PROCEEDINGS OF THE ROYAL SOCIETY B | BIOLOGICAL SCIENCES

Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature

Ian R. Bradbury, Sophie Hubert, Brent Higgins, Tudor Borza, Sharen Bowman, Ian G. Paterson, Paul V. R. Snelgrove, Corey J. Morris, Robert S. Gregory, David C. Hardie, Jeffrey A. Hutchings, Daniel E. Ruzzante, Chris T. Taggart and Paul Bentzen

Supplementary data	"Data Supplement" http://rspb.royalsocietypublishing.org/content/suppl/2010/06/27/rspb.2010.0985.DC1.html
References	**This article cites 53 articles, 12 of which can be accessed free** http://rspb.royalsocietypublishing.org/content/early/2010/06/24/rspb.2010.0985.full.html#ref-list-1
P<P	Published online 30 June 2010 in advance of the print journal.
Subject collections	Articles on similar topics can be found in the following collections ecology (1798 articles) environmental science (417 articles) evolution (2062 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Proc. R. Soc. B* go to: **http://rspb.royalsocietypublishing.org/subscriptions**

Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature

Ian R. Bradbury[1,*,†], Sophie Hubert[2], Brent Higgins[2], Tudor Borza[2], Sharen Bowman[2], Ian G. Paterson[1], Paul V. R. Snelgrove[3], Corey J. Morris[4], Robert S. Gregory[4], David C. Hardie[1], Jeffrey A. Hutchings[1], Daniel E. Ruzzante[1], Chris T. Taggart[1] and Paul Bentzen[1]

[1]*Marine Gene Probe Laboratory, Department of Biology, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia, Canada B3H4J1*
[2]*Atlantic Genome Centre, Halifax, Nova Scotia, Canada B3H3Z1*
[3]*Ocean Sciences Center and Biology Department, Memorial University of Newfoundland, St John's, NL, PO Box 4200, Newfoundland, Canada A1C 5S7*
[4]*Department of Fisheries and Oceans Canada, St John's, Newfoundland, Canada A1C 5X1*

Despite the enormous economic and ecological importance of marine organisms, the spatial scales of adaptation and biocomplexity remain largely unknown. Yet, the preservation of local stocks that possess adaptive diversity is critical to the long-term maintenance of productive stable fisheries and ecosystems. Here, we document genomic evidence of range-wide adaptive differentiation in a broadcast spawning marine fish, Atlantic cod (*Gadus morhua*), using a genome survey of single nucleotide polymorphisms. Of 1641 gene-associated polymorphisms examined, 70 (4.2%) tested positive for signatures of selection using a Bayesian approach. We identify a subset of these loci ($n = 40$) for which allele frequencies show parallel temperature-associated clines ($p < 0.001$, $r^2 = 0.89$) in the eastern and western north Atlantic. Temperature associations were robust to the statistical removal of geographic distance or latitude effects, and contrasted 'neutral' loci, which displayed no temperature association. Allele frequencies at temperature-associated loci were significantly correlated, spanned three linkage groups and several were successfully annotated supporting the involvement of multiple independent genes. Our results are consistent with the evolution and/or selective sweep of multiple genes in response to ocean temperature, and support the possibility of a new conservation paradigm for non-model marine organisms based on genomic approaches to resolving functional and adaptive diversity.

Keywords: Atlantic cod; natural selection; local adaption; temperature; Atlantic Ocean; single nucleotide polymorphism

1. INTRODUCTION

Marine biocomplexity and diversity have been associated with sustainability and persistence of marine fisheries in the face of climatic and anthropogenic threats (Hilborn *et al.* 2003; Ruzzante *et al.* 2006; Schindler *et al.* 2010). Accordingly, a better understanding of the spatial scale of population structuring and adaptation is paramount to successful conservation in the oceans (Hutchings 2004; Sale *et al.* 2005; Conover *et al.* 2006). Historically, the prevalence of intraspecific diversity, both neutral and adaptive, in many marine species has been unclear (Hellberg 2009). On the one hand, observations of weak genetic structure (Hauser & Carvalho 2008) and

the ubiquity of a pelagic larval stage among marine taxa have fostered the postulate that gene flow is extensive (Levin 2006) and, by extension, that local adaptation is rare (Conover *et al.* 2006). On the other hand, very large population sizes typical of marine species should favour the predominance of natural selection over genetic drift, suggesting that the potential for adaptive diversity in marine populations requires re-evaluation (Hauser & Carvalho 2008), a need that has been reinforced by recent studies that challenge preconceptions of the spatial scales of gene flow and adaptation in the ocean (Levin 2006; Hutchings *et al.* 2007; Gaggiotti *et al.* 2009).

Atlantic cod, *Gadus morhua*, is historically one of the most exploited and studied marine species; however, the degrees of population (e.g. Fox *et al.* 2009) and adaptive structuring (Hutchings *et al.* 2007) remain contentious. Tagging and oceanographic studies indicate a high proportion of resident non-migratory stocks (Robichaud & Rose 2004) as well as discrete spawning locations within

Figure 1. Map of sample locations of Atlantic cod. See electronic supplementary material, table S1 for sample details and figure S1 for fine-scale map of Newfoundland locations. Note that two locations are represented by a single symbol (i.e. Smith Sound A and B, and Georges Bank A and B). Scale bar, 0–500 km.

single management units (Fox *et al.* 2009); both observations support a hypothesis of unrecognized diversity. In contrast, studies of neutral microsatellite loci and mtDNA generally indicate weak spatial structuring over large spatial scales (Bentzen *et al.* 1996; O'Leary *et al.* 2007; Carr & Marshall 2008) consistent with a significant mismatch between the spatial scale of genetic differentiation and management (Reiss *et al.* 2009). Nonetheless, several loci have been identified that display elevated divergence or clinal structure (e.g. Sick 1965; Pogson 2001; Case *et al.* 2005; Nielsen *et al.* 2006), supporting a role for adaptation in the spatial structuring of cod populations. More recently, both Moen *et al.* (2008) and Nielsen *et al.* (2009) used single nucleotide polymorphisms (i.e. SNPs) to examine the presence of local adaptation primarily in the eastern Atlantic, and both identified several potential candidate genes, some of which are associated with ocean climate. In conjunction with molecular studies, common garden experiments suggest adaptive differences in growth and survival in the western Atlantic are also associated with temperature (e.g. Hutchings *et al.* 2007). Nonetheless, it remains unanswered just how pervasive the influence of selection is across the Atlantic cod genome, and the degree to which ocean temperature acts as a selective and population defining agent.

