle, Asgeir Aglen, and Torild Johansen

Abstract: Coastal Atlantic cod (*Gadus morhua*) in the Northeast Atlantic has seen a continuous decline since the industrialization of the coastal fishery, and management needs to address the spatial and temporal complexities of coexisting cod stocks. Toward that end, genetic analyses and oceanographic modelling of coastal and oceanic cod larval drift patterns were combined to elucidate the mechanisms responsible for an observed genetic cline over a >1500 km stretch along the coast of Norway. The results indicate that the north–south cline in coastal cod represents an extended contact zone between genetically divergent North Sea and Northeast Arctic cod and is maintained by two-way gene flow: by northward drift of pelagic eggs and larvae and by southward spawning migrations of Northeast Arctic cod. Computer simulations verify that the genetic cline can be established rapidly if gene flow into coastal populations is substantial. The shape of the cline, on the other hand, was found to be largely insensitive to the total amount of gene flow and therefore carries little information on extent of gene flow into and among coastal populations.

Résumé : La morue (*Gadus morhua*) côtière dans le nord-est de l'océan Atlantique connaît un déclin soutenu depuis l'industrialisation de la pêche côtière, et sa gestion doit tenir compte des complexités spatiales et temporelles associées aux stocks de morues coexistants. À cette fin, des analyses génétiques et la modélisation océanographique des motifs de dérive côtière et océanique de larves de morue ont été combinées dans le but d'élucider les mécanismes à l'origine d'un cline génétique observé sur une étendue de >1500 km le long de la côte norvégienne. Les résultats indiquent que le cline du nord au sud de la morue côtière représente une vaste zone de contact de morues de la mer du Nord et de morues du nord-est de l'océan Arctique, divergentes sur le plan génétique, qui est maintenue par un flux génétique bidirectionnel produit par la dérive vers le nord d'œufs et de larves pélagiques et par les migrations de frai vers le sud de morues du nord-est de l'océan Arctique. Des simulations informatiques confirment que l'établissement du cline génétique peut être rapide si le flux génétique entrant dans les populations côtières est important. La forme du cline, par contre, s'avère largement insensible à la quantité totale du flux génétique et offre donc peu d'information sur l'ampleur du flux génétique entrant dans les populations côtières et entre ces dernières. [Traduit par la Rédaction]

Introduction

A major role for genetic analyses of commercially important species is to uncover population structuring and aid management by providing information on the appropriate geographic partitioning of the resource for assessment and harvesting or protection (Ryman and Utter 1987; Waples et al. 2008). This goal is challenging when, as is often the case for many marine organisms, for example, genetic patterns take the form of gradual trends or clines rather than distinct geographic patches (Spies et al. 2015). Genetic clines, the manifestation of gradual shifts in allele frequencies over an often considerable geographic distance, can arise through several evolutionary mechanisms. These include isolation by distance, whereby local populations diverge through random genetic drift and gene flow among neighbours creates the cline (Wright 1943; Kimura and Weiss 1964; Slatkin 1993); contact between formerly genetically divergent populations (secondary contact: Barton and Hewitt 1985); mechanical mixtures in samples from the overlap zone (Hemmer-Hansen et al. 2019); gene surfing during range expansion (Edmonds et al. 2004; Excoffier et al. 2009); and natural selection in response to an environmental gradient acting directly on the observed genes (Haldane 1948; Schmidt et al. 2008) or indirectly through hitchhiking with selected genes (Maynard Smith and Haigh 1974; Barton 2000). The existence of a genetic cline in population genetic data may therefore owe to a wide array of causes, and the interpretation of such data in terms of management implications is not obvious. A way forward in such situations is to try to clarify the mechanism(s) behind the genetic cline and thereby gain insights into patterns of population connectivity and isolation and use that insight to inform management.

Genetic gradients or clines spanning hundreds to several thousands of kilometres have been described in several marine species, including the Atlantic cod (*Gadus morhua*) (Mork et al. 1985; Bradbury et al. 2010; Dahle et al. 2018*b*). In the Northeast Atlantic, the pantophysin I (*Pan* I) locus has been found to display a marked north–south cline with the *B* allele increasing in frequency from

Received 2 October 2020. Accepted 8 February 2021.

P.E. Jorde. Institute of Marine Research, Nye Flødevigveien 20, NO-4817 His, Norway.
M.B.O. Huserbråten, B.B. Seliussen, M.S. Myksvoll, F.B. Vikebø, G. Dahle, and A. Aglen. Institute of Marine Research, Nordnesgaten 50, NO-5005 Bergen, Norway.
T. Johansen. Institute of Marine Research, Framsenteret, Hjalmar Johansensgate 14, NO-9007 Tromsø, Norway.
Corresponding author: Per Erik Jorde (email: p.e.jorde@hi.no).

nearly null in the North Sea to reach a high frequency in the Barents Sea (Sarvas and Fevolden 2005). Less marked clinal patterns are seen also in putative selective neutral microsatellites (Skarstein et al. 2007; Dahle et al. 2018*b*). Recent genomic studies have revealed that much of the observed spatial genetic structure in Atlantic cod is confined to four large (several megabases) chromosomal inversions (Bradbury et al. 2010; Berg et al. 2016; Kirubakaran et al. 2016, 2020; Sodeland et al. 2016). These inversion polymorphisms have a wide geographic distribution on both sides of the Atlantic (Berg et al. 2017), display highly divergent sequences (Sodeland et al. 2016), and are likely to be evolutionarily quite old and predating the postglacial colonization of North Atlantic coastal waters (Berg et al. 2016). The *Pan* I locus is located within one of these inversions (on chromosome LG1: Kirubakaran et al. 2016) and so are many, but not all, of the other genes that share the north–south clinal pattern in the Northeast Atlantic (Johansen et al. 2020).

To draw relevant management implication of these observations, we investigated the mechanism(s) responsible for generating and maintaining the clinal genetic pattern. Toward that end, we combined genetic screening of offshore and coastal cod along the Northeast Atlantic (Norwegian) coast with modelling of putative gene flow patterns and tests for genotype–environment associations. The findings are discussed in relation to pressing management issues for coastal cod in this region.

Material and methods

The species

The Atlantic cod is a commercially important demersal species found across the North Atlantic in both offshore and coastal regions from 0 to 600 m depth (Brander 1995; Berg et al. 2017). In the North Atlantic some 20 cod stocks are assessed and managed separately. Many of these have been severely overfished, and experience dictates that recovery can be slow (Hutchings 2000; Pedersen et al. 2017). The Norwegian fisheries relate to three cod stocks: North Sea cod, Norwegian coastal cod (CC: divided in two management units north and south of 62°N), and the Northeast Arctic cod (NEAC) stock from the Barents Sea. While the CC spawn all along the coast and perform only short migrations (Jakobsen 1987: Michalsen et al. 2014), the NEAC perform spawning migrations from the Barents Sea southward to the coast of Norway in February to April (Nordeide 1998; Olsen et al. 2010; Michalsen et al. 2014; Johansen et al. 2018). The pelagic NEAC eggs and larvae are carried northwards with the Norwegian Coastal Current back to nursery and feeding areas in the Barents Sea (Bergstad et al. 1987).

The main fishery for CC is during spawning season in spring, and north of 62°N this is in a mixed fishery with the much larger NEAC stock component. The CC is at its historical low, and since 2004, ICES (the International Council for the Exploration of the Sea) has urged rebuilding the stock (ICES 2019). A rebuilding plan specifying reductions in fishing mortality for Norwegian CC has been in operation since 2011, but the stock remains at low levels (ICES 2019).

Sampling

This project utilized previously available samples from the "CODBIOBANK" project, carried out during the years 2002 to 2007, and targeted adult cod from spawning locations along the entire Norwegian coast (Dahle et al. 2018*a*). Sampling took place from February to May during the spawning period, in close cooperation with local fishermen and Institute of Marine Research (IMR) annual surveys. For the present study, we selected a subset of 28 locations (Fig. 1; also refer to online Supporting Information, Table S1[1]), intended to represent cod in outer coastal areas, which are the main target of commercial coastal catches, while avoiding migratory NEAC in our samples as much as possible. We refer to the sampled cod as CC while acknowledging that individuals of other stock component(s), including NEAC and (or) putative fjord populations, may inadvertently be present in the catches. Temporal replicates were obtained for 2 years at six localities (cf. Supporting Information, Table S1[1]) or nearby localities at most a few kilometres apart, for a total of 34 coastal samples. Each coastal sample typically consisted of 48 adults (range 15 to 96). All individuals were measured, sex and maturity stage were determined based on visual inspection of gonads, and otoliths were collected for age and shape analysis. Mean age of sampled cod was 5.9 years (range 2 to 13 years).

