








## Article

# Population Genomics Reveals the Underlying Structure of the Small Pelagic European Sardine and Suggests Low Connectivity within Macaronesia

Rute R. da Fonseca <sup>1,2,\*</sup> , Paula F. Campos <sup>2,3</sup> , Alba Rey-Iglesia <sup>4</sup>, Gustavo V. Barroso <sup>5</sup>, Lucie A. Bergeron <sup>6</sup>, Manuel Nande <sup>3</sup>, Fernando Tuya <sup>7</sup> , Sami Abidli <sup>8</sup>, Montse Pérez <sup>9</sup> , Isabel Riveiro <sup>9</sup>, Pablo Carrera <sup>9</sup>, Alba Jurado-Ruzafa <sup>10</sup> , M. Teresa G. Santamaría <sup>10</sup>, Rui Faria <sup>3,†</sup>, André M. Machado <sup>3</sup> , Miguel M. Fonseca <sup>3</sup>, Elsa Froufe <sup>3</sup> and L. Filipe C. Castro <sup>3,11,\*</sup> 

- <sup>1</sup> Center for Global Mountain Biodiversity, GLOBE Institute, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen, Denmark
  - <sup>2</sup> The Bioinformatics Centre, Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark; campos.f.paula@gmail.com
  - <sup>3</sup> CIIMAR—Interdisciplinary Centre of Marine and Environmental Research, University of Porto, 4050-123 Porto, Portugal; mnande@ciimar.up.pt (M.N.); r.macieiradefaria@sheffield.ac.uk (R.F.); andre.machado@ciimar.up.pt (A.M.M.); mig.m.fonseca@gmail.com (M.M.F.); elsafrufe@gmail.com (E.F.)
  - <sup>4</sup> Centre for GeoGenetics, Natural History Museum Denmark, University of Copenhagen, Østervoldgade 5-7, 1350 Copenhagen, Denmark; ardelaiglesia@sund.ku.dk
  - <sup>5</sup> Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA; gvbarroso@gmail.com
  - <sup>6</sup> Section for Ecology and Evolution, University of Copenhagen, 2100 Copenhagen, Denmark; lucie.a.bergeron@gmail.com
  - <sup>7</sup> Grupo en Biodiversidad y Conservación, IU-ECOQUA, Universidad de Las Palmas de Gran Canaria, 35017 Las Palmas, Spain; ftuya@yahoo.es
  - <sup>8</sup> Laboratory of Environment Bio-Monitoring, Faculty of Sciences of Bizerte, University of Carthage, Bizerte 7021, Tunisia; abidli\_sami@yahoo.fr
  - <sup>9</sup> Centro Oceanográfico de Vigo, Instituto Español de Oceanografía, IEO-CSIC, 36390 Vigo, Spain; montse.perez@ieo.csic.es (M.P.); isabel.riveiro@ieo.csic.es (I.R.); pablo.carrera@ieo.csic.es (P.C.)
  - <sup>10</sup> Centro Oceanográfico de Canarias, Instituto Español de Oceanografía, IEO-CSIC, 38180 Santa Cruz de Tenerife, Spain; alba.jurado@ieo.csic.es (A.J.-R.); mtgarsantamaria@gmail.com (M.T.G.S.)
  - <sup>11</sup> Department of Biology, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal
- \* Correspondence: rfonseca@sund.ku.dk (R.R.d.F.); filipe.castro@ciimar.up.pt (L.F.C.C.)  
† Current address: BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal.



**Citation:** da Fonseca, R.R.; Campos, P.F.; Rey-Iglesia, A.; Barroso, G.V.; Bergeron, L.A.; Nande, M.; Tuya, F.; Abidli, S.; Pérez, M.; Riveiro, I.; et al. Population Genomics Reveals the Underlying Structure of the Small Pelagic European Sardine and Suggests Low Connectivity within Macaronesia. *Genes* **2024**, *15*, 170. <https://doi.org/10.3390/genes15020170>

Academic Editors: Zissis Mamuris and Maria-Anna Kyrgiagini

Received: 27 November 2023

Revised: 8 January 2024

Accepted: 23 January 2024

Published: 27 January 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The European sardine (*Sardina pilchardus*, Walbaum 1792) is indisputably a commercially important species. Previous studies using uneven sampling or a limited number of markers have presented sometimes conflicting evidence of the genetic structure of *S. pilchardus* populations. Here, we show that whole genome data from 108 individuals from 16 sampling areas across 5000 km of the species' distribution range (from the Eastern Mediterranean to the archipelago of Azores) support at least three genetic clusters. One includes individuals from Azores and Madeira, with evidence of substructure separating these two archipelagos in the Atlantic. Another cluster broadly corresponds to the center of the distribution, including the sampling sites around Iberia, separated by the Almeria–Oran front from the third cluster that includes all of the Mediterranean samples, except those from the Alboran Sea. Individuals from the Canary Islands appear to belong to the Mediterranean cluster. This suggests at least two important geographical barriers to gene flow, even though these do not seem complete, with many individuals from around Iberia and the Mediterranean showing some patterns compatible with admixture with other genetic clusters. Genomic regions corresponding to the top outliers of genetic differentiation are located in areas of low recombination indicative that genetic architecture also has a role in shaping population structure. These regions include genes related to otolith formation, a calcium carbonate structure in the inner ear previously used to distinguish *S. pilchardus* populations. Our results provide a baseline for further characterization of physical and

genetic barriers that divide European sardine populations, and information for transnational stock management of this highly exploited species towards sustainable fisheries.

**Keywords:** European sardine; low coverage sequencing; population structure; marine fish

## 1. Background

Population structuring in the absence of obvious physical barriers has puzzled biologists for centuries. A lack of or reduced genetic structure is expected in an oceanic environment. This is because most marine organisms are capable of exchanging migrants across large distances in long larval pelagic phases, plus they have high fecundity, large population sizes and adult migratory behavior [1,2]. Yet, many studies have shown that several species have higher spatial genetic differentiation than expected considering their high dispersal potential [3,4]. In the case of marine fish, structure can range from a lack of differentiation between oceans to significant structure within an ocean basin [2], challenging the simple concept of “open seas” and the assumption of high connectivity in marine environments [5]. Assessing the existence of population structure in marine species capable of long-distance dispersal is essential to identify the various factors involved in population differentiation and diversification in the absence of complete physical barriers [6]. This is especially relevant for conservation efforts, including stock management of commercially important species [1,2,7].

The Mediterranean Sea and the contiguous Northeastern Atlantic Ocean have been the focus of several phylogeographic and population genetic studies on marine fish (e.g., [8,9]). The Almeria–Oran Front, a well-defined oceanographic break situated east of the Strait of Gibraltar, has been suggested to be responsible for hindering gene flow between Mediterranean and Atlantic fish populations of many fish species but it is far from being an universal barrier [8]. The less studied Macaronesian region, a group of archipelagos (Azores, Madeira and Canaries) separated from the Euro-African mainland by c. 100–1900 km, has also been the target of several phylogeographic studies (e.g., [10,11]). This region is characterized by the presence of several oceanographic currents, e.g., the North Atlantic Current, the Azores Current and the Canary Current [12], that together with the apparent lack of physical barriers can strengthen the potential for gene flow. Therefore, it is not surprising that several studies have reported low population genetic differentiation within the Macaronesian region for different taxa [1], including fishes [13,14]. Species distributed across these regions can thus inform us about the existence of cryptic substructure and possible barriers to gene flow among populations.

One of the most important pelagic fish resources in Atlantic waters is the European sardine, *Sardina pilchardus* Walbaum, 1792. This species has an enormous economic value, especially in Southern Europe and Morocco [15], where it is a main target of the purse-seine fleets in Portugal and Spain, representing a major source of income for local economies [16]. Recently reported low biomass levels [17] led to a recommendation for reduced fishing in Southern Europe, with great economic impact. It also prompted us to reevaluate the current population structure of *S. pilchardus*, aiming at the ongoing discussion on how genetic information can contribute to stock delineation for management purposes [7].

