*Molecular Ecology*

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE OPEN ACCESS

# Resolving the Population Structure and Demographic History of the European Anchovy in the Northeast Atlantic: Tracking Historical and Contemporary Environmental Changes

José Martin Pujolar[1] | Courtney E. C. Gardiner[2] | Sophie von der Heyden[2] | Joana I. Robalo[3] | Rita Castilho[4,5] | Regina Cunha[4] | Romina Henriques[6] | Einar E. Nielsen[7]

[1]Centre for Gelatinous Plankton Ecology and Evolution, DTU Aqua – Technical University of Denmark, Kongens Lyngby, Denmark | [2]Department of Botany and Zoology, Stellenbosch University, Matieland, South Africa | [3]MARE – Marine and Environmental Sciences Centre, ARNET-Aquatic Research Network, ISPA Instituto Universitario, Lisbon, Portugal | [4]Centre of Marine Sciences (CCMAR/CIMAR LA), Universidade Do Algarve, Faro, Portugal | [5]Pattern Institute, Faro, Portugal | [6]Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa | [7]Section for Marine Living Resources, DTU Aqua – Technical University of Denmark, Silkeborg, Denmark

**Correspondence:** José Martin Pujolar (jmapu@aqua.dtu.dk) | Einar E. Nielsen (een@aqua.dtu.dk)

**Received:** 1 April 2025 | **Revised:** 15 May 2025 | **Accepted:** 24 May 2025

**Handling Editor:** J A H Benzie

**Funding:** This work was supported by BioDivClim Grant number 0156-00018B Innovation Fund Denmark (BiodivERsA call 2019–2020) and by the Department of Science and Innovation (DSI), South Africa for the project 'Biodiversity on the run: evolutionary and socio-economic consequences of shifting distribution ranges in commercially exploited marine fishes'. J.I.R. was supported by the FCT-Foundation for Science and Technology, projects MARE/UIDB/MAR/04292, MARE/UIDP/04292, and LA/P/0069/2020 granted to MARE and ARNET. R.C. and R.C. received Portuguese national funds from FCT-Foundation for Science and Technology through projects UIDB/04326/2020, UIDP/04326/2020, and LA/P/0101/2020.

**Keywords:** climate change | genetic diversity | marine ecosystems | population genomics | range expansion

## ABSTRACT

The spatial distribution of the European anchovy has expanded in the northern part of its range in the Northeast Atlantic in recent decades. However, whether this results from a northward range shift of southern conspecifics or the expansion of a local northern population is unknown. Using for the first time whole-genome sequencing, we explore current patterns of genetic diversity and population sub-structuring of European anchovy in the Northeast Atlantic, with special focus on recently expanded North Sea areas. Genomic data suggested three distinct groups: Northern (North Sea and Kattegat), Southern (Ireland and Central Portugal) and Cadis (South Portugal). Despite most of the genome being homogenised by high levels of gene flow characteristic of small pelagic fish, several large regions of high genetic differentiation were observed. This suggests that genomic population boundaries might be maintained by local adaptation within chromosome structural variants (inversions). Admixture analysis indicates that the ongoing northern range shift involves both migrants of southern origin and expansion of the local North Sea population. Historical demographic inference suggests that anchovies survived the last glacial period with small population sizes, followed by a split into the current Northern and Southern groups at the end of the last glacial maximum. The Southern group then expanded into the North Sea as the ice sheets retreated, in an expansion involving a large number of individuals, which is consistent with the retention of most of the genetic diversity. In comparison with other small pelagic fish, the genetic patterns found in anchovies (deeply divergent groups, no loss of genetic diversity during expansion, mixing between

Romina Henriques and Einar E. Nielsen should be considered joint senior author.

groups) align well with those found in European sprat, while sardines fit the pattern of expansion of a leading-edge population, with reduced genetic diversity and much shallower divergence between populations. This study contributes to a better understanding of population structure, range shifts and local adaptation in small pelagic fish under climate change, informing conservation and management efforts.

## 1 | Introduction

Small marine pelagic fish, including anchovies, sardines, herring and sprat, are key components of the functioning of marine ecosystems (Alheit and Peck 2019). Characterised by relatively short life-cycles, fast growth, large biomass and high mobility, they are a food source for many marine species and underpin major food webs (Cury et al. 2000; Palomera et al. 2007; Pikitch et al. 2014). Given their key trophic positioning, fluctuations in abundance and range of small pelagic fish populations can have significant ecological and socio-economic impacts (Alheit et al. 2009, 2012). Due to their high sensitivity to environmental conditions, collapses of small pelagic fish have been observed in all oceans, such as the regime shift observed in the Benguela Current ecosystem in the 1990s (Hutchings et al. 2009) or the collapse of the European anchovy in the Bay of Biscay in the 2000s (Uriarte et al. 2023). The high vagility of small pelagic fish through their life cycle facilitates gene flow over large geographical scales due to the lack of apparent barriers to dispersal (Palumbi 1994), which is expected to promote range shifts rather than local adaptation as a response to environmental changes. Associated with high gene flow, low levels of genetic structuring are expected, which might suggest shallow evolutionary trajectories and limited adaptive divergence (Nielsen et al. 2009). Despite this high dispersal potential, genetic studies on small pelagic fish have been found to show sharp breaks in genetic structure among populations, as seen in the recent whole-genome studies on Atlantic horse mackerel *Trachurus trachurus* (Fuentes-Pardo et al. 2023), European sardine *Sardina pilchardus* (Da Fonseca et al. 2024) or Pacific sardine *Sardinops sagax* in South Africa (Teske et al. 2021).

While most of the genome might be permeable to gene flow via hybridisation and introgression, large chromosome structural variants (SVs) can protect divergent haplotypes from recombination, thus facilitating reproductive isolation despite ongoing gene flow (Zhang et al. 2021). SVs, inversions in particular, often underlie ecologically important traits and contribute to local adaptation in many species (Jones et al. 2012; Ayala et al. 2013). In Atlantic cod *Gadus morhua*, chromosomal inversions comprising genes associated with low salinity adaptation have been identified as distinguishing features between migratory and nonmigratory ecotypes (Hemmer-Hansen et al. 2013; Berg et al. 2015, 2016, 2017; Barth et al. 2017). In Atlantic horse mackerel, a large putative chromosomal inversion underlies a latitudinal genetic cline (Fuentes-Pardo et al. 2023).

The European anchovy *Engraulis encrasicolus* is a short-lived (3–4 years) small pelagic fish widely distributed across the eastern Atlantic Ocean, the Mediterranean Sea and the Black Sea, where it is a major target of commercial fisheries (Mutalipassi et al. 2024). Off the Atlantic coast of Europe and Africa, the range of the species extends from the south of Norway to Namibia, where it mixes with the South African anchovy *E. capensis* (Van der Lingen et al. 2021). Recent observations indicate a northward range expansion of the European anchovy populations in the North Sea since the 1990s (Petitgas et al. 2012). Survey data show that anchovies are consistently caught in the entire North Sea and adjacent waters, such as the Irish Sea, Skagerrak, Kattegat and the Baltic Sea (Alheit et al. 2012). The range expansion is predicted to result from ongoing anthropogenic-mediated climate change (Raybaud et al. 2017) and can be explained by either a range shift of populations coming from the South or the expansion of remnant local North Sea populations (Alheit et al. 2012).

Population genetic studies on the European anchovy have been steadily conducted since the 1990s, with the Mediterranean Sea and the Bay of Biscay being the focus for most studies due to the regional importance of these fisheries (Magoulas et al. 2006; Sanz et al. 2008). Moreover, many studies have focussed on the occurrence of two distinct ecotypes, a marine ecotype found in open-sea habitats and an estuarine ecotype found inshore, which show high genetic differentiation (Le Moan et al. 2016; Montes et al. 2016). Previous genetic studies using Single Nucleotide Polymorphism (SNP) panels have shown a clear genetic boundary between the Bay of Biscay and the northern stock in the North Sea/English Channel (Zarraonaindia et al. 2012; Silva et al. 2014; Montes et al. 2016; Huret et al. 2020). The population structure along the Atlantic coast of the Iberian Peninsula is less clear since genetic differences might stem from the inclusion of estuarine populations (Zarraonaindia et al. 2009, 2012; Montes et al. 2016).

Our study is the first to use whole-genome sequencing to assess the genetic differentiation and diversity of the European anchovy from the Northeastern Atlantic, with particular focus on the northern expanded range of the species. Connectivity at the northern edge of the distribution has yet to be evaluated, especially in a context of climate change and shifting populations, which can affect fishery dynamics and the management of small pelagic fisheries. Historical population changes inferred from mitochondrial DNA show strong links with paleoceanographic conditions (Silva et al. 2014), suggesting that ongoing climate change is actively shaping the eco-evolutionary history of the species. By using millions of genome-wide SNPs, we inferred current patterns of population structure and admixture between regions and tested whether the northern range shift involves migrants of southern origin or the expansion of a local North Sea population. Using a coalescent approach, we reconstructed the past demographic history of the species including demographic changes and divergence times, which are discussed within an evolutionary biogeographic context. Finally, we tested the paradox that small pelagic fish with high dispersal capacity and ongoing gene flow may still undergo local adaptation due to reduced recombination within chromosome inversions.

## 2 | Materials and Methods

### 2.1 | Sampling

A total of 40 European anchovy were collected during 2021 as part of the Swedish, Danish, Irish and Portuguese national surveys. Samples broadly covered the distribution of the species in the Northeast Atlantic both north–southwards as well as east-(Table S1) westwards (Figure 1a). All individuals in our study correspond to the marine ecotype found in open-sea habitats. 'Swedish' individuals were taken from the Kattegat, at the exit of the Baltic Sea. 'Danish' individuals were taken across the North Sea, including Dogger Bank, German Bight and the West Coast of Denmark. 'Irish' individuals were taken from South and South-eastern Ireland in the Celtic Sea. 'Portuguese' individuals were taken from the Silver Coast and the Lisbon Coast in central Portugal and the Gulf of Cadis in South Portugal/Spain. Sampling details are summarised in Table S1.