Additional samples, not previously published, of offshore cod were collected as reference samples for comparative analyses, representing the major ocean-spawning cod in the region: North Sea cod (Vikingbank, n = 93 adults) and the NEAC, the latter collected as spawning migrants off western Lofoten (n = 48) and as juveniles from the Barents Sea (n = 139) (Fig. 1; Supporting Information, Table S1[1]) the reference samples were collected by longline in close cooperation with fishermen in the North Sea and by IMR annual trawl surveys in Lofoten and the Barents Sea.

Ageing and otolith typing

All fish were aged from otoliths, based on break and burn as described by Berg and Albert (2003). Type classification of CC and NEAC is based on the second annual ring when the cod is 2 years old (Rollefsen 1933; Berg and Albert 2003). The cod were classified into types 1 and 2 as "certain" and "uncertain" CC, respectively, and types 4 and 5 as "uncertain" and "certain" NEAC, respectively. Otoliths from North Sea cod are indistinguishable from the CC type. The NEAC reference samples were collected at known NEAC spawning (adult sample) or nursery (juvenile) areas and were not otolith-typed.

Genetic analyses

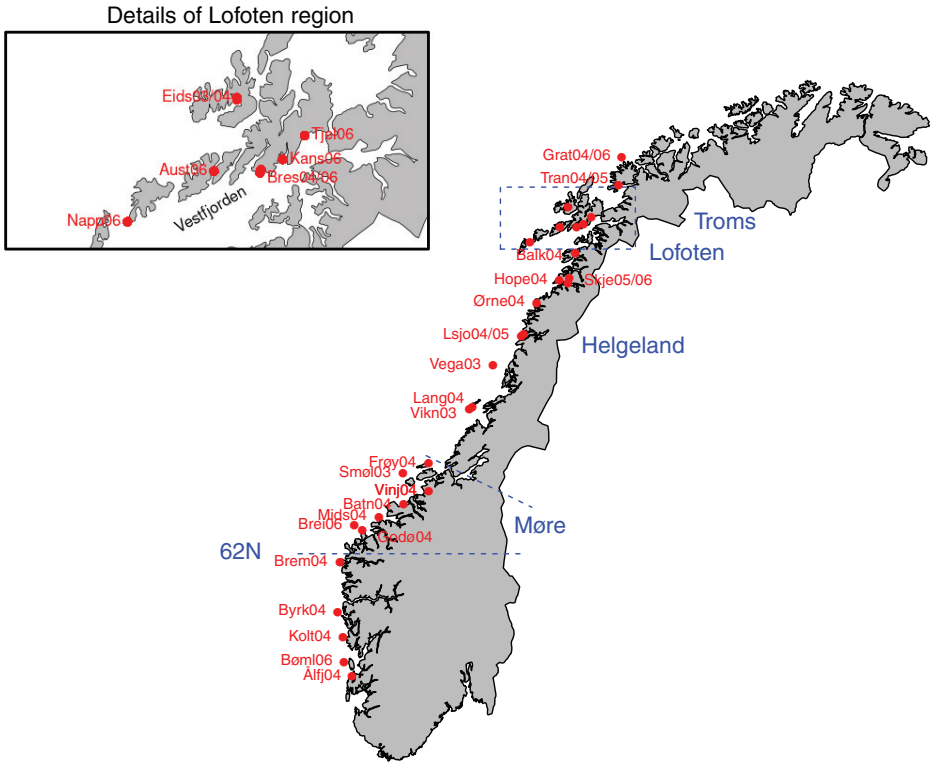
Fin clips from all individuals were stored in 96% ethanol prior to DNA extraction. DNA extraction was performed using the DNeasy Blood & Tissue Kit (Qiagen). 42 single-nucleotide polymorphisms (SNPs), including *Pan* I, were selected to distinguish between NEAC and CC and uncover structuring within CC (Johansen et al. 2018). SNPs were genotyped using matrix-assisted laser desorption–ionization time-of-flight mass spectroscopy (MALDI-TOF MS) assays (Agena Bioscience Inc., Hamburg Germany). Genotyping was performed using the iPLEX protocol following manufacturer's instructions (Agena Bioscience). The MassARRAY Typer software was used for automated genotype calling (Agena Bioscience). SNPs with more than 20% missing data per sample or a minor allele frequency below 0.05 were discarded, resulting in 40 SNPs for subsequent statistical analyses. Of these 40 SNPs, 27 were situated within one of the four chromosome inversions on chromosomes LG1 (15 SNPs, including *Pan* I), LG2 (four SNPs), LG7 (four SNPs), and LG12 (four SNPs), while 13 SNPs were outside inversion regions on these or on other chromosomes (Supporting Information, Table S2[1]). Missing values among the total 1935 individuals (including reference samples) averaged 150.9 per SNP, or 7.8%.

Statistical analyses

Genotype frequencies and allele frequencies for the 40 SNPs were estimated and used to characterize genetic variability patterns within and among samples or groups of samples. Deviations from Hardy–Weinberg genotype proportions was characterized by F_{IS} following Weir and Cockerham (1984). Correlations of genotypes within pairs of markers (i.e., linkage disequilibrium) were calculated as the square (r^2) of the ratio between the composite disequilibrium coefficient (delta) and Weir's denominator (Weir 1996; p. 137) using inhouse software. Approximate tests for significance of F_{IS} and r^2

[1]Supplementary data are available with the article at https://doi.org/10.1139/cjfas-2020-0380.

Fig. 1. Map of Norway showing named geographic regions and sample localities (red dots). Sample abbreviations are as in Supporting Information, Table S1[1]. Map was created using R (R Core Team 2020) with the maps and mapproj packages (data from CIA World Databank II, https://www.evl.uic.edu/pape/data/WDB/; original S code by Becker and Wilks 1993, 1995; R version by Brownrigg 2018; McIlroy 2020). [Colour online.]



were calculated as $X^2 = n \times F_{IS}{}^2$ and $X^2 = n \times r^2$, respectively, where n is the sample size, and evaluated against the χ^2 distribution with one degree of freedom (Nei 1987; p. 156 and eq. 7.52, respectively). Genetic differentiation among samples and regional groups of samples was quantified by F_{ST} calculated according to following Weir and Cockerham (1984). F_{ST} was also calculated pairwise among coastal samples and between coastal samples and oceanic references. We used pairwise F_{ST} estimates with the North Sea reference to elucidate mechanisms behind CC genetic structure by comparing spatial differentiation pattern along the coast with modelled oceanographic gene flow patterns and possible selection gradients (water temperature profiles), as described later. Standard errors (SE) and 95% confidence intervals (CI) for F_{ST} were calculated by the leave-one-out jackknife procedure over SNPs (Efron and Tibshirani 1993; eqs. 12.6 and 11.5).

Genotype data were further analysed at the individual level to evaluate possible population mixing within samples. Two different clustering algorithms based on individual genotypes (all 40 SNPs) were used, as implemented in the STRUCTURE software (version 2.3.4; Pritchard et al. 2000) and in the DAPC (discriminant analysis of principal components; Jombart et al. 2010) routine in the adegenet package (version 2.1.3; Jombart 2008) under the R statistical environment (version 3.6.3; R Core Team 2020). In the DAPC we first rand the find.clusters function on all samples, including reference samples, but this failed to resolve an optimal number of clusters (K). The dapc function were subsequently run by retaining 20 principal component axes, and the first two principal components were plotted, colour-coding individuals according to sample type (coastal or reference sample). STRUCTURE was run with $K = 2$ with 100 000 burn-ins and 200 000 replicates, exploring both the admixture and the no-admixture models with correlated allele frequencies, and with higher K to check for evidence for further population structuring.

Oceanographic modelling

Ocean modelling was used to predict patterns of movements of pelagic early life-history stages as probable pathways for gene flow into and among CC populations (Myksvoll et al. 2014). The hydrodynamic model used to represent the ocean currents in the study area was based on the Regional Ocean Modeling System (ROMS, http://myroms.org), a free-surface, hydrostatic, primitive equation ocean general circulation model (Shchepetkin and McWilliams 2005; Haidvogel et al. 2008). The ROMS model was run with a horizontal resolution of 800 m × 800 m (for details see Albretsen et al. 2011), high-resolution wind fields (Skamarock et al. 2008), and realistic freshwater discharge from all rivers in the model domain (provided by the Norwegian Water Resources and Energy Directorate; see Beldring et al. 2003). Advection of particles in the horizontal plane were modelled by applying the fourth-order Runge–Kutta advection scheme in an open-source Lagrangian particle tracking model (LADIM, https://github.com/bjornaa/ladim.git) coupled with the velocity fields from the ROMS model.

