McGill, L, McDevitt, AD, Hellemans, B, Neat, F, Knutsen, H, Mariani, S, Christiansen, H, Johansen, T, Volckaert, FAM and Coscia, I

Population structure and connectivity in the genus Molva in the Northeast Atlantic

ICES Journal of Marine Science, 2023, **80**, 1079–1086
DOI: 10.1093/icesjms/fsad040
Advance access publication date: 17 March 2023
Original Article

Population structure and connectivity in the genus *Molva* in the Northeast Atlantic

L. McGill [1,2], A. D. McDevitt[1,3], B. Hellemans [4], F. Neat [5], H. Knutsen[6,7], S. Mariani[8], H. Christiansen[4], T. Johansen [9], F. A. M. Volckaert [4], and I. Coscia [1,4,10,*]

[1]School of Science, Engineering and Environment, University of Salford, Salford M4 5WT, UK
[2]Inverness College, University of the Highlands and Islands, Inverness IV2 5NA, UK
[3]Department of Natural Sciences and the Environment, Atlantic Technological University, Galway H91 T8NW, Ireland
[4]Laboratory of Biodiversity and Evolutionary Genomics, KU Leuven, Leuven B3000, Belgium
[5]World Maritime University, Malmö WMU: SE-201 24, Sweden
[6]Institute of Marine Research, His 4817, Norway
[7]Centre for Coastal Research, Department of Natural Sciences, University of Agder, Kristiansand 4604, Norway
[8]School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool L3 5UG, UK
[9]Institute of Marine Research, Tromsø 9019, Norway
[10]Marine Institute, Rinville, Oranmore, Co. Galway H91R673, Ireland

[*] *Corresponding author:tel: +353 (0)91 387584;* e-mail: Ilaria.Coscia@marine.ie.

In fisheries, operational management units and biological data often do not coincide. In many cases, this is not even known due to the lack of information about a species' population structure or behaviour. This study focuses on two such species, the common ling *Molva molva* and the blue ling *M. dypterygia*, two Northeast Atlantic gadoids with overlapping geographical distribution, but different depth habitats. Heavily exploited throughout their ranges, with declining catches, little is known about their population structure. Genotyping-by-sequencing at thousands of genetic markers indicated that both species are separated into two major groups, one represented by samples from the coasts of western Scotland, Greenland, and the Bay of Biscay and the other off the coast of Norway. This signal is stronger for the deeper dwelling blue ling, even though adult dispersal was also identified for this species. Despite small sample sizes, fine-scale patterns of genetic structure were identified along Norway for common ling. Signatures of adaptation in blue ling consisted in signs of selections in genes involved in vision, growth, and adaptation to cold temperatures.

Keywords: deep sea, fisheries, fjords, genomics, ling, *Molva dypterygia*, *Molva molva*, SNPs.

Introduction

The first step towards the development of efficient conservation and management plans is the reliable delimitation of biological units that translate into management units, or *stocks* in the case of exploited marine species (Cadrin, 2020). In many instances, fisheries stocks are delineated and managed not on a biological basis, but on a political one (Reiss *et al.*, 2009), which hampers efforts towards sustainable exploitation. In addition, it is paramount to understand how these units are formed and maintained at various temporal and spatial scales (Hidalgo *et al.*, 2017) and to estimate the extent of connectivity between them. The latter is often a stumbling block, especially in marine fishes, as large effective population sizes and large dispersal capabilities result in weak genetic population structure (Waples, 1998). Genomic approaches represent a significant improvement in our ability to reliably identify populations, assess connectivity, and estimate adaptive potentials (Bernatchez *et al.*, 2017). Methods like genomic scans produce thousands of markers that can be screened in hundreds of individuals in non-model organisms, often from across environmental gradients, and have the potential to estimate connectivity and the effects of natural selection on populations (Maroso *et al.*, 2018; Barth *et al.*, 2019). When paired with a reference genome, genomic scans can be useful in identifying regions under selection in response to environmental variability (Manel *et al.*, 2016). The availability of a large number of neutral and adaptive markers from these techniques allows the discovery of barriers to gene flow in a range of marine habitats, from intertidal (Coscia *et al.*, 2020) to deep-sea ecosystems (Gaither *et al.*, 2018; Gonçalves da Silva *et al.*, 2020).

Fishes living in the deeper layers of the oceans (<200 m) are characterized by slow metabolic and growth rates, K-selection with late maturation, and long generation times, which make them particularly vulnerable to overexploitation (Clarke *et al.*, 2015). Many of these fish stocks have declined since the onset of deep-sea fisheries in the 1970s, but sustainable levels of exploitation could be achieved or maintained using science-based management (Hilborn *et al.*, 2020). Until the advent of genomic scans, understanding population structure and connectivity in deep-sea fishes has been difficult, leading to the hypothesis that these populations are genetically more homogeneous than their shallower counterparts (Cunha *et al.*, 2012; Coscia *et al.*, 2018). Recent and growing evidence has shown so far that high-resolution markers, such as single-nucleotide polymorphisms (SNPs), can provide important information about population structure across space and depth ranges (Gonçalves da Silva *et al.*, 2020).

Received: 23 November 2022; **Revised:** 24 February 2023; **Accepted:** 28 February 2023