The European sardine has a broad distribution from the Eastern Mediterranean to the North-East Atlantic, including the Azores, Madeira and the Canary Islands’ Archipelagos, and is found along the African coast down to Senegal [18]. Like other marine pelagic fish, *S. pilchardus* shows schooling and migratory behavior and high dispersal capabilities, both at the larval and adult stages. In agreement, low levels of genetic differentiation were detected across the species distribution using allozymes [19–22], mitochondrial DNA (mtDNA) [23,24], and microsatellites [10,25]. Nevertheless, phenotypic variation in gill raker counts and head length [18,26], as well as mitochondrial haplotype frequency differences [23] led to the proposal of two subspecies: *S. pilchardus pilchardus* (North Sea to south-

ern Portugal) and *S. pilchardus sardina* (Mediterranean Sea and northwest African coast). Accordingly, otolith shapes differ between Atlantic and Mediterranean sardines [27,28], and further suggest a subdivision between the Northern Mediterranean and the Alboran–Algero-Provençal basin [27,29]. A study using 15 allozymes supports the latter [30], but, unlike the otolith shapes, these markers suggest discontinuity caused by the Almeria–Oran front. When considering a large fraction of the European sardine Atlantic range, allozymes and microsatellites suggest that Madeira and Azores form a significantly differentiated group [10]. This mosaic of regional population structure built by several independent studies has been mostly justified by geographical barriers that potentially hinder gene flow, expected to be high for the abundant and mobile *S. pilchardus*. The phenotypic differences between groups might also have arisen from retention of adaptive phenotypes, as population structure in the Mediterranean was found to be associated with environmental variables [31]. This prompted us to raise questions about the contributions of genomic architecture to the basis of the observed present-day population structure.

In this study, we produced a European sardine genomic data set consisting of whole genome nuclear data and complete mitochondrial genomes for 88 individuals that were analyzed together with data from 20 sardine individuals from a previous study [32], in a total of 108 samples from 15 locations across 5000 km of the species distribution range. This enabled us to investigate previously suggested barriers to gene flow, map the major genetic clusters that characterize *S. pilchardus* in a large part of its distribution, compare markers with different modes of inheritance, and also reveal the first insights into the genomic barriers contributing to the observed population structure.

## 2. Materials and Methods

### 2.1. Sample Collection, DNA Extraction and Sequencing

Samples originate from 15 different geographical locations encompassing a large part of the species' current distribution range (Figure 1A, Table 1). Samples from the Alboran Sea and from Central and Western Bay of Biscay ( $n = 15$  individuals) were collected during the PELACUS oceanographic surveys. The remaining 73 specimens were collected at local markets from ten distinct geographic locations (Table S1). Additionally, sequence data for samples from the Gulf of Cadiz, Mar Menor, the Gulf of Lion, and a second set for Central Bay of Biscay, were obtained from Barry et al. (2022) [32].

Total genomic DNA was extracted using Qiagen's DNeasy Blood & Tissue Kit (Hilden, Germany) according to the manufacturer's instructions, with the following modifications: prior to elution in 100  $\mu$ L AE buffer, samples were incubated at 37 °C for 10 min to increase DNA yield. DNA concentration and purity were verified using a Nanodrop Spectrophotometer and a Qubit Fluorometer. A commercial service (Novogene, Guangzhou City, China) produced Truseq Nano DNA libraries and sequenced paired-end reads (150 base pairs (bp)) in a Novaseq6000. To assess the patterns of genetic differentiation of the European sardine, 81 samples were sequenced to 3X sequencing depth (i.e., each position of the genome is covered by 3 reads on average) and 7X to 20X sequencing depth (details in Supplementary File S1: Table S1). Raw data for 20 sardine individuals from Barry et al. [32] were further processed using the same procedure as described in the next section (sequencing depth between 15 and 22X). Table S1 indicates the assignment of samples to the different subsets considered for further analysis.

### 2.2. Assembly Filtering and Re-Sequencing Data Pre-Processing

Individual contigs in the reference genome (GenBank assembly accession: GCA\_900499035.1) of *S. pilchardus* [33] matching mitochondrial DNA (unique assignment to mitochondrial DNA without nuclear genome flanking regions) were identified via BLAST (version 2.6.0+) [34] using the mitochondrial genome (mtDNA) assembled by Machado et al. [35] as a query. Matching contigs were removed from the assembly file and replaced by the mtDNA of Machado et al. (2018) [35] to enable the extraction of individual mtDNA sequences from all individuals after mapping of resequencing data.

**Table 1.** Sampling information including our assignment to the three observed genetic clusters (Figure 1B).

Location	Short ID	Latitude	Longitude	Source	Collection Date (Number of Individuals)	Genetic Cluster	Tissue
Azores	AZO	37.8	−26.7	Fish market	June 2017 (10)	West	Muscle
Madeira	MAD	32.0	−16.9	Fish market	July 2017 (10)	West	Muscle
Canary Islands	CAN	28.8	−15.0	Collaborators	July 2017 (5), June 2022 (5)	East	Muscle
Bretagne	BRE	48.3	−4.9	Fish market	July 2017 (3)	Center	Muscle
Bay Biscay C	BCX	43.3	−1.9	[32]	December 2018 (5)	Center	Pectoral fin
Bay Biscay C	BC0	43.5	−1.7	PELACUS	Spring 2015 (5)	Center	Muscle
Bay Biscay S	BS0	43.8	−7.6	PELACUS	Spring 2015 (5)	Center	Fin
Iberia NW	IB1	42.2	−9.6	Fish market	June 2017 (5)	Center	Muscle
Iberia NW	IB2	41.1	−9.6	Fish market	June 2017 (5)	Center	Muscle
Gulf of Cádiz	CAD	37.0	−7.9	[32]	October 2018 (5)	Center	Pectoral fin
Morocco	MOR	34.5	−8.2	Collaborators	July 2017 (5)	Center	Muscle
Alboran Sea	ALB	36.1	−3.8	PELACUS	2017 (5)	Center	Muscle
Mar Menor	MME	38.0	−0.7	[32]	November 2018 (5)	East	Dorsal fin
Gulf of Lion	LIO	43.4	3.7	[32]	November 2018 (5)	East	Pectoral fin
Tunisia	TUN	37.7	10.8	Fish market	August 2017 (5)	East	Muscle
Adriatic Sea	ADR	44.9	13.1	Fish market	August 2017 (10)	East	Muscle
Aegean Sea	AEG	40.2	22.9	Fish market	August 2017 (10)	East	Muscle

Regions of low complexity in the reference genome of *S. pilchardus* [33] were detected with GenMap (version v1.3.0) [36] using a  $k$ -mer of 100 bp. We calculated the normalized depth per scaffold using the sequencing depth of scaffold 1 as a reference to identify potential misassemblies (e.g., unmerged haploid scaffolds or collapsed repeats regions). Regions of mappability below 1 (meaning that more than one 100 bp kmer matched the region, indicating duplications) or identified as repeats in all the other scaffolds, and sites with data missing in more than 25% of the individuals were excluded from all subsequent analyses.

Raw Illumina reads for all 108 samples were first processed with Trimmomatic (version 0.36) [37] for removal of adapter sequences, trimming bases with quality <20 and discarding reads with length <80. Clean reads were mapped to the genome assembly using bwa-mem version: 0.7.17-r1188 [38], and samtools version: 1.7 [39] was used to retain reads with mapping quality >25. PCR duplicates were removed with Picard MarkDuplicates (version 1.95; <http://broadinstitute.github.io/picard/>) and only reads where both pairs were retained were considered for the local realignment around indels with GATK version 3.6–0-g89b7209 [40] and further analyses. The mapping and base quality options -minQ 20 -minMapQ 30 were used in all subsequent analyses with ANGSD [41].

### 2.3. Population Structure

Beagle files with the nuclear genome positions of single nucleotide polymorphisms (SNPs) were produced by ANGSD [41] using the following options: -GL 1 -doGlf 2 -minMaf 0.05 -C 50 -baq 2 -remove\_bads 1 -uniqueOnly 1 -SNP\_pval 1e−6. Linkage disequilibrium (LD) was estimated as  $r^2$  values for all SNP pairs minimum 500 kbp apart with ngsLD v1.1.1 [42], and an LD decay curve was plotted using 0.05% of all estimated  $r^2$  values. This indicated that a distance threshold of 2000 bp was adequate for linkage pruning. A total of 560,735 SNPs were obtained using all samples ( $n = 108$ , Table S1), and the 319,236 SNPs located in putatively neutral regions of the genome (50 kb genomic windows with PBS values below the 90th percentile, details below; Figure S2) were used in all the population genomic analyses. Admixture proportions were estimated by running NGSadmix version 32 [43] for  $K$  equal to 2 and 3 with 300 seed values, ensuring convergence (convergence was not reached for  $K = 4$  and above). A principal component analysis (PCA) using the same SNP set was obtained with PCAngsd version 0.1 [44].