### 2.2 | Sequencing

Muscle tissue was removed from each individual and stored at −80°C until DNA extraction. Genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit (using the manufacturer's protocol), with subsequent quality control via electrophoresis on an agarose gel and measurement of the DNA concentration using a Qubit fluorometer. DNA extracts were used to construct short-insert libraries with an insert size of ca. 350 bp. Paired-end sequencing libraries (PE150) were constructed and all 40 individuals were sequenced on a DNBseq platform at BGI, China.

### 2.3 | Data Filtering and SNP Discovery

Data quality of the reads was initially inspected using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx-toolkit). Next, adapters and low-quality bases were removed with Trimmomatic v0.36 (Bolger et al. 2014) using a minimum Phred score of 15. Subsequently, paired and unpaired reads were aligned separately to the European anchovy genome assembly IST_EnEncr_1.0 (NCBI Accession Number GCF_034702125.1) using BWA mem v0.7.1 (Li and Durbin 2009). Aligned data were pre-processed using the toolkit Picard v2.6.0 (http://broadinstitute.github.io/picard), which included merging, coordinate-sorting and removal of duplicate reads that may have arisen during library preparation due to PCR amplification.

Subsequently, a vcf file with all SNP variants for all 40 individuals was constructed using the mpileup and call commands in BCFtools v1.9 (Li et al. 2009). We excluded sites for which average read depth was less than 10 or more than 100, quality was

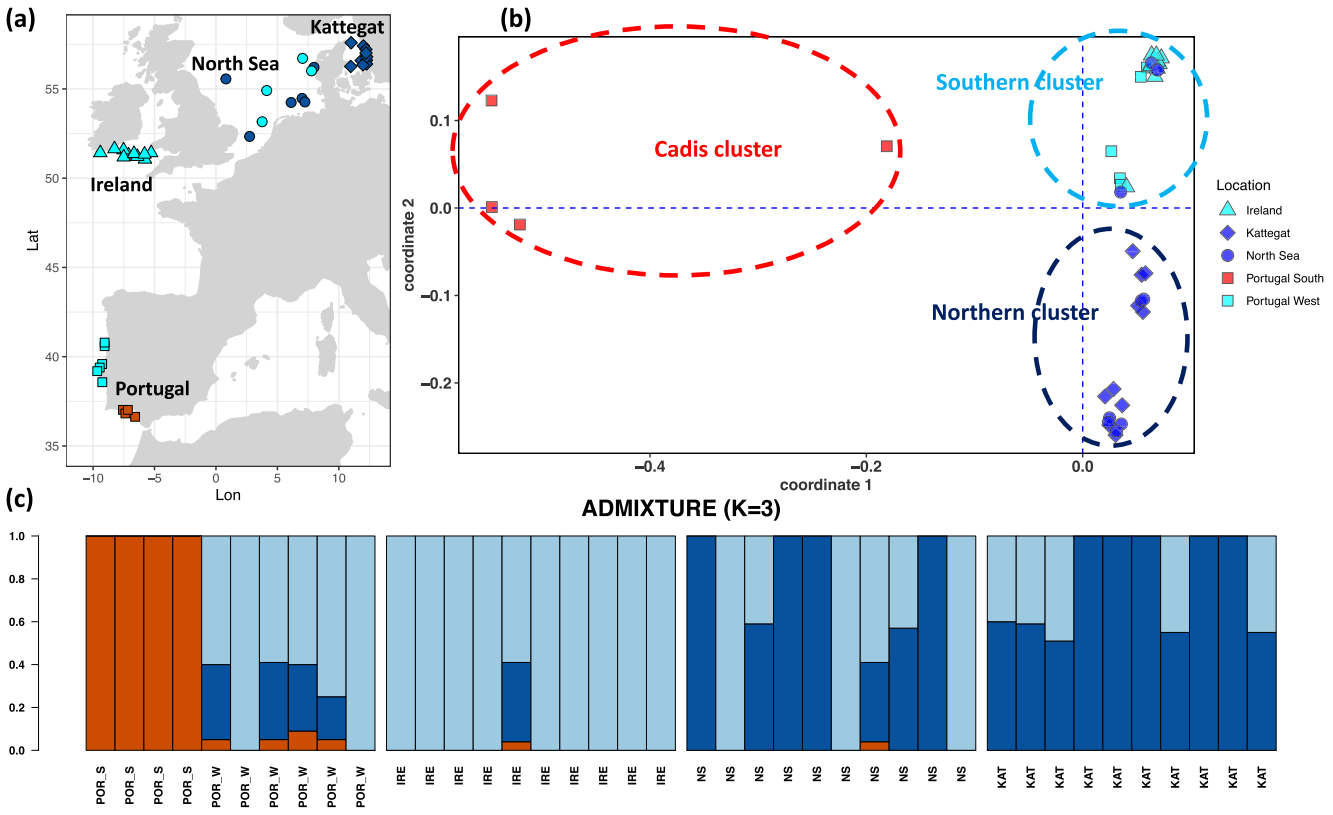


**FIGURE 1** | (a) Sampling locations of European anchovy including distribution of Northern (dark blue), Southern (light blue) and Cadis (red) clusters. For sampling details see Table S1. (b) Visualisation of population structure using a Principal Component Analysus (PCA) of all 40 individuals clustering in three separate clusters: Northern (dark blue), Southern (light blue) and Cadis (red). (c) ADMIXTURE analysis with most likely scenario of three clusters ($K = 3$). Individual admixture proportions are shown for each individual. Locations: IR_S, Ireland; KAT, Kattegat; NS, North Sea; POR_S, Portugal-South; POR_W, Portugal-West.

below 20 and in proximity to indels. As a final step, SNP variants were filtered using vcftools v0.1.14 (Danecek et al. 2011) so that only biallelic SNPs present in all 40 individuals sequenced were retained.

### 2.4 | Genome-Wide Diversity and Population Structure Analyses

All genomic diversity and population structure analyses were conducted using a filtered SNP data set after linkage disequilibrium-based variant pruning implemented in PLINK v1.9 (Purcell et al. 2007). The indep-pairwise option was used with a window size of 100, a variant count of 5 to switch window at the end of each step, and a $r^2$ threshold of 0.5. Levels of genetic diversity were assessed using Nei and Li (1979) nucleotide diversity ($\Pi$) and individual observed heterozygosity ($H_o$) calculated using vcftools v0.1.14 (Danecek et al. 2011). Standardised genetic differentiation ($F_{ST}$) statistics between population pairs were calculated using vcftools v0.1.14 following Weir and Cockerham (1984). In order to assess the significance of pairwise $F_{ST}$ values, we carried out bootstrapping over loci to generate a confidence interval around the observed $F_{ST}$. $p$ values were calculated using a one-sample $t$-test for each pairwise $F_{ST}$.

Population structure was initially explored with a Principal Component Analysis (PCA) in smartPCA from the Eigensoft package (Patterson et al. 2006). Population structure was further investigated using ADMIXTURE (Alexander and Lange 2011), which uses a clustering algorithm to infer the most likely number of groups ($K$) and then calculates the proportion of ancestry components (admixture proportion) for each individual. The analysis was performed with $K = 1$–6. For each $K$, 10 independent runs were conducted to check the consistency of the results. Following the software guidelines, the cross-validation (CV) procedure was applied to choose the best $K$ for the model.

### 2.5 | Tests for Selection

The classic $F_{ST}$-based genome-scan method was used to detect signatures of positive selection and uncover local adaptation. The method is based on locus-specific $F_{ST}$ estimates averaged across genome regions, with regions with very high $F_{ST}$ compared to the rest of the genome suspected to be under strong local adaptation. Genome-wide $F_{ST}$ was estimated using VCFtools v0.1.14 (Danecek et al. 2011) with the unpruned data set by calculating a sliding-window $F_{ST}$ with a window size of 10 kb and a step size of 10 kb. Several combinations of window and step sizes were tested (From 5 to 50 kb), monitoring SNP counts per window. We settled on a 10-kb window with a 10-kb step size, which had sufficient SNP density to provide reliable $F_{ST}$ estimates and the size was large enough to encompass regions where linkage disequilibrium has decayed, allowing for the detection of independent loci. A minimum of 10 SNPs per window was chosen to make the $F_{ST}$ estimate reliable. Windows with average window-$F_{ST}$ values above the 98th percentile of the empirical distribution were considered as candidate regions putatively under selection. Sliding-window $F_{ST}$ along chromosomes was visualised with a Manhattan plot using CMPlot in R.

Based on the list of candidate regions under selection, we localised all corresponding SNPs and genes. A SNP was considered to be located within a gene when found in exonic or intronic regions or within 10 kb of a gene, which is based on the estimated decay of linkage disequilibrium in fish (Hemmer-Hansen et al. 2014). Genes were then annotated using the gene predictions for the European anchovy genome IST_EnEncr_1.0 available on NCB. Gene descriptions were obtained from the Zebrafish Information Network (ZFIN) database (http://zfin.org), together with human orthologs, which were used in the GeneCards human gene database (http://genecards.org) to retrieve the NCBI gene summary, the UniProtKB/Swiss-Prot summary as well as the Ensembl Gene IDs. Functional interpretation of the set of candidate genes was conducted using Gene Ontology (GO) term classification and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis implemented in DAVID v6.8 (Dennis et al. 2003). Zebrafish was used as reference and standard settings of gene count and ease in DAVID were used. Functional Gene Ontology terms of the candidate genes were visualised with a Lollipop chart showing the number of genes involved in enriched GO terms on the x-axis using R. GO terms were further analysed with the ClueGo plug-in of Cytoscape v3.8.2 (Bindea et al. 2009), which visualises the non-redundant biological terms for large clusters of genes in a functionally grouped network. Analyses were conducted using the Ensembl gene IDs, only considering GO terms with corrected $p$ values.

### 2.6 | Demographic History Inference

Coalescent methods can be applied to infer past population dynamics and demographic changes over historical time scales with high resolution. We inferred the demographic history of the European anchovy populations in our study using the Pairwise Sequentially Markovian Coalescent (PSMC) method (Li and Durbin 2011). PSMC extracts demographic information about a given population/species from the distribution of heterozygous sites across the genome and infers the distribution of time since the most recent common ancestor (TMRCA) between each pair of alleles at all loci across the whole genome of one single individual. This provides information about changes in effective population sizes over time, hence allowing us to assess responses to paleo-oceanographic conditions.