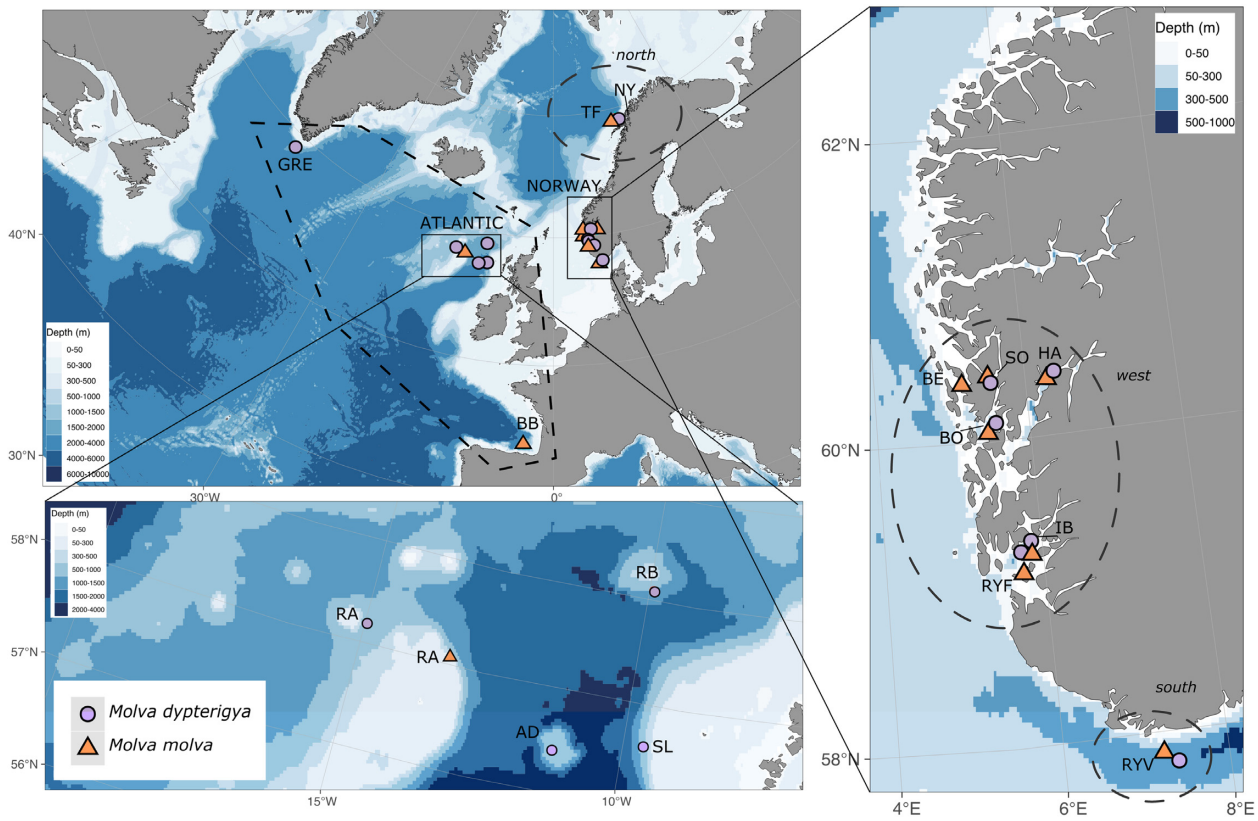

Figure 1. Maps of the sampling sites for common ling, *M. molva* (orange triangle), and blue ling, *M. dypterygia* (purple circles), across the Northeast Atlantic.

The common ling *Molva molva* (Linnaeus, 1758) and the blue ling *M. dypterygia* (Pennant, 1784) (Lotidae; Gadiformes) are distributed across the North Atlantic Ocean with largely overlapping geographical distributions but with significant differences in depth preference and behaviour, especially in reproduction. *Molva molva* is common on rocky bottoms between 50 and 500 m, reaches maturity at 5–6 years of age, and spawns in spring and early summer (Cohen *et al.*, 1990). *Molva dypterygia* dwells between 350 and 1500 m, reaches maturity at 9–11 years of age, forms spawning aggregations at depths of 500–1000 m between April and May (Large *et al.*, 2010), and are targeted by fisheries. Landings peaked in the 1980s at around 10000 tonnes, but have since declined to <1000 tonnes in recent years (Vieira *et al.*, 2019). Exploited for decades throughout their range, both as target and as bycatch, the two species have the potential to support sustainable exploitation, if appropriate management is in place (Edwards *et al.*, 2012; ICES, 2020); hence, stakeholders and managers call for genetic information to better understand stock structure and sustainability (ICES, 2020). Genetic structure has been previously investigated using microsatellites only for *M. molva*, and weak structure was identified between populations off the west coast of Scotland and Iceland and along the Norwegian coast (Blanco Gonzalez *et al.*, 2015).

Here, we report an analysis of the population genomics of *M. molva* and *M. dypterygia* throughout most of their geographical range. Using genotyping-by-sequencing (GBS) and mapping thousands of SNPs against a recently published *M. molva* reference genome (Malmstrøm *et al.*, 2017), we aim to explore genetic structure in both species in the context of their contrasting life history and ecological traits. The goals of this study are (1) to test whether genomic approaches can uncover patterns of neutral variation linked to dispersal and genetic connectivity and (2) to investigate genetic signatures of local adaptation as a further tool for exploring diverging units. This information is relevant to the management of *Molva* spp. fisheries and expands on our understanding of the evolutionary dynamics that shape marine connectivity between the continental shelf and the deep ocean.

Material and methods

Sampling and laboratory procedures

Samples were collected between 2005 and 2015 by both fishing and research vessels (Figure 1, Table 1). In total, 190 individuals of *M. dypterygia* (blue ling) and 83 of *M. Molva* (common ling) were available. After DNA extraction, libraries were prepared using a GBS protocol using the *ApeKI* enzyme and including a final size-selection step. Details about DNA extraction and library preparation are available in the Supplementary Material.

Bioinformatics analysis

Details of the bioinformatics are provided in the Supplementary Material. The raw data were quality controlled using *FastQC* (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and demultiplexed and filtered using *STACKS* 2.4 (Rochette *et al.*, 2019), with both *de novo* and the reference mapping approaches. After testing different combinations, the STACKS parameters chosen were $m = 3$, $M = 4$, and $n = 5$.

Table 1. Samples included in this study.