The mitochondrial genome (mtDNA) for each individual was obtained as a consensus sequence of the reads mapped to the mtDNA sequence included in the reference genome using the option `-doFasta 2` and removing positions with sequencing depth below 10X (`-setMinDepth 10`) in ANGSD [41]. A haplotype network was reconstructed using mitochondrial SNPs with minor allele frequency >30% (total of 26 SNPs) in POPART [45] with the Median Joining Network algorithm [46].

#### 2.4. Population Differentiation

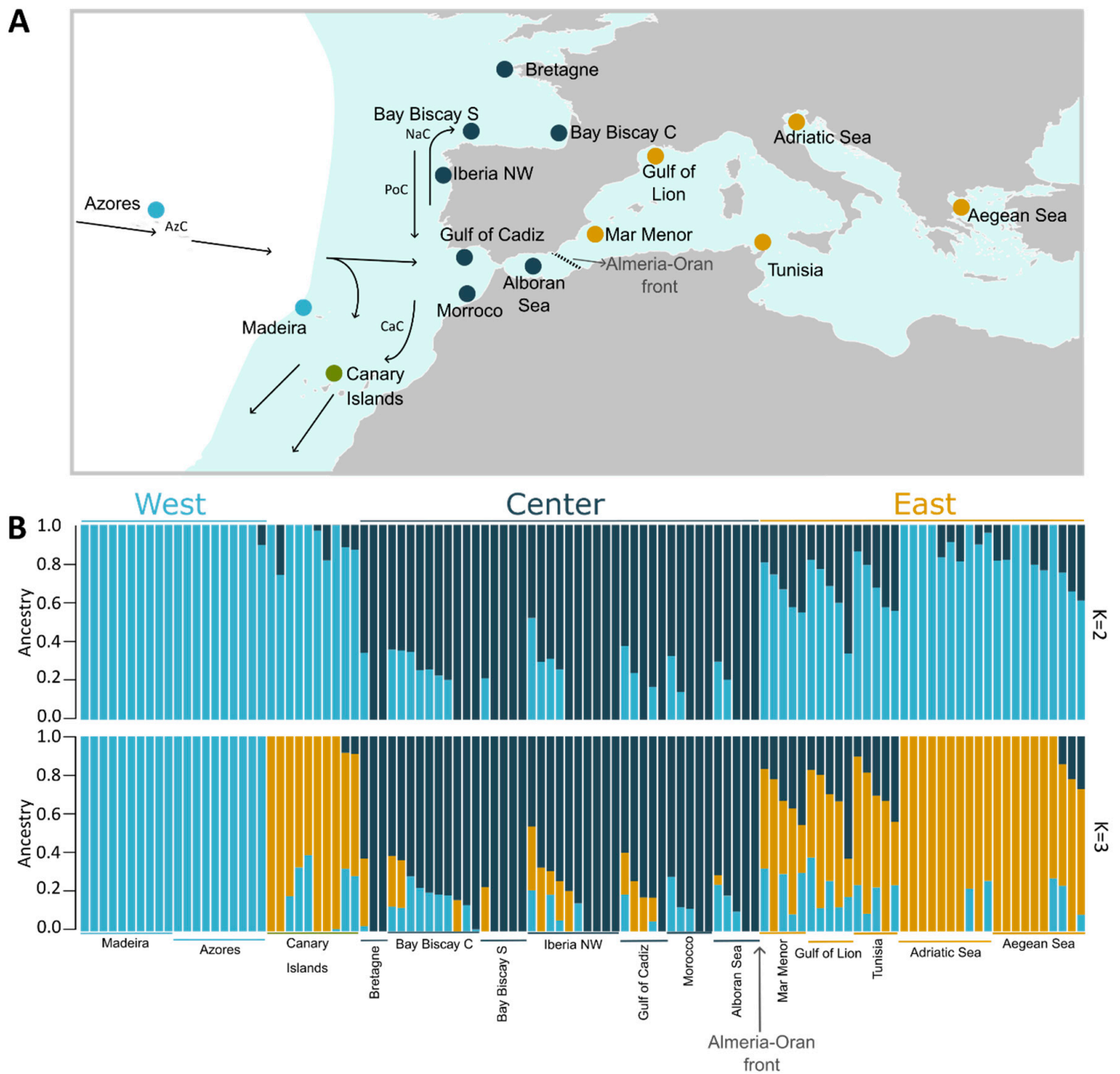
We used methods based on the site frequency spectrum (SFS) [47,48] to obtain the genome-wide fixation index ( $F_{ST}$ ) values in ANGSD [41]. We calculated  $F_{ST}$  for the populations containing at least 10 individuals with each geographical region represented by two locations (West: Madeira and Azores; East: Adriatic and Aegean; Center: Bay of Biscay and Gulf of Cádiz/Morocco; Table S2; Figure 2; Figure 3). First, we generated unfolded SAF files (`angsd -bam bamList -doSaf 1 -anc ANC -GL 1`) and then we estimated the folded SFS for each pair of populations (`realSFS safidx1 safidx2 -fold 1`). Each joint folded SFS was then used to estimate  $F_{ST}$  (`-whichFst 1 -fold 1`). Table S2 also includes values for the 25th and 75th percentile of the distribution of  $F_{ST}$  values calculated per 50 kb windows along the genome. To detect genomic windows of high differentiation in each region, we estimated the population branch statistic [49] for non-overlapping windows of 50 kb in ANGSD using this same approach with three populations with 10 individuals each (Tables S1–S3). The individuals were chosen from those that had 100% assignment to one of the three ancestral populations defined by NGSadm in preliminary analysis using all genomic positions that passed the filters described above.

Maritime geographical pairwise distances (<https://sea-distances.org/>, accessed on 2 June 2022) were calculated using the seaport nearest to the sampling location. Average distances were considered for merged populations. A Mantel test implemented in `ade4` [50] was used to test the statistical significance of the correlation between the geographic and the genetic distance matrices.

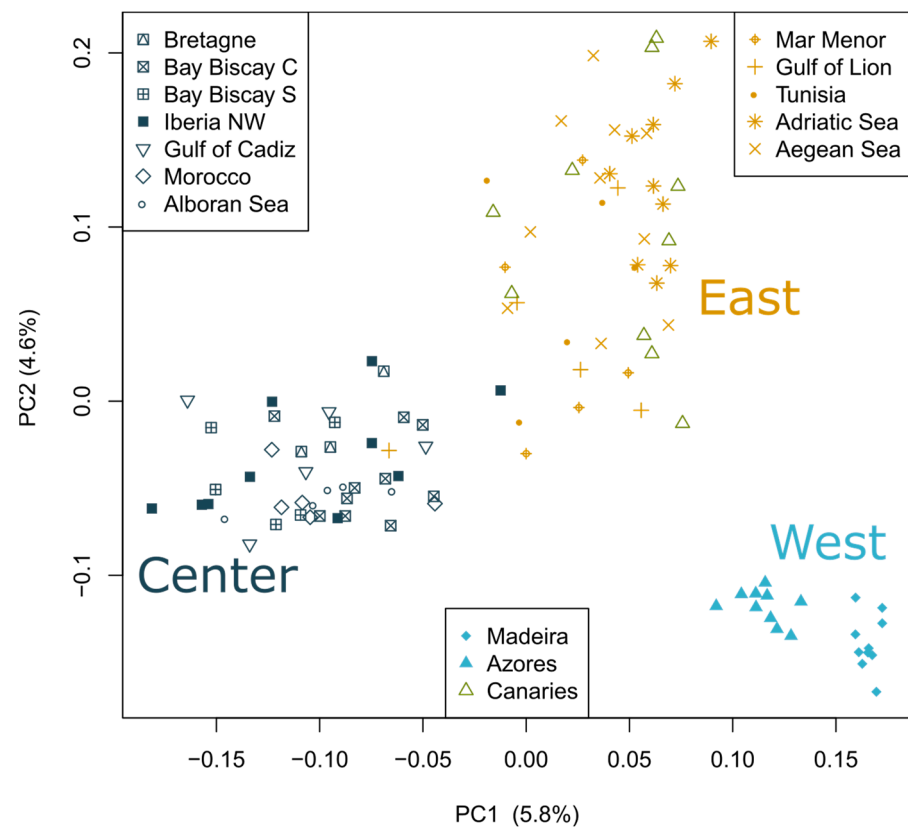
We assessed the gene content of the top outlier PBS windows (outliers in the 99th percentile; Figure S2) for each region by running `tblastn` [34] (BLAST version 2.6.0+) of the zebrafish proteome (ENSEMBL version GRCz11\_pep) against the reference genome using the option `"-evalue 0.000001"`. Phenotypes associated with each gene were extracted from ENSEMBL using Biomart [51].