Here, we apply a population genomics approach that combines a genome-wide survey of genetic polymorphism in Atlantic cod with range-wide sampling. This study encompasses the scope of expected environmental and ecological influences in order to explore the geographical scale and magnitude of locus-specific molecular variation. The inclusion of samples from both sides of the North Atlantic that span a large latitudinal gradient provides natural replication of population response to environmental gradients (i.e. clines in allele frequency) and represents an effective approach to the identification of candidate genes associated with ocean climate (Schmidt *et al.* 2008). Such genomic-based understanding of adaptive divergence and the influence of selective forces, both natural and anthropogenic, will be of focal importance as future successful management identifies the appropriate spatial scales of conservation (Hilborn *et al.* 2003; Conover *et al.* 2006; Reiss *et al.* 2009).

2. MATERIAL AND METHODS

(a) *Sample collection and DNA extraction*

In total, 14 locations were sampled during the period 1996–2007 (figure 1 and electronic supplementary material, figure S1), spanning the entire North Atlantic. On average, 22 fish (electronic supplementary material, table S1) were collected per location by handlines or bottom trawl. Fin clips or blood samples were taken and immediately placed in 95 per cent ethanol. DNA was extracted following the protocol of Elphinstone *et al.* (2003) modified to work with a 96-well filter plate and automated on a robotic liquid handling system (Perkin-Elmer). See electronic supplementary material for further details on sample collection and populations sampled.

(b) *Expressed sequence tag libraries, SNP identification, annotation and linkage*

Expressed sequence tag (EST) library creation and sequencing were performed as part of the Cod Genomics and Broodstock Development Project (CGP) and have been described by Bowman *et al.* (2007, 2010). Libraries used 884 individuals (Bowman *et al.* 2010) from Eastern Canada (Bay Bulls, Newfoundland and Cape Sable, Nova Scotia), for which cDNAs were sequenced from the 3′-end (UTR), and contigs were searched for SNPs using Polyphred (Nickerson *et al.* 1997). A total of 3072 potential biallelic SNPs with a minimum of $4\times$ read coverage and a minor allele frequency greater than 25 per cent were selected for analysis. The minor allele frequency of greater than 0.25 was used to ensure that the putative SNPs were not sequencing errors, and were likely to be suitable for mapping purposes (see Hubert *et al.* 2010). Genotyping was performed by the Genome Quebec Innovation Center using the GoldenGate assay (Fan *et al.* 2003). Of the 3072 putative SNPs, 2284 (approx. 74%) were selected following screening, of which 1641 were informative (53%). Further details regarding SNP development, genotyping and linkage mapping are contained in the electronic supplementary material and available in Hubert *et al.* (2009, 2010) and Bowman *et al.* (2007, 2010). To ensure that our samples were representative and to estimate the frequency of rare alleles, two (temperature-associated, see below) SNPs (cgpGmo-S866 and cgpGmo-S419) were genotyped in an additional 500 individuals (adults and juveniles) from southern Newfoundland, using the KASpar genotyping system (KBiosciences).

(c) *Data analysis*

A wide variety of methods have been developed to identify regions of the genome that have been subject to natural selection (Nielsen 2005). We used a Bayesian approach (Beaumont & Balding 2004) that directly estimates the posterior probability of a given locus being under selection by defining two alternative models, one with and one without the effect of selection. The respective posterior probabilities of each model are estimated using a reversible jump Markov chain Monte Carlo approach as implemented in the software BAYESCAN (Foll & Gaggiotti 2008) and compared using the ratio of the two models (i.e. log Bayes Factor). Because geographic isolation or demographic structure may contribute to false-positives (Foll & Gaggiotti 2008), we ran the analysis for the entire dataset and then assessed population structure using principal coordinate analysis (PCA) with significant outliers removed. The BAYESCAN analysis was then repeated separately on each of the dominant clusters with all loci. To evaluate the potential for false-positives in tests for selection, simulated datasets were generated using EASYPOP (Balloux 2001) and analysed using BAYESCAN (see electronic supplementary material for details). We compared the BAYESCAN results to an alternate method for outlier identification using a hierarchical island model to generate the distribution of genetic within and among populations as implemented in ARELQUIN (Excoffier & Lischer 2010).

All SNPs identified as outliers using BAYESCAN were examined for spatial trends in allele frequency. However, given the natural replication of environmental gradients on either side of the Atlantic, and the opportunity this presents for the identification of climate-associated candidate genes (Schmidt *et al.* 2008), we use a conservative approach and focus primarily on loci which were identified as outliers and which display parallel clines in allele frequency on either side of the Atlantic. Multiple regression and simple Mantel tests were used to examine environmental and ecological associations with allele frequency (see electronic supplementary material). We have used mean annual environmental values (e.g. temperatures) as proxies for the regimes experienced by each of the populations analysed here. Notwithstanding its inherent uncertainties, this approach is one that has been used repeatedly in studies of Atlantic cod (Brander 1994; Myers *et al.* 2001; Hutchings *et al.* 2007). Because associations with environmental variables may be by a result of neutral or chance associations, we accounted for the influence of geography by using both partial Mantel tests and the residuals from the isolation by distance relationship (IBD). Linkage among outliers was examined using estimates of D' and significance based on Fisher's exact test calculated in TASSEL 3.0 (Bradbury *et al.* 2007).

To mathematically describe clinal variation, we generated sigmoidal curves of cline data, and calculated clinal width using the inverse of the maximum slope. As allele frequencies were not fixed at 0 and 1 on either side of the cline, clinal width equalled Δp/slope, where Δp is the change in allele frequencies from one side of the cline to the other. We used the gradient model of environmental change (Endler 1977), in which dispersal distance is much smaller than the cline width. See the electronic supplementary material for details.

3. RESULTS

In total, 1641 informative loci were genotyped. On average 87.2 per cent of loci were polymorphic at each geographical location; however, the Baltic Sea sample was successfully genotyped for only 1405 loci owing to the failure of some assays. Locus-specific F_{ST} values were as high as 0.60 for some loci. To test the representative nature of our samples and to examine the frequency of rare alleles, we genotyped 500 individuals of two life-history stages (adult and juvenile young of year) from southern Newfoundland at two loci, cgpGmo-S866 and cgp-GmoS419. The results from these large samples were consistent with those obtained with the smaller samples used in our broad-scale survey: the frequency of rare alleles at cgpGmo-S419 and cgpGmo-S866 were 0.005 and 0.000, respectively.