Species	Group	Location	Code	ID	Lat (DD)	Long (DD)	Year	Depth	N
Molva molva	Atlantic	Bay of Biscay	BB	MM-BB15	44	−3	2015		2
	Atlantic	Rockall	RA	MM-RA08	58.08	−13.35	2008		8
	Atlantic	Rockall	RA	MM-RA14	57.79	−13.44	2014	160	6
	West Norway	Bergen	BE	MM-BE08	60.39	5.16	2008		8
	West Norway	Bømlafjorden	BO	MM-BO14	60.06	5.45	2014		10
	West Norway	Hardangerfjord	HA	MM-HA14	60.38	6.28	2014		10
	West Norway	Indre Boknafjord	IB	MM-IB13	59.28	5.85	2014	110	1
	North Norway	Nygrunnen	NY	MM-NY13	69.18	14.5	2013	144–620	10
	West Norway	Ryfylke	RYF	MM-RYF13	59.13	5.73	2013		2
	South Norway	Ryvingen	RYV	MM-RYV14	57.88	7.21	2014	300–380	10
	West Norway	Sørfjorden	SO	MM-SO14	60.43	5.51	2014		6
	North Norway	Tromsoflaket	TF	MM-TF05	69.02	13.44	2005		10
Molva dypterygia	Atlantic	Anton Dohrn bank	AD	MD-AD07	57.42	−11.22	2007	650	1
	Atlantic	Greenland	GRE	MD-GRE15	59	−44	2015		10
	Atlantic	Rockall	RA	MD-RA07	56.95	−13.43	2007	680	5
	Atlantic	Rockall	RA	MD-RA10	56.01	−14.87	2010	805–830	40
	Atlantic	Rockall	RA	MD-RA11	58.2	−14.96	2011	500–620	47
	Atlantic	Rosemary Bank	RS	MD-RS07	59.1	−9.92	2007	820–910	13
	Atlantic	Atlantic slope	SL	MD-SL07	57.61	−9.63	2007	760–1000	19
	Atlantic	Atlantic slope	SL	MD-SL11	58.5	−9	2011	800	22
	Atlantic	Atlantic slope	SL	MD-SL14	56.72	−9.1325	2014	940	4
	West Norway	Bømlafjorden	BO	MD-BO14	60.06	5.45	2014		4
	West Norway	Hardangerfjord	HA	MD-HA14	60.39	6.28	2014		3
	West Norway	Indre Boknafjord	IB	MD-IB13	59.28	5.85	2013	40–150	1
	North Norway	Nygrunnen	NY	MD-NY13	69.18	14.5	2013	144–620	12
	West Norway	Ryfylke	RYF	MD-RYF13	59.27	5.72	2013	125–400	2
	South Norway	Ryvingen	RYV	MD-RYV14	57.88	7.21	2014		6
	West Norway	Sørfjorden	SO	MD-SO13	6.43	5.51	2013		1

Group represents the pool that each *Location* falls into; each *Location* has an associated *Code* and *ID*; sampling coordinates in decimal degrees (*Lat*. and *Long*. DD); *Year* and *Depth* denote the year and depth of sampling; and *N* denotes the number of individuals sampled.

Markers under selection

Neutrality was tested with two approaches implemented in *pcadapt* (Luu *et al*., 2017) and the *lfmm* function in *LEA* (Frichot and François, 2015). *pcadapt* is based on a principal component analysis (PCA) of individual genotypes and performs particularly well in the presence of weak structure, admixture, or range expansions (Luu *et al*., 2017). We chose the most appropriate number of clusters *K* in the scree plot, which displays in decreasing order the percentage of variance explained by each principal component. *Q*-values were used to control the false discovery rate, and SNPs were considered as significant outliers at alpha values ≤0.01. To minimize the detection of false positives, we considered as outliers only SNPs selected by both methods. The flanking regions of these SNPs were extracted and mapped against the *M. molva* genome (assembled to the scaffold level). The selected scaffolds were then BLASTed, including 10 kb up- and downstream. *MEGAblast* and *blastn* were used to search the scaffolds with outlier SNPs against the NCBI database (cut-off *e-value* of 1×10^{-10}).

Neutral population structure

For population analyses, single locations were grouped to increase the number of individuals per sample after preliminary investigations failed to detect significant substructure. "Atlantic" included all non-Norwegian locations, whereas Norway was split into three subgroups: "South Norway", "West Norway", and "North Norway" (Figure 1, Table 1). Expected (H_E) and observed (H_O) heterozygosities were estimated in *hierfstat* (Goudet, 2005; Goudet and Jombart, 2015), and pairwise F_{ST} (Weir and Cockerham, 1984) and relative 95% confidence interval (1000 bootstraps) were estimated with *assigner* (Gosselin *et al*., 2016). Population structure was first estimated with individual-based multivariate approaches like PCA and Discriminant Analysis of Principal Components (DAPC), performed using, respectively, the function *dudi.pca* in *ade4* (Dray and Dufour, 2007) and the function *dapc* in *adegenet* (Jombart, 2008; Jombart and Ahmed, 2011). Data were further explored with the Bayesian-based approach in *Structure* (Pritchard *et al*., 2000), with a burn-in of 20000 and 180000 iterations, to determine optimal *K* (between 1 and 5).

Results

The first FastQC quality control revealed that sequencing quality was high, the drop-off towards the end of the reads (126 bp) was minimal, and no reads were flagged as "poor" across both species (Supplementary Table S1). Three libraries were prepared, one for common ling and two for blue ling: After demultiplexing, 480039207 reads for the common ling and 1019877898 (from the two libraries) for the blue ling were retained in total.

The final datasets included 2983 SNPs for 83 individuals of common ling for the *de novo* and 6569 SNPs for the reference-based analysis and 2118 SNPs for 190 individuals of blue ling for the *de novo* and 3078 SNPs for the reference-based analysis. Subsequent statistical analyses were carried out on both datasets (*de novo* and reference-based) and compared. No differences in the inferred structure were detected; hence, we retained (as detailed below) the reference-based data, as more SNPs were retained, and reference-mapped data can be used to identify genes potentially under selection.

Table 2. Genetic diversity and F_{ST} for common and blue ling.

Common ling	*N*	H_O	H_E	*Ar*	F_{IS}	Atlantic	South Norway	West Norway	North Norway
Atlantic	16	0.27	0.25	1.84	**0.05**		**0.009**	**0.007**	**0.007**
South Norway	10	0.26	0.25	1.84	**−0.015**			**0.002**	**0.003**
West Norway	37	0.25	0.25	1.85	**−0.004**				0.001
North Norway	20	0.25	0.25	1.84	−0.004				
Blue ling									
Atlantic	161	0.27	0.27	1.27	−0.006		**0.013**	**0.015**	**0.014**
South Norway	6	0.27	0.27	1.27	**−0.012**			0.001	**0.006**
West Norway	10	0.26	0.27	1.27	**0.010**				**0.005**
North Norway	12	0.30	0.28	1.28	**−0.055**				

N, number of individuals remaining after filtering for each location; H_E, expected heterozygosity; H_O, observed heterozygosity; A_r, allelic rishness; F_{IS}, inbreeding coefficient. Pairwise F_{ST} is above diagonal. F_{IS} and F_{ST}, values in bold are significant (95% confidence interval).

