Quantifying genetic differentiation and population assignment between two contingents of Atlantic mackerel (Scomber scombrus) in the Northwest Atlantic

Bourret, Audrey; Smith, Andrew; Van Beveren, Elisabeth; Plourde, Stephane; Curti, Kiersten L.; Jansen, Teunis; Richardson, David E.; Castonguay, Martin; Rodriguez-Ezpeleta, Naiara; Parent, Genevieve J.

Published in:
Canadian Journal of Fisheries and Aquatic Sciences

Link to article, DOI:
10.1139/cjfas-2022-0232

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Canadian Journal of Fisheries and Aquatic Sciences

OPEN ACCESS | Article

Quantifying genetic differentiation and population assignment between two contingents of Atlantic mackerel (*Scomber scombrus*) in the Northwest Atlantic

Audrey Bourret[a], **Andrew Smith**[b], **Elisabeth Van Beveren**[b], **Stéphane Plourde**[b], **Kiersten L. Curti**[c], **Teunis Jansen**[d,e], **David E. Richardson**[f], **Martin Castonguay**[b], **Naiara Rodriguez-Ezpeleta**[g], **and Geneviève J. Parent**[a]

[a]Laboratory of Genomics, Maurice Lamontagne Institute, Fisheries and Oceans Canada, 850 Route de la Mer, Mont-Joli, QC G5H 3Z4, Canada; [b]Maurice Lamontagne Institute, Fisheries and Oceans Canada, 850 Route de la Mer, Mont-Joli, QC G5H 3Z4, Canada; [c]Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02556, USA; [d]Greenland Institute of Natural Resources (GINR), Kivioq 2, PO Box 570, Nuuk 3900, Greenland; [e]DTU AQUA (National Institute of Aquatic Resources), Kemitorvet, 2800 Kgs., Lyngby, Denmark; [f]Northeast Fisheries Science Center, National Marine Fisheries Service, 28 Tarzwell Drive, Narragansett, RI 02879, USA; [g]AZTI, Marine Research, Basque Research and Technology Alliance (BRTA), Txatxarramendi ugartea z/g, Sukarrieta - Bizkaia, 48395, Spain

Corresponding authors: **Audrey Bourret** (email: audrey.bourret@dfo-mpo.gc.ca) and **Geneviève Parent** (email: genevieve.parent@dfo-mpo.gc.ca)

Abstract

In the Northwest Atlantic (NWA), the Atlantic mackerel (*Scomber scombrus*) has a northern and a southern contingent, which spawn in Canada and the United States (U.S.), respectively. Both contingents mix mostly along the U.S. shelf during overwintering. The discrimination of individuals from each contingent in fisheries could improve the management of this depleted species in both countries. Here, we used genome-wide markers (>10 000 single nucleotide polymorphisms (SNPs)) to assess genomic differences between mackerel of both contingents, and possibly infer the proportions of each contingent in NWA management units. Small but significant genetic differentiation was observed between the northern and southern contingents (F_{ST} = 0.0010). Genetic assignments to reference samples from the two contingents were performed with predictive accuracy > 85%. Fish from both contingents were present along the U.S. shelf during late winter and early spring but also, without prior evidence of this, likely in Canadian waters from late spring to fall. Genetic assignments could be used as a stock discrimination tool so that fishery removals can be effectively determined and managed on a contingent level.

Key words: population genomics, fisheries, *Scombridae*, migratory, pelagic, fish

Introduction

Defining the population structure of marine migratory fish species is challenging. Long-distance migrants may cross multiple management jurisdictions each year and have multiple spawning areas (e.g., Puncher et al. 2018; Rodríguez-Ezpeleta et al. 2019; Gíslason et al. 2020). Within such species-specific aggregations, the discrimination of biological units is essential for their sustainable management (Ovenden et al. 2015; Cadrin 2020). Many techniques have been used to advance this goal, including the comparisons of life-history traits, morphometrics, meristics, parasite loads, fatty acid signatures, otolith chemical compositions, and genetics (Begg and Waldman 1999; Cadrin et al. 2014). Recently, genomic analyses have become more accessible and have been successful in resolving mismatches between management and biological units (Mullins et al. 2018). For instance, genomics has been a useful tool for assessing the population structure of several species with long-distance migratory patterns such as Atlantic bluefin tuna (*Thunnus thynnus*, Puncher et al. 2018), yellowfin tuna (*Thunnus albacares*, Mullins et al. 2018; Pecoraro et al. 2018), albacore (*Thunnus alalunga*, Vaux et al. 2021), and Atlantic cod (*Gadus morhua*, Berg et al. 2017).

The Atlantic mackerel (*Scomber scombrus*) is another long-distance migratory species (seasonal migrations of over 2000 km; Parsons and Moores 1973; Moores et al. 1975), with genetically distinct populations on each side of its transatlantic distribution. Specifically, mackerel from the Northwest Atlantic (NWA) and Northeast Atlantic (NEA) are generally considered to be discrete populations (Nesbø et al. 2000; Rodríguez-Ezpeleta et al. 2016; Gíslason et al. 2020). In the NEA, Atlantic mackerel has, in recent years, expanded northwestwards to Iceland and Greenland waters from their traditional summer feeding habitat in Europe (Jansen et al. 2016; Gíslason et al. 2020). There is no evidence of migration between both sides of the Atlantic (Rodríguez-Ezpeleta et al. 2016; Gíslason et al. 2020).

Within the NWA, two contingents (spawning components) have long been recognized (Sette 1943, 1950)— a northern contingent that spawns predominantly in the southern Gulf of St. Lawrence and a southern contingent that spawns mostly in southern New England and the western Gulf of Maine. Atlantic mackerel are indeterminate batch spawners (Morse 1980; Jansen et al. 2021) and spawning site fidelity to the two main spawning grounds has been suggested in the NWA (Studholme et al. 1999), with recent evidence coming from otolith stable isotope applications (Redding et al. 2020; Arai et al. 2021). Both contingents overwinter along the edge of the continental shelf from the Scotian Shelf (Canada) to the Mid-Atlantic Bight (United States (U.S.), Fig. 1). They mix to an unknown degree, making the efficacy of management measures for this overfished species difficult to evaluate. They also differ in their seasonal migration routes based on observations through tagging studies and seasonal patterns in landings (Sette 1943, 1950; Parsons and Moores 1973; Beckett et al. 1974; Moores et al. 1975; Stobo 1976). In the spring, the southern contingent moves inshore and, in recent years, spawns in May (average peak day of spawning: 19 May; Sette 1943; Berrien 1978; McManus et al. 2018). The northern contingent also migrates inshore from the continental shelf, but most fish veer northwestwards into the southern Gulf of St. Lawrence, where mature adults spawn in June and July (average peak day of spawning: 21 June; Grégoire et al. 2010; Brosset et al. 2020).

Despite distinct spawning regions and seasonal migration circuits of both NWA contingents, most phenotypic markers (e.g., meristics, MacKay and Garside 1969; otolith shape, Castonguay et al. 1991) have had low discrimination success. Previous genetic studies relying on allozyme (Maguire et al. 1987), mitochondrial deoxyribonucleic acid (mtDNA) control region (Lambrey de Souza et al. 2006), and microsatellite (Gíslason et al. 2020) markers have likewise been unsuccessful in distinguishing between the northern and southern contingents. Recently, otolith stable isotope compositions ($\delta^{18}O/\delta^{13}C$ values) have been applied with success, achieving moderate levels of discrimination (75%–92%; Arai et al. 2021). The authors of this study showed complete discrimination between NEA and NWA populations using otolith stable isotope compositions and were able to estimate, within the NWA, the relative interannual variations in the proportions of northern and southern fish of different ages sampled from the U.S. winter fishery. However, this method needs baseline samples every year for calibration, and relies on field laboratory procedures dedicated to otolith processing and stable isotope analyses ($\geq$1 h otolith^{-1}) which require substantial human and material resources (Arai et al. 2021). We hypothesized that genome-wide markers could provide tools that increase our capacity to discriminate individuals from the two NWA contingents, without the need for baseline samples every year. Further, inferences on population structure using genomics will carry greater weight in stock assessments as a proxy of population rather than natal homing, as for otolith stable isotope composition.

Mackerel stock assessment and management are performed independently by the U.S. and Canada, although scientists from both countries collaborate on stock assessments. The whole NWA stock is considered overfished with overfishing continuing to occur (NEFSC 2018). Since the 2010s, the spawning stock biomass has declined below the limit reference points, with generally low recruitment (NEFSC 2018; DFO 2021). The northern contingent alone is considered in the Critical Zone of Canada's Precautionary Approach framework (DFO 2021). The difficulty in separating both contingents has led the U.S. to assess the two contingents as a single NWA stock (NEFSC 2018). This should lead to suboptimal management as fishing pressure might not be applied proportionally to contingent size and hence one contingent might risk higher overexploitation (Cope and Punt 2011; Ying et al. 2011). In contrast, Canada assesses only the northern contingent (DFO 2021) but explicitly acknowledges that the U.S. fleet removes a highly uncertain proportion of this contingent. The large uncertainty in total contingent-specific fishery removals has a substantial impact on the efficiency of management actions (Van Beveren et al. 2020*a*, *b*). A stock discrimination method that can be scaled up to resolve the relative importance of each contingent in Canadian and U.S. fisheries with reasonable precision would improve management of mackerel in both countries (e.g., Van Beveren et al. 2019), which is perceived as urgent for a stock that requires rebuilding.

Here, we aimed to assess the genetic distinctiveness between the southern and northern contingents with a large sample size and genome-wide markers. We first assessed the population structure the Atlantic mackerel, then we tested genetic assignments as a stock discrimination tool for improving assessment and management of the species in the NWA.

Methods

Sampling

Length-stratified samples of Atlantic mackerel from the NWA were collected from scientific surveys and the commercial fishery, between March and November 2018 and 2019 (see commercial sampling protocol, Smith et al. 2020; Fig. 1, Table 1). The sampling covered U.S. coastal zones in early spring (March–April, $n = 278$), and roughly reflected patterns of migration across Canadian waters from late spring to fall (May–October, $n = 399$), with fish entering the southern Gulf St. Lawrence (4T) in late spring–early summer to spawn, then moving into the northern Gulf St. Lawrence (4S) and around Newfoundland (4R, 3K) later in the season for feeding purposes (Van Beveren et al. 2022). Individuals were frozen at -20 °C or colder and were shipped on ice to the Maurice Lamontagne Institute. Defrosted individuals were measured (fork length, rounded to the nearest millimeter) and weighed, sexed, staged for maturity, and aged using otoliths (Table 1). Atlantic mackerel from the NEA were sampled during scientific surveys ($n = 81$, Table 1; see Nøttestad et al. 2016*b* and Rodríguez-Ezpeleta et al. 2016 for more details). For all individuals, a piece of muscle or fin tissue was excised and preserved in ethanol for later deoxyribonucleic acid (DNA) extraction.

