Check for updates

RESEARCH ARTICLE

REVISED Atlantic mackerel population structure does not support genetically distinct spawning components

[version 2; peer review: 4 approved, 2 approved with reservations]

Alice Manuzzi[1], Imanol Aguirre-Sarabia[1], Natalia Díaz-Arce[1], Dorte Bekkevold[2], Teunis Jansen[2,3], Jessica Gomez-Garrido[4,5], Tyler S. Alioto[4,5], Marta Gut[4,5], Martin Castonguay[6], Sonia Sanchez-Maroño[1], Paula Álvarez[1], Naiara Rodriguez-Ezpeleta[1]

[1]AZTI, Marine Research, Basque Research and Technology Alliance (BRTA), Sukarrieta, Spain
[2]DTU Aqua, National Institute of Aquatic Resources, Section for Marine Living Resources, Silkeborg, Denmark
[3]GINR, Greenland Institute of Natural Resources, Nuuk, Greenland
[4]Centro Nacional de Análisis Genómico (CNAG), Barcelona, Spain
[5]Universitat de Barcelona (UB), Barcelona, Spain
[6]Maurice Lamontagne Institute, Fisheries and Oceans Canada, Mont-Joli, Canada, Mont-Joli, Canada

v2 **First published:** 24 Apr 2024, **4**:82
https://doi.org/10.12688/openreseurope.17365.1
Latest published: 20 Jan 2025, **4**:82
https://doi.org/10.12688/openreseurope.17365.2

Abstract

Background

The Atlantic mackerel, *Scomber scombrus* (Linnaeus, 1758) is a commercially valuable migratory pelagic fish inhabiting the northern Atlantic Ocean and the Mediterranean Sea. Given its highly migratory behaviour for feeding and spawning, several studies have been conducted to assess differentiation among spawning components to better define management units, as well as to investigate possible adaptations to comprehend and predict recent range expansion northwards.

Methods

Here, the genome of *S. scombrus* was sequenced and annotated, as an increasing number of population genetic studies have proven the relevance of reference genomes to investigate genomic markers/regions potentially linked to differences at finer scale. Such reference genome was used to map Restriction-site-associated sequencing (RAD-seq) reads for SNP discovery and genotyping in more than 500 samples distributed along the species range. The resulting genotyping tables have been used to perform connectivity

Open Peer Review

Approval Status ? ? ✓ ✓ ✓ ✓

	1	2	3	4	5	6
version 2 (revision) 20 Jan 2025				✓ view	✓ view	✓ view
version 1 24 Apr 2024	? view	? view	✓ view			

1. **Natalia Bayona**, Emory University, Oxford, USA
2. **Sven Winter**, University of Veterinary Medicine, Vienna, Austria
Senckenberg Biodiversity and Climate Research Centre (BiK-F), Frankfurt am Main, Germany
3. **Julia Tovar Verba**, Ludwig-Maximilians-Universität München, Munich, Germany
4. **Yusuf Bektas**, Karadeniz Technical University, Ortahisar, Turkey

and adaptation analyses.

Results

The assembly of the reference genome for *S. scombrus* resulted in a genome of 741 Mb. Our population genetic results show that the Atlantic mackerel consist of three previously known genetically isolated units (Northwest Atlantic, Northeast Atlantic, Mediterranean), and provide no evidence for genetically distinct spawning components within the Northwest or Northeast Atlantic.

Conclusions

Therefore, our findings resolved previous uncertainties by confirming the absence of genetically isolated spawning components in each side of the northern Atlantic, thus rejecting homing behaviour and the need to redefine management boundaries in this species. In addition, no further genetic signs of ongoing adaptation were detected in this species.

Plain language summary
The Atlantic mackerel is a commercially valuable migratory pelagic fish inhabiting the northern Atlantic Ocean and the Mediterranean Sea. In both sides of the Atlantic, this species spawns at several locations (spawning components) while migrating north for feeding. It has been hypothesized that the mackerel has a homing behaviour, meaning that they return to spawn to where they were born, in which case, genetic differences should be visible between spawning components. Here, we have sequenced the complete genome of Atlantic mackerel and performed population genetic analyses based on thousands of markers. Our results provide no evidence for genetically distinct spawning components which implies that, for fisheries management, this species should be managed as one unit in each side of the Atlantic.

Keywords
Atlantic mackerel, complete genome, RAD-seq, population structure, genome-wide SNPs, fisheries management

5. **Gaetano Catanese**, IRFAP LIMIA (Laboratorio de Investigaciones Marinas y Acuicultura) – Govern de les Illes Balears, Balearic Islands, Spain
6. **John S Hargrove**, Pacific States Marine Fisheries Commission, Trout Rd, USA

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Earth and Environmental Sciences gateway.

This article is included in the Horizon 2020 gateway.

Corresponding author: Naiara Rodriguez-Ezpeleta (nrodriguez@azti.es)

Author roles: Manuzzi A: Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Aguirre-Sarabia I**: Data Curation, Formal Analysis, Methodology; **Díaz-Arce N**: Formal Analysis, Methodology, Resources, Software, Visualization, Writing – Review & Editing; **Bekkevold D**: Investigation, Validation, Writing – Review & Editing; **Jansen T**: Investigation, Validation, Writing – Review & Editing; **Gomez-Garrido J**: Formal Analysis, Investigation, Methodology, Visualization, Writing – Review & Editing; **Alioto TS**: Formal Analysis, Methodology, Software, Visualization, Writing – Review & Editing; **Gut M**: Formal Analysis, Methodology, Resources; **Castonguay M**: Resources, Writing – Review & Editing; **Sanchez-Maroño S**: Investigation, Writing – Review & Editing; **Álvarez P**: Conceptualization, Funding Acquisition, Resources, Supervision, Validation, Writing – Review & Editing; **Rodriguez-Ezpeleta N**: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 869300 (Climate Change and Future Marine Ecosystem Services and Biodiversity [FutureMARES]). This project was also funded by the Department of Agriculture and Fisheries of the Basque government (projects CABALL, GENPES and GENGES), by the General Secretary of Fisheries of the Spanish Government (project MACKSG), and by the Spanish Ministry of Science and Innovation (project GIFAMAN with reference PID2020-115656RB-I00). This manuscript is contribution number 1211 from the Marine Research Division of AZTI. This project was supported by the Department of Agriculture, Fisheries and Food of the Basque Country and the General Secretary of Fisheries of the Spanish Government. Institutional support to CNAG was from the Spanish Ministry of Science and Innovation through the Instituto de Salud Carlos III and Generalitat de Catalunya through the Departament de Salut and the Departament de Recerca I Universitats.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 Manuzzi A *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Manuzzi A, Aguirre-Sarabia I, Díaz-Arce N *et al.* **Atlantic mackerel population structure does not support genetically distinct spawning components [version 2; peer review: 4 approved, 2 approved with reservations]** Open Research Europe 2025, **4**:82 https://doi.org/10.12688/openreseurope.17365.2

First published: 24 Apr 2024, **4**:82 https://doi.org/10.12688/openreseurope.17365.1

REVISED **Amendments from Version 1**

This new revision of the manuscript includes changes requested by reviewers. In particular, we have modified Figure 2 to specify the age stratum of each individual sample, added p values to Fst table and related text, included relevant references that were omitted in a previous version, added more details of the reagents and software (including versions) used, and modify a few sentences in the discussion section for improving clarity in our message.

Any further responses from the reviewers can be found at the end of the article

Introduction

The Atlantic mackerel, *Scomber scombrus* L., is a highly migratory coastal pelagic fish widely distributed throughout the northern Atlantic Ocean, from Labrador, Canada, to Cape Lookout, United States, in the west, and from Iceland and Norway to as far south as Mauritania in the east, as well as in the Mediterranean Sea (MED) and in the Black Sea (Iversen, 2002; Lockwood, 1988; Trenkel *et al.*, 2014). Within the northwest Atlantic (NWA), two spawning components have been recognized: a southern component, spawning along the US East coast in March/April, and a northern component spawning around the Gulf of St. Lawrence in June/July (Grégoire *et al.*, 2010; O'Brien *et al.*, 1993; Sette, 1950). Within the northeast Atlantic (NEA), three spawning components have been defined: a southern component, spanning at the Cantabrian Sea and Atlantic Iberian waters, a western component, including the Bay of Biscay, Celtic Seas and West of Scotland and a North Sea component, including the North Sea, Skagerrak and Kattegat, with spawning starting in February off Portugal, and ending in July north of Scotland. Summer northward feeding migrations occur on both sides of the Atlantic (Godø *et al.*, 2004; Holst & Iversen, 1992; Nesbø *et al.*, 2000; Radlinski *et al.*, 2013; Studholme, 1999; Uriarte & Lucio, 2001).

In the past two decades distribution of Atlantic mackerel has shifted poleward, increasing abundance around Iceland, Greenland, and the Faroe Islands (Nøttestad *et al.*, 2016). As a consequence of this this climate change-related expansion (Jansen *et al.*, 2016), the resulting disagreement over quotas escalated into the 'mackerel crisis' (Jensen *et al.*, 2015), so called for having exacerbated relations among the countries involved. This poleward expansion has been associated to changing water temperatures and food availability (Jansen *et al.*, 2012; Olafsdottir *et al.*, 2019; Overholtz *et al.*, 2011; Payne *et al.*, 2022), environmental conditions affecting interannual changes in post- and pre-spawning patterns and migration timing (Jansen & Gislason 2011; Trenkel *et al.*, 2014). In fact, a recent study forecasted a 370km northward shift in mackerel spawning distribution for every degree of sea temperature increase (Chust *et al.*, 2023).

The Atlantic mackerel holds significant economic importance for numerous countries harvesting the species (Jensen *et al.*, 2015); only in the NEA, catches of 1084 thousand tonnes (ICES, 2023) are reported for the species in the ICES areas, corresponding to approximately 1048 million € in sales value (mean yearly value per kg from EU sale; https://eumofa.eu/first-sale-monthly-data). In the NEA, ICES recognizes the existence of different spawning components within a unique stock. Similarly, the species is also considered part of a single stock in the NWA but, despite a collaborative effort, assessment and management are performed independently by the United States and Canada (DFO, 2021; NEFSC, 2018). Consequently, the Atlantic mackerel has been the subject of numerous studies aimed at understanding the connectivity within its range, especially regarding the migratory dynamics (spawning, feeding, overwintering) of the species that may contribute to population structure and require consideration in fishery assessment. Despite high dispersal rates suggesting potentially high connectivity throughout the species' range (Lockwood, 1988; Uriarte & Lucio, 2001; Uriarte *et al.*, 2001), several studies have confirmed the division between NWA and NEA mackerel (Gíslason *et al.*, 2020; Nesbø *et al.*, 2000; Rodriguez-Ezpeleta *et al.*, 2016), yet contrasting results still exist for both sides of the Atlantic on the existence of genetic differentiation linked to spawning contingents (Gíslason *et al.*, 2020; Jansen & Gislason, 2013; NEFSC, 2018; Uriarte *et al.*, 2001).

It has been hypothesized that the two NWA components show fidelity to their respective spawning sites (Studholme, 1999), and studies based on tagging and seasonal landing patterns further suggest that the individuals from the different components undergo distinct seasonal migrations (Beckett *et al.*, 1974; Sette, 1950). While one recent study based on otolith chemistry appears to have moderate discriminatory power (Arai *et al.*, 2021), other studies using meristic markers (Mackay & Garside, 1969), otolith shape (Castonguay *et al.*, 1991) as well as genetic markers from allozymes (Maguire *et al.*, 1987), mitochondrial DNA markers (Lambrey de Souza *et al.*, 2006) and microsatellites (Gíslason *et al.*, 2020) have failed to discriminate between spawning components. Nevertheless, a recent genetic study based on a higher number of genome-wide markers (single nucleotide polymorphisms; SNPs) has been able to perform individual genetic assignment to component of origin, despite the lack of a clear genetic population differentiation (Bourret *et al.*, 2023). Likewise, weak spawning site fidelity has been indicted for the NEA based on growth data (Jansen *et al.*, 2013) while additional studies in support of this hypothesis based on alternative methodologies (e.g. otoliths; Dawson (1991), and parasite infections; Somdal and Schram (1992)) were found inconclusive due to time-space biased sampling. Mixing of the components occurs, as demonstrated by tagging studies where the same mature individuals were detected in more than one component during spawning season (Uriarte *et al.*, 2001), pointing towards the hypothesis of a panmictic population (Jansen & Gislason, 2013). Originally, site fidelity was proposed based on the analyses of two mitochondrial DNA regions (Nesbø *et al.*, 2000). However, this analysis was based on relatively few individuals and it did not account for differences between year classes (Jansen & Gislason, 2013). Another genetic analysis based on microsatellites did not find any support for separation between spawning components (Gíslason *et al.*, 2020), suggesting lack of homing behaviour. This lack of a clear differentiation, coupled with the highly

migratory behaviour of the species, may indicate spawning in the NEA could occur within a single large spatiotemporal continuum, being close to panmictic but with a weak dynamic cline rather than distinct components (Jansen & Gislason, 2013).

Given the economic value of the Atlantic mackerel and the recent changes in its distribution, the need to resolve remaining doubts about its connectivity has become urgent. Genomic approaches can provide valuable information to answer the questions presented so far, even more so if a reference genome is used (Lu & Luo, 2020; Stapley *et al.*, 2010). Recently, there have been major developments in the production of accessible reference genomes (Fan *et al.*, 2020) as they are recognised as a key tool for improving conservation and management approaches (Hohenlohe *et al.*, 2021; Rhie *et al.*, 2021; Thorburn *et al.*, 2023). Indeed, the use of species-specific genomes might help to answer fundamental questions about genetic diversity, population structure or the genetic basis of adaptation (Fuentes-Pardo & Ruzzante, 2017; Whitacre *et al.*, 2022), for instance by identifying regions under selection due to environmental change (Merot *et al.*, 2023), and improves the feasibility of studies monitoring and predicting effects of climate change. This type of studies may be of great relevance for marine species, as marine diversity is particularly vulnerable to climate change (Cheung *et al.*, 2009) and have proven to be relevant for fisheries management and conservation (Ovenden *et al.*, 2015).

Here, we present a high-quality scaffold-level reference genome for Atlantic mackerel, as well as the first population genetic study using reference-based genome-wide markers and across most of the distribution range of the species. Single nucleotide polymorphisms (SNPs) were obtained from samples collected throughout the species whole range, including recent areas of expansion (e.g. Greenland), in order to: (1) assess fine-scale population structure and connectivity existing within the genetically distinct NWA, NEA and MED populations; (2) investigate the origin of the Greenlandic/Icelandic mackerel and possible adaptive markers associated with the recent northward migration; and finally (3) evaluate the possible link between spawning components and genetics by identifying markers that could be used to estimate components proportions within management units.

Methods

Tissue sampling and nucleotide extraction

S. scombrus larvae, juvenile as well as spawning and non-spawning and adult samples were collected from Northwest (NWA) and Northeast Atlantic Ocean (NEA) and Mediterranean Sea (MED) locations using scientific surveys and commercial fisheries (Supplementary Table S1 in *Extended data* (Manuzzi *et al.*, 2024); Figure 2.A). Larvae were collected using Bongo 40 or Gulf VII plankton nets and stored in 96% molecular grade ethanol at -20°C. Likewise, from each individual a ~1 cm^3 muscle tissue piece was excised and immediately stored in 96% molecular grade ethanol at -20°C until DNA extraction. Maturity of samples was assigned according to six scale categories (Walsh & Hopkins, 1990). For one female adult individual captured in the northeastern Atlantic (ICES Area VIIj2), muscle, gill, heart and gonad tissues were excised and stored in RNAlater™ (Ambion) until RNA extraction and was used for genome assembly and annotation. DNA was extracted from about 20 mg of muscle tissue or whole larvae using the Wizard® Genomic DNA Purification kit (cat # A1120, Promega, WI, USA) following the manufacturer's instructions. Extracted DNA was suspended in Milli-Q® IQ 7000 (Millipore) water, 20 µl in the case of larvae and 100 µl for tissues. Concentration was determined with the Quant-iT dsDNA HS assay kit using a Qubit® 2.0 Fluorometer (Life Technologies) and integrity by migrating about 100 ng of GelRed™-stained DNA on an agarose 1.0% gel. DNA used for genome sequencing was subjected to quality/purity control using the UV/Vis measurements of the Nanodrop™ 2000 (Thermo Fisher Scientific™), quantified using the Qubit DNA BR Assay kit (cat # Q32850, Thermo Fisher Scientific), and its integrity assessed with the Femto Pulse Genomic DNA 165 kb kit (Agilent). RNA was extracted with mirVana™ miRNA Isolation Kit (cat # AM1560, Invitrogen™). RNA was quantified by Qubit RNA BR Assay kit (Thermo Fisher Scientific) and its integrity was estimated by using RNA 6000 Nano Bioanalyzer 2100 Assay (Agilent).

