Microgeographical population structure of cod *Gadus morhua* in the North Sea and west of Scotland: the role of sampling loci and individuals

Einar Eg Nielsen[1],*, **Peter John Wright[2]**, **Jakob Hemmer-Hansen[1]**,
Nina Aagaard Poulsen[1, 3], **Iain Monro Gibb[2]**, **Dorte Meldrup[1]**

[1]**Technical University of Denmark, National Institute of Aquatic Resources, DK-8600 Silkeborg, Denmark**
[2]**Fisheries Research Services Marine Laboratory, PO Box 101, 375 Victoria Road, Aberdeen, AB11 9DB, UK**
[3]**University of Aarhus, Department of Biological Sciences, Ny Munkegade, DK-8000 Aarhus C, Denmark**

ABSTRACT: We investigated potential microgeographical population structure among spatial and temporal samples of cod *Gadus morhua* L., collected in the northern North Sea and around Scotland, using microsatellite genetic markers. Results were highly dependent on the samples and microsatellite loci included. Analysis of molecular variance (AMOVA) revealed significant spatial (p = 0.04) and temporal (p = 0.02) variance when including samples of juveniles and the microsatellite Gmo 132, which is known to be subject to selection. However, neither spatial nor temporal variance components were significant (p = 0.15 and 0.23, respectively) after exclusion of juvenile samples and Gmo 132. Patterns of genetic differentiation showed a similar sensitivity to the sampling of loci. No apparent pattern was identified when only using suspected neutral microsatellites. In contrast, analysis of Gmo132 alone revealed a clear isolation of 2 samples collected at Viking and pairwise grouping of temporal adult samples from the same location. On a northeast Atlantic regional scale, inferences on local populations and patterns of population structuring were more robust to the inclusion of the microsatellite under selection. Our results demonstrate that, without cautious consideration of biased samples of individuals and loci, apparent microgeographical patterns of spatial genetic differentiation could be caused by sampling non-randomly distributed individuals or hitch-hiking selection at presumed neutral marker loci. However, while loci subject to selection may provide biased results in relation to identifying populations based on an evolutionary paradigm, they may prove valuable for separating populations on ecological time scales.

KEY WORDS: Atlantic cod · Genetic · Microsatellites · Population structure · Selection

INTRODUCTION

The application of highly variable microsatellite genetic markers has revolutionized the perception of population structure and evolution in the sea (e.g. Hoelzel et al. 2007, Zbinden et al. 2007). Microsatellite studies have challenged the assumption that the majority of marine organisms constitute homogeneous entities, following common evolutionary trajectories, due to the apparent lack of physical or environmental boundaries in the oceans. More than a decade of application of these highly variable markers for a diverse array of marine taxa has provided inferences of population structure on different geographical levels, ranging from large scale transoceanic divergence to genetic differentiation among local demes separated by a few kilometers (e.g. Kusumo et al. 2006, Mathews 2007). For marine and estuarine fish, a number of recent studies have strongly suggested 'microgeographical' genetic differentiation (e.g. Pampoulie et al. 2004, Bradbury et al. 2008). However, the existence of population structure on a geographical scale much smaller than expected from passive

dispersal of pelagic eggs and larvae and/or adult dispersal abilities is somewhat controversial (Palumbi 2003).

Inferences of small-scale local populations critically depend not only on the ability to partition statistically significant proportions of the total genetic variance among geographical localities, but also the ability to demonstrate the biological significance of the observed genetic differences (Waples 1998, Waples & Gaggiotti 2006). The levels of genetic structuring observed on a microgeographical scale in marine fish are on the very brink of detection, applying between 10 and 20 highly variable microsatellites and relatively large sample sizes (N > 50) (Waples & Gaggiotti 2006). Moreover, many potentially confounding artificial sources of genetic variance, such as non-random sampling in space and time, family structure and technical variance, can lead to erroneous inferences of local populations. For example, a population of adults may distribute randomly among a number of different spawning areas. Following reproduction and random genetic drift, the juveniles sampled at different areas may display statistically significant but biologically insignificant genetic differentiation (the Allendorf-Phelps effect; Allendorf & Phelps 1981, Waples 1998, Waples & Gaggiotti 2006). Finally, microsatellite loci affected not only by demographic factors, but also by direct or hitch-hiking selection (i.e. where the presumed neutral marker is physically linked to a gene subject to selection) have been identified in marine fish (see Nielsen et al. 2006b), thereby potentially obscuring estimates of gene flow. Accordingly, inferences of population structure in high gene flow species should be made with caution. Thus, to evaluate potentially biased sampling and to substantiate stability of the observed patterns of spatial genetic differentiation over time, previous studies have suggested the use of temporally replicated sampling (Waples 1998) and tests for selection at individual loci (Nielsen et al. 2006b).

The Atlantic cod *Gadus morhua* L. is among the most well studied marine organisms in terms of genetic population structure using microsatellites. There is a general consensus that a strong genetic differentiation exists among cod populations from the eastern and western side of the Atlantic Ocean (e.g. Bentzen et al. 1996, Nielsen et al. 2006b, O'Leary et al. 2007). On a smaller geographical scale, microsatellites have demonstrated genetic structuring among the major cod 'stocks'. For example, Ruzzante et al. (1998) found genetic differentiation among most of the large spawning aggregations along the Northwest Atlantic continental shelf. Similarly, studies by Nielsen et al. (2001) and O'Leary et al. (2007) have demonstrated a pattern of marked genetic structure among cod from most of the major spawning areas in the central and eastern parts of the North Atlantic.

Microgeographical differentiation has been inferred in a number of studies of cod (e.g. Hutchinson et al. 2003, Knutsen et al. 2003, Nielsen et al. 2003). For example, Knutsen et al. (2003) found fine-scaled geographic population structuring for cod sampled along a segment of the Norwegian Skagerrak coast. The geographical extent of these Norwegian populations has been estimated to be approximately 30 km or maybe even less (Jorde et al. 2007). Microgeographical population structure in cod is not only supported by genetic data, but also ecological data on adult site and spawning fidelity (Wright et al. 2006a) as well as egg retention (Espeland et al. 2007, Knutsen et al. 2007), juvenile segregation (Gibb et al. 2007) and nursery origin of spawning fish (Wright et al. 2006b).

Here, we investigate potential microgeographical population structure among cod sampled in the northern North Sea and around Scotland. To place the local results in a broader context and to identify major regional genetic breaks, we include samples from other populations from adjacent sea areas within the Northeast Atlantic, in particular from the North Sea region. We critically evaluate the potential role of biased sampling of individuals and loci, specifically with respect to the inclusion of juveniles and loci subject to hitch-hiking selection. Finally, we discuss our results in relation to ecological and evolutionary paradigms of population structure and evaluate the prospects for the application of genetic markers subject to selection for studies of population structure and evolution in marine organisms.

MATERIALS AND METHODS

Sample collection. We analyzed samples of gill tissue or fin-clips stored in ethanol from adult cod collected at 6 locations off the West coast of Scotland and in the northern North Sea (Table 1, Fig. 1). We aimed at attaining sample sizes above 50 individuals and used almost exclusively spawning and/or mature individuals. Temporally replicated sampling was conducted for all areas except Bergen to evaluate the stability of the patterns of genetic differentiation observed in one year. From 2 of the areas (Clyde and Shetland), we also included juvenile (0 year group) samples to investigate the potential role of non-random or family sampling for inferring microgeographical differentiation. Additional samples from adjacent areas, the Faeroe Plateau and Faeroe Bank, Central and Eastern North Sea, Baltic Sea including transitional areas (Belt Sea and Kattegat/Skagerrak), as well as an out-group sample of Northeast Arctic cod, were included for regional comparison. The additional samples have been subject to previous microsatellite analysis (see Nielsen et al. 2003, 2006b, 2007).

