Historical analysis of *Pan* I in Atlantic cod (*Gadus morhua*): temporal stability of allele frequencies in the southeastern part of the species distribution

Einar Eg Nielsen, Brian R. MacKenzie, Eydfinn Magnussen, and Dorte Meldrup

Abstract: We investigated temporal genetic differentiation at the pantophysin (*Pan* I) locus in four Atlantic cod (*Gadus morhua*) populations from the southeastern part of the species distribution: the Baltic Sea, the North Sea, the Faroe Plateau, and the Faroe Bank. Historical otolith collections enabled investigation of allele frequency variation over time periods up to 69 years employing *Pan* I primers specifically designed for partially degraded DNA. Small and nonsignificant temporal changes in *Pan* I allele frequencies were observed in the four populations. Simultaneous microsatellite analysis revealed similar temporal genetic stability with temporal F_{ST} values ranging from 0 to 0.006, suggesting limited demographic changes. Sea surface temperature, which has been suggested as the primary driver for the geographical distribution of *Pan* I alleles in cod, showed no long-term trend although temperature has increased since the mid-1990s. Our study demonstrates that populations in the southeastern part of the species range has been characterized by very high frequencies of the *Pan* I^A allele for many decades, and accordingly, *Pan* I serves as a reliable marker for genetic stock identification on a macrogeographical scale.

Résumé : Nous étudions la différenciation génétique au locus *Pan* I chez quatre populations de morues franche (*Gadus morhua*) du sud-est de l'aire de répartition de l'espèce, soit celles de la Baltique, de la Mer du Nord, du plateau de Féroé et du banc de Féroé. Des collections historiques d'otolithes ont permis d'analyser la variation de la fréquence des allèles sur des périodes pouvant atteindre 69 ans à l'aide d'amorces *Pan* I spécialement conçues pour l'ADN partiellement dégradé. Nous observons des changements temporels faibles et non significatifs dans la fréquence des allèles de *Pan* I dans les quatre populations. Des analyses simultanées des microsatellites montrent aussi une stabilité génétique temporelle, les valeurs temporelles de F_{ST} variant de 0–0,006, ce qui indique de faibles changements démographiques. La température de surface de la mer, qu'on a suggéré comme principal facteur explicatif de la répartition géographique des allèles de *Pan* I chez la morue, ne montre aucune tendance à long terme, bien que la température ait augmenté depuis le milieu des années 1990. Notre étude démontre que les populations du sud-est de l'aire de répartition de l'espèce sont caractérisées depuis plusieurs décennies par de hautes fréquences de l'allèle *Pan* I^A; c'est pourquoi *Pan* I peut servir de marqueur fiable de l'identification génétique des stocks à l'échelle macrogéographique.

[Traduit par la Rédaction]

Introduction

Noncoding DNA markers such as microsatellites have become the preferred option for inferring population structure in marine organisms. In particular for marine fishes, the application of microsatellites for population genetic studies has generated many new insights (see reviews by Waples 1998; DeWoody and Avise 2000). The discovery of genetic structure on most geographical levels investigated in marine fishes (e.g., Nielsen et al. 2003; O'Reilly et al. 2004; Jørgensen et al. 2005) has opened a wealth of new opportunities for genetic stock identification (GSI) in marine fishes. Among the most promising possibilities are the assignment of individuals to population (reviewed by Hansen et al. 2001; Manel et al. 2005) and estimation of the proportion contributed by potential donor populations to a mixed fishery, so-called mixed stock analysis (e.g., see Pella and Masuda 2001). The power of these statistical methods depends critically on the level of genetic divergence among populations. If microsatellite differentiation is low, as commonly found among populations of marine fishes, the number of loci has to be increased to improve exclusion power

Received 7 November 2006. Accepted 24 May 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 12 October 2007. J19649

E.E. Nielsen[1] and D. Meldrup. Department of Inland Fisheries, Danish Institute for Fisheries Research, DK-8600 Silkeborg, Denmark.
B.R. MacKenzie. Department of Marine Ecology and Aquaculture, Danish Institute for FisheriesResearch, Kavalergården 6, DK-2920 Charlottenlund, Denmark.
E. Magnussen. Faculty of Science and Technology, University of the Faroe Islands, 3 Nóatún, FO 100, Tórshavn, Faroe Islands.

[1]Corresponding author (e-mail: een@difres.dk).

 doi:10.1139/F07-104 © 2007 NRC Canada

(Manel et al. 2005). Unfortunately, also the time and cost of conducting the genetic analysis increases, making them less attractive for real time and routine application.

Alternatively, it is possible to turn to genes under directional selection and (or) genetic markers associated with genes under selection, which are expected to show higher population divergence than neutral markers. In particular for organisms characterized by high effective population sizes, such as marine fishes, selection is a much more powerful force than random genetic drift for driving population divergence (see Endler 1986).