To model the inflow of settling pelagic juveniles into the sampled CC populations from the three hypothesized sources, offshore North Sea, other CC, and offshore spawning NEAC, we released particles in three distinct scenarios. First, to represent cod spawning in the northern North Sea, we released particles at random locations along the Norwegian Trench slope where larger cod have been known to aggregate during the spawning season in spring (Fox et al. 2008; Huserbråten et al. 2018). Second, to represent other CC populations, we released particles from the geographical positions of all the coastal genetic sampling locations (cf. Fig. 1). Third, to represent

NEAC spawning, we released particles within spawning areas along the Norwegian coast known to be used by the NEAC (Sundby et al. 2013; Supporting Information, Fig. S1[1]). Particles were released every day during 1–31 March to cover the approximately common peak spawning period of North Sea and NEAC cod (ICES 2005) and drifted at a fixed depth uniformly distributed between 1 and 20 m. The particles were allowed to drift until 20 July, as settlement period of cod in this area has been found to last from start of June to mid-July (cf. Huserbråten et al. 2018). We repeated the modelled drift scenarios over a 10-year time period (2008–2017) and summed particles over years in a connectivity matrix, which was used in subsequent genetic simulations described below.

Temperature analyses

Annual mean sea water temperatures at 20 m depth at each sample locality, based on daily average values from the ROMS-model over 5 years, were tested for potential impact on levels of genetic divergence with linear regression (lm function in R; R Core Team 2020). Because of the strong correlation of temperature on latitude (Supporting Information, Fig. S2[1]), we first regressed temperature on latitude and used the residual temperature to test for effect on the genetic structure of CC. We used the pairwise F_{ST} estimates between coastal samples and the North Sea to quantify CC spatial genetic structure.

Genetic simulations

The consequences of drift of pelagic eggs and larvae, as summarized in the connectivity matrix from the oceanographic modelling, on CC genetic structure was inferred from computer simulations. We compared observed and simulated genetic patterns, as quantified by pairwise F_{ST} measures between CC samples and North Sea cod, to find the simulation parameters that best fitted observations. Simulation parameters included local population sizes and number of particles (larvae) arriving from each of the three gene flow sources: North Sea, NEAC, and other CC populations. Because these numbers are unknown, we evaluated a range of numerical scalings of the modelled connectivity matrix, representing a wide range of average levels of gene flow (Supporting Information, Table S3[1]).

Briefly, using an inhouse simulation program (see Data availability statement below), we simulated gene flow and random genetic drift in an array of 33 "coastal" (N = 5000 individuals each) and two "oceanic" populations treated as infinite size (i.e., with no drift or immigration). The population array was initiated with two selectively neutral alleles, in frequencies chosen to give the observed differentiation between the North Sea and NEAC reference samples (F_{ST} = 0.431 on average over 40 SNPs; see Results). Generations were discrete, and all individuals were immediately replaced after reproduction by the offspring. In each computer run, the required number of migrants were drawn at random from each population and exchanged with the others according to the connectivity matrix. Because gene flow in and out of populations was asymmetric, excess emigration was made up for by increased reproduction, whereas excess immigration was made up for by culling during the reproduction stage, preserving constant population sizes at the time of census. The connectivity matrix from the oceanographic modelling was scaled to yield a range of average total proportion of immigrants of 0.2%, 2%, and 20% in coastal populations. The relative contribution from other coastal populations, the North Sea and the NEAC, was varied to determine appropriate scaling of the connectivity matrix that yielded results that best fitted the observed F_{ST} values. We judged "best fit" by eye, partly aided by the sum of squared differences between simulated and observed F_{ST}. The cycle of reproduction, emigration, immigration, and census was repeated each generation for up to 5000 generations, and pairwise F_{ST} estimates with the North Sea population were calculated for each coastal population from the simulated allele frequencies. Replicated computer runs (1000 for each parameter set) was taken to represent independent gene loci, and F_{ST} was averaged over runs.

Results

Genetic analyses included a total of 1655 cod from 28 coastal locations and two oceanic reference populations (Supporting Information, Table S1[1]) screened for genetic variation at 40 SNPs (Supporting Information, Table S2[1]). Heterozygosity (H_S) varied among the 40 SNPs from 0.121 to 0.496, for a mean of 0.365 (Table 1). The mean H_S in the North Sea sample (0.333) was similar to the coastal ones, while it was considerably lower in the NEAC (0.189). Overall, there was a tendency for correlation in genotypes among markers (i.e., linkage disequilibrium) in the pooled coastal samples. This correlation averaged 0.0238 for all $40 \times 39/2 = 780$ marker pairs, reduced to 0.0078 when considering only pairs situated on different linkage groups and thus physically unlinked (612 pairs; Table 1). Both of these estimates were highly significant, as judged by the χ^2 test.

The CC samples differed genetically from NEAC and North Sea cod as well as among themselves. The mean F_{ST} among coastal locations (temporal replicates pooled) was 0.028 over the 40 SNPs, with a 95% CI of 0.017–0.038. Pooling all coastal samples, the mean F_{ST} between coastal and North Sea cod was 0.036, whereas between CC and NEAC mean F_{ST} was 0.232 (Table 2). The North Sea and NEAC (adults and juveniles pooled) samples in turn differed from each other by F_{ST} = 0.431 (95% CI: 0.299–0.564).

Differentiation among CC varied greatly among geographic regions, with CC south of 62°N being similar to North Sea cod (F_{ST} = 0.010), while CC further north differed substantially from North Sea cod (in region Troms: F_{ST} = 0.127). On average, F_{ST} between coastal localities and the North Sea reference increased by 0.009 for every degree latitude north (cf. Supporting Information, Fig. S2[1], top left panel), verifying the previously described genetic cline in CC. Comparison with NEAC displayed the opposite trend, with CC being more similar to, yet still differing substantially from, NEAC in the north (Table 2). The transition or cline in differentiation was not strictly linear, however, and there was a notable deviation from a linear trend in F_{ST} for the Lofoten Inner (Vestfjorden) samples at ~68°N, where CC was more similar (i.e., had lower F_{ST}) to the those in the North Sea than were neighbours immediately south and north (Fig. 2). Of the six temporal replicated coastal sample pairs, two differed significantly between the two time points (at localities Eidsfjord (years 2003 and 2004) and Tranøybotn (2004 and 2005)), whereas four did not (Lille Sjona, Skjerstadfjord, Bresja and Gratangen; Supporting Information, Table S4[1]).

The magnitude of genetic divergence with the North Sea decreased with mean water temperature (slope = −0.021, P = 0.001; Supporting Information, Fig. S2[1]). However, water temperature was strongly correlated with latitude ($P \sim 0.000$), and after removing this relationship the residual temperatures had no significant effect on F_{ST} (P = 0.334: Supporting Information, Fig. S2[1], top right panel).

Leaving out suspected NEAC individuals from coastal samples (i.e., individuals carrying otolith types 4 or 5) typically had little effect on the F_{ST} estimates (cf. Fig. 2). In particular, the northernmost samples had very few such types (region Troms; Table 1), and F_{ST} hardly changed for them. Only in the samples Godøy 2004 and Nappstraumen 2006 did F_{ST} change substantially after removal of individuals with NEAC-type otoliths. In both cases, removal of individuals led to reduced F_{ST} with the North Sea, indicating that some NEAC were mixed in with CC in the samples. However, the Godøy sample displayed an anomalously high F_{ST} relative to neighbouring locations also after removing all suspected NEAC individuals from this locality.

Genetic differences at the individual level allowed tentative classification to coastal and NEAC types. Using STRUCTURE with the no-admixture model, the majority of CC clustered with the North Sea reference sample when the software was run for K = 2 groups (Fig. 3,

Table 1. Regional genetic variability and disequilibrium in 40 single-nucleotide polymorphisms (SNPs).

Geographic region	n	H_S	H_{obs}	F_{IS}	r^2_{all}	$r^2_{nonlinked}$	NEAC / CC otolith type
Coast							
South of 62°N	238	0.358	0.346	0.032	0.0219	0.0061	0 / 218
Møre	296	0.357	0.344	0.036	0.0267	0.0106	10 / 286
Helgeland	553	0.362	0.347	0.041	0.0205	0.0043	16 / 535
Lofoten Inner	238	0.361	0.348	0.037	0.0236	0.0083	10 / 227
Lofoten Outer	140	0.355	0.322	0.092	0.0369	0.0204	17 / 123
Troms	190	0.340	0.333	0.023	0.0284	0.0123	3 / 183
All coast pooled	1655	0.365	0.343	0.059	0.0238	0.0078	56 / 1572
Oceanic reference samples							
North Sea	93	0.333	0.326	0.023	0.0273	0.0124	1 / 92
NEAC (pooled)	187	0.189	0.177	0.067	0.0279	0.0077	NA

Note: H_S is the average gene diversity (expected heterozygosity) corrected for finite sample size (Nei and Roychoudhury 1974), and H_{obs} is the observed proportion of heterozygotes. $F_{IS} = (H_{obs} - H_S)/H_S$ (Nei 1987, eq. 7.31), and r^2 (Weir 1996, p. 126) estimate the average deviation from Hardy–Weinberg genotype proportions and average linkage disequilibrium, respectively, the latter separately among all 780 locus pairs (all) and among 612 pairs situated at different linkage groups (nonlinked). n are the sample sizes, and the number of individuals carrying coastal cod (CC: types 1 and 2) or Northeast Arctic cod (NEAC)-type (4 and 5) otoliths are also given. The NEAC reference samples and 27 coastal individuals lacked otolith type.