Signatures of selection

For the common ling, *pcadapt* identified 43 SNPs under selection and the *lfmm* function identified 5. Only three were shared between both methods, and these were chosen as true outliers and removed from neutral population structure analysis. For blue ling, *pcadapt* identified 8 SNP outliers, and the *lfmm* function identified 11. Five were shared between the two approaches and were considered as outliers. After BLASTing these three and five outlier SNPs for common and blue ling, respectively, no significant matches were found for common ling. Significant matches were found for outlier SNPs for blue ling. Those with the smallest *e*-values included genes associated with opsin-coding (vision), hepcidin-coding (regulation of iron absorption), antifreeze glycoprotein, HGDFL3 and glucose transporter, and somatotropin-2 (important for osmoregulation; Supplementary Tables S3–S7). Population structure could not be explored using this dataset, given the small number of outliers identified.

Population structure with neutral markers

The final neutral dataset for common ling included the remaining 6566 SNPs. For the blue ling, a final, putatively neutral, dataset of 3073 SNPs was used. Population H_E and H_O heterozygosities were similar across species and across groups within species (Figure 1, Table 1), varying between 0.25 and 0.30 (Table 2). Inbreeding (F_{IS}) values were small but significantly larger than zero for most groups in both species. F_{IS} ranged between −0.004 (West Norway) and 0.05 (Atlantic) for common ling, and between −0.05 (North Norway) and 0.01 (West Norway) for blue ling. Pairwise F_{ST} were significant but small for both species (Table 2). The largest pairwise F_{ST} occurred between Atlantic and Norwegian populations in both species. Overall F_{ST} was 0.007 for the blue ling and 0.004 for the common ling (Table 2).

PCA resolved two groups within each species, one Atlantic and one Norwegian. This pattern was more striking in blue ling (Figure 2a) than in common ling, where the two groups were closer in multivariate space (Figure 2b). In common ling, individuals from Bergen (BE) and Sørfjorden (SO) are included in a third group (Figure 2a), also highlighted in other approaches below. For *Structure*, the optimal *K* was 2 in blue ling (Supplementary Figure S1). Both the PCA and *Structure* revealed migration in blue ling, with two F_0 migrants of Norwegian ancestry caught in the Atlantic (one in the Slope, SL, and one Rosemary Bank, RB) and one Atlantic individual caught in the North Norway region (Nygrunnen, NY).

The Norwegian samples of common ling were spread out on the PCA, whereas the Atlantic samples grouped more closely together. This was not due to single populations diverging, but rather to specific individuals. Fish farther from the centroid of the distribution belonged to the West Norway Bergen (BE) and Sørfjorden (SO). They are more evident in the *Structure* barplot (Figure 2d). In the first barplot ($K = 2$), the second cluster is made up of three individuals from Bergen and, to some extent, to three individuals from Sørfjorden (SO). The second barplot ($K = 3$), further separated them from the rest, which then were split into Atlantic and Norwegian samples. The chosen *K* value for the common ling according to Mean(K) was $K = 4$ (Supplementary Figures S1 and S2), but the "fourth" cluster was not biologically meaningful, hence the most parsimonious choice (also considering the outcome of the PCA) was to select three as the most likely *K*. The DAPC analysis also confirmed the observed pattern and supported separation between North Norway and the West-South Norwegian samples, as observed in pairwise F_{ST} (Supplementary Figure S3).

Discussion

This study explored patterns of population connectivity and adaptive divergence in two congeneric fishes, the common ling, *Molva molva*, and the blue ling, *M. dypterygia*. Results point towards two genetic units in the deeper blue ling (Atlantic and Norway), and three in common ling (Atlantic, Norway, and Bergen and Sørfjorden), providing managers with crucial information to define stock boundaries and also offering important insights into fine-scale processes, specifically within Norwegian fjords. Outlier loci reveal signatures of selection at genes involved in regulating osmosis and physiological responses to cold temperatures, hypoxia, and starvation.

Broad and fine scale population structure

The use of thousands of markers scattered across the genome are more likely to detect patterns of weak population structure and local adaptation in marine fish populations than "traditional" markers, such as microsatellites (Allendorf *et al.*, 2010; Coscia *et al.*, 2012). This is especially useful for