Fig. 1. Sampling location and month of reference (star) and nonreference samples (circle) in NAFO division (North Atlantic Fisheries Organization) of the Northwest Atlantic. Bathymetry data were imported from GEBCO and the NAFO divisions shapefile was taken from NAFO (https://www.nafo.int/Data/GIS). Map projection is NAD83 (CSRS)/Quebec Lambert (EPSG:6622).

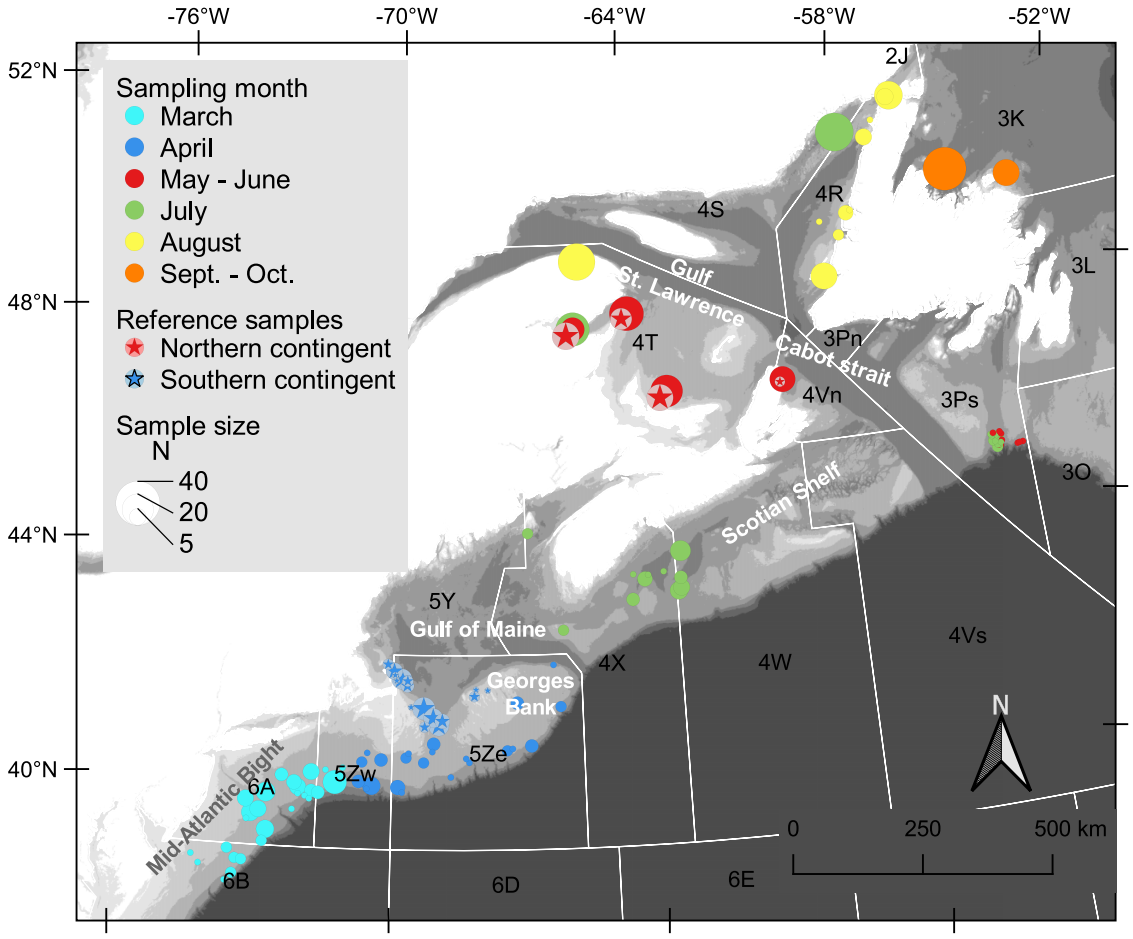


Table 1. Properties (age, sex) and number (N) of the Atlantic mackerel samples by country, NAFO division (North Atlantic Fisheries Organization), source (scientific survey versus commercial sampling program), year, and month.

Country	NAFO	Survey	Year	Month	Mean age	Sex ratio	N_G	N_A	N_{REF}
Canada	4W	Scientific	2018	July	3.1 (2–4)	12:13:0	25	22	0
	4X	Scientific	2018	July	3.0 (3–3)	8:6:0	14	14	0
	3Ps	Scientific	2019	May–July	3.2 (2–5)	22:11:0	33	26	0
	4Vn	Scientific	2019	June	2.6 (2–4)	9:8:0	17	15	2
	4T	Commercial	2018	June–August	3.8 (3–7)	61:45:0	106	104	31
	4T	Commercial	2019	June–July	4.3 (2–8)	32:27:0	59	54	20
	4S	Commercial	2018	July	4.1 (1–7)	19:11:0	30	30	0
	4R	Scientific	2019	August	3.6 (2–5)	29:31:0	60	49	0
	3K	Commercial	2018	September–October	3.3 (3–5)	28:27:0	55	52	0
U.S.	6A	Scientific	2019	March	2.6 (1–5)	43:27:7	77	56	0
	6B	Scientific	2019	March	2.9 (1–5)	9:10:0	19	11	0
	5Zw	Scientific	2019	March–April	3.1 (1–6)	24:19:1	44	32	0
	5Ze	Scientific	2019	April	2.0 (1–5)	72:44:16	132	90	61
	5Y	Scientific	2019	April	2.5 (1–4)	1:5:0	6	2	2
Spain	-	Scientific	2013	NA	NA	0:0:38	38	36	-
Greenland	-	Scientific	2019	July–August	NA	0:0:43	43	36	-

Note: The mean age is in years (range in parentheses) and the sex ratio represented female:male:unknown. The number of individuals genotyped (N_G) and analyzed following quality control (N_A), and the number of individuals included as reference (N_{REF}) are presented. Abbreviation: NA, not available.

SNP detection

DNA was extracted from the preserved tissues using the DNeasy blood and tissues kit and an RNAse treatment (QIAGEN, Germantown, MD, U.S.). DNA quality was visually checked on 2% agarose gels and DNA concentration was assessed with a Nanodrop (ND-1000, Thermo Fisher Scientific, Waltham, MA, U.S.). Only extracts with high-molecular weight DNA and DNA concentrations between 5 and 20 ng·mL^{-1} were used for library preparation (N individuals genotyped = 758; Table 1).

Double digest restriction site-associated DNA sequencing (ddRADseq) libraries were prepared at the Plateforme d'analyse génomique (IBIS, Université Laval) using 20 ng of DNA and *PstI* and *MspI* restriction enzymes, following Poland et al. (2012). ddRADseq is a robust and cost-effective method to obtain a reduced representation of the genome for non-model organisms (Peterson et al. 2012). Two sequencers were used in this project, namely HiSeq and NovaSeq. For the HiSeq sequencer, samples were indexed with a unique sequence on a maximum of 96 extracts per plate. Each plate of 96 indexed DNA extracts was then pooled in a single library which was sequenced in a single lane of Illumina HiSeq 4000 PE100 (four single libraries in four lanes) at Genome Québec. For the NovaSeq sequencer, samples were first indexed as described above, then a second index was added to each plate (seven plates here, including one previously sequenced on HiSeq) and all samples were pooled as one library and sequenced in a single flow cell of Illumina NovaSeq 6000 S4 PE100 at Genome Quebec, with 10% PhiX. A total of 96 extracts were sequenced with both sequencing technologies to ensure reproducibility.

Read quality was checked with FastQC (v0.11.8, Andrews 2010) and MultiQC (v1.7, Ewels et al. 2016). Adapters were removed with Trimmomatic (v0.39, Bolger et al. 2014). Reads were then processed with Stacks (v2.4, Catchen et al. 2013; Rochette et al. 2019). Demultiplexing and filtering were performed with the process_radtags module, with a truncation at 85 bp. Paired-end reads were assembled de novo as no suitable reference genomes were available, and single nucleotide polymorphisms (SNPs) were called using the *ustacks*, *cstacks*, *sstacks*, *tsv2bam*, and *gstacks* modules. Parameters *m*, *M*, and *n* were chosen following the recommendation of Rochette and Catchen (2017). In short, we tested different parameter values on a subset of 22 representative individuals (minimum depth coverage (m) = 3–5, number of mismatches allowed between putative alleles (M) = 1–8, and number of mismatches allowed between putative loci during the construction of the catalogue (n) = $M - 1$ to $M + 1$), using *denovo_map.pl* pipeline. Parameters $m = 3$, $M = 4$, and $n = 5$ showed the highest number of loci shared by 80% of samples (r80) and were thus retained. The catalogue was assembled with cstacks on a subset of 210 representative individuals showing a mean coverage between 15× and 40×. Only samples with a mean coverage above 10× were kept in the analysis (N = 630).

Two panels of SNPs were created with the populations module of Stacks, one with individuals from both sides of the Atlantic (NEA–NWA panel) and one with only West Atlantic individuals (NWA panel). For each panel, only SNPs with minor allele frequencies higher than 0.10, missing data lower than 10%, and expected heterozygosity (*He*) lower than 0.5 were retained. One individual with more than 30% missing data was removed. He and observed heterozygosity (*Ho*) were computed for each library independently to check for potential batch effects, including potential effects of the two sequencing strategies. Outlier SNPs showing both $He > 0.45$ and $Ho < 0.05$ were removed. Relatedness coefficients were computed with VCFtools (Danecek et al. 2011), using the algorithm of Manichaikul et al. (2010). Only the first SNP per locus was retained to reduce inclusion of SNPs in linkage disequilibrium. Specific data formats were obtained with VCFtools, plink (Purcell et al. 2007), VCFr (Knaus and Grünwald 2017), dartR (Gruber et al. 2018), and genepopedit (Stanley et al. 2017). An analysis of molecular variance (AMOVA) was performed on the subset of samples sequenced with both HiSeq and NovaSeq using the poppr R package (v2.9.2, ade4 method, NEA–NWA SNP panel; Kamvar et al. 2014) to assess the proportion of variance explained by the sequencers.