Library preparation and sequencing

The *S. scombrus* genome assembly was produced using a combination of genomic and transcriptome long and short reads. Whole genome long-read sequencing libraries were prepared using the SQK-LSK110 1D sequencing kit from Oxford Nanopore Technologies™ (ONT). The sequencing run was performed on a PromethION™ 24 instrument (ONT) using a flow cell R9.4.1 FLO-PRO002 (ONT). Transcriptome long read sequencing libraries were prepared using the cDNA-PCR Sequencing Kit SQK-PCS111 (ONT), following the manufacturer's instructions. Sequencing runs were performed on GridION™ Mk1 (ONT) using a Flowcell R9.4.1 FLO-MIN106D (ONT). The quality parameters of the sequencing runs, for both long reads and cDNA, were monitored in real time using the MinKNOW™ platform (version 21.10.8 for long reads and version 22.05.7 for cDNA), and the base calling was performed using Guppy (ONT, version 6.2.7 and version 6.1.5 respectively, https://community.nanoporetech.com/downloads) "super-accuracy" mode ("dna_r9.4.1_450bps_sup.cfg" config file). The free alternative open-source high performance base calling software Dorado (https://github.com/nanoporetech/dorado) can also be used to base call the reads. Whole genome short-read sequencing Illumina platform compatible libraries were prepared using the PCR-free protocol from KAPA HyperPrep kit (Roche). Following end-repair and adenylation, Illumina platform-compatible adapters containing unique dual indexes and unique molecular identifiers (Integrated DNA Technologies) were ligated. The transcriptome short-read libraries were prepared with KAPA Stranded mRNA-Seq Illumina Platforms Kit (Roche) following the manufacturer´s recommendations starting with 500 ng of total RNA as the input material. The libraries were quality controlled on an Agilent 2100 Bioanalyzer with the DNA 7500 assay (Agilent) to assess size and quantified using the Kapa Library Quantification Kit for Illumina platforms (Roche). Whole genome and

transcriptome short-read libraries were sequenced on an Illumina NovaSeq™ 6000 with a read length of 2x150bp, following the manufacturer's protocol for dual indexing. Image analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA 3.4.4). Generated raw long and short genome and transcriptome paired-end reads are available at the European Nucleotide Archive (ENA; accession number: PRJEB70238).

Restriction-site-associated DNA (RAD) libraries were prepared following the methods of Etter *et al.* (2011). Namely, between 50 and 750 ng of starting DNA (depending on availability and integrity) was digested with the SdaI (SbfI) restriction enzyme (Thermo Fisher Scientific, cat # ER1191) in 20 µl reactions including 17 µl DNA, 1 µl Enzyme SdaI (SbfI) and 2 µl of 10X buffer. The mix was incubated 2 hours at 37°C, followed 20 minutes at 80°C to inactivate the enzyme. After digestion, Illumina-compatible P1 adapter, which contains 5bp sample-specific barcodes for samples demultiplexing, was ligated. The reaction was done in 30 µl using T4 DNA Ligase (Thermo Fisher Scientific, cat# EL0013), with 20 µl of digested DNA, 0.5 µl NEB Buffer 2 (10X, cat# B7002S)), 0.3 µl rATP (100mM, Thermo Scientific™ cat# R0441), 1 µl Ligase (30U/ µl, Thermo Scientific™ cat# EL0013), 2 µl adapter P1 (100nM) and water. The mix was incubated 16 hours at 22°C, followed 10 minutes at 65°C. Pools of 32 individuals were created and purified using Genejet PCR Purification kit (Thermo Fisher Scientific, cat# K0701). Each pool was sheared using the Covaris® M220 Focused-ultrasonicator™ Instrument (Life Technologies) with a default method "DNA_0500_bp_130_µl_Snap_Cap_microTUBE" (PowerPeak 50.0; Duty Factor 10.0; Cycles/Burst 200; Time 00h:01m:00s) and size selected to 300-500 bp by cutting agarose gel migrated DNA. The bands were extracted using the Genejet Purification kit (Thermo Fisher Scientific, cat# K0701) and eluting in 25.5 µl of Milli-Q water. Purified products were then End repaired (Thermo Fisher Scientific, cat# K0771) incubating at 20°C for 20 minutes the following mix: 25.5 µl of pool, 3 µl End Repair Reaction Mix (10X) and 1.5 µl End Repair Enzyme. Following, 3'-dA Overhang Addition was performed by incubating the end-repaired products at 37°C for 30 minutes (30 µl) with 5 µl of NEB Buffer 2 (10X), 1 µl of dATP (10mM, Thermo Fisher Scientific, cat# R0141) and 3 Klenow Fragment, exo– (5 U/µL, Thermo Fisher Scientific, cat# EP0422). Before amplification, Illumina different P2 adaptors (10 µM) were ligated to each Pool in the same conditions as the P1 ligation step.

Finally, each library was amplified in triplicates in volume of 25µl, using 5 µl of sample, 12.5 µl High Fidelity Phusion Master mix (Thermo Fisher Scientific, cat# F531S) and 1 µl of PCR primers (10 µM, PCR-For 5´-AATGATACGGCGACCACCGA-3´ and PCR-Rev 5´-CAAGCAGAAGACGGCATACGA-3´). The PCR profile was 98°C- 3minutes and 14 cycles of 98°C-15s; 65°C-30s;72°C-30s. Each replicate was mixed and purified in ratio 1:1.8 (sample/beads) with Axyprep MAG PCR Clean (Axygen®, cat # Mag-PCR-CL-50). Batches of three pools were paired end sequenced (100 bp) on an Illumina HiSeq2000.

Genome assembly and annotation

The genome assembly and annotation workflows are schematized in Supplementary Figures S1 and S2 in *Extended data* (Manuzzi *et al.*, 2024). Prior to assembly, the DNA Illumina short reads were trimmed for adaptors using TrimGalore v0.6.7 (https://github.com/FelixKrueger/TrimGalore) and the ONT long read genomic data was filtered to remove short and low-quality reads with Filtlong2 v0.2.1 (https://github.com/rrwick/Filtlong) with parameters: --minlen 700 --min_mean_q 80. Resulting ONT reads were assembled with Flye v2.9 (Kolmogorov *et al.*, 2019) using the 'nano-raw' mode and a minimum overlap of 1000. Illumina short reads and ONT were mapped against the generated assembly using BWA-mem v2.2.1 (Li & Durbin, 2009) and MINIMAP2 v2.24 (Li, 2018) respectively to assess contigs continuity and completeness as well as to identify possible base accuracy errors. The assembly was polished with HyPo v1.0.3 (Kundu *et al.*, 2019) to improve base accuracy using both Illumina and ONT data. Finally, the polished assembly was purged with purge_dups v1.2.6 (Guan *et al.*, 2020) to remove alternate haplotypes and other artificially duplicated repetitive regions. The consensus quality (QV) of the final assembly was estimated by Merqury v1.3 (Rhie *et al.*, 2020) and the gene completeness by BUSCO v5.4.0 (Manni *et al.*, 2021) using the odb10_actinopterygii database. QV represents the probability of error on a logarithmic scale for the consensus of the called bases. The higher the QV the higher the consensus, e.g. Q40 corresponds to a 99.99% accuracy (Rhie *et al.*, 2020). The histogram of the k-mer counting distribution was plotted in GenomeScope 2.0 v2.0 (Ranallo-Benavidez *et al.*, 2020).

The annotation of the assembly was obtained by combining transcript alignments, protein alignments and *ab initio* gene predictions. Repeats present in the genome assembly were annotated with RepeatMasker v4-1-2 (http://www.repeatmasker.org) using the custom repeat library available for *Danio rerio*. Moreover, a new repeat library specific for our assembly was made with RepeatModeler v1.0.11. After excluding those repeats that were part of repetitive protein families (performing a BLASTref search against Uniprot) from the resulting library, RepeatMasker was run again with this new library to annotate the specific repeats. The transcriptome short Illumina and long ONT reads were aligned to the genome using, respectively, STAR v-2.7.10a (Dobin *et al.*, 2013) and MINIMAP2 v2.24 (Li, 2018) with the splice option after which high-quality junctions to be used during the annotation process were obtained by running Portcullis v1.2.4 (Mapleson *et al.*, 2018). Transcript models were subsequently generated using Stringie v2.2.1 (Pertea *et al.*, 2015) on each BAM file and then all the models produced were combined using TACO v0.7.3 (Niknafs *et al.*, 2017). Finally, PASA assemblies were produced with PASA v2.5.2 (Haas *et al.*, 2008), and the *TransDecoder* program, which is part of the PASA package, was run on the PASA assemblies to detect coding regions in the transcripts. The complete proteomes of *Carassius auratus, Cynoglossus semilaevis, Danio rerio, Oryzias latipes, Parambassis ranga, Sparus aurata* and *Scopthalmus maximus* were downloaded from Uniprot in March 2022 and aligned to the

genome using Miniprot 0.6 (Li, 2023). *Ab initio* gene predictions were performed on the repeat-masked assembly with three different programs: GeneIDv1.4 (Alioto *et al.*, 2018), Augustus v3.5.0 (Stanke *et al.*, 2006) and Genemark-ET v4.71 (Lomsadze *et al.*, 2014) with and without incorporating evidence from the RNAseq data. The gene predictors were run with trained parameters for human, except Genemark, which runs in a self-trained mode. Finally, all the data were combined into consensus CDS models using EvidenceModeler-1.1.1 (EVM; Haas *et al.* (2008)). Additionally, UTRs and alternative splicing forms were annotated via two rounds of PASA annotation updates. Functional annotation was performed on the annotated proteins with Blast2go v1.3.3 (Conesa *et al.*, 2005) after a Blastp (Diamond Blastp v2.0.15, Buchfink *et al.*, 2021) search against the nr database (last accessed March 2023) and an Interproscan v5.55_88.0 (Jones *et al.*, 2014) run to detect protein domains on the annotated proteins. The annotation of non-coding RNAs was obtained by running cmsearch v1.1.4 (Cui *et al.*, 2016), that is part of the Infernal (Nawrocki & Eddy, 2013) package, against the RFAM database of RNA families v12.0 (Nawrocki & Eddy, 2013). Additionally, tRNAscan-SE v2.0.11 (Chan & Lowe, 2019) was run in order to detect the transfer RNA genes present in the masked genome assembly. Identification of long non-coding RNAs was done by first filtering the set of PASA-assemblies that had not been included in the annotation of protein-coding genes to retain those longer than 200bp and not covered more than 80% by a small ncRNA. The resulting transcripts were clustered into genes using shared splice sites or significant sequence overlap as criteria for designation as the same gene. BlobToolKit INSDC pipeline (Challis *et al.*, 2020) was run on the assembled genome using the NCBI nt database (updated on May 2023) and the following BUSCO odb10 databases: Actinopterygii, Vertebrata, Metazoa, Eukarya, Fungi and Bacteria. The genome assembly and annotation are available in the ENA under the accession number "GCA_963921475.1". Additionally, all the files resulting from the annotation process and a genome browser can be found in https://denovo.cnag.cat/fScoSco.

RAD-loci assembly and genotype table generation

Generated RAD-tags were analysed using *Stacks* version 2.4 (Catchen *et al.*, 2013). Quality filtering and demultiplexing were performed using the module *process_radtags* with default parameters, removing reads with adaptor sequences and truncating to 90 bp maximum length to further exclude low quality bases. Only reads whose forward and reverse pair passed quality filtering were kept and the module *clone_filter* was applied to remove PCR duplicates. Clean reads were aligned to the newly assembled Atlantic mackerel genome (see above) using BWA-mem 0.7.17 (Li & Durbin, 2009) with default parameters. Mapped reads were then deduplicated using Picard 2.0.1 MarkDuplicates (http://broadinstitute.github.io/picard/). Samples exceeding a 25% duplication level were excluded from downstream analysis. The module *gstacks* was used to assemble paired-end reads into contigs, merging them to the single-end loci and identifying and genotyping SNPs. The module *populations* was used to select the RAD-loci found in at least 90% of the individuals, sorted by reference order to ensure output is sorted according to the genome coordinates of the reference genome and possibly overlapping sites are not maintained (variant sites are printed only once for every genomic position). VCFtools v.0.1.13 (Danecek *et al.*, 2011) was used to select samples and SNPs with a minimum of 0.75 and 0.95 genotyping rate respectively and SNPs with minimum allele frequency (MAF) larger than 0.05. Kin pairs were detected through a genetic relatedness matrix estimated using VCFtools based on the values of relatedness calculated between all pairs of individuals using the Akj model (Yang *et al.*, 2010). Seven pairs involving overall 14 individuals with relatedness coefficients between 0.66 to 1 were found, of which individuals with the highest level of missing data from each pair were removed for downstream analysis. An additional pair of samples with 0.16 relatedness coefficient was also removed for being divergent with respect to the remaining pairwise comparisons (all < 0.08). After removal of these eight individuals, the module *populations* as well as downstream filtering steps were run again with the same parameters as above on the overall dataset as well as for each subset of individuals to be analysed.

Population structure analyses

Spatial structure inferences and neutrality tests were performed for the overall dataset (East and West Atlantic, and Mediterranean Sea; N=515, after filtering) as well as for the different spatial subsets (Atlantic only N=417, NWA N=150, NEA N=273). In addition, two further subsets were created including only reference samples, namely larvae and spawning adults (LSA), to investigate connectivity between spawning component (NEA-LSA N=170, NWA-LSA N=59; Supplementary Table S1 in *Extended data* (Manuzzi *et al.*, 2024). The genetic ancestry of each individual was estimated on the overall dataset filtered for physical linkage (1 SNP each 1kb window) using the model-based clustering method implemented in ADMIXTURE 1.3.0 (Alexander *et al.*, 2009) assuming from 2 to 18 ancestral populations (K) and setting 1,000 bootstrap runs. The linkage filter was performed by using a customized R script ("ld_pruning.r"; Manuzzi *et al.* (2024)) as suggested by ADMIXTURE manual since the software does not account for LD and thus it may produces false positives. The value of K with the lowest associated error value was identified using ADMIXTURE's cross-validation (CV) criterion. Principal Component Analyses (PCAs) were performed using the package *adegenet* 2.1.10 (Jombart & Ahmed, 2011) in R version 4.1.2 (Team, 2022) and visualized using *ggplot2* 3.5.1 (Wickham, 2016) and *factoextra* 1.0.7 (Kassambara & Mundt, 2020) packages. Estimates of pairwise F_{ST} between groups were performed using the R package StAMPP 1.6.3 (Pembleton *et al.*, 2013) after removing populations with reduced sample size (<7, GREEN N=2) and their significance was assessed with 1,000 permutations over loci. For all FST comparisons the p-values were adjusted using a False Discovery Rate (FDR) method. Loci potentially under selection were screened for the different spatial subsets (ATL, NEA, NWA) and for the subsets including only reference samples (NEA-LSA, NWA-LSA) (Supplementary Table S2; Manuzzi *et al.* (2024)) using the multivariate approach implemented in the R package *pcadapt* 4.3.3 (Luu *et al.*, 2017) setting alpha value to 0.05 as cut-off.