Recent studies have shown that it is possible to identify microsatellite loci under selection in marine fishes (Nielsen et al. 2006); however, the selected genetic marker that has attracted most attention recently for marine fishes is the pantophysin (*Pan* I) locus in gadoid fishes. Pogson et al. (1995) sequenced the anonymous cDNA clone GM798, which showed elevated levels of genetic differentiation compared with a number of other cDNA clones and identified the gene as synaptophysin (*Syp* I) — later revised to *Pan* I (Pogson 2001). *Pan* I is an integral membrane protein of unknown function (see Pogson and Mesa 2004 and references therein for further details about protein structure and suspected function). By investigating inter- and intra-specific sequence divergences, positive Darwinian selection has been invoked (Pogson 2001; Canino and Bentzen 2004; Pogson and Mesa 2004). In cod, two highly different allele classes have been identified, the *Pan* I^A and the *Pan* I^B alleles respectively, which differ by four amino acid substitutions and can be differentiated by a *Dra* I restriction site. The two alleles appear to have evolved prior to the split between cod and its sister taxon walleye pollock (*Theragra chalcogramma*) (Pogson and Mesa 2004). Within the two classes, additional sequence variation can be found (see Pogson 2001), but the major focus of population genetic studies has been on the distribution of the diallelic system only, revealing extensive *Pan* I differentiation in cod (for an overview see Case at al. 2005). Whether this distributional pattern is caused by contemporary or historical patterns of selection is difficult to demonstrate (Pogson 2001), but contemporary selection at *Pan* I has been invoked as responsible for growth variation among populations (Fevolden and Pogson 1995) and individuals (Case et al. 2006), as well as inducing temporal instability in allele frequencies (Karlsson and Mork 2003).

Potential environmental drivers responsible for selection at *Pan* I were investigated in a recent paper by Case et al. (2005). They found *Pan* I allele frequencies to be correlated with temperature, salinity, and depth, of which the latter was also suggested recently by Pampoulie et al. (2006). Case et al. (2005) concluded that "In areas of 'normal' salinity, temperature may be the determining factor in *Pan* I allele frequencies." Short-term (interannual) fluctuations in ocean temperatures are well documented as well as longer term trends (Stenseth et al. 2004; B.R. MacKenzie and D. Schiedek, unpublished data), which may be due to global warming (Barnett et al. 2005). Therefore, the expectations are temporally changing patterns of selection, which could lead to fluctuating or gradually changing allele frequencies, as suggested by Karlsson and Mork (2003). Temporal changes in allele frequencies may have an important bearing for invoking population of origin of individuals or groups of individuals, thereby leading to biased or erroneous GSI estimates. Accordingly, to employ genetic markers under selection for GSI, it is important to identify long-term patterns of stability to determine the effects of the environmental driver(s) of selection, but also to rule out demographic factors (i.e., migration and genetic drift as sources of temporal variation).

DNA analyses from historical collections (scales and otoliths; e.g., see Nielsen et al. 1997; Hutchinson et al. 1999) of fish have shown much promise for investigating patterns of temporal genetic variation. In marine fishes, archived material has been used to evaluate long-term (putative) neutral variation by estimation of genetically effective population sizes and migration (Hauser et al. 2002; Hutchinson et al. 2003; Poulsen et al. 2006). Until now, however, no attempts have been made to evaluate historical patterns of selection on specific candidate loci in marine fishes.

Here we employ historical cod otolith samples spanning up to 69 years to evaluate long-term temporal stability in *Pan* I allele frequencies in populations from the southeastern part of the species' distributional range. We use new primers specifically designed for analysis of partially degraded DNA from otoliths. We investigate the relative roles of neutral (drift and migration) and selective evolutionary forces by simultaneously analyzing noncoding microsatellites. We also explore the potential role of thermal selection in shaping patterns of *Pan* I allele frequencies by correlating historical temperature regimes with temporal development of *Pan* I allele frequencies. Finally, based on the current results, we evaluate the suitability of *Pan* I and other selected markers for GSI in marine fishes.

Materials and methods

Sample collection

A total of 442 cod from four populations was analyzed for *Pan* I variation: the Baltic proper (Bornhom Basin), the North Sea (Moray Firth), the Faroe Plateau, and the Faroe Bank (see Fig. 1 and Table 1). The four areas are important spawning grounds for cod, but have all suffered from overexploitation in recent years. Contemporary samples were collected from primarily adult mature individuals and consisted of gill tissue stored in ethanol. Samples from the Baltic Sea and the North Sea were collected in relation to previous research projects (for details about samples, see Nielsen et al. 2001, 2003; Poulsen et al. 2006), while the contemporary samples from Faroese waters were collected by the University of the Faroe Islands in cooperation with the Faroese Fisheries Laboratory. Historical samples consisted of saggital otoliths stored in paper bags and provided by the Danish Institute for Fisheries Research, Fisheries Research Services Scotland, and the Faroese Fisheries Laboratory.

Sea surface (0–2 m) temperature (SST) data were obtained from the International Council for the Exploration of the Sea (ICES) Hydrographic Data Centre and the Faroese Fisheries Laboratory for three regions corresponding to spawning areas for the sampled cod populations. The three areas were the

© 2007 NRC Canada

Fig. 1. Location of temporal Atlantic cod (*Gadus morhua*) samples and *Pan* I allele frequencies. Grey: *Pan* I^A; black: *Pan* I^B. For list of sample abbreviations, see Table 1.