Table 2. Pairwise F_{ST} (with 95% confidence intervals, CI) between coastal regions and oceanic reference samples.

		Oceanic references	
Geographic region	n	North Sea, $n = 93$	NEAC, $n = 187$
Coastal samples			
South of 62°N	238	0.010 (0.004–0.017)	0.364 (0.246–0.482)
Møre	296	0.010 (0.004–0.016)	0.325 (0.214–0.436)
Helgeland	553	0.050 (0.030–0.070)	0.223 (0.135–0.310)
Lofoten Inner	238	0.038 (0.017–0.059)	0.279 (0.167–0.391)
Lofoten Outer	140	0.072 (0.043–0.101)	0.231 (0.128–0.333)
Troms	190	0.127 (0.076–0.179)	0.170 (0.080–0.260)
All	1655	0.036 (0.020–0.052)	0.232 (0.143–0.320)

Note: Estimates are based on 40 SNPs, and CIs were calculated by jackknifing over SNPs. n are sample sizes.

upper panel, blue bars), but an appreciable number clustered with higher probability to the NEAC reference (orange bars) instead. The proportion of CC that assigned to the NEAC reference was especially high in northern samples (Eidsfjord and Tranøybotn), but such individuals occurred in all coastal samples. Under the admixture model many coastal individuals had intermediate Q values and were possibly of mixed origin (Fig. 3, lower panel, blue and orange bars of intermediate height). For $K = 3$ and higher, the estimated probability of data (Ln Prob of Data) increased, but without revealing any obvious geographic patterns among coastal samples (cf. Supporting Information, Fig. S3[1]).

The correlation between otolith type and STRUCTURE assignments of individual cod into CC and NEAC (no-admixture model) was highly significant ($p < 0.0001$; Table 3) yet only moderately strong, with a Pearson product-moment correlation coefficient of $r = 0.295$. The major reason for this rather low value was that a large number (274 out of 351, or 78%) of cod that STRUCTURE assigned with high probability (>0.8) to NEAC instead had CC (type 1) otoliths. An appreciable fraction (8 out of 26, or 31%) of individuals carrying uncertain NEAC otoliths (type 4) were genetically assigned with high probability to CC cod, whereas only a single individual with certain NEAC otoliths (type 5) was assigned to CC.

An alternative visualization of individuals genetic composition in a DAPC plot (Fig. 4) provided some insight into the interrelationship of coastal (blue dots) and reference cod (red or green dots) and their otolith classification. Overall, there was a broad overlap in the DAPC plot among cod sampled at coastal localities and those sampled offshore, with no evidence for distinct clusters within the former. Individuals with otoliths classified as "certain" NEAC (type 5, encircled in Fig. 4) tended to cluster with NEAC reference cod (red dots), whereas those with the "uncertain" NEAC (type 4) otoliths showed less tendency to do so (triangles).

Oceanographic modelling

The particle simulation model yielded a matrix (Fig. 5), whose off-diagonal elements represented number of particles exchanged between the CC localities (northwards above diagonal, southward below), and with the North Sea (Vikingbank) and the 15 pooled NEAC spawning grounds (unidirectional to the coast; rightmost columns). The matrix was strongly asymmetric, with most particles moving northwards (above diagonal), reflecting transport by the northward moving Norwegian Coastal Current. As a result, most southern localities contributed particles (eggs or larvae) to downstream localities in the north with very little movement in the opposite direction, except for NEAC, which contributed to all but the very southernmost coastal localities through spawning migration (spawning areas 39 to 51; Supporting Information, Fig. S1[1]).

Genetic simulations

The effort to replicate observed F_{ST} values by means of genetic simulations based on the (scaled) connectivity matrix from oceanographic modelling recovered the gross features of the observed north–south genetic cline (Fig. 6). Best fit between observed and simulated F_{ST}, as measured pairwise between the North Sea and coastal localities, occurred when the relative contributions from the three gene flow components (i.e., other coastal sites, the North Sea, and pooled NEAC spawning grounds) were scaled to yield similar amounts (Supporting Information, Fig. S4[1]). Under almost all simulated conditions, however, a cline evolved along the coast with increasing F_{ST} towards the north irrespectively of the total amount of gene flow (0.2%, 2%, or 20%). When the total amount was very high (here 20%), the cline was fully established within a modest number of generations (10–50), whereas when it was low (0.2%) many more generations (>1000) were required (Fig. 6). While this cline resembled the observed one (gray line) in the broad sense, it failed in the details, in particular regarding the observed dip in F_{ST} in the Vestfjorden area.

Fig. 2. Observed pairwise F_{ST} between the North Sea (Vikingbank) reference sample and 34 coastal samples. Vertical bars represent 95% confidence intervals (CI) for the observed estimates, calculated from jackknifing over single-nucleotide polymorphisms (SNPs). Open symbols represent estimates after individuals carrying NEAC type otoliths (types 4 and 5) were excluded.

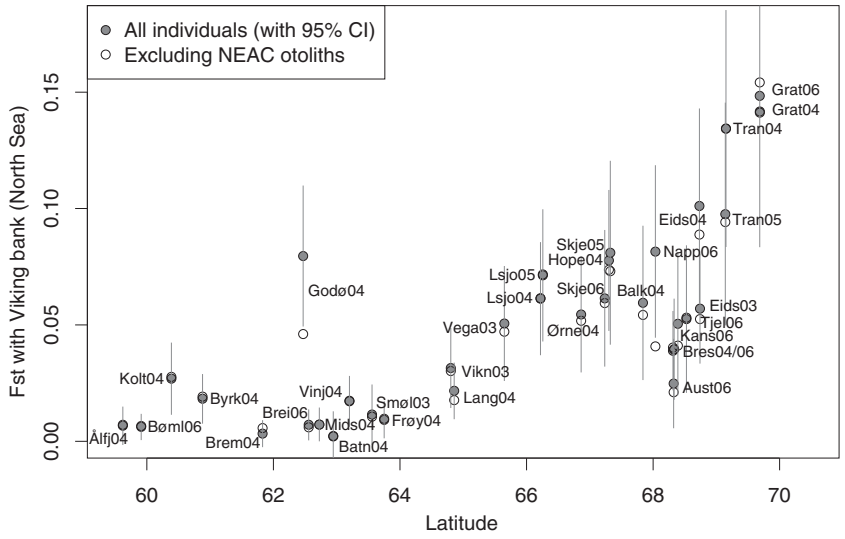
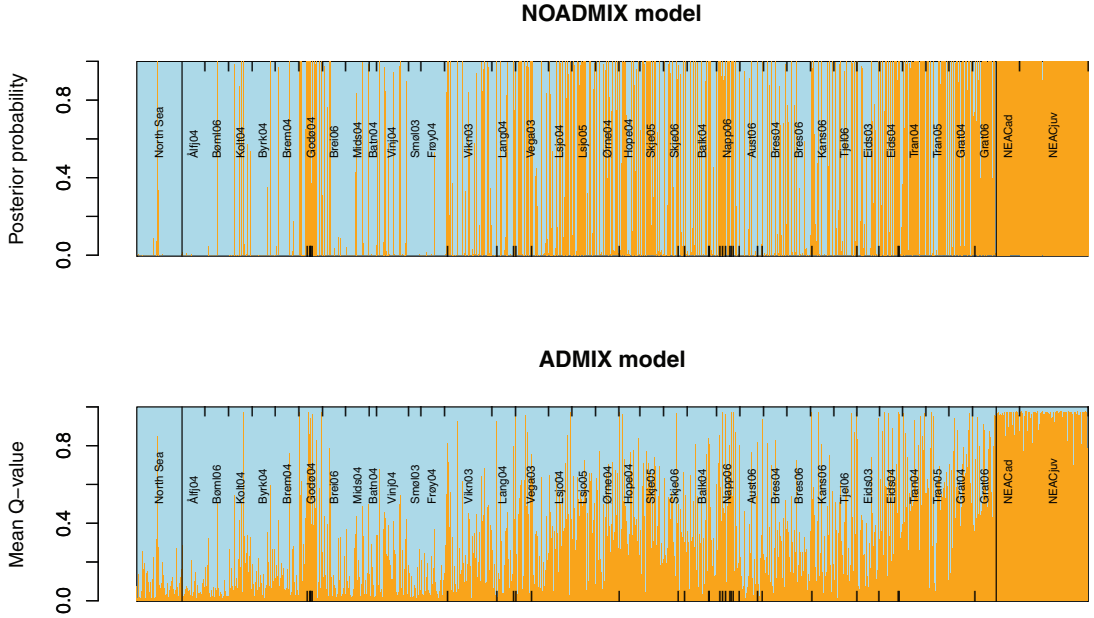


Fig. 3. STRUCTURE results for coastal cod and reference samples, calculated from 40 SNPs, for $K = 2$. The upper panel gives the posterior probability for individual membership to the North Sea (blue) or Northeast Arctic cod (NEAC; orange) group or population under the no-admixture model, and the lower panel gives mean Q values under the admixture model. Tick marks along the top edge of the plots separates coastal samples, and marks along the bottom edge indicate individuals carrying NEAC (type 5) otoliths. Abbreviated sample names are as in Supporting Information, Table S1[1]. [Colour online.]