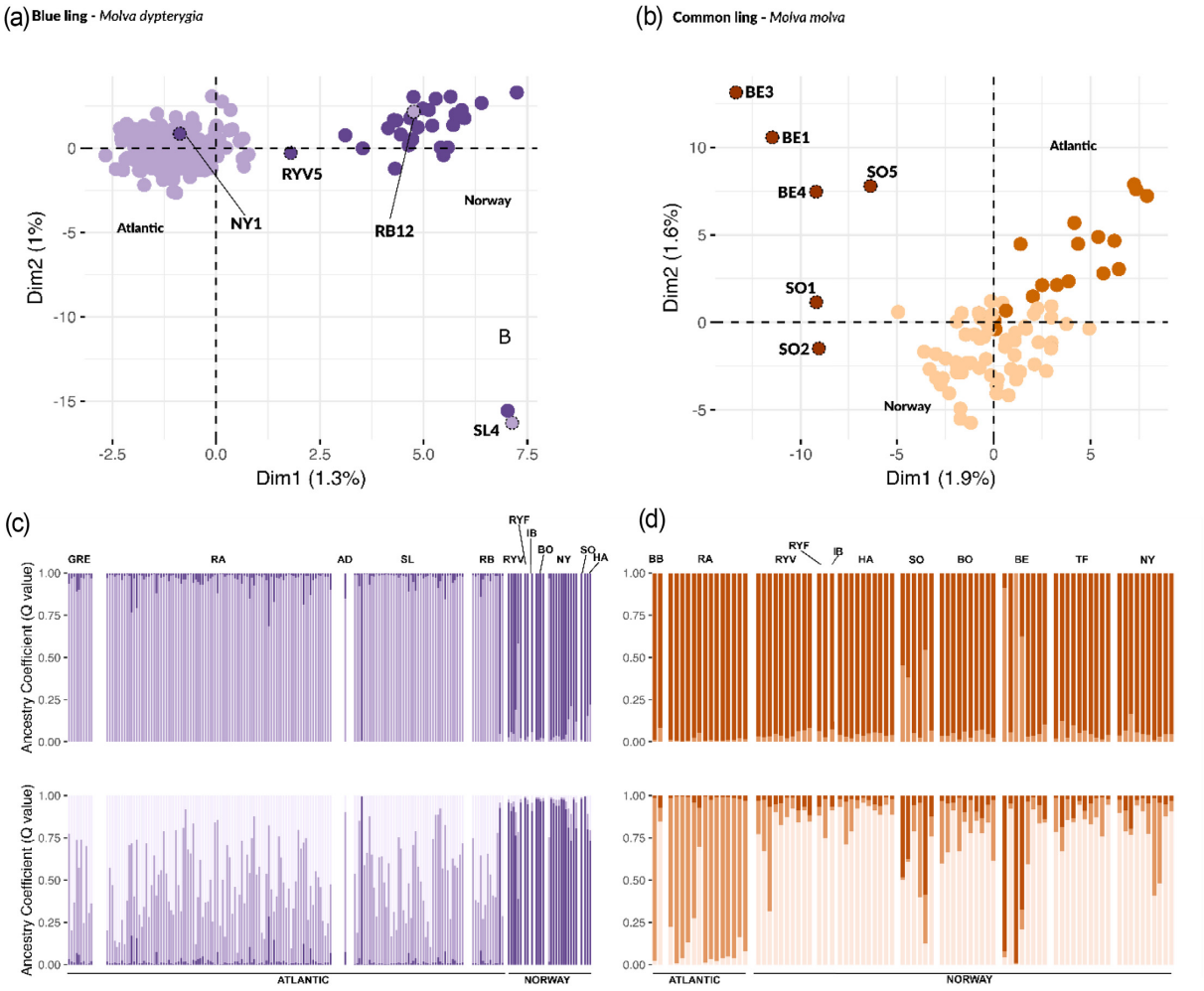

Figure 2. PCA plots and Structure plots for the blue ling (a, c) and common ling (b, d). The upper barplots show $k = 2$, and the lower ones show $k = 3$. Individuals of particular importance are singularly labelled in the PCA plots so that they can be found in the Structure plots. The colours match the number of clusters found, two for blue ling (corresponding to Atlantic and Norway) and three for common ling [Atlantic, Norwegian coast, and individuals from Bergen (BE) and Sørfjorden (SO)].

studying species living in the deep sea, where habitat homogeneity and relatively longer generation times were thought to be the major drivers for the lack of population structure observed (Cunha *et al.*, 2012; Coscia *et al.*, 2018).

A previous microsatellite study of population structure in *M. molva* found that individuals in the Atlantic Ocean were weakly divergent from their Norwegian counterpart (Blanco Gonzalez *et al.*, 2015). The authors hypothesized that larval retention played a large role in maintaining this pattern, with the Rockall area having less connectivity due to its bathymetric profile restricting adult migration. Subsequent studies showed that other species with similar life histories and spatial distributions, such as tusk *Brosme brosme* (Knutsen *et al.*, 2009) and saithe *Pollachius virens* (Saha *et al.*, 2015; Myksvoll *et al.*, 2021), have an analogous pattern of genetic differentiation.

This study strengthens the perception of a clear phylogeographic break between the Atlantic (Rockall, Anton Dohrn, and Rosemary Banks areas) and Norway, for both these species. The divergence is stronger for deeper living blue ling, than for common ling, with F_{ST} values an order of magnitude larger. The two species differ in their depth distribution and spawning behaviour: the common ling is not thought to aggregate for spawning, but the blue ling does so in several locations across the North Atlantic (Large *et al.*, 2010). Large *et al.* (2010) identified five spawning aggregations within the Atlantic area around Rockall, whereas only one aggregation is thought to be present off the coast of Norway (Helle *et al.*, 2020). The Atlantic aggregations are likely not reproductively isolated as they do not separate into genetic units. Connectivity/gene flow could occur at early life stages, with gametes/eggs or post-larvae being transported by ocean currents between spawning grounds, or as a result of adult dispersal within the Atlantic Ocean.

Important small-scale patterns of genetic structure were detected within the Norwegian fjords. Common ling sampled in Bergen and Sørfjorden presented a unique genomic make-up (Figure 2d). Sørfjorden (Osterøy) is a long and sheltered fjord, which differs markedly from the surrounding fjords, being shallower and less saline. Wind-driven currents can flow towards the land (Dam, 2017; Aksnes *et al.*, 2019). It is possible that a population survives and possibly spawns here and extends as far as Bergen, where genetically similar individuals have been found. Genetic differentiation between fjord and open-water fish populations has been detected in Atlantic cod (Johansen *et al.*, 2020; Jorde *et al.*, 2021), but we should be

mindful of the limited number of fjords and individuals sampled in the present study. It would be important to clarify this for population management.

Adult dispersal

Adult movement over large spatial scales, albeit rare, has been detected in this study. The admixture analysis did not detect any intra-specific hybrids, individuals of mixed Atlantic/Norwegian ancestry, in either species. Yet, the analysis detected blue ling migrants between Norway and the Atlantic Ocean (Figure 2c): two Norwegian individuals from the Atlantic slope (SL) and Rosemary Bank (RB), and one Atlantic individual from the northernmost Norwegian area, Nygrunnen (NY). These are adult-F_0 migrants, assigned with admixture coefficients >95% in different approaches. One of the Norwegian fish caught in the Atlantic was mature and ready to spawn (maturity was not recorded for the other individuals). No dispersing adults were identified for the common ling between the Atlantic and Norway, but a significantly smaller number of samples were screened for this species, and less genetic diversity may have hampered the ability to detect migrants, if present. Marine fish occurring at greater depths have greater dispersal ability than shallower species (Baco *et al.*, 2016), which seems to be the case for blue ling. Behaviour may play a role in limiting gene flow, by reducing the ability of adult dispersers to spawn in a different spawning ground. The simplest explanation invokes spawning time. Intraspecific variation in spawning time (isolation-by-time) has been described in several taxa (Hendry and Day, 2005), such as North Atlantic cod in response to latitudinal or temporal thermal regimes (Oomen and Hutchings, 2015). Local adaptation to environmental conditions is another possible explanation. Although not many outliers were detected in this study, we found more SNPs near genes putatively under selection in blue ling than in common ling. Deeper layers in the open Atlantic Oceans, known for greater physical homogeneity, likely differ from Norwegian fjords, which are influenced by the many rivers and streams flowing in them. In this case, individuals moving between the two environments are adapted to different conditions (as explained below) that may hinder their ability to spawn in the new habitat. A third mechanism that may play a role in maintaining genetic separation despite adult dispersal is chromosomal inversions (Barth *et al.*, 2019). For example, rearranged portions of the genomes in Atlantic cod constitute a post-zygotic barrier through reproductive incompatibilities between groups, or through reduced fitness in intraspecific hybrids (Barth *et al.*, 2019). Future work should take this into account.