Tests for selection identified multiple loci potentially experiencing directional selection. PCA with the preliminary outliers removed identified three clusters of individuals, two of which comprised multiple locations (electronic supplementary material, figure S2). Tests for selection using BAYESCAN were repeated using all loci ($n = 1641$) but for each side of the Atlantic separately, excluding the land-locked Ogac Lake populations owing to clear isolation (electronic supplementary material, figure S2) and the Baltic sample owing to reduction in the number of SNPs. The analysis identified 4.2 per cent (70) of the SNPs examined, as potentially influenced by directional selection (figure 2a,b), of which 94 per cent were outliers in both the east and west Atlantic. Simulations were used to evaluate the potential for false positives. Across all simulations, the portion of false positives was consistently less than 0.5 per cent.
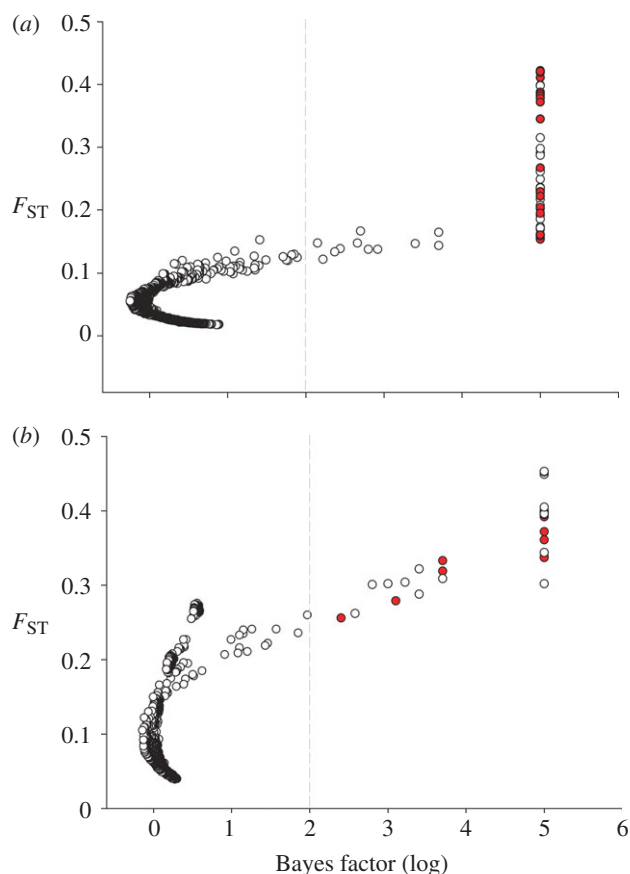
Figure 2. Bayesian test for selection on individual SNPs. Approach as implemented in BAYESCAN for the (*a*) western and (*b*) eastern Atlantic separately. Outliers were defined as those loci that possess a Bayes factor > 100. Red symbols represent outliers significantly associated with temperature. The Ogac Lake and Baltic Sea samples were omitted from outlier analysis (see §2).

In comparison, the test based on the hierarchical island model identified the majority of outliers identified using BAYESCAN (90% at $\alpha = 0.05$, or 75% at $\alpha = 0.01$; electronic supplementary material, figure S3) including 80–90% of the clinal loci.

Individual examination of the allele frequencies at these 70 loci revealed 40 that displayed parallel clines in allele frequency on both sides of the North Atlantic (figure 3*a* and electronic supplementary material, figure S4*a*). In addition to loci displaying parallel clines, a further 14 of the outliers displayed clines in the west but were fixed in the eastern Atlantic (electronic supplementary material, figure S4*b*). Allele frequencies at the remaining SNPs highlighted the isolation of the Barents Sea sample (electronic supplementary material, figure S4*c*). The parallel clines revealed regions of clear transition in allele frequency ($F_{ST} > 0.20$), between the Scotian Shelf and Newfoundland in the west and between Ireland and Iceland in the east (figure 3*a*), where non-outlier loci were only weakly structured ($F_{ST} \sim 0.01$). Clinal loci displayed no variation ($F_{ST} \sim 0.00$) across large geographical areas in the north where non-outlier loci were highly structured ($F_{ST} \sim 0.10$), such as between the isolated landlocked Ogac Lake and the Barents Sea, which are separated by a geographical distance of more than 5000 km (figure 3*a* and electronic supplementary

material, figure S4*a*). Estimates of linkage among clinal loci based on D' or Fisher's exact test revealed that most clinal loci were significantly linked (electronic supplementary material, figure S5).

Multiple regression of the allele frequencies at the parallel clinal loci with seven environmental and biological characters (see §2) identified average annual bottom temperature as the dominant explanatory factor ($r^2 = 0.90$, $p < 0.001$; figure 3*b* and electronic supplementary material, table S3), which is consistent with a link between this subset of loci and ocean temperature. The association between genetic differentiation at these parallel clinal loci and temperature was further supported by simple Mantel tests (electronic supplementary material, table S4), which indicated a significant correlation ($r = 0.84$, $p < 0.001$; electronic supplementary material, table S4). We considered the possibility that the observed association with temperature might be influenced by covariates such as geography or latitude, which reflect colonization history or demography. Average bottom temperature was not associated with latitude directly ($r^2 = 0.02$, $p = 0.677$), and allele frequency remained strongly associated with temperature even after the geographical effect was removed either by analysing the residuals of the IBD relationship or by using partial Mantel tests. Using the residuals of the IBD relationship, the association with temperature remained significant ($r^2 = 0.73$, $p < 0.0001$; electronic supplementary material, figure S6*a*,*b*). When the geographical distance matrix was held constant with a partial Mantel test, the correlation between genetic differentiation and temperature remained strong and highly significant ($r = 0.83$, $p < 0.001$; electronic supplementary material, table S4).