Table 1. *Gadus morhua*. Summary statistics for collected samples, showing geographic sampling locality and position (mean), year and month of sampling, proportion of maturing and mature individuals and total number of individuals collected. NA: not applicable; –: no data

| Geographic locality	Abbreviation	Life stage	Position (mean)	Year	Month	Proportion of maturing/ mature (%)	Total no. individuals
Northern North Sea and west of Scotland							
Clyde	CL02	Adults	55.37° N, 5.37° W	2002	Feb	0/100	45
	CL03	Adults	55.33° N, 5.26° W	2003	Mar	0/100	50
	CL01J	Juveniles	55.60° N, 4.71° W	2001	Nov–Dec	NA	50
Butt of Lewis	BL02	Adults	59.15° N, 5.93° W	2002	Mar	22/74	50
	BL03	Adults	59.58° N, 5.15° W	2003	Mar	54/42	50
Shetland	SH02	Adults	60.36° N, 1.07° W	2002	Feb–Mar	63/21	50
	SH03	Adults	60.42° N, 2.08° W	2003	Mar	6/82	50
	SH01J	Juveniles	60.31° N, 1.59° W	2001	Nov–Dec	NA	50
Moray Firth	MF02	Adults	57.85° N, 2.23° W	2002	Feb–Mar	58/4	75
	MF03	Adults	58.00° N, 3.00° W	2003	Jan	2/96	63
Viking	VI02	Adults	61.10° N, 1.20° E	2002	Mar	4/4[a]	50
	VI03	Adults	60.66° N, 1.69° E	2003	Feb	36/31	61
Bergen	BE06	Adults	60.15° N, 4.47° E	2006			48
Adjacent sea-areas							
Faeroe Plateau	FP	Adults	62.53° N, 6.18° W	2002	Apr	–	69
Faeroe Bank	FB	Adults	60.56° N, 8.52° W	2002	Apr	5/58	50
Central North Sea	CNS	Adults	55.17° N, 03.39° E	1996	Feb–Mar	18/68	82
Eastern North Sea	ENS	Adults	57.10° N, 08.20° E	1999	Feb–Mar	32/46	76
Kattegat/Skagerrak	KS	Adults	57.15° N, 11.35° E	1996	Feb–Mar	85/10	50
Belt Sea	BS	Adults	55.11° N, 10.28° E	1996	Feb–Mar	38/50	88
Eastern Baltic	EB	Adults	55.19° N, 15.54° E	1996	Apr	63/24	80
Northeast Arctic	NEAC	Adults	68.41° N, 12.82° E	1995	Aug	–	69

[a]92 % were spent individuals
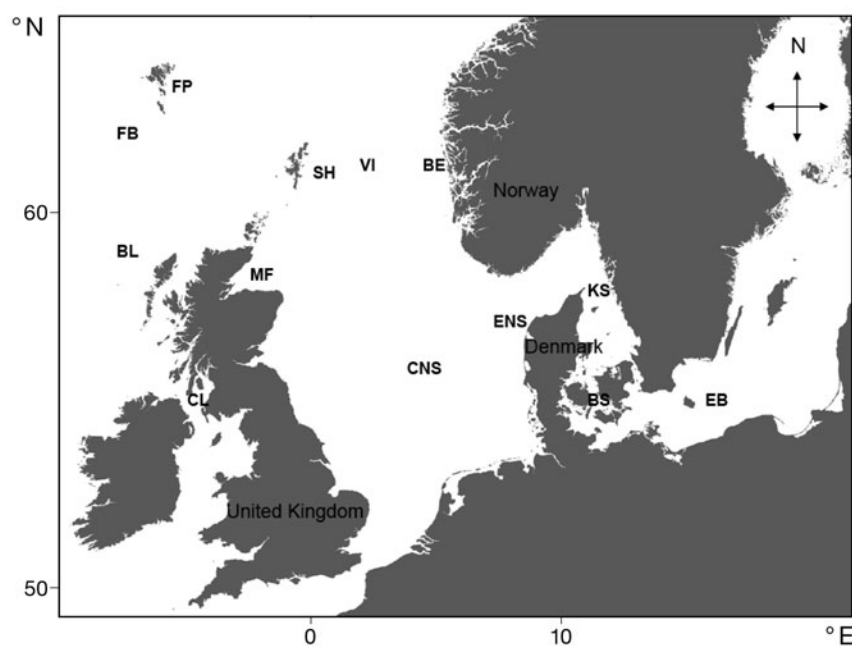
Fig. 1. *Gadus morhua*. Sampling localities of adult and juvenile cod. See Table 1 for location abbreviations and details

DNA analysis. DNA was extracted using a proteinase K/chelex procedure following Estoup et al. (1996). We PCR amplified 10 highly polymorphic di-, tri- and tetranucleotide microsatellite loci: Gmo 02 (di-), Gmo 132 (di-) (Brooker et al. 1994), Gmo 19 (tetra-), Gmo 34 (tetra-), Gmo 35 (tri-), Gmo 37 (tetra-) (Miller et al. 2000), Tch 5 (tetra-), Tch 11 (tetra-), Tch 14 (tetra-) (O'Reilly et al. 2000) and GADM 1 (di-) (Hutchinson et al. 2001). Hitch-hiking selection has previously been demonstrated for the locus Gmo 132 (Nielsen et al. 2006b) and has also been suggested for Gmo 34 and Gmo 37 (e.g. Nielsen et al. 2006b, Skarstein et al. 2007, Westgaard & Fevolden 2007). The microsatellites were analyzed on a Basestation51 automated sequencer (MJ Research) according to manufacturer's recommendations. In all gel runs, individuals with known geno-

types were included as positive controls. Individuals were genotyped using the Cartographer Sequencing and Genotyping Analysis Software (MJ Research).

Statistical analysis. Tests for deviations from Hardy-Weinberg equilibrium for individual microsatellite loci in each sample were conducted using the program FSTAT (Goudet 1995). FSTAT was also used to provide unbiased estimates of pairwise F_{ST} values (Wright's fixation index, a measure of genetic differentiation) for microsatellites following Weir & Cockerham (1984). Heterogeneity in allele frequencies was tested by permuting alleles among samples 10 000 times, generating contingency tables of alleles within samples and using the log-likelihood test statistic G (Goudet et al. 1996). To visualize genetic relationships among spatial and temporal samples around Scotland and from the northern North Sea, as well as elucidating regional patterns, we used VISTA 5.6.3 (Young 1996) for multidimensional scaling (MDS) analysis of Nei's D_A distances (Nei et al. 1983). Spatial and temporal molecular variance components and their significance (10 000 permutations) for samples from the northern North Sea and west of Scotland were estimated using an analysis of molecular variance (AMOVA) approach (Excoffier et al. 1992) with the program Arlequin1.1 (Schneider et al. 2000). To identify geographic patterns of genetic differentiation among regional cod samples, including both the samples from the northern North Sea and west of Scotland as well as the additional samples from adjacent areas, we applied a landscape genetics approach with the program BARRIER (Manni et al. 2004). This approach combines genetic and geographic information to identify the location of major genetic breaks. Only barriers supported by significant pairwise neutral F_{ST} values were recorded. To identify outlier loci likely to be influenced by selection, we used the simulation-based approach by Beaumont & Nichols (1996) implemented in the LOSITAN Selection detection Workbench (available at popgen.eu/soft/lositan/). Based on mean 'neutral' F_{ST} values from the data, we simulated distributions of loci (10 000 permutations) with associated F_{ST} values and heterozygosities. This was achieved by running an initial simulation to identify and remove outliers from the neutral estimates. Outlier loci in the real dataset were subsequently identified by comparison with the simulated neutral distributions, and the probabilities of the simulated F_{ST} values being smaller than the sample locus F_{ST} values were given. For comparison, simulations were conducted using both the Stepwise Mutation Model (SMM) and the Infinite Alleles Model. As also reported by Beaumont & Nichols (1996), no qualitative differences were observed between the 2 mutation models. We attempted identification of full-sibs in our juvenile samples from Shetland and Clyde by using

the program Colony 1.2 (Wang 2004). The program uses a maximum likelihood method to assign individuals to family groups. We tested whether the most likely number of full-sib families identified in the 2 samples of juveniles deviated from the distributions in a sample of adults from Shetland (SH02) and 2 simulated samples of unrelated individuals based on the allele frequencies of the Shetland population. This was done using the program HYBRIDLAB (Nielsen et al. 2006b).

For all analyses related to microgeographical population structure in the northern North Sea and west of Scotland, spatial, temporal and juvenile samples were included individually. For analyses of regional population structure, temporal samples of adults from Clyde, Butt of Lewis, Shetland, Moray Firth and Viking were pooled for convenience and samples of juveniles were excluded.

RESULTS

No significant deviations from Hardy-Weinberg equilibrium were found for individual loci in any sample following sequential Bonferroni correction for multiple testing (results not shown) (Rice 1989).