Discussion

Our observations verify the notable north–south cline in allele frequencies in CC in the Northeast Atlantic, as first described by Mork et al. (1985) in isozymes, by Skarstein et al. (2007) and Dahle et al. (2018*b*) in microsatellites, and by Johansen et al. (2020) in SNPs, while our modelling and simulations explore possible mechanisms for its generation and maintenance.

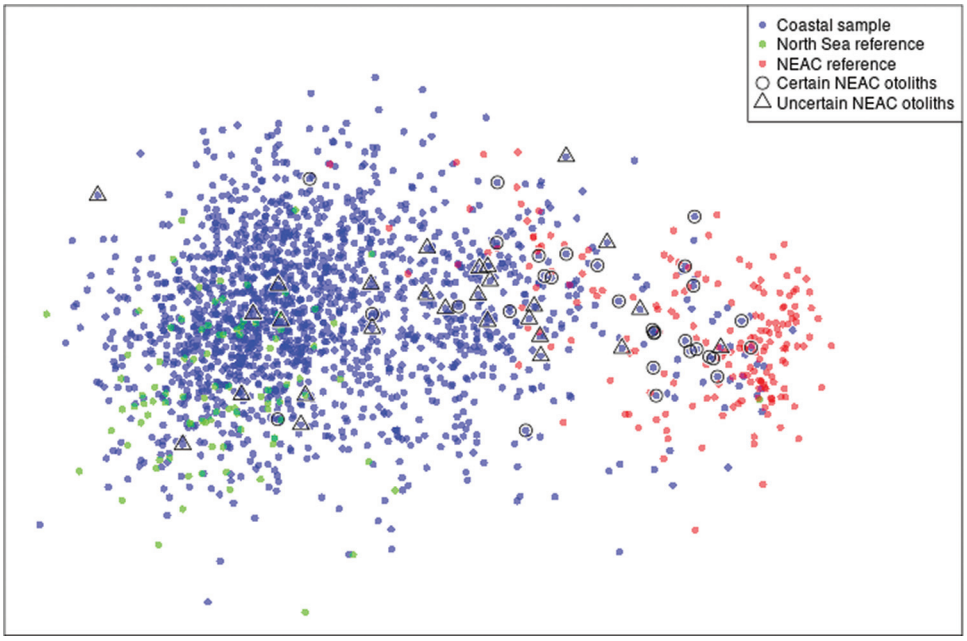
As compared with Johansen et al. (2020), we sacrificed number of SNPs to expand the number of individuals, localities, and years sampled, allowing a more fine-scaled spatiotemporal analysis. Most of the SNPs that were genotyped herein are located within the four known chromosome inversions in the Atlantic cod genome. Indeed, these SNPs were chosen largely for their resolving power in distinguishing NEAC from CC. Because of restricted recombination, SNPs located within inversion tended to have correlated genotypes and thus be in linkage disequilibrium. There would therefore be little if any advantage of increasing the number of SNPs within these genomic regions for our specific purpose of resolving the genetic cline in CC. Conversely, a much more extensive genomics approach would be useful in future studies to try to

Table 3. Cross-tabulation of otolith type and STRUCTURE assignment (posterior probability; no-admixture model).

	STRUCTURE posterior probability to NEAC reference			
Otolith type	<0.2	0.2–0.8	>0.8	Sum
CC (type 1)	1135	85	274	1494
CC uncertain (type 2)	40	6	32	78
NEAC uncertain (type 4)	8	1	17	26
NEAC (type 5)	1	1	28	30
Sum	1184	93	351	1628

Note: Otoliths were classified as coastal (CC) or NEAC type (27 individuals were not classified and are not included in the table). Posterior probabilities refer to STRUCTURE $K = 2$ output for the NEAC cluster and is binned as low (<0.2), intermediate (0.2–0.8), and high (>0.8). Contingency chi-square test: $\chi^2 = 149.98$, df = 6, $P \ll 0.0001$.

Fig. 4. Discriminant analysis of principal components (DAPC) plot of coastal and reference cod. Individuals are colour-coded according to sample (blue = coastal samples, green = North Sea, red = NEAC). Individuals carrying NEAC-type otoliths are enclosed in triangles (for uncertain NEAC, type 4) or circles (for certain NEAC, type 5: the NEAC references themselves are not encircled, as many were collected as juveniles and not otolith-typed). [Colour online.]



uncover genetic factors behind Atlantic cod life-history strategies. A plausible, but as yet unsubstantiated, hypothesis is that one or more of the hundreds of genes residing in these inversions (e.g., Berg et al. 2016) dispose the cod to either stay and settle on the coast, thus being a "coastal" cod, or leave the coast and take up a more oceanic and migratory lifestyle, thus being a "NEAC".

Causes for the cline

The observed genetic cline in Northeast Atlantic cod may be interpreted as a gradual transition in CC from North Sea-like in the south towards more NEAC-like in the north. It so happens that NEAC differ from CC and North Sea cod primarily in the relative frequencies of inversion haplotypes, with much less divergence in the rest of the genome (Berg et al. 2016; Johansen et al. 2020). This state of affairs most likely reflects the result of natural selection acting on one or more of the hundreds of genes located in these inversions, including the *Pan* I locus (Case et al. 2006; Skarstein et al. 2007) and other loci (Kirubakaran et al. 2016; Berg et al. 2017; Sinclair-Waters et al. 2018). This presumed selective divergence in inversion haplotype frequencies may be historical and (or) it may be an ongoing process. In the latter case, selective response to an environmental gradient (e.g., temperature or others) may be the direct cause of the observed genetic cline. However, our computer simulations demonstrate that ongoing selection is not a necessary prerequisite for the generation and maintenance of a north–south genetic cline. Instead, the genetic cline may have arisen through intercrossing and gene flow between two already divergent stocks (i.e., the ancestors of present NEAC and North Sea–CC, respectively). If so, CC in the Northeast Atlantic represents a geographically extensive hybrid zone, most likely established in postglacial times when the Northeast Atlantic coast became available for colonization (Bigg et al. 2008). Such a secondary contact or hybrid zone scenario provides a parsimonious explanation for the north–south cline in all genes for which the two original stocks differ.

Extent of gene flow

The combined oceanographic modelling and genetic simulations indicated that the cline can be established rapidly by gene flow, in a matter of a few tens of generations if gene flow from oceanic populations is extensive. Moreover, the cline can be upheld indefinitely as an equilibrium between gene flow from opposing directions from the North Sea and Northeast Arctic. Interestingly, the shape of the cline was found to be fairly insensitive to the total amount of gene

Fig. 5. Visual presentation of the connectivity matrix obtained by oceanographic modelling. Squares represent sample localities coloured according to number of particles received (log-scale) and are arranged from south (Ålfjord) to north (Gratangen). The matrix shows that all coastal sample localities received particles from several other coastal sites and from the two offshore stocks (North Sea and NEAC), except for the very southernmost samples, which did not receive particles from the NEAC. Three coastal localities did not leave particles at any other locality. [Colour online.]