For all re-sequenced individuals, a consensus sequence was generated with BCFtools v1.9, following the same approach as Nadachowska-Brzyska et al. (2015). First, variants were called for each individual separately with the mpileup and call commands. Subsequently, the consensus command was used to create a consensus sequence. The resulting vcf file was converted to the psmc input format with vcfutils.pl (distributed with BCFtools). Next, the consensus sequence was divided into non-overlapping 10-bp bins, with bins scored as either heterozygous or homozygous. Then, PSMC v0.6.5 (Li and Durbin 2011) was run with 25 iterations, $T_{max}$ ($-t$) of 15, initial mutation/recombination ratio ($-r$) of 5 and time bin parameter ($-$p) set to the standard '4 + 25*2 + 4 + 6'. For parameter conversion, results were scaled using a generation time of 1 year and a conservative average mutation rate of $1 \times 10^{-8}$ per nucleotide per generation (Nikolic et al. 2020).

## 3 | Results

### 3.1 | Genome-Wide Diversity and Patterns of Population Structure

We conducted whole-genome re-sequencing on 40 European anchovy individuals from four regions along the Atlantic coast of Europe. On average, 36.2 million reads of 90 bp per individual were generated (Table 1). After quality filtering, on average 10.8% were eliminated. Retained sequences (89.2%) have a mean depth of 15.1× (9.8–20.3×) and a mean quality score of 36.1. On average, 98.4% of paired sequences and 96.6% of unpaired sequences aligned to the European anchovy genome, with a mean mapping quality of 24.9. After mapping, a total of 110.5 million SNPs were obtained. After filtering for minimum depth, quality and biallelic SNPs genotyped in all individuals, a total of 30,262,784 SNPs was obtained (unpruned data set). After filtering for linkage disequilibrium, a final data set of 1,073,507 high-quality SNPs was retained for a pruned data set that was used in all downstream analyses except the tests for selection, which required the unpruned data set.

Genome-wide diversity indices were similar across the four regions studied (Table S2), including nucleotide diversity ($\Pi = 0.236–0.246$) and heterozygosity ($H_o = 0.248–0.264$). The results of the PCA based on the whole-genome pruned SNP data set showed that individuals clustered into three major clades (Figure 1b). The first principal component (PC1) distinguished the first cluster, which consisted of all individuals from the Gulf of Cadis in South Portugal ('Cadis'). The second principal component (PC2) separated 'Northern' and 'Southern' cluster. All individuals from Kattegat were part of the Northern cluster, while all individuals from Ireland and Central Portugal were part of the Southern cluster. By contrast, individuals from the North Sea appeared in both clusters. Within each cluster, some individuals were intermediate between clusters, pointing to possible admixture between populations. Calculation of genetic differentiation between clusters showed a high and significant $F_{ST}$ between Cadis-Northern ($F_{ST} = 0.055$, $p$ value $< 0.05$) and between Cadis-Southern ($F_{ST} = 0.054$, $p$ value $< 0.05$). A lower but significant $F_{ST}$ value of 0.012 was found between the Northern and Southern clusters (Table 2), which was higher when excluding intermediate individuals ($F_{ST} = 0.021$, $p$ value $< 0.05$). Re-examination of genome-wide diversity levels taking into account the groups observed in the PCA showed a higher diversity in the Southern cluster relative to the Northern cluster (Table S2), in terms of both nucleotide diversity (Southern: $\Pi = 0.240$; Northern: $\Pi = 0.234$) and heterozygosity (Southern: $H_o = 0.263$; Northern: $H_o = 0.249$). The Cadis cluster showed the lowest values overall ($\Pi = 0.227$; $H_o = 0.241$). However, none of the comparisons were significantly different from each other.

When exploring the overall number of genetic groups present, ADMIXTURE suggested a scenario with three clusters ($K = 3$) as the most likely (Figure 1c). The three clusters suggested by admixture analysis supported the PCA results, with a Cadis, a Southern and a Northern cluster. Individuals from South Portugal were all pure individuals of the Cadis cluster, while individuals from Central Portugal belonged to the Southern cluster (2 pure, 4 admixed). Individuals from Ireland all belonged to the Southern cluster (9 pure, 1 admixed), while individuals from Kattegat all belonged to the Northern cluster (5 pure, 5 admixed). The North Sea was the most mixed region with six individuals belonging to the Northern cluster (4 pure, 2 admixed) and four belonging to the Southern cluster (3 pure, 1 admixed).

### 3.2 | Detection of Selective Sweeps

We detected strong signatures of selection in the genome based on sliding-window $F_{ST}$. When comparing the Northern and Southern clusters, regions with high $F_{ST}$ were apparent (Figure 2). Rather than being found all along the genome, high $F_{ST}$ windows were restricted to five chromosomes: 3, 4, 6, 18 and 19. Particularly for chromosome 3, 30.8% of windows showed $F_{ST}$ higher than the 98% percentile and 16.5% higher than the 99% percentile. At chromosome 4, 10.4% of windows showed $F_{ST}$ higher than the 98% percentile and at chromosome 6 the percentage was 4.2%. These regions or structural variants with $F_{ST}$ higher than the 98% percentile encompassed most of chromosome 3 (between 10.4 and 43.9 Mbp) and chromosome 6 (between 3.5 and 51.9 Mbp), while encompassing only the second half of chromosome 4 (between 29.5 and 58.4 Mbp; Figure 2).

When comparing the Cadis cluster versus the Northern (Figure S1a) and the Southern (Figure S1b) clusters, many regions with high $F_{ST}$ were also apparent, generally more spread across the genome. Still, some apparent structural variants were observed in both comparisons at chromosomes 2, 3, 6, 13, 16 and 18.

For the rest of the selection analysis, we focussed on the comparison of the Northern and Southern clusters. A total of 981 genes were located in the windows with $F_{ST}$ higher than the 98% percentile. Out of these, 633 were fully characterised and were selected as candidate genes for local adaptation for GO analyses (Figure 3). Those included ATP production, reproduction, transmembrane transporter activity, integrated component of membrane, Golgi apparatus and vesicle trafficking, DNA repair and chromosomal DNA replication, WNT signalling pathway and circadian rhythms among many others. A list of all annotated genes putatively under selection including gene description in the Zfin database, human orthologs and summaries from Entrez and UniProtKB/Swiss-Prot is provided in File S1.

### 3.3 | Demographic Inference Using PSMC

We had data of sufficient resolution to infer for the first time the demographic history of European anchovy using PSMC. The analysis revealed that the three clusters inferred by PCA/ADMIXTURE (Northern, Southern and Cadis) share common ancestry encompassing the last 3 million years, as they converge into one single ancestral population (Figure 4). Initial effective population sizes were estimated between 100,000 and 200,000 individuals and remained relatively stable for most of the early and mid-Pleistocene. The split of the Cadis cluster occurred ca. 250,000 years ago during the late-Pleistocene. After this first split, all clusters remained with relatively low population sizes until ca. 20,000 years ago. At this point, both the Cadis cluster and the Northern/Southern cluster started a period of steady demographic expansion. The split between the Northern and Southern clusters occurred ca.

**TABLE 1** | Statistics describing the distribution of different properties of whole-genome sequences after each step of filtering and alignment to the European anchovy genome assembly using FastQC, Trimmomatic, Bowtie and Samtools.

| **FastQC** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Quality_1.fq** | | | | | | **Quality_2.fq** | | | | | |
| ***N* reads** | **Median** | **Mean** | **10% Percentile** | **25% Percentile** | **75% Percentile** | **90% Percentile** | **Median** | **Mean** | **10% Percentile** | **25% Percentile** | **75% Percentile** | **90% Percentile** |
| 72,312,241 | 37 | 35.01 | 38.55 | 37.47 | 35.35 | 29.60 | 36.91 | 34.04 | 37.93 | 37.47 | 34.05 | 26.57 |
| **Trimmomatic** | | | | | | | | | | | | |
| ***N* reads** | **Both surviving** | | **%** | **Forward only surviving** | | **%** | **Reverse only surviving** | | **%** | **Retained** | **%** | **Dropped** | **%** |
| 72,312,241 | 50,971,323 | | 70.41 | 7,822,863 | | 10.80 | 5,756,609 | | 8.01 | 64,550,794 | 79.20 | 7,771,111 | 10.80 |
| **BWA** | | | | | | | | | | | | |
| ***N* paired reads** | | | ***N* mapped** | | **%** | | ***N* unpaired reads** | | | ***N* mapped** | | **%** | |
| 105,991,129 | | | 104,274,470 | | 98.38 | | 13,583,870 | | | 13,129,314 | | 96.62 | |
| **Samtools** | | | | | | | | | | | | |
| **Coverage** | | | | **Mean depth** | | | **Mean base *Q*** | | | | **Mean map *Q*** | | |
| 57.27 | | | | 15.07 | | | 36.05 | | | | 14.94 | | |

15,000 years ago, after which the Southern cluster continued to increase moderately, while the Northern cluster showed a slight decline. While the PSMC is less accurate for recent periods, it is apparent that the Southern cluster retained a larger effective population size over time than the Northern cluster.