Population structure analyses

Principal component analysis (PCA) was performed on both SNP panels using the adegenet R package (Jombart 2008), with alleles renamed 0, 1, and 2 (representing the count of the minor allele), centered, and missing values at a given SNP replaced by the mean genotype across all individuals. The impact of missing values on PCA inferences was explored by plotting per sample missingness using a color gradient (Yi and Latch 2022). We also performed a discriminant analysis of principal components (DAPC) with adegenet. Groups were formed a priori (NWA–NEA: by country, NWA: by NAFO division, North Atlantic Fisheries Organization) and optimal numbers of PC axes to retain were determined with the *optim.a.score* function to avoid overfitting (Miller et al. 2020). Bayesian clustering analyses were performed with Structure 2.3.4 (Pritchard et al. 2000) running the admixture model, with a burn-in step of 100 000 and 200 000 iterations, testing K = 1–10 for the NEA–NWA panel, and K = 1–8 for the NWA panel, both with 10 runs per K value. To reduce run time, we used the R package ParallelStructure (Besnier and Glover 2013) to distribute runs on multiple cores, and we reduced the SNP panels to keep only SNPs with less than 5% missing values (NEA–NWA = 4140 SNPs, NWA = 3930 SNPs). We used Ln P($X|K$) (Pritchard et al. 2000) and ΔK (Evanno et al. 2005) as criteria to infer the most likely number of clusters (K), and CLUMPP (Jakobsson and Rosenberg 2007) to aggregate multiple Structure runs. F_{ST} estimates were computed between both sides of the Atlantic (NWA–NEA panel) and among NAFO divisions (NWA panel) with dartR, following the equation of Weir and Cockerham (1984) and with 95% confidence intervals (CIs) calculated via bootstrapping (n = 999).

We identified SNPs potentially under selection with PCAdapt (v4.3.3, Privé et al. 2020) and Bayescan (v2.1, Foll and Gaggiotti 2008) to characterize qualitatively the contribution of putative adaptive SNPs to the observed population structure. PCAdapt is an individual-based approach which identifies SNPs significantly associated with genetic structure from important principal components (PCs) using Mahalanobis

distance. We chose the number of PCs following a visual observation of the scree plot, as recommended ($K = 1$ for the NEA–NWA panel, $K = 2$ for the NWA panel; Cattell 1966). Bayescan is a Bayesian population-based approach which relies on the distribution of F_{ST} estimates. With the NEA–NWA panel, individuals were grouped by ocean side (NEA vs. NWA), while for the NWA panel, individuals were grouped by country (Canada vs. U.S.) or by NAFO division (eight groups). SNPs with a q value < 0.05 were identified as outliers.

Samples from the different NAFO divisions within the NWA are putatively admixed due to the highly migratory nature of the species. We attempted to characterize the genetic structure between the contingents based on what we will refer to as reference samples characterizing best the northern and southern contingents. The most representative fish of each contingent would be ripe individuals caught in their respective spawning areas. For the northern contingent, we selected 53 ripe individuals across all available samples (maturity stage 5 following Maguire 1981) from the southern Gulf of St. Lawrence caught in June (NAFO 4TVn, mean age = 4.8 years, age range = 3–8 years; Table 1). For the southern contingent, ripe individuals could not be sampled because no scientific survey or fishery occurs during the spawning season (i.e., May). Instead, we selected all 63 coastal U.S. samples caught in April (the latest samples available) as the reference for the southern contingent (NAFO 5YZe, mean age = 1.7 years, age range = 1–4 years; Table 1). The latter samples likely represent the southern contingent migrating to the coast for spawning and should thus be the least admixed with the northern contingent (Fig. 1). Those samples are, hereafter, named as the reference samples for the northern and southern contingents.

Assignment validation and composition estimations

We used cross-validation tests to assess the discriminatory capacity of the reference samples. Monte Carlo cross-validation tests were performed with the assignPOP R package (Chen et al. 2018), using the *assign.MC* function with 70% of individuals from the reference samples randomly selected per training set (train.ind parameters), and 30 iterations. Five different machine-learning classification algorithms (linear discriminant function; support vector machine, SVM; decision tree; random forest; naïve Bayes) were tested using subsets of SNPs showing higher F_{ST} between training set groups (upper 10%, 25%, 50%, or all SNPs). The SVM algorithm using all loci was the best predictive model based on the mean and variance of assignment accuracy of the testing sets (Fig. S1).

To further test the validity of the reference samples with the assignment tool, we randomly reassigned the 116 individuals from reference samples, without replacement, to create 100 random data sets of two reference groups. We performed on each data set a cross-validation test with the best predictive model parameters (SVM algorithm, 30 iterations, all SNPs) to determine the mean and variance of assignment accuracy of reference groups using random reference samples. We also simulated admixed individuals from the reference samples to determine the capacity of our assignment tool to identify any bias in assignment of admixed individuals. We used the hybriddetective R package (Wringe et al. 2017) to simulate pure individuals, first (F1) and second (F2) generation hybrids, and backcrossed (i.e., pure × F1) individuals from the reference populations, with allele sampling (function *freqbasedsim_AlleleSample*, three simulations of three repetitions). Assignment of simulated individuals to reference samples was performed using the *assign.X* function from assignPOP and the SVM algorithm. For each simulated individual, a membership probability to each reference sample was first estimated. The individual was then assigned to the reference sample with the highest membership probability (> 50%). Note that membership probability to the southern contingent = 1 − membership probability to the northern contingent.

We performed genetic assignments of all adults from the NWA (excluding reference samples) with the *assign.X* function from assignPOP and the SVM algorithm. A membership probability to each reference sample was computed for each individual. We estimated the effect of missing data on assignment results by rerunning cross-validation tests and genetic assignments with data sets limiting missing data per individual or increasing randomly the proportion of missing data per individual. We limited missing data per individuals by removing individuals that were missing more than 10%, 15%, and 20% data, compared to the initial 30%. We also simulated 300 data sets with increasing thresholds of missing data per individual, i.e., 100 data sets with 10%, 20%, or 30%. These simulated data sets were created by the imputation of missing values to reach the target thresholds and the removal of individuals with observed missing data above the threshold prior to the imputation. We also tested the effect of the sequencing technology by running a cross-validation test and genetic assignments with a reduced data set including only samples sequenced using NovaSeq (n = 396 individuals, including 36 and 63 northern and southern references, respectively).

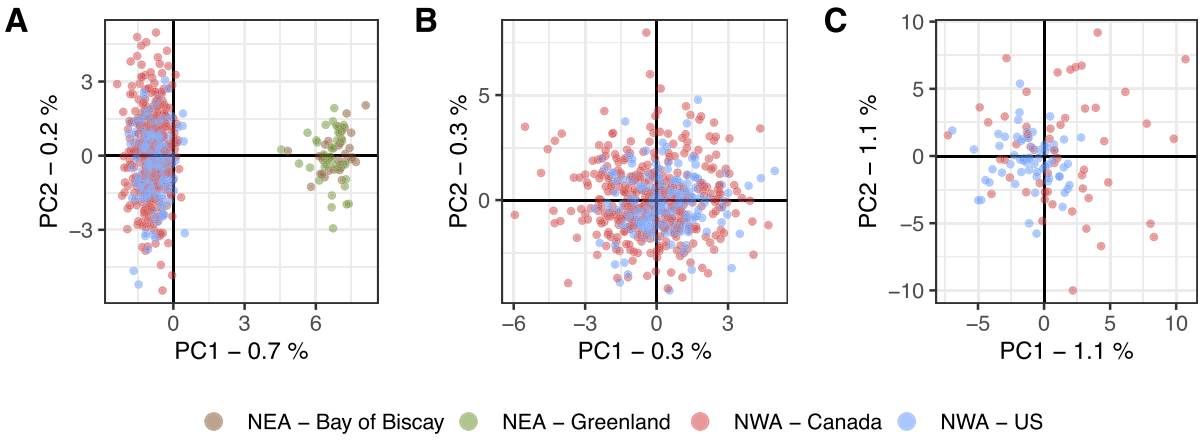
Finally, we estimated the composition of northern and southern contingent individuals by NAFO division using the genetic assignment results of all adults from the NWA (excluding reference samples). The membership probability to each reference sample computed for each individual was used to assign individuals to the reference sample based on > 50% or 70% membership probability threshold.

Results

SNP panels

A total of 2864 million reads were obtained from the 11 libraries in HiSeq (4 libraries, mean 3.6 million reads per sample) and NovaSeq (7 libraries, mean 2.2 million reads per sample). Following the demultiplexing step, 128 individuals with a mean coverage < 10× were excluded from further analyses (Table 1, N_A). The Stack pipeline generated a catalogue of 1.2 million loci, and the effective per sample coverage for remaining samples of 20.5 ± 12.3×. SNPs were filtered for diversity and over several steps of quality filtering (Table S1). The batch-effect filtration step removed up to 65 SNPs that were likely due to difference between sequencing

Fig. 2. Principal component analyses (PCAs) for (A) all samples (NEA, Northeast Atlantic; NWA, Northwest Atlantic), (B) all NWA samples, and (C) NWA contingent-specific subset of reference samples only. PCAs were performed using the full panel including both neutral and outlier SNPs. Each dot represents one sample and the color indicates the sampling country. The percentage of genetic variance explained by each axis is shown.



technologies. Following their removal, no significant proportion of the genetic variance could be explained by the differences between sequencers used (AMOVA: $P > 0.999$, Table S2, also see Fig. S2). We also excluded one sample showing a relatedness coefficient value larger than 0.25 with four other samples (e.g., potential siblings). We created two distinct SNP panels to test for population structure of Atlantic mackerel (1) across the North Atlantic Ocean (NEA–NWA; $N_{individuals} = 629$, $N_{SNPs} = 10\,832$) and (2) within the NWA (NWA; $N_{individuals} = 557$, $N_{SNPs} = 10\,703$; Table S2). The overall percentage of missing data was 5.95% and 6.03% for the NEA–NWA and NWA panels, respectively. The highest level of missing data was associated with U.S. samples (NEA–NWA panel, U.S. = 10.0 ± 5.6%, non-U.S. = 4.2% ± 4.0%; NWA panel, U.S. = 9.9% ± 5.5%, non-U.S. = 4.0% ± 3.9%; Fig. S3).