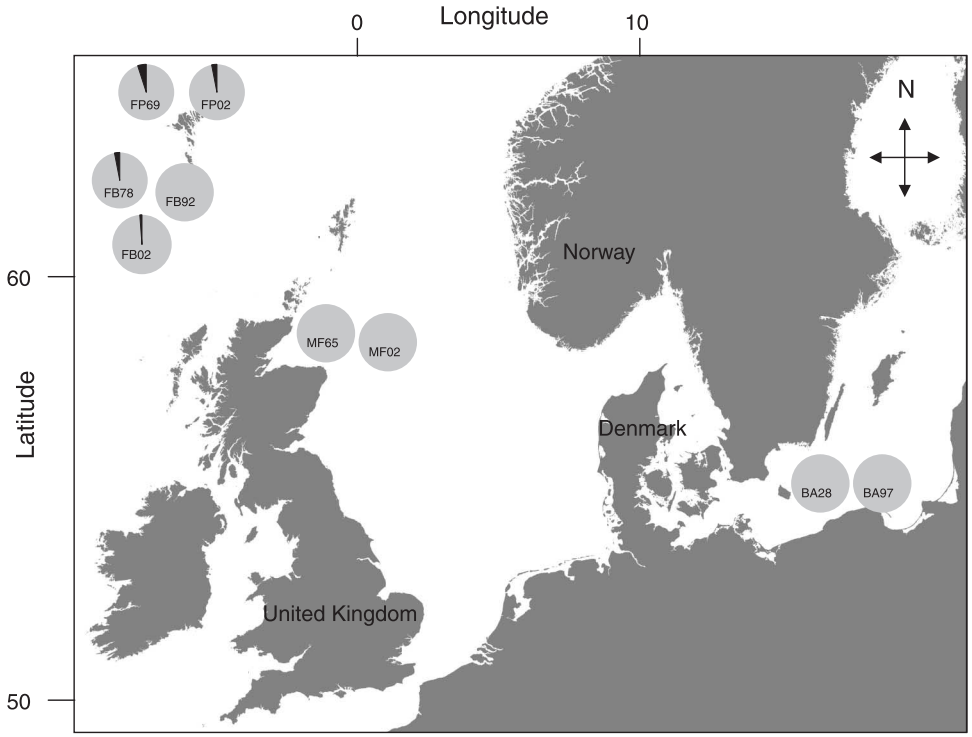


Table 1. Summary statistics of contemporary (and historical in parentheses) Atlantic cod (*Gadus morhua*) samples analyzed for *Pan* I (microsatellites) and frequencies of the *Pan* I^A allele.

Geographical locality	Position	Sampling year	Abbreviation	Month of sampling	No. of individuals	*Pan* I^A
Baltic Sea	Bornholm Basin	1928	BA28	NA	50 (50)	1
	(54.51°N, 15.28°E)	1997	BA97	April	50 (62)	1
North Sea	Moray Firth	1965	MF65	May	49 (50)	1
	(57.95°N, 2.20°W)	2002	MF02	March	50 (71)	1
Faroe Plateau	Faeroe Plateau	1969	FP69	April	50 (33)	0.95
	(62.52°N, 6.18°W)	2002	FP02	April	48 (45)	0.97
Faroe Bank	Faroe Bank	1978	FB78	April	56 (46)	0.97
	(60.94°N, 8.57°W)	1992	FB92	April	49 (0)	1
		2002	FB02	April	50 (71)	0.99

Bornholm Basin (55.9°N–54.9°N, 16.3°E–14.7°E), northwest North Sea (5°W–0°; 55°N–60°N), and the Faroe Plateau. The Faroese data are based on daily measurements at Mykines (7.672°W, 62.097°N) during 1914–1969 and at Oyragjógv (7.170°W, 62.116°N) 1990–2005. Data for the intervening years are derived from a validated oceanographic model of the local circulation (Hatun et al. 2005). Temperatures on the Faroe Bank are 1–2 °C warmer than those on the Plateau (Westerberg 1990; Lastein 1992) and are closely correlated with Plateau temperatures (E. Magnussen, unpublished data). The month chosen for comparison with genetic data was June, which approximates the timing of peak abundance of early life history stages (eggs and newly hatched larvae) for most of these regions (but not Bornhom Basin). It can be debated if June temperature per se could be the selective agent responsible for geographic differences in *Pan* I allele frequencies. However, reporting of June temperatures enabled direct comparison of our results with the most extensive study of relationships between temperatures and *Pan* I allele frequencies reported by Case et al. (2005).

DNA analysis

DNA was extracted from both contemporary and historical samples using a proteinase K – chelex procedure following Estoup et al. (1996). Owing to the degraded nature of DNA from archival material (Nielsen et al. 1999), we were unable to use previously designed primers for amplification of the *Pan* I locus. Based on one of the sequences from Pogson and Fevolden (2003) deposited in GenBank (Accession number AY146692), we designed primers for a 142-basepair (bp) segment spanning the diagnostic *Dra* I restriction site discriminating between the *Pan* I^A and the *Pan* I^B alleles (forward: 5′-GGCAAATGAAACCCAGAAAA; reverse: 5′- ATGACACTTGTGGCAAGCAG).

Polymerase chain reaction (PCR) amplification of the *Pan* I segment was conducted using standard reagents and an an-

© 2007 NRC Canada

nealing temperature of 55° with 30 and 40 PCR cycles for contemporary and historical samples, respectively. Restriction analysis was carried out according to manufacturer's recommendations (Sigma), and fragments were visualized on a 2% agarose gel (NuSieve, Cambrex Corp.) stained with ethidium bromide.

PCR amplification of nine highly polymorphic microsatellites were conducted for the Baltic Sea and North Sea samples: *Gmo*2 (Brooker et al. 1994); *Gmo*19, *Gmo*34, *Gmo*35 (Miller et al. 2000); *Tch*5, *Tch*11, *Tch*14, *Tch*22 (O'Reilly et al. 2000); and GADM1 (Hutchinson et al. 2001). These results have previously been used to infer effective population sizes in cod populations (Poulsen et al. 2006). We also employed nine loci for the Faroe Bank samples; however, *Tch*22 was exchanged with *Gmo*8 (Miller et al. 2000). We did not perform microsatellite analysis on the intermediate (1992) samples, since our results from 1978 and 2002 revealed lack of temporal differentiation. The historical Faroe Plateau samples only allowed consistent amplification of three microsatellite loci with short maximum allele sizes because of bad DNA quality. These were *Gmo*2, with a maximum observed allele size of 135 bp; *Gmo*34 (maximum 119 bp); and *Gmo*35 (maximum 142 bp). Accordingly, the temporal tests for genetic differentiation at the Faroe Plateau are only conducted on three loci (see Table 2).