Results

Genome assembly and annotation

The Atlantic mackerel estimated genome size was of 769Mb with 0.68% heterozygosity. The raw assembly we obtained comprises a total of 765Mb in 9,496 contigs, while the final assembly (after polishing and purging), 741Mb in 4,482 contigs. The N50, namely the length of the shortest contig at 50% of the total assembly length, of the final assembly is 1.7 Mb and the L50, the smallest number of contigs required to make up half of genome size, is 110 with the longest contig measuring 10,208,092 bp. The consensus quality (QV) of the final assembly was estimated at 41.4 with 98.6% gene completeness reported. A statistical representation of the genome assembly is shown in Figure 1 in the form of a snail plot. In total, we annotated 26,428 protein-coding genes that produce 38,000 transcripts (1.44 transcripts per gene) and encode for 35,860 unique protein products. We were able to assign functional labels to 94.2% of the annotated proteins. The annotated transcripts contain 10.88 exons on average, with 93% of them being multi-exonic (Table 2). In addition, 7,871

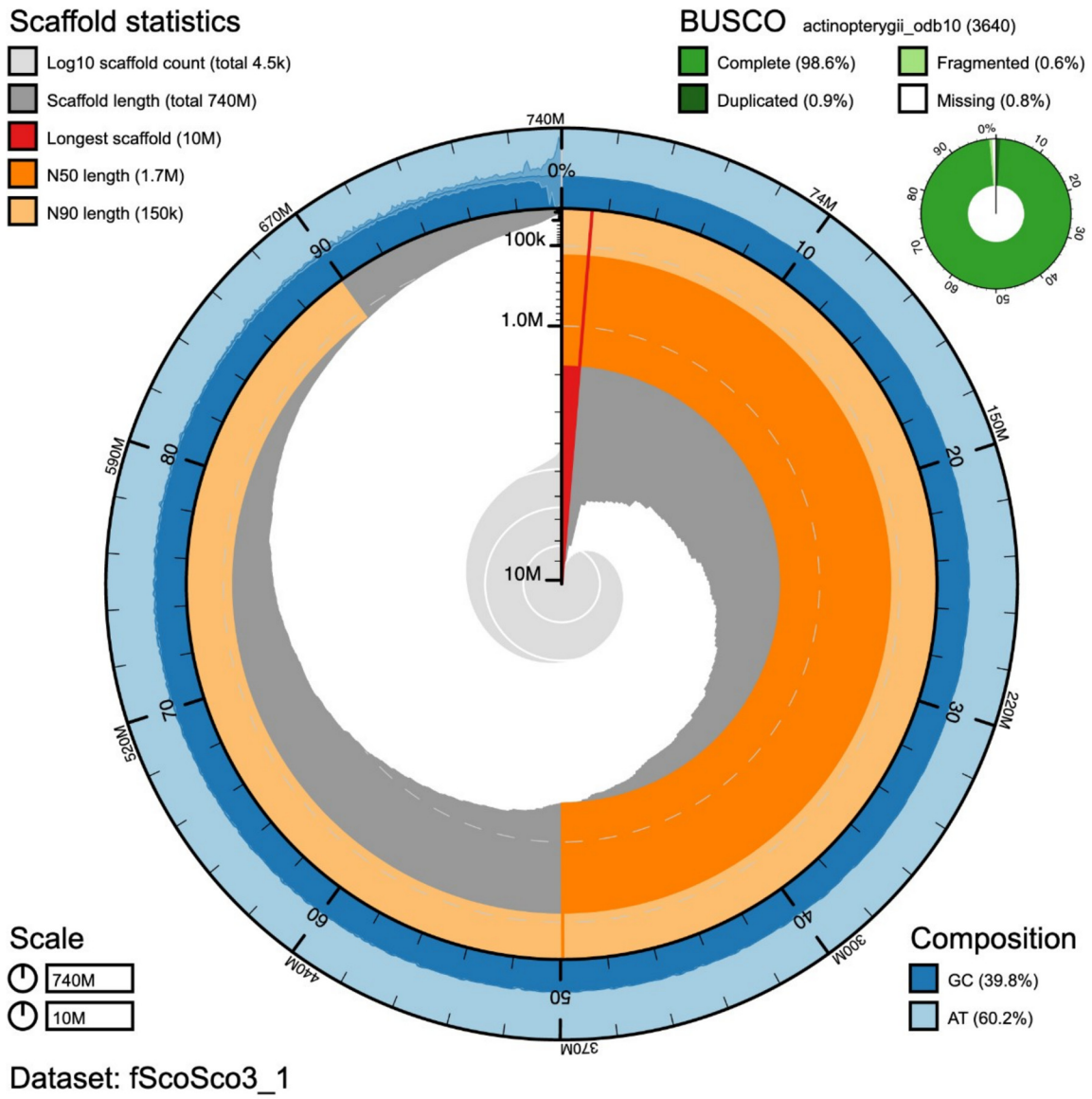

Figure 1. Snailplot summary of the assembly metrics for the *Scomber scombrus* genome. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 741,290,950 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (10,208,092 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (1,748,220 and 152,403 bp), respectively. The pale grey spiral shows the cumulative scaffold count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated and missing BUSCO genes in the actinopterygii_odb10 set is shown in the top right.

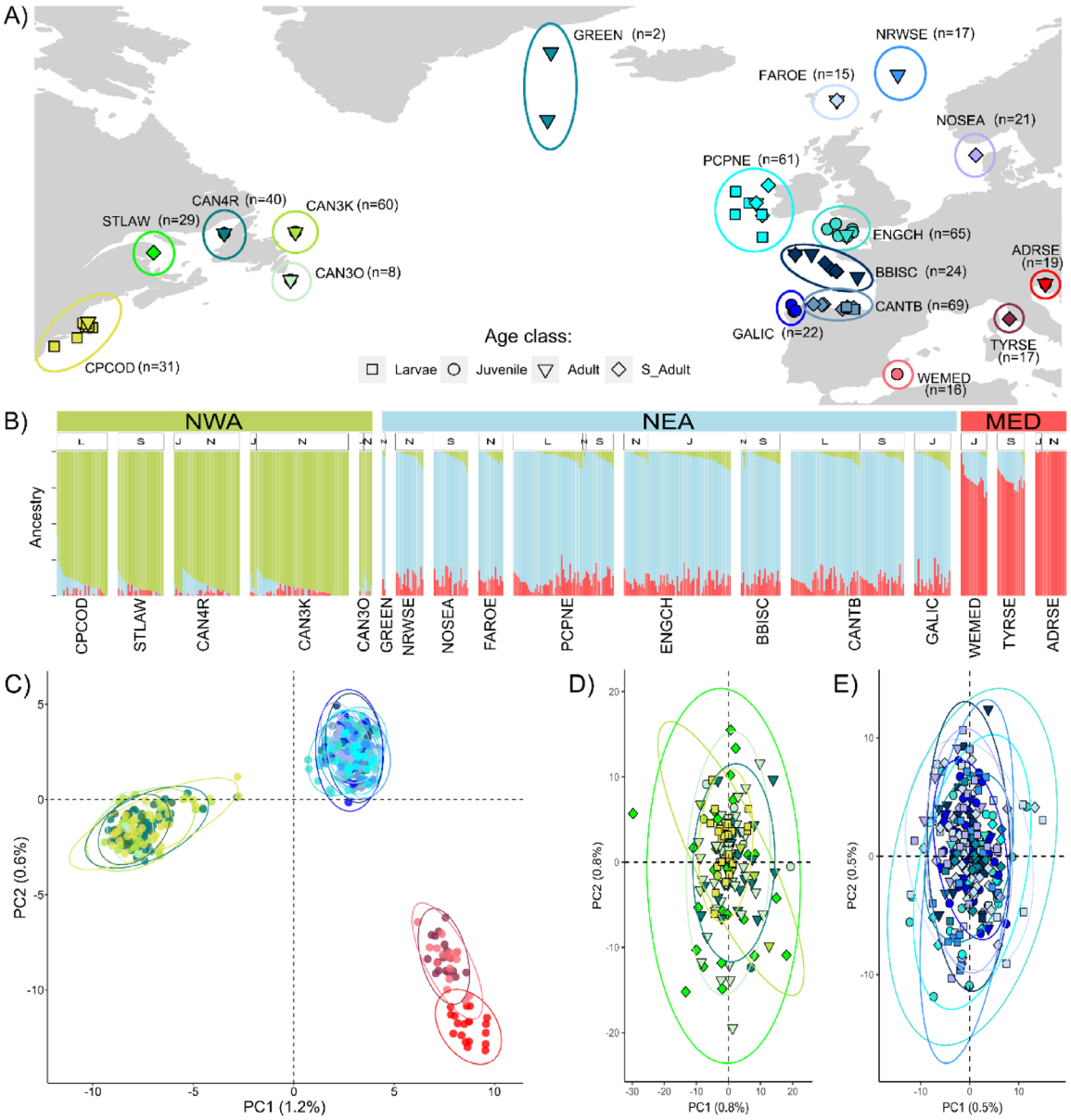

Figure 2. Population structure analyses. (**A**) Map of sampling collections included in the study. Each colour represents a sampling area. Names and sample sizes of each group are reported in the map. Different shapes are used to identify different life stages (larvae, juvenile, adult, and spawning adult). (**B**) Admixture results for best K=3. Levels of admixture are shown on the y-axis and are ordered by Q-values within each group and for each life stage. Life stages are coded on top of the admixture plot as: L=larvae, J=juvenile, N=adult not-spawning, S=spawning adult. In green, the western Atlantic ancestry component, in sky-blue, the eastern Atlantic, and in red the Mediterranean one as estimated by ADMIXTURE. (**C**) Principal component analyses (PCA) performed on individuals of the Atlantic mackerel based on the genomic dataset of 24,271 SNPs and 515 individuals in the overall dataset and for the different age classes within the NWA (**D**; 150 individuals, 39,541 SNPs) and NEA (**E**; 273 individuals, 39,528 SNPs).

non-coding transcripts were annotated, of which 6,063 and 1,808 are long and short non-coding RNA genes, respectively. Summarizing assembly and annotation statistics are reported in Table 1 and Table 2, respectively, and the results on the genome coverage distribution are reported in Supplementary Figure S3 (Manuzzi *et al.*, 2024) and show two 20-mer peaks of coverage at 33 (1n, heterozygous) and at 66 (2n, homozygous).

RAD-loci assembly and genotype table generation

After sequencing, we obtained an average of 6,108,766 raw reads per sample (minimum=411,982; maximum=26,351,634).

Table 1. Genome assembly statistics. Results obtained after polishing steps performed using HyPo4 with both Illumina and ONT data, and purging using purge_dups5 to remove alternate haplotype and duplicated repetitive regions. The final assembly comprised of 741Mb and 4,483 contigs scaffolds.

Assembly	Flye + hypo + purged
N50	1,748,220 bp
L50	110
Total sequences	4,482
Total length	741,290,950 bp
BUSCO* complete	98.6%
BUSCO* duplicated	0.9%
QV	41.4058
Kmer completeness	88.873

Table 2. Genome annotation statistics. The gene annotation of the mackerel genome assembly was obtained by combining transcript alignments, protein alignments and ab initio gene predictions.

	Annotation
Number of protein-coding genes	26,428
Median gene length (bp)	7,720
Number of transcripts	38,000
Number of exons	278,035
Number of coding exons	264,221
Median UTR length (bp)	1,142
Median intron length (bp)	431
Exons/transcript	10.88
Transcripts/gene	1.44
Multi-exonic transcripts	93%
Gene density (gene/Mb)	35.65

Following mapping, 60 specimens were excluded from downstream analyses because of > 50% duplicated reads, < 20% of properly paired reads and/or median insert size <10bp. For the remaining individuals, an average of 99.14% of the filtered reads aligned to the reference genome. After SNPs calling, eight additional samples were removed due to high relatedness values (Methods) and further 23 due to missing data >25%. Finally, seven genotype tables were generated with a variable number of samples (59–515) and SNP markers (10,888–39,541) depending on the purpose of the analysis (Supplementary Table S2; Manuzzi *et al.* (2024)) and with an average SNP coverage, calculated over all samples for each dataset, of 16X–23X.

Population structure analyses

Based on the results of the ADMXITURE test, K=2–3 were equally likely numbers of genetic clusters within our collection, corresponding to the lowest cross-validation errors (CV; Supplementary Figure S4.A; Manuzzi *et al.* (2024)). From the ADMIXTURE plot, the eastern Atlantic appears to more likely have shared ancestral origin with the Mediterranean samples, with the greatest difference existing between the Mediterranean and the western Atlantic collections (K=2, Supplementary Figure S4.B; Manuzzi *et al.* (2024)). Increasing the number of ancestral populations (K=3, Figure 2.B) showed three clusters with distinct genetic ancestry, corresponding to the geographical separations, namely NWA (green), NEA (sky-blue) and MED (red). No differentiation was detected within each cluster, and the three lineages coinciding with the geographical division described the full extent of the structure with individual samples showing high ancestry proportion (Q) within each respective cluster. Individual ancestry proportions do not appear to be correlated with the latitudinal gradients nor with the different age classes (L=larvae, J=juveniles, N=non-reproductive adults and S=reproductive adults) visible in the upper bar of the ADMXITURE plot. This geographical separation is confirmed by the results of the PCA (Figure 2.C), where a clear differentiation between the same three genetically separated groups is apparent, with the first and second components of the PCA explaining 1.2% and 0.6% of the variance respectively. All pairwise comparisons between the three groups resulted in high F_{ST} values (p-values=0.000), with the highest value reported for MED vs NWA (F_{ST}=0.035), as expected given their geographical distance, and the lowest between NEA and NWA (F_{ST}=0.014), closely followed by the NEA vs MED comparison (F_{ST}=0.019). However, beyond the NWA-NEA Atlantic divide, the PCAs further confirm the lack of further substructure within the respective Atlantic basins (NWA Figure 2.D; NEA Figure 2.E), where the variance is minimal and explained in equal proportions by the first two axes in both within-basin analyses (NWA, PC1=PC2=0.8%; NEA, PC1=PC2=0.5%). Pairwise F_{ST} values ranged from 0.000 to 0.001 within the NEA even for the more spatially distant areas, and from 0.000 to 0.003 within NWA (Supplementary Tables S3–S4; Manuzzi *et al.* (2024)). In the Mediterranean Sea, instead, local structure is discernible in both the ADMIXTURE and the PCA analyses where the Adriatic Sea appears genetically distinct from the remaining western Mediterranean collections and is composed of genotypes of pure Mediterranean ancestry. This differentiation within the Mediterranean Sea is also seen in the F_{ST} pairwise comparisons, with significant (p-values=0.000) F_{ST} values of 0.003 and 0.004 for the Adriatic Sea versus the Tyrrhenian Sea and Western Mediterranean Sea

respectively. Finally, all population structure analyses clearly cluster the Greenlandic samples with the NEA (Figure 2.B-C-D) confirming the eastern Atlantic origin of the samples currently harvested in the relatively recent northward migration ground. Although outlier markers were identified for the Atlantic (275 SNPs), Northwest Atlantic (65) and Northeast Atlantic (38), lack of sub structuring was maintained when analysing only these markers (Supplementary Figure S5; Manuzzi *et al.* (2024)). Additionally, even when including only reference samples, no evidence of a genetically based differentiation was identified neither by neutral nor putative outlier markers (Supplementary Figure S6; Manuzzi *et al.* (2024)), with significant (p-value=0.000) F_{ST} values of 0.0004 and 0.001 between larvae and spawning adults within the NEA and NWA, respectively.

Discussion

Here, we present a reference genome for the Atlantic mackerel, as well as the most comprehensive SNP dataset and most broadly distributed samples used to resolve the questions left unanswered regarding the population structure and connectivity of this species.

Three main genetically isolated populations in Atlantic mackerel

Our results confirmed the presence of isolated populations between the two sides of the North Atlantic Ocean, as well as between the Atlantic and the Mediterranean Sea with no or reduced genetic exchange between the three distinct groups, which settles previous studies (Bourret *et al.*, 2023; Gíslason *et al.*, 2020; Nesbø *et al.*, 2000; Rodriguez-Ezpeleta *et al.*, 2016; Zardoya *et al.*, 2004). Moreover, we confirm that mackerel found in Greenland are part of the NEA population (Bourret *et al.*, 2023; Gíslason *et al.*, 2020). Despite inter-population F_{ST} values being low (F_{ST}=0.014–0.035), these are in the range of what is usually observed in marine fish species with large population sizes since F_{ST} estimates are affected by both effective population sizes and migration rates, which may derive into insufficient statistical power provided by neutral markers to reject the panmictic hypothesis (Waples, 1998). Furthermore, low F_{ST} values do not always indicate high migration rates between populations and may also reflect the presence of a large metapopulation with high effective population size and low migration rates (Gagnaire *et al.*, 2015); further studies are required to decipher these dynamics.

Thus, despite the use of a large dataset with tens of thousands of SNP markers and a reference genome, no further signs of population structure were detected within the NEA or the NWA. Only in the Mediterranean Sea, differentiation between the Adriatic Sea and the western Mediterranean was found, confirming previous studies (Rodriguez-Ezpeleta *et al.*, 2016). This supports the hypothesis of high mixing and large effective population size.