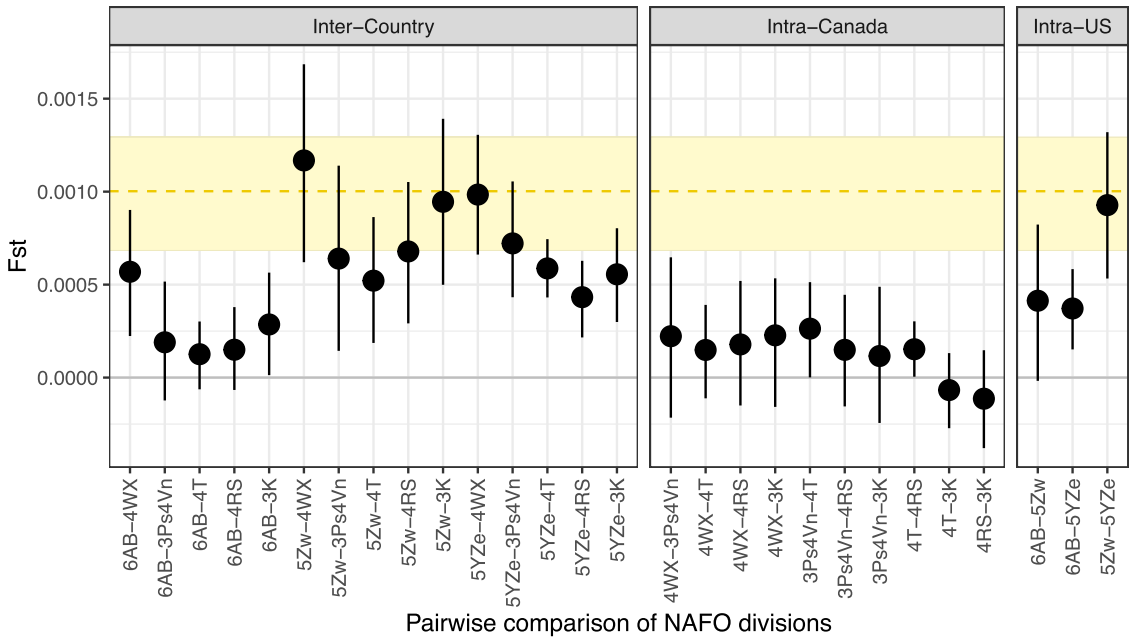
Transatlantic and NWA population structure

NWA and NEA populations were genetically distinct using the first SNP panel. The PCA showed two nonoverlapping clusters differentiating NWA and NEA samples (Fig. 2A). The DAPC showed similar results on the first discrinimant function (Fig. S4A). With the Bayesian clustering analysis, the best supported number of clusters (K value) was two and these two clusters matched those of the PCA (Fig. S4A). In the PCA, DAPC, and Bayesian clustering analyses, the Greenland samples grouped with the Bay of Biscay samples (Figs. 2A, S4A, and S5A). The F_{ST} between the NWA and the NEA populations was estimated as 0.0131 (CI = 0.0125–0.0138). PCAdapt and Bayescan identified 114 and 179 SNPs as outliers, respectively, with 69 that were common to both the methods. PCA without the 224 SNPs identified as outliers lead to similar results (Fig. S6A). Note that individuals with higher levels of missing values were located at the center of the NWA cluster and at the left edge of the NEA cluster (Fig. S7A).

Canadian and U.S. samples were genetically different based on F_{ST} and DAPC results only. The PCA showed overlap between Canadian and U.S. samples (Fig. 2B), while the first discriminant function of the DAPC showed incomplete overlap between Canadian and U.S. NAFO divisions, mostly for the 3K-4T-4WX Canadian samples (Fig. S4B). For the Bayesian clustering analyses, the best supported number of cluster in the NWA differed between the two criteria used, i.e., $K = 1$ from Pritchard Ln P($X|K$) and $K = 2$ inferred from Evanno ΔK (Fig. S5B). However, no relevant biological pattern was detected with $K = 2$ (Fig. S5B). F_{ST} between Canadian and U.S. NAFO divisions ranged between 0.0001 and 0.0012 (Fig. 3). Most intercountry F_{ST} had their 95% CIs excluding the null value, except three F_{ST} comprising NAFO division 6AB. In Canadian samples, the F_{ST} estimates between NAFO divisions were $\leq$0.0003, with only one division's 95% CI excluding the null value (4T-4RS; Fig. 3). Larger F_{ST} values were observed between U.S. NAFO divisions, with two out of three 95% CI excluding the null value (F_{ST} range: 0.0004–0.0009; Fig. 3). Three outliers SNPs were identified with Bayescan (1 SNP when grouping by country and 3 SNPs when grouping by NAFO division), and 72 different SNPs with PCAdapt. PCA without these 75 outlier SNPs did not differ meaningfully from the analyses presented (Fig. S6B). Individuals in the PCA with more missing data were located at the center of the cluster (Fig. S7B).

We selected the most representative individuals of the northern and southern contingents as reference samples (see methods). The PCA including only these individuals showed that the southern contingent cluster overlapped only partially with the northern contingent cluster, which was more widespread (Fig. 2C). A similar pattern was observed in a PCA conducted on SNPs data set excluding the 75 SNPs identified as outliers (Fig. S6C). F_{ST} between these reference samples was 0.0010 (CI = 0.0007–0.0013), an order of magnitude smaller than the transatlantic genetic differentiation. This F_{ST} value was expectedly larger than most intercountry comparisons (Fig. 3), which might include a mixture of fish from both contingents, therefore supporting our choice of reference individuals.

Fig. 3. Genetic differentiation (F_{ST} with 95% confidence interval, CI) between NAFO divisions (North Atlantic Fisheries Organization) across countries (intercountry), within Canada (intra-Canada), and within the U.S. (intra-US). The orange dashed line and yellow area represents the F_{ST} value and 95% CI between the reference samples from each contingent, respectively.



Assignment validation and composition estimation

We assessed the power of the reference samples from the northern and southern contingents for genetic assignments using cross-validation tests and simulations of admixed individuals (Fig. 4). The assignment accuracy (i.e., the mean proportion of correct classifications of individuals within testing sets) was greater than 85% for both the reference samples from the northern (mean ± SD = 85.8% ± 7.1%) and the southern contingent (mean ± SD = 88.2% ± 7.9%) (Fig. 4A). As a baseline comparison, the mean assignment accuracy of the 100 random data sets was 50%, as expected (mean ± SD = 50.0% ± 3.0%, range = 45.5%–62.4%). The genetic assignment of simulated individuals to the northern or southern contingent varied between genotypes (Fig. 4B). Pure individuals were all classified to the appropriate reference contingent, with a membership probability over 99.6%. All backcrossed individuals with the southern contingent were classified consistently with their genetic composition, with a mean membership probability of 99.98% ± 0.01%, higher than the expected probability of 75% (Fig. 4B). Most of backcrossed individuals with the northern contingent (98.6%) were classified consistently, with a mean membership probability of 79.0% ± 10.5%, which was close to predictions (Fig. 4B). All F1 and F2 hybrids were classified as southern individuals, with membership probability > 89.9%, a value higher than expected under randomness (i.e., 50%; Fig. 4B).

The U.S. samples on an average had a greater proportion of missing data, as highlighted earlier, and we observed that membership probability and proportion of missing data were negatively correlated (Spearman's rank correlation: $\rho = -0.86$, $P < 0.001$; Fig. S8). Thus, we evaluated the possible effect of missing data on assignment results by selecting only individuals of the northern and southern contingents with a lower proportion of missing data, or by inserting missing data randomly in all individuals. In data sets filtered for missing data per individual (maximum of 10%, 15%, or 20%), the overall assignment accuracy was highly similar to the one observed using the original full data set (Fig. S9A). The same negative correlations between membership probability and missing data were observed at the three thresholds maximum missing data threshold (Fig. S9B). Membership probabilities between the original and subset data sets were highly correlated (all $\rho > 0.96$; Fig. S9C). Increasing artificially the proportion of missing data per individual to 10%, 20%, or 30% through simulations showed that the relationship between original and simulated membership probability persisted (Fig. S10). More stringent filtering parameter or increased proportion of missing data per individual lead to similar results, showing that the observed assignment patterns were not mainly caused by the unbalanced missing data.

The composition of the northern versus the southern contingent estimated through genetic assignments varied between countries and NAFO divisions (Fig. 5). The proportion of northern contingent fish in all spring samples from U.S. waters was 33.6% or 25.6% with > 50% or > 70% membership probability thresholds, respectively (Fig. 5A). The proportion of the northern contingent varied among U.S. NAFO divisions between 24.1% in NAFO 5YZe and 37.3% in NAFO 6AB with > 50% membership probability or between 20.8% and 30.2% with > 70% membership probability thresholds (Figs. 5A and 5B). In Canadian waters, the proportion of the southern contingent in samples from late spring to autumn was 18.2% or 12.3% with > 50% or > 70% membership probability

Fig. 4. Genetic assignment validation using (A) cross-validation using northern and southern contingent reference samples and the best predictive model (support vector machine, SVM) and (B) membership probability to the northern contingent of simulated genotypes ("pure" northern and southern contingent, backcrossed with northern or southern contingent, F1 and F2 hybrids). See Fig. S1 for complete cross-validation results and methods for more details.