Based on clinal width (Endler 1977) and dispersal distances of 100–1000 km, we estimated that the strength of selection acting on these SNPs (temperature-associated SNPs) in the western Atlantic ranges between 0.00054 and 0.04800 (electronic supplementary material, figure S7). We observed a positive association between pair-wise genetic differentiation (F_{ST}) and temperature difference for the clinal/temperature-associated loci (electronic supplementary material, figure S6*a*). Finally, we noted a temperature-associated decline in the ratio of the estimated gene flow at these loci to the neutral loci to the point that they reached a 90–95% reduction at the maximum temperature differential (electronic supplementary material, figure S6*c*).

Linkage mapping revealed 23 major linkage groups (see Hubert *et al.* 2010). Out of the 40, 38 parallel clinal temperature-associated loci were assigned to three linkage groups (figure 4). Multiple temperature-associated SNPs share a linkage group with the haemoglobin β1 gene, where the polymorphisms associated with the HbI-1/1, HbI-1/2 and HbI-2/2 phenotypes were suggested to occur (Andersen *et al.* 2009; Borza *et al.* 2009), though they are not tightly linked (figure 4). Genotypes of the temperature-associated loci were strongly correlated (electronic supplementary material, figure S5), resulting in three highly discrete classes of multi-locus genotypes that were easily observed using PCA of just the outlier loci (electronic supplementary material, figure S8) each of which are supported by the three linkage groups.
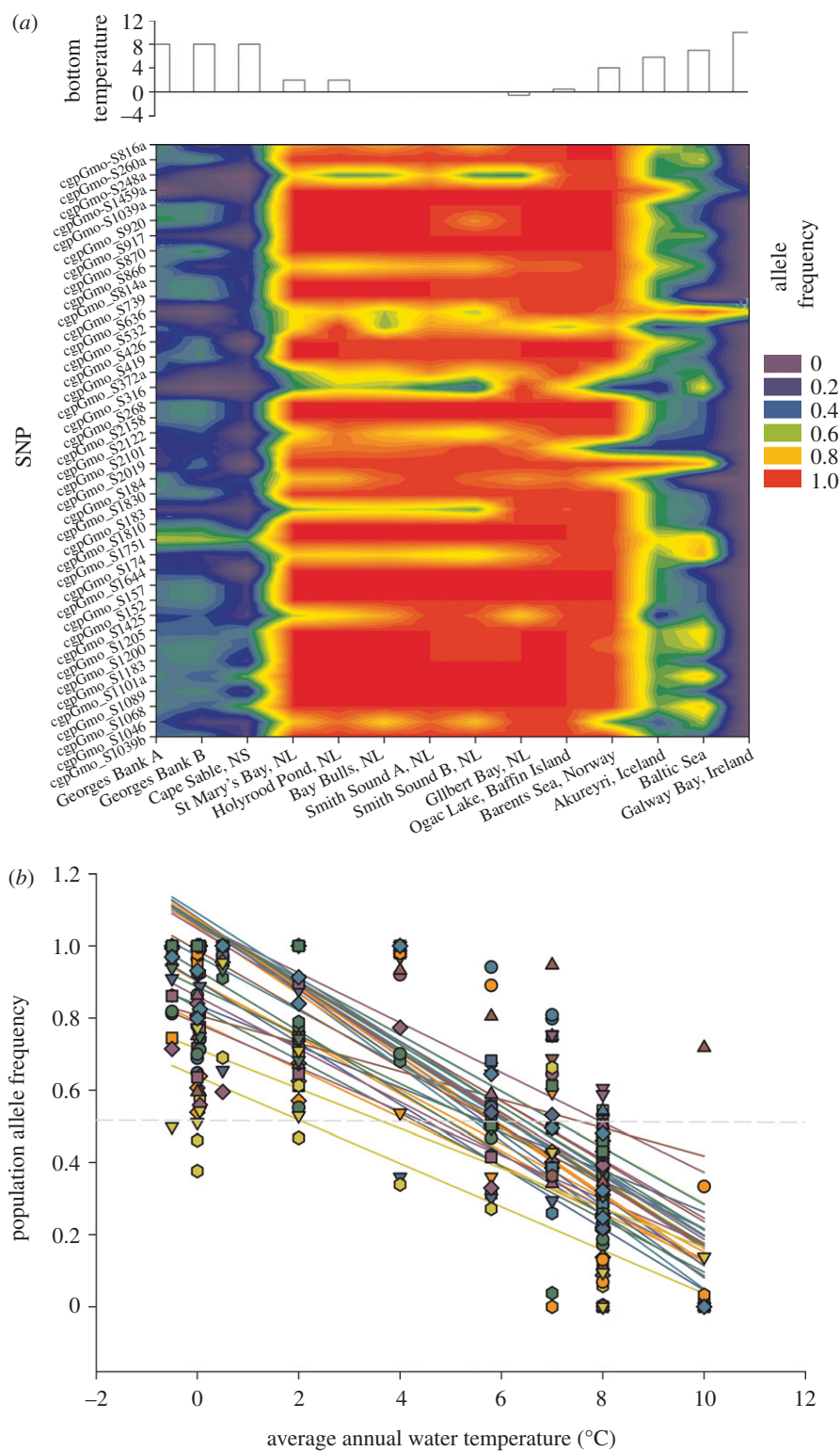
Figure 3. Spatial changes in allele frequency and associations with ocean temperature. (*a*) Histogram of average annual bottom temperature and plot of allele frequencies at clinal outlier SNPs and (*b*) association between allele frequencies for parallel clinal outlier SNPs and average annual bottom temperatures. Data include all locations (see §2). Overall relationship: $r^2 = 0.76$; $p < 0.001$; $y = -0.08 + 0.98x$.