A single population within the Northeast Atlantic mackerel: implications for management

The lack of fine-scale genetic structure allowed us to reject the hypothesis of strict homing behaviour genetically isolating the NEA mackerel spawning areas. Under this hypothetical scenario, genetic differentiation would be maintained despite the demographic connectivity supported by the migratory behaviour as movement would not derive into gene flow (Dittman & Quinn, 1996). Unlike the NWA (Bourret *et al.*, 2023), none of the analyses conducted to date trying to disentangle the population structure of NEA mackerel have been based on thousands of genome-wide markers, nor have been successful in providing compelling evidence for the presence or absence of genetically isolated spawning components (Nesbø *et al.*, 2000). Our large dataset, both in number of individuals and markers, settles doubts on the population structure of Atlantic mackerel in the NEA, and reject the hypothesis of homing behaviour. Indeed, our analyses find no evidence of genetically isolated spawning components even when using only larvae and adults, which are supposed to represent the spawning population at one location, or outlier markers, that is, those potentially under selection. These results do not support a mixed stock fishery scenario as that found in other species (e.g. Atlantic bluefin tuna; Rodríguez-Ezpeleta *et al.* (2019)), which would require the need for developing a genetic stock identification method. Indeed, our study supports the use of a single stock for NEA mackerel assessment, which is the approach currently used.

Yet, it should be noted that our work does not include samples from southern Europe (e.g. south of Portugal) which would allow to further investigate the connectivity between southern NEA and and the North Atlantic African coast and for which no conclusion can be drawn. Additionally, we cannot exclude the possibility that our dataset fails to include a reduced number of genetic markers that might be involved in differentiating contingents, that have not been targeted by our RAD sequencing approach. This has been seen, for example, in the Atlantic herring (Han *et al.*, 2020), for which early genetic studies were unable to disentangle population differentiation, nor to discern adaptive markers from the neutral background, due to the high levels of gene flow and large effective population sizes. However, once whole genome sequencing (WGS) data were used, differences associated with a reduced number of genomic regions were detected in both Atlantic herring and Atlantic horse mackerel. For the Atlantic herring (Han *et al.*, 2020), fewer regions were found to be associated with an adaptation to seasonal reproduction allowing the differentiation of reproductive components despite the overall weak genomic differentiation, whereas in the Atlantic horse mackerel (Fuentes-Pardo *et al.*, 2023) fewer regions under selection revealed geographic differences associated with environmental variables.

Forecasting changes

The highly migratory behaviour, coupled with the wide species distribution, entails that the Atlantic mackerel can occupy several different environments during its life, suggesting that the species may be adapted to a wide range of temperatures. The lack of differentiation within the NEA supports the hypothesis of a high dispersal capacity of the species, further confirmed by the Greenland samples that appear to have originated from a north-westward migration of NEA

mackerel (Jansen *et al.*, 2016). In addition, the species shows high plasticity as it can tolerate up to 5–15°C (Olafsdottir *et al.*, 2019) despite its optimal temperature range (9–13°C), which is possibly linked to the high genomic heterozygosity detected at the species level. Here, our findings do not show signs of selection associated to this shift, and no putative outlier markers were detected. These results, together with observed northward shift of distribution (Olafsdottir *et al.*, 2019) and spawning area (Chust *et al.*, 2023; dos Santos Schmidt *et al.*, 2023) linked to water temperature variations suggest a future displacement of NEA mackerel under climate scenarios of a progressively warmer Atlantic with forecasted spawning distribution shifts both westwards and northwards (Bruge *et al.*, 2016). Despite the low sample size, the Greenlandic samples in our collection are clearly part of the NEA population, and no connectivity with NWA was detected here. However, if the species' range keeps expanding it cannot be excluded that a bridge in connectivity at northern latitudes may happen, since mackerel from the NWA population have been found as far north as Labrador (Parsons, 1970). Indeed, warmer climate creating thermal niches suitable for expansion could bring isolated populations closer thus reducing the existing barrier to gene flow and allowing for secondary contact to occur. This is usually referred to as an interspecific contact, however it could occur between naturally isolated populations of the same species (intraspecific). Reported cases of recent interspecific secondary contact are so far limited (e.g. hybridization by human-driven translocation of *Ciona robusta* to its congeners' distribution areas, Le Moan *et al.* (2021)), especially as a result of recent climate change (e.g. hybridization due to recent distribution shift causing contact during spawning events *Argyrosomus coronus*, Potts *et al.* (2014)) and none is reported for intraspecific contact due to the disappearance of a previously existing barrier. Indeed, population shifts and the creation of contact zones between populations and/or species usually occur over geological timescales, but this could change and accelerate with the current acceleration and magnitude of ocean warming (Pinsky *et al.*, 2020).

Future Atlantic mackerel genomic studies

The availability of a species reference genome could help to further investigate the presence of distinct evolutionary significant units for management linked to adaptive potential, thereby identifying associated gene families, and to assess the demographic history of identified areas of differentiation (e.g. the one that caused the Adriatic split). In recent years, high-throughput sequencing has made it more cost-effective to produce reference genomes even for non-model organisms. Here, we have used three generally accepted metrics to assess genome quality: contiguity, completeness, and correctness (Lewin *et al.*, 2018; Wang & Wang, 2023) and found that the Atlantic mackerel genome produced here follows high quality standards. This resource will be highly valuable for future studies on this species, such as those based on whole genome resequencing data, which may reveal further complexity not detected when using reduced representation datasets. The evidence for the northward expansion of the distribution range of the species, and potential secondary contact between NEA and NWA, highlight the need for future genetic monitoring of this species, especially on newly occupied areas, to detect and estimate potential population mixing (Schwartz *et al.*, 2007).

Ethics and consent

Ethical approval and consent were not required.

Data availability

Underlying data

The data used in this study are publicly available at:

Sequence Read Archive (SRA, NCBI): Demultiplexed RAD sequences of Atlantic mackerel included in this study. Accession number: PRJNA1081273; https://www.ncbi.nlm.nih.gov/bioproject/1081273 (Manuzzi *et al.*, 2024).

European Nucleotide Archive (ENA): Long and short genome and transcriptome paired-end reads used for the genome assembly and annotation. Accession number: PRJEB70238; https://www.ebi.ac.uk/ena/browser/view/PRJEB70238 (Manuzzi *et al.*, 2024).

Extended data

Zenodo: Data from: Atlantic mackerel population structure does not support genetically distinct spawning components. https://zenodo.org/doi/10.5281/zenodo.10684820 (Manuzzi *et al.*, 2024).

This project contains the following supplementary data:

- "Table S1.xlsx" including all metadata for the samples included in the population structure and adaptation analyses performed in the manuscript. The table contains 20 columns: AnalisisID, Barcode, Pool, Area, Region, Pop, Age, Size(mm), Sex, Maturity stage, Latitude, Longitude, Collection date, and each dataset tested reporting the list of samples included.

- "Extended_data.docx" containing Supplementary Tables and Figures. In term of tables, it includes: (1) Table S2: Table reporting the samples' numbers, SNPs' number and filtering steps of each genotype tables (Dataset) used. (2) Table S3. . Pairwise FST (lower diagonal) and respective p-values (upper diagonal) between the different populations within the NEA. (3) Table S4. Pairwise FST (lower diagonal) and respective p-values (upper diagonal) between the different populations within the NWA. Additionally, it includes (5) Figure S1. Workflow of the genome assembly process, (6) Figure S2. Workflow of the genome annotation process, (7) Figure S3. Genoscope transformed linear plot, (8) Figure S4. Cross-validation for ADMIXTURE test, (9) Figure S5. PCA plots of the outlier SNPs detected by pcadapt and (10) Figure S6. PCA plots of neutral and outlier SNPs detected using reference samples from NEA and NWA

- Scripts used to perform the analyses described in this manuscript.

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Authors contributions

N.R-E. and P. A. conceived the study and obtained the funding. N.R-E. and A.M. wrote the manuscript; J.G.G., T.S.A., M.G. generated and sequenced the genome libraries and assembled the genome. A.M. conducted the population genomics analyses with help from I.A-S., N.D-A. and N.R-E. D.B., T.J, and M.C. contributed samples. A. M., N. D-A., D. B., T. J., S.S-M., P. A. and N. R-E. interpreted results. All authors commented and contributed to the final version of the manuscript.

Acknowledgements

We wish to thank Iñaki Mendibil for excellent technical assistance and David Richardson (NOAA), Pablo Carrera (IEO), Iñigo Krug (AZTI), Høgni Debes (Faroe Marine Research Institute), Marte Schei (SINTEF), Steve Milligan, Jeroen Van Der Kooij and Joana Silva (CEFAS), and Walter Golet (University of Maine) for their assistance in providing the samples included in this study.

References

Alexander DH, Novembre J, Lange K: **Fast model-based estimation of ancestry in unrelated individuals.** *Genome Res.* 2009; **19**(9): 1655–1664.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Alioto T, Blanco E, Parra G, *et al.*: **Using geneid to identify genes.** *Curr Protoc Bioinformatics.* 2018; **64**(1): e56.
PubMed Abstract | **Publisher Full Text**

Arai K, Castonguay M, Secor DH: **Multi-decadal trends in contingent mixing of Atlantic mackerel (*Scomber scombrus*) in the Northwest Atlantic from otolith stable isotopes.** *Sci Rep.* 2021; **11**(1): 6667.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Beckett JS, Stobo WT, Dickson CA: **Southwesterly migration of Atlantic mackerel, *Scomber scombrus*, tagged off Nova Scotia.** International Commission for The Northwest Atlantic Fisheries. 1974.
Reference Source

Bourret A, Smith A, Van Beveren E, *et al.*: **Quantifying genetic differentiation and population assignment between two contingents of Atlantic mackerel (*Scomber scombrus*) in the Northwest Atlantic.** *Can J Fish Aquat Sci.* 2023; **80**(7): 1084–1097.
Publisher Full Text

Bruge A, Alvarez P, Fontán A, *et al.*: **Thermal Niche tracking and future distribution of Atlantic mackerel spawning in response to ocean warming.** *Front Mar Sci.* 2016; **3**.
Publisher Full Text

Buchfink B, Reuter K, Drost HG, *et al.*: **Sensitive protein alignments at tree-of-life scale using DIAMOND.** *Nat Methods.* 2021; **18**(4): 366–368.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Castonguay M, Simard P, Gagnon P: **Usefulness of Fourier analysis of otolith shape for Atlantic mackerel (*Scomber scombrus*) stock discrimination.** *Can J Fish Aquat Sci.* 1991; **48**(2): 296–302.
Publisher Full Text

Catchen J, Hohenlohe PA, Bassham S, *et al.*: **Stacks: an analysis tool set for population genomics.** *Mol Ecol.* 2013; **22**(11): 3124–3140.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Challis R, Richards E, Rajan J, *et al.*: **BlobToolKit – Interactive quality assessment of genome assemblies.** *G3 (Bethesda).* 2020; **10**(4): 1361–1374.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Chan PP, Lowe TM: **tRNAscan-SE: searching for tRNA genes in genomic sequences.** *Methods Mol Biol.* 2019; **1962**: 1–14.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Cheung WWL, Lam VWY, Sarmiento JL, *et al.*: **Projecting global marine biodiversity impacts under climate change scenarios.** *Fish Fish.* 2009; **10**(3): 235–251.
Publisher Full Text

Chust G, Taboada FG, Alvarez P, *et al.*: **Species acclimatization pathways: latitudinal shifts and timing adjustments to track ocean warming.** *Ecol Indic.* 2023; **146**: 109752.
Publisher Full Text

Conesa A, Götz S, García-Gómez JM, *et al.*: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics.* 2005; **21**(18): 3674–3676.
PubMed Abstract | **Publisher Full Text**

Cui X, Lu Z, Wang S, *et al.*: **CMsearch: simultaneous exploration of protein sequence space and structure space improves not only protein homology detection but also protein structure prediction.** *Bioinformatics.* 2016; **32**(12): i332–i340.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Danecek P, Auton A, Abecasis G, *et al.*: **The variant call format and VCFtools.** *Bioinformatics.* 2011; **27**(15): 2156–2158.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Dawson WA: **Otolith measurement as a method of identifying factors affecting first-year growth and stock separation of mackerel (*Scomber scombrus* L.).** *ICES J Mar Sci.* 1991; **47**(3): 303–317.
Publisher Full Text

DFO: **Assessment of the northern contingent of Atlantic mackerel (*Scomber scombrus*) in 2020.** In: *Canadian Science Advisory Secretariat Science Advisory Report.* 2021.
Reference Source

Dittman A, Quinn T: **Homing in Pacific salmon: mechanisms and ecological basis.** *J Exp Biol.* 1996; **199**(Pt 1): 83–91.
PubMed Abstract | **Publisher Full Text**

Dobin A, Davis CA, Schlesinger F, *et al.*: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics.* 2013; **29**(1): 15–21.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

dos Santos Schmidt TC, Slotte A, Olafsdottir AH, *et al.*: **Poleward spawning of Atlantic mackerel (*Scomber scombrus*) is facilitated by ocean warming but triggered by energetic constraints.** *ICES J Mar Sci.* 2023; **81**(3): 600–615, fsad098.
Publisher Full Text

Etter PD, Bassham S, Hohenlohe PA, *et al.*: **SNP discovery and genotyping for evolutionary genetics using RAD sequencing.** *Methods Mol Biol.* 2011; **772**: 157–178.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Fan G, Song Y, Yang L, *et al.*: **Initial data release and announcement of the 10,000 fish genomes project (Fish10K).** *GigaScience.* 2020; **9**(8): giaa080.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Fuentes-Pardo AP, Farrell ED, Pettersson ME, *et al.*: **The genomic basis and environmental correlates of local adaptation in the Atlantic horse mackerel (*Trachurus trachurus*).** *Evol Appl.* 2023; **16**(6): 1201–1219.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Fuentes-Pardo AP, Ruzzante DE: **Whole-genome sequencing approaches for conservation biology: advantages, limitations and practical recommendations.** *Mol Ecol.* 2017; **26**(20): 5369–5406.
PubMed Abstract | **Publisher Full Text**

Gagnaire PA, Broquet T, Aurelle D, *et al.*: **Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era.** *Evol Appl.* 2015; **8**(8): 769–786.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Gíslason D, Helyar SJ, Óskarsson GJ, *et al.*: **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.* 2020; **77**(2): 604–612.
Publisher Full Text

Godø OR, Hjellvik V, Iversen SA, *et al.*: **Behaviour of mackerel schools during summer feeding migration in the Norwegian Sea, as observed from fishing vessel sonars.** *ICES J Mar Sci.* 2004; **61**(7): 1093–1099.
Publisher Full Text

Grégoire F, Castonguay M, Maguire JJ, *et al.*: **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, 2010.
Reference Source

Guan D, McCarthy SA, Wood J, *et al.*: **Identifying and removing haplotypic duplication in primary genome assemblies.** *Bioinformatics.* 2020; **36**(9): 2896–2898.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Haas BJ, Salzberg SL, Zhu W, *et al.*: **Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced**

alignments. *Genome Biol.* 2008; **9**(1): R7.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Han F, Jamsandekar M, Pettersson ME, *et al.*: **Ecological adaptation in Atlantic herring is associated with large shifts in allele frequencies at hundreds of loci.** *eLife.* 2020; **9**: e61076.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Hohenlohe PA, Funk WC, Rajora OP: **Population genomics for wildlife conservation and management.** *Mol Ecol.* 2021; **30**(1): 62–82.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Holst JC, Iversen SA: **Distribution of Norwegian spring-spawning herring and mackerel in the Norwegian Sea in late summer, 1991.** In: *ICES Council Meeting 1992/H:* 1992; **13**: 1–13.
Reference Source

ICES: **Mackerel (*Scomber scombrus*) in subareas 1-8 and 14, and in Division 9.a (Northeast Atlantic and adjacent waters)**. In: *ICES Advice: Recurrent Advice*. 2023.