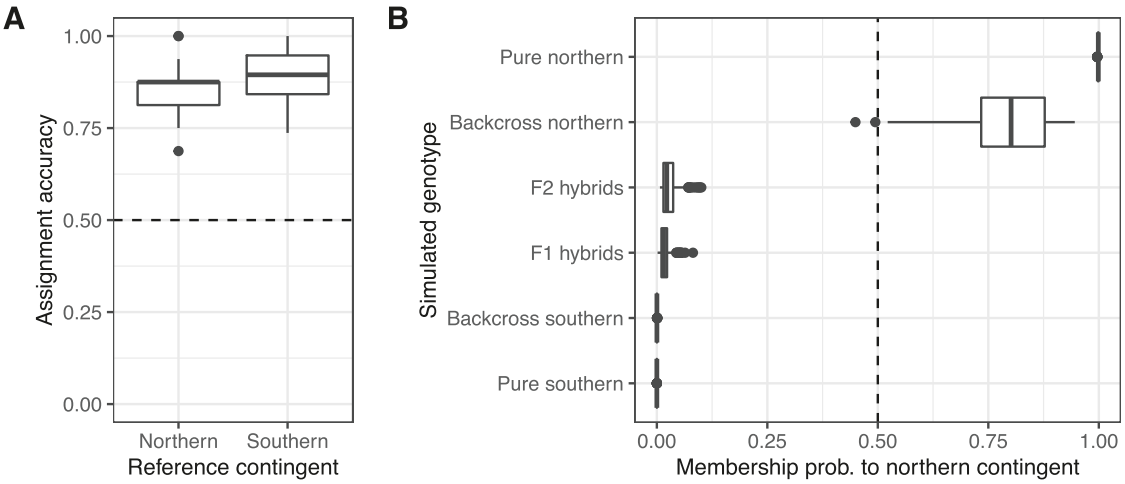
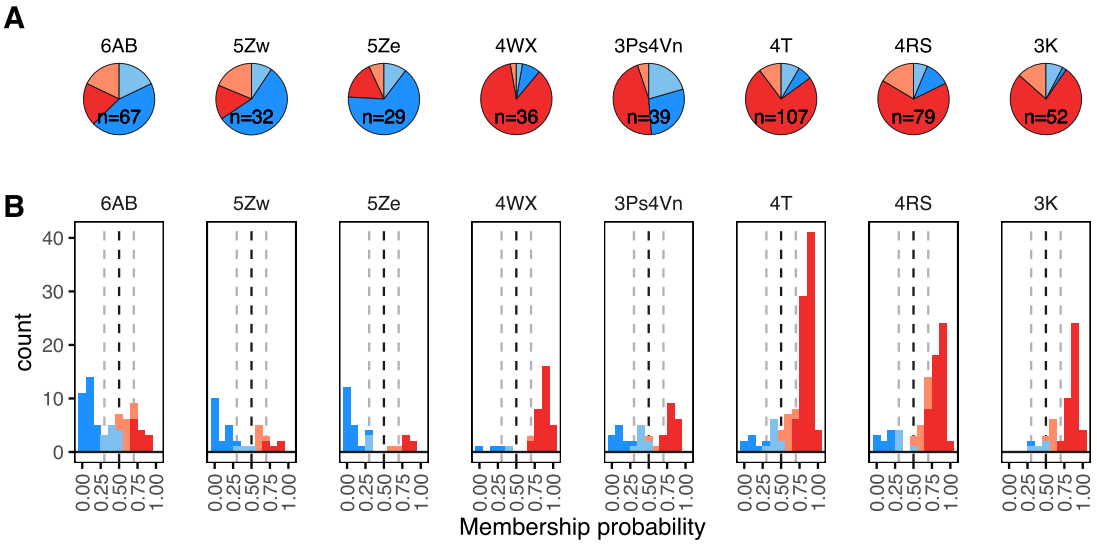


Fig. 5. Genetic assignment results for all nonreference samples to northern (red) and southern (blue) contingents, with indication of membership probability to contingents (light colors: > 50%, dark colors: > 70%). (A) Composition estimates in each NAFO division (North Atlantic Fisheries Organization). (B) Distribution of individual membership probabilities to the northern contingent, with dashed lines representing thresholds of membership probability.



thresholds, respectively. Southern contingent fish were detected in all Canadian NAFO divisions (> 50% and > 70% membership probability threshold; Fig. 5A). The proportion of the southern contingent was as much as 48.8% or 37.9% in the Cabot Strait area (NAFO 3Ps4Vn) but below 17.7% or 14.8% in all other Canadian zones, with the > 50% or 70% membership probability thresholds, respectively (Figs. 5A and 5B).

The genetic assignments and derived composition estimates were not an artefact of the sequencing technology used. Rerunning the same analysis using only one technology (NovaSeq) led to a slightly lower assignment accuracy for the northern contingent (mean ± SD = 70.6% ± 14.3%; Fig. S11A) and an overall larger uncertainty as the number of reference individuals dropped by 30%, from 53 to 36. Genetic assignments and hence composition estimates were not meaningfully impacted (Figs. S11B and S11C).

Discussion

This study confirms the distinctiveness of NWA and NEA mackerel and presents the first genetic evidence for rejecting panmixia between the NWA northern and southern contingents. The genome-wide markers also allowed the development of a genomic tool for reliable genetic assignment of NWA mackerel to contingents. Our results suggest the mixing of the northern and southern contingents on the U.S.

continental shelf area during at least March and April, and provide the first indication that the southern contingent might be present in Canadian waters during some part of the fishing season.

Transatlantic and NWA population structure

No migrants were observed between the NWA and NEA populations of Atlantic mackerel, as evidenced by our SNP markers. This corroborates the results of previous genetic studies based on microsatellites (Gíslason et al. 2020) and SNPs (Rodríguez-Ezpeleta et al. 2016), suggesting limited migration between the two sides of the north Atlantic Ocean. Interestingly, our F_{ST} estimate for transatlantic samples (0.0131) was similar to that of Rodríguez-Ezpeleta et al. (2016) (0.0157), who used a genotyping approach comparable to ours (i.e., based on restriction enzyme site-directed amplification), but applied a different method to generate the libraries (RADseq in Rodríguez-Ezpeleta et al. 2016 versus ddRADseq in this study). Although some of the similarities can be explained by the shared use of certain samples from the Bay of Biscay (38 samples here and 22 samples in Rodríguez-Ezpeleta et al. 2016), the congruence of the results provides more confidence in the conclusion. Transatlantic mackerel migration is currently also not expected, as water temperatures in the Labrador Sea (Yashayaev et al. 2021) are generally below mackerel's thermal preference and even tolerance (Olla et al. 1975, 1976; Studholme et al. 1999). The NEA mackerel population has, in the recent decade, also undergone a spatial expansion resulting in the seasonal colonization of the seas off Southeast Greenland (Jansen et al. 2016; Nøttestad et al. 2016*a*; Olafsdottir et al. 2019). Although initially there was some doubt about the origin of these northern individuals, we provide additional evidence (see Gíslason et al. 2020) based on genetic similarity that these mackerel originate from the NEA population.

We also observed a genetic differentiation between the NWA northern and southern contingents using all samples (F_{ST} and DAPC results) and the reference samples (PCA, F_{ST}, and cross-validation results). The level of genetic differentiation observed is low, but suggests the presence of distinct biological units in the NWA. Low levels of genetic differentiation are not surprising for marine fish with large population sizes and few migration barriers (Nielsen et al. 2009). The genetic differentiation observed between the two contingents in the NWA is in agreement with the suggested philopatric behaviour in Atlantic mackerel (Studholme et al. 1999); a large proportion of each contingent is expected to return to its respective breeding ground, reinforcing relative isolation and genetic differentiation. For several marine migratory fish species with similar philopatric behaviour, genomic studies have likewise demonstrated some degree of differentiation between spawning components (e.g., Atlantic salmon, Bradbury et al. 2018; yellowfin tuna, Pecoraro et al. 2018).

The lack of clear distinctiveness between contingents observed with the PCA and Structure analyses might be explained by the very low levels of genetic differentiation observed. PCA summarizes genetic information in a low dimension space and is useful for the exploration of the structure contained within a data set (Helyar et al. 2011). However, PCA plots may not represent all of the genetic structure (Francois et al. 2010; Alanis-Lobato et al. 2015) and are sensitive to missing values (Novembre and Stephen 2008; Yi and Latch 2022). Similarly, Structure has limited discriminatory power in the presence of low genetic differentiation and moderate gene flow (Latch et al. 2006; Waples and Gaggiotti 2006), which could explain why clusters were not detected in this study. It can also be argued that the PCA and Structure analyses indicate correctly that there is no distinction between the contingents, but that the observed significant pairwise F_{ST} values represent false-positives due to large data sets (Helyar et al. 2011), an argument in line with similar discussions on the importance of focusing on the effect size rather than the significance of P values (Halsey et al. 2015; Berner and Amrhein 2022). In this study, effect sizes matched expectations; the largest F_{ST} values (effect size) were observed between the reference sets and the intercountry comparisons while the smallest ones were mainly within Canadian NAFO division comparisons. Moreover, the observed accuracy of genetic assignments was above random expectation in the cross-validation results (see also the assignment validation section below). Thus, the evidence presented for rejecting panmixia in the NWA seems legitimate and aligns with the ecological knowledge of this species. However, studies that explicitly link F_{ST} to the discrimination capacity of likewise commonly used PCA and admixture analyses would be extremely useful, as low differentiation detections are expected to be encountered more frequently with the advancement of genomic tools.

The observed genetic differentiation between the northern and southern contingents of Atlantic mackerel in the NWA was an order of magnitude lower than the estimated transatlantic difference. The population structure observed within the NWA was driven by neutral SNPs and not by SNPs putatively associated with selection. This genetic differentiation is consequently expected to be the result of current limited gene flow between the two contingents. The relatively recent colonization of the southern Gulf of St. Lawrence following the Last Glacial Maximum between 26.5 and 19.0 thousand years ago (Clark et al. 2009) can explain the small genetic differentiation observed between the two contingents. The recent post-glacial population history coupled with large population size of many marine fish leave little time for populations to accumulate large genetic differentiation (Nielsen et al. 2009). Ongoing effective migration between the two NWA contingents may also contribute to the low genetic differentiation observed.

The small yet significant genetic differentiation between the contingents may explain why previous attempts to clearly discriminate them based on fewer genetic markers were unsuccessful (Maguire et al. 1987; Lambrey de Souza et al. 2006; Gíslason et al. 2020). The level of genetic differentiation between populations of marine fish species varies widely between studies using SNPs, both due to methodological and evolutionary differences (Hemmer-Hansen et al. 2014). In Atlantic mackerel, higher genetic differentiation is observed in the NEA between the Bay of Biscay and the Mediterranean Sea ($F_{ST} = 0.0201$, Rodríguez-Ezpeleta et al. 2016) compared

to those reported here for the NWA. Gene flow between the Atlantic Ocean and Mediterranean Sea is likely more limited compared to that between the contingents of the NWA. Populations from within and outside the Mediterranean Sea are well differentiated for many marine species (e.g., Zane et al. 2000; Patarnello et al. 2007; Luttikhuizen et al. 2008). This is explained by a 5.5 million years old colonization of the Mediterranean Sea and little gene flow through the Gibraltar Strait (Krijgsman et al. 1999).