For both *Pan* I and microsatellite analysis, all runs included negative and positive controls, the latter consisting of individuals with known genotypes. Approximately 25% of the historical samples and 15% of contemporary samples were rerun to check for consistency of results. No signs of contamination were observed.

Statistical analysis

Tests for deviations from Hardy–Weinberg equilibrium for *Pan* I and individual microsatellite loci in each sample were conducted using the program FSTAT (Goudet 1995). The same program was also used to provide unbiased estimates of temporal, pairwise F_{ST} values for microsatellites, following Weir and Cockerham (1984), and associated 95% confidence intervals by bootstrapping over loci. Accordingly, no confidence interval was calculated for the Faroe Plateau data (i.e., only three loci; see Table 2). Finally, temporal heterogeneity in allele frequencies was tested by permuting alleles among samples (10 000 times).

The raw SST data were averaged to monthly values and plotted as time series to enable visual inspection for overall trends and variations at multiannual time scales. We also used statistical methods to investigate whether significant ($\alpha = 0.05$) variations occurred at multiannual time scales. These methods included linear regression for the detection of trends over the entire time period and general additive models (GAMs; Swartzman et al. 1992; Begg and Marteinsdottir 2002) for detection of variations at shorter (multiannual) time scales. The methodology used for fitting the GAMs to the SST data is described elsewhere (B.R. MacKenzie and D. Schiedek, unpublished data); in brief, the GAMs were fitted using a generalized cross-validation procedure applied to a locally weighted least squares regression. The variance (deviance) explained by these models was tested for statistical significance using the pseudo-R^2 criterion (Swartzman et al. 1992; B.R. MacKenzie and D. Schiedek, unpublished data).

Table 2. Temporal estimates of microsatellite intrapopulation differentiation (θ_{ST}), 95% confidence intervals (CI), and associated *p* values for pairwise exact tests of differentiation.

Temporal comparison	Baltic Sea	North Sea	Faroe Plateau	Faroe Bank
θ_{ST}	0.006	0.001	0.000	0.006
95% CI	0–0.012	0–0.06	—	0–0.023
p	<0.01	0.02	0.41	0.15

Results

As expected from the sequence information, the restriction analysis of the PCR-amplified *Pan* I segment provided a single band of 142 bp for *Pan* I^A homozygotes, two bands of 40 and 102 bp for *Pan* I^B homozygotes, and all three bands for heterozygotes. Generally, high amplification success was achieved from close to 100% for contemporary samples down to 96% for the FP69 sample. No significant deviations from Hardy–Weinberg proportions were found for *Pan* I; however, a single *Pan* I^B homozygote was found in two historic samples (FP69 and FB78). The amplification success of microsatellites was generally good, with less than 5% missing single locus genotypes. Only the FP69 sample had a high frequency of missing genotypes. For this sample, one-third of the individuals did not provide any amplification product for two or all three loci after two rounds of attempted amplification, and all results from these individuals were discarded accordingly. For the remaining individuals, 17% missing single locus genotypes were present. No deviations from Hardy-Weinberg proportions were evident after Bonferroni correction (Rice 1989) in any of the samples.

Samples from the two populations from the Faroe Islands were generally polymorphic for *Pan* I (Fig. 1), except for the Faroe Bank 1992 sample, while all population samples from the Baltic Sea and the North Sea were monomorphic for the *Pan* I^A allele. Despite an apparent tendency for a reduction in the *Pan* I^B allele frequency over time in the Faroese samples, none of the changes were significant (FB78–FB92: $p = 0.10$; FB78–FB02: $p = 0.34$; FP69–FP02: $p = 0.48$). Likewise, the microsatellite analysis revealed a general pattern of temporal stability (Table 2), with very low F_{ST} values between temporal samples.

SST in June was usually warmer in the Bornholm Basin and northwestern North Sea than on the Faroe Plateau (Fig. 2; Table 3). Long-term mean temperatures in these areas were 11.4 and 10.6 °C, and therefore 2–3 °C warmer than on the Faroe Plateau (8.3 °C; Fig. 2, Table 3). Temperatures in the Bornholm Basin were more variable than those at the other sites (Fig. 2, Table 3).

There is no evidence of an overall long-term increase or decrease in temperature in the three regions (Table 3). Temperatures in two areas (Bornholm Basin and Faroe Plateau) have increased since the mid-1990s until the early 2000s; the increase is largest in the Bornholm Basin where temperatures rose 4.8 °C between 1997 and 2003 (Fig. 2). The GAM analysis explained significant multiannual temperature variability only for the Faroe Plateau (pseudo-R^2 = 18%, $p <$

© 2007 NRC Canada

0.001; $p > 0.05$ for the other regions); temperature variations at the other sites occurred at short time scales (i. e., year-to-year), which could not be resolved by the GAMs.