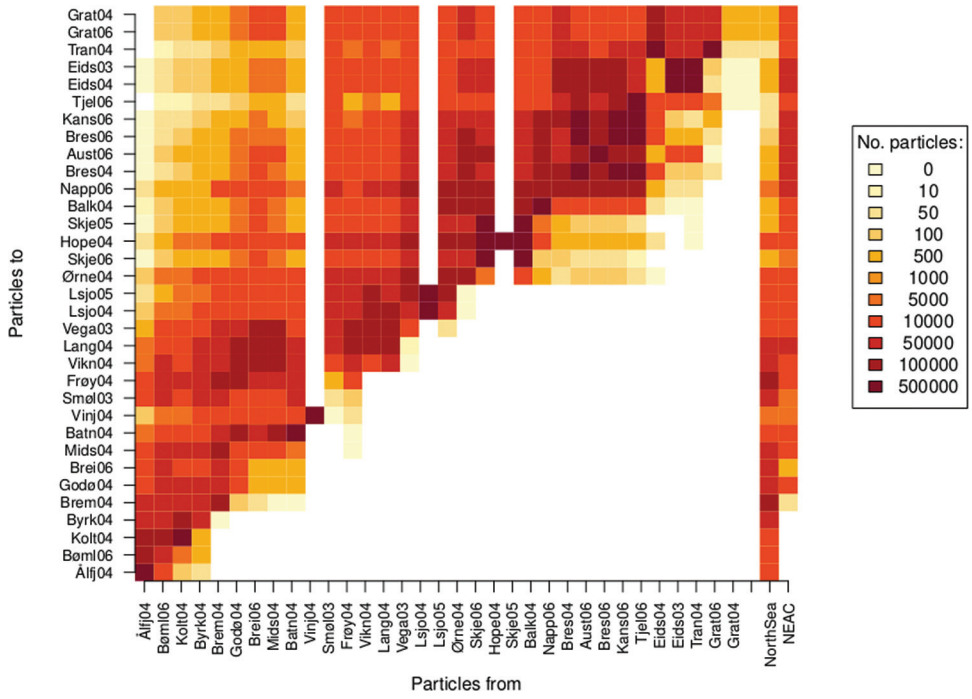
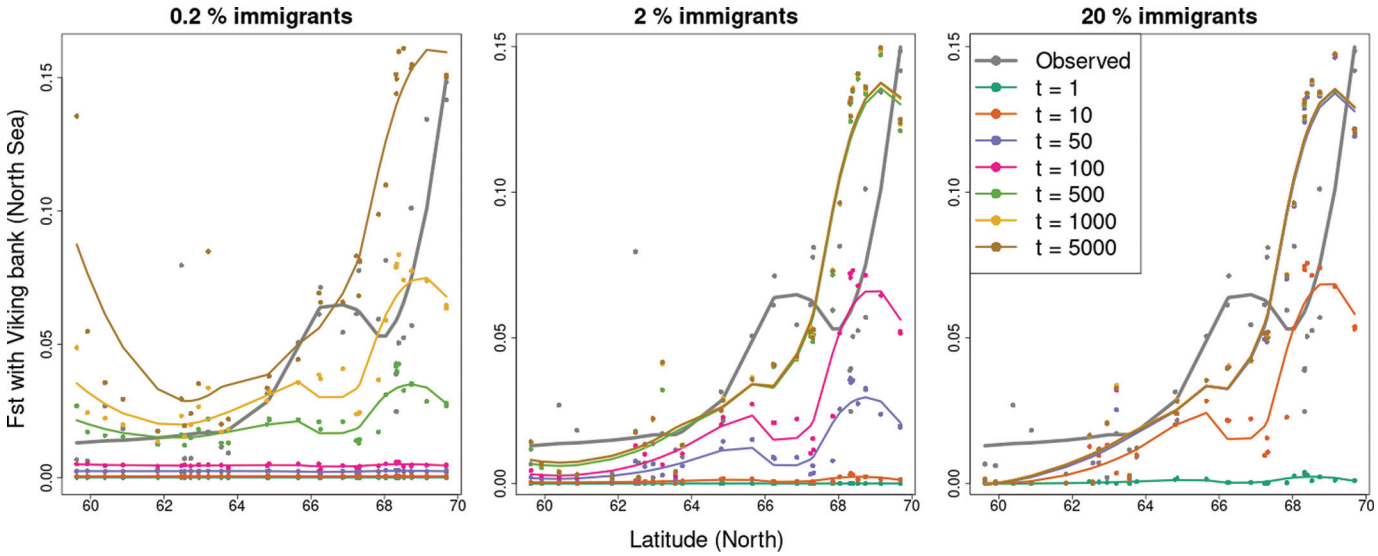


Fig. 6. Comparison of observed (gray) and simulated (coloured) F_{ST} values, calculated between coastal samples and the North Sea reference population. Simulated values were generated by population genetic computer simulations based on gene flow patterns inferred from oceanographic modelling (cf. Fig. 5). The three panels represent a 100-fold range in geneflow, representing on average 0.2% to 20% of coastal populations size, and show the emergence of a north–south cline in genetic divergence over time (*t*: generations). [Colour online.]



flow, whether extensive (here 20%) or low (0.2%). Best fit to observations occurred when the total contributions from North Sea cod, from NEAC, and from other coastal populations were even, but qualitatively similar clines occurred in simulations whenever the relative contributions were not extremely skewed. The existence of the cline itself therefore does not tell us much, if anything, about the magnitude of gene flow into coastal populations; it could be high and the cline established rapidly or low and taking longer to reach the equilibrium shape (cf. Fig. 6). While our computer simulations build on a number of simplifying assumptions that clearly are not realistic in all their details, we expect this finding of near independence of magnitude of gene flow to be robust. Much of the real-life

complexities, such as age structure and overlapping generations, are summarized in population genetics by the use of the so-called effective populations size (N_e) or the harmonic mean effective size over generations if population size changes. Thus, these details matter little for simulations covering hundreds to thousands of generations.

Genetic introgression

The putative presence of NEAC cod in coastal samples creates problems for the analysis of spatial genetic structure in CC. We found a strong tendency for otolith type 5 to be associated with individuals that clustered genetically to NEAC, indicating the presence of at least some NEAC cod in our coastal samples. However, removing all individuals of putative NEAC origin (i.e., those carrying otolith types 4 or 5) had little effect on the observed genetic cline, which remained essentially unchanged (cf. Fig. 2). The cline is therefore not driven by the presence of NEAC individuals in the coastal samples but, rather, by introgression of genes of NEAC (and North Sea) origins into CC. The observed mismatches between genotype and otolith type may be related to such introgression events. Otolith shape in cod is believed to largely reflect environmental conditions during early life stages (e.g., Stransky et al. 2008), and the observed mismatches indicate that some NEAC offspring, or possibly NEAC × CC hybrids, settle along the coast and could become incorporated in coastal populations. Such a scenario provides a likely mechanism for the inferred introgression of NEAC genes into coastal populations.

Current versus historic gene flow

Situated between the historically large, oceanic stocks in the North Sea and the Barents Sea, the potential for on-coast gene flow and introgression is obvious. Less clear is the extent to which gene flow is presently ongoing or is largely a historical phenomenon. One line of evidence for ongoing gene flow is that individuals with NEAC-type otoliths occur in all coastal regions where NEAC are known to spawn (see Supporting Information, Fig. S1[1] for NEAC spawning grounds) and that also match regions where genetic introgression of NEAC is observed. In contrast, no NEAC otoliths were observed on the coast south of 62°N (Table 1), and CC in this southern region were found to be genetically similar to North Sea cod. Spawning of NEAC cod is not constant over time, however, and while the epicentre of spawning is in Lofoten, the NEAC can either spawn farther south in Møre or farther north in Finnmark depending on cold or warm years (Sundby and Nakken 2008). Opdal (2010) hypothesized, based on historical catch statistics going back to the mid-19th century, that migratory NEAC historically spawned along the entire west coast of Norway (i.e., all the way south to 59°N). If these southern spawners were indeed genetically NEAC, we might expect that they had left a genetic signature in present-day CC. We found little evidence for genetic impact of NEAC in the samples south of 62°N, however, and if such spawning actually took place in recent history (up until about 1930 according to Opdal 2010), the NEAC genes must have been swamped by later gene flow from the North Sea. The apparent lack — or nearly so — of NEAC genes in the south could therefore be regarded as indirect evidence for gene flow into coastal populations being generally high and that genetic introgression is recent and may still be ongoing.

Local deviations from a smooth cline

While modelling and simulations could explain the gross shape of the northward cline, conspicuous deviations were seen for the group of samples located in Vestfjorden (northward of 68°N) and for the occasional large temporal (or spatially fine-scaled) genetic differences around Godøy and in Eidsfjord and Tranøybotn (cf. Fig. 2). With respect to the former, it is possible that the release point of particles (larvae) in the oceanographic modelling did not adequately reflect actual NEAC spawning sites (Supporting Information, Fig. S1[1]), causing simulated particles to enter the Vestfjorden rather than following currents towards the Barents Sea as they should. Alternatively, it may be that NEAC larvae have poorer settling success in the Vestfjorden area or that samples in this area inadvertently have included other stock components, referred to as "fjord cod" (Sarvas and Fevolden 2005; Westgaard and Fevolden 2007; Nordeide et al. 2011; Jorde et al. 2018). The observation of large genetic shifts among years or among closely situated sites are indications that there exist genetically divergent cod populations along the coast and some of these could display migratory behaviour. Clearly, much remains to be uncovered about population structure, genomics, and behaviour of cod in coastal waters.

Management implications

An important finding of the present study is that the genetic cline in CC by itself carries very little information on the magnitude of gene flow in the CC population system. The traditional interpretation of a pattern of gradual increase in genetic differentiation with geographic distance is that such a pattern has arisen within the population system due to accumulation of genetic drift and with gene flow among the constituent populations being restricted both in distance and in amount. This interpretation stems from the predicted relationship between level of genetic divergence and number of migrants in simplified models of gene flow in linear habitats ("isolation-by-distance"; Slatkin 1993; Rousset 1997). Such predictions can be highly misleading, however, when the assumed model is wrong, as in the present case. Instead of restricted gene flow, we find that a large part of the western coast, from the southernmost locality (Ålfjord, at 59.3°N) to at least Frøya (63.4°N), is characterized by high internal connectivity. The present management border at 62°N could therefore be regarded as a conservative northern border for the southern CC component. And due to the high connectivity of this area with the northern North Sea, this component should probably be managed in coordination with offshore cod along the European continental slope and the Norwegian Trench.