Assignment validation

The low genetic differentiation between northern and southern contingents within the NWA was reliable for assignments to reference groups. Specifically, the proportion of correctly reclassified individuals was high for both contingents (> 85.5%). The same range of assignment accuracy was observed for two spawning groups of Atlantic cod in the Gulf of Maine, exhibiting a relatively low level of genetic differentiation (F_{ST} = 0.006, O'Donnell and Sullivan 2021). In this study, a correct assignment rate of up to 88.5% was observed when using a small set of highly differentiated loci (25 loci, including 9 under selection). In our case, the highest assignment rate was reached when using all loci (Fig. S1). This is similar to what was observed in Greenland halibut (*Reinhardtius hippoglossoides*), where all loci were necessary to reach the highest assignment rate between two groups showing an F_{ST} of 0.0015, comparable to what we observed for mackerel (Carrier et al. 2020). These results suggest that even in the case of low genetic differentiation, powerful assignment tools based on supervised machine-learning algorithms could reach reliable assignment accuracy. However, assignment tools are not exempted from potential bias in detecting artefacts such as family structure and data errors rather than "true" genetic signals (Waples and Gaggiotti 2006). In this study, we restricted the impact of family structure by using reference individuals from multiple sampling dates, ages (cohorts), and sampling locations. We also tested different sources of data errors such as the presence of missing values and the sequencing technology, and found that our assignment tool was reliable in both contexts. The capacity of the assignment tool to perform genetic assignments with an accuracy above 85% suggested that there is a distinct genetic signature between both contingents, and is in line with F_{ST} and DAPC results.

Improvements to the reference samples used for the southern contingent could further minimize potential bias in assignments. The primary commercial fishery in the U.S. does not overlap with mackerel's reproductive season, and hence there were no samples available from fish that can with absolute certainty be assigned to the southern contingent. We, therefore, assumed that samples from the Gulf of Maine (NAFO 5YZe) caught in April largely consisted mostly of southern contingent individuals, despite the fact that those may have been admixed with the northern contingent (Arai et al. 2021). This would explain the biased classification of admixed individuals (F1 and F2) to the southern contingent, as well as higher than expected membership probability for backcrossed individuals within the southern contingent (Fig. 4B). Consequently, the assignment results may overestimate the U.S. composition in both the Canadian and U.S. waters if admixed individuals were present in the samples. To overcome this potential bias, improvement to the southern contingent reference samples are needed through sampling ripe mackerel during May in U.S. waters.

Composition estimations

Northern and southern contingents of Atlantic mackerel have long been known to mix along the U.S. Atlantic shelf during late fall, winter, and spring (Sette 1950). The intensity of this overlap in space and time is, however, poorly understood, as individuals should not only be discriminable but large amounts of samples with a large coverage would be necessary to fully understand this process. Our study nonetheless indicated that in March–April 2019, a relatively high percentage of individuals from across the U.S. shelf showed a northern population genetic signature (up to 33.6% of nonreference individuals when considering a membership probability > 50%). Previous studies using otolith stable isotope analysis have also confirmed that along the U.S. coast, northern and southern contingents are mixing during this period, with individuals from the northern contingent representing between 16.7% and 76.9% of U.S. samples, depending on the year (study period: 1998–2000 and 2011–2016; Arai et al. 2021). This range included our estimate, despite that these ratio estimates from otolith stable isotope analyses were based on relatively small samples (N between 16 and 60 depending on the year), collected over a longer period (between January and May) and in years prior to ours so that, for instance, the range of ages and lengths of the fish could also differ. Regardless of the method used, larger sample sizes covering a wider spatiotemporal range will be necessary to understand the migratory dynamics and overlap of both contingents and to decompose landings data.

The genetic assignment also indicated that the southern contingent may visit Canadian waters from May to October, for which so far no evidence existed. In the Cabot Strait (NAFO 3Ps4Vn), the proportion of southern contingent fish varied between 37.9% and 48.8% (out of 39 fish) depending on the classification threshold used (i.e., > 50% or 70% membership probability). This proportion was much greater than the error rate of assignments estimated from a cross-validation test, and thus the presence of the southern contingent in this area in 2019 is likely. In other areas within Canadian waters, only up to 17.7% of sampled fish were assigned to the southern contingent, which present more uncertainties. Those results, although unexpected, should not be completely surprising in light of the common extensive northward migrations mackerel can undertake, and the shift in distribution of the southern contingent towards warmer northerly waters in recent years (Overholtz et al. 2011; McManus et al. 2018).

Conclusion

The genetic distinctiveness of both contingents in the NWA observed with genome-wide markers combined with genetic assignment approaches could allow a better understanding of mixing of both contingents through space, time, and

mackerel's life-history, which is largely unknown. While improvements such as a reference genome and better reference samples are desirable, the approach tested here proved to be an interesting alternative or complement to the previously developed discriminant methods based on otolith stable isotopes. Incorporating information on population structure into stock management could reduce overfishing and improve stock assessment (Waples et al. 2008; Spies and Punt 2015). Genetic information could allow Canadian and U.S. fisheries management agencies to infer the proportion of each contingent in landings throughout the fishing season, an information that could help rebuild these stocks by developing future harvesting strategies that are sustainable for each contingent. Especially the southern contingent, which is at a much lower biomass level than the northern contingent (Richardson et al. 2020) and is not managed separately (NEFSC 2018), risks to be unsustainably harvested without population-specific advice and management (e.g., Ying et al. 2011).

Acknowledgements

We are deeply grateful to Linda Girard and Mélanie Boudreau who phenotyped each individual and prepared samples for genetic analyses, and to Grégoire Cortial, Laury-Ann Dumoulin, Jade Larivière, and Eric Parent from the MLI Laboratory of Genomics who extracted and prepared DNA for sequencing. We also thank David Secor, Arai Kohma, Matthew Craig, and two anonymous reviewers for truly helpful comments on a previous version of the manuscript.

Article information

History dates

Received: 3 October 2022
Accepted: 2 February 2023
Accepted manuscript online: 23 February 2023
Version of record online: 23 March 2023

Copyright

Data availability

R scripts and data sets (including both SNPs' panels) are available in the GenomicsMLI-DFO/Mackerel_NWA_2021 github repository, https://github.com/GenomicsMLI-DFO/Mackerel_NWA_2021.

Author information

Author ORCIDs

Audrey Bourret https://orcid.org/0000-0002-5387-3011
Teunis Jansen https://orcid.org/0000-0001-8722-4493

Author contributions

Conceptualization: SP, DER, MC, GJP
Data curation: AB, AS
Formal analysis: AB
Funding acquisition: MC, GJP
Investigation: MC, GJP
Methodology: MC, GJP
Project administration: MC, GJP
Resources: AB, AS, EVB, KLC, TJ, DER, MC, NR
Software: AB
Supervision: GJP
Validation: AB
Visualization: AB
Writing – original draft: AB, GJP
Writing – review & editing: AB, AS, EVB, SP, KLC, TJ, DER, MC, NR, GJP

Competing interests

The authors declare there are no competing interests.

Funding information

This project was funded by the DFO International Governance Strategy Science Program (2018–2020) and DFO Fisheries Management, either directly or through the catch monitoring program (2019, 2020).

Supplementary material

Supplementary data are available with the article at https://doi.org/10.1139/cjfas-2022-0232.

References

Alanis-Lobato, G., Cannistraci, C.V., Eriksson, A., Manica, A., and Ravasi, T. 2015. Highlighting nonlinear patterns in population genetics datasets. Sci. Rep. **5**: 1–8. doi:10.1038/srep08140.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [accessed 4 October 2018].

Arai, K., Castonguay, M., and Secor, D.H. 2021. Multi-decadal trends in contingent mixing of Atlantic mackerel (*Scomber scombrus*) in the Northwest Atlantic from otolith stable isotopes. Sci. Rep. **11**(1): 1–13. doi:10.1038/s41598-021-86116-2. PMID: 33414495.

Beckett, J.S., Stobo, W.T., and Dickson, C.A. 1974. Southwesterly migration of Atlantic mackerel, *Scomber scombrus*, tagged off Nova Scotia. International Commission for the Northwest Atlantic Fisheries. Vol. **74**(94), pp. 4.

Begg, G.A., and Waldman, J.R. 1999. An holistic approach to fish stock identification. Fish. Res. **43**(1–3): 35–44. doi:10.1016/S0165-7836(99)00065-X.

Berg, P.R., Star, B., Pampoulie, C., Bradbury, I.R., Bentzen, P., Hutchings, J.A., et al. 2017. Trans-oceanic genomic divergence of Atlantic cod ecotypes is associated with large inversions. Heredity, **119**(6): 418–428. doi:10.1038/hdy.2017.54. PMID: 28930288.

Berner, D., and Amrhein, V. 2022. Why and how we should join the shift from significance testing to estimation. J. Evol. Biol. **35**: 777–787. doi:10.1111/jeb.14009. PMID: 35582935.

Berrien, P.L. 1978. Eggs and larvae of *Scomber scombrus* and *Scomber japonicus* in continental shelf waters between Massachusetts and Florida. Fish. Bull. **76**: 95–115.

Besnier, F., and Glover, K.A. 2013. ParallelStructure: A R Package to Distribute Parallel Runs of the Population Genetics Program STRUCTURE on Multi-Core Computers. PLoS ONE **8**(7): e70651. doi:10.1371/journal.pone.0070651.

Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, **30**(15): 2114–2120. doi:10.1093/bioinformatics/btu170. PMID: 24695404.

Bradbury, I.R., Wringe, B.F., Watson, B., Paterson, I., Horne, J., Beiko, R., et al. 2018. Genotyping-by-sequencing of genome-wide microsatellite loci reveals fine-scale harvest composition in a coastal Atlantic salmon fishery. Evol. Appl. **11**(6): 918–930. doi:10.1111/eva.12606. PMID: 29928300.

Brosset, P., Smith, A.D., Plourde, S., Castonguay, M., Lehoux, C., and Van Beveren, E. 2020. A fine-scale multi-step approach to understand fish recruitment variability. Sci. Rep. **10**: 1–14. doi:10.1038/s41598-020-73025-z. PMID: 31913322.

Cadrin, S.X. 2020. Defining spatial structure for fishery stock assessment. Fish. Res. **221**: 105397. Elsevier. doi:10.1016/j.fishres.2019.105397.

Cadrin, S.X., Kerr, L.A., and Mariani, S. 2014. Stock identification methods: applications in fisheries science. 2nd ed., Elsevier Academic Press. doi:10.1016/C2011-0-07625-1.

Carrier, E., Ferchaud, A.L., Normandeau, E., Sirois, P., and Bernatchez, L. 2020. Estimating the contribution of Greenland Halibut (*Reinhardtius hippoglossoides*) stocks to nurseries by means of genotyping-by-sequencing: sex and time matter. Evol. Appl. **13**(9): 2155–2167. doi:10.1111/eva.12979. PMID: 33005216.