Discussion

Analysis of *Pan* I from historical otolith collections

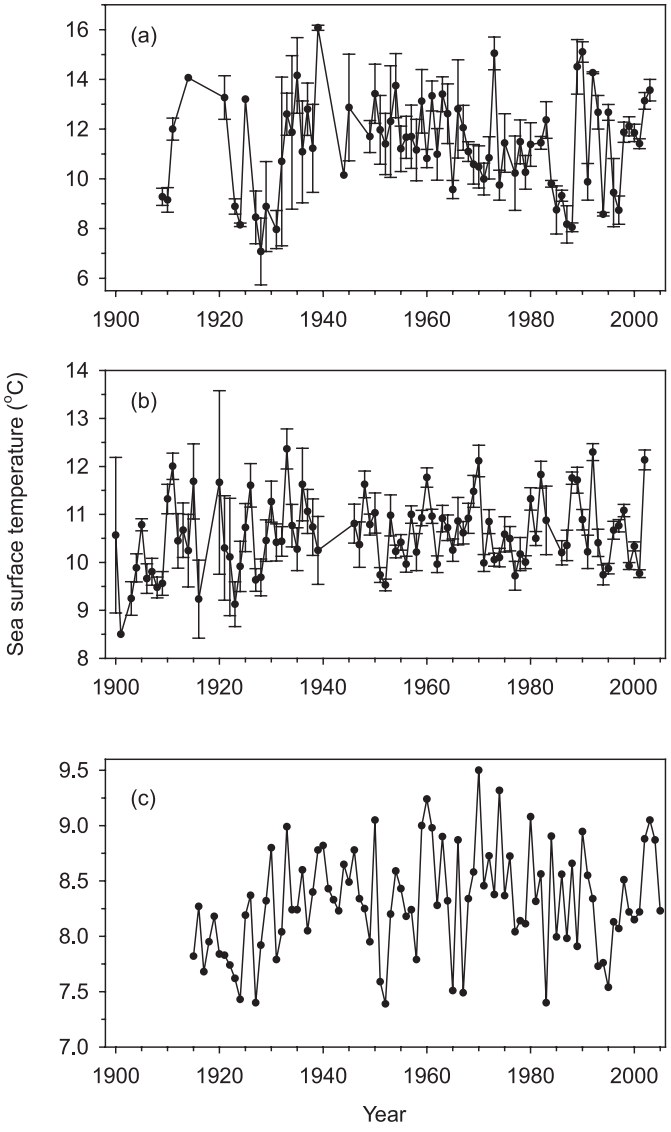
Our study demonstrates that analysis of *Pan* I genetic variation can be achieved from historical otolith collections in cod, as has been demonstrated previously for microsatellites (Hutchinson et al. 2003; Poulsen et al. 2006). Compared with microsatellite analysis, the method appears to be more robust for analysis of partially degraded DNA, since high PCR amplification success was possible even for the FB69 sample, where only three microsatellite loci could be amplified. Even compared with those loci, which have an amplified fragment size similar to the *Pan* I fragment spanned by our newly designed primers, we had a higher amplification success for *Pan* I, which was almost similar to contemporary samples. A possible explanation for the different amplification success could be microsatellite-specific technical problems, such as short allele dominance and large allele dropout, causing inconsistent scoring and subsequent need to discard affected loci or individuals for further analysis. We also think the method would be well suited for even older material, since the most severe degradation of DNA from archived material appears to occur during the first years of storage (see Nielsen et al. 1999). In short, we think that the method described here is very robust and serves as a highly reliable, low-tech alternative to other *Pan* I genotyping methods (Fevolden and Pogson 1995; Case et al. 2005 and Stenvik et al. 2006), not only for historical tissue but also, for example, for real-time analysis onboard research vessels.

Evolutionary drivers of *Pan* I allele frequencies

The present study demonstrates that the *Pan* I locus has not been subject to any major changes in allele frequencies in the populations sampled over the last three to seven decades. This is consistent with neutral expectations based on our microsatellite analysis, showing similar limited genetic differentiation over time. The previously published results of cod from the North Sea and the Baltic Sea (Poulsen et al. 2006) are corroborated by the new results from the Faroe Islands in suggesting "large" effective population sizes in cod. By large, we do not necessarily mean large compared with census size, but large enough to render random genetic drift an evolutionary driver of low importance (see Poulsen et al. 2006 for a more detailed discussion on effective population sizes in cod).

Geographic comparison of *Pan* I^A frequencies with SST shows that the temperature at which *Pan* I^A frequencies begins to increase is about 6–7 °C; at lower temperatures, this allele was rare, whereas at temperatures >10 °C, *Pan* I^A frequency was nearly 100% (Case et al. 2005). In our study, temperatures at no time during the last 70 years in any of our areas reached the low temperatures that are typical of habitats occupied by cod populations dominated by very high frequencies of the *Pan* I^B allele. Generally, we observed very low frequencies of this allele in all of our samples, regardless of when or where the samples were collected. Our

Fig. 2. Temporal development of June sea surface temperatures (mean ± two standard errors) for the Baltic Sea (Bornholm Basin), northwest North Sea, and Faroe Plateau. Means, variances, and results of linear regression analyses are available in Table 3.



results, when considered together with those of Case et al. (2005), demonstrate that the temporal variation in *Pan* I allele frequencies is much lower than the geographic variation, at least in the populations sampled and under the given historic temperature variability observed in and among our sites. However, given that temperatures in many areas of the northeast Atlantic, including the Faroe Plateau, North Sea, and Baltic Sea all have increased in the last 10–15 years (B.R. MacKenzie and D. Schiedek, unpublished data) and are expected to increase further because of future global warming (IPCC 2001; BACC 2006), *Pan* I^A frequencies will likely remain high in these regions during the coming decades given that temperature is actually an important factor determining allele frequencies at the *Pan* I locus. Both the results of Case et al. (2005) and Pampoulie et al. (2006) also suggest a correlation between *Pan* I allele frequencies and depth, implying that other factors than temperature per se

© 2007 NRC Canada

Table 3. Summary statistics and results of linear regression of June sea surface temperature measured at three cod spawning areas in the northeast Atlantic.