#### 2.5. Recombination Rate

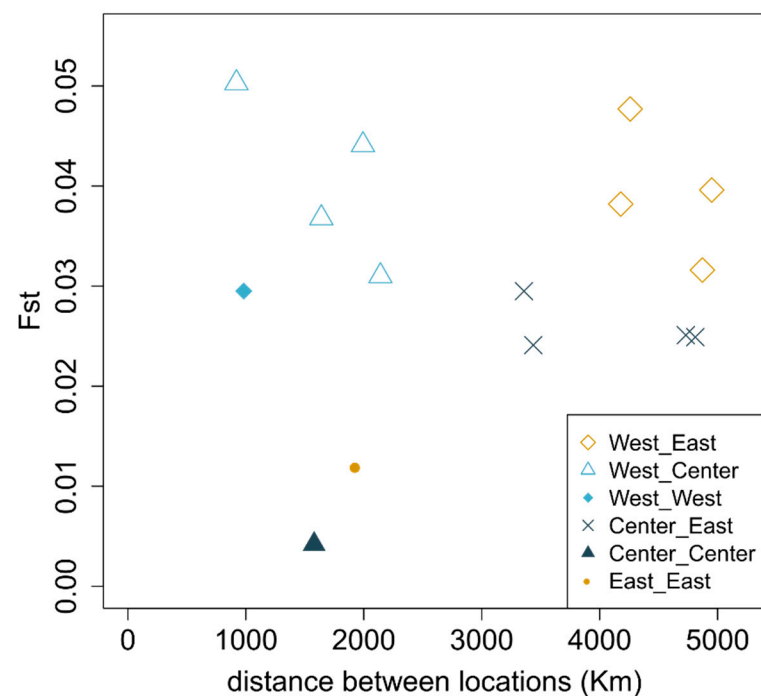
Variants were called using GATK version 4.0.7.0 [40] for one representative individual per region (Table S1). Briefly, variants were first called for each individual with HaplotypeCaller in BP-RESOLUTION mode; then those GVCF files for each sample were combined into a single one using `CombineGVCFs` per scaffold of interest, followed by joint genotyping with `GenotypeGVCF`. The default filter of GATK (`--phred-scaled-global-read-mismapping-rate 45 --base-quality-score-threshold 18 --min-base-quality-score 10`) was used. Recombination rates for 100 kb non-overlapping windows along the genome were estimated using the iSMC approach from [52]. We fitted an iSMC model with 40-time intervals and five categories of recombination rates to the samples from each population and optimized parameters in composite likelihood fashion [53]. We then obtained recombination landscapes of single-nucleotide resolution by performing posterior decoding in each diploid using the estimated parameters and computed a consensus map for each sample by averaging (for each site) the posterior estimates of  $\rho = 4 * N_e * r$  from all diploids. The final map of 100 kb resolution was obtained by further averaging the single-nucleotide estimates over 100 kb in non-overlapping windows.



**Figure 1.** (A) Sampling sites across the species distribution (blue area, adapted from FAO). The color of each circle represents the most frequent genetic cluster for  $K = 3$  (light blue: West; dark blue: Center; yellow: East; the Canary Islands are depicted in green). Surface currents are represented by arrows: Azores Current (AzC); Canary Current (CaC); Portugal Current (PoC); and Navidad Current (NaC). The Almeria–Oran Front (AO) is shown as a dashed line. (B) Population structure plot showing the ancestry of each individual (vertical bar) to two (above) and three (below) genetic clusters.



**Figure 2.** Distribution of individuals based on the first three components of the principal component analysis. Variance explained by each component is shown in parentheses.



**Figure 3.** Average pairwise  $F_{ST}$  between populations with minimum of 10 individuals (Table S2) calculated across 50 kb windows located in putatively neutral regions of the genome. The correlation between distance and  $F_{ST}$  is not significant (Mantel test;  $p$ -value = 0.32).

### 3. Results

#### 3.1. Population Structure

The admixture analysis conducted in NGSadmix showed that European sardines comprise three genetic clusters (Figure 1B). When setting the number of expected clusters to two ( $k = 2$ ; Figure 1B, top), one of the clusters is prevalent in the Center region, while the other is more frequent in both Western and Eastern regions, as well as the Canary Islands. Individuals showing patterns compatible with admixed ancestry between these two clusters were observed at all sampling sites, except at Madeira. For  $k = 3$  (Figure 1B, bottom), one of these clusters (West–East–Canaries) splits into two: one frequent in the Mediterranean and Canaries and the other in the West (Madeira and Azores). Patterns compatible with admixed ancestry among the three main clusters were observed in individuals from the Central region, Eastern Iberia, South of France, Tunisia and the Aegean Sea. A few individuals from Canaries and Adriatic show some admixed ancestry with the Western cluster. The proportion of individuals with admixed ancestry is lower in populations located in the extremes of the distribution: Madeira, Azores in the West, Aegean and Adriatic Sea in the East, as well as Bretagne and Bay of Biscay S in the central group.

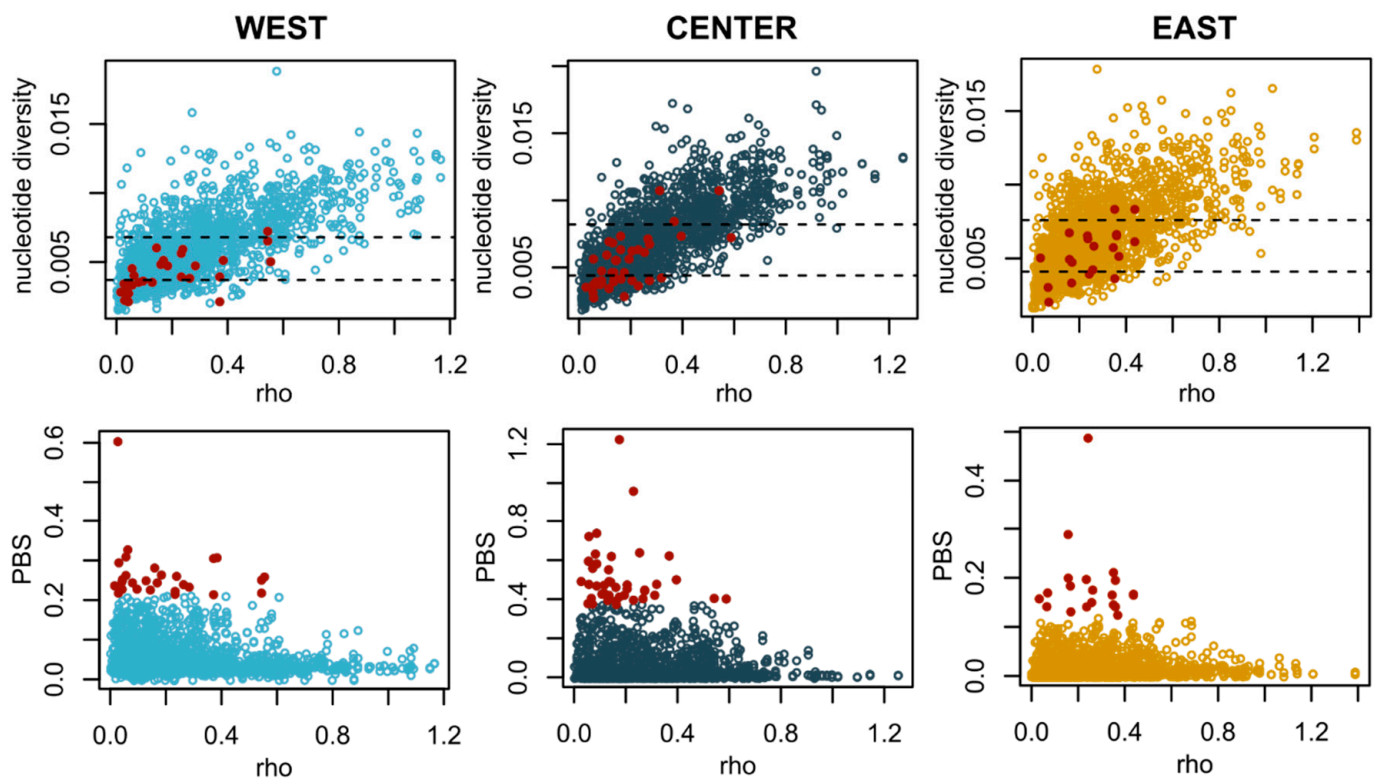
Three separate clusters of populations can also be observed in the principal component analysis (PCA; Figure 2). The first two PCs explained 5.8% and 4.6% of the total variation. Within the three main clusters, individuals from the same sampling location do not cluster together, likely reflecting admixture, as observed in Figure 1B, except for Madeira and Azores, which appear separated in the PC1.