Castonguay, M., Simard, P., and Gagnon, P. 1991. Usefulness of Fourier analysis of otolith shape for Atlantic mackerel (*Scomber scombrus*) stock discrimination. Can. J. Fish. Aquat. Sci. **48**(2): 296–302. doi:10.1139/f91-041.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and Cresko, W.A. 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. **22**(11): 3124–3140. doi:10.1111/mec.12354. PMID: 23701397.

Cattell, R.B. 1966. The scree test for the number of factors. Multivar. Behav. Res. **1**(2): 245–276. doi:10.1207/s15327906mbr0102. PMID: 26828106.

Chen, K.Y., Marschall, E.A., Sovic, M.G., Fries, A.C., Gibbs, H.L., and Ludsin, S.A. 2018. assignPOP: an r package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. Methods Ecol. Evol. **9**(2): 439–446. doi:10.1111/2041-210X.12897.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., et al. (2009). The last glacial maximum. Science, **325**(5941): 710–714. doi:10.1126/science.1172873. PMID: 19661421.

Cope, J.M., and Punt, A.E. 2011. Reconciling stock assessment and management scales under conditions of spatially varying catch histories. Fish. Res. **107**(1–3): 22–38. Elsevier B.V. doi:10.1016/j.fishres.2010.10.002.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., et al. 2011. The variant call format and VCFtools. Bioinformatics **27**(15): 2156–2158. doi:10.1093/bioinformatics/btr330. PMID: 21653522.

DFO. 2021. Assessment of the northern contingent of Atlantic mackerel (*Scomber scombrus*) in 2020. Canadian Science Advisory Secretariat Science Advisory Report 2021/029.

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. **14**(8): 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x. PMID: 15969739.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics, **32**(19): 3047–3048. doi:10.1093/bioinformatics/btw354. PMID: 27312411.

Foll, M., and Gaggiotti, O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics, **180**(2): 977–993. doi:10.1534/genetics.108.092221. PMID: 18780740.

Francois, O., Currat, M., Ray, N., Han, E., Excoffier, L., and Novembre, J. 2010. Principal component analysis under population genetic models of range expansion and admixture. Mol. Biol. Evol. **27**(6): 1257–1268. doi:10.1093/molbev/msq010. PMID: 20097660.

Gíslason, D., Helyar, S.J., Óskarsson, G.J., Ólafsdóttir, G., Slotte, A., Jansen, T., et al. 2020. The genetic composition of feeding aggregations of the Atlantic mackerel *(Scomber scombrus*) in the central north Atlantic: a microsatellite loci approach. ICES J. Mar. Sci. **77**(2): 604–612. doi:10.1093/icesjms/fsaa003.

Grégoire, F., Castonguay, M., Maguire, J.J.-J., Gregoire, F., Castonguay, M., and Maguire, J.-J. 2010. Is the index from the NEFSC spring research bottom trawl surveys representative of the abundance of the so-called northern contingent of Atlantic mackerel (*Scomber scombrus* L.)? Transboundary Resources Assessment Committee Documents 2010/11.

Gruber, B., Unmack, P.J., Berry, O.F., and Georges, A. 2018. dartr: an r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Mol. Ecol. Resour. **18**(3): 691–699. doi:10.1111/1755-0998.12745. PMID: 29266847.

Halsey, L.G., Curran-Everett, D., Vowler, S.L., and Drummond, G.B. 2015. The fickle P value generates irreproducible results. Nat. Methods, **12**(3): 179–185. doi:10.1038/nmeth.3288. PMID: 25719825.

Helyar, S.J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M.I., Ogden, R., Limborg, M.T., et al. 2011. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Mol. Ecol. Resour. **11**(SUPPL. 1): 123–136. doi:10.1111/j.1755-0998.2010.02943.x. PMID: 21429169.

Hemmer-Hansen, J., Therkildsen, N.O., and Pujolar, J.M. 2014. Population genomics of marine fishes: next-generation prospects and challenges. Biol. Bull. **227**(2): 117–132. doi:10.1086/BBLv227n2p117. PMID: 25411371.

Jakobsson, M., and Rosenberg, N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics, **23**(14): 1801–1806. doi:10.1093/bioinformatics/btm233. PMID: 17485429.

Jansen, T., Post, S., Kristiansen, T., Oskarsson, G.J., Boje, J., MacKenzie, B.R., et al. 2016. Ocean warming expands habitat of a rich natural resource and benefits a national economy. Ecol. Appl. **26**: 2021–2032. doi:10.1002/eap.1384. PMID: 27755730.

Jansen, T., Slotte, A., Christina dos Santos Schmidt, T., Reedtz Sparrevohn, C., Arge Jacobsen, J., and Sigurd Kjesbu, O. 2021. Bioenergetics of egg production in Northeast Atlantic mackerel changes the perception of fecundity type and annual trends in spawning stock biomass. Prog. Oceanogr. **198**: 102658. doi:10.1016/j.pocean.2021.102658.

Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics, **24**(11): 1403–1405. doi:10.1093/bioinformatics/btn129. PMID: 18397895.

Kamvar, Z.N., Tabima, J.F., and Grünwald, N.J. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ **2**: e281. doi:10.7717/peerj.281.

Knaus, B.J., and Grünwald, N.J. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. **17**(1): 44–53. doi:10.1111/1755-0998.12549. PMID: 27401132.

Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., and Wilson, D.S. 1999. Chronology, causes and progression of the Messinian salinity crisis. Nature, **400**(6745): 652–655. doi:10.1038/23231.

Lambrey de Souza, J., Sévigny, J., Chanut, J., Barry, W.F., and Grégoire, F. 2006. High genetic variability in the mtDNA control region of a Northwestern Atlantic teleost, *Scomber scombrus* L. Can. Tech. Rep. Fish. Aquat. Sci. 33.

Latch, E.K., Dharmarajan, G., Glaubitz, J.C., and Rhodes, O.E. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conserv. Genet. **7**: 295–302. doi:10.1007/s10592-005-9098-1.

Luttikhuizen, P.C., Campos, J., Bleijswijk, J., Peijnenburg, K.T.C.A., and van der Veer, H.W. 2008. Phylogeography of the common shrimp, Crangon crangon (L.) across its distribution range. Mol. Phylogenet. Evol. **46**(3): 1015–1030. doi:10.1016/j.ympev.2007.11.011. PMID: 18207428.

MacKay, K.T., and Garside, E.T. 1969. Meristic analyses of Atlantic mackerel, *Scomber scombrus*, from the North American coastal populations. J. Fish. Res. Board Can. **26**(9): 2537–2540. doi:10.1139/f69-248.

Maguire, J.-J. 1981. Maturité, fécondité, ponte et évaluation de la taille du stock reproducteur du maquereau atlantique (*Scomber scombrus*) dans le golfe de St.-Laurent. Master's thesis, Université Laval. pp. 137.

Maguire, J.-J., Chagnon, Y.C., Castonguay, M., and Mercille, B. 1987. A review of mackerel management areas in the Northwest Atlantic. CAFSAC Res. Doc. **87–71**: 31.

Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. 2010. Robust relationship inference in genome-wide

association studies. Bioinformatics, **26**(22): 2867–2873. doi:10.1093/bioinformatics/btq559. PMID: 20926424.

McManus, M.C., Hare, J.A., Richardson, D.E., and Collie, J.S. 2018. Tracking shifts in Atlantic mackerel (*Scomber scombrus*) larval habitat suitability on the Northeast U.S. continental shelf. Fish. Oceanogr. **27**(1): 49–62. doi:10.1111/fog.12233.

Miller, J.M., Cullingham, C.I., and Peery, R.M. 2020. The influence of a priori grouping on inference of genetic clusters: simulation study and literature review of the DAPC method. Heredity, **125**(5): 269–280. doi:10.1038/s41437-020-0348-2. PMID: 32753664.

Moores, J.A., Winters, G.H., and Parsons, L.S. 1975. Migrations and biological characteristics of Atlantic mackerel (*Scomber scombrus*) occurring in Newfoundland Waters. J. Fish. Res. Board Can. **32**(8): 1347–1357. doi:10.1139/f75-155.

Morse, W.W. 1980. Spawning and fecundity of Atlantic mackerel, *Scomber scombrus*, in the Middle Atlantic Bight. Fish. Bull. **78**: 103–108.

Mullins, R.B., McKeown, N.J., Sauer, W.H.H., Shaw, P.W., and Grant, W.S. 2018. Genomic analysis reveals multiple mismatches between biological and management units in yellowfin tuna (*Thunnus albacares*). ICES J. Mar. Sci. **75**(6): 2145–2152. doi:10.1093/icesjms/fsy102.

NEFSC. 2018. 64th Northeast regional stock assessment workshop (64th SAW) assessment report. U.S. Department of Commerce, Northeast Fisheries Science Center Reference Document 18–04. pp. 529.

Nesbø, C.L., Rueness, E.K., Iversen, S.A., Skagen, D.W., and Jakobsen, K.S. 2000. Phylogeography and population history of Atlantic mackerel (*Scomber scombrus* L.): a genealogical approach reveals genetic structuring among the eastern Atlantic stocks. Proc. R. Soc. B Biol. Sci. **267**(1440): 281–292. doi:10.1098/rspb.2000.0998.

Nielsen, E.E., Hemmer-Hansen, J., Larsen, P.F., and Bekkevold, D. 2009. Population genomics of marine fishes: Identifying adaptive variation in space and time. Mol. Ecol. **18**(15): 3128–3150. doi:10.1111/j.1365-294X.2009.04272.x. PMID: 19627488.

Nøttestad, L., Diaz, J., Penã, H., Søiland, H., Huse, G., and Fernö, A. 2016a. Feeding strategy of mackerel in the Norwegian Sea relative to currents, temperature, and prey. ICES J. Mar. Sci. **73**(4): 1127–1137. doi:10.1093/icesjms/fsv239.

Nøttestad, L., Utne, K.R., Óskarsson, G.J., Jónsson, S.Þ., Jacobsen, J.A., Tangen, Ø., et al. 2016b. Quantifying changes in abundance, biomass, and spatial distribution of Northeast Atlantic mackerel (*Scomber scombrus*) in the Nordic seas from 2007 to 2014. ICES J. Mar. Sci. **73**(2): 359–373. doi:10.1093/icesjms/fsv218.

Novembre, J., and Stephens, M. 2008. Interpreting principal component analyses of spatial population genetic variation. Nat. Genet. **40**(5): 646–649. doi:10.1038/ng.139.

