


Article

Targeted Sequencing of Mitochondrial Genes Reveals Signatures of Molecular Adaptation in a Nearly Panmictic Small Pelagic Fish Species

Miguel Baltazar-Soares *, André Ricardo de Araújo Lima and Gonçalo Silva 

MARE-Marine and Environmental Sciences Centre, ISPA-Instituto Universitário, Rua Jardim do Tabaco 34, 1149-041 Lisboa, Portugal; alima@ispa.pt (A.R.d.A.L.); gsilva@ispa.pt (G.S.)

* Correspondence: miguelalexsoares@gmail.com

Abstract: Ongoing climatic changes, with predictable impacts on marine environmental conditions, are expected to trigger organismal responses. Recent evidence shows that, in some marine species, variation in mitochondrial genes involved in the aerobic conversion of oxygen into ATP at the cellular level correlate with gradients of sea surface temperature and gradients of dissolved oxygen. Here, we investigated the adaptive potential of the European sardine *Sardina pilchardus* populations offshore the Iberian Peninsula. We performed a seascape genetics approach that consisted of the high throughput sequencing of mitochondria's *ATP6*, *COI*, *CYTb* and *ND5* and five microsatellite loci on 96 individuals coupled with environmental information on sea surface temperature and dissolved oxygen across five sampling locations. Results show that, despite sardines forming a nearly panmictic population around Iberian Peninsula, haplotype frequency distribution can be explained by gradients of minimum sea surface temperature and dissolved oxygen. We further identified that the frequencies of the most common *CYTb* and *ATP6* haplotypes negatively correlate with minimum sea surface temperature across the sampled area, suggestive of a signature of selection. With signatures of selection superimposed on highly connected populations, sardines may be able to follow environmental optima and shift their distribution northwards as a response to the increasing sea surface temperatures.

Keywords: small pelagic fishes; OXPHOS complex; adaptive potential; climate change



Citation: Baltazar-Soares, M.; de Araújo Lima, A.R.; Silva, G. Targeted Sequencing of Mitochondrial Genes Reveals Signatures of Molecular Adaptation in a Nearly Panmictic Small Pelagic Fish Species. *Genes* **2021**, *12*, 91. <https://doi.org/10.3390/genes12010091>

Received: 4 December 2020

Accepted: 11 January 2021

Published: 13 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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/>).

1. Introduction

Ongoing climatic change translates into a series of environmental fluctuations with the potential to transform the current distribution of global biodiversity. As biological responses fall in the spectrum of “adapt, move or go extinct”, the range shifts, biological invasions and re-shaping of entire ecosystems, are thus expected to increase in frequency [1]. The marine environment is particularly susceptible to climatic shifts [1], and phenomena such as rising sea surface temperatures, increases in CO₂ concentrations (ocean acidification) and depletion of O₂ in deep oceanic waters, translate into tropicalization of fish assemblages in temperate zones [2,3] and hypoxia in coastal regions and estuaries [4,5]. Investigating the adaptive potential of species to cope with environmental shifts relies on the identification of phenotypic or molecular changes, expected to emerge as evolutionary responses [6]. Advances in sequencing technologies have greatly improved research on the molecular variation across non-model organisms [7]. Coupled with the extensively available environmental data that exists in marine regions, seascape genetic approaches are frequently being used to identify signatures of evolutionary processes [8,9].

On top of the environmental pressures imposed by climate change, marine species of commercial relevance are also subject to overexploitation. Unsustainable harvesting compromises the viability of stocks in multiple ways. For example, underestimating stock sizes leads to a logical reduction in the critical mass of spawners necessary to maintain