The main clusters observed in the mitochondrial haplotype network combine haplotypes that are not geographically confined to a region, suggestive of gene flow mainly between Center and East (Figure S1), which is consistent with their proximity in PC2 (Figure 2). The West group dominates a centrally branching haplogroup.

#### 3.2. Population Differentiation

The levels of nucleotide diversity were comparable across regions, although slightly lower for populations in the West (Table S2). In general, we observed lower values of genetic differentiation as measured by  $F_{ST}$  for comparisons within regions (distances ranging from 982 to 1943 km show  $F_{ST}$  values between 0.004 and 0.03; Figure 3 and Table S3). The highest values of  $F_{ST}$  included comparisons with Madeira (West) and sampling locations in the Center region and the Aegean Sea ( $0.044 < F_{ST} < 0.05$ ). We do not find evidence of isolation by distance across the sampled range (*Mantel* test;  $p$ -value = 0.32). Areas of the genome that correspond to the top PBS outliers (99th percentile, corresponding to the genomic regions where one population is the most dissimilar to the other two) show lower recombination rates, and areas of very low differentiation are associated with the highest recombination rates (Figure 4). The gene content of the windows with the top individual PBS values per geographical location (Figure S2) includes proteins involved in otolith formation (East), vasculature and organ formation (Center) and blood coagulation (West) (Table S4).





**Figure 4.** Summary statistics calculated across 50 kb windows located in putatively neutral regions of the genome located in comparison with the top outliers (99th percentile) of the PBS analysis for each region (red dots). **Top:** Nucleotide diversity vs. recombination rate. **Bottom:** PBS vs. recombination rate.

#### 4. Discussion

In this study, we present the first analysis of population structure in European sardine across a large part of its distribution range using whole-genome sequencing data. A number of mechanisms have been suggested to explain how population structure can evolve in an environment without any complete physical barrier to gene flow, including local adaptation, habitat discontinuity, different habitat preferences and behavior, sexual selection, oceanographic currents, isolation by distance and limited dispersal capabilities [6,8,54–56].

Altogether, the assessment of nuclear genome sequences by means of individual ancestry information, principal component analysis (Figures 1B and 2) and differentiation ( $F_{ST}$ ) among populations from different geographic regions (Figure 3), supports that the European sardine comprises three main stocks: “West”, which includes individuals from Azores and Madeira (part of the Macaronesian region in the Atlantic), “Central”, which corresponds to Iberia (the center of the sampling distribution) and Northern Morocco, and “East”, which gathers the Mediterranean samples and those from the Canary Islands. The observed genetic differentiation between Mediterranean and Atlantic populations (except the Canary Islands) is in agreement with previous phenotypic and genetic studies based on mtDNA [18,23,26], suggesting the existence of a phylogeographic break between the South of Portugal and Mediterranean populations. The Almeria–Oran Front is likely mainly responsible for the reduction in gene flow among populations on each side, as previously observed in sardine [30] and other species [57,58]. In fact, the shared pattern of ancestry of the individuals from the Alboran Sea and the Gulf of Cadiz indicates that the Strait of Gibraltar is not such a strong barrier as previously suggested [27,29,30]. However, a small proportion of ancestry associated with the Central cluster can be observed across the Mediterranean, except for the Adriatic Sea (Figure 1B). The singularity of the sardine populations in the Northern Adriatic Sea has been suggested to result from a bottleneck

linked to heavy fishing combined with oceanographic isolation [59]. The impact of gene flow between the Center and Mediterranean populations could also explain the sharing of mitochondrial haplotypes (Figure S1).

The Western group has also been observed with microsatellite data [10], which revealed high differentiation between Azores/Madeira and the other Atlantic populations. Notably, populations from these two archipelagos cluster together genetically, despite Madeira being geographically closer to Canary Islands and almost at the same distance to Iberia as it is to Azores. This strongly suggests a barrier to gene flow between the region formed by these two archipelagos and the other populations analyzed in this study, including Canary Islands and Iberia. This genetic division can be caused by currents, isolation by distance and a lack of suitable habitat between these regions, local adaptation to different environmental conditions, or other reasons. The fact that we did not observe a pattern consistent with isolation by distance and that we excluded markers putatively under selection argues against the latter. Nonetheless, this needs further investigation.

The higher differentiation of sardine populations from Azores and Madeira is also clear in the mitogenome network (Figure S1). Although two other main clades are observed, they are formed by haplotypes from individuals with a very different nuclear-based ancestry. Thus, it is not easy to objectively pinpoint the geographic origin of these mtDNA clades.

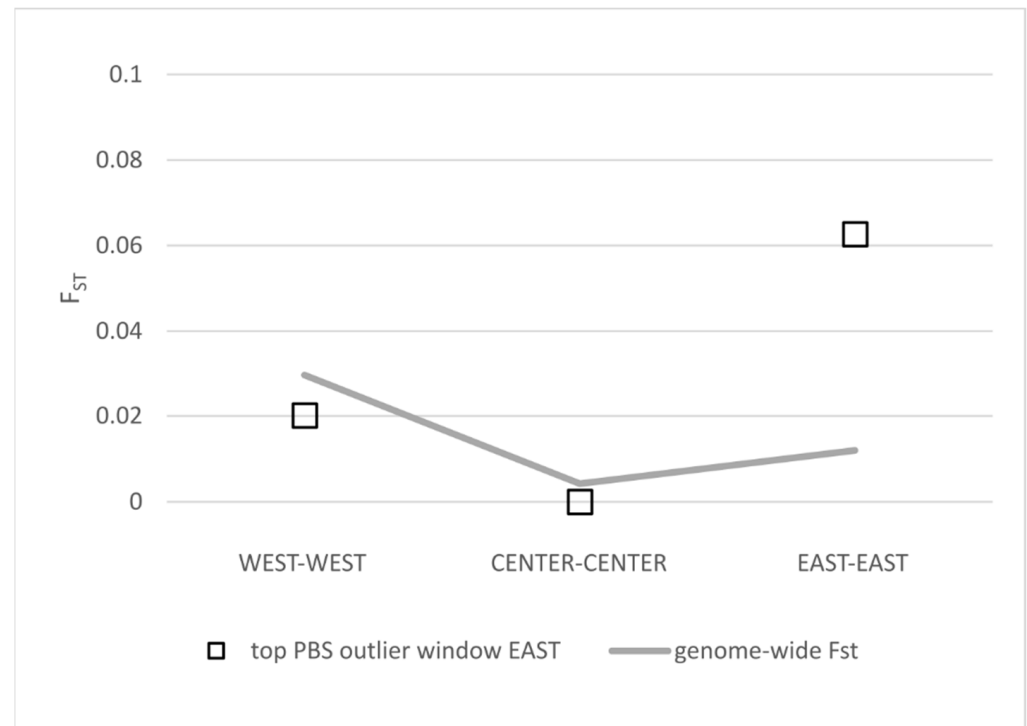
Discordance between differentially inherited markers can simply result from stochastic patterns of lineage sorting, but it can also be indicative of introgression [60]. Patterns suggesting admixture between the three genetic clusters were also observed with the nuclear data in all populations except Madeira. Given the lower effective population size of mtDNA when compared to nuclear DNA, we would expect to see it more sorted within each region.

An important piece of information that can help us to disentangle the role of gene flow versus shared ancestral polymorphism is the geographic pattern of differentiation. Genetic differentiation is lower between closer geographic populations within the East and Center clusters (Figure 3). Furthermore, we observed that the proportion of individuals with pure nuclear ancestry is higher in populations that are geographically more distant from populations with a different ancestry, suggesting that at least some of the patterns observed with nuclear and mtDNA genomes can indeed be created by gene flow between populations from these genetic clusters. Although this needs to be further confirmed using model-based approaches, if true, it provides additional support that the barriers involved in the differentiation between these three genetic clusters are only partial. Furthermore, the ancestry patterns observed between populations from the Central and Eastern clusters could suggest bidirectional gene flow between populations from Iberia and Mediterranean populations outside Iberia, which is also supported by the more even distribution of the two main mitochondrial haplotypes between these regions.