Region	Latitude, longitude	*N*	Mean	Variance	R^2	*p*
Bornholm Basin, Baltic Sea	55.9°N–54.9°N, 16.3°E–14.7°E	75	11.4	3.7	0.00	0.60
Northwest North Sea	55°N–60°N, 5°W–0°	91	10.6	0.6	0.04	0.05
Faroe Plateau	62.097°N, 7.672°W; 62.116°N, 7.170°W	91	8.3	0.2	0.05	0.04

may be important for shaping allele frequencies at this locus. However, as temperature is also highly correlated with depth, disentangling direct and correlated effects may prove difficult.

The absence of a significant multidecadal June temperature increase does not allow us to come to any clear conclusions regarding the relative importance of contemporary temperature selection on shaping the distribution of *Pan* I allele frequencies as suggested by Case et al. (2005). However, demonstration of selection in the wild is difficult (Endler 1986), in particular when many environmental clines are correlated (Sarup et al. 2006) and when the specific causes of mortality are poorly known. Even if temperature is the driver for allele frequency differences, other aspects than mean June surface temperatures (e.g., maximum and minimum) could prove more important, particularly if recruitment in some years is determined after the larval stage. For example, processes occurring after settlement to the benthos can affect mortality rates during the juvenile stages (Heath and Gallego 1997), and carrying capacity for juveniles in most cod populations in the northeast and northwest Atlantic has been shown to be correlated with bottom temperature in juvenile habitat (Myers et al. 2001). Still, we see evidence of year-to-year fluctuations in temperature, which could explain why Case et al. (2005) in 1997 observed a lower frequency of the *Pan* I^A allele at the Faroe Plateau than we did (0.88 versus 0.95–0.97, *p* values of 0.023 and 0.020, respectively), also supporting the hypothesis that short-term variation in allele frequencies, as observed by Karlsson and Mork (2003), may be caused by fluctuating contemporary temperature selection (or that samples were taken at different depths). This could suggest that the polymorphism at *Pan* I observed in many cod populations are maintained by fluctuating directional selection regimes as suggested by Hedrick (2002) for major histocompatibility complex genes. However, if relatively large changes in allele frequencies caused by fluctuating selection are common, there is a very high chance that one of the alleles could be lost over time, leading to fixation. Alternatively, the observed pattern of *Pan* I genetic differentiation in cod could be explained by a combination of gene flow and contemporary directional or fluctuating selection, explaining why the most northerly cod populations from the Barents Sea are not completely fixed for the *Pan* I^B allele (see e.g., Case et al. 2005). Also, in relation to the results of this study, the few individuals with *Pan* I^B alleles found around the Faroe Islands could be migrants or descendants from migrants (hybrids) from the proximate population at the Iceland–Faroese Ridge, which according to Case et al. (2005) has a very high frequency of *Pan* I^B alleles. This would also explain the occurrence of the unexpected *Pan* I^B homozygotes in our samples. Such a gene flow selection hypothesis for the Faroe Plateau could be investigated by including samples from populations potentially donating migrants. Individuals or proportions of an individual's genotype could then be assigned to population of origin using microsatellite analysis (see Pritchard et al. 2000), and the potential migrant origin of *Pan* I^B genotypes could be determined.

Application of *Pan* I for GSI

Our findings of temporal stability of *Pan* I allele frequencies over several decades illustrate that this locus can be used as a reliable genetic marker to distinguish between cod populations from the northern and southern part of its eastern distributional range in the Northeast Atlantic and most likely also among well-defined genetically isolated cod populations on a smaller geographical scale. However, using *Pan* I as a single genetic marker on a microgeographical scale cannot be recommended before more knowledge on the relative importance of different contemporary evolutionary factors affecting this locus has been generated. Accordingly, the temporal and spatial dynamics in other regions, preferably with a higher degree of polymorphism, should be investigated using a combination of noncoding genetic markers and *Pan* I, including information on potential drivers of evolution.

The *Pan* I locus has not only been used to identify local cod populations, but also to investigate the origin of individual cod in mixed aggregations (Berg et al. 2005). However, compared with results obtained by microsatellites (e.g., Nielsen et al. 2001), this locus alone has less power for individual assignment. Instead, combining information from a number of loci subject to selection is expected to provide extreme power for individual assignment. Until now, only a few genes likely to be under selection have been identified in cod (Anonymous 2006); however, with the initiation of large-scale cod genomic projects such as the CGP (Atlantic cod Genomics and Broodstock Development), the knowledge and sequence information of genes under selection is expected to grow tremendously in the near future. Still, careful evaluation of the temporal stability of other selected genes should be conducted before they can be applied for GSI.

Pan I in cod in the southeastern part of the species distribution

The pattern of temporal genetic differentiation at the *Pan* I locus in cod from the southeastern part of the distributional range shows stability over the sampled period. There is no evidence that the apparently cold-adapted *Pan* I^B allele has been present in high frequencies in North Sea and (or) Baltic Sea cod populations for many decades. Similarly, there is little evidence of a recent reduction in *Pan* I^B allele frequencies in cod populations at the Faroe Plateau and the Faroe

© 2007 NRC Canada

Bank. The pattern of stability is apparently caused by the very limited genetic drift (large effective population sizes) in the four populations, elucidated by the temporal stability in microsatellite allele frequencies. Similarly, there is no clear evidence of a dramatically changed temperature selection regime with respect to mean June temperatures, which potentially could have given rise to a selective sweep at *Pan* I. This illustrates that cod populations in the southeastern part of its distributional range are — and have been for a long time — characterized by a very high frequency of the *Pan* I^A allele, even though we do not have sufficient power to be able to test our initial hypothesis of a relationship between increasing temperatures and *Pan* I allele frequencies. Accordingly, *Pan* I should be considered a reliable marker for GSI in cod on a macrogeographical scale, while on a smaller geographical scale a better understanding of the evolutionary drivers and their temporal dynamics are required. To elucidate these associations, more genetic studies on cod in time and space, simultaneously applying genetic markers with different evolutionary properties, should be conducted as well as more common garden experiments of the effect of genotype on fitness at different temperatures. We expect that genes under selection, such as *Pan* I, will be the preferred high resolution markers for GSI in the future.