Signatures of selection

The three outlying SNPs found in the common ling did not match known genes, whereas the five outliers detected in blue ling did (Supplementary Tables S3–S7). Both sets of outliers were used to assess spatial structuring with PCA and DAPC but failed to produce a meaningful pattern (Supplementary Figure S4) due to the small number of markers. This number of markers limits inferences of local adaptation but offers insights into local adaptation in deep-sea fishes. Overall, the blue ling outlier SNPs were embedded in genes coding for *opsins*, *hepcidin*, antifreeze *glycoproteins*, *HGDFL3* *glucose transporter*, and *somatotropin*-2. One of the most prominent links was with the opsin protein, matching twice in our searches (Supplementary Tables S3 and S7). This gene is commonly associated with varying light environments, affecting the sensitivity of photoreceptors (Pampoulie *et al.*, 2015; Luehrmann *et al.*, 2018). In fish species, these adaptations may be due to the attenuation of light at depth, or to water colour/turbidity variations, and can be short term (altering gene expression within the life of an individual) or long term, affecting the evolution of the species (Lin *et al.*, 2017). The Atlantic and Norwegian populations described here for the two ling species live in habitats with different light regimes. The former inhabits the open ocean where shorter wavelengths penetrate deeper compared to coastal areas, especially fjords, where considerable spring and summer turbidity reduces light penetration (Mascarenhas *et al.*, 2017).

The second hit for blue ling outliers was with Hepcidin proteins, regulating iron physiology. They have been well characterised in fish, and positive selection has been explained as a response to infections, hypoxia and excess iron (Mu *et al.*, 2018; Barroso *et al.*, 2021). These proteins are particularly important in cold-water fishes (Xu *et al.*, 2008), whose innate immune response might be reduced by low temperatures (Bly and William Clem, 1991). Hepcidins are also involved in the response to hypoxia in teleosts (Neves *et al.*, 2015). Norwegian fjords' supply of dissolved oxygen depends on the rate of water renewal, which is often limited: Available oxygen is quickly consumed by the microbial community, which creates hypoxia, or even anoxia (Aksnes *et al.*, 2019).

Another positive hit was with genes coding for antifreeze glycoproteins. These promote cold adaptation and are described in detail in northern gadoids as a crucial family of genes that prevent the growth of ice crystals in the bloodstream (Baalsrud *et al.*, 2018; Zhuang *et al.*, 2019).

Finally, strong hits were found with growth-related genes, *HGDFL3* and somatotropin-2 (*Growth Hormone*-2). For example, Arctic cod may be vulnerable to warming, which can adversely affect growth (Laurel *et al.*, 2016). Growth-related factors are also indirectly linked to osmoregulation, with a particular role in the adaptation to seawater, by promoting growth of gills (Sakamoto and McCormick, 2006). Atlantic populations of blue ling live in saline offshore habitats, whereas habitats along the Norwegian coast are influenced by freshwater from terrestrial runoff. Enhanced gill growth may also reflect adaptation to hypoxic habitats.

Conclusions

Sustainable exploitation of marine fishes needs to rely on the separation of meaningful biological units. Genomic approaches can aid this, while also providing information about evolutionary dynamics, connectivity, adaptation, and dispersal. This genomic study on two species of *Molva* with different depth profiles and life histories has identified patterns of population structure and connectivity that were not evident using traditional genetic markers (microsatellites), and these data further allowed us to investigate signatures of putative selection. It would be beneficial to implement routine genetic surveys of other marine resources to complement established assessment protocols for both the sustainable exploitation and conservation of populations and species.

Acknowledgements

The authors are grateful to all the fishermen and scientists who have collected and shared their samples: Sylvia Frantzen and Kristin Helle (IMR), Michele Salaun (IFREMER), and Rasmus Nygaard (Greenland Institute of Natural Resources).

Supplementary material

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Conflict of interest

The authors have no conflicts of interest to declare.

Funding

This project was partially funded by the European Union's FP7 research and innovation programme under the Marie Skłodowska–Curie grant agreement no. PIEF-GA-2013–625131 *DEEMdeep* awarded to I.C.

Author contributions

I.C. designed the study and secured the initial funding. A.D.McD. and S.M. partially funded further sequencing of the samples. I.C. and B.H. carried out laboratory work. I.C., L.McG., and A.D.McD. analysed the data and wrote the first draft of the manuscript. I.C., L.McG., A.D.McD., S.M., H.K., F.N., B.H., H.K., T.J., and F.A.M.V. interpreted the data and performed critical reading from the early stages of the manuscript.

Data availability

The datasets used in this study are available from Zenodo (DOI: 10.5281/zenodo.7675329).

References

Aksnes, D. L., Aure, J., Johansen, P.-O., Johnsen, G. H., and Vea Salvanes, A. G. 2019. Multi-decadal warming of Atlantic water and associated decline of dissolved oxygen in a deep fjord. Estuarine, Coastal and Shelf Science, 228: 106392.

Allendorf, F. W., Hohenlohe, P. A., and Luikart, G. 2010. Genomics and the future of conservation genetics. Nature Reviews Genetics, 11: 697–709.

Baalsrud, H. T., Tørresen, O. K., Solbakken, M. H., Salzburger, W., Hanel, R., Jakobsen, K. S., and Jentoft, S. 2018. *De novo* gene evolution of antifreeze glycoproteins in codfishes revealed by whole genome sequence data. Molecular Biology and Evolution, 35: 593–606.

Baco, A. R., Etter, R. J., Ribeiro, P. A., von der Heyden, S., Beerli, P., and Kinlan, B. P. 2016. A synthesis of genetic connectivity in deep-sea fauna and implications for marine reserve design. Molecular Ecology, 25: 3276–3298.

Barroso, C., Carvalho, P., Nunes, M., Gonçalves, J. F. M., Rodrigues, P. N. S., and Neves, J. V. 2021. The era of antimicrobial peptides: use of hepcidins to prevent or treat bacterial infections and iron disorders. Frontiers in Immunology, 12: 4018.

Barth, J. M. I., Villegas-Ríos, D., Freitas, C., Moland, E., Star, B., André, C., Knutsen, H., *et al.* 2019. Disentangling structural genomic and behavioural barriers in a sea of connectivity. Molecular Ecology, 28: 1394–1411.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., Maes, G. E., *et al.* 2017. Harnessing the power of genomics to secure the future of seafood. Trends in Ecology and Evolution, 32: 665–680.