Western cluster ancestry is also observed in populations from the Center, Canary Islands and mainly in the Western Mediterranean populations. Although these patterns are compatible with admixture, gene flow between populations from the Eastern and Western clades is more difficult to explain. This discordance between molecular markers can also reflect the fact that regional populations of sardines seem to undergo periodic extinctions and recolonizations [61]. A recolonization of the Mediterranean from a refugium on the West African coast, as it has been suggested for anchovies [62], a species that shares several traits with sardines [63], could potentially explain the admixed ancestry of the Canary Islands and the Eastern cluster (Figure 3). The Canary Current upwelling system [64] could afterwards have been acting as a partial barrier to gene flow, maintaining the affiliation between the Canary Islands and the Mediterranean individuals. Some isolation between the archipelagos and the continent has also been found in other pelagic fish species using morphological traits [65–67].

Finally, we found that genomic regions corresponding to the top outliers of genetic differentiation are located in areas of low recombination (Figure 4), suggesting that genetic architecture can be contributing to some extent to the observed pattern of population

structure. Interestingly, one of these regions includes genes related to otolith formation. Otolith shapes have been found to divide Atlantic and Mediterranean sardines [27] and support subdivisions within the Mediterranean populations [68]. We observed that the  $F_{ST}$  values for this genomic region are very high compared to the average genome-wide value when including the Aegean population (Figure S3). This is also valid when comparing with the Adriatic population (Figure 5), which could potentially contribute to the observed regional anatomical divergence within the Mediterranean [68]. Future work, including larger sampling sizes and associated phenotype information, is required to assess how variation in specific regions of the genome affects phenotypes of interest for fisheries and stock assessment.



**Figure 5.** Comparison of  $F_{ST}$  values for the top PBS outlier window putatively associated with otolith development within three geographical regions: WEST (Azores vs. Madeira), CENTER (Bay of Biscay vs. Gulf of Cádiz) and EAST (Adriatic vs. Aegean).

## 5. Conclusions

Our main results provide evidence for three main genetic clusters of sardine populations across the analyzed specimens, suggesting at least two important barriers to gene flow. Although these do not seem complete, with gene flow possibly occurring between the three main phylogeographic regions identified, they seem to be strong enough to maintain populations genetically differentiated following their evolutionary trajectory. Our results thus offer an important baseline for further studies trying to identify the nature of these and other possible barriers between sardine populations, which can be compared with the phylogeographic patterns of other organisms with a similar distribution. Finally, the differentiation patterns reported here, together with the genetic resources generated for this commercially important species, offer information of strategic importance for transnational stock management of this highly exploited species towards sustainable fisheries.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes15020170/s1>.

**Author Contributions:** R.R.d.F. and L.F.C.C. designed the study; F.T., M.N., S.A., M.P., I.R., P.C., A.J.-R. and M.T.G.S. organized and executed the sample collection; P.F.C., A.R.-I. and E.F. performed

the laboratory work; R.R.d.F. analyzed the data with contributions from G.V.B., L.A.B., R.F., A.M.M., M.M.F.; R.R.d.F., P.F.C., E.F. and L.F.C.C. wrote the manuscript with contributions from all authors. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors gratefully acknowledge the following for funding their research: Villum Fonden Young Investigator Grant VKR023446 (R.D.F.); Fundação para a Ciência e a Tecnologia (FCT), Portugal, Scientific Employment Stimulus Initiative, grants CEECINST/00027/2021/CP2789/CT0003 to E.F., 10.54499/CEECINST/00133/2018/CP1510/CT0004 to L.F.C.C. and CEECIND/01799/2017 to P.F.C.; R.D.F. thanks the VILLUM FONDEN for the Center for Global Mountain Biodiversity (grant no 25925). M.P. and I.R. thank the Axencia Galega de Innovación (GAIN), Xunta de Galicia, Spain, for its funding of the AQUACOV and MERVEX Research Groups (grants IN607B 2018/14 and IN607-A 2018/4) and IMPRESS project supported by Spanish MICINN through grant RTI2018-099868-B-I00. E.F. research was funded by the project The Sea and the Shore, Architecture and Marine Biology: The Impact of Sea Life on the Built Environment Project No. POCI-01-0145-FEDER-029537, co-financed by COMPETE 2020, Portugal 2020 and the European Union through the European Regional Development Fund (ERDF). L.F.C.C. research was funded by: project VALORMAR (reference nr. 24517), supported by COMPETE2020, LISBOA2020, ALGARVE2020, PORTUGAL2020, through ERDF; strategic funding UIDB/04423/2020 through FCT and ERDF, in the framework of the program PT2020. We thank the scientific and technical staff and the crew of the PELACUS0315 and SARLINK oceanographic surveys conducted by the Instituto Español de Oceanografía (IEO, CSIC). Alboran Sea samples were collected during the SARLINK oceanographic survey. Samples from Galicia (Bay of Biscay South), Cantabrian Sea and Bay of Biscay were collected during the PELACUS 0315 Oceanographic survey, funded by the EU through the European Maritime and Fisheries Fund (EMFF) within the National Program of collection, management and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The Illumina fastq files raw reads are deposited in *NCBI Sequence Read Archive* (Bioproject accession no.: PRJNA688514, samples accession no.: SRR13325046, SRR13324980, SRR13324981, SRR13324982, SRR13324985, SRR13324986, SRR13324988, SRR13324989, SRR13324991, SRR13324992, SRR13324999, SRR13325000, SRR13325002, SRR13325003, SRR13325007, SRR13325008, SRR13325010, SRR13325011, SRR13325013, SRR13325014, SRR13325015, SRR13325018, SRR13325019, SRR13325021, SRR13325022, SRR13325023, SRR13325024, SRR13325025, SRR13325026, SRR13325027, SRR13325028, SRR13325029, SRR13325030, SRR13325031, SRR13325032, SRR13325033, SRR13325034, SRR13325035, SRR13325036, SRR13325037, SRR13325038, SRR13325039, SRR13325040, SRR13325041, SRR13325042, SRR13325043, SRR13325044, SRR13325045, SRR13325047, SRR13325048, SRR13325049, SRR13325050, SRR13325051, SRR13325052, SRR13325053, SRR13325054, SRR13325055, SRR13325056, SRR13325057, SRR13325058, SRR13325059, SRR13324977, SRR13324978, SRR13324979, SRR13324983, SRR13324984, SRR13324987, SRR13324990, SRR13324993, SRR13324994, SRR13324995, SRR13324996, SRR13324997, SRR13324998, SRR13325001, SRR13325005, SRR13325006, SRR13325004, SRR13325009, SRR13325012, SRR13325016, SRR13325017, SRR13325020, SRR13325060). The data will be public upon acceptance of the manuscript for publication.

**Acknowledgments:** All figures were edited in Inkscape (<http://www.inkscape.org/>). Thanks to Alessandro Laio, Amélia Fonseca, Ludovic Dickel, Patrícia Campos, Sara Rocha, Yorgos Athanasidis, and Sabour Brahim for supplying tissue samples. We would also like to thank Anders Albrechtsen, Katherine Richardson, Lounes Chikhi, Jonas Meisner, Jørgen Bendtsen, Rasmus Heller, Ricardo Pereira, and Stephen Sabatino for advice.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Faria, J.; Froufe, E.; Tuya, F.; Alexandrino, P.; Pérez-Losada, M. Panmixia in the endangered slipper lobster *scyllarides latus* from the Northeastern Atlantic and Western Mediterranean. *J. Crustac. Biol.* **2013**, *33*, 557–566. [[CrossRef](#)]
2. Andersson, L.; Bekkevold, D.; Berg, F.; Farrell, E.D.; Felkel, S.; Ferreira, M.S.; Fuentes-Pardo, A.P.; Goodall, J.; Pettersson, M. How Fish Population Genomics Can Promote Sustainable Fisheries: A Road Map. *Annu. Rev. Anim. Biosci.* **2023**, *12*. [[CrossRef](#)] [[PubMed](#)]