## 4 | Discussion

### 4.1 | Identification of Three European Anchovy Groups in the Northeast Atlantic

Characterising the population structure of small pelagic fish, which support an array of associated biodiversity as well as human fisheries, is vital in order to better understand the impact of climate change (ICES 2012). This is particularly important for systems where species are shifting their ranges in response to environmental changes, such as temperature (Alheit et al. 2007, 2012). Whole-genome sequencing data of European anchovy individuals in the Northeast Atlantic suggest three distinct groups: a Northern cluster including all individuals from the Kattegat and most individuals from the North Sea; a Southern cluster including all individuals from Ireland, Central Portugal and few from the North Sea; and a Cadis cluster including all individuals from South Iberia. The Northern cluster likely represents the population that spawns in the German Bight and the German Wadden Sea in the southern North Sea from June to August with a peak in July (Alheit et al. 2007). The Southern cluster likely represents the population that spawns in the Bay of Biscay from April to August with a peak in May–June (Motos 1996). The Cadis cluster appears to spawn locally in the Gulf of Cadis from early spring to early autumn with a peak in July–August (Baldó et al. 2006). The clear genetic divergence between Northern and Southern clusters is consistent with previous studies using SNP panels (Zarraonaindia et al. 2012; Silva et al. 2014; Montes et al. 2016; Huret et al. 2020) and microsatellites (Van der Kooij et al. 2024), showing significant genetic differentiation between samples from the Bay of Biscay and the North Sea. Interestingly, while whole-genome data place samples from the Celtic Sea within the Southern cluster, samples from the adjacent Irish Sea and English Channel are part of the Northern cluster, according to SNP (Zarraonaindia et al. 2012; Montes et al. 2016) and microsatellite data (Van der Kooij et al. 2024). Future studies could use whole-genome sequencing to assess individual levels of introgression in these regions, given the high levels of admixture observed in the southern North Sea in our study (see below). Additionally, whole-genome sequencing efforts should cover the unsampled area in our study between the Celtic Sea and North Portugal, specifically the Bay of Biscay and the Cantabric Sea, to confirm the genetic cohesion reported here. On the other hand, our study is the first to clearly delineate the genetic boundary between the Southern and Cadis clusters, as previous genetic studies

**TABLE 2** | Pairwise genetic differentiation between the Northern, Southern and Cadis clusters.

| | Northern | Southern | Cadis |
|---|---|---|---|
| Northern | ***** | | |
| Southern | 0.012* | ***** | |
| Cadis | 0.055* | 0.054* | ***** |

*Note:* Statistical significance: *$p < 0.05$.

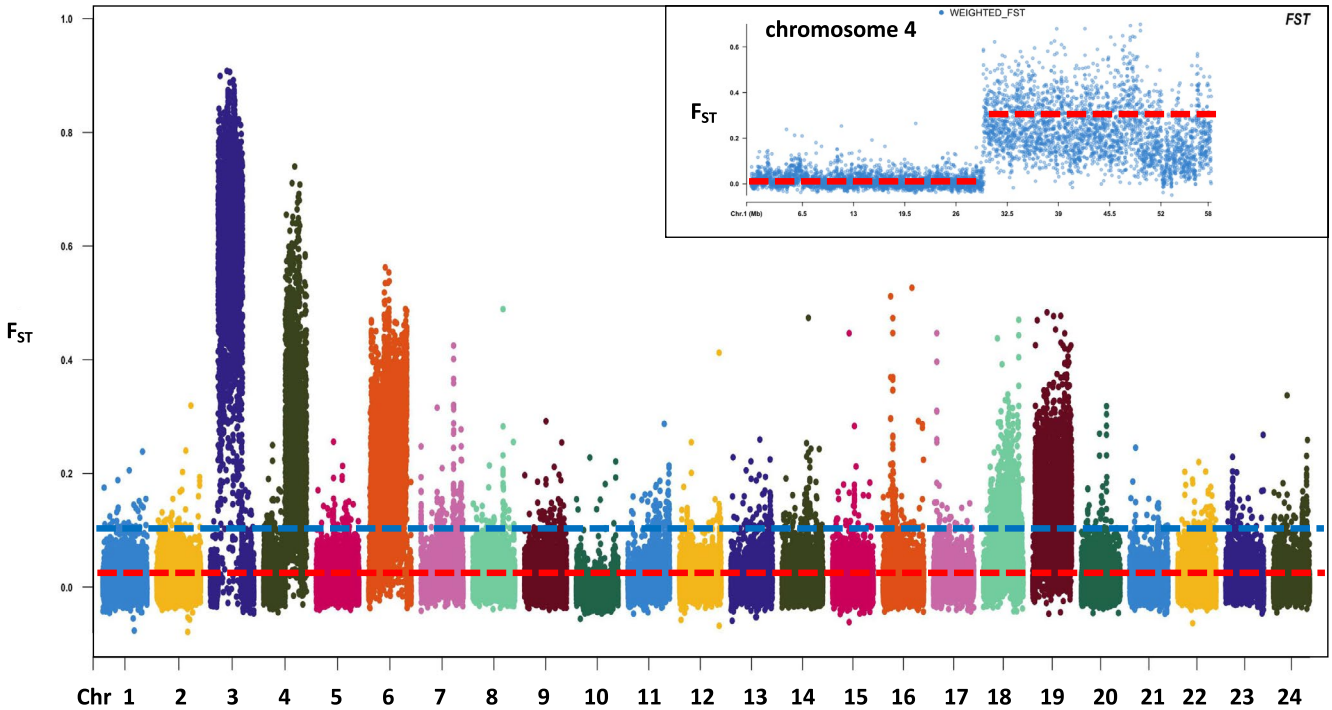


**FIGURE 2** | Manhattan plot based on sliding-window genetic differentiation FST comparing the Northern and Southern clusters. Inner panel shows detailed plot for chromosome 4. Regions above the 98% percentile of the empirical distribution are marked with discontinuous blue line. The average FST is represented by a discontinuous red line.

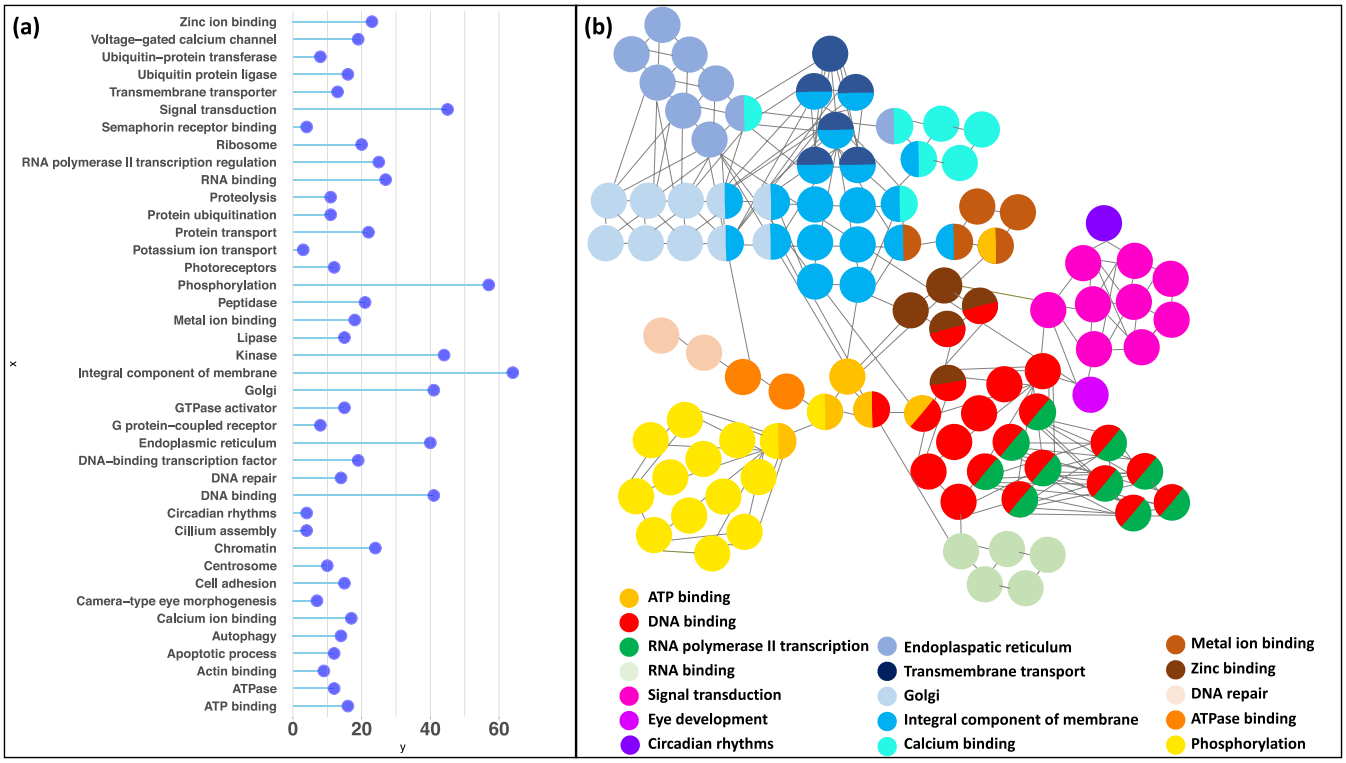


**FIGURE 3** | Visualisation of the ontology terms detected during Gene Ontology and KEGG analysis when comparing the Northern and Southern European anchovy clusters, including biological processes, cellular components and metabolic and signalling pathways. (a) Lollipop chart with number of genes involved in enriched GO terms on the *x*-axis and (b) ClueGO network for a better biological interpretation of functionally grouped genes. Analyses were conducted using the zebrafish Ensembl Gene IDs, only considering GO terms with corrected *p* values < 0.05 and selecting the major significant GO term as the representation of the group.

in the region focused on narrow-shelf populations (Zarraonaindia et al. 2012; Montes et al. 2016).

The finding of distinct genetic clusters in our study points to strong spawning site fidelity in European anchovy. Philopatry allows fish to place their offspring in the same habitat and environmental conditions of the parental generation, ensuring larval survival and reproductive success. However, the low (but albeit significant) genetic differentiation found suggests ongoing gene flow between groups. Population dynamics patterns show that in spring–summer during the spawning season, anchovies present a restricted distribution, limited to the Bay of Biscay and the southern North Sea (Huret et al. 2020). By contrast, in autumn-winter, anchovies show a continuous distribution and are also found in the Celtic Sea, Irish Sea, English Channel, and the entire North Sea. Overlapping distribution in overwintering habitats gives ample opportunity for both groups to mix in autumn-winter. As a result, some individuals are expected to cross over and spawn occasionally with the other spawning group. Ongoing gene flow, as suggested by the high number of introgressed individuals found in our study, would explain the low genetic differentiation found.