population dynamics [10]; capturing larger individuals imposes an early age at maturation, reducing the overall quality and quantity of brood stock [11,12]; disproportionately harvesting of unequal stocks may eliminate unique genetic diversity, thus reducing the species with adaptive potential, and therefore limiting the breath of possible responses to environmental pressures [13]. Coastal species with commercial value are heavily exposed to a synergistic effect of environmental and anthropogenic pressures. In this context, stocks of small pelagic fishes such as sardines, anchovies or herrings are among those that are particularly vulnerable [14]. “Small pelagic fishes” defines a group of ray-finned and fusiform short size fishes that tend to form large shoals that inhabit the water column near coastal locations. They mostly aggregate at the edge of continental shelves in order to benefit from the oxygen and nutrient-rich waters that emerge during upwelling events [15,16]. Due to their abundance, biomass, and trophic level occupied, those fish are an important component of an ecosystem that expands beyond the coastal limitations of their distribution [17–19].

The European sardine *Sardina pilchardus* is a small epipelagic fish inhabiting the north eastern Atlantic coast, the Mediterranean, and the Black Sea. It is a key component of the Mediterranean blue economy, supported by centuries-old tradition of consumption, and more recently, production of canned products [20]. Both the Mediterranean and Eastern Atlantic stocks along a coastal line from Senegal to Galicia, in Spain, are in a particularly fragile state [21–23]. Consequently, sardine fisheries have a recent, yet long record of fishing restrictions, quota limitations, and premature closures of fishing seasons.

In this work, we aim to investigate the adaptive potential of sardine populations surrounding the temperate zone of the Iberia Peninsula to fluctuations on sea surface temperature and dissolved oxygen. We focused on genetic variation at the mitochondrial level, as the mitochondrial genome encodes for genes involved on the oxidative phosphorylation (OXPHOS) pathway [24,25]. The central role of mitochondria in organismal metabolism questions longstanding assumptions of the organelle’s evolutionary neutrality [26]. It is thus not surprising that signatures of selection have been identified among the mitochondrial-encoded OXPHOS genes in several fish species, including Atlantic and Pacific salmons, anchovies, or sardines in the Indian Ocean, and associated with thermal gradients or dissolved oxygen concentration [26–28]. We followed a target-gene approach to screen the molecular diversity of sardine’s mitochondrial ATP6, CYTB, ND5 and COI, which code for complexes I,III–V of the OXPHOS pathway, and specifically sequenced key functional regions where signatures of selection have been detected in other marine fishes [26,27]. We also sampled a broad geographic area, enveloping the coastal waters of all the Iberian Peninsula, and consequently a gradient of coastal environments and putative barriers to dispersal. To explore the hypothesis that the frequency of mitochondrial variants among mitochondrial-encoded OXPHOS genes is related to the spatial variation of temperature and dissolved oxygen concentration, we collated environmental information from sampling locations. Finally, we complemented the survey of the species’ genetic diversity with five microsatellite markers, which allowed us to inspect the connectivity among sampled sites.

2. Methods

2.1. Sample Collection and DNA Extraction

A total of 96 individuals from 5 locations, Bay of Biscay (Spain), Gulf of Lion (France), Sesimbra (Portugal), Olhão (Portugal) and Tarragona (Spain) were collected, spanning the Atlantic-Mediterranean region surrounding the Iberian Peninsula (Figure 1). DNA was extracted with EZ-DNA Genomic DNA Isolation kit© (Zymo Research, California, CA, USA) following manufacturer instructions. Quantity and quality of extracted DNA were first assessed with Nanodrop© and later with Qubit© fluorometer.