3. Palero, F.; Abelló, P.; Macpherson, E.; Gristina, M.; Pascual, M. Phylogeography of the European spiny lobster (*Palinurus elephas*): Influence of current oceanographical features and historical processes. *Mol. Phylogenet. Evol.* **2008**, *48*, 708–717. [[CrossRef](#)] [[PubMed](#)]
4. Pérez-Ruzafa, Á.; González-Wangüemert, M.; Lenfant, P.; Marcos, C.; García-Charton, J.A. Effects of fishing protection on the genetic structure of fish populations. *Biol. Conserv.* **2006**, *129*, 244–255. [[CrossRef](#)]
5. Graves, J. Molecular insights into the population structures of cosmopolitan marine fishes. *J. Hered.* **1998**, *89*, 427–437. [[CrossRef](#)]
6. Faria, R.; Johannesson, K.; Stankowski, S. Speciation in marine environments: Diving under the surface. *J. Evol. Biol.* **2021**, *34*, 4–15. [[CrossRef](#)] [[PubMed](#)]
7. Caballero-Huertas, M.; Frigola-Tepe, X.; Coll, M.; Muñoz, M.; Viñas, J. The current knowledge status of the genetic population structure of the European sardine (*Sardina pilchardus*): Uncertainties to be solved for an appropriate fishery management. *Rev. Fish. Biol. Fish.* **2022**, *32*, 745–763. [[CrossRef](#)]
8. Patarnello, T.; Volckaert, F.A.M.J.; Castilho, R. Pillars of Hercules: Is the Atlantic-Mediterranean transition a phylogeographical break? *Mol. Ecol.* **2007**, *16*, 4426–4444. [[CrossRef](#)]
9. Tine, M.; Kuhl, H.; Gagnaire, P.A.; Louro, B.; Desmarais, E.; Martins, R.S.; Hecht, J.; Knaust, F.; Belkhir, K.; Klages, S.; et al. European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nat. Commun.* **2014**, *5*, 5770. [[CrossRef](#)]
10. Kasapidis, P.; Silva, A.; Zampicinini, G.; Magoulas, A. Evidence for microsatellite hitchhiking selection in European sardine (*Sardina pilchardus*) and implications in inferring stock structure. *Sci. Mar.* **2012**, *76*, 123–132. [[CrossRef](#)]
11. Sá-Pinto, A.; Branco, M.; Sayanda, D.; Alexandrino, P. Patterns of colonization, evolution and gene flow in species of the genus *Patella* in the Macaronesian Islands. *Mol. Ecol.* **2008**, *17*, 519–532. [[CrossRef](#)]
12. Sala, I.; Caldeira, R.M.A.; Estrada-Allis, S.N.; Froufe, E.; Couvelard, X. Lagrangian transport pathways in the northeast Atlantic and their environmental impact. *Limnol. Oceanogr. Fluids Environ.* **2013**, *3*, 40–60. [[CrossRef](#)]
13. Francisco, S.M.; Faria, C.; Lengkeek, W.; Vieira, M.N.; Velasco, E.M.; Almada, V.C. Phylogeography of the shanny *Lipophrys pholis* (Pisces: Blenniidae) in the NE Atlantic records signs of major expansion event older than the last glaciation. *J. Exp. Mar. Biol. Ecol.* **2011**, *403*, 14–20. [[CrossRef](#)]
14. Stefanni, S.; Castilho, R.; Sala-Bozano, M.; Robalo, J.I.; Francisco, S.M.; Santos, R.S.; Marques, N.; Brito, A.; Almada, V.C.; Mariani, S. Establishment of a coastal fish in the Azores: Recent colonisation or sudden expansion of an ancient relict population? *Heredity* **2015**, *115*, 527–537. [[CrossRef](#)]
15. Hilmi, N.; Farahmand, S.; Lam, V.W.Y.; Cinar, M.; Safa, A.; Gilloteaux, J. The impacts of environmental and socio-economic risks on the fisheries in the mediterranean region. *Sustainability* **2021**, *13*, 10670. [[CrossRef](#)]
16. Braga, H.O.; Azeiteiro, U.M.; Oliveira, H.M.F.; Pardal, M.A. Evaluating fishermen’s conservation attitudes and local ecological knowledge of the European sardine (*Sardina pilchardus*), Peniche, Portugal. *J. Ethnobiol. Ethnomed.* **2017**, *13*, 25. [[CrossRef](#)] [[PubMed](#)]
17. ICES Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA). 2020. Available online: [https://ices-library.figshare.com/articles/report/Working\\_Group\\_on\\_Southern\\_Horse\\_Mackerel\\_Anchovy\\_and\\_Sardine\\_WGHANSA\\_/18621557](https://ices-library.figshare.com/articles/report/Working_Group_on_Southern_Horse_Mackerel_Anchovy_and_Sardine_WGHANSA_/18621557) (accessed on 22 January 2024).
18. Parrish, R.H.; Serra, R.; Grant, W.S. The Monotypic Sardines, *Sardina* and *Sardinops*: Their Taxonomy, Distribution, Stock Structure, and Zoogeography. *Can. J. Fish. Aquat. Sci.* **1989**, *46*, 2019–2036. [[CrossRef](#)]
19. Chlaida, M.; Laurent, V.; Kifani, S.; Benazzou, T.; Jaziri, H.; Planes, S. Evidence of a genetic cline for *Sardina pilchardus* along the Northwest African coast. *ICES J. Mar. Sci.* **2009**, *66*, 264–271. [[CrossRef](#)]
20. Chlaida, M.; Kifani, S.; Lenfant, P.; Ouragh, L. First approach for the identification of sardine populations *Sardina pilchardus* (Walbaum 1792) in the Moroccan Atlantic by allozymes. *Mar. Biol.* **2006**, *149*, 169–175. [[CrossRef](#)]
21. Laurent, V.; Caneco, B.; Magoulas, A.; Planes, S. Isolation by distance and selection effects on genetic structure of sardines *Sardina pilchardus* Walbaum. *J. Fish. Biol.* **2007**, *71*, 1–17. [[CrossRef](#)]
22. Spanakis, E.; Tsimenides, N.; Zouros, E. Genetic differences between populations of sardine, *Sardina pilchardus*, and anchovy, *Engraulis encrasicolus*, in the Aegean and Ionian seas. *J. Fish. Biol.* **1989**, *35*, 417–437. [[CrossRef](#)]
23. Atarhouch, T.; Rüber, L.; Gonzalez, E.G.; Albert, E.M.; Rami, M.; Dakkak, A.; Zardoya, R. Signature of an early genetic bottleneck in a population of Moroccan sardines (*Sardina pilchardus*). *Mol. Phylogenet. Evol.* **2006**, *39*, 373–383. [[CrossRef](#)]
24. Tinti, F.; Di Nunno, C.; Guarniero, I.; Talenti, M.; Tommasini, S.; Fabbri, E.; Piccinetti, C. Mitochondrial DNA sequence variation suggests the lack of genetic heterogeneity in the Adriatic and Ionian stocks of *Sardina pilchardus*. *Mar. Biotechnol.* **2002**, *4*, 163–172. [[CrossRef](#)]
25. Gonzalez, E.G.; Zardoya, R. Relative role of life-history traits and historical factors in shaping genetic population structure of sardines (*Sardina pilchardus*). *BMC Evol. Biol.* **2007**, *7*, 197. [[CrossRef](#)]
26. Andreu, B.B. Las branquispinas en la caracterización de las poblaciones de *Sardina pilchardus* (Walb). *Investig. Pesquera.* **1969**, *33*, 425–607.
27. Jemaa, S.; Bacha, M.; Khalaf, G.; Dessailly, D.; Rabhi, K.; Amara, R. What can otolith shape analysis tell us about population structure of the European sardine, *Sardina pilchardus*, from Atlantic and Mediterranean waters? *J. Sea Res.* **2015**, *96*, 11–17. [[CrossRef](#)]