Acknowledgements

The project has been carried out in the frame of the EU-funded project METACOD (Grant No. Q5RS-2001-00953), and we thank Nina Poulsen, Daniel Ruzzante, Peter Wright, Fiona Gibb, and Gudrun Marteinsdottir and the rest of the project team for providing samples and stimulating discussions. We are also grateful to the EU MARBEF Network of Excellence "Marine Biodiversity and Ecosystem Functioning" (contract No. GOCE-CT-2003-505446). We thank ICES for providing hydrographic data.

References

Anonymous. 2006. ICES Working Group on Application of Genetics in Fisheries and Mariculture report 2006. ICES CM 2006/MCC:04.

BACC. 2006. International Conference on Assessment of Climate Change for the Baltic Sea Basin, Göthenberg, Sweden, 22–23 May 2006. GKSS, Gestacht, Germany.

Barnett, T.P., Pierce, D.W., AchutaRao, K.M., Gleckler, P.J., Santer, B.D., Gregory, J.M., and Washington, W.M. 2005. Penetration of human-induced warming into the world's oceans. Science (Washington, D.C.), **309**: 284–287.

Begg, G.A., and Marteinsdottir, G. 2002. Environmental and stock effects on spatial distribution and abundance of mature cod *Gadus morhua*. Mar. Ecol. Prog. Ser. **229**: 245–262.

Berg, E., Sarvas, T.H., Harbitz, A., Fevolden, S.E., and Salberg, A.B. 2005. Accuracy and precision of stock separation of northeast Arctic cod and Norwegian coastal cod by otoliths — comparing readings, image analyses and a genetic method. Mar. Freshw. Res. **56**: 753–762.

Brooker, A.L., Cook, D., Bentzen, P., Wright, J.M., and Doyle R.W. 1994. Organization of microsatellites differs between mammals and cold-water teleost fishes. Can. J. Fish. Aquat. Sci. **51**: 1959–1966.

Canino, M.F., and Bentzen, P. 2004. Evidence for positive selection at the pantophysin (Pan I) locus in walleye pollock, *Theragra chalcogramma*. Mol. Biol. Evol. **21**: 1391–1400.

Case, R.A.J., Hutchinson, W.F., Hauser, L., Van Oosterhout, C., and Carvalho, G.R. 2005. Macro- and micro-geographic variation in pantophysin (*Pan*I) allele frequencies in NE Atlantic cod *Gadus morhua*. Mar. Ecol. Prog. Ser. **301**: 267–278.

Case, R.A.J, Hutchinson, W.F., Hauser, L, Buehler, V. Clemmesen, C., Dahle, G., Kjesbu, O.S., Moksness, E., Ottera, H., Paulsen, H., Svasand, T., Thorsen, A., and Carvalho, G.R. 2006. Association between growth and Pan I genotype within Atlantic cod full-sibling families. Trans. Am. Fish. Soc. **135**: 241–250.

DeWoody, J.A., and Avise, J.C. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. J. Fish Biol. **56**: 461–473.

Endler, J.A. 1986. Natural selection in the wild. Princeton University Press, Princeton, N.J.

Estoup, A., Largiader, C.R., Perrot, E., and 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.

Fevolden, S.E., and Pogson, G.H. 1995. Differences in nuclear DNA RFLPs between the Norwegian coastal cod and the Northeast Arctic population of Atlantic cod. *In* Ecology of fjords and coastal waters. *Edited by* H.R. Skjoldahl, C. Hopkins, K.E. Erikstad, and H.P. Leinaas. Elsevier Science, Amsterdam. pp. 403–415.

Goudet, J. 1995. FSTAT. Version 1.2. A computer program to calculate F-statistics. J. Hered. **86**: 485–486.

Hansen, M.M., Kenchington, E., and Nielsen, E.E. 2001. Assigning individual fish to populations using microsatellite DNA markers: methods and applications. Fish Fish. **2**: 93–112.

Hatun, H., Sandø, A.B., Drange, H., and Bentsen, M. 2005. Seasonal to decadal temperature variations in the Faroe–Shetland inflow waters. *In* The Nordic Seas: an integrated perspective. American Geophysical Union, Washington, D.C. pp. 239–250.

Hauser, L., Adcock, G.J., Smith, P.J., Bernal-Ramírez, J.H., and Carvalho, G.R. 2002. Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). Proc. Natl. Acad. Sci. U.S.A. **99**: 11742–11747.

Heath, M., and Gallego, A. 1997. From the biology of the individual to the dynamics of the population; bridging the gap in fish early life studies. J. Fish Biol. **51**: 121–134.

Hedrick, P.W. 2002. Pathogen resistance and genetic variation at MHC loci. Evolution, **56**: 1902–1908.

Hutchinson, W.F., Carvalho, G.R., and Rogers, S.I. 1999. A nondestructive technique for the recovery of DNA from dried fish otoliths for subsequent molecular genetic analysis. Mol. Ecol. **8**: 893–894.