4. DISCUSSION

Despite the enormous economic and ecological importance of marine fish stocks, the appropriate spatial scales for their conservation remain poorly understood (e.g. Reiss *et al.* 2009). Understanding the geographical scale of population structuring and adaptation is paramount to successful conservation as natural populations face unprecedented rates of environmental

change. This study supports previous work hypothesizing that natural selection shapes marine populations on smaller spatial scales than expected, given the high dispersal capability of most marine organisms (e.g. Ruzzante *et al.* 2006; Gaggiotti *et al.* 2009). We show that ocean temperature may be a strong evolutionary force across the genomes of cold-blooded organisms by demonstrating multi-gene patterns of temperature-
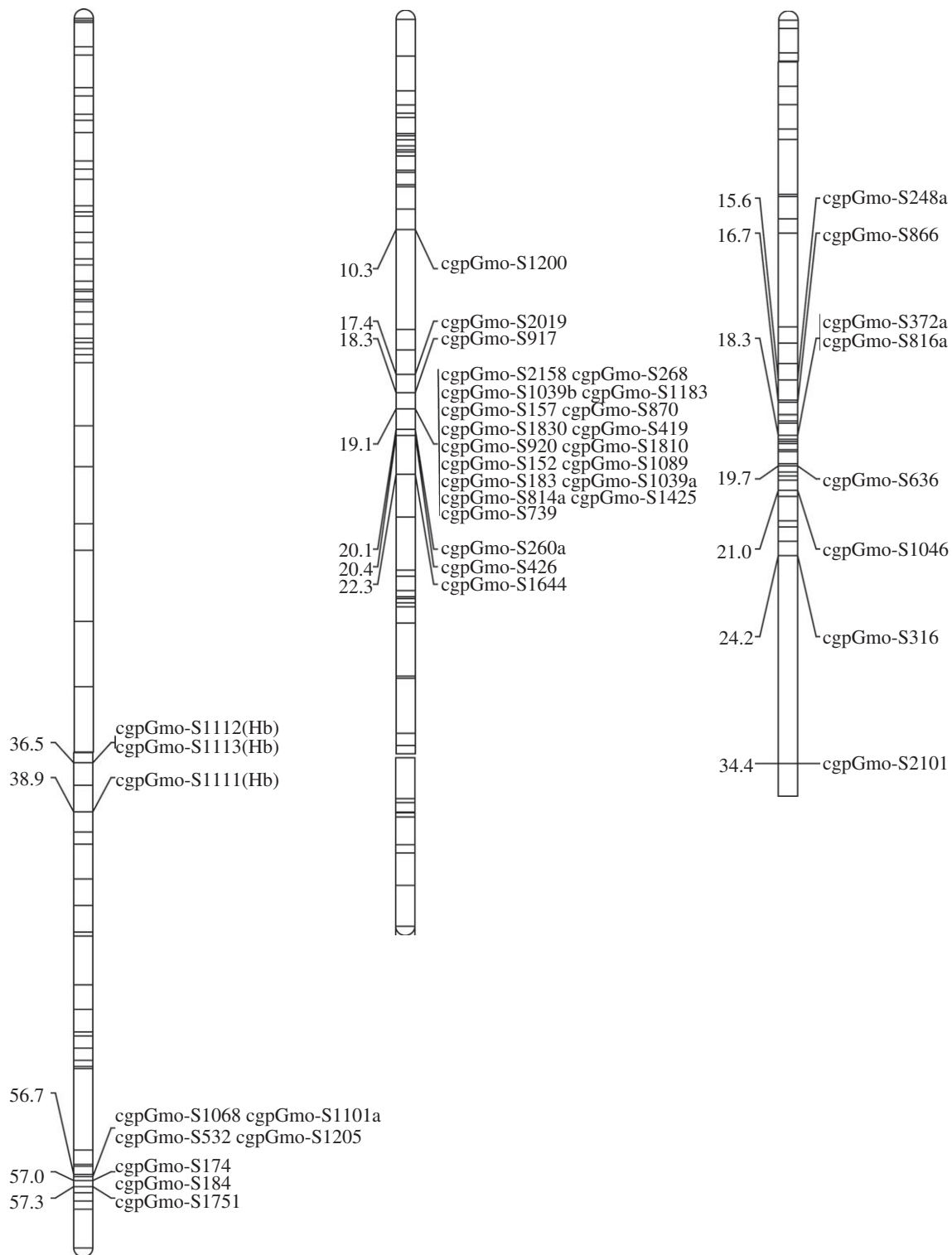
Figure 4. Location of temperature-associated SNPs across the cod genome. Temperature-associated loci spanned three linkage groups (Hb, haemoglobin β1-associated SNPs). See Hubert *et al.* (2010) for details regarding mapping crosses and Borza *et al.* (2009) about the SNPs associated to Hb β1. Only clinal loci shown, and two SNPs could not be placed on the map as they were not informative in any of the families used for mapping. Linkage groups are arranged LG2, LG7 and LG12 from left to right.

related variation in Atlantic cod that occur in parallel on both sides of the Atlantic. We identify a subset of gene-associated polymorphisms for which allele frequencies show parallel temperature-associated clines in otherwise genetically distinct populations on both sides of the Atlantic and test positive for signatures of selection. These SNPs map to three linkage groups and several were successfully annotated consistent with a hypothesis of multiple independent genes. Taken together, these

results suggest that discrete regions of the Atlantic cod genome are subject to directional selection and associated with temperature-related local adaptation.

Previous studies of cod have identified environmental correlations with phenotype or evidence of single molecular loci under selection such as pantophysin (*Pan* I, Pogson & Fevolden 2003), haemoglobin (haemoglobin βI, Andersen *et al.* 2009; Borza *et al.* 2009) and blood type E (Møller 1969) primarily in the eastern Atlantic. Notably, the

clinal loci identified here were not associated with either *Pan*I or haemoglobin β1, though several share a linkage group with the haemoglobin β1 gene. SNP clines recently identified by Moen *et al.* (2008) were similar in nature to *Pan*I clines at four sites in the northeast Atlantic, and these SNPs were associated with several ribosomal proteins and other proteins involved in either immune response or muscle contraction (Moen *et al.* 2008). In addition, Nielsen *et al.* (2009) identified several SNPs ($n = 8$) displaying signatures of directional selection primarily in the eastern Atlantic, none of which overlapped with this study. The annotation of the temperature-associated outliers identified here ($n = 7$; see electronic supplementary material, table S5) suggests highly differing functions and supports a hypothesis that temperature is influencing a wide range of physiological processes. A significant influence of temperature is expected given the precedent set by previous studies and the regulatory role ambient temperature plays in all metabolic and enzymatic reactions of poikilothermic organisms. Nonetheless, the fact that a large portion of outlier SNPs identified here remain unannotated (electronic supplementary material, table S5) highlights the need for further evaluation of the nature of temperature-associated selection across the cod genome.