Blanco Gonzalez, E., Knutsen, H., Jorde, P. E., Glover, K. A., and Bergstad, O. A. 2015. Genetic analyses of ling (*Molva molva*) in the Northeast Atlantic reveal patterns relevant to stock assessments and management advice. ICES Journal of Marine Science, 72: 635–641.

Bly, J. E., and William Clem, L. 1991. Temperature-mediated processes in teleost immunity: *in vitro* immunosuppression induced by *in vivo* low temperature in channel catfish. Veterinary Immunology and Immunopathology, 28: 365–377.

Cadrin, S. X. 2020. Defining spatial structure for fishery stock assessment. Fisheries Research, 221: 105397.

Clarke, J., Milligan, R. J., Bailey, D. M., and Neat, F. C. 2015. A scientific basis for regulating deep-sea fishing by depth. Current Biology, 25: 2425–2429.

Cohen, D., Inada, T., Iwamoto, T., and Scialabba, N. 1990. FAO species catalogue vol.10. Gadiform fishes of the world (order Gadiformes). An annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to date. FAO Fisheries Synopsis, 10(0014–5602): 442.

Coscia, I., Castilho, R., Massa-Gallucci, A., Sacchi, C., Cunha, R. L., Stefanni, S., Helyar, S. J., *et al.* 2018. Genetic homogeneity in the deep-sea grenadier *Macrourus berglax* across the North Atlantic Ocean. Deep Sea Research Part I, 132: 60–67.

Coscia, I., Vogiatzi, E., Kotoulas, G., Tsigenopoulos, C. S., and Mariani, S. 2012. Exploring neutral and adaptive processes in expanding populations of gilthead sea bream, *Sparus aurata* L., in the North-East Atlantic. Heredity, 108: 537–546.

Coscia, I., Wilmes, S. B., Ironside, J. E., Goward-Brown, A., O'Dea, E., Malham, S. K., McDevitt, A. D., *et al.* 2020. Fine-scale seascape genomics of an exploited marine species, the common cockle *Cerastoderma edule*, using a multimodelling approach. Evolutionary Applications, 13: 1854–1867.

Cunha, R. L., Coscia, I., Madeira, C., Mariani, S., Stefanni, S., and Castilho, R. 2012. Ancient divergence in the trans-oceanic deep-sea shark *Centroscymnus crepidater*. PLoS One, 7: e49196.

Dam, G. 2017. Simulation of spreading of fine sediment in Sørfjorden due to rock dumping (No. O17009-P3-C; p. 42). Dam Engineering. https://www.vegvesen.no/globalassets/vegprosjekter/utbygging/e16banearnastanghelle/vedlegg/fagrapporter/tidligere-rapporter/report-simulating-spreading-fine-sediment-in-sorfjorden-2017-4-mb-pdf.pdf (last accessed 18 October 2022).

Dray, S., and Dufour, A.-B. 2007. The *ade4* Package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22: 1–20.

Edwards, C. T. T., Hillary, R. M., Levontin, P., Blanchard, J. L., and Lorenzen, K. 2012. Fisheries assessment and management: a synthesis of common approaches with special reference to deep-water and data-poor stocks. Reviews in Fisheries Science, 20: 136–153.

Frichot, E., and François, O. 2015. LEA: an R package for landscape and ecological association studies. Methods in Ecology and Evolution, 6: 925–929.

Gaither, M. R., Gkafas, G. A., de Jong, M., Sarigol, F., Neat, F., Regnier, T., Moore, D., *et al.* 2018. Genomics of habitat choice and adaptive evolution in a deep-sea fish. Nature Ecology and Evolution, 2: 680–687.

Gonçalves da Silva, A., Barendse, W., Kijas, J., England, P. R., and Hoelzel, A. R. 2020. Genomic data suggest environmental drivers of fish population structure in the deep sea: a case study for the orange roughy (*Hoplostethus atlanticus*). Journal of Applied Ecology, 57: 296–306.

Gosselin, T., Anderson, E. C., and Bradbury, I. 2016. assigner: assignment analysis with GBS/RAD data using R. R Package Version 0.4, 1(10.5281). http://thierrygosselin.github.io/assigner/ (last accessed 12 December 2019).

Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. Molecular Ecology Notes, 5: 184–186.

Goudet, J., and Jombart, T. 2015. Estimation and tests of hierarchical *F*-statistics. R Core Team. https://CRAN.R-project.org/package=hierfstat (last accessed 21 August 2019).

Helle, K., Quintela, M., Taggart, J. B., Seliussen, B., Dahle, G., Glover, K. A., and Johansen, T. 2020. Development of SNP for the deep-sea fish blue ling, *Molva dypterygia* (Pennant, 1784) from ddRAD sequencing data. Conservation Genetics Resources, 12: 231–237.

Hendry, A. P., and Day, T. 2005. Population structure attributable to reproductive time: isolation by time and adaptation by time. Molecular Ecology, 14: 901–916.

Hidalgo, M., Kaplan, D. M., Kerr, L. A., Watson, J. R., Paris, C. B., and Browman, H. I. 2017. Advancing the link between ocean connectivity, ecological function and management challenges. ICES Journal of Marine Science, 74: 1702–1707.

Hilborn, R., Amoroso, R. O., Anderson, C. M., Baum, J. K., Branch, T. A., Costello, C., de Moor, C. L., *et al.* 2020. Effective fisheries management instrumental in improving fish stock status. Proceedings of the National Academy of Sciences, 117: 2218–2224.

ICES. 2020. Working Group on the Biology and Assessment of Deep-sea Fisheries Resources (WGDEEP). ICES Scientific Reports. 2:38. 928 pp. http://doi.org/10.17895/ices.pub.6015 (last accessed 13 October 2022).

Johansen, T., Besnier, F., Quintela, M., Jorde, P. E., Glover, K. A., Westgaard, J.-I., Dahle, G., *et al.* 2020. Genomic analysis reveals neutral and adaptive patterns that challenge the current management regime for East Atlantic cod *Gadus morhua* L. Evolutionary Applications, 13: 2673–2688.

Jombart, T. 2008. *adegenet*: a R package for the multivariate analysis of genetic markers. Bioinformatics, 24: 1403–1405.