O'Donnell, T.P., and Sullivan, T.J. 2021. Low-coverage whole-genome sequencing reveals molecular markers for spawning season and sex identification in Gulf of Maine Atlantic cod (*Gadus morhua*, Linnaeus 1758). Ecol. Evol. **11**(15): 10659–10671. doi:10.1002/ece3.7878.

Olafsdottir, A.H., Utne, K.R., Jacobsen, J.A., Jansen, T., Óskarsson, G.J., Nøttestad, L., et al. 2019. Geographical expansion of Northeast Atlantic mackerel (Scomber scombrus) in the Nordic Seas from 2007 to 2016 was primarily driven by stock size and constrained by low temperatures. Deep-Sea Res. II Top. Stud. Oceanogr. **159**: 152–168. doi:10.1016/j.dsr2.2018.05.023.

Olla, B.L., Bejda, A.J., and Studholme, A.L. 1976. Swimming speeds of Atlantic mackerel, *Scomber scombrus*, under laboratory conditions: relation to capture by trawling. International Commission for the Northwest Atlantic Fisheries Document 76/XII/143. pp. 6.

Olla, B.L., Studholme, A.L., Bejda, A.J., Samet, C., and Martin, A.D. 1975. The effect of temperature on the behaviour of marine fishes: a comparison among Atlantic mackerel, *Scomber scombrus*, bluefish, *Pomatomus saltatrix*, and tautog, *Tautoga onitis*. *In* Combined effects of radioactive, chemical and thermal releases to the environment. International Atomic Energy Agency, Vienna. pp. 299–308.

Ovenden, J.R., Berry, O., Welch, D.J., Buckworth, R.C., and Dichmont, C.M. 2015. Ocean's eleven: a critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. Fish Fish. **16**(1): 125–159. doi:10.1111/faf.12052.

Overholtz, W.J., Hare, J.A., and Keith, C.M. 2011. Impacts of interannual environmental forcing and climate change on the distribution of Atlantic mackerel on the U.S. Northeast continental shelf. Mar. Coast. Fish. **3**(1): 219–232. doi:10.1080/19425120.2011.578485.

Parsons, L.S., and Moores, J.A. 1973. Long-distance migration of an Atlantic mackerel, *Scomber scombrus*, tagged in Newfoundland waters. International Commission for the Northwest Atlantic Fisheries Document. International Commission for the Northwest Atlantic Fisheries, Dartmouth, NS. Vol. **73**(82), pp. 2.

Patarnello, T., Volckaert, F.A.M.J., and Castilho, R. 2007. Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? Mol. Ecol. **16**(21): 4426–4444. doi:10.1111/j.1365-294X.2007.03477.x.

Pecoraro, C., Babbucci, M., Franch, R., Rico, C., Papetti, C., Chassot, E., et al. 2018. The population genomics of yellowfin tuna (*Thunnus albacares*) at global geographic scale challenges current stock delineation. Sci. Rep. **8**(1): 13890. doi:10.1038/s41598-018-32331-3.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., and Hoekstra, H.E. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE, **7**: 37135. doi:10.1371/journal.pone.0037135.

Poland, J.A., Brown, P.J., Sorrells, M.E., and Jannink, J.L. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE, **7**(2). doi:10.1371/journal.pone.0032253.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, **155**(2): 945–959. doi:10.1093/genetics/155.2.945.

Privé, F., Luu, K., Vilhjálmsson, B.J., Blum, M.G.B., and Rosenberg, M. 2020. Performing highly efficient genome scans for local adaptation with R Package pcadapt Version 4. Mol. Biol. Evol. **37**(7): 2153–2154. doi:10.1093/molbev/msaa053.

Puncher, G.N., Cariani, A., Maes, G.E., Van Houdt, J., Herten, K., Cannas, R., et al. 2018. Spatial dynamics and mixing of bluefin tuna in the Atlantic Ocean and Mediterranean Sea revealed using next-generation sequencing. Mol. Ecol. Resour. **18**(3): 620–638. doi:10.1111/1755-0998.12764.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., et al. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. **81**(3): 559–575. doi:10.1086/519795.

Redding, S.G., Cooper, L.W., Castonguay, M., Wiernicki, C., and Secor, D.H. 2020. Northwest Atlantic mackerel population structure evaluated using otolith δ18O composition. ICES J. Mar. Sci. **77**(7): 2582–2589. doi:10.1093/icesjms/fsaa117.

Richardson, D.E., Carter, L., Curti, K., Marancik, K.E., and Castonguay, M. 2020. Changes in the spawning distribution and biomass of Atlantic mackerel (*Scomber scombrus*) in the western Atlantic over four decades. Fish. Bull. **118**: 120–134. doi:10.7755/FB.118.2.2.

Rochette, N.C., and Catchen, J.M. 2017. Deriving genotypes from RAD-seq short-read data using Stacks. Nat. Protoc. **12**(12): 2640–2659. doi:10.1038/nprot.2017.123.

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. Mol. Ecol. **28**(21): 4737–4754. doi:10.1111/mec.15253.

Rodríguez-Ezpeleta, N., Bradbury, I.R., Mendibil, I., Álvarez, P., Cotano, U., and Irigoien, X. 2016. Population structure of Atlantic mackerel inferred from RAD-seq-derived SNP markers: effects of sequence clustering parameters and hierarchical SNP selection. Mol. Ecol. Resour. **16**(4): 991–1001. doi:10.1111/1755-0998.12518.

Rodríguez-Ezpeleta, N., Díaz-Arce, N., Walter, J.F., Richardson, D.E., Rooker, J.R., Nøttestad, L., et al. 2019. Determining natal origin for improved management of Atlantic bluefin tuna. Front. Ecol. Environ. **17**(8): 439–444. doi:10.1002/fee.2090.

Sette, O.E. 1943. Biology of the Atlantic mackerel (*Scomber scombrus*) of north America. Part I: early life history, including the growth, drift, and mortality of the egg and larval populations. Fish. Bull. **50**: 149–237.

Sette, O.E. 1950. Biology of the Atlantic mackerel (*Scomber scombrus*) of North America. Part II: migration and habits. Fish. Bull. **51**: 251–358.

Smith, A.D., Van Beveren, E., Girard, L., Boudreau, M., Brosset, P., Castonguay, M., and Plourde, S. 2020. Atlantic mackerel (*Scomber scombrus* L.) in NAFO Subareas 3 and 4 in 2018. Canadian Science Advisory Secretariat Science Research Document 2020/013. pp. iv + 37.

Spies, I., and Punt, A.E. 2015. The utility of genetics in marine fisheries management: a simulation study based on Pacific cod off Alaska. Can. J. Fish. Aquat. Sci. **72**: 1415–1432. doi:10.1139/cjfas-2014-0050.

Stanley, R.R.E., Jeffery, N.W., Wringe, B.F., DiBacco, C., and Bradbury, I.R. 2017. genepopedit: a simple and flexible tool for manipulating multilocus molecular data in R. Mol. Ecol. Resour. **17**(1): 12–18. doi:10.1111/1755-0998.12569.

Stobo, W.T. 1976. Movements of mackerel tagged in Subarea 4. International Commission for the Northwest Atlantic Fisheries Document. Vol. **76**(49), pp. 5.

Studholme, A.L., Packer, D.B., Berrien, P.L., Johnson, D.L., Zetlin, C.a., and Morse, W.W. 1999. Atlantic Mackerel, *Scomber scombrus*, life history and habitat characteristics. NOAA Technical Report NMFS-NE-141 (September). pp. 44.

Van Beveren, E., Duplisea, D.E., Brosset, P., and Castonguay, M. 2019. Assessment modelling approaches for stocks with spawning components, seasonal and spatial dynamics, and limited resources for data collection. PLoS ONE, **14**(9). doi:10.1371/journal.pone.0222472.

Van Beveren, E., Duplisea, D.E., Marentette, J.R., Smith, A., and Castonguay, M. 2020a. An example of how catch uncertainty hinders effective stock management and rebuilding. Fish. Res. **224**: 105473. doi:10.1016/j.fishres.2019.105473.

Van Beveren, E., Marentette, J.R., Smith, A.D., Castonguay, M., and Duplisea, D.E. 2020b. Evaluation of rebuilding strategies for northwestern Atlantic mackerel (NAFO subareas 3 and 4). Canadian Science Advisory Secretariat Science Research Document 2020/021. pp. v + 56.

Van Beveren, E., Plourde, S., Pepin, P., Cogliati, K., and Castonguay, M. 2022. A review of the importance of various areas for northern contingent West-Atlantic mackerel spawning. ICES J. Mar. Sci. 1–15. doi:10.1093/icesjms/fsac211.

Vaux, F., Bohn, S., Hyde, J.R., and O'Malley, K.G. 2021. Adaptive markers distinguish North and South Pacific Albacore amid low population differentiation. Evol. Appl. **14**(5): 1343–1364. doi:10.1111/eva.13202.

Waples, R.S., and Gaggiotti, O. 2006. INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol. Ecol. **15**(6): 1419–1439. doi:10.1111/j.1365-294X.2006.02890.x.

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

Weir, B.S., and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution (N.Y.), **38**(6): 1358–1370.

Wringe, B.F., Stanley, R.R.E., Jeffery, N.W., Anderson, E.C., and Bradbury, I.R. 2017. hybriddetective: a workflow and package to facilitate the detection of hybridization using genomic data in r. Mol. Ecol. Resour. **17**(6): e275–e284. doi:10.1111/1755-0998.12704.

Yashayaev, I., Peterson, I., and Wang, Z. 2021. Meteorological, sea ice, and physical oceanographic conditions in the Labrador sea during 2018. Canadian Science Advisory Secretariat Science Research Document 2021/042. pp. iv + 26.

Yi, X., and Latch, E.K. 2022. Nonrandom missing data can bias principal component analysis inference of population genetic structure. Mol. Ecol. Resour. **22**(2): 602–611. doi:10.1111/1755-0998.13498.

Ying, Y., Chen, Y., Lin, L., and Gao, T. 2011. Risks of ignoring fish population spatial structure in fisheries management. Can. J. Fish. Aquat. Sci. **68**(12): 2101–2120. doi:10.1139/F2011-116.

Zane, L., Ostellari, L., Maccatrozzo, L., Bargelloni, L., Cuzin-Roudy, J., Buchholz, F., and Patarnello, T. 2000. Genetic differentiation in a pelagic crustacean (*Meganyctiphanes norvegica*: Euphausiacea) from the North East Atlantic and the Mediterranean Sea. Mar. Biol. **136**(2): 191–199. doi:10.1007/s002270050676.