Iversen SA: **Changes in the perception of the migration pattern of Northeast Atlantic mackerel during the last 100 years.** *ICES Mar Sc.* 2002; **215**: 382–390.
Reference Source

Jansen T, Campbell A, Brunel T, *et al.*: **Spatial segregation within the spawning migration of North Eastern Atlantic mackerel (*Scomber scombrus*) as indicated by juvenile growth patterns.** *PLoS One.* 2013; **8**(2): e58114.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Jansen T, Campbell A, Kelly C, *et al.*: **Migration and fisheries of north east Atlantic mackerel (*Scomber scombrus*) in Autumn and Winter.** *PLoS One.* 2012; **7**(12): e51541.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Jansen T, Gislason H: **Temperature affects the timing of spawning and migration of North Sea mackerel.** *Cont Shelf Res.* 2011; **31**(1): 64–72.
Publisher Full Text

Jansen T, Gislason H: **Population structure of Atlantic mackerel (*Scomber scombrus*).** *PLoS One.* 2013; **8**(5): e64744.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Jansen T, Post S, Kristiansen T, *et al.*: **Ocean warming expands habitat of a rich natural resource and benefits a national economy.** *Ecol Appl.* 2016; **26**(7): 2021–2032.
PubMed Abstract | **Publisher Full Text**

Jensen F, Frost H, Thøgersen T, *et al.*: **Game theory and fish wars: the case of the Northeast Atlantic mackerel fishery.** *Fish Res.* 2015; **172**: 7–16.
Publisher Full Text

Jombart T, Ahmed I: ***adegenet 1.3-1*: new tools for the analysis of genome-wide SNP data.** *Bioinformatics.* 2011; **27**(21): 3070–3071.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Jones P, Binns D, Chang HY, *et al.*: **InterProScan 5: genome-scale protein function classification.** *Bioinformatics.* 2014; **30**(9): 1236–1240.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kassambara A, Mundt F: **factoextra: extract and visualize the results of multivariate data analyses (R package version 1.0.7)**. 2020.
Reference Source

Kolmogorov M, Yuan J, Lin Y, *et al.*: **Assembly of long, error-prone reads using repeat graphs.** *Nat Biotechnol.* 2019; **37**(5): 540–546.
PubMed Abstract | **Publisher Full Text**

Kundu R, Casey J, Sung WK: **HyPo: super fast & accurate polisher for long read genome assemblies.** *bioRxiv.* 2019; 2019.2012.2019.882506.
Publisher Full Text

Lambrey de Souza J, Sévigny J, Chanut JP, *et al.*: **High genetic variability in the mtDNA control region of a Northwestern Atlantic teleost, *Scomber scombrus* L.** *Can Tech Rep Fish Aquat Sci.* 2006; 33.
Reference Source

Le Moan A, Roby C, Fraïsse C, *et al.*: **An introgression breakthrough left by an anthropogenic contact between two ascidians.** *Mol Ecol.* 2021; **30**(24): 6718–6732.
PubMed Abstract | **Publisher Full Text**

Lewin HA, Robinson GE, Kress WJ, *et al.*: **Earth BioGenome Project: sequencing life for the future of life.** *Proc Natl Acad Sci U S A.* 2018; **115**(17): 4325–4333.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Li H: **Minimap2: pairwise alignment for nucleotide sequences.** *Bioinformatics.* 2018; **34**(18): 3094–3100.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Li H: **Protein-to-genome alignment with miniprot.** *Bioinformatics.* 2023; **39**(1): btad014.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.** *Bioinformatics.* 2009; **25**(14): 1754–1760.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Lockwood SJ: **The mackerel, its biology, assessment and the management of a fishery**. Farnham, Surrey, England: Fishing News Books Ltd, 1988.
Reference Source

Lomsadze A, Burns PD, Borodovsky M: **Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm.** *Nucleic Acids Res.* 2014; **42**(15): e119.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Lu G, Luo M: **Genomes of major fishes in world fisheries and aquaculture: status, application and perspective.** *Aquac Fish.* 2020; **5**(4): 163–173.
Publisher Full Text

Luu K, Bazin E, Blum MGB: ***pcadapt*: an R package to perform genome scans for selection based on principal component analysis.** *Mol Ecol Resour.* 2017; **17**(1): 67–77.
PubMed Abstract | **Publisher Full Text**

Mackay KT, Garside ET: **Meristic Analyses of Atlantic Mackerel, *Scomber scombrus*, from the North American Coastal Populations.** *J Fish Res Board Can.* 1969; **26**(9): 2537–2540.

Maguire JJ, Chagnon YC, Castonguay M, *et al.*: **A review of mackerel management areas in the Northwest Atlantic**. In: *CAFSAC Res.* Doc.87/71, 1987; 32.
Reference Source

Manni M, Berkeley MR, Seppey M, *et al.*: **BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes.** *Mol Biol Evol.* 2021; **38**(10): 4647–4654.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Manuzzi A, Aguirre-Sarabia I, Díaz-Arce N, *et al.*: **Data from: Atlantic mackerel population structure does not support genetically distinct spawning components.** [Dataset]. In: Zenodo. 2024.
http://www.doi.org/10.5281/zenodo.10684820

Mapleson D, Venturini L, Kaithakottil G, *et al.*: **Efficient and accurate detection of splice junctions from RNA-seq with Portcullis.** *GigaScience.* 2018; **7**(12): giy131.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Mérot C, Stenløkk KSR, Venney C, *et al.*: **Genome assembly, structural variants, and genetic differentiation between lake whitefish young species pairs (*Coregonus* sp.) with long and short reads.** *Mol Ecol.* 2023; **32**(6): 1458–1477.
PubMed Abstract | **Publisher Full Text**

Nawrocki EP, Eddy SR: **Infernal 1.1: 100-fold faster RNA homology searches.** *Bioinformatics.* 2013; **29**(22): 2933–2935.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

NEFSC: **64th Northeast regional stock assessment workshop (64th SAW) assessment report.** In: U.S. Department of Commerce, Northeast Fisheries Science Center Reference Document, 2018; 529.
Reference Source

Nesbø CL, Rueness EK, Iversen SA, *et al.*: **Phylogeography and population history of Atlantic mackerel (*Scomber scombrus* L.): a genealogical approach reveals genetic structuring among the eastern Atlantic stocks.** *Proc Biol Sci.* 2000; **267**(1440): 281–292.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Niknafs YS, Pandian B, Iyer HK, *et al.*: **TACO produces robust multisample transcriptome assemblies from RNA-seq.** *Nat Methods.* 2017; **14**(1): 68–70.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Nøttestad L, Utne KR, Óskarsson GJ, *et al.*: **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, J du Conseil.* 2016; **73**(2): 359–373.
Publisher Full Text

O'Brien L, Burnett J, Mayo RK: **Maturation of nineteen species of finfish off the Northeast coast of the United States, 1985-1990.** US Department of Commerce. NOAA Technical Report NMFS, 1993; **113**: 66.
Reference Source

Olafsdottir AH, Utne KR, Jacobsen JA, *et al.*: **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 Research Part II: Topical Studies in Oceanography.* 2019; **159**: 152–168.
Publisher Full Text

Ovenden JR, Berry O, Welch DJ, *et al.*: **Ocean's eleven: a critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries.** *Fish Fish.* 2015; **16**(1): 125–159.
Publisher Full Text

Overholtz WJ, Hare JA, Keith CM: **Impacts of interannual environmental forcing and climate change on the distribution of Atlantic mackerel on the U.S. Northeast continental shelf.** *Mar Coast Fish.* 2011; **3**(1): 219–232.
Publisher Full Text

Parsons LS: **Northern range extension of the Atlantic mackerel, *Scomber scombrus*, to Black Island, Labrador.** *J Fish Res Board Can.* 1970; **27**(3): 610–613.
Publisher Full Text

Payne MR, Danabasoglu G, Keenlyside N, *et al.*: **Skilful decadal-scale prediction of fish habitat and distribution shifts.** *Nat Commun.* 2022; **13**(1): 2660.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Pembleton LW, Cogan NOI, Forster JW: **StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations.**

Mol Ecol Resour. 2013; **13**(5): 946–952.
PubMed Abstract | **Publisher Full Text**

Pertea M, Pertea GM, Antonescu CM, *et al.*: **StringTie enables improved reconstruction of a transcriptome from RNA-seq reads.** *Nat Biotechnol.* 2015; **33**(3): 290–295.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Pinsky ML, Selden RL, Kitchel ZJ: **Climate-driven shifts in marine species ranges: scaling from organisms to communities.** *Ann Rev Mar Sci.* 2020; **12**: 153–179.
PubMed Abstract | **Publisher Full Text**

Potts WM, Henriques R, Santos CV, *et al.*: **Ocean warming, a rapid distributional shift, and the hybridization of a coastal fish species.** *Glob Chang Biol.* 2014; **20**(9): 2765–2777.
PubMed Abstract | **Publisher Full Text**

Radlinski MK, Sundermeyer MA, Bisagni JJ, *et al.*: **Spatial and temporal distribution of Atlantic mackerel (*Scomber scombrus*) along the northeast coast of the United States, 1985–1999.** *ICES J Mar Sci.* 2013; **70**(6): 1151–1161.
Publisher Full Text

Ranallo-Benavidez TR, Jaron KS, Schatz MC: **GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes.** *Nat Commun.* 2020; **11**(1): 1432.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Rhie A, McCarthy SA, Fedrigo O, *et al.*: **Towards complete and error-free genome assemblies of all vertebrate species.** *Nature.* 2021; **592**(7856): 737–746.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Rhie A, Walenz BP, Koren S, *et al.*: **Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies.** *Genome Biol.* 2020; **21**(1): 245.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Rodriguez-Ezpeleta N, Bradbury IR, Mendibil I, *et al.*: **Population structure of Atlantic mackerel inferred from RAD-seq-derived SNP markers: effects of sequence clustering parameters and hierarchical SNP selection.** *Mol Ecol Resour.* 2016; **16**(4): 991–1001.
PubMed Abstract | **Publisher Full Text**

Rodríguez-Ezpeleta N, Díaz-Arce N, Walter JF, *et al.*: **Determining natal origin for improved management of Atlantic bluefin tuna.** *Front Ecol Environ.* 2019; **17**(8): 439–444.
Publisher Full Text

RStudio: **Integrated development environment for R.** RStudio, PBC, Boston, MA, 2022.
Reference Source

Schwartz MK, Luikart G, Waples RS: **Genetic monitoring as a promising tool for conservation and management.** *Trends Ecol Evol.* 2007; **22**(1): 25–33.
PubMed Abstract | **Publisher Full Text**

Sette OE: **Biology of the Atlantic mackerel (*Scomber scombrus*) of North America: part II—migration and habits.** *Fisheries Bulletin of the Fish and Wildlife Service.* 1950; **110**.
Reference Source

Somdal O, Schram TA: **Ectoparasites on northeast Atlantic mackerel (*Scomber scombrus* L.) from western and north sea stocks.** *Sarsia.* 1992; **77**(1): 19–31.
Publisher Full Text

Stanke M, Schöffmann O, Morgenstern B, *et al.*: **Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources.** *BMC Bioinformatics.* 2006; **7**: 62.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Stapley J, Reger J, Feulner PG, *et al.*: **Adaptation genomics: the next generation.** *Trends Ecol Evol.* 2010; **25**(12): 705–712.
PubMed Abstract | **Publisher Full Text**

Studholme AL: **Essential fish habitat source document. Atlantic mackerel, *Scomber scombrus*, life history and habitat characteristics.** Northeast Fisheries Science Center (U.S.) NOAA technical memorandum NMFS-NE, 1999; 141.
Reference Source

Thorburn DMJ, Sagonas K, Binzer-Panchal M, *et al.*: **Origin matters: using a local reference genome improves measures in population genomics.** *Mol Ecol Resour.* 2023; **23**(7): 1706–1723.
PubMed Abstract | **Publisher Full Text**

Trenkel VM, Huse G, MacKenzie BR, *et al.*: **Comparative ecology of widely distributed pelagic fish species in the North Atlantic: implications for modelling climate and fisheries impacts.** *Prog Oceanogr.* 2014; **129**: 219–243.
Publisher Full Text

Uriarte A, Alvarez P, Iversen S, *et al.*: **Spatial pattern of migration and recruitment of North East Atlantic mackerel.** *ICES Annual Science Conference.* ICES CM 2001/O: 17. 2001.
Reference Source

Uriarte A, Lucio P: **Migration of adult mackerel along the Atlantic European shelf edge from a tagging experiment in the south of the Bay of Biscay in 1994.** *Fish Res.* 2001; **50**(1–2): 129–139.
Publisher Full Text

Walsh M, Hopkins P: **Estimation of total potential fecundity and atresia in the western mackerel stock, 1989.** *International Council for the Exploration of the Sea*. ICES CM 1990/H: 31 Pelagic Fish Committee, 1990.
Reference Source

Wang P, Wang F: **A proposed metric set for evaluation of genome assembly quality.** *Trends Genet.* 2023; **39**(3): 175–186.
PubMed Abstract | **Publisher Full Text**

Waples RS: **Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species.** *J Hered.* 1998; **89**(5): 438–450.
Publisher Full Text

Whitacre LK, Wildhaber ML, Johnson GS, *et al.*: **Exploring genetic variation and population structure in a threatened species, *Noturus placidus*, with whole-genome sequence data.** *G3 (Bethesda).* 2022; **12**(4): jkac046.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Wickham H: **ggplot2: elegant graphics for data analysis.** Springer-Verlag New York, 2016.
Publisher Full Text

Yang J, Benyamin B, McEvoy BP, *et al.*: **Common SNPs explain a large proportion of the heritability for human height.** *Nat Genet.* 2010; **42**(7): 565–569.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Zardoya R, Castilho R, Grande C, *et al.*: **Differential population structuring of two closely related fish species, the mackerel (*Scomber scombrus*) and the chub mackerel (*Scomber japonicus*), in the Mediterranean Sea.** *Mol Ecol.* 2004; **13**(7): 1785–1798.
PubMed Abstract | **Publisher Full Text**

Open Peer Review

Current Peer Review Status: ? ? ✓ ✓ ✓ ✓

Version 2

Reviewer Report 20 February 2025

https://doi.org/10.21956/openreseurope.20953.r50707

© **2025 Hargrove J.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✓ **John S Hargrove**
Pacific States Marine Fisheries Commission, Trout Rd, Eagle, USA

In their article titled "Atlantic mackerel population structure does not support genetically distinct spawning components", Manuzzi et al. generated an annotated genome for the Atlantic mackerel, collected reduced representation sequence data consisting of tens of thousands of SNPs, and used these data to test for the presence of genetic structuring within the spawning components of the northeast Atlantic Ocean and northwest Atlantic Ocean. The authors provide excellent justification for the study, noting that Atlantic mackerel are of significant economic importance, that climate change-related range expansion has led to management complications, and that ambiguity remains regarding the genetic composition of individual spawning components despite previous research studies. The authors note the utility of a reference genome as this resource facilitates novel forms of downstream analysis (e.g. detection of genomic regions under selection) which may further advance conservation and management efforts.

Overall, I was very impressed with the breadth and quality of this study. The construction of a reference genome is a major accomplishment in and of itself, and the methodologies used to sequence, assemble, and annotate the genome were well-established and appropriate. Access to a reference genome minimized potential complications associated with *de novo* locus identification, thus bolstering the quality of the data used for population structure analysis. I found the statistical approaches to assess population structure and the genetic relationships between spawning components were appropriate and correctly applied. Lastly, I thought the authors did an excellent job interpreting the results and placing them in the context of existing knowledge. I applaud the researchers on their work!

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: My research focuses broadly on population genetics with a particular emphasis on genetic stock structure, mating systems, and inter/intraspecific hybridization.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 11 February 2025

https://doi.org/10.21956/openreseurope.20953.r50706

© 2025 Catanese G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Gaetano Catanese
IRFAP LIMIA (Laboratorio de Investigaciones Marinas y Acuicultura) – Govern de les Illes Balears, Balearic Islands, Spain

In this study, the authors sequenced and annotated the genome of *Scomber scombrus* to investigate population structure and potential genetic adaptations. Using RAD-seq data from over 500 samples, the analysis of SNPs confirmed the existence of three genetically distinct units (Northwest Atlantic, Northeast Atlantic, and Mediterranean) but found no evidence of separate spawning components within the Atlantic. These findings could clarify previous uncertainties, rejecting homing behavior and the need to redefine management boundaries for the species.