Conversely, a single northern CC component above 62°N as presently defined is not supported by our findings. Instead, CC along the northern coast display marked genetic heterogeneities with at least one point around 68°N departing from the general genetic cline. As discussed above, this departure could imply that the Vestfjorden area harbour genetically differentiated, local CC population(s) of "fjord cod". Because of the proximity to important NEAC spawning and fishing areas, CC in this area may be particularly susceptible to unintentional catch in the NEAC fishery and in need for special protection. Indeed, particular localities (Henningsværboksen; Dahle et al. 2018*a*) situated in this area are already targeted for special treatment in the form of temporal fishery closures, but management measures should probably be extended to the whole area. Indeed, bycatch of CC in NEAC fisheries is a recurring problem along the coast from Møre (Johansen et al. 2018) and northwards.

Funding statement

Norges Forskningsråd: 299554/F40
Nærings- og Fiskeridepartementet
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability

All genotypes, data and software used to carry out the computer simulations are deposited at datadryad.org with doi: 10.5061/dryad.zs7h44j7t.

Acknowledgements

A major motivation for the present and related studies from the Institute of Marine Research (IMR) has been to resolve population structure in Norwegian coastal cod as an aid in providing advice for management of this declining fishery resource. This study was supported by the project SPACEFISH in the IMR program "Biologiske

prosesser", funded by the Ministry of Trade, Industry and Fisheries, and the "CoastRisk" project (No. 299554/F40) funded by the Research Council of Norway. We thank local fishermen as well as crew and technical staff on board our research vessels for collecting cod samples. Two anonymous reviewers provided valuable comments on a previous version of this paper.

References

Albretsen, J., Sperrevik, A.K., Sandvik, A.D., and Asplin, L. 2011. NorKyst-800 Report. User manual and technical descriptions. Fisken og Havet 2-2011, Institute of Marine Research, Nordnes, Bergen, Norway. pp. 1–48.

Barton, N.H. 2000. Genetic hitchhiking. Philos. Trans. R Soc. B Biol. Sci. **355**: 1553–1562. doi:10.1098/rstb.2000.0716. PMID:11127900.

Barton, N.H., and Hewitt, G.M. 1985. Analysis of hybrid zones. Annu. Rev. Ecol. Syst. **16**: 113–148. doi:10.1146/annurev.es.16.110185.000553.

Becker, R.A., and Wilks, A.R. 1993. Maps in S. AT&T Bell Laboratories Statistics Research Report 2/93.

Becker, R.A., and Wilks, A.R. 1995. Constructing a geographical database. AT&T Bell Laboratories Statistics Research Report No. 2/95.

Beldring, S., Engeland, K., Roald, L. A., Sælthun, N.R., and Voksø, A. 2003. Estimation of parameters in a distributed precipitation-runoff model for Norway. Hydrol. Earth Syst. Sci. **7**: 304–316. doi:10.5194/hess-7-304-2003.

Berg, E., and Albert, O.T. 2003. Cod in fjords and coastal waters of North Norway: distribution and variation in length and maturity at age. ICES J. Mar. Sci. **60**: 787–797. doi:10.1016/S1054-3139(03)00037-7.

Berg, P.R., Star, B., Pampoulie, C., Sodeland, M., Barth, J.M.I., Knutsen, H., et al. 2016. Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. Sci. Rep. **6**: 23246. doi:10.1038/srep23246. PMID:26983361.

Berg, P.R., Star, B., Pampoulie, C., Bradbury, I.R., Bentzen, P., Hutchings, J.A., et al. 2017. Trans-oceanic genomic divergence of Atlantic cod ecotypes is associated with large inversions. Heredity, **119**: 418–428. doi:10.1038/hdy.2017.54. PMID:28930288.

Bergstad, O.A., Jørgensen, T., and Dragesund, O. 1987. Life history and ecology of the gadoid resources of the Barents Sea. Fish. Res. **5**: 119–161. doi:10.1016/0165-7836(87)90037-3.

Bigg, G.R., Cunningham, C.W., Ottersen, G., Pogson, G.H., Wadley, M.R., and Williamson, P. 2008. Ice-age survival of Atlantic cod: agreement between palaeoecology models and genetics. Proc. R Soc. B Biol. Sci. **275**: 163–173. doi:10.1098/rspb.2007.1153. PMID:17999951.

Bradbury, I.R., Hubert, S., Higgins, B., Borza, T., Bowman, S., Paterson, I.G., et al. 2010. Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. Proc. R Soc. B Biol. Sci. **277**: 3725–3734. doi:10.1098/rspb.2010.0985. PMID:20591865.

Brander, K.M. 1995. The effect of temperature on growth of Atlantic cod (*Gadus morhua* L.). ICES J. Mar. Sci. **52**: 1–10. doi:10.1016/1054-3139(95)80010-7.

Brownrigg, R. 2018. maps: draw geographical maps. Available from https://CRAN.R-project.org/package=maps.

Case, R., Hutchinson, W.F., Hauser, L., Buehler, V., Clemmesen, C., Dahle, G., et al. 2006. Association between growth and *Pan I* genotype within Atlantic cod (*Gadus morhua*) full sibling families. Trans. Am. Fish. Soc. **135**: 241–250. doi:10.1577/T05-130.1.

Dahle, G., Johansen, T., Westgaard, J.-I., Aglen, A., and Glover, K.A. 2018*a*. Genetic management of mixed-stock fisheries "real-time": The case of the largest remaining cod fishery operating in the Atlantic in 2007–2017. Fish. Res. **205**: 77–85. doi:10.1016/j.fishres.2018.04.006.

Dahle, G., Quintela, M., Johansen, T., Westgaard, J.-I., Besnier, F., Aglen, A., et al. 2018*b*. Analysis of coastal cod (*Gadus morhua* L.) sampled on spawning sites reveals a genetic gradient throughout Norway's coastline. BMC Genet. **19**: 42. doi:10.1186/s12863-018-0625-8.

Edmonds, C.A., Lillie, A.S., and Cavalli-Sforza, L.L. 2004. Mutations arising in the wave front of an expanding population. Proc. Natl. Acad. U.S.A. **101**: 975–979. doi:10.1073/pnas.0308064100. PMID:14732681.

Efron, B., and Tibshirani, R.J. 1993. An introduction to the bootstrap. Chapman & Hall, New York.

Excoffier, L., Foll, M., and Petit, R.J. 2009. Genetic consequences of range expansions. Annu. Rev. Ecol. Evol. Syst. **40**: 481–501. doi:10.1146/annurev.ecolsys.39.110707.173414.

Fox, C.J., Taylor, M., Dickey-Collas, M., Fossum, P., Kraus, G., Rohlf, N., et al. 2008. Mapping the spawning grounds of North Sea cod (*Gadus morhua*) by direct and indirect means. Proc. R Soc. B Biol. Sci. **275**: 1543–1548. doi:10.1098/rspb.2008.0201. PMID:18397869.

Haidvogel, D.B., Arango, H., Budgell, W.P., Cornuelle, B.D., Curchitser, E., Di Lorenzo, E., et al. 2008. Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System. J. Comp. Phys. **227**: 3595–3624. doi:10.1016/j.jcp.2007.06.016.

Haldane, J.B.S. 1948. The theory of a cline. J. Genet. **48**: 277–284. doi:10.1007/BF02986626. PMID:18905075.

Hemmer-Hansen, J., Hüssy, K., Baktoft, H., Huwer, B., Bekkevold, D., Haslob, H., et al. 2019. Genetic analyses reveal complex dynamics within a marine fish management area. Evol. Appl. **12**: 830–844. doi:10.1111/eva.12760. PMID:30976313.

Huserbråten, M.B.O., Moland, E., and Albretsen, J. 2018. Cod at drift in the North Sea. Prog. Oceanogr. **167**: 116–124. doi:10.1016/j.pocean.2018.07.005.

Hutchings, J.A. 2000. Collapse and recovery of marine fishes. Nature, **406**: 882–885. doi:10.1038/35022565. PMID:10972288.

ICES. 2005. Spawning and life history information for North Atlantic cod stocks. ICES Cooperative Research Report 274. pp. 1–162.

ICES. 2019. Advice on fishing opportunities, catch, and effort: Arctic Ocean, Barents Sea, and Norwegian Sea ecoregions. Published 13 June 2019. doi:10.17895/ices.advice.4711.

Jakobsen, T. 1987. Coastal cod in northern Norway. Fish. Res. **5**: 223–234. doi:10.1016/0165-7836(87)90042-7.

Johansen, T., Westgaard, J.-I., Seliussen, B.B., Nedreaas, K., Dahle, G., Glover, K.A., et al. 2018. "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. **75**: 1093–1104. doi:10.1093/icesjms/fsx224.