Northern North Sea and west of Scotland samples

The overall F_{ST} estimate including juvenile samples and all 10 loci was low (0.002), but highly significant heterogeneity in allele frequencies was observed (p < 0.001). None of the pairwise F_{ST} values among samples were significant after sequential Bonferroni correction (Table A1 in Appendix 1). Still, levels of genetic differentiation were relatively high, ranging between –0.0013 and 0.0082, and 20 of 78 individual p-values were below 0.05, of which 9 were associated with comparisons including Viking samples. When Gmo 132 was exluded, the pairwise F_{ST} values were generally reduced, although the range was similar (F_{ST} values ranging between –0.0018 and 0.0083) (Table A1). Only 8 individual p-values were below 0.05 and none of them included Viking samples. Pairwise F_{ST} values for Gmo 132 alone provided elevated levels of genetic differentiation (F_{ST} values ranging between –0.0071 and 0.0505, all values above 0.01 involved Viking samples) (Table A1). Eight values were significant, all including Viking samples, and 28 individual p-values were below 0.05.

AMOVA revealed that both spatial and temporal variance components were significant when both juvenile samples and the locus Gmo 132 were included (Table 2). When excluding Gmo 132, the spatial variance was reduced and no longer significant. Similarly, when juvenile samples were excluded, the temporal

Table 2. *Gadus morhua*. Analysis of molecular variance (AMOVA) among samples collected around Scotland and in the northern North Sea. Estimates of spatial (V_s) and temporal (V_t) variance proportions (and their significance) are provided for 4 different data combinations of inclusion (+) and exclusion (–) of juvenile samples and the microsatellite locus Gmo132. *Significant at the 5% level

Data combination		V_s	p	V_t	p
Juveniles	+	0.14	0.04*	0.13	0.02*
Gmo132	+				
Juveniles	+	0.05	0.20	0.14	0.04*
Gmo132	–				
Juveniles	–	0.17	0.02*	0.10	0.18
Gmo132	+				
Juveniles	–	0.07	0.15	0.09	0.23
Gmo132	–				

variance was reduced and became non-significant. When both juvenile samples and Gmo 132 were excluded, neither spatial nor temporal variance components were significant.

The MDS plot for the northern North Sea and west of Scotland samples including all 10 loci (Fig. 2a) did not reveal any clear overall spatial patterns corresponding to the geography. However, the 2 samples from Viking appeared to group together, slightly separating from the other samples along both Dimensions 1 and 2 (explaining 31 and 24% of the variance, respectively) (Fig. 2a). When Gmo 132 was excluded from the analysis, the apparent isolation of the Viking samples disappeared (Dimensions 1 and 2 explained 33 and 16% of the variance, respectively) (Fig. 2b). The application of only Gmo 132 provided a clear pattern of isolation of the Viking samples, but also a relatively clear grouping of temporal samples collected at the same location, except for the juvenile samples (Dimensions 1 and 2 explained 63 and 23% of the variance, respectively) (Fig. 2c).

The simulation-based test for selection in the northern North Sea and west of Scotland adult samples (Fig. 3a) identified Gmo 132 as an outlier locus (simulated F_{ST} smaller than Gmo 132 F_{ST}, p = 0.99), likely to be influenced by 'positive' (directional) selection. When all samples were included, Gmo 34 appeared also to be influenced by positive selection, while Gmo 02 and Tch 5 may be influenced by balancing selection (Fig. 3b).

The assignment of individuals to full-sib groups in the 2 samples of juveniles from Shetland and Clyde provided the highest likelihood for 40 and 41 families in 50 individuals (9 and 10 pairs of full sibs, respectively). However, the number of likely families did not deviate significantly (p = 0.977) from the number of families identified in the adult sample from Shetland (39 families) or the 2 simulated samples of unrelated individuals (39 and 37 families, respectively).
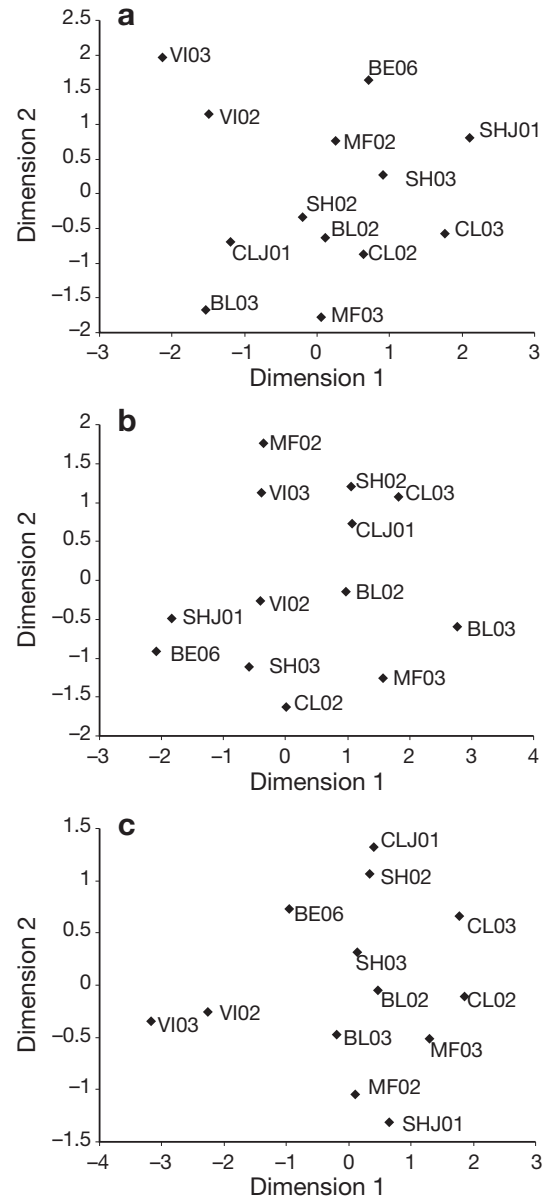

Fig. 2. *Gadus morhua.* Multidimensional scaling (MDS) plot of Nei's D_A distances (Nei et al. 1983) among spatial and temporal samples of adults and juveniles collected in the northern North Sea and west of Scotland, (a) all 10 loci, (b) 9 loci (excluding Gmo 132) and (c) Gmo 132 only. Abbreviations are listed in Table 1