First of all, I would like to commend the authors for their outstanding work in developing this manuscript. The study is well-structured, presenting a clear and logical progression from the introduction to the conclusions. The introduction effectively contextualizes the research question, providing a solid theoretical background and a well-articulated justification for the study. The discussion is equally well-developed, offering a critical interpretation of the findings in light of existing literature, while also highlighting their significance and potential implications.
A particularly strong aspect of this work is the robustness of the methodology and the comprehensiveness of the analyses. The results are supported by rigorous statistical and

analytical approaches, which enhance the reliability and credibility of the conclusions drawn. The authors have demonstrated great attention to detail in ensuring that the findings are presented clearly and transparently.
Additionally, I would like to acknowledge the thorough and detailed review process carried out by the other reviewers. Their insightful comments and suggestions have contributed significantly to refining the manuscript. The authors have responded thoughtfully to all critiques, implementing the requested revisions in a precise and constructive manner. Their efforts in addressing ambiguities and improving clarity are evident, further strengthening the overall quality of the work.

However, I would like to offer a comment and recommendation to enhance the manuscript, specifically regarding the genetic differentiation of the samples.
In response to another reviewer's request, the authors included pairwise FST comparisons between Mediterranean and Atlantic groups in the Results section. However, I believe it would be valuable to provide a more comprehensive FST table that includes all sampling sites and to expand the discussion (also just one sentence) by considering differentiation within the Mediterranean Sea and comparing them with Atlantic locations.
Specifically, I have two questions. First, what is the authors' hypothesis regarding genetic differentiation between the Adriatic and Tyrrhenian/Western Mediterranean regions? Second, could some genetic markers (outliers) show similarities between geographically distant sampling sites in the Atlantic and Mediterranean Sea?
As in this study, similar findings of Mediterranean genetic differences were reported by Rodríguez-Ezpeleta et al. (2016), but also that study did not explore in depth non-neutral markers of this region.

I understand that this could be not the primary objective of the manuscript and I also recognize the potential limitations arising from the relatively low number of sampling sites and individuals in the Mediterranean, which could affect the interpretation of the results. However, some attempt to discuss or comment on the observed differences in the considered areas could be insightful.

In this work, the authors conducted an analysis for outlier detection of the two Atlantic regions (indicating lack of sub structuring), but not for the Mediterranean—or at least, no mention of such analysis is made in the manuscript. If evidence of selective pressures related to environmental factors are found between Mediterranean populations and in relation to Atlantic ones, it could provide a highly interesting (even if not conclusive) perspective that might offer new insights into genetic similarities or differences among spawning components worldwide.

To clarify, as noted by the authors, previous studies on other species have suggested that genetic differentiation can arise due to adaptation to specific environmental factors (such as bathymetry, salinity, dissolved oxygen....etc.). This is typically reflected in genetic markers under selective pressure (outliers). For example, in *Engraulis encrasicolus*, the same genes associated with lower salinity habitat have been shown to drive the evolution of ecotypes inhabiting estuarine areas, even in geographically distant regions like the Bay of Biscay, the Gulf of Cádiz, and the Tyrrhenian Sea (Catanese G, et al., 2017 [Ref-1]). Considering this, the authors may want to explore whether similar patterns of adaptive differentiation are present in their dataset in relation to different areas.

In conclusion, I believe this manuscript is a valuable contribution to the field, both for its scientific

rigor and its potential impact. I appreciate the authors' efforts to enhance the clarity and depth of their study and, although I had some suggestions, consider the manuscript suitable for indexing.

Updated Note:
I agree with the authors' response and I appreciate their clarification. Their explanation effectively addresses my curiosity regarding the inclusion of Mediterranean data and its role as an outgroup in the study. I also understand that further analyses on Mediterranean populations would shift the focus away from the main objectives of the research. This answer fully satisfies my inquiries.

References
1. Catanese G, Watteaux R, Montes I, Barra M, et al.: Insights on the drivers of genetic divergence in the European anchovy.*Sci Rep*. 2017; **7** (1): 4180 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I mainly work on the application of molecular techniques to fisheries and aquaculture studies, investigating the genetic structure and evaluating the current status of genetic variability of wild populations of fisheries resources and marine organisms. Moreover, I work on authenticity of fishery products, on the detection of pathogens affecting marine organisms and on strategies to minimize by-catches of traditional fishing for a gradual elimination of discards in European fisheries

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 11 Feb 2025

Naiara Rodriguez Ezpeleta

We sincerely thank the reviewer for their thorough examination of our work and the positive feedback provided. While we appreciate the interest in the Mediterranean population, it is important to clarify that this is not the primary focus of our study. The Mediterranean data were included as an outgroup to improve understanding of mackerel populations on both sides of the Atlantic Ocean. Additionally, as noted by the reviewer, the sample size from the Mediterranean is limited and does not consider spawning and non-spawning individuals from various areas. Including the suggested analyses and speculating on the evolutionary history of Mediterranean populations would divert attention from the main focus of our research. The data have been made publicly available and is available at the GenBank SRA archive since 2016, allowing any interested party to analyze them independently or in conjunction with additional samples. We hope this clarification is satisfactory.

Competing Interests: No competing interests were disclosed.

Reviewer Response 18 Feb 2025

GAETANO CATANESE

I agree with the authors' response and I appreciate their clarification. Their explanation effectively addresses my curiosity regarding the inclusion of Mediterranean data and its role as an outgroup in the study. I also understand that further analyses on Mediterranean populations would shift the focus away from the main objectives of the research. This answer fully satisfies my inquiries.

Competing Interests: No competing interests were disclosed.

Reviewer Report 11 February 2025

https://doi.org/10.21956/openreseurope.20953.r50346

© 2025 Bektas Y. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Yusuf Bektas
Faculty of Arts and Sciences, Department of Molecular Biology and Genetics, Karadeniz Technical University, Ortahisar, Turkey

Dear Editor,

This manuscript investigates the population structure of Atlantic mackerel (*Scomber scombrus*) by genetic analysis to investigate whether there are genetically distinct spawning components. Whole genome sequencing (WGS) and RAD-seq methods were used to test whether there are genetic differences between the Northwest Atlantic (NWA), Northeast Atlantic (NEA) and Mediterranean

(MED) populations. It was determined that the Atlantic mackerel population is genetically divided into three main units and that there are no genetically distinct spawning components within NWA and NEA. Mackerel in Greenland is genetically identical to the NEA population, i.e., a northward expansion was observed. Climate change has been shown to affect the distribution of mackerel, but no genetic adaptation signals were detected. The authors confirmed these results using ADMIXTURE, PCA, FST and SNP under selection analyses. The study has important implications for population genetics, fisheries management and climate change impacts.

The research question is clear and meaningful. The genetic structure of Atlantic mackerel and an important issue in terms of fisheries management are addressed. The sampling includes sufficient geographical coverage for population genetics. A lack of samples from Southern Europe (south of Portugal, North Africa) was noted. Additional samples from this region could have been collected to expand the analyses. The current sampling strategy can be validated by supporting it with fisheries records.

The combination of whole genome sequencing (WGS) and RAD-seq provides a robust genetic approach.

Population genetic analyses were appropriate and meticulously implemented. Linkage disequilibrium (LD) control was performed and appropriate SNP filtering was applied for the analyses. Statistical significance was achieved using multiple test correction (FDR). Permutation tests and statistical cross-validations were performed. Adaptive SNPs can be validated with additional selection analyses such as BAYESCAN or FLK.

The results are consistent with the literature and scientifically sound. Although the FST values are low, they are correctly interpreted as significant. The claim that Atlantic mackerel does not exhibit homing behavior is supported by genetic data. The result that the mackerels in Greenland are the same as the NEA population has been analyzed in accordance with the principles of population genetics. Could functional analyses be performed for the SNPs under selection? It can be tested whether these SNPs are associated with genetic adaptation processes.
Potential adaptive mechanisms can be examined by determining which genes are located in the relevant genome regions. Can genetic data be integrated with ecological factors such as sea temperatures and migration routes?
All genetic data and analysis codes have been shared as open access. The bioinformatics analysis pipeline has been documented and reproducibility has been ensured.

In conclusion
The study is scientifically sound, meticulously prepared and strong in terms of genetic analyses. The results are sufficiently supported by the existing data, but certain improvements can be made. It is strong in terms of reproducibility and is a transparent study thanks to the open data policy.

Best regards,

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Phylogenetic systematic and Population Genetics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 07 November 2024

https://doi.org/10.21956/openreseurope.18768.r44860

© 2024 Verba J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Julia Tovar Verba
[1] Division of Evolutionary Biology, Faculty of Biology, Ludwig-Maximilians-Universität München, Munich, Germany
[2] Division of Evolutionary Biology, Faculty of Biology, Ludwig-Maximilians-Universität München, Munich, Germany

In the manuscript entitled “Atlantic mackerel population structure does not support genetically distinct spawning components”, the authors aim to examine the genetic differentiation and connectivity of the species, adding to and clarifying conflicting results in previous studies. The authors utilize broadly distributed samples, an annotated reference genome and thousands of genomic markers, which are appropriate for addressing questions related to population structure of marine species, with large and highly connected populations. The authors also make a good effort in predicting future changes, and in providing clear information for management. Overall, the manuscript is clearly articulated, and the methodology is robust. However, I believe the paper

can be improved by addressing some minor points.
In the introduction, could you please include since when the historical records related to Atlantic mackerel distribution exist? This could include information about early documentation of the species distribution patterns and any observed shifts prior to recent changes - if existent. Also in the introduction, the authors mention that the whole range of the species was sampled, although as mentioned later on in the manuscript, there are no samples from Portugal or from the North Atlantic African coast. In the methods section, I would recommend a more thorough explanation regarding the inclusion of all age classes. Additionally, it is not clear if there was no dataset with only Mediterranean samples. In the results section, I recommend the integration of the K = 2 plot within the main text since it might be relevant to see the shared ancestry between the Mediterranean and Northeast Atlantic. Also, I would like to see the other Ks (4 and maybe 5) in the supplementary material, to see if the division in the Mediterranean populations appears. In the discussion, could you please clarify the sentence “In addition, the species shows high plasticity as it can tolerate up to 5-15°C (Olafsdottir et al., 2019) (ref-1) despite its optimal temperature range (9-13°C), which is possibly linked to the high genomic heterozygosity detected.”? It is hard to understand if both plasticity and heterozygosity are on the species or individual level. In the discussion, I recommend making it clear that the sample distribution of this study is more broadly distributed in comparison with other studies. Since there were previous papers with a similar objective and methods, it would be good to see it clear what is the addition of this manuscript. The discussion surrounding adaptation mechanisms is currently limited and could be improved. But mainly I would like to see a more thorough examination of Mediterranean population divergence in the discussion.

References

1. Olafsdottir A, Utne K, Jacobsen J, Jansen T, et al.: 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 Research Part II: Topical Studies in Oceanography*. 2019; **159**: 152-168 Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and does the work have academic merit?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Ecology, population genomics, marine biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 09 Jan 2025

Naiara Rodriguez Ezpeleta

We would like to thank the reviewer for assessing our manuscript and for offering valuable insights. Below we provide answers to the comments raised:

COMMENT: In the introduction, could you please include since when the historical records related to Atlantic mackerel distribution exist? This could include information about early documentation of the species distribution patterns and any observed shifts prior to recent changes - if existent.
ANSWER: We have decided not to add this information to the introduction section as we do not think it is essential for the purpose of the study.

COMMENT: Also in the introduction, the authors mention that the whole range of the species was sampled, although as mentioned later on in the manuscript, there are no samples from Portugal or from the North Atlantic African coast.
ANSWER: replaced "across the distribution" by "across most of the distribution"

COMMENT: In the methods section, I would recommend a more thorough explanation regarding the inclusion of all age classes.
ANSWER: we have included the fact that adults included spawning and non-spawning individuals; details about age of each individual are included in a supplementary table.

COMMENT: Additionally, it is not clear if there was no dataset with only Mediterranean samples.
ANSWER: The purpose of this study is to assess connectivity within the Atlantic and in particular within each side of the Atlantic. This is why a dataset with only mediterranean samples was not created. Also, the Mediterranean samples included had already been analyzed in a previous study (Rodríguez-Ezpeleta et al. 2016 Molecular Ecology Resources)

COMMENT: In the results section, I recommend the integration of the K = 2 plot within the main text since it might be relevant to see the shared ancestry between the Mediterranean and Northeast Atlantic.
ANSWER: We have kept K=2 in supplementary material not to overload Figure 1 which is already quite busy.

COMMENT: Also, I would like to see the other Ks (4 and maybe 5) in the supplementary material, to see if the division in the Mediterranean populations appears.

ANSWER: The division between Mediterranean populations has already been assessed in a previous study (Rodriguez-Ezpeleta et al. 2016 Molecular Ecology Resources) and no new Mediterranean samples have been included in this study.

COMMENT: In the discussion, could you please clarify the sentence "In addition, the species shows high plasticity as it can tolerate up to 5-15°C (Olafsdottir et al., 2019) (ref-1) despite its optimal temperature range (9-13°C), which is possibly linked to the high genomic heterozygosity detected."? It is hard to understand if both plasticity and heterozygosity are on the species or individual level.
ANSWER: We have added, "at the species level"

COMMENT: In the discussion, I recommend making it clear that the sample distribution of this study is more broadly distributed in comparison with other studies. Since there were previous papers with a similar objective and methods, it would be good to see it clear what is the addition of this manuscript.
ANSWER: added "and most broadly distributed samples" to the first sentence of introduction.

COMMENT: The discussion surrounding adaptation mechanisms is currently limited and could be improved. But mainly I would like to see a more thorough examination of Mediterranean population divergence in the discussion.
ANSWER: As mentioned above, the Mediterranean sea population is out of the scope of this study.

Competing Interests: No competing interests were disclosed.

Reviewer Report 16 October 2024

https://doi.org/10.21956/openreseurope.18768.r44524

© 2024 Winter S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Sven Winter
[1] University of Veterinary Medicine, Vienna, Austria
[2] Senckenberg Biodiversity and Climate Research Centre (BiK-F), Frankfurt am Main, Germany
[3] University of Veterinary Medicine, Vienna, Austria
[4] Senckenberg Biodiversity and Climate Research Centre (BiK-F), Frankfurt am Main, Germany

In this study, the authors sequenced, assembled, and annotated the reference genome of the Atlantic mackerel from Oxford Nanopore long-read and Illumina short-read data.
They further used this reference to map newly generated population-level RAD-sequencing data and study population structure in the North Atlantic, aiming to identify the impact of spawning components on the genetic structure of the species and identify the origin of individuals of the

recent northward expansion.
The authors provide a vital genetic resource, which, quite frankly, I expected to be generated years ago due to the species' economic value.
The chosen sequencing, assembly, and annotation methods are all very sound and follow well-established procedures. However, they fall just a little short of the current standards set by most genome consortia due to the lack of proximity-ligation scaffolding, which results in an assembly that is not (near) chromosome level.
The methods used for the population genomic analyses are also sound and well-established. Yet, they do not exhaust the possibilities the dataset grants them, especially for the task set by the aims mentioned by the authors. If the task was to study spawning components, then the authors need to define those locations and analyze the genetic makeup of individuals in more detail. A standard procedure of PCA and Admixture is a nice start, but I would suggest including more analyses on genetic connectivity, gene flow, and, ideally (if the density of Rad loci permits it), structural variation, which has been shown to influence population structure and behavior of populations drastically (e.g., in quails but also scallops). In addition, even though this is a species with large numbers, I am wondering if the current state of fisheries management would benefit from analyses focused on effective population size estimates (e.g., GONE, Ne-Estimator etc.) to see if there is a genetic impact over recent decades.
I want to congratulate the authors on generating a great dataset. I feel that the study so far is well-written and well-conducted, but it would benefit from some more in-depth analyses to support the conclusions and gain additional information useful for the management of the species.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: genome assembly, population genomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have

significant reservations, as outlined above.

Author Response 09 Jan 2025

Naiara Rodriguez Ezpeleta

We would like to thank the reviewer for the insightful comments provided on our manuscript. Below, we provide an answer to each of them:

COMMENT: The chosen sequencing, assembly, and annotation methods are all very sound and follow well-established procedures. However, they fall just a little short of the current standards set by most genome consortia due to the lack of proximity-ligation scaffolding, which results in an assembly that is not (near) chromosome level.