Hutchinson, W.F., Carvalho, G.R., and Rogers, S.I. 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, W.F., Van Oosterhout, C., Rogers, S.I., and Carvalho, G.R. 2003. Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*). Proc. R. Soc. Lond. B. Biol. Sci. **270**: 2125–2132.

IPCC. 2001. Climate change 2001: impacts, adaptation, and vulnerability (Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change). Cambridge University Press, Cambridge, UK.

Jørgensen, H.B.H., Hansen, M.M., Bekkevold, D., Ruzzante, D.E., and Loeschcke, V. 2005. Marine landscapes and population ge-

© 2007 NRC Canada

netic structure of herring (*Clupea harengus* L.) in the Baltic Sea. Mol. Ecol. **14**: 3219–3234.

Karlsson, S., and Mork, J. 2003. Selection-induced variation at the pantophysin locus (*Pan*I) in a Norwegian fjord population of cod (*Gadus morhua* L.). Mol. Ecol. **12:** 3265–3274.

Lastein, L. 1992. Hydrographic investigations on the Faroe Bank 1985–1992. ICES CM/C:13.

Manel, S., Gaggiotti, O.E., and Waples, R.S. 2005. Assignment methods: matching biological questions with appropriate techniques. Trends Ecol. Evol. **20**: 136–142.

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

Myers, R.A., MacKenzie, B.R., Bowen, K.G., and Barrowman, N.J. 2001. What is the carrying capacity of fish in the ocean? A meta-analysis of population dynamics of North Atlantic cod. Can. J. Fish. Aquat. Sci. **58**: 1464–1476.

Nielsen, E.E., Hansen, M.M., and Loeschcke, V. 1997. Analysis of microsatellite DNA from old scale samples of Atlantic salmon: a comparison of genetic composition over sixty years. Mol. Ecol. **6**: 487–492.

Nielsen, E.E., Hansen, M.M., and Loeschcke, V. 1999. Analysis of DNA from old scale samples; technical aspects, applications and perspectives for conservation. Hereditas, **130**: 265–276.

Nielsen, E.E., Hansen, M.M., Schmidt, C., Meldrup, D., and Grønkjær, P. 2001. Population of origin of Atlantic cod. Nature (London), **413**: 272.

Nielsen, E.E., Hansen, M.M., Ruzzante, D.E., Meldrup, D., and 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, E.E., Hansen, M.M., and Meldrup, D. 2006. 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.

O'Reilly P.T., Canino, M.F., Bailey, K.M., and Bentzen, P. 2000. Isolation of twenty low stutter di- and tetranucleotide microsatellites for population analysis of walleye pollock and other gadoids. J. Fish Biol. **56**: 1074–1086.

O'Reilly, P.T., Canino, M.F., Bailey, K.M., and Bentzen, P. 2004. Inverse relationship between F_{ST} and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. Mol. Ecol. **13**: 1799–1814.

Pampoulie, C., Ruzzante, D.E., Chosson V, Jörunsdóttir, T.D., Taylor, L., Thorsteinsson, V., Danielsdottir, A.K., and Marteinsdottir, G. 2006. The genetic structure of Atlantic cod (*Gadus morhua*) around Iceland: insight from microsatellites, the *Pan* I locus, and tagging experiments. Can. J. Fish. Aquat. Sci. **63**: 2660–2674.

Pella, J., and Masuda, M. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. Fish. Bull., **99**: 151–167.

Pogson, G.H. 2001. Nucleotide polymorphism and natural selection at the pantophysin (*Pan* I) locus in the Atlantic cod, *Gadus morhua* (L.). Genetics, **157**: 317–330.

Pogson, G.H., and Fevolden, S.E. 2003. Natural selection and the genetic differentiation of coastal and Arctic populations of the Atlantic cod in northern Norway: a test involving nucleotide sequence variation at the pantophysin (*Pan*I) locus. Mol. Ecol. **12**: 63–74.

Pogson, G.H., and Mesa, K.A. 2004. Positive Darwinian selection at the pantophysin (*Pan* I) locus in marine gadid fishes.Mol. Biol. Evol. **21**: 65–75.

Pogson, G.H., Mesa, K.A., and Boutilier, R.G. 1995. Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. Genetics, **139**: 375–385.

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

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

Rice, W.R. 1989. Analysing tables of statistical tests. Evolution, **43**: 223–225.

Sarup, P., Sørensen, J.G., Dimitrov, K., Barker, J.S.F., and Loeschcke, V. 2006. Climatic adaptation of *Drosophila buzzatii* populations in southeast Australia. Heredity, **96**: 479–486.

Stenseth, N.C., Ottersen, G., Hurrell, J.W., and Belgrano, A. 2004. Marine ecosystems and climate variation. Oxford University Press, Oxford, UK.

Stenvik, J., Wesmajervi, M.S., Damsgard, B., and Delghandi, M. 2006. Genotyping of pantophysin I (*Pan* I) of Atlantic cod (*Gadus morhua* L.) by allele-specific PCR. Mol. Ecol. Notes, **6**: 272–275.

Swartzman, G., Huang, C., and Kaluzny, S. 1992. Spatial analysis of Bering Sea groundfish survey data using generalized additive models. Can. J. Fish. Aquat. Sci. **49**: 1366–1378.

Waples, R.S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J. Hered. **89**: 439–450.

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

Westerberg, H. 1990. Benthic temperature in the Faroe area. Rep. No. 5. Department of Oceanography, University of Göthenburg, Sweden.

© 2007 NRC Canada