Regional samples

When adult cod samples from adjacent areas were included, the overall regional F_{ST} for all 10 loci was 0.012 and highly significant (p < 0.0001). All pairwise comparisons including Eastern Baltic and Northeast Arctic were significant following Bonferroni correction (Table A2). A number of pairwise comparisons including Viking, Faeroe Plateau and Faeroe Bank were also
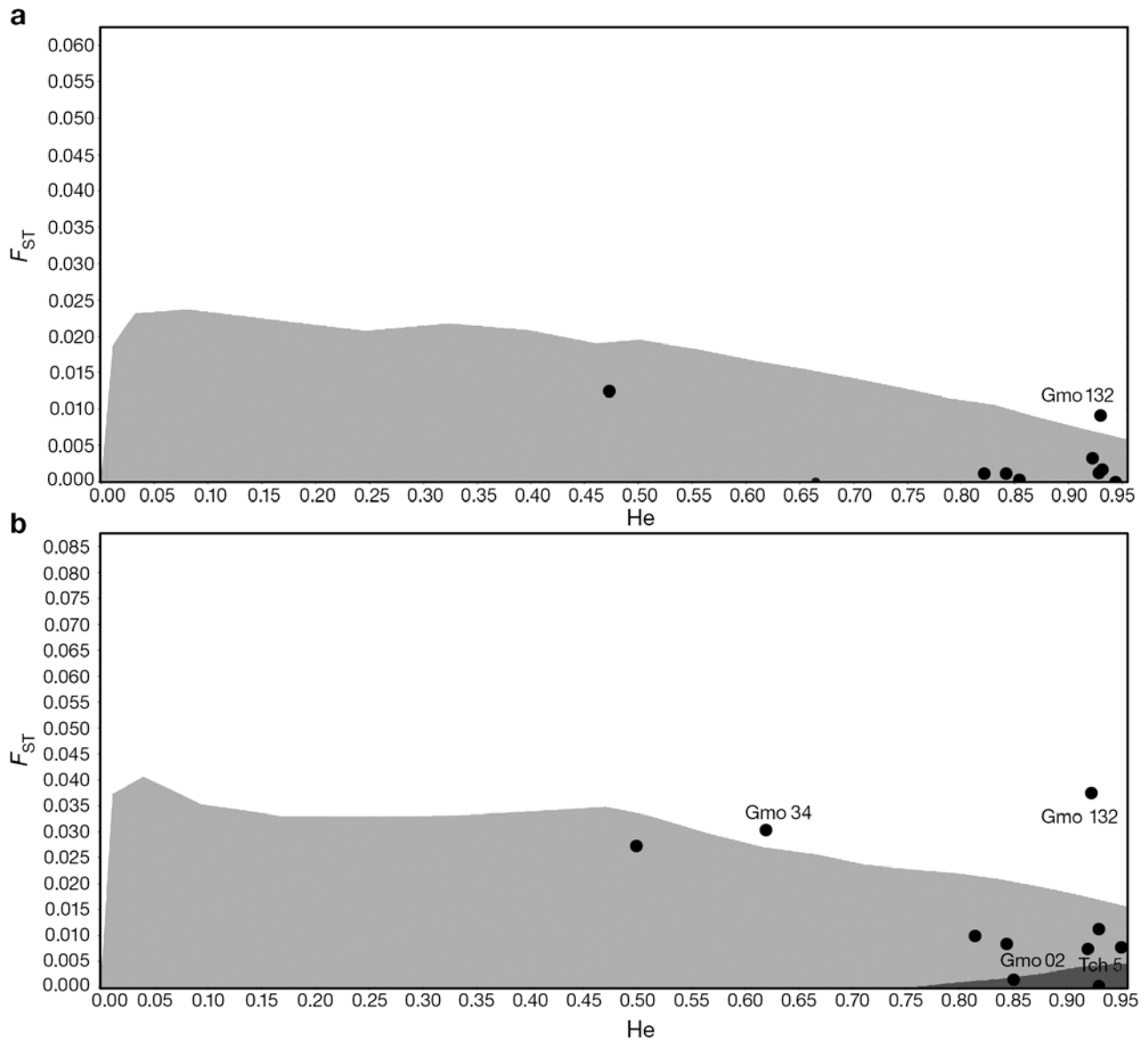
Fig. 3. *Gadus morhua.* Plot of F_{ST} versus heterozygosity (He) to identify loci (●) subject to hitch-hiking selection. Light grey area: 95% CI for neutrality. Only outliers identified as candidates for directional selection (white area, Gmo 34 and Gmo 132) or balancing selection (dark grey area, Gmo 02 and Tch 5) are labelled. Analysis is based on (a) northern North Sea and west of Scotland samples, (b) samples from the North Sea region and adjacent sea areas

significant; in total, 67 of 91 individual p-values were significant, i.e. <0.05. After exclusion of Gmo 132 (Table A2 in Appendix 1), all pairwise comparisons including Eastern Baltic and Northeast Arctic were still significant, while pairwise comparisons including Viking, Faeroe Plateau and Faeroe Bank were no longer significant, except for the comparison between Faeroe Bank and Belt Sea. Still, a relatively high proportion of the remaining individual p-values were <0.05 (22/65). When specifically testing the hypothesis of differentiation between Faeroe Plateau and Faeroe Bank, the F_{ST} value was relatively high (0.0077) and

with strong statistical support (p = 0.00385). Pairwise comparisons for Gmo 132 only (Table A2) yielded a high number of significant tests including all comparisons with Eastern Baltic and Northeast Arctic, and a high number of Viking comparisons. Overall, 59 individual tests had p-values below 0.05.

The MDS plot of all 10 loci (Fig. 4a) clearly identified the Eastern Baltic and the Northeast Arctic as highly distinct, while the remaining samples grouped together on this regional scale without any clear geographical signal (Dimensions 1 and 2 explaining 52 and 26% of the variance, respectively). Exclusion of
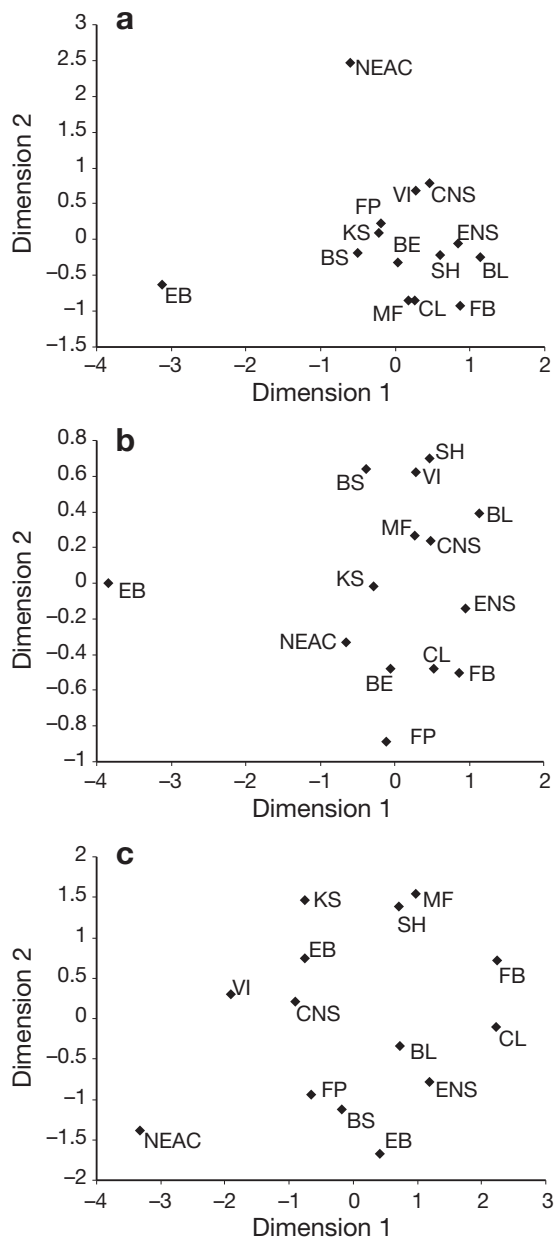
Fig. 4. *Gadus morhua.* Multidimensional scaling (MDS) plot of Nei's D$_A$ distances (Nei et al. 1983) among samples of adults from the North Sea region and adjacent sea areas, (a) for all 10 loci, (b) for 9 loci (excluding Gmo 132) and (c) for Gmo 132 only. Abbreviations are listed in Table 1

Gmo 132 reduced the apparent distinctness of Northeast Arctic along both dimensions (Dimensions 1 and 2 explaining 74 and 13% of the variance, respectively), while no other clear changes of pattern were apparent (Fig. 4b), except that the Faeroe Plateau sample appeared somewhat isolated along Dimension 2. Finally, the plot of Gmo 132 alone (Fig. 4c) identified the Northeast Arctic as very distinct, while no clear

correspondences with geography for the other samples were apparent (Dimensions 1 and 2 explaining 60 and 37% of the variance, respectively).

BARRIER analysis identified 3 apparent genetic breaks, supported by the pairwise F_{ST} analysis, between the single population samples Eastern Baltic, Northeast Arctic and Faeroe Plateau and all other samples, irrespective of whether Gmo 132 was included or not. All 3 barriers were supported by 5 or more loci. A fourth barrier separated the Scottish samples Moray Firth and Shetland from the North Sea samples Central North Sea and Viking only when Gmo 132 was included.

The simulation-based identification of outlier loci (Fig. 3b) revealed that Gmo 132 is most likely affected by positive (directional) selection in this region of the species distribution (simulated F_{ST} smaller than Gmo 132 F_{ST}, p = 0.99). However, Gmo 34 also appeared to be subject to directional selection (p = 0.98), and Gmo 02 and Tch 5 appeared to be under stabilizing selection (p = 0.015 and p = 0.001, respectively).