Jombart, T., and Ahmed, I. 2011. *adegenet* 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics, 27: 3070–3071.

Jorde, P. E., Huserbråten, M. B., Seliussen, B. B., Myksvoll, M. S., Vikebø, F. B., Dahle, G., Aglen, A., *et al.* 2021. The making of a genetic cline: introgression of oceanic genes into coastal cod populations in the Northeast Atlantic. Canadian Journal of Fisheries and Aquatic Sciences, 78: 958–968.

Knutsen, H., Jorde, P. E., Sannæs, H., Rus Hoelzel, A., Bergstad, O. A., Stefanni, S., Johansen, T., *et al.* 2009. Bathymetric barriers promoting genetic structure in the deepwater demersal fish tusk (*Brosme brosme*). Molecular Ecology, 18: 3151–3162.

Large, P. A., Diez, G., Drewery, J., Laurans, M., Pilling, G. M., Reid, D. G., Reinert, J., *et al.* 2010. Spatial and temporal distribution of spawning aggregations of blue ling (*Molva dypterygia*) west and northwest of the British Isles. ICES Journal of Marine Science, 67: 494–501.

Laurel, B. J., Spencer, M., Iseri, P., and Copeman, L. A. 2016. Temperature-dependent growth and behavior of juvenile Arctic cod (Boreogadus saida) and co-occurring North Pacific gadids. Polar Biology, 39: 1127–1135.

Lin, J.-J., Wang, F.-Y., Li, W.-H., and Wang, T.-Y. 2017. The rises and falls of opsin genes in 59 ray-finned fish genomes and their implications for environmental adaptation. Scientific Reports, 7: 15568.

Luehrmann, M., Stieb, S. M., Carleton, K. L., Pietzker, A., Cheney, K. L., and Marshall, N. J. 2018. Short-term colour vision plasticity on the reef: changes in opsin expression under varying light conditions differ between ecologically distinct fish species. Journal of Experimental Biology, 221: jeb175281.

Luu, K., Bazin, E., and Blum, M. G. B. 2017. *pcadapt*: an R package to perform genome scans for selection based on principal component analysis. Molecular Ecology Resources, 17: 67–77.

Malmstrøm, M., Matschiner, M., Tørresen, O. K., Jakobsen, K. S., and Jentoft, S. 2017. Whole genome sequencing data and *de novo* draft assemblies for 66 teleost species. Scientific Data, 4: 160132.

Manel, S., Perrier, C., Pratlong, M., Abi-Rached, L., Paganini, J., Pontarotti, P., and Aurelle, D. 2016. Genomic resources and their influence on the detection of the signal of positive selection in genome scans. Molecular Ecology, 25: 170–184.

Maroso, F., Hillen, J. E. J., Pardo, B. G., Gkagkavouzis, K., Coscia, I., Hermida, M., Franch, R., *et al.* 2018. Performance and precision of double digestion RAD (ddRAD) genotyping in large multiplexed datasets of marine fish species. Marine Genomics, 39: 64–72.

Mascarenhas, V. J., Voß, D., Wollschlaeger, J., and Zielinski, O. 2017. Fjord light regime: bio-optical variability, absorption budget, and hyperspectral light availability in Sognefjord and Trondheimsfjord, Norway. Journal of Geophysical Research: Oceans, 122: 3828–3847.

Mu, Y., Huo, J., Guan, Y., Fan, D., Xiao, X., Wei, J., Li, Q., *et al.* 2018. An improved genome assembly for *Larimichthys crocea* reveals hepcidin gene expansion with diversified regulation and function. Communications Biology, 1: 1–12.

Myksvoll, M. S., Devine, J., Quintela, M., Geffen, A. J., Nash, R. D. M., Sandvik, A., Besnier, F., *et al.* 2021. Linking dispersal connectivity to population structure and management boundaries for saithe in the Northeast Atlantic. Marine Ecology Progress Series, 680: 177–191.

Neves, J. V., Caldas, C., Vieira, I., Ramos, M. F., and Rodrigues, P. N. S. 2015. Multiple hepcidins in a teleost fish, *Dicentrarchus labrax*: different hepcidins for different roles. The Journal of Immunology, 195: 2696–2709.

Oomen, R. A., and Hutchings, J. A. 2015. Variation in spawning time promotes genetic variability in population responses to environmental change in a marine fish. Conservation Physiology, 3: cov027.

Pampoulie, C., Skirnisdottir, S., Star, B., Jentoft, S., Jónsdóttir, I. G., Hjörleifsson, E., Thorsteinsson, V., *et al.* 2015. Rhodopsin gene polymorphism associated with divergent light environments in Atlantic Cod. Behavior Genetics, 45: 236–244.

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

Reiss, H., Hoarau, G., Dickey-Collas, M., and Wolff, W. J. 2009. Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish and Fisheries, 10: 361–395.

Rochette, N. C., Rivera-Colón, A. G., and Catchen, J. M. 2019. Stacks 2: analytical methods for paired-end sequencing improve RADseq-based population genomics. Molecular Ecology, 28: 4737–4754.

Saha, A., Hauser, L., Kent, M., Planque, B., Neat, F., Kirubakaran, T. G., Huse, I., *et al.* 2015. Seascape genetics of saithe (*Pollachius virens*) across the North Atlantic using single nucleotide polymorphisms. ICES Journal of Marine Science, 72: 2732–2741.

Sakamoto, T., and McCormick, S. D. 2006. Prolactin and growth hormone in fish osmoregulation. General and Comparative Endocrinology, 147: 24–30.

Vieira, R. P., Trueman, C. N., Readdy, L., Kenny, A., and Pinnegar, J. K. 2019. Deep-water fisheries along the British Isles continental slopes: status, ecosystem effects and future perspectives. Journal of Fish Biology, 94: 981–992.

Waples, R. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. Journal of Heredity, 89: 438–450.

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

Xu, Q., Cheng, C.-H. C., Hu, P., Ye, H., Chen, Z., Cao, L., Chen, L., *et al.* 2008. Adaptive evolution of hepcidin genes in antarctic notothenioid fishes. Molecular Biology and Evolution, 25: 1099–1112.

Zhuang, X., Yang, C., Murphy, K. R., and Cheng, C.-H. C. 2019. Molecular mechanism and history of non-sense to sense evolution of antifreeze glycoprotein gene in northern gadids. Proceedings of the National Academy of Sciences, 116: 4400–4405.

Handling Editor: W. Stewart Grant