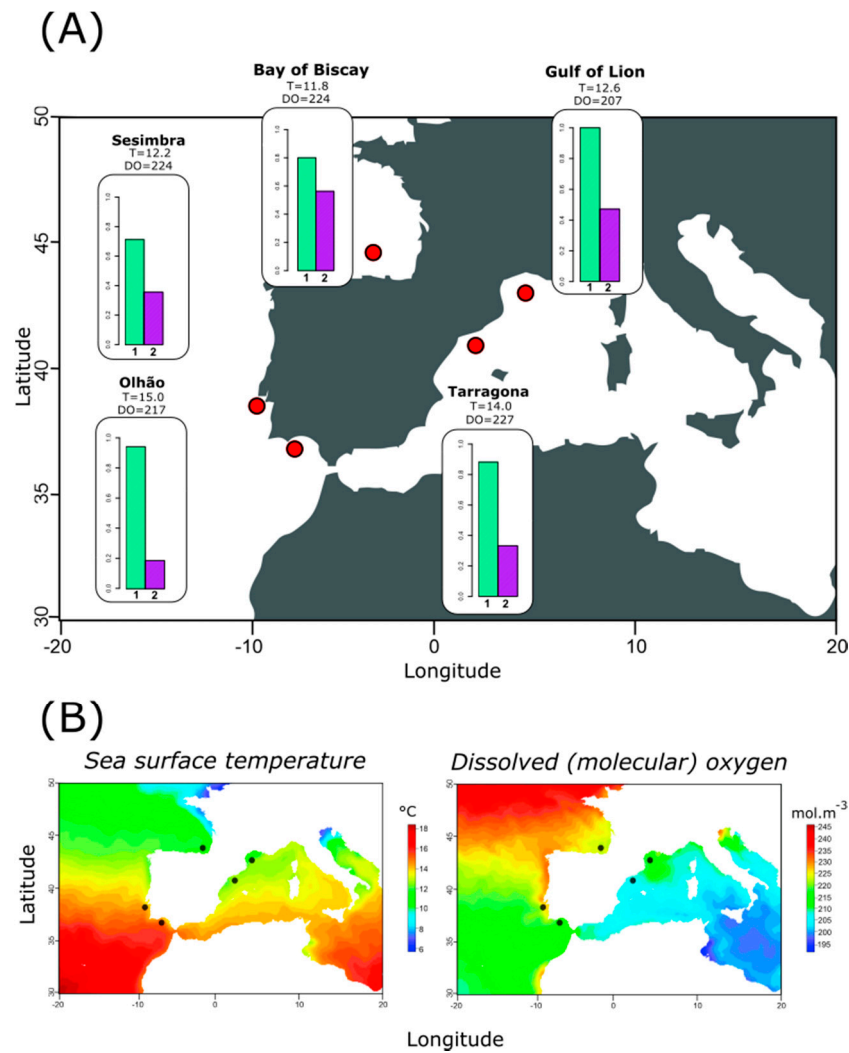


Figure 1. Sampling locations and associated environmental variation. **(A)** Sampling locations and plots of the frequencies (y -axis = frequency of the most common haplotype among ATP6, ATP6-Hap1 (1) and CYTB, CYTB-Hap3 (2)). The minimum sea surface temperature (T °C) and minimum dissolved oxygen ($\text{mol}\cdot\text{m}^{-3}$) obtained from each location is also shown. **(B)** Heatmap of minimum sea surface temperature and minimum dissolved oxygen distribution in the marine area surrounding the Iberian Peninsula. Black dots represent the sampling locations.

2.2. Mitochondrial DNA: Defining Target-Regions for Amplicon Sequencing

To explore whether the mitochondrial regions identified as putative under selection on other small pelagic fishes would exhibit similar signatures on sardines, 4 mitochondrial genes, each encoding different subunits involved in specific complexes of the mitochondrial respiratory chain were used: ND5 (NADH: ubiquinone oxidoreductase core subunit 5) for respiratory complex I; CYTB (cytochrome b) for respiratory complex III; COI (cytochrome c oxidase I) for respiratory complex IV; ATP6 (ATP synthase membrane subunit 6) for respiratory complex V. To facilitate the identification which within-gene regions should target sardines, we collected sequences from NCBI of other small pelagic species on which signatures of selection have been reported on our target genes. Those were then aligned with BioEdit [29] against the sardine mitogenome (reference NCBI: NC_009592.1) to verify where matching occurs. The conservation of chosen regions was further checked against other fish species. Primers were designed on the sardine genome to flank putative regions of interest. Reference numbers of utilized sequences are available on Table S1.