ANSWER: Yes, we are aware: the reason for this was the difficulty of obtaining high quality DNA from mackerel tissue samples despite using freshly caught individuals and following similar procedures than with other fish species for which we were successful. In response to this comment and to a similar one by reviewer 1, we have replaced “high-quality genome” by “genome” throughout the manuscript.

COMMENT: The methods used for the population genomic analyses are also sound and well-established. Yet, they do not exhaust the possibilities the dataset grants them, especially for the task set by the aims mentioned by the authors. If the task was to study spawning components, then the authors need to define those locations and analyze the genetic makeup of individuals in more detail. A standard procedure of PCA and Admixture is a nice start, but I would suggest including more analyses on genetic connectivity, gene flow, and ideally (if the density of Rad loci permits it), structural variation, which has been shown to influence population structure and behaviour of populations drastically (e.g., in quails but also scallops).
ANSWER: By using analyses based on outliers, we have implicitly tested presence of potential structural variants, as presence of these should reveal alternative groupings in PC analyses as we have seen in other studies (see for example Diaz-Arce et al. 2024 Molecular Ecology). Here, despite doing analyses based on outlier markers, we do not find any additional pattern suggesting a structural variant that is large enough to be detectable with the SNPs included in our study.

COMMENT: In addition, even though this is a species with large numbers, I am wondering if the current state of fisheries management would benefit from analyses focused on effective population size estimates (e.g., GONE, Ne-Estimator etc.) to see if there is a genetic impact over recent decades.

ANSWER: The most critical question for Atlantic mackerel at the moment is that about the presence of isolated spawning components, which is the one we mostly tackle in this study. Of course, analyses looking at the effect of effective population size over time would be interesting, but we do not think that the samples we have allow for such analyses, given that we have a small number of individuals per year/location/cohort. However, we are sure that the genetic resources we have produced will be extremely valuable for other studies wishing to continue on this matter.

Competing Interests: No competing interests were disclosed.

Author Response 13 Jan 2025

Naiara Rodriguez Ezpeleta

We would like to thank the reviewer for the insightful comments provided on our manuscript. Below, we provide an answer to each of them:

COMMENT: The chosen sequencing, assembly, and annotation methods are all very sound and follow well-established procedures. However, they fall just a little short of the current standards set by most genome consortia due to the lack of proximity-ligation scaffolding, which results in an assembly that is not (near) chromosome level.

ANSWER: Yes, we are aware: the reason for this was the difficulty of obtaining high quality DNA from mackerel tissue samples despite using freshly caught individuals and following similar procedures than with other fish species for which we were successful. In response to this comment and to a similar one by reviewer 1, we have replaced “high-quality genome” by “genome” throughout the manuscript.

COMMENT: The methods used for the population genomic analyses are also sound and well-established. Yet, they do not exhaust the possibilities the dataset grants them, especially for the task set by the aims mentioned by the authors. If the task was to study spawning components, then the authors need to define those locations and analyze the genetic makeup of individuals in more detail. A standard procedure of PCA and Admixture is a nice start, but I would suggest including more analyses on genetic connectivity, gene flow, and ideally (if the density of Rad loci permits it), structural variation, which has been shown to influence population structure and behaviour of populations drastically (e.g., in quails but also scallops).

ANSWER: By using analyses based on outliers, we have implicitly tested presence of potential structural variants, as presence of these should reveal alternative groupings in PC analyses as we have seen in other studies (see for example Diaz-Arce et al. 2024 Molecular Ecology). Here, despite doing analyses based on outlier markers, we do not find any additional pattern suggesting a structural variant that is large enough to be detectable with the SNPs included in our study.

COMMENT: In addition, even though this is a species with large numbers, I am wondering if the current state of fisheries management would benefit from analyses focused on effective population size estimates (e.g., GONE, Ne-Estimator etc.) to see if there is a genetic impact over recent decades.

ANSWER: The most critical question for Atlantic mackerel at the moment is that about the presence of isolated spawning components, which is the one we mostly tackle in this study. Of course, analyses looking at the effect of effective population size over time would be interesting, but we do not think that the samples we have allow for such analyses, given

that we have a small number of individuals per year/location/cohort. However, we are sure that the genetic resources we have produced will be extremely valuable for other studies wishing to continue on this matter.

Competing Interests: No competing interests were disclosed.

Reviewer Report 07 June 2024

https://doi.org/10.21956/openreseurope.18768.r40457

© 2024 Bayona N. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Natalia Bayona
[1] Division of Natural Science and Mathematics, Oxford College, Emory University, Oxford, Georgia, USA
[2] Division of Natural Science and Mathematics, Oxford College, Emory University, Oxford, Georgia, USA

In the manuscript entitled “Atlantic Mackerel Population Structure Does Not Support Genetically Distinct Spawning Components,” the authors aim to sequence, assemble, and annotate the whole genome of *Scomber scombrus* and use such to map restriction-enzyme associated fragments to call SNPs from individuals across most of the current geographic distribution range. These methods were used to assess contrasting hypotheses of either fine genetic differences (based on otolith, meristic, allozymes, mitochondrial, and microsatellite markers) or mixing (based on tagging and a recent genomic study) between spawning components present within each of the recognized populations/stocks of the species (NEA, NWA, and MED). The authors also aimed to identify genetic markers associated with local adaptation processes of these spawning groups. Finally, researchers assessed the origin of the population expansion that species suffer northwards and the potential loci associated with this shift.

I applaud the researchers' work and effort in generating a brand-new genome assembly and annotation for the species. This is a valuable resource that is long due for this and other marine species with high economic and ecological relevance in their distribution range. I found the methods and results regarding genome assembly and annotation sounding and appropriate. Minor suggestions are given below.

In terms of the population genetics analyses, the methods used are also suitable for the system. However, in terms of the hypotheses laid in the introduction, and thus the ones I believe the authors aim to address, I have some recommendations. First, the authors, in the title and introduction, state that they aim to assess if spawning components correspond to genetic discrete or mixed groups. However, the study falls short in delineating the currently recognized spawning components, assigning the samples collected to these components, and explicitly considering these in the data analyses. For example, in the Northeast Atlantic, there have been recognized

three components in the spawning seasons (Southern: Spain and Portugal, Western: Bay of Biscay and Ireland, and North Sea; Jansen and Gislason., 2013; Collette and Graves 2019). In the Northwest Atlantic, two distinct spawning components, northern: Canada-Gulf of St. Lawrence, and southern: US-Mid-Atlantic Bight to Gulf of Main, have been proposed (Moura et al., 2020). This study has the potential to directly test for genetic differentiation between spawning components if samples are grouped and assigned accordingly for analyses. I understand that given the results from ADMIXTURE and PCAs, there seems to be no relevant signal, and it is likely that the pattern will be maintained, but to make strong conclusions (included in the title), specific analyses between spawning components should be performed. For example, if the color or shape in the PCAs could represent the spawning components rather than (or additionally to) sampling locations, there could be a more conspicuous and straightforward signal (lack of), and the conclusion would be more obvious. Similarly, it is interesting that FST values are very low for different stocks within the Mediterranean. If the significance of these values can be included, along with those for comparisons of components in reference samples, that would also aid the audience in understanding the evidence of differences/similarities between spawning components. Also, under this context, it may be relevant for the authors to acknowledge that no samples are obtained from what would represent the spawning Southern NEA component, and thus, conclusions should be taken carefully given this underrepresentation. Finally, I would recommend authors assess if there are fine-scale differences within spawning seasons (years), given that they collected samples that span across multiple years (2011-2017) in different components and the grouping of all of these just by location, regardless of timing, may cofound genetic effects may have within seasons. While doing this, it is important to also omit individuals from the respective spawning areas that were not in the process of spawning (including only certain spawning adults; if larvae are included, it is important to maybe segregate by season).

Additionally, other studies have argued population connectivity between the western Mediterranean (Barcelona) and the NEA (Zardoya et al., 2004), which is not evident in this study even though the westernmost sample from the Mediterranean (WESTMED) is even closer to the Atlantic than the sample from Zardoya et al., 2004. This is valuable evidence that should be highlighted in the discussion.

Overall, the study has a good set of analyses and presents and discusses trends in a logical and pertinent fashion. I celebrate the effort and great amount of work and thought the authors have placed into the study. My main concern was highlighted above, and minor suggestions and comments on the study can be found below, which include some suggestions in scientific writing as well.

Major comments:

1. Figure S2 is identical to Figure 3 from Cruz et al., 2023, Refer ref [4]. Although the two papers have one co-author in common, in this case, I suggest the authors remove the figure from the supplementary and instead provide proper attributions to the published paper.
2. The title and introduction of the paper suggest that the core analyses rely on detecting whether spawning components are genetically different. However, such comparisons are not presented in methods or results. For example, the colors and shapes of the PCA and ADMIXTURE analyses correspond to geographically distinct locations but not to spawning components. In my letter above, I presented examples of how to consider the established spawning components in the analyses and the year of the samples.

Minor comments:

Title

"Atlantic mackerel population structure does not support genetically distinct spawning components." It is a strong title, but conclusions should be drawn from explicit analyses, suggested in major comments.

Abstract

1. The background information could be enhanced by including the current understanding of the number of populations and spawning components that have been characterized in the distribution range.
2. I would not consider this a "high-quality" genome assembly.

Intro

1. I would add to the first sentence that the species is **coastal** pelagic so that geographic distribution breaks in the North Atlantic can be more easily understood.
2. I would add "… Labrador, **Canada** to Cape Lookout, **United States** in the west." If countries are given for the eastern population instead of cities, it would be good to maintain some consistency.
3. It is relevant that the authors highlight that the species is also present in the Black Sea.
4. When citing the spawning groups in the Northwest, it is important that authors recognize the original work of O'Brien et al. 1993.
5. When citing the spawning groups in the Northeast, it is important that authors recognize the original work of Nesbø et al., 1999-2000.
6. I celebrate the initiative of the authors to generate a reference genome for the species, but I would not call this assembly a "high-quality" given the large number of scaffolds still present.
7. The authors portray one of the objectives to assess the origin of the Greenlandic/Islandic population and markers associated with the recent northward migration. I found it relevant here to cite hypotheses based on Jansen et al., 2016.

Materials & Methods

1. I suggest the authors differentiate in this section the samples used for genome assembly and annotation from the samples used for population analyses. For example, it is unclear which sample was used for genome sequencing. Was it the same female used for its annotation? If so, what tissue was used for DNA extraction and sequencing? Where was the female captured? Also, no further information is given regarding the amount and quality of DNA used for library preparation.
2. I recommend that authors include either in Figure 2A or in a summary table the groups or samples representing the spawning components and even maybe stages and years of

collection per locality. The latter is because, in the introduction, the authors make the case that other studies did not account for differences between year classes/generations, which limits the scope of conclusions. Although this information is present in Supplemental Materials, it should be easy to access and summarized by the authors in this section so the audience can grasp the power of these analyses.

3. When using Guppy for base calling, which mode/parameters were used? For example, was the mode “high-accuracy” used for both data types?
4. The authors claimed, “The free alternative open-source high performance base calling software Dorado can also be used to base call the reads”. Do they mean that they use this software? If so, for which reads and what version?
5. Remove the extra SbfI in the fourth line of the third paragraph of methods and include that the enzyme is at 10 U/uL.
6. Correct typo “Enzime” for “Enzyme” on line six of the third paragraph of methods.
7. If the authors did not use the 10X T4 DNA ligase buffer from ThermoFisher, but an NEB buffer instead, please provide the catalog number from NEB of the buffer used.
8. The ATP typically used for libraries is at 10 mM, but here authors state that they used a 100 mM, could you please provide manufacturer and catalogue number?
9. What were the number of cycles and times for the shearing of DNA fragments in Covaris?
10. Correct typo “1,5” for “1.5” on line 18 of the first paragraph of page 6.
11. Revise that throughout the text, spaces should separate numbers and their units. There is inconsistency throughout.
12. In the second line of the second paragraph of page six, the 5 uL is not per sample but per ligated pool, right? Also, please mention the primers used here, including the manufacturer. Revise spaces in the PCR profile.
13. The document containing Fig S1 and S2 is not named “Extended Data” but “Supplementary Material.”
14. Please include the versions of some software under “Genome Assembly and Annotation”: TrimGalore, Fitlong2, BWA, HyPo, purge_dups, etc.
15. What version of Piccard was used for deduplication?
16. Missing a parenthesis closure when explaining the data set used to assess connectivity between spawning components.
17. Please include versions of the ADMIXTURE and R packages used.
18. Please specify what subset or subsets of individuals were used for pcadapat and its version.

Results

1. Table 2 has a column title “SCO1A Annotation” However, the genome assembly label is fScoSco3.1, what does SCO1A means in this context?
2. There is one minor discrepancy between Table 1 and the ENA project information according to the accession number provided by the authors. Table 1 states a total number of sequences of 4,483, and ENA states 4,482. The total length in the table is 741,290,963 bp and in ENA 741,290,950 bp.
3. The use of GenomeScope for genome profiling included in FigS3 should be included in the methods and properly cited.
4. In table S2, the thousandth separators under the column contigs should be a comma and not a quote symbol. Also, the “Coverage” column should indicate units, which I believe are “X.”

5. The title of Figure S4 should say that the colors match those from Figure 2, not 1.
6. Revise writing of sentence “FST values further validated this split, as all pairwise comparisons between the three groups resulted **high** with the **highest** FST value reported for MED vs NWA (FST=0.035)”. Also, the NEA vs MED should be FST=0.019 instead of 0.19.
7. The authors mentioned in the methods that the statistical significance of Fst values is assessed and adjusted, but these values are not presented in the results. Please include these values. They are particularly relevant for the Fst values presented within the Mediterranean Sea.
8. Remove the close parenthesis in the sentence: “This differentiation within the Mediterranean Sea is also seen in the FST pairwise comparisons, with FST 0.003 and 0.004 for the Adriatic Sea versus the Tyrrhenian Sea and Western Mediterranean Sea, respectively**)**.”

Discussion

1. When discussing the origin of the Greenlandic/Islandic population, it is relevant to cite the original work of Jansen et al., 2016.
2. In the paragraph: “Despite inter-population FST values being low (FST=0.014-0.035), these are in the range of what is usually observed in marine fish species with large population sizes since FST estimates are affected by both effective population sizes and migration rates, which may derive into insufficient statistical power provided by neutral markers to reject the panmictic hypothesis” I am not sure if this is arguing that the values observed are not statistically significant. Please clarify and make sure to include the p-values of the tests in the results.
3. The paragraph: “Furthermore, the high rate of heterozygosity detected at genome level may reflect the presence of a large metapopulation, for which patterns of low diversity would persist even in the presence of uneven migration along its distribution… ” results are confusing to me. High values of heterozygosity indicate high and not low genetic diversity. What the cited study argues is that in populations (not an individual), high levels of heterozygosity may be associated with large effective population sizes at a metapopulation level rather than very high local effective population sizes, but migration helps homogenize these values in the metapopulation. To further address this, the authors of this study should then present the values of heterozygosity observed globally and locally with the SNPs and address such hypotheses rather than making such relevant final remarks based on the heterozygosity of one individual for which the origin was not depicted. Finally, to draw conclusions on metapopulation models, I recommend that the authors test such hypotheses, perhaps by estimating the effective number of migrants between spawning components and the three established populations.

I congratulate the authors for the hard work they have put in so far, and I believe that with the proper focus/direction, the manuscript will be able to be indexed in the near future.

Thanks for allowing me to review the manuscript.