DISCUSSION

Results from the present study show that careful evaluation of included loci and samples can be of paramount importance for drawing conclusions on genetic population structure on a microgeographical scale in marine fish. Without cautious consideration of potentially biased samples due to the inclusion of juveniles, apparent patterns of spatial genetic differentiation could be caused by sampling non-randomly distributed individuals, clearly showing that the Allendorf-Phelps effect (Allendorf & Phelps 1981) should also be considered in marine fish. However, in our case, the occurrence of full-sib families could not be verified. This is most likely an effect of low statistical power or a result of a high proportion of related individuals in the juvenile samples that are not full-, but half-sibs. Likewise, the inclusion of microsatellite loci, like Gmo 132, subject to direct or hitch-hiking selection can lead to erroneous inferences on genetic population structure and the degree of demographic and evolutionary isolation of local populations. At present, there is no information on genes linked to Gmo 132. Sequencing efforts in our lab of associated genomic regions have not revealed any proximate genes. Likewise, no links between environmental drivers and selection at this locus have been established (see Nielsen et al. 2006b). However, loci subject to selection can still provide useful information on microgeographical population separation along ecological dimensions and time scales, as evidenced by the clear local geographical patterns of differentiation identified by Gmo 132.

Defining local populations

When evaluating microgeographical population structure in cod from genetic data, it is important to state the population definition to which the interpretation of data relates (see Waples & Gaggiotti 2006). In the following sections, we have adopted the least strict criterion of Waples & Gaggiotti (2006) for defining a population from an evolutionary paradigm, EV4: $N_e m$ < 25 (island model of migration), where N_e is the effective population size and m the migration rate. The rationale behind choosing this value is somewhat arbitrary; however, in practice it reflects the statistical power for detecting departures from panmixia with 10 microsatellite loci and sample sizes of 50 (see Waples & Gaggiotti 2006). Similarly, we have adopted the criterion EC1: m < 0.1 for defining a population from an ecological paradigm.

Population structure of cod in the northern North Sea and west of Scotland

Using the above mentioned criteria, we do not find it likely that the samples of cod collected from the northern North Sea and west of Scotland in the present study represent local populations following independent evolutionary trajectories. This is supported by different lines of evidence. Firstly, there is no significant spatial genetic variance, and the spatial variance does not appear to be larger than the temporal variance when using only microsatellites, conforming to neutral expectations. This is also corroborated by the apparent lack of a spatial pattern, where population samples in time and space do not seem to group according to their geographic locality. Further, there are no significant pairwise tests of genetic differentiation after adjusting to table-wide levels of significance, and the few low p-values seem to appear sporadically among samples, and hence are not associated with particular geographical localities. Finally, a subsequent analysis using the coalescent based program MIGRATE (Beerli & Felsenstein 1999) showed that estimates of migration among demes were almost exclusively between 20 and 60 (results not shown), suggesting high levels of historic gene flow.

In contrast, the evidence from the locus under selection, Gmo 132, suggests that at least cod from Viking represent an isolated cod population. Local isolation is evidenced by many significant pairwise tests for genetic differentiation and the apparent grouping of temporal samples. However, there are a number of potential explanations for this pattern of genetic differentiation. At one end of the scale, the isolation could be an artefact involving no reproductive isolation at all. If eggs and larvae from various spawning grounds are dispersed, mixed and subsequently retained in nursery areas with different ecological selection regimes each generation, then different juvenile mortalities associated with different Gmo 132 genotypes could create the observed pattern. However, the pattern would not reflect reproductive isolation, but rather selection at some life stage(s) followed by random mating each generation. At the other end of the scale, genetic differentiation at Gmo 132 may represent true temporally stable reproductive isolation, and therefore indicate that the number of migrants is much lower than the 25 used in our evolutionary definition of populations. This could be the case if populations are not at migration drift equilibrium. This would occur if migration rates among natural populations are reduced over ecological time scales, which is particularly relevant for overexploited species like cod, where recent demographic changes have been observed (Dulvy et al. 2006). Accordingly, genetically effective migration rates for neutrally evolving loci would be overestimated using an island model of migration. Even for fairly large migration rates and, for marine fish, small population sizes ranging in the 100s, the number of generations before F_{ST} approaches equilibrium could substantially exceed ecological time scales relevant for population management and conservation (see Whitlock & McCauley 1999 for a thorough discussion on inferring migration from F_{ST}). Estimates of migration among putative populations of cod in the area based on traditional tagging, data storage tags, otolith shape and microchemistry suggest limited adult straying among spawning areas and site attachment of juveniles in nursery areas (Wright et al. 2006b, Gibb et al. 2007). However, even limited effective migration rates on an ecological timescale would be sufficient to homogenize populations genetically on an evolutionary timescale (e.g. Waples 1998). Furthermore, although there is evidence of local selection at Gmo 132 in the Viking samples, there is little additional evidence of population structuring on a microgeographical scale at this or other loci. Consequently, if we combine the ecological and genetic data on cod in the northern North Sea and west of Scotland, the population structure which is, at present, best supported by both lines of evidence is an ecological and not an evolutionary metapopulation.

Population structure of cod in the North Sea and adjacent areas

In general, the regional pattern of genetic differentiation was less sensitive to the inclusion of Gmo 132. This was not surprising given the generally larger

genetic differentiation and associated support from a number of neutral loci. Overall, there is little support for evolutionary significant population structuring for the sampled locations within the North Sea region, including the northern North Sea and west of Scotland ecological metapopulation. Samples from different areas cluster together irrespective of their geographical proximity, and small and non-significant pairwise F_{ST} values are apparent. This is in contrast to the study by Hutchinson et al. (2001), which reported marked genetic structuring within the North Sea region. However, their study also included Gmo 132, and their conclusions of genetic differentiation within the central and northern North Sea may have been influenced by this locus. Further, they found the clearest patterns of genetic differentiation towards the southern North Sea/English Channel, an area not sampled within the present study. In the present study, the North Sea ecological metapopulation is bordered to the east by the Baltic Sea cod populations, including the Belt Sea transition area. Evidence comes from the major genetic breaks identified, the pattern of population structure and consistent significant pairwise F_{ST} values, even when only applying loci expected to conform to neutral expectations. Towards the north, we also observed evolutionary significant structuring, illustrated by the clear genetic barrier towards Northeast Arctic cod. This genetic break may be caused by an apparent gap in our sampling scheme, as we did not include any samples of Norwegian coastal cod (NCC) along the Norwegian coast. Using a smaller but overlapping set of microsatellites, Skarstein et al. (2007) also found a break between North Sea samples, including Bergen and NCC samples collected further to the north. To a lesser degree, but still on an evolutionary scale, northern North Sea and west of Scotland samples from the present study were separated from Faeroe Plateau and Faeroe Bank cod populations. Although a significant barrier and the pattern of population differentiation clearly only supported the Faeroe Plateau as an evolutionary distinct unit (see also Pampoulie et al. 2008), the majority of pairwise tests for differentiation also supported the isolation of the Faeroe Bank population. The pairwise F_{ST} between the 2 Faeroe samples is also marginally significant, which may support the occurrence of 2 distinct populations in Faeroese waters.

Microgeographical population structure in cod

The present study suggests that evolutionary significant microgeographical population structuring may not be a common phenomenon in cod, at least within the North Sea region, while separation on ecological time scales is strongly suggested by loci subject to

selection and ecological data. We do not, however, argue that previous evidence of small-scale genetic structure within the North Sea region, as suggested by Hutchinson et al. (2001), Knutsen et al. (2003) and Jorde et al. (2007), is erroneous. First of all, we did not conduct detailed sampling at the same geographical locations; therefore, evolutionary separated units of cod may still be found within the region. In addition, we did not attempt to distinguish between ecological and evolutionary populations. On the other hand, although tests for cohort stability within samples were conducted (i.e. whether different cohorts from the same locality were genetically more similar than the same cohort from different localities, see Jorde et al. 2007), previous studies did not explicitly evaluate the temporal stability of the observed pattern of population structure by conducting repeated sampling in different years. Similarly, no tests for outlier loci subject to selection were conducted, although both locus Gmo 132 and other loci suspected to be under selection in cod were included. Accordingly, the apparent stability of allele frequencies observed among cohorts sampled at one location could, at least partly, be an effect of similar selection regimes in time and space without necessarily having to invoke reproductive isolation. Consequently, we call for further investigations of microgeographical population structure of cod in the North Sea region at both evolutionary and ecological time scales, but similar studies for high gene flow marine species in general are warranted. Local-scale studies should be conducted using large samples of mature adult individuals collected at the spawning areas to minimize non-random sampling. Temporal sampling should be conducted, preferentially with longer time intervals than in the present study to avoid potential unintentional sampling of the same dominating cohorts. Another option is to use historical otolith collections (e.g. Hutchinson et al. 2003, Poulsen et al. 2006) to investigate temporal stability over several decades. Furthermore, we recommend using a high number of markers to allow rigorous testing of selective neutrality. However, we still encourage including genetic markers under selection, since they can provide important evidence of ecological isolation (see Westgaard & Fevolden 2007). However, a 'candidate gene approach' using markers associated with genes of known function involved in physiological processes could in general prove to be better than using coincidentally identified markers associated with unknown genes or genes of unknown function. This would not only allow recognition of population structure, but also identification of genetically based adaptations to local environmental conditions. For example, Hemmer-Hansen et al. (2007) showed that genetic differentiation at an insertion–deletion associated with the heat-