2.2.1. Sequencing, Filtering and Variant Calling at Targeted Mitochondrial Regions

Extracted DNA was sent to CIBIO (Centro de Investigação em Biodiversidade e Recursos Genéticos) in Vairão, Portugal, to primer optimization, library preparation and amplicon sequencing. The primer-pairs specifically designed for this study are available on Table S2. All obtained sequences are available in Text S1. Detailed protocols on library preparation and amplicon sequencing are available in Appendix A.1.

Demultiplexing, barcode removal and QC control were performed in Illumina's BaseSpace online platform. Only reads with Phred quality score (Q) > 30 were utilized in this study. Demultiplexed reads were aligned against a reference, i.e., sardine's mitogenome (reference NCBI: NC_009592.1) with *bwa-mem* algorithm [30] implemented in *bwa* [31]. Only reads that were properly paired were kept, mapping qualities above 10 and coverage above 2 were kept for variant calling. Variant calling was performed with Stacks v 1.48 [32] through imported .bam files. Specifically, we ran *ref_map.pl* pipeline with the sardine mitogenome as reference and default parameters.

2.2.2. Diversity and Distribution of Genetic Variation at Targeted Mitochondrial Regions

Haplotype diversity (Hd), number of haplotypes (nH), nucleotide diversity (π) and segregation sites (S) were calculated both independently for each gene in DNAsp [33]. In total, 4 datasets, corresponding to each sequenced mitochondrial region were built. To assess the connectivity among sampled locations, the population structure in Arlequin for all datasets was estimated via F_{ST} calculation (10,000 permutations). The level of significance was set to $\alpha = 0.05$ [34]. Isolation by distance was inferred by using log-transformed distance-by-sea among sampling locations and F_{ST} pairwise values in the R package *ecodist*, with the level of significance set to $\alpha = 0.05$ [35]. The number of shared haplotypes among locations was calculated to inspect whether geographic distance could play a role in the distribution of haplotype diversity. Finally, median-joining networks were calculated and drawn to better understand the genealogy among identified haplotypes in POPART [36].

2.2.3. Relationship between Haplotype Composition and Environmental Variables at Each Location

All computations were performed in the R environment [37]. A Redundancy Analysis (RDA) was performed to investigate whether haplotype distribution could be explained by environmental conditions found at each collection site. Environmental parameters (annual monthly means) were extracted from the Bio-Oracle database [38] with scripts customized to collect data on the maximum, minimum, mean and the range of values of sea surface temperature (SST) and dissolved oxygen (DO) given the approximate GPS positions of exact sampling locations. Thus, a matrix of haplotype frequencies was utilized as a response variable in an RDA [39], while a matrix of values related to surface temperature and dissolved oxygen were utilized as putative explanatory variables. Haplotypes unique to single locations were excluded from the dataset for this analysis.

The overall statistical significance of the RDA model was assessed using an ANOVA-like permutation function, *anova.cca*, implemented in the in the R package *vegan* [40]. Significance of constrained axis and effects and significance of each variable were also assessed with *anova.cca*, defining the "axis" and "margin" to avoid the sequential test of terms [41]. Lastly, a putative association of specific haplotype frequencies and environmental variation was inferred by considering, as significant, those relationships that diverged 2.5 standard deviations from the mean distribution on the RDA plot (i.e., two-tailed p -value = 0.0125).

2.2.4. Inferring Historical Events of Selection across Phylogenies with dN/dS Ratios-Based Tests

As we were interested in screening for signatures of selection on each specific respiratory complex, a phylogenetic inference of events of selection was performed independently

for each gene-specific dataset. As such, we utilized methods included in Hyphy via the Datamonkey web interface, which are based on the inference of nonsynonymous and synonymous substitution rates across phylogenies [42]. A single-likelihood ancestor counting (SLAC) [43], a fixed effect likelihood (FEL) [43] and fast unconstrained Bayesian approximation (FUBAR) [