References

1. Jansen T, Gislason H: Population structure of Atlantic mackerel (Scomber scombrus).*PLoS One*. 2013; **8** (5): e64744 PubMed Abstract | Publisher Full Text
2. Moura A, Muniz A, Mullis E, Wilson J, et al.: Population structure and dynamics of the Atlantic mackerel (Scomber scombrus) in the North Atlantic inferred from otolith chemical and shape

signatures. *Fisheries Research*. 2020; **230**. Publisher Full Text
3. Zardoya R, Castilho R, Grande C, Favre-Krey L, et al.: Differential population structuring of two closely related fish species, the mackerel (Scomber scombrus) and the chub mackerel (Scomber japonicus), in the Mediterranean Sea.*Mol Ecol*. 2004; **13** (7): 1785-98 PubMed Abstract | Publisher Full Text
4. Cruz F, Gómez-Garrido J, Gut M, Alioto TS, et al.: Chromosome-level assembly and annotation of the Xyrichtys novacula (Linnaeus, 1758) genome.*DNA Res*. 2023; **30** (5). PubMed Abstract | Publisher Full Text
5. Nesbø CL, Rueness EK, Iversen SA, Skagen DW, et al.: Phylogeography and population history of Atlantic mackerel (Scomber scombrus L.): a genealogical approach reveals genetic structuring among the eastern Atlantic stocks.*Proc Biol Sci*. 2000; **267** (1440): 281-92 PubMed Abstract | Publisher Full Text
6. Jansen T, Post S, Kristiansen T, Óskarsson GJ, et al.: Ocean warming expands habitat of a rich natural resource and benefits a national economy.*Ecol Appl*. 2016; **26** (7): 2021-2032 PubMed Abstract | Publisher Full Text
7. Collette, B.B. and Graves, J: Tunas and Billfishes of the World. *Johns Hopkins University Press:Baltimore, Maryland*. 2019.
8. O'Brien, L., Burnett, J. and Mayo, R.K.: Maturation of nineteen species of finfish off the northeast coast of the United States, 1985-1990. *National Oceanographic and Atmospheric Administration (NOAA)*. 1993.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Genomics, bioinformatics, population genetics, conservation genetics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 13 Jan 2025

Naiara Rodriguez Ezpeleta

We would like to thank the reviewer for the insightful comments to our work. Please find below the comments provided and corresponding answers.

COMMENT: In terms of the population genetics analyses, the methods used are also suitable for the system. However, in terms of the hypotheses laid in the introduction, and thus the ones I believe the authors aim to address, I have some recommendations. First, the authors, in the title and introduction, state that they aim to assess if spawning components correspond to genetic discrete or mixed groups. However, the study falls short in delineating the currently recognized spawning components, assigning the samples collected to these components, and explicitly considering these in the data analyses. For example, in the Northeast Atlantic, there have been recognized three components in the spawning seasons (Southern: Spain and Portugal, Western: Bay of Biscay and Ireland, and North Sea; Jansen and Gislason., 2013; Collette and Graves 2019). In the Northwest Atlantic, two distinct spawning components, northern: Canada-Gulf of St. Lawrence, and southern: US-Mid-Atlantic Bight to Gulf of Main, have been proposed (Moura et al., 2020). This study has the potential to directly test for genetic differentiation between spawning components if samples are grouped and assigned accordingly for analyses. I understand that given the results from ADMIXTURE and PCAs, there seems to be no relevant signal, and it is likely that the pattern will be maintained, but to make strong conclusions (included in the title), specific analyses between spawning components should be performed. For example, if the colour or shape in the PCAs could represent the spawning components rather than (or additionally to) sampling locations, there could be a more conspicuous and straightforward signal (lack of), and the conclusion would be more obvious.
ANSWER: Larvae, juveniles and spawning adults are the samples representing spawning components; thus instead of coloring by spawning locations, which are not easy to clearly delimit, we have labelled samples according to age class also in the Figures 2D and E to better show the lack of differentiation, even when considering only reference samples corresponding to potential spawning components.

COMMENT: Similarly, it is interesting that FST values are very low for different stocks within the Mediterranean. If the significance of these values can be included, along with those for comparisons of components in reference samples, that would also aid the audience in understanding the evidence of differences/similarities between spawning components.
ANSWER: P-values of pairwise FST between Mediterranean locations and Atlantic reference samples comparisons now are included in the text (results section).

COMMENT: Also, under this context, it may be relevant for the authors to acknowledge that no samples are obtained from what would represent the spawning Southern NEA component, and thus, conclusions should be taken carefully given this underrepresentation.
ANSWER: Lack of comprehensive sampling of the southern NEA locations consequences of this gap are disclosed and commented in the Discussion section.

COMMENT: Finally, I would recommend authors assess if there are fine-scale differences within spawning seasons (years), given that they collected samples that span across

multiple years (2011-2017) in different components and the grouping of all of these just by location, regardless of timing, may cofound genetic effects may have within seasons. While doing this, it is important to also omit individuals from the respective spawning areas that were not in the process of spawning (including only certain spawning adults; if larvae are included, it is important to maybe segregate by season).
ANSWER: Analysis of differentiation between reference samples spanning different years and across different locations warrant the same results of no differentiation between spawning components, both at individual based level (see ADMIXTURE and PCA analysis in Fig. 2) as well as for Fst inferences as shown below (one group (CPCOD2015L) had to be removed from pop-based inference as only one larva was captured in that location/year. NEA (FST\p-values) PCPNE2013L (42) CANTB2013L (42) PCPNE2013SA (17) BBISC2013 SA (21) CANTB2013SA (27) NOSEA2017SA (21) PCPNE2013L (42) 0,449 0,000 0,000 0,023 0,002 CANTB2013L (42) 0,000 0,000 0,000 0,003 0,000 PCPNE2013SA (17) 0,002 0,001 0,000 0,000 0,001 BBISC2013SA (21) 0,001 0,001 0,000 0,000 0,001 CANTB2013SA (27) 0,000 0,000 0,000 0,000 0,000 NOSEA2017SA (21) 0,001 0,001 0,001 0,001 0,000 NWA (FST\p-values) CPCOD2004L (6) CPCOD2013L (7) CPCOD2016L (17) STLAW2012SA (28) CPCOD2004L (6) - 0,086 1,000 1,000 CPCOD2013L (7) 0,002 - 1,000 1,000 CPCOD2016L (17) 0,000 0,000 - 1,000 STLAW2012SA (28) 0,000 0,000 0,000 -

COMMENT: Additionally, other studies have argued population connectivity between the western Mediterranean (Barcelona) and the NEA (Zardoya et al., 2004), which is not evident in this study even though the westernmost sample from the Mediterranean (WESTMED) is even closer to the Atlantic than the sample from Zardoya et al., 2004. This is valuable evidence that should be highlighted in the discussion.
ANSWER: Reference Zardoya et al., 2004 has been added to the discussion section; however, we have decided not to put more emphasis on this because our study focuses on the Atlantic spawning components and because we already showed and discussed this in a previous work (Rodriguez-Ezpeleta et al. 2016).

COMMENT: Figure S2 is identical to Figure 3 from Cruz et al., 2023, Refer ref [4]. Although the two papers have one co-author in common, in this case, I suggest the authors remove the figure from the supplementary and instead provide proper attributions to the published paper. **ANSWER: This figure highlights the workflow followed for genome assembly as support for the methodological approach. Our work has no connection with the work of Cruz et al. 2023 other than the fact that both studies used the same standard workflow for genome assembly (as many others); thus, we have decided to maintain the figure as Supplementary Material.**

COMMENT: The title and introduction of the paper suggest that the core analyses rely on detecting whether spawning components are genetically different. However, such comparisons are not presented in methods or results. For example, the colors and shapes of the PCA and ADMIXTURE analyses correspond to geographically distinct locations but not to spawning components. In my letter above, I presented examples of how to consider the established spawning components in the analyses and the year of the samples**.**
ANSWER: Our analyses do show comparisons among components because they use reference samples labelled by both location and age class. In order to make this

clearer, as requested by the reviewer above, the main figure has been modified accordingly.

COMMENT: Title - “Atlantic mackerel population structure does not support genetically distinct spawning components.” It is a strong title, but conclusions should be drawn from explicit analyses, suggested in major comments.
ANSWER: We believe that our analyses strongly support what the title states. We have used reference samples (larvae, juveniles and spawning adults) from the different recognized spawning locations and no differentiation is seen between them. This implies that spawning components are not genetically distinct.

COMMENT: Abstract - The background information could be enhanced by including the current understanding of the number of populations and spawning components that have been characterized in the distribution range.
ANSWER: We believe that the information currently in the abstract is sufficient

COMMENT: I would not consider this a “high-quality” genome assembly.
ANSWER: “High-quality” was removed as suggested.

COMMENT: Intro I would add to the first sentence that the species is coastal pelagic so that geographic distribution breaks in the North Atlantic can be more easily understood.
ANSWER: Done

COMMENT: I would add “... Labrador, Canada to Cape Lookout, United States in the west.” If countries are given for the eastern population instead of cities, it would be good to maintain some consistency.
ANSWER: Done

COMMENT: It is relevant that the authors highlight that the species is also present in the Black Sea.
ANSWER: Done

COMMENT: When citing the spawning groups in the Northwest, it is important that authors recognize the original work of O’Brien et al. 1993. ANSWER: Done COMMENT: When citing the spawning groups in the Northeast, it is important that authors recognize the original work of Nesbø et al., 1999-2000.
ANSWER: Done

COMMENT: I celebrate the initiative of the authors to generate a reference genome for the species, but I would not call this assembly a "high-quality" given the large number of scaffolds still present.
ANSWER: we replaced “high-quality genome” by “genome” throughout the text

COMMENT: The authors portray one of the objectives to assess the origin of the Greenlandic/Islandic population and markers associated with the recent northward migration. I found it relevant here to cite hypotheses based on Jansen et al., 2016.
ANSWER: mentioned and cited.

COMMENT: Materials & Methods - I suggest the authors differentiate in this section the samples used for genome assembly and annotation from the samples used for population analyses. For example, it is unclear which sample was used for genome sequencing. Was it the same female used for its annotation? If so, what tissue was used for DNA extraction and sequencing? Where was the female captured? Also, no further information is given regarding the amount and quality of DNA used for library preparation.
ANSWER: Information on tissue type and quantity for samples used for population structure is already specified in the section "Tissue sampling and nucleotide extraction" and in "Library preparation and sequencing". Information on the sample used for assembly and annotation are now included in "Tissue sampling and nucleotide extraction".

COMMENT: Materials & Methods - I recommend that authors include either in Figure 2A or in a summary table the groups or samples representing the spawning components and even maybe stages and years of collection per locality. The latter is because, in the introduction, the authors make the case that other studies did not account for differences between year classes/generations, which limits the scope of conclusions. Although this information is present in Supplemental Materials, it should be easy to access and summarized by the authors in this section so the audience can grasp the power of these analyses.
ANWER: This information is already included in the supplementary material as the reviewer stated and we do not see the need to include a summary table as part of the main manuscript

COMMENT: Materials & Methods - When using Guppy for base calling, which mode/parameters were used? For example, was the mode "high-accuracy" used for both data types?
ANSWER: Parameters used are now reported in the text.

COMMENT: Materials & Methods - The authors claimed, "The free alternative open-source high performance base calling software Dorado can also be used to base call the reads". Do they mean that they use this software? If so, for which reads and what version?
ANSWER: This software was not used, but it is a request from Ocean Research Europe (ORC journal) to provide free available alternative to replicate analyses in the case the program used as a paying wall, as in our case. Thus, the software was included as a comparable suggestion to the one we used.

COMMENT: Materials & Methods - Remove the extra SbfI in the fourth line of the third paragraph of methods and include that the enzyme is at 10 U/uL.
ANSWER: done

COMMENT: Materials & Methods - Correct typo "Enzime" for "Enzyme" on line six of the third paragraph of methods.
ANSWER: done

COMMENT: Materials & Methods - If the authors did not use the 10X T4 DNA ligase buffer

from ThermoFisher, but an NEB buffer instead, please provide the catalog number from NEB of the buffer used.
ANSWER: done

COMMENT: Materials & Methods - The ATP typically used for libraries is at 10 mM, but here authors state that they used a 100 mM, could you please provide manufacturer and catalogue number?
ANSWER: done

COMMENT: Materials & Methods - What were the number of cycles and times for the shearing of DNA fragments in Covaris?
ANSWER: done

COMMENT: Materials & Methods - Correct typo “1,5” for “1.5” on line 18 of the first paragraph of page 6.
ANSWER: done

COMMENT: Materials & Methods - Revise that throughout the text, spaces should separate numbers and their units. There is inconsistency throughout.
ANSWER: Done

COMMENT: Materials & Methods - In the second line of the second paragraph of page six, the 5 uL is not per sample but per ligated pool, right? Also, please mention the primers used here, including the manufacturer. Revise spaces in the PCR profile.
ANSWER: After samples are pooled, all reagents used are reported per ligated pool. Primers and PCR profile have been included.

COMMENT: Materials & Methods - The document containing Fig S1 and S2 is not named “Extended Data” but “Supplementary Material.”
ANSWER: Document has been renamed

COMMENT: Materials & Methods - Please include the versions of some software under “Genome Assembly and Annotation”: TrimGalore, Fitlong2, BWA, HyPo, purge_dups, etc.
ANSWER: Done

COMMENT: Materials & Methods - What version of Piccard was used for deduplication?
ANSWER: Done

COMMENT: Materials & Methods - Missing a parenthesis closure when explaining the data set used to assess connectivity between spawning components.
ANSWER: corrected

COMMENT: Materials & Methods - Please include versions of the ADMIXTURE and R packages used.
ANSWER: Done

COMMENT: Materials & Methods - Please specify what subset or subsets of individuals were

used for pcadapat and its version.
ANSWER: Revised

COMMENT: Results - Table 2 has a column title “SCO1A Annotation” However, the genome assembly label is fScoSco3.1, what does SCO1A means in this context?
ANSWER: Removed SCO1A and left only Annotation

COMMENT: Results - There is one minor discrepancy between Table 1 and the ENA project information according to the accession number provided by the authors. Table 1 states a total number of sequences of 4,483, and ENA states 4,482. The total length in the table is 741,290,963 bp and in ENA 741,290,950 bp.
ANSWER: We thank the reviewer for highlighting the discrepancy. The correct stats are the ones in the ENA, thus Table 1 now reports those values. The difference is due to a 13 nt sequence that was removed when preparing the final assembly files.

COMMENT: Results - The use of GenomeScope for genome profiling included in FigS3 should be included in the methods and properly cited.
ANSWER: done

COMMENT: Results - In table S2, the thousandth separators under the column contigs should be a comma and not a quote symbol. Also, the “Coverage” column should indicate units, which I believe are “X.”
ANSWER: changed

COMMENT: Results - The title of Figure S4 should say that the colors match those from Figure 2, not 1.
ANSWER: changed

COMMENT: Results - Revise writing of sentence “FST values further validated this split, as all pairwise comparisons between the three groups resulted high with the highest FST value reported for MED vs NWA (FST=0.035)”. Also, the NEA vs MED should be FST=0.019 instead of 0.19.
ANSWER: revised

COMMENT: Results - The authors mentioned in the methods that the statistical significance of Fst values is assessed and adjusted, but these values are not presented in the results. Please include these values. They are particularly relevant for the Fst values presented within the Mediterranean Sea.
ANSWER: P-values of all Fst pairwise comparisons are now included in the Results as well as in Supplementary Tables.

COMMENT: Results - Remove the close parenthesis in the sentence: “This differentiation within the Mediterranean Sea is also seen in the FST pairwise comparisons, with FST 0.003 and 0.004 for the Adriatic Sea versus the Tyrrhenian Sea and Western Mediterranean Sea, respectively).”
ANSWER: done

COMMENT: Discussion - When discussing the origin of the Greenlandic/Islandic population, it is relevant to cite the original work of Jansen et al., 2016.
ANSWER: work cited

COMMENT: Discussion - In the paragraph: “Despite inter-population FST values being low (FST=0.014-0.035), these are in the range of what is usually observed in marine fish species with large population sizes since FST estimates are affected by both effective population sizes and migration rates, which may derive into insufficient statistical power provided by neutral markers to reject the panmictic hypothesis” I am not sure if this is arguing that the values observed are not statistically significant. Please clarify and make sure to include the p-values of the tests in the results.
ANSWER: All p-values are now included in the results.

COMMENT: Discussion -The paragraph: “Furthermore, the high rate of heterozygosity detected at genome level may reflect the presence of a large metapopulation, for which patterns of low diversity would persist even in the presence of uneven migration along its distribution…” results are confusing to me. High values of heterozygosity indicate high and not low genetic diversity. What the cited study argues is that in populations (not an individual), high levels of heterozygosity may be associated with large effective population sizes at a metapopulation level rather than very high local effective population sizes, but migration helps homogenize these values in the metapopulation. To further address this, the authors of this study should then present the values of heterozygosity observed globally and locally with the SNPs and address such hypotheses rather than making such relevant final remarks based on the heterozygosity of one individual for which the origin was not depicted. Finally, to draw conclusions on metapopulation models, I recommend that the authors test such hypotheses, perhaps by estimating the effective number of migrants between spawning components and the three established populations.
ANSWER: We have reworked this sentence to accommodate the reviewer´s comment.

Competing Interests: No competing interests were disclosed.