shock gene Hsc70 in the European flounder vastly exceeded neutral genetic differentiation, thus suggesting local adaptation. Likewise, a combination of neutral and selected gene markers can be used to evaluate genetically based adaptation over time in response to climate change (Nielsen et al. 2007). Hundreds of such gene-associated markers have already been developed for cod (see Moen et al. 2008) and will most likely soon be available for a number of other important marine fish species. We expect that the future exploitation of 'genes that matter' in population genetics of marine fishes will dramatically increase our knowledge and understanding of evolution in the sea, thereby enabling us to better set priorities for biodiversity conservation and fisheries management on a variety of geographical scales.

LITERATURE CITED

Allendorf FW, Phelps SR (1981) Use of allelic frequencies to describe population structure. Can J Fish Aquat Sci 38: 1507–1514

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc R Soc Lond B 263:1619–1626

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics 152: 763–773

Bentzen P, Taggart CT, Ruzzante DE, Cook D (1996) Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. Can J Fish Aquat Sci 53:2706–2721

Bradbury IR, Campana SE, Bentzen P (2008) Estimating contemporary early life-history dispersal in estuarine fish: integrating molecular and otolith elemental approaches. Mol Ecol 17:1438–1450

Brooker AL, Cook D, Bentzen P, Wright JM, Doyle RW (1994) Organization of microsatellites differs between mammals and cold-water teleost fishes. Can J Fish Aquat Sci 51: 1959–1966

Dulvy NK, Jennings S, Rogers SI, Maxwell DL (2006) Threat and decline in fishes: an indicator of marine biodiversity. Can J Fish Aquat Sci 63:1267–1275

Espeland SH, Gundersen AF, Olsen EM, Knutsen H, Gjøsæter J, Stenseth NC (2007) Home range and elevated egg densities within an inshore spawning ground of coastal cod. ICES J Mar Sci 64:920–928

Estoup A, Largiader CR, Perrot E, Chourrout D (1996) Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. Mol Mar Biol Biotechnol 5:295–298

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Gibb FM, Gibb IM, Wright PJ (2007) Isolation of Atlantic cod (*Gadus morhua*) nursery areas. Mar Biol 151:1185–1194

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Goudet J, Raymond M, Demeeus T, Rousset F (1996) Testing differentiation in diploid populations. Genetics 144: 1933–1940

Hemmer-Hansen J, Nielsen EE, Frydenberg J, Loeschcke V (2007) Adaptive divergence in a high gene flow environment: *Hsc70* variation in the European flounder (*Platicthys flesus* L.). Heredity 99:592–600

Hoelzel AR, Hey J, Dahlheim ME, Nicholson C, Burkanov V, Black N (2007) Evolution of population structure in a highly social top predator, the killer whale. Mol Biol Evol 24:1407–1415

Hutchinson WF, Carvalho GR, Rogers SI (2001) Marked genetic structuring in localised spawning populations of cod *Gadus morhua* in the North Sea and adjoining waters, as revealed by microsatellites. Mar Ecol Prog Ser 223: 251–260

Hutchinson WF, Van Oosterhou C, Rogers SI, Carvalho GR (2003) Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*). Proc R Soc Lond B 270:2125–2132

Jorde PE, Knutsen H, Espeland SH, Stenseth NC (2007) Spatial scale of genetic structuring in coastal cod *Gadus morhua* and geographic extent of local populations. Mar Ecol Prog Ser 343:229–237

Knutsen H, Jorde PE, André C, Stenseth NC (2003) Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. Mol Ecol 12: 385–394

Knutsen H, Olsen EM, Ciannelli L, Espeland SH and others (2007) Egg distribution, bottom topography and small-scale cod population structure in a coastal marine system. Mar Ecol Prog Ser 333:249–255

Kusumo HT, Pfister CA, Wooton JT (2006) Small-scale genetic structure in the sea palm Postelsia *palmaeformis* Ruprecht (*Phaeophyceae*). Mar Biol 149:731–742

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. Hum Biol 76:173–190

Mathews LM (2007) Evidence for restricted gene flow over small spatial scales in a marine snapping shrimp *Alpheus angulosus*. Mar Biol 152:645–655

Miller KM, Le KD, Beacham TD (2000) Development of tri- and tetranucleotide repeat microsatellite loci in Atlantic cod (*Gadus morhua*). Mol Ecol 9:238–239

Moen T, Hayes B, Nilsen F, Delghandi M and others (2008) Identification and characterisation of novel SNP markers in Atlantic cod and their application in genetic mapping and population genetics. BMC Genet 9:18

Nei M, Tajima F, Tatano Y (1983) Accuracy of estimated phylogenetic trees from molecular data. J Mol Evol 19:153–170

Nielsen EE, Hansen MM, Schmidt C, Meldrup D, Grønkjær P (2001) Determining the population of origin of individual cod in the Northeast Atlantic. Nature 413:272

Nielsen EE, Hansen MM, Ruzzante DE, Meldrup D, Grønkjær P (2003) Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea, revealed by individual admixture analysis. Mol Ecol 12:1497–1508

Nielsen EE, Bach LA, Kotlicki P (2006a) HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. Mol Ecol Notes 6:971–973

Nielsen EE, Hansen MM, Meldrup D (2006b) Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): Implications for inferring population structure in non-model organisms. Mol Ecol 15:3219–3229

Nielsen EE, MacKenzie BR, Magnussen E, Meldrup D (2007) Historical analysis of Pan I in cod (*Gadus morhua*); no evidence of a recent selective sweep in populations from the southern part of the species distribution. Can J Fish Aquat Sci 64:1448–1455

O'Leary DB, Coughlan J, Dillane E, McCarthy TV, Cross TF (2007) Microsatellite variation in cod *Gadus morhua* throughout its geographic range. J Fish Biol 70:310–335

O'Reilly PT, Canino MF, Bailey KM, Bentzen P (2000) Isolation of twenty low stutter di- and tetra nucleotide microsatellites for population analysis of walleye pollock and other gadoids. J Fish Biol 56:1074–1086

Palumbi SR (2003) Population genetics, demographic connectivity and the design of marine protected areas. Ecol Appl 13:146–158

Pampoulie C, Gysels ES, Maes GE, Hellemans B, Leentjes V, Jones AG, Volckaert FAM (2004) Evidence for fine-scale genetic structure and estuarine colonisation in a potential high gene flow marine goby (*Pomatoschistus minutus*). Heredity 92:434–445

Pampoulie C, Steingrund P, Stefánsson MÖ, Daníelsdóttir AK (2008) Genetic divergence among East Icelandic and Faroese populations of Atlantic cod provides evidence for historical imprints at neutral and non-neutral markers. ICES J Mar Sci 65:65–71

Poulsen NA, Nielsen EE, Schierup MH, Loeschcke V, Grønkjær P (2006) Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*). Mol Ecol 15:321–331

Rice WR (1989) Analysing tables of statistical tests. Evolution 43:223–225

Ruzzante DE, Taggart CT, Cook D (1998) A nuclear DNA basis for shelf- and bank-scale population structure in Northwest Atlantic cod (*Gadus morhua*): Labrador to Georges Bank. Mol Ecol 7:1663–1680

Schneider S, Kueffer JM, Roessli D, Excoffier L (2000) Arlequin, version 2.000. Genetics and Biometry Lab, Department of Anthropology, University of Geneva, Switzerland, available at www.anthropologie.unige.ch/arlequin/

Skarstein TH, Westgaard JI, Fevolden SE (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 70:271–290

Wang J (2004) Sibship reconstruction from genetic data with typing errors. Genetics 166:1963–1979