Johansen, T., Besnier, F., Quintela, M., Jorde, P.E., Glover, K.A., Westgaard, J.I., et al. 2020. Genomics reveal neutral and adaptive patterns that challenge the current management regime for East Atlantic cod *Gadus morhua* L. Evol. Appl. **13**: 2673–2688. doi:10.1111/eva.13070.

Jombart, T. 2008. adegenet: an R package for the multivariate analysis of genetic markers. Bioinformatics, **24**: 1403–1405. doi:10.1093/bioinformatics/btn129. PMID:18397895.

Jombart, T., Devillard, S., and Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. **11**: 94. doi:10.1186/1471-2156-11-94. PMID:20950446.

Jorde, P.E., Kleiven, A.R., Sodeland, M., Olsen, E.M., Ferter, K., Jentoft, S., and Knutsen, H. 2018. Who is fishing on what stock: population-of-origin of individual cod (*Gadus morhua*) in commercial and recreational fisheries. ICES J. Mar. Sci. **75**: 2153–2162. doi:10.1093/icesjms/fsy080.

Kimura, M., and Weiss, G.H. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics, **49**: 561–576. doi:10.1093/genetics/49.4.561. PMID:17248204.

Kirubakaran, T.G., Grove, H., Kent, M.P., Sandve, S.R., Baranski, M., Nome, T., et al. 2016. Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. Mol. Ecol. **25**: 2130–2143. doi:10.1111/mec.13592. PMID:26923504.

Kirubakaran, T.G., Andersen, Ø., Moser, M., Árnyasi, M., McGinnity, P., Lien, S., et al. 2020. A nanopore based chromosome-level assembly representing Atlantic cod from the Celtic Sea. Genes Genom. Genet. **10**: 2903–2910. doi:10.1534/g3.120.401423.

Maynard Smith, J., and Haigh, J. 1974. The hitch-hiking effect of a favourable gene. Genet. Res. **23**: 23–35. doi:10.1017/S0016672300014634.

McIlroy, D., 2020. mapproj: map projections. Available from https://CRAN.R-project.org/package=mapproj.

Michalsen, K., Johansen, T., Subbey, S., and Beck, A. 2014. 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. **71**: 1417–1432. doi:10.1093/icesjms/fsu083.

Mork, J., Ryman, N., Ståhl, G., Utter, F., and Sundnes, G. 1985. Genetic variation in Atlantic cod (*Gadus morhua*) throughout its range. Can. J. Fish. Aquat. Sci. **42**(10): 1580–1587. doi:10.1139/f85-198.

Myksvoll, M.S., Jung, K.-M., Albretsen, J., and Sundby, S. 2014. Modelling dispersal of eggs and quantifying connectivity among Norwegian coastal cod subpopulations. ICES J. Mar. Sci. **71**: 957–969. doi:10.1093/icesjms/fst022.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nei, M., and Roychoudhury, A.K. 1974. Sampling variances of heterozygosity and genetic distance. Genetics, **76**: 379–390. doi:10.1093/genetics/76.2.379.

Nordeide, J.T. 1998. Coastal cod and north-east Arctic cod — do they mingle at the spawning grounds in Lofoten? Sarsia, **83**: 373–379. doi:10.1080/00364827.1998.10413696.

Nordeide, J.T., Johansen, S.D., Jørgensen, T.E., Karlsen, B.O., and 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. doi:10.3354/meps09232.

Olsen, E., Aanes, S., Mehl, S., Holst, J.C., Aglen, A., and Gjøsæter, H. 2010. Cod, haddock, saithe, and capelin in the Barents Sea and adjacent waters: a review of the biological values of the area. ICES J. Mar. Sci. **67**: 87–101. doi:10.1093/icesjms/fsp229.

Opdal, A.F. 2010. Fisheries change spawning ground distribution of northeast Arctic cod. Biol. Lett. **6**: 261–264. doi:10.1098/rsbl.2009.0789. PMID:19923140.

Pedersen, E.J., Thompson, P.L., Ball, R.A., Fortin, M.-J., Gouhier, T.C., Link, H., et al. 2017. Signatures of the collapse and incipient recovery of an overexploited marine ecosystem. R. Soc. Open Sci. **4**: 170215. doi:10.1098/rsos.170215.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, **155**: 945–959. PMID:10835412.

R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/.

Rollefsen, G. 1933. The otoliths of the cod. Fiskeridir. Skr. Ser. Havunders. **4**(3): 1–18.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics, **145**: 1219–1228. doi:10.1093/genetics/145.4.1219. PMID:9093870.

Ryman, N., and Utter, F. 1987. Population Genetics and Fishery Management. Washington Sea Grant Program, University of Washington Press, Seattle.

Sarvas, T.H., and Fevolden, S.E. 2005. Pantophysin (*Pan I*) locus divergence between inshore v. offshore and northern v. southern populations of Atlantic cod in the north-east Atlantic. J. Fish Biol. **67**: 444–469. doi:10.1111/j.0022-1112.2005.00738.x.

Schmidt, P.S., Serrão, E.A., Pearson, G.A., Riginos, C., Rawson, P.D., Hilbish, T.J., et al. 2008. Ecological genetics in the North Atlantic: environmental gradients and adaptation at specific loci. Ecology, **89**: S91–S107. doi:10.1890/07-1162.1. PMID:19097487.

Shchepetkin, A.F., and McWilliams, J.C. 2005. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Model. **9**: 347–404. doi:10.1016/j.ocemod.2004.08.002.

Sinclair-Waters, M., Bradbury, I.R., Morris, C.J., Lien, S., Kent, M.P., and Bentzen, P. 2018. Ancient chromosomal rearrangement associated with local adaptation of a postglacially colonized population of Atlantic Cod in the northwest Atlantic. Molecular Ecology. **27**(2): 339–351. doi:10.1111/mec.14442.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W., and Powers, J.G. 2008. A description of the advanced research WRF version 2. NCAR Tech Note NCAR/TN-468+STR. Available at https://apps.dtic.mil/dtic/tr/fulltext/u2/a487419.pdf

Skarstein, T.H., Westgaard, J.-I., and Fevolden, S.-E. 2007. Comparing microsatellite variation in north-east Atlantic cod (*Gadus morhua* L.) to genetic structuring as revealed by the pantophysin (*Pan I*) locus. J. Fish Biol. **70**(suppl. C): 271–290. doi:10.1111/j.1095-8649.2007.01456.x.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution, **47**: 264–279. doi:10.1111/j.1558-5646.1993.tb01215.x. doi:10.2307/2410134. PMID:28568097.

Sodeland, M., Jorde, P.E., Lien, S., Jentoft, S., Berg, P.R., Grove, H., et al. 2016. 'Islands of divergence' in the Atlantic cod genome represent polymorphic chromosomal rearrangements. Genome Biol. Evol. **8**: 1012–1022. doi:10.1093/gbe/evw057. PMID:26983822.

Spies, I., Spencer, P.D., and Punt, A.E. 2015. Where do we draw the line? A simulation approach for evaluating management of marine fish stocks with isolation-by-distance stock structure. Can. J. Fish. Aquat. Sci. **72**(7): 968–982. doi:10.1139/cjfas-2014-0366.

Stransky, C., Baumann, H., Fevolden, S.-E., Harbitz, A., Høie, H., Nedreaas, K.H., et al. 2008. Separation of Norwegian coastal cod and Northeast Arctic cod by outer otolith shape analysis. Fish. Res. **90**: 26–35. doi:10.1016/j.fishres.2007.09.009.

Sundby, S., and Nakken, O. 2008. Spatial shifts in spawning habitats of Arcto-Norwegian cod related to multidecadal climate oscillations and climate change. ICES J. Mar. Sci. **65**: 953–962. doi:10.1093/icesjms/fsn085.

Sundby, S., Fossum, P., Sandvik, A., Vikebø, F.B., Aglen, A., Buhl-Mortensen, L., et al. 2013. KunnskapsInnhenting Barentshavet–Lofoten–Vesterålen (KILO). *In* Fisken og Havet. pp. 3–2013. [In Norwegian.]

Waples, R.S., Punt, A.E., and Cope, J.M. 2008. Integrating genetic data into management of marine resources: how can we do it better? Fish Fish. **9**: 423–449. doi:10.1111/j.1467-2979.2008.00303.x.

Weir, B.S. 1996. Genetic data analysis II. Sinauer Associates, Sunderland, Mass.

Weir, B.S., and Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution, **38**: 1358–1370. doi:10.2307/2408641. PMID:28563791.

Westgaard, J.-I., and Fevolden, S.-E. 2007. Atlantic cod (*Gadus morhua* L.) in inner and outer coastal zones of northern Norway display divergent genetic signature at non-neutral loci. Fish. Res. **85**: 306–315. doi:10.1016/j.fishres.2007.04.001.

Wright, S. 1943. Isolation by distance. Genetics, **28**: 139–156. PMID:17247074.