Many examples exist demonstrating how the thermal environment inhabited by natural populations may directly impose strong selection pressure and result in temperature-associated adaptations (see Angilletta 2009). In cod, temperature has been directly linked to polymorphism at the haemoglobin gene β1 that influences the binding properties of oxygen (Andersen *et al.* 2009), resulting in temperature-associated polymorphism clines in the Atlantic (Sick 1965; Brix *et al.* 1998; Andersen *et al.* 2009). Moreover, temperature is known to influence geographical variation in traits, such as antifreeze production (Goddard *et al.* 1999) and early life-history growth rate (Hutchings *et al.* 2007) in cod. The role of temperature-associated selection as the dominant force in shaping the clines observed here is supported by the positive tests for selection and the strong correlations with ocean climate. Our observation of parallel clines in allele frequencies of the same loci in both the eastern and western Atlantic, despite longstanding historical separation (Bigg *et al.* 2008; Carr & Marshall 2008), is consistent with the hypothesis of natural selection along a replicated, broad-scale environmental gradient (e.g. Schmidt *et al.* 2008). Our estimates of selection based on clinal width (0.0005–0.05), though dependent on the magnitude of gene flow, were comparable if not lower than those reported for other natural populations (e.g. Endler 1986).

Admittedly, clinal patterns may be formed either by selection along an environmental gradient, through secondary contact between previously isolated populations, or some combination of these (e.g. lactate dehydrogenase in *Fundulus*, Schulte 2001). Therefore, an alternate hypothesis for the formation of the temperature-associated clines observed here is that they were formed through secondary contact (see Case *et al.* 2005). During the last glacial cycle, the geographical distribution of cod was likely fragmented, and restricted to as little as 20 per cent of the present day range during the LGM, particularly in the western Atlantic where suitable shelf habitat became very limited as sea levels dropped

(Bigg *et al.* 2008). Nonetheless, it is clear that cod have been continuously present on both sides of the Atlantic for at least 100 000 years, with estimates of divergence between the east and west Atlantic that predate the last glacial maximum (approx. 100–150 kyr, Bigg *et al.* 2008; Carr & Marshall 2008). However, there is no evidence to date of replicated mtDNA clines on either side of the Atlantic coincident with our clines (Árnason 2004; Carr & Marshall 2008). Recent cod whole mtDNA genome data spanning the SNP cline in the western Atlantic has revealed no evidence of clinal structure, and the major haplogroups are distributed across the region (S. M. Carr and H. D. Marshall 2009, personal communication). Moreover, our putatively neutral SNPs show little differentiation within the east and west (electronic supplementary material, figure S2). Although it is not possible to rule out an allopatric mechanism in this context, the replicated associations with temperature fields in the presence of strong east–west isolation support a selective/adaptive hypothesis. Moreover, the observation that contemporary geography (latitude or geographical distance) could not explain the clines on either side of the Atlantic supports the hypothesis that our clines and associations with temperature are robust.

The fact that many temperature-associated SNPs were highly correlated and clustered in several linkage groups leave the number of genes involved in question. Nonetheless, their distribution over several linkage groups, and the successful annotation of the genes encompassing several temperature-associated SNPs, two of which are non-synonymous, supports the hypothesis of coevolution of multiple genes in response to ocean temperature. Moreover, the SNPs were derived from sequencing of the 3′-end of cDNAs, and these expressed sequence tags show no sequence similarity to each other with the exception of S1039a and S1039b, which were identified from a single contig. The hypothesis of independent genes thus seems likely. It remains possible however that some of these loci are not independent and are outliers as a result of a selective sweep, where neighbouring loci display significant environment associations that result from close proximity to one advantageous variant (i.e. hitch-hiking). Although this scenario could explain the correlation of alleles and environmental associations at loci within tight clusters such as in LG7 (figure 4), it is unlikely to explain the correlation of alleles at loci in differing linkage groups (electronic supplementary material, figures S5 and S8). On the other hand, the low dispersion of temperature-associated SNPs across the genome is consistent with a tendency for temperature-associated genes to share regions of the genome and a model of divergent selection with gene flow (Nosil *et al.* 2009).

Because the SNPs in the present study were identified from libraries constructed using DNA from western Atlantic cod, some polymorphisms may post-date the isolation of eastern and western Atlantic cod populations (approx. 100 kyr based on studies of nuclear and mitochondrial DNA markers, Bigg *et al.* 2008), and therefore ascertainment bias could influence inferences made about other regions (e.g. eastern Atlantic). We did observe lower genetic diversity in the eastern Atlantic, which represents a likely effect of ascertainment bias. Moreover, the observation of several outliers displaying clines in one side of the Atlantic while being fixed on

the other may also support the presence of ascertainment issues. However, *trans*-Atlantic differences in allele frequencies are expected and have been observed previously in mtDNA (Árnason 2004; Bigg *et al.* 2008; Carr & Marshall 2008) and nuclear DNA studies (Bentzen *et al.* 1996; Pogson 2001; O'Leary *et al.* 2007). More critically, the fact that the temperature-associated clines were replicated in the western and eastern Atlantic suggests that ascertainment has not biased our interpretation of temperature-associated SNPs as the major conclusion of this study. Furthermore, the clear clustering of individuals by genotype (based on temperature-associated SNPs; electronic supplementary material, figure S8) rather than locality strongly supports the hypothesis that the temperature-associated SNPs are robust to ascertainment bias. Finally, there is some similarity in the trends observed in the temperature-associated outlier SNPs and elsewhere at the microsatellite Gmo132. Previous studies have noted clinal structure at Gmo132 in the eastern Atlantic (Skarstein *et al.* 2007) and western Atlantic (Bentzen *et al.* 1996), which has been attributed to selection (Nielsen *et al.* 2006). There is also clear overlap in allele frequencies at Gmo132 between samples from the Baltic Sea and Georges Bank (Taggart & Cook 1996) as observed here at our clinal SNPs, despite strong isolation.

The conservation of marine species in the face of global climatic change will be dependent on our understanding of species and ecosystem responses to changing selection pressures. Though it remains unclear of how the nature and strength of temperature-dependent selection will change with shifting marine climates, change is nonetheless expected (Vermeij & Roopnarine 2008). Our results support the hypothesis that some warm water alleles are present and available to selection in cold water environments but at low frequencies. Alternatively, a response to ocean warming may involve a northward expansion of southern populations, noting that there is substantial evidence that marine fish move in response to climatic variation (Rose 2005). The genetic clines identified here at a large number of marker loci provide unprecedented insight into both the geographical and genome scale at which selection shapes the adaptive landscape of poikilothermic marine species. Further examinations of these environmental associations, likely through common garden experiments, are required to evaluate the associations observed and the mechanisms involved. Nonetheless, these results support the possibility of a new management and conservation paradigm based on functional diversity in marine systems, as has been established for Pacific salmonids (e.g. Waples 1995; Allendorf *et al.* 1997). Genomic approaches to resolving the spatial scale of functional and adaptive diversity could inform spatial management strategies, and reduce demographic and fisheries variability through the so-called portfolio effects (Hilborn *et al.* 2003; Schindler *et al.* 2010).