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J Hered 89:438–450

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol 15:1419–1439

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Westgaard JI, Fevolden SE (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

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. Heredity 82:117–125

Wright PJ, Galley E, Gibb IM, Neat FC (2006a) Fidelity of adult cod to spawning grounds in Scottish waters. Fish Res 77:148–158

Wright PJ, Neat FC, Gibb FM, Gibb IM, Thordason H (2006b) Evidence for metapopulation structuring in cod from the west of Scotland and North Sea. J Fish Biol 69:181–199

Young FW (1996) ViSta: the visual statistics system. Research Memorandum 94-1(b) (2nd edn). L. L. Thursone Psychometric Laboratory, University of North Carolina, Chapel Hill

Zbinden JA, Largiadèr CR, Leippert F, Margaritoulis D, Arlettaz R (2007) High frequency of multiple paternity in the largest rookery of Mediterranean loggerhead sea turtles. Mol Ecol 16:3703–3711

Appendix 1. Pairwise genetic differentiation among cod samples

Table A1. Estimates of pairwise genetic differentiation (F_{ST}) (below diagonal) and p-values (above diagonal) for all 10 microsatellites, 9 microsatellites (Gmo 132 excluded) and only the microsatellite Gmo 132, among spatial and temporal cod samples collected in the northern North Sea and west of Scotland. See Table 1 for sample location abbreviations

Sample	CL02	CL03	CL01J	BL02	BL03	SH02	SH03	SH01J	MF02	MF03	VI02	VI03	BE06
All 10 microsatellites													
CL02	–	0.21923	0.26667	0.03205	0.00962	0.87051	0.51026	0.13718	0.13141	0.33910	0.06282	0.00513	0.19231
CL03	0.0006	–	0.57244	0.22308	0.01795	0.94808	0.82115	0.77628	0.27949	0.32115	0.14231	0.00128	0.30833
CL01J	0.0008	0.0012	–	0.44872	0.37756	0.86282	0.38141	0.13141	0.55256	0.23397	0.63462	0.08333	0.20128
BL02	0.0020	0.0014	0.0008	–	0.15192	0.90641	0.16154	0.38205	0.02179	0.01538	0.14423	0.00705	0.40769
BL03	0.0043	0.0051	0.0017	0.0018	–	0.34231	0.00705	0.00192	0.11603	0.06859	0.22372	0.00192	0.00128
SH02	–0.0008	–0.0001	–0.0007	–0.0002	0.0035	–	0.51923	0.42115	0.89808	0.95641	0.55833	0.13141	0.61474
SH03	–0.0004	–0.0013	0.0019	0.0003	0.0036	0.0009	–	0.62564	0.75705	0.26667	0.13397	0.11154	0.75321
SH01J	0.0031	0.0009	0.0051	0.0014	0.0074	0.0045	0.0002	–	0.20064	0.11154	0.08718	0.00385	0.04615
MF02	0.0033	0.0003	0.0011	0.0015	0.0031	–0.0002	–0.0003	0.0007	–	0.03910	0.09936	0.00513	0.10577
MF03	0.0004	0.0023	0.0015	0.0009	0.0005	–0.0010	0.0002	0.0039	0.0012	–	0.00897	0.00577	0.04231
VI02	0.0025	0.0037	0.0006	0.0025	0.0032	0.0018	0.0021	0.0043	0.0016	0.0053	–	0.75192	0.32821
VI03	0.0059	0.0075	0.0047	0.0046	0.0057	0.0043	0.0037	0.0075	0.0040	0.0053	–0.0008	–	0.01795
BE06	0.0022	0.023	0.0028	0.0015	0.0082	0.0034	–0.0007	0.0025	0.0020	0.0052	0.0018	0.0040	–

Table A1 (continued)

Sample	CL02	CL03	CL01J	BL02	BL03	SH02	SH03	SH01J	MF02	MF03	VI02	VI03	BE06
9 microsatellites (Gmo 132 excluded)													
CL02	–	0.20577	0.24936	0.26090	0.25833	0.91987	0.65385	0.34551	0.08590	0.86474	0.67308	0.69551	0.71346
CL03	0.0012	–	0.04038	0.03654	0.00321	0.26923	0.72436	0.81410	0.01923	0.08462	0.32115	0.07628	0.16218
CL01J	0.0020	0.0026	–	0.24551	0.49103	0.43333	0.01923	0.08205	0.13846	0.13397	0.79936	0.35833	0.11731
BL02	0.0014	0.0019	0.0015	–	0.56667	0.78654	0.22949	0.14487	0.01218	0.30385	0.09551	0.13269	0.20833
BL03	0.0026	0.0048	0.0018	–0.0004	–	0.33974	0.03205	0.09423	0.20897	0.71795	0.42244	0.05641	0.00256
SH02	–0.0006	0.0017	0.0003	0.0004	0.0034	–	0.52374	0.51987	0.36667	0.96090	0.62244	0.59808	0.58654
SH03	–0.0001	–0.0013	0.0039	0.0003	0.0029	0.0016	–	0.80577	0.37628	0.48846	0.69359	0.49936	0.90962
SH01J	0.0028	0.0012	0.0055	0.0016	0.0053	0.0048	–0.0001	–	0.08141	0.27244	0.33974	0.20577	0.37179
MF02	0.0044	0.0008	0.0025	0.0021	0.0031	0.0012	0.0007	0.0009	–	0.12372	0.42179	0.14615	0.01154
MF03	0.0000	0.0034	0.0024	–0.0002	–0.0014	–0.0003	0.0005	0.0039	0.0019	–	0.28269	0.38974	0.18526
VI02	–0.0002	0.0002	–0.0006	0.0011	0.0022	0.0004	–0.0001	0.0019	–0.0005	0.0014	–	0.94487	0.76026
VI03	–0.0003	0.0013	0.0009	0.0005	0.0032	0.0005	–0.0002	0.0022	0.0003	0.0007	–0.0018	–	0.44167
BE06	0.0012	0.0023	0.0035	0.0020	0.0083	0.0039	–0.0007	0.0016	0.0030	0.0045	0.0007	0.0012	–
Gmo 132 only													
CL02	–	0.51859	0.46859	0.22692	0.08846	0.62629	0.66282	0.13013	0.80577	009744	0.00064	0.00064	0.00577
CL03	–0.0008	–	0.95513	0.58910	0.19872	0.99295	0.56282	0.32436	0.75833	0.84936	0.00064	0.00064	0.02564
CL01J	–0.0017	–0.0045	–	0.68077	0.43141	0.90128	0.99423	0.54744	0.96859	0.79551	0.21346	0.00064	0.25641
BL02	0.0024	0.0036	–0.0019	–	0.56346	0.73462	0.42372	0.20513	0.43654	0.04936	0.00321	0.00064	0.31859
BL03	0.0081	0.0091	0.0017	0.0086	–	0.24167	0.08205	0.08654	0.35064	0.05192	0.21346	0.00256	0.19936
SH02	–0.0016	–0.0060	–0.0037	–0.0007	0.0052	–	0.49744	0.52628	0.76859	0.80192	0.01923	0.00256	0.13910
SH03	–0.0048	–0.0019	–0.0071	–0.0007	0.0029	–0.0041	–	0.48077	0.85962	0.31218	0.05577	0.00321	0.04038
SH01J	0.0015	–0.0039	–0.0029	0.0018	0.0041	–0.0038	–0.0030	–	0.24551	0.04487	0.03141	0.00064	0.00385
MF02	–0.0027	–0.0013	–0.0056	–0.0024	0.0025	–0.0044	–0.0056	0.0001	–	0.45705	0.05064	0.00192	0.31859
MF03	–0.0017	–0.0022	–0.0011	0.0049	0.0090	–0.0039	–0.0024	0.0015	–0.0022	–	0.00385	0.00064	0.00128
VI02	0.0307	0.0283	0.0164	0.0233	0.0047	0.0190	0.0151	0.0175	0.0139	0.0220	–	0.70449	0.01923
VI03	0.0505	0.0497	0.0372	0.0333	0.0226	0.0340	0.0330	0.0357	0.0295	0.0416	0.0000	–	0.00385
BE06	0.0095	0.0088	0.0007	0.0028	–0.0004	0.0014	0.0015	0.0078	–0.0016	0.0077	0.0059	0.0150	–