28. Neves, J.; Silva, A.A.; Moreno, A.; Veríssimo, A.; Santos, A.M.; Garrido, S. Population structure of the European sardine *Sardina pilchardus* from Atlantic and Mediterranean waters based on otolith shape analysis. *Fish. Res.* **2021**, *243*, 106050. [[CrossRef](#)]
29. Alemany, F.; Álvarez, F. Growth differences among sardine (*Sardina pilchardus* Walb.) populations in Western Mediterranean. *Sci. Mar.* **1993**, *57*, 229–234.
30. Ramon, M.M.; Castro, J.A. Genetic variation in natural stocks of *Sardina pilchardus* (sardines) from the western Mediterranean Sea. *Heredity* **1997**, *78*, 520–528. [[CrossRef](#)]
31. Antoniou, A.; Manousaki, T.; Ramírez, F.; Cariani, A.; Cannas, R.; Kasapidis, P.; Magoulas, A.; Albo-Puigserver, M.; Lloret-Lloret, E.; Bellido, J.M.; et al. Sardines at a junction: Seascape genomics reveals ecological and oceanographic drivers of variation in the NW Mediterranean Sea. *Mol. Ecol.* **2023**, *32*, 1608–1628. [[CrossRef](#)]
32. Barry, P.; Broquet, T.; Gagnaire, P.-A. Age-specific survivorship and fecundity shape genetic diversity in marine fishes. *Evol. Lett.* **2022**, *6*, 46–62. [[CrossRef](#)] [[PubMed](#)]
33. Louro, B.; De Moro, G.; Garcia, C.; Cox, C.J.; Veríssimo, A.; Sabatino, S.J.; Santos, A.M.; Canario, A.V. A haplotype-resolved draft genome of the European sardine (*Sardina pilchardus*). *Gigascience* **2019**, *8*, giz059. [[CrossRef](#)]
34. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [[CrossRef](#)] [[PubMed](#)]
35. Machado, A.M.; Tørresen, O.K.; Kabeya, N.; Couto, A.; Petersen, B.; Felício, M.; Campos, P.F.; Fonseca, E.; Bandarra, N.; Lopes-Marques, M.; et al. “Out of the Can”: A Draft Genome Assembly, Liver Transcriptome, and Nutrigenomics of the European Sardine, *Sardina pilchardus*. *Genes* **2018**, *9*, 485. [[CrossRef](#)] [[PubMed](#)]
36. Pockrandt, C.; Alzamel, M.; Iliopoulos, C.S.; Reinert, K. GenMap: Ultra-fast computation of genome mappability. *Bioinformatics* **2020**, *36*, 3687–3692. [[CrossRef](#)] [[PubMed](#)]
37. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
38. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
39. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
40. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; Del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **2011**, *43*, 491–498. [[CrossRef](#)]
41. Korneliussen, T.S.; Albrechtsen, A.; Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinform.* **2014**, *15*, 356. [[CrossRef](#)]
42. Fox, E.A.; Wright, A.E.; Fumagalli, M.; Vieira, F.G. ngsLD: Evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics* **2019**, *35*, 3855–3856. [[CrossRef](#)]
43. Skotte, L.; Korneliussen, T.S.; Albrechtsen, A. Estimating individual admixture proportions from next generation sequencing data. *Genetics* **2013**, *3*, 693–702. [[CrossRef](#)]
44. Meisner, J.; Albrechtsen, A. Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics* **2018**, *210*, 719–731. [[CrossRef](#)]
45. Leigh, J.W.; Bryant, D. POPART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **2015**, *6*, 1110–1116. [[CrossRef](#)]
46. Bandelt, H.J.; Forster, P.; Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **1999**, *16*, 37–48. [[CrossRef](#)]
47. Korneliussen, T.S.; Moltke, I.; Albrechtsen, A.; Nielsen, R. Calculation of Tajima’s D and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinform.* **2013**, *14*, 289. [[CrossRef](#)]
48. Nielsen, R.; Korneliussen, T.; Albrechtsen, A.; Li, Y.; Wang, J. SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. *PLoS ONE* **2012**, *7*, e37558. [[CrossRef](#)] [[PubMed](#)]
49. Yi, X.; Liang, Y.; Huerta-Sanchez, E.; Jin, X.; Cuo, Z.X.; Pool, J.E.; Xu, X.; Jiang, H.; Vinckenbosch, N.; Korneliussen, T.S.; et al. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **2010**, *329*, 75–78. [[CrossRef](#)]
50. Dray, S.; Dufour, A.B. The ade4 Package: Implementing the Duality Diagram for Ecologists. *J. Stat. Softw.* **2007**, *22*, 1–20. [[CrossRef](#)]
51. Smedley, D.; Haider, S.; Ballester, B.; Holland, R.; London, D.; Thorisson, G.; Kasprzyk, A. BioMart—Biological queries made easy. *BMC Genom.* **2009**, *10*, 22. [[CrossRef](#)] [[PubMed](#)]
52. Barroso, G.V.; Puzović, N.; Dutheil, J.Y. Inference of recombination maps from a single pair of genomes and its application to ancient samples. *PLoS Genet.* **2019**, *15*, e1008449. [[CrossRef](#)]
53. Barroso, G.V.; Dutheil, J.Y. The landscape of nucleotide diversity in *Drosophila melanogaster* is shaped by mutation rate variation. *Peer Community J.* **2023**, *3*, e40. [[CrossRef](#)]
54. Alvarado Bremer, J.R.; Viñas, J.; Mejuto, J.; Ely, B.; Pla, C. Comparative phylogeography of Atlantic bluefin tuna and swordfish: The combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Mol. Phylogenet. Evol.* **2005**, *36*, 169–187. [[CrossRef](#)]

55. Díaz-Jaimes, P.; Uribe-Alcocer, M.; Rocha-Olivares, A.; García-de-León, F.J.J.; Nortmoon, P.; Durand, J.D.D. Global phylogeography of the dolphinfish (*Coryphaena hippurus*): The influence of large effective population size and recent dispersal on the divergence of a marine pelagic cosmopolitan species. *Mol. Phylogenet. Evol.* **2010**, *57*, 1209–1218. [[CrossRef](#)] [[PubMed](#)]
56. Kumar, R.; Kumar, V. A review of phylogeography: Biotic and abiotic factors. *Geol. Ecol. Landsc.* **2018**, *2*, 268–274. [[CrossRef](#)]
57. Pita, A.; Pérez, M.; Balado, M.; Presa, P. Out of the Celtic cradle: The genetic signature of European hake connectivity in South-western Europe. *J. Sea Res.* **2014**, *93*, 90–100. [[CrossRef](#)]
58. Ouagajjou, Y.; Presa, P. The connectivity of *Mytilus galloprovincialis* in northern Morocco: A gene flow crossroads between continents. *Estuar. Coast. Shelf Sci.* **2015**, *152*, 1–10. [[CrossRef](#)]
59. Ruggeri, P.; Splendiani, A.; Bonanomi, S.; Arneri, E.; Cingolani, N.; Santojanni, A.; Belardinelli, A.; Giovannotti, M.; Caputo, V. Temporal genetic variation as revealed by a microsatellite analysis of European sardine (*Sardina pilchardus*) archived samples. *Can. J. Fish. Aquat. Sci.* **2012**, *69*, 1698–1709. [[CrossRef](#)]
60. Lavretsky, P.; McCracken, K.G.; Peters, J.L. Phylogenetics of a recent radiation in the mallards and allies (Aves: Anas): Inferences from a genomic transect and the multispecies coalescent. *Mol. Phylogenet. Evol.* **2014**, *70*, 402–411. [[CrossRef](#)]
61. Grant, W.S.; Bowen, B.W. Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. *J. Hered.* **1998**, *89*, 415–426. [[CrossRef](#)]
62. Magoulas, A.; Castilho, R.; Caetano, S.; Marcato, S.; Patarnello, T. Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (*Engraulis encrasicolus*). *Mol. Phylogenet. Evol.* **2006**, *39*, 734–746. [[CrossRef](#)] [[PubMed](#)]
63. Checkley, D.M.; Asch, R.G.; Rykaczewski, R.R. Climate, Anchovy, and Sardine. *Ann. Rev. Mar. Sci.* **2017**, *9*, 469–493. [[CrossRef](#)] [[PubMed](#)]
64. Benazzouz, A.; Mordane, S.; Orbi, A.; Chagdali, M.; Hilmi, K.; Atillah, A.; Pelegrí, J.L.; Hervé, D. An improved coastal upwelling index from sea surface temperature using satellite-based approach—The case of the Canary Current upwelling system. *Cont. Shelf Res.* **2014**, *81*, 38–54. [[CrossRef](#)]
65. ICES Second workshop on Atlantic chub mackerel (*Scomber colias*) (WKCOLIAS2). *ICES Sci. Rep.* **2021**, *3*, 1–231.
66. Domínguez-Petit, R.; Navarro, M.R.; Cousido-Rocha, M.; Tornero, J.; Ramos, F.; Hernández, C.; Landa, J.; Jurado-Ruzafa, A.; Nunes, C.; Silva, A.V. Spatial variability of life-history parameters of the Atlantic chub mackerel (*Scomber colias*), an expanding species in the northeast Atlantic. *Sci. Mar.* **2022**, *86*, 4. [[CrossRef](#)]
67. Moreira, C.; Froufe, E.; Vaz-Pires, P.; Triay-Portella, R.; Correia, A.T. Landmark-based geometric morphometrics analysis of body shape variation among populations of the blue jack mackerel, *Trachurus picturatus*, from the North-East Atlantic. *J. Sea Res.* **2020**, *163*, 101926. [[CrossRef](#)]
68. Neves, J.; Veríssimo, A.; Múrias Santos, A.; Garrido, S. Comparing otolith shape descriptors for population structure inferences in a small pelagic fish, the European sardine *Sardina pilchardus* (Walbaum, 1792). *J. Fish Biol.* **2023**, *102*, 1219–1236. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.