Our results highlight the North Sea as a mixing zone for anchovy populations. Interestingly, recent studies have reported a northward expansion of anchovies since the 1990s, extending into the entire North Sea and even the Baltic Sea (Alheit et al. 2012). Our data suggest that individuals found in these newly occupied areas primarily originate from the local population spawning in the southern North Sea. For example, individuals collected in Kattegat in our study, where survey data indicate anchovies were missing in the 1970s and 1980s but present from the mid-1990s onward (Alheit et al. 2012), were all assigned to the Northern cluster. However, 40% of North Sea individuals were assigned to the Southern cluster, indicating that some migrant individuals of southern origin move northward into the North Sea during winter. Genetically pure southern individuals were found in the Dutch Dogger Bank, the central North Sea and the Danish West Coast. Hence, genetic data suggest a mix of migrants and local North Sea populations in newly occupied habitats. As the North Sea warms due to ongoing climate change, our results suggest that the Southern population is expanding into previously unsuitable habitats, tracking evolving environmental conditions.

In contrast to the North Sea, population mixing is much lower in Portugal, where the Cadis cluster is genetically more distinct, showing 5% genetic differentiation and no signs of admixture. This suggests Cape San Vicente in South Portugal as the limit between the two groups, which would act as a hard barrier separating West and South Portuguese populations. The west coast of Portugal typically has low anchovy abundances with occasional population surges, whereas the south coast is characterised by more stable large abundances (ICES 2018). Fisheries data indicate that the Gulf of Cadis hosts a self-sustained population, with the main spawning near the Guadalquivir River mouth in Spain (Baldó et al. 2006). Morphological studies also showed a gradual shift in head-to-body ratio along the western coast and the Gulf of Cadis (Caneco et al. 2004). Lagrangian simulations

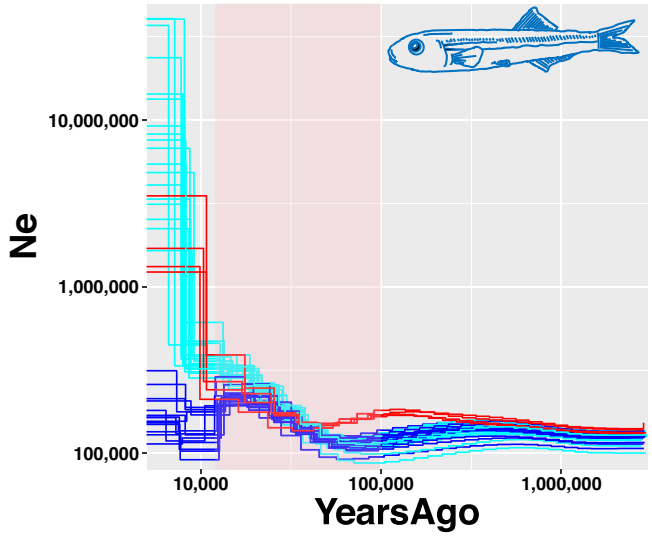


**FIGURE 4** | Demographic history reconstruction of European anchovy populations using the PSMC (Pairwise Sequentially Markovian Coalescent) method. A universal mutation rate of $1 \times 10^{-8}$ and generation time of 1 year were used. All individuals were included in the analysis and each PSMC curve corresponds to a single individual. The three clusters are colour-coded: Northern (dark Blue), Southern (light blue) and Cadis (red). The shaded area indicates the Last Glacial Period (LGP).

suggest that dispersal north from the Gulf of Cadis to the western Portuguese coast is possible but affected by low probability of early stages survival (Casaucao et al. 2021). Collectively, fisheries data, population dynamics and morphological evidence all support our genetic findings that there is little to no gene flow between South and West Portugal.

### 4.2 | Structural Variants Facilitating Local Adaptation Despite Gene Flow

Chromosomal inversions between populations may facilitate local adaptation despite opportunities for gene flow (Wellenreuther and Bernatchez 2018). Spatially diversifying selection favours the build-up of linkage disequilibrium between genes involved in local adaptation and reproductive isolation. However, gene flow and recombination can rapidly break these associations (Coyne and Orr 2004; Schluter and Rieseberg 2022), unless divergent haplotypes are protected from recombination by structural variants (Zhang et al. 2021).

Our tests of selection revealed extensive areas of high genetic differentiation concentrated in specific parts of the genome, likely representing structural variants. Those regions encompass genes under local adaptation that characterise and discriminate the three population clusters (Northern, Southern and Cadis) despite gene flow. It is apparent in our study that most of the genome appears permeable to introgression, as deduced by the low overall genetic differentiation and the widespread presence of admixed individuals, ranging from 10% in Ireland to 50% in Kattegat and Central Portugal. However, certain genomic regions appear to be impermeable and may act as barriers to gene flow. We hypothesise that the 'genomic islands of differentiation' on most of chromosome 3, the second half of chromosome 4 and most of chromosome 6 are putative structural variants, likely chromosomal inversions. This aligns with a well-documented paradox in small pelagic species, where high dispersal capacity and associated gene flow should theoretically prevent local adaptation, yet structural variants allow adaptive variants to persist by preventing recombination (Zhang et al. 2021). Exceptionally large chromosomal regions suggesting suppressed recombination consistent with megabase inversions/supergenes have been detected in Atlantic herring (Han et al. 2020), king scallop (Hollenbeck et al. 2022) and European sprat (Pettersson et al. 2024), contributing to environmental adaptation.

Putative candidate genes under selection were found for a wide array of traits, including energy production, circadian rhythmicity and reproduction. For example, the gene ATP5F1D (ATP synthase F1 subunit delta) encodes the delta subunit of the catalytic core of the mitochondrial ATP synthase that produces energy for use in many cellular processes. Differences in allele frequencies between populations at this gene might reflect local adaptation to different metabolic demands linked to different environmental conditions such as water temperature or oxygen concentration. Average sea surface temperature data from winter 1990 to 2011 in the North East Atlantic (Jansen et al. 2012) reveal significant regional variation, ranging from 10°C to 12°C in the Celtic Sea to 3°C–7°C in the Wadden Sea in the southern North Sea and as low as 1°C–2°C in the Kattegat. Populations in colder waters such as the North Sea and Kattegat could benefit from ATP5F1D alleles associated with enhanced energy production and higher metabolic rates. This suggests divergent selection across temperature gradients, consistent with local adaptation in energy metabolism.

Interestingly, genes from the circadian clock pathway appeared to be under selection. Circadian rhythms are internal biological processes that maintain approximately a 24-h periodicity and regulate an array of activities such as feeding or spawning. The adaptive significance of circadian rhythms is to synchronise physiological and behavioural processes to occur at advantageous times and better capitalise on the available resources (Sharma 2003). We identified footprints of local adaptation at the CLOCK gene, which controls the persistence and periodicity of circadian rhythms, and CRY1, which regulates the circadian clock network (Hardin 2000). These genes may be under local selection due to differences in local photoperiod associated with the difference of 10° in latitude and 10° in longitude between the spawning site in the Bay of Biscay (45° N, 3° W) and the spawning site in the North Sea (55° N, 7° E). Circadian rhythm genes contribute to optimal reproductive performance, and the CLOCK gene has been associated with the timing of reproduction in a number of species (Boden and Kennaway 2006). Reproductive genes were also putatively under selection, including BSG (basigin) and TSSK6 (Testis-specific serine kinase 6), both located on chromosome 6. Both genes play essential roles in fertilisation success, with a role in sperm-egg fusion (Harada et al. 2013). Selection on these genes might reflect local adaptation to the particular environmental conditions of the spawning sites in the North Sea and Gulf of Biscay, with notable differences in temperature, salinity and oceanographic conditions. The southern

North Sea is characterised by colder temperatures and low salinities around 30 PSU, while the Bay of Biscay is characterised by warmer, more saline > 35 PSU and more stable environments (ICES 2022, 2023). Interestingly, the transcriptomic study of Montes et al. (2016) also identified BSG and TSSK6 to be under putative selection when comparing marine and coastal anchovy populations in the Bay of Biscay, further supporting their adaptive significance.

An alternative explanation for the highly differentiated markers being under ecologically based divergent selection could be intrinsic genetic incompatibilities (Bierne et al. 2011). According to the coupling hypothesis, increased differentiation is not driven by local adaptation to different environments but by negative interactions between genes within the genome that do not depend on the environment. Intrinsic genomic incompatibilities act as endogenous barriers to gene flow, even in the absence of strong ecological differences.

Given that large structural variants in other fish species have been linked to salinity adaptation (Hemmer-Hansen et al. 2013; Berg et al. 2016; Barth et al. 2017), we tested whether the structural variants identified in our study might be associated with osmoregulation and salinity adaptation. In our studied area in the eastern Atlantic, salinity is highly variable, ranging from values > 37 PSU in the Gulf of Cadis and > 35 PSU in the Celtic Sea to lower salinities around 30 PSU in the southern North Sea and between 18 and 26 PSU in the Kattegat (ICES 2022, 2023; Pacariz et al. 2014). We searched for key genes involved in these functions, including the sodium pump $Na^+/K^+$ ATPase (ATP1A1), $Na^+H^+$ antiporter (NHE), V-type $H^+$ ATPase (VHA), carbonic anhydrase and acid-sensing ion channels (Stern and Lee 2020). However, none of these genes were located within the structural variants or showed signatures of selection in our data. Further research is needed to determine whether the identified structural variants play a role in salinity adaptation.

### 4.3 | Demographic History in the Northeast Atlantic

We carried out the first reconstruction of the demographic history of the European anchovy in the Northeast Atlantic using the PSMC coalescent approach. The exact timing of events should be interpreted with caution, given that the use of a universal mutation rate and factors such as gene flow, introgression and selection can bias the estimates of population sizes and divergence times (Leaché et al. 2014; Adams et al. 2018).

The three groups detected in our study (Northern, Southern and Cadis) share a common ancestor dating back as far as 3 million years ago (MYA) in the Pliocene. During the early Pleistocene (from 2.58 MYA) the ancestral population remained low but stable, without major expansions or declines. The split between the Cadis group and the Northern/Southern group appears to have occurred ca. 250,000 years ago during the 'Quaternary Glaciation'. This period was characterised by alternating glacial and interglacial periods, where ice sheets expanded and contracted (Hewitt 2004). During such periods, the Cadis group might have become isolated for an extended period, leading to the accumulation of genetic differences.