REFERENCES

Allendorf, F. W. *et al.* 1997 Prioritizing Pacific salmon stocks for conservation. *Conserv. Biol.* **11**, 140–152. (doi:10.1046/j.1523-1739.1997.95248.x)

Andersen, Ã. I., Wetten, O. F., De Rosa, M. C., Andre, C., Carelli Alinovi, C., Colafranceschi, M., Brix, O. & Colosimo, A. 2009 Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations. *Proc. R. Soc. B* **276**, 833–841. (doi:10.1098/rspb.2008.1529)

Angilletta, M. J. 2009 *Thermal adaptation. A theoretical and empirical synthesis.* New York, NY: Oxford University Press.

Árnason, E. 2004 Mitochondrial cytochrome b DNA variation in the high-fecundity Atlantic cod: trans-Atlantic clines and shallow gene genealogy. *Genetics* **166**, 1871–1885. (doi:10.1534/genetics.166.4.1871)

Balloux, F. 2001 EASYPOP (Version 1.7): a computer program for population genetics simulations. *J. Hered.* **92**, 301–302. (doi:10.1093/jhered/92.3.301)

Beaumont, M. & Balding, D. 2004 Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* **13**, 969–980. (doi:10.1111/j.1365-294X.2004.02125.x)

Bentzen, P., Taggart, C. T., Ruzzante, D. E. & 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. (doi:10.1139/cjfas-53-12-2706)

Bigg, G. R., Cunningham, C. W., Ottersen, G., Pogson, G. H., Wadley, M. R. & Williamson, P. 2008 Ice-age survival of Atlantic cod: agreement between paleoecological models and genetics. *Proc. R. Soc. B* **275**, 163–172. (doi:10.1098/rspb.2007.1153)

Borza, T., Stone, C., Gamperl, A. K. & Bowman, S. 2009 Atlantic cod (*Gadus morhua*) hemoglobin genes: multiplicity and polymorphism. *BMC Genet.* **10**, 51. (doi:10.1186/1471-2156-10-51)

Bowman, S., Higgins, B., Stone, C., Kozera, C., Curtis, B. A., Tarrant Bussey, J., Kimball, J., Verheul, H. & Johnson, S. C. 2007 Generation of genomics resources for Atlantic cod (*Gadus morhua*): progress and plans. *Bull. Aquacult. Assoc. Can.* **105**, 24–30.

Bowman, S. *et al.* 2010 An integrated approach to gene discovery and marker development in Atlantic cod (*Gadus morhua*). *Mar. Biotechnol. (Online first).* (doi:10.1007/s10126-010-9285-z)

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y. & Buckler, E. S. 2007 TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635. (doi:10.1093/bioinformatics/btm308)

Brander, K. M. 1994 The location and timing of cod spawning around the British Isles. *ICES J. Mar. Sci.* **51**, 71–89. (doi:10.1006/jmsc.1994.1007)

Brix, O., Forås, E. & Strand, I. 1998 Genetic variation and functional properties of Atlantic cod hemoglobins: introducing a modified tonometric method for studying fragile hemoglobins. *Comp. Biochem. Physiol.* **119A**, 575–580. (doi:10.1016/S10956433(97)00469-8)

Carr, S. M. & Marshall, H. D. 2008 Intraspecific phylogeographic genomics from multiple complete mtDNA genomes in Atlantic cod (*Gadus morhua*): origins of the 'Codmother' trans-Atlantic vicariance, and mid-glacial population expansion. *Genetics* **108**, 381–389.

Case, R. A. J., Hutchinson, W. F., Hauser, L., Van Oosterhout, C. & 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. (doi:10.3354/meps301267)

Conover, D. O., Clarke, L. M., Munch, S. B. & Wagner, G. N. 2006 Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *J. Fish Biol.* **69**, 21–47. (doi:10.1111/j.1095-8649.2006.01274.x)

Elphinstone, M. S., Hinten, G. N., Anderson, M. J. & Nock, C. J. 2003 An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Mol. Ecol. Notes* **3**, 317–320. (doi:10.1046/j.1471-8286.2003.00397.x)

Endler, J. A. 1977 *Geographical variation, speciation, and clines.* Princeton, NJ: Princeton University Press.

Endler, J. A. 1986 *Natural selection in the wild.* Princeton, NJ: Princeton University Press.

Excoffier, L. & Lischer, H. E. L. 2010 Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567.

Fan, J. B. *et al.* 2003 Highly parallel SNP genotyping. *Cold Spring Harb. Symp. Quant. Biol.* **68**, 69–78. (doi:10.1101/sqb.2003.68.69)

Foll, M. & Gaggiotti, O. E. 2008 A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977–993. (doi:10.1534/genetics.108.092221)

Fox, C. J. *et al.* 2009 Mapping the spawning grounds of North Sea cod (*Gadus morhua*) by direct and indirect means. *Proc. R. Soc. B* **275**, 1543–1548. (doi:10.1098/rspb.2008.0201)

Gaggiotti, O., Bekkevold, D., Jorgensen, H. *et al.* 2009 Disentangling effects of evolutionary, demographic and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. *Evolution* **63**, 2939–2951. (doi:10.1111/j.1558-5646.2009.00779.x)

Goddard, S. V., Kao, M. & Fletcher, G. L. 1999 Population differences in antifreeze production cycles of juvenile Atlantic cod (*Gadus morhua*) reflect adaptations to overwintering environment. *Can J. Fish. Aquat. Sci.* **56**, 1991–1999. (doi:10.1139/cjfas-56-11-1991)

Hauser, L. & Carvalho, G. R. 2008 Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish Fish.* **9**, 333–362. (doi:10.1111/j.1467-979.2008.00299.x)

Hellberg, M. E. 2009 Gene flow and isolation among populations of marine animals. *Annu. Rev. Ecol. Evol. Syst.* **40**, 291–310. (doi:10.1146/annurev.ecolsys.110308.120223)