Table A2. Estimates of pairwise genetic differentiation (F_{ST}) (below diagonal) and p-values (above diagonal) for all 10 microsatellites, 9 microsatellites (excluding Gmo 132), and only the microsatellite Gmo 132, among spatial samples of cod collected in the northern North Sea, west of Scotland and in adjacent sea areas. See Table 1 for sample location abbreviations

Sample	CL	BL	SH	MF	VI	BE	FP	FB	CNS	ENS	KS	BS	EB	NEAC
All 10 microsatellites														
CL	–	0.00659	0.99011	0.37692	<0.00055	0.25330	0.01429	0.00220	0.00055	0.04780	0.02308	0.00165	<0.00055	<0.00055
BL	0.0024	–	0.15165	0.01209	<0.00055	0.01923	0.00110	<0.00055	0.00330	0.03626	0.01758	0.00330	<0.00055	<0.00055
SH	0.0011	0.0011	–	0.99341	0.03956	0.84176	0.09451	0.03516	0.11429	0.90110	0.15440	0.02527	<0.00055	<0.00055
MF	0.0011	0.0008	–0.0008	–	<0.00055	0.08022	0.07527	0.01374	0.01319	0.03901	0.07582	0.08516	<0.00055	<0.00055
VI	0.0052	0.0038	0.0031	0.0037	–	0.03352	0.00165	<0.00055	0.22527	0.01099	0.75275	0.04011	<0.00055	<0.00055
BE	0.0021	0.0045	0.0010	0.0031	0.0032	–	0.02418	0.04396	0.16374	0.21374	0.50275	0.12198	<0.00055	<0.00055
FP	0.0052	0.0043	0.0034	0.0017	0.0044	0.0053	–	<0.00055	<0.00055	<0.00055	0.07637	<0.00055	<0.00055	<0.00055
FB	0.0030	0.0031	0.0034	0.0022	0.0083	–0.0002	0.0085	–	0.00110	0.00385	0.03901	<0.00055	<0.00055	<0.00055
CNS	0.0049	0.0051	0.0032	0.0055	0.0016	0.0045	0.0059	0.0098	–	0.03462	0.20220	0.01593	<0.00055	<0.00055
ENS	0.0007	0.0005	–0.0015	0.0000	0.0019	0.0009	0.0034	0.0009	0.0035	–	0.08077	0.00604	<0.00055	<0.00055
KS	0.0032	0.0021	0.0020	0.0009	0.0003	–0.0003	0.0025	0.0012	0.0023	0.0005	–	0.57912	<0.00055	<0.00055
BS	0.0024	0.0026	0.0022	0.0018	0.0020	0.0037	0.0041	0.0067	0.0032	0.0023	–0.0001	–	<0.00055	<0.00055
EB	0.0397	0.0456	0.0419	0.0399	0.0405	0.0350	0.0367	0.0483	0.0421	0.0438	0.0333	0.0327	–	<0.00055
NEAC	0.0447	0.0455	0.0419	0.0429	0.0260	0.0382	0.0321	0.0530	0.0266	00396	0.0332	0.0341	0.0619	–

Table A2 (continued)

Sample	CL	BL	SH	MF	VI	BE	FP	FB	CNS	ENS	KS	BS	EB	NEAC
9 microsatellites (Gmo 132 excluded)														
CL	–	0.01923	0.95934	0.27198	0.28242	0.76923	0.03571	0.00220	0.04835	0.09835	0.41154	0.01099	<0.00055	<0.00055
BL	0.0022	–	0.26868	0.09451	0.05934	0.05330	0.00220	0.01813	0.02912	0.05330	0.12747	0.02363	<0.00055	<0.00055
SH	−0.0009	0.0011	–	0.98736	0.61593	0.97637	0.12473	0.01923	0.26154	0.86044	0.35824	0.02088	<0.00055	<0.00055
MF	0.0014	0.0006	−0.0006	–	0.12253	0.25275	0.15275	0.01429	0.06319	0.10275	0.36978	0.12033	<0.00055	<0.00055
VI	0.0006	0.0016	0.0000	0.0006	–	0.25275	0.00165	0.00110	0.20275	0.09396	0.75440	0.10659	<0.00055	<0.00055
BE	0.0012	0.0049	0.0008	0.0031	0.0023	–	0.07747	0.20495	0.29341	0.47473	0.10879	0.10879	<0.00055	<0.00055
FP	0.0046	0.0045	0.0032	0.0015	0.0038	0.0058	–	0.00385	0.00275	0.00165	0.08297	0.00165	<0.00055	<0.00055
FB	0.0031	0.0021	0.0038	0.0023	0.0042	−0.0010	0.0077	–	0.00824	0.04066	0.29066	<0.00055	<0.00055	<0.00055
CNS	0.0022	0.0042	0.0018	0.0042	0.0018	0.0047	0.0060	0.0076	–	0.19670	0.41538	0.00385	<0.00055	<0.00055
ENS	0.0004	0.0007	−0.0015	0.0000	−0.0003	0.0011	0.0038	0.0004	0.0024	–	0.13462	0.00879	<0.00055	<0.00055
KS	0.0017	0.0021	0.0014	0.0002	−0.0002	−0.0003	0.0030	−0.0005	0.0022	0.0005	–	0.81758	<0.00055	<0.00055
BS	0.0014	0.0022	0.0020	0.0015	0.0011	0.0045	0.0048	0.0061	0.0033	0.0024	−0.0004	–	<0.00055	<0.00055
EB	0.0399	0.0486	0.0429	0.0410	0.0407	0.0364	0.0392	0.0476	0.0428	0.0462	0.0349	0.0336	–	<0.00055
NEAC	0.0168	0.0243	0.0174	0.0190	0.0166	0.0193	0.0139	0.0252	0.0135	0.0173	0.0153	0.0161	0.0425	–
Gmo 132 only														
CL	–	0.04615	0.95275	0.66703	<0.00055	0.00330	0.02582	0.22802	<0.00055	0.09341	<0.00055	0.00385	<0.00055	<0.00055
BL	0.0038	–	0.10495	0.01484	<0.00055	0.05769	0.05055	<0.00055	0.00330	0.12582	0.00220	0.00604	<0.00055	<0.00055
SH	0.0023	0.0006	–	0.72473	<0.00055	0.05495	0.24176	0.48132	0.05824	0.71813	0.06374	0.30659	<0.00055	<0.00055
MF	0.0012	0.0017	0.0026	–	<0.00055	0.04231	0.11374	0.23352	0.02692	0.03132	0.00165	0.21319	<0.00055	<0.00055
VI	0.0410	0.0214	0.0271	0.0279	–	0.00165	<0.00055	<0.00055	0.51209	0.00385	0.59176	0.03901	<0.00055	<0.00055
BE	0.0091	0.0009	0.0023	0.0031	0.0111	–	0.02527	0.00824	0.08791	0.03242	0.03462	0.45055	<0.00055	<0.00055
FP	0.0096	0.0027	0.0049	0.0032	0.0094	0.0015	–	0.02088	0.12802	0.18132	0.26538	0.39121	<0.00055	<0.00055
FB	0.0024	0.0108	0.0006	0.0009	0.0408	0.0063	0.0148	–	0.00110	0.01044	0.00220	0.01429	<0.00055	<0.00055
CNS	0.0257	0.0116	0.0147	0.0159	0.0001	0.0028	0.0047	0.0271	–	0.00989	0.07637	0.61593	<0.00055	<0.00055
ENS	0.0023	0.0014	0.0018	0.0003	0.0194	−0.0010	0.0007	0.0045	0.0124	–	0.21538	0.40110	<0.00055	<0.00055
KS	0.0154	0.0021	0.0060	0.0066	0.0047	−0.0007	−0.0011	0.0146	0.0031	0.0004	–	0.08956	<0.00055	<0.00055
BS	0.0105	0.0056	0.0037	0.0039	0.0092	−0.0027	−0.0012	0.0110	0.0026	0.0009	0.0017	–	<0.00055	<0.00055
EB	0.0384	0.0213	0.0346	0.0313	0.0394	0.0236	0.0170	0.0532	0.0363	0.0246	0.0206	0.0256	–	<0.00055
NEAC	0.2454	0.2073	0.2202	0.2139	0.1074	0.1969	0.1781	0.2667	0.1366	0.2109	0.1833	0.1754	0.2141	–