The results also suggest that the European anchovy populations survived with low population sizes during the Last Glacial Period (LGP), from ca. 120,000 to 12,000 years ago, when massive ice sheets covered large parts of the Northern Hemisphere, including Scandinavia, the British Isles and parts of mainland Europe (Svendsen et al. 2004). Coinciding with the end of the LGP, anchovies split into the current Northern and Southern groups. Many temperate species survived the Last Glacial Maximum (LGM) in ice-free southern European refugia along the northern Atlantic coast of the Iberian Peninsula, Southern France, Italy and the Balkans (Hewitt 1996; Taberlet et al. 1998). This suggests that the warm water-adapted European anchovy may have been confined to refugia during glacial maxima as seen in many other temperate species (Stewart et al. 2010). It is feasible that during the LGM the Northern/Southern group could have settled around the Iberian Peninsula, even possibly spawning in the Bay of Biscay, as many fish species exhibit strong fidelity to ancestral spawning sites (i.e., Atlantic cod, Skjæraasen et al. 2011; European eel, Miller et al. 2019). Towards the end of the LGM ca. 20,000 years ago, the Irish, English, North Sea and Fenno-Scandian ice sheets retreated and new favourable habitats became available for the European anchovy. At this timepoint, the size of all genetic groups increased. We hypothesise that a subset of the ancestral Southern/Northern group (possibly surviving the LGM in an in situ glacial refugium in northern Spain) expanded into the North Sea, which had become ice-free, establishing a new spawning population. Soon after the population expansion, the two groups split at ca. 15,000 years ago.

Following this split, the Southern group appeared to experience a notable population size increase ca. 10,000 years ago, which might reflect a further range expansion (e.g., into the Celtic Sea). A similar pattern was observed for the Cadis population, while the North Sea population remained relatively stable after the split. From this point onward, PSMC analyses become less accurate for estimating population sizes, as evidenced by the large variability of population size trajectories among individuals within groups. Nevertheless, it is apparent that the Southern group has a larger population size than the Northern group. This is concordant with fisheries and population dynamics data, as the Bay of Biscay hosts the largest aggregation of European anchovy in the Northeast Atlantic (Uriarte et al. 2023).

### 4.4 | A Northern Expansion After the LGM With no Significant Loss of Genetic Diversity

Comparison of genome-wide diversity using millions of high-quality SNPs reveals similar levels of genetic diversity in the Northern and Southern clusters. This indicates that most of the genetic diversity was retained during the expansion into the North Sea at the end of the LGM. This pattern is consistent with a large expansion event driven by a large number of individuals in a relatively short period of time.

Small pelagic clupeoid fish including anchovies, sardines, herring or sprat are characterised by the ability to respond quickly to environmental changes because of their high reproductive plasticity as they can rapidly alter their reproductive traits when environmental conditions require it (Alheit 1989). Moreover, their extreme vagile nature allows them to travel long distances

and facilitates range shifts in their geographical distribution, such as the northward expansion of European anchovy after the LGM inferred by the PSMC analysis. A large number of founders, as suggested from our genetic data for the North Sea population, is likely crucial in preventing inbreeding associated with low population sizes and providing sufficient standing genetic variation for natural selection to act upon. This is particularly relevant for the expansion of European anchovy in the North Sea, where environmental conditions differ significantly from their original range, particularly in terms of temperature. The European anchovy is a warm water-adapted fish, as seen in the Mediterranean Sea, where spawning occurs exclusively in summer. The cooler North Sea waters are less favourable for spawning (Alheit et al. 2007). Nevertheless, anchovies have continued their northward expansion to the Skagerrak–Kattegat region and the Baltic Sea (Alheit et al. 2012). In the southern areas around the Bay of Biscay, surface salinities over $33‰_0$ are characteristic, whereas in the Baltic Sea they drop as low as $6‰_0$. Hence, substantial standing genetic variation in expanding populations could be crucial in adapting and thriving in new environmental conditions encountered after shifting from native into new habitats.

## 5 | Conclusions

Poleward and upward range shifts have been reported for marine species in response to changing environmental conditions, including contemporary climate change (Pinsky et al. 2020). However, responses vary depending on species tolerance to environmental factors such as temperature and oxygen, dispersal capacity, biotic interactions, habitat availability and human activities. A common pattern often seen in species shifting range is populations expanding at the leading edge, which tend to exhibit reduced genetic variability due to founder effects (Fifer et al. 2022). Among small pelagic species expanding in the North Sea, anchovies and sardines are showing range shifts due to different demographic processes. In contrast with the mixing reported here for anchovies in the North Sea and the retention of most of the genetic diversity, sardines fit the pattern of expansion of a leading-edge population, with reduced genetic diversity in the expanded populations (McKeown et al. 2024). While the expansion of sardines consists of new colonisers only, our study shows that anchovies in newly colonised areas in the North Sea are a mix of remnant local populations and southern migrants. Anchovies and sardines also differ in their level of population structure, with our study showing clear distinct evolutionary trajectories for the Northern and Southern clusters, as shown in the demographic reconstruction analysis, whereas North Sea and Bay of Biscay sardine populations exhibit a much shallower divergence (McKeown et al. 2024, 2025). The genetic patterns found for anchovies align more closely with those reported for European sprat *Sprattus sprattus*, another small pelagic fish that shows similar levels of genetic diversity across the Northeastern Atlantic (McKeown et al. 2020). While Bay of Biscay and the North Sea sprat constitute a single population, there are deeply divergent populations in the Norwegian fjords and the Baltic Sea, with some evidence of mixing between the Baltic and North Sea populations (McKeown et al. 2020; Quintela et al. 2020). This study contributes to elucidating population structure, potential range shifts and adaptive responses in small pelagic fish under climate change. Ultimately, the findings will contribute to a better understanding of evolutionary processes in marine ecosystems, which in turn will be a key factor for the conservation and management of these species in response to global environmental change.

### Author Contributions

J.M.P. wrote the original draft of the manuscript, performed all bioinformatics and statistical analyses, and prepared the visualisation and data presentation of the work. E.E.N., R.H., S.H., J.I.R. and R.C. contributed to conceptualisation and funding acquisition. E.E.N. and R.H. supervised the study. All authors reviewed and edited the manuscript and approved the final submission.

### Acknowledgements

We are grateful to Kai Ulrich Wieland, Susan Mærsk Lusseau, Bastian Huwer and Helle Rasmussen at DTU and Jon-Ivar Westgaard and Staby Arved at the Institute of Marine Research in Norway for samples. We thank Maj-Britt Jacobsen and Dorte Meldrup for laboratory assistance. We also acknowledge Susana Garrido at IPMA. This study received Portuguese national funds from FCT - Foundation for Science and Technology through projects UIDB/04326/2020 (DOI:10.54499/UIDB/04326/2020), UIDP/04326/2020 (DOI;10.54499/UIDP/04326/2020) and LAP/P/0101/2020 (DOI:10.54499/LA/P/0101/2020).

### Conflicts of Interest

The authors declare no conflicts of interest.

### Data Availability Statement

The data that support the findings of this study are openly available on Genbank after publication, with all raw sequences submitted to the Sequence Read Archive SRA Submission SUB15191179. All code and scripts used for data analysis is available on Github at https://github.com/martipujolar/scripts.

### References

Adams, R. H., D. R. Schield, D. C. Card, and T. A. Castoe. 2018. "Assessing the Impacts of Positive Selection on Coalescent-Based Tree Estimation and Species Delimitation." *Systematic Biology* 67: 1076–1090. https://doi.org/10.1093/sysbio/syy034.

Alexander, D. H., and K. Lange. 2011. "Enhancements to the ADMIXTURE Algorithm for Individual Ancestry Estimation." *BMC Bioinformatics* 12: 246. https://doi.org/10.1186/1471-2105-12-246.

Alheit, J. 1989. "Conparative Spawning Biology of Anchovies, Sardines and Sprats." *ICES Rapports et Procès-Verbaux Des Réunions* 191: 7–14.

Alheit, J., and M. A. Peck. 2019. "Drivers of Dynamics of Small Pelagic Fish Resources; Biology, Management and Human Factors." *Marine Ecology Progress Series* 617–618: 1–6. https://doi.org/10.3354/meps129852019.

Alheit, J., T. Pohlmann, M. Casini, et al. 2012. "Climate Variability Drives Anchovies and Sardines Into the North and Baltic Seas." *Progress in Oceanography* 96: 128–139. https://doi.org/10.1016/j.pocean.2011.11.015.

Alheit, J., C. Roy, and S. Kifani. 2009. "Decadal-Scale Variability in Populations." In *Climate Change and Small Pelagic Fish*, edited by D. M. Checkley, J. Alheit, Y. Oozeki, and C. Roy. Cambridge University Press.

Alheit, J., R. Voss, V. Mohrholz, and R. Hinrichs. 2007. "Climate Drives Anchovies and Sardines Into the North Sea." *GLOBEC International Newsletter* 13: 77–78.

Ayala, D., R. F. Guerrero, and M. Kirkpatrick. 2013. "Reproductive Isolation and Local Adaptation Quantified for a Chromosome Inversion in a Malaria Mosquito." *Evolution* 67: 946–958. https://doi.org/10.1111/j.1558-5646.2012.01836.x.

Baldó, F., E. García-Isarch, M. P. Jiménez, Z. Romero, A. Sánhez-Lamadrid, and I. A. Catalán. 2006. "Spatial and Temporal Distribution of the Early Life Stages of Three Commercial Fish Species in the Northeastern Shelf of the Gulf of Cadiz." *Deep-Sea Research Part II: Topical Studies in Oceanography* 53: 1391–1401. https://doi.org/10.1016/j.dsr2.2006.04.004.

Barth, J. M. I., P. R. Berg, P. R. Jonsson, et al. 2017. "Genome Architecture Enables Local Adaptation of Atlantic Cod Despite High Connectivity." *Molecular Ecology* 26: 4452–4466. https://doi.org/10.1111/mec.14207.

Berg, P. R., S. Jentoft, B. Star, et al. 2015. "Adaptation to Low Salinity Promotes Genomic Divergence in Atlantic Cod (*Gadus morhua*)." *Genome Biology and Evolution* 7: 1644–1663. https://doi.org/10.1093/gbe/evv093.