Hilborn, R., Quinn, T. P., Schindler, D. E. & Rogers, D. E. 2003 Biocomplexity and fisheries sustainability. *Proc. Natl Acad. Sci. USA* **100**, 6564–6568. (doi:10.1073/pnas.1037274100)

Hubert, S., Tarrant Bussey, J., Higgins, B., Curtis, B. A. & Bowman, S. 2009 Development of single nucleotide polymorphism markers for Atlantic cod (*Gadus morhua*) using expressed sequences. *Aquaculture* **296**, 7–14. (doi:10.1016/j.aquaculture.2009.07.027)

Hubert, S., Higgins, B., Borza, T. & Bowman, S. 2010 Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*). *BMC Genom.* **11**, 191. (doi:10.1186/1471-2164-11-191)

Hutchings, J. A. 2004 The cod that got away. *Nature* **428**, 899–900. (doi:10.1038/428899a)

Hutchings, J. A., Swain, D. P., Rowe, S., Eddington, J. D., Puvanendran, V. & Brown, J. A. 2007 Genetic variation in life-history reaction norms in a marine fish. *Proc. R. Soc. B* **274**, 1693–1699. (doi:10.1098/rspb.2007.0263)

Levin, L. A. 2006 Recent progress in understanding larval dispersal: new directions and digressions. *Integr. Comp. Biol.* **46**, 282–297. (doi:10.1093/icb/icj024)

Moen, T., Hayes, B., Nilsen, F., Delghandi, M., Fjalestad, K. T., Fevolden, S.-E., Berg, P. & Lien, S. 2008 Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. *Genetics* **9**, 218–223. (doi:10.1186/1471-2156-9-18)

Moller, D. 1969 The relationship between Arctic cod and coastal cod in their immature stages illustrated by frequencies of genetic characters. *Fiskridir. Skr. Havunders* **15**, 220–223.

Myers, R. A., Mackenzie, B. R., Bowen, K. G. & Barrowman, N. J. 2001 What is the carrying capacity for fish in the ocean? A meta-analysis of population dynamics of North Atlantic cod. *Can. J. Fish. Aquat. Sci.* **58**, 1464–1476. (doi:10.1139/cjfas-58-7-1464)

Nickerson, D. A., Tobe, V. O. & Taylor, S. L. 1997 PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucl. Acids Res.* **25**, 2275–2745. (doi:10.1093/nar/25.14.2745)

Nielsen, R. 2005 Molecular signatures of natural selection. *Annu. Rev. Genet.* **39**, 197–218. (doi:10.1146/annurev.genet.39.073003.112420)

Nielsen, E., Hansen, M. & Meldrup, D. 2006 Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Mol. Ecol.* **15**, 3219–3229. (doi:10.1111/j.1365-294X.2006.03025.x)

Nielsen, E. *et al.* 2009 Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evol. Biol.* **9**, 276. (doi:10.1186/1471-2148-9-276)

Nosil, P., Funk, D. J. & Ortiz-Barrientos, D. 2009 Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **18**, 375–402. (doi:10.1111/j.1365-294X.2008.03946.x)

O'Leary, D. B., Coughlan, J., Dillane, E., McCarthy, T. V. & Cross, T. F. 2007 Microsatellite variation in cod *Gadus morhua* throughout its geographic range. *J. Fish Biol.* **70**, 310–335. (doi:10.1111/j.1095-8649.2007.01451.x)

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. & Fevolden, S. E. 2003 Natural selection and the genetic differentiation of coastal and Arctic populations of Atlantic cod in northern Norway: a test involving nucleotide sequence variation at the pantophysin (*Pan*I) locus. *Mol. Ecol.* **12**, 63–74. (doi:10.1046/j.1365-294X.2003.01713.x)

Reiss, H., Hoarau, G., Dickey-Collas, M. & Wolff, W. J. 2009 Genetic population structure of marine fish: mismatch between biological and fisheries management units. *Fish Fish.* **10**, 361–395. (doi:10.1111/j.1467-2979.2008.00324.x)

Robichaud, D. & Rose, G. A. 2004 Migratory behaviour and range in the Atlantic cod: inference from a century of tagging. *Fish Fish.* **5**, 185–214. (doi:10.1111/j.1467-2679.2004.00141.x)

Rose, G. A. 2005 On the distributional responses of North Atlantic fish to climate change. *ICES J. Mar. Sci.* **62**, 1360–1374. (doi:10.1016/j.icesjms.2005.05.007)

Ruzzante, D. E. *et al.* 2006 Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. *Proc. R. Soc. B* **273**, 1459–1464. (doi:10.1098/rspb.2005.3463)

Sale, P. F. *et al.* 2005 Critical science gaps impede use of no-take fishery reserves. *Trends Ecol. Evol.* **20**, 74–80. (doi:10.1016/j.tree.2004.11.007)

Schindler, D. E., Hilborn, R., Chasco, B., Boatright, C. P., Quinn, T. P., Rogers, L. A. & Webster, M. S. 2010 Population diversity and the portfolio effect in an exploited species. *Nature* **465**, 609–612. (doi:10.1038/nature09060)

Schmidt, P. S. *et al.* 2008 Ecological genetics in the North Atlantic: environmental gradients and adaptation at specific loci. *Ecology* **89**, S91–S107. (doi:10.1890/07-1162.1)

Schulte, P. M. 2001 Environmental adaptations as windows on molecular evolution. *Comp. Biochem. Phys. B* **128**, 597–611. (doi:10.1016/S1096-4959(00)00357-2)

Sick, K. 1965 Haemoglobin polymorphism of cod in the North Sea and North Atlantic Ocean. *Hereditas* **54**, 49–69. (doi:10.1111/j.1601-5223.1965.tb02005.x)

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

Taggart, C. T. & Cook, D. 1996 Microsatellite genetic analysis of Baltic cod: a preliminary report. Dalhousie University, Halifax, NS. Manuscript Report, Marine Gene Probe Laboratory, 13p.

Vermeij, G. J. & Roopnarine, P. D. 2008 Ecology: the coming Arctic invasion. *Science* **321**, 780–781. (doi:10.1126/science.1160852)

Waples, R. S. 1995 Evolutionarily significant units and the conservation of biological diversity under the Endangered Species Act. *Am. Fish. Soc. Symp.* **17**, 8–27.