Berg, P. R., B. Star, C. Pampoulie, et al. 2017. "Trans-Oceanic Genomic Divergence of Atlantic Cod Ecotypes Is Associated With Large Invasions." *Heredity* 119: 418–428. https://doi.org/10.1038/hdy.2017.54.

Berg, P. R., B. Star, C. Pampoulie, et al. 2016. "Three Chromosomal Rearrangements Promote Genomic Divergence Between Migratory and Stationary Ecotypes of Atlantic Cod." *Scientific Reports* 6: 23246. https://doi.org/10.1038/srep23246.

Bierne, N., J. Welch, E. Loire, F. Bonhomme, and P. David. 2011. "The Coupling Hypothesis: Why Genome Scans May Fail to Map Local Adaptation Genes." *Molecular Ecology* 20, no. 10: 2044–2072. https://doi.org/10.1111/j.1365-294X.2011.05080.x.

Bindea, G., B. Mlecnik, H. Hackl, et al. 2009. "ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks." *Bioinformatics* 25: 1091–1093. https://doi.org/10.1093/bioinformatics/btp101.

Boden, M. J., and D. J. Kennaway. 2006. "Circadian Rhythms and Reproduction." *Reproduction* 132: 379–392. https://doi.org/10.1530/rep.1.00614.

Bolger, A. M., M. Lohse, and B. Usabel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence." *Bioinformatics* 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Caneco, C., A. Silva, and A. Morais. 2004. "Morphometric Variation Among Anchovy (*Engraulis encrasicolus*) Populations From the Bay of Biscay and Iberian Waters." ICES CM 2004/EE:24. https://doi.org/10.17895/ices.pub.25349287.

Casaucao, A., E. González-Ortegón, M. P. Jiménez, et al. 2021. "Assessment of the Spawning Habitat, Spatial Distribution, and Lagrangian Dispersion of the European Anchovy (*Engraulis encrasicolus*) Early Stages in the Gulf of Cadiz During an Apparent Anomalous Episode in 2016." *Science of the Total Environment* 781: 146530. https://doi.org/10.1016/j.scitotenv.2021.146530.

Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer.

Cury, P., A. Bakun, R. J. M. Crawford, et al. 2000. "Small Pelagics in Upwelling Systems: Patterns of Interaction and Structural Changes in 'Wasp-Waist' Ecosystems." *ICES Journal of Marine Science* 57: 603–618. https://doi.org/10.1006/jmsc.2000.0712.

Da Fonseca, R. R., P. F. Campos, A. Rey-Iglesia, et al. 2024. "Population Genomics Reveals the Underlying Structure of the Small Pelagic European Sardine and Suggests Low Connectivity With Macaronesia." *Genes* 15: 170. https://doi.org/10.3390/genes15020170.

Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Dennis, G., B. T. Sherman, D. Hosack, et al. 2003. "DAVID: Database for Annotation, Visualization and Integrated Discovery." *Genome Biology* 4, no. 9: R60. https://doi.org/10.1186/gb-2003-4-9-r60.

Fifer, J. E., N. Yasuda, T. Yamakita, C. B. Bove, and S. W. Davies. 2022. "Genetic Divergence and Range Expansion in a Western North Pacific Coral." *Science of the Total Environment* 813: 152423. https://doi.org/10.1016/j.scitotenv.2021.152423.

Fuentes-Pardo, A. P., E. D. Farrell, M. E. Pettersson, C. G. Sprehn, and L. Andersson. 2023. "The Genomic Basis and Environmental Correlates of Local Adaptation in the Atlantic Horse Mackerel (*Trachurus trachurus*)." *Evolutionary Applications* 16: 1201–1219. https://doi.org/10.1111/eva.13559.

Han, F., M. Jamsandekar, M. E. Pettersson, et al. 2020. "Ecological Adaptation in Atlantic Herring Is Associated With Large Shifts in Allele Frequencies at Hundreds of Loci." *eLife* 9: e61076. https://doi.org/10.7554/eLife.61076.

Harada, Y., K. Yoshida, N. Kawano, and K. Miyado. 2013. "Critical Role of Exosomes in Sperm-Egg Fusion and Virus-Induced Cell-Cell Fusion." *Reproductive Medicine and Biology* 12: 117–126. https://doi.org/10.1007/s12522-013-0152-2.

Hardin, P. E. 2000. "From Biological Clock to Biological Rhythms." *Genome Biology* 1: 1023. https://doi.org/10.1186/gb-2000-1-4-reviews1023.

Hemmer-Hansen, J., E. E. Nielsen, N. O. Therkildsen, et al. 2013. "A Genomic Island of Divergence Linked to Ecotype Divergence in Atlantic Cod." *Molecular Ecology* 22: 2653–2667. https://doi.org/10.1111/mec12284.

Hemmer-Hansen, J., N. O. Therkildsen, and J. M. Pujolar. 2014. "Population Genomics of Marine Fishes: Next-Generation Prospects and Challenges." *Biological Bulletin* 227: 117–132. https://doi.org/10.1086/BBLv227n2p117.

Hewitt, G. M. 1996. "Some Genetic Consequences of Ice Ages, and Their Role in Divergence and Speciation." *Biological Journal of the Linnean Society* 58: 247–276. https://doi.org/10.1006/bijl.1996.0035.

Hewitt, G. M. 2004. "Genetic Consequences of Climatic Oscillations in the Quaternary." *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 359: 183–195. https://doi.org/10.1098/rstb.2003.1388.

Hollenbeck, C. M., D. S. Portnoy, D. Garcia de la Serrano, T. Magnesian, I. Mateus, and I. A. Johnston. 2022. "Temperature-Associated Selection Linked to Putative Chromosomal Inversions in King Scallop (*Pecten maximus*)." *Proceedings of the Royal Society B: Biological Sciences* 289: 20221573. https://doi.org/10.1098/rspb.2022.1573.

Huret, M., C. Lebigre, M. Iriondo, I. Montes, and A. Estonba. 2020. "Genetic Population Structure of Anchovy (*Engraulis encrasicolus*) in North-Western Europe and Variability in the Seasonal Distribution of the Stocks." *Fisheries Research* 229: 105619. https://doi.org/10.1016/j.fishres.2020.105619.

Hutchings, L., C. D. van der Lingen, L. J. Shannon, et al. 2009. "The Benguela Current: An Ecosystem of Four Components." *Progress in Oceanography* 83: 15–32. https://doi.org/10.1016/j.pocean.2009.07.046.

ICES. 2012. "Report of the Working Group on Small Pelagic Fish, Their Ecosystems and Climate Impact (WGSPEC), 27 February–2 March 2012, Fuengirola, Spain."

ICES. 2018. "Report of the Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA), 26–30 June 2018, Lisbon, Portugal."

ICES. 2022. "Working Group for the Bay of Biscay and the Iberian Waters Ecoregion (WGBIE)." ICES Scientific Report 4 52: 847. https://doi.org/10.17895/ices.pub.20068988.

ICES. 2023. "Report on Ocean Climate." (Eds. González-Pola, C., Larsen, K. M. H., Fratantoni, P., and Beszczynska-Möller, A.). ICES Cooperative Research Reports 358. 123 pp.

Jansen, T., A. Campbell, C. Kelly, H. Hatún, and M. R. Payne. 2012. "Migration and Fisheries of North East Atlantic Mackerel (*Scomber scombrus*) in Autumn and Winter." *PLoS One* 7: e51541. https://doi.org/10.1371/journal.pone.0051541.

Jones, F. C., M. G. Grabherr, Y. F. Chan, et al. 2012. "The Genomic Basis of Adaptive Evolution in Threespine Sticklebacks." *Nature* 484: 55–61. https://doi.org/10.1038/nature10944.

Le Moan, A., P. A. Gagnaire, and F. Bonhomme. 2016. "Parallel Genetic Divergence Among Coastal-Marine Ecotype Pairs of European Anchovy Explained by Differential Introgression After Secondary Contact." *Molecular Ecology* 25: 3187–3202. https://doi.org/10.1111/mec.13627.

Leaché, A. D., R. B. Harris, B. Rannala, and Z. Yang. 2014. "The Influence of Gene Flow on Species Tree Estimation: A Simulation Study." *Systematic Biology* 63: 17–30. https://doi.org/10.1093/sysbio/syt049.

Li, H., and R. Durbin. 2009. "Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform." *Bioinformatics* 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Li, H., and R. Durbin. 2011. "Inference of Human Population History From Individual Whole-Genome Sequences." *Nature* 475: 493–496. https://doi.org/10.1038/nature10231.

Li, H., B. Handsaker, A. Wysoker, et al. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Magoulas, A., R. Castilho, S. Caetano, S. Marcato, and T. Patarnello. 2006. "Mitochondrial DNA Reveals a Mosaic Pattern of Phylogeographical Structure in Atlantic and Mediterranean Populations of Anchovy (*Engraulis encrasicolus*)." *Molecular Phylogenetics and Evolution* 39: 734–746. https://doi.org/10.1016/j.ympev.2006.01.016.

McKeown, N. J., F. Campanella, J. F. Silva, et al. 2025. "Hierarchical Neutral and Non-Neutral Spatial Genetic Structuring in the European Sardine (*Sardina pilchardus*) Revealed by Genetic Analysis: Implications for Management." *Evolutionary Applications* 18: e70080. https://doi.org/10.1111/eva.70080.

McKeown, N. J., P. Carpi, J. F. Silva, A. J. E. Healey, P. W. Shaw, and J. van der Kooij. 2020. "Genetic Population Structure and Tools for the Management of European Sprat (*Sprattus sprattus*)." *ICES Journal of Marine Science* 77: 2134–2143. https://doi.org/10.1093/icesjms/fsaa113.

McKeown, N. J., C. Lebigre, J. van der Kooij, and M. Huret. 2024. "Genomic Analysis of NE Atlantic Sardine (*Sardina pilchardus*) Reveals Reduced Variation in a Recently Established North Sea Population and Directs Reconsideration of Management Units." *Ecology and Evolution* 14: e70101. https://doi.org/10.1002/ece3.70101.

Miller, M. J., H. Westerberg, H. Sparholt, et al. 2019. "Spawning by the European Eel Across 2000 Km of the Sargasso Sea." *Biology Letters* 15: 20180835. https://doi.org/10.1098/rsbl.2018.0835.

Montes, I., I. Zarraonaindia, M. Irondo, et al. 2016. "Transcriptome Analysis Deciphers Evolutionary Mechanisms Underlying Genetic Differentiation Between Coastal and Offshore Anchovy Populations in the Bay of Biscay." *Marine Biology* 163: 205. https://doi.org/10.1007/s00227-016-2979-7.

Motos, L. 1996. "Reproductive Biology and Fecundity of the Bay of Biscay Anchovy Population (*Engraulis encrasicolus*)." *Scientia Marina* 60: 195–207.

Mutalipassi, M., E. D'Anza, M. Pugliano, et al. 2024. "Casting Light on the European Anchovy: From Biology to Conservation and Industry." *Frontiers in Ecology and Evolution* 12: 1352055. https://doi.org/10.3389/fevo.2024.1352055.

Nadachowska-Brzyska, K., C. Li, L. Smeds, G. Zhang, and H. Ellegren. 2015. "Temporal Dynamics of Avian Populations During Pleistocene Revealed by Whole-Genome Sequences." *Current Biology* 25: 1375–1380. https://doi.org/10.1016/j.cub.2015.03.047.

Nei, M., and W. H. Li. 1979. "Mathematical Model for Studying Genetic Variation in Terms of Restriction Endonucleases." *Proceedings of the National Academy of Sciences of the United States of America* 76: 5269–5273. https://doi.org/10.1073/pnas.76.10.5269.

Nielsen, E. E., J. Hemmer-Hansen, N. A. Poulsen, et al. 2009. "Genomic Signatures of Local Directional Selection in a High Gene Flow Marine Organism: The Atlantic Cod (*Gadus morhua*)." *BMC Evolutionary Biology* 9: 276. https://doi.org/10.1186/1471-2148-9-276.

Nikolic, N., S. Liu, M. W. Jacobsen, et al. 2020. "Speciation History of European (*Anguilla anguilla*) and American Eel (*A. rostrata*) Analysed Using Genomic Data." *Molecular Ecology* 29: 565–577. https://doi.org/10.1111/mec.15342.

Pacariz, S., G. Björk, and H. Svedäng. 2014. "Interannual Variability in the Transport of Fish Eggs in the Kattegat and Öresund." *ICES Journal of Marine Science* 71: 1706–1716. https://doi.org/10.1093/icesjms/fsu044.

Palomera, I., M. P. Olivar, J. Salat, A. Sabates, A. Garcia, and B. Morales-Nin. 2007. "Small Pelagic Fish in the NW Mediterranean Sea: An Ecological Review." *Progress in Oceanography* 74: 377–396. https://doi.org/10.1016/j.pocean.2007.04.12.

Palumbi, S. R. 1994. "Genetic Divergence Reproductive Isolation, and Marine Speciation." *Annual Review of Ecology, Evolution, and Systematics* 25: 547–572. https://doi.org/10.1146/annurev.es.25.110194.002555.

Patterson, N., A. L. Price, and D. Reich. 2006. "Population Structure and Eigenanalysis." *PLoS Genetics* 2: e190. https://doi.org/10.1371/journal.pgen.0020190.

Petitgas, P., J. Alheit, M. Peck, et al. 2012. "Anchovy Population Expansion in the North Sea." *Marine Ecology Progress Series* 444: 1–13. https://doi.org/10.3354/meps09451.

Pettersson, M. E., M. Quintela, F. Besnier, et al. 2024. "Limited Parallelism in Genetic Adaptation to Brackish Water Bodies in European Sprat and Atlantic Herring." *Genome Biology and Evolution* 16: evae133. https://doi.org/10.1093/gbe/evae133.

Pikitch, E. K., K. J. Rountos, T. E. Essington, et al. 2014. "The Global Contribution of Forage Fish to Marine Fisheries and Ecosystems." *Fish and Fisheries* 15: 43–64. https://doi.org/10.1111/faf.12004.

Pinsky, M. L., R. L. Selden, and Z. J. Kitchel. 2020. "Climate-Driven Shifts in Marine Species Ranges: Scaling From Organisms to Communities." *Annual Review of Marine Science* 12: 153–179. https://doi.org/10.1146/annurev-marine-010419-010916.

Purcell, S., B. Neale, K. Todd-Brown, et al. 2007. "PLINK: A Toolset for Whole-Genome Association and Population Based Linkage Analysis." *American Journal of Human Genetics* 81: 559–575. https://doi.org/10.1086/519795.

Quintela, M., C. Kvamme, D. Bekkevold, et al. 2020. "Genetic Analysis Redraws the Management Boundaries for the European Sprat." *Evolutionary Applications* 13: 1906–1922. https://doi.org/10.1111/eva12942.

Raybaud, V., M. Bacha, R. Amara, and G. Beaugrand. 2017. "Forecasting Climate-Driven Changes in the Geographical Range of European Anchovy (*Engraulis encrasicolus*)." *ICES Journal of Marine Science* 74: 1288–1299. https://doi.org/10.1093/icesjms/fsx003.

Sanz, N., J. L. García-Marín, J. Vinas, M. I. Roldán, and C. Plà. 2008. "Spawning Groups of European Anchovy: Population Structure and

Management Implications." *ICES Journal of Marine Science* 65: 1635–1644. https://doi.org/10.1093/icesjms/fsn128.

Schluter, D., and L. H. Rieseberg. 2022. "Three Problems in the Genetics of Speciation by Selection." *Proceedings of the National Academy of Sciences of the United States of America* 119: e2122153119. https://doi.org/10.1073/pnas.2122153119.

Sharma, V. K. 2003. "Adaptive Significance of Circadian Clocks." *Chronobiology International* 20: 901–919. https://doi.org/10.1081/cbi-120026099.

Silva, G., J. B. Horne, and R. Castilho. 2014. "Anchovies Go North and West Without Losing Diversity: Post-Glacial Range Expansions in a Small Pelagic Fish." *Journal of Biogeography* 41: 1171–1182. https://doi.org/10.1111/jbi.12275.

Skjæraasen, J. E., J. J. Meager, Ø. Karlsen, J. A. Hutchings, and A. Fernö. 2011. "Extreme Spawning-Site Fidelity in Atlantic Cod." *ICES Journal of Marine Science* 68: 1472–1477. https://doi.org/10.1093/icesjms/fsr055.

Stern, D. B., and C. E. Lee. 2020. "Evolutionary Origins of Genomic Adaptations in an Invasive Copepod." *Nature Ecology & Evolution* 4: 1084–1094. https://doi.org/10.1038/s41559-020-1201-y.

Stewart, J. R., A. M. Listen, I. Barnes, and L. Dalén. 2010. "Refugia Revisited: Individual Responses of Species in Space and Time." *Proceedings of the Royal Society B: Biological Sciences* 277: 661–671. https://doi.org/10.1098/srpb.2009.1272.

Svendsen, J. I., H. Alexanderson, V. Astakhov, et al. 2004. "Late Quaternary Ice Sheet History of Northern Eurasia." *Quaternary Science Reviews* 23: 1229–1271. https://doi.org/10.1016/j.quascirev.2003.12.008.

Taberlet, P., L. Fumigalli, A. G. Wust-Saucy, and J. F. Cosson. 1998. "Comparative Phylogeography and Postglacial Colonization Routes in Europe." *Molecular Ecology* 7: 453–464. https://doi.org/10.1046/j.1365-294x.1998.00289x.

Teske, P. R., A. Emami-Khoyi, T. R. Golla, et al. 2021. "The Sardine Run in Southeastern Africa Is a Mass-Migration Into an Ecological Trap." *Science Advances* 7: eabf4514. https://doi.org/10.1126/sciadv.abf4514.

Uriarte, A., L. Ibaibarriaga, S. Sanchez-Marono, et al. 2023. "Lessons Learnt on the Management of Short-Lived Fish From the Bay of Biscay Anchovy Case Study: Satisfying Fishery Needs and Sustainability Under Recruitment Uncertainty." *Marine Policy* 150: 105512. https://doi.org/10.1016/j.marpol.2023.1055122023.

Van der Kooij, J., N. McKeown, F. Campanella, et al. 2024. "Northward Range Expansion of Bay of Biscay Anchovy Into the English Channel." *Marine Ecology Progress Series* 741: 217–236. https://doi.org/10.3354/meps14603.

Van der Lingen, C. D., L. Hutchings, D. Merkle, J. J. Van der Werhuizen, and J. Nelson. 2021. "Comparative Spawning Habitats of Anchovy (*Engraulis capensis*) and Sardine (*Sardinops sagax*) in the Southern Benguela Upwelling Ecosystem." Proceedings of the Symposium on Spatial Processes and Management of Marine Populations. University of Alaska Sea Grant, Fairbanks.

Weir, B. S., and C. C. Cockerham. 1984. "Estimating F-Statistics for the Analysis of Population Structure." *Evolution* 38: 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x.

Wellenreuther, M., and L. Bernatchez. 2018. "Eco-Evolutionary Genomics of Chromosomal Inversions." *Trends in Ecology & Evolution* 33: 427–440. https://doi.org/10.1016/j.tree.2018.04.002.

Zarraonaindia, I., M. Iriondo, A. Albaina, et al. 2012. "Multiple SNP Markers Reveal Fine-Scale Population and Deep Phylogeographic Structure in European Anchovy (*Engraulis encrasicolus*)." *PLoS One* 7: e42201. https://doi.org/10.1371/journal.pone.0042201.

Zarraonaindia, I., M. A. Pardo, M. Iriondo, C. Manzano, and A. Estonba. 2009. "Microsatellite Variability in European Anchovy (*Engraulis encrasicolus*) Calls for Further Investigation of Its Genetic Structure and Biogeography." *ICES Journal of Marine Science* 66: 2176–2182. https://doi.org/10.1093/icesjms/fsp187.

Zhang, L., R. Reifova, Z. Halenkova, and Z. Gompert. 2021. "How Important Are Structural Variants for Speciation?" *Genes* 12: 1084. https://doi.org/10.3390/genes12071084.

**Supporting Information**

Additional supporting information can be found online in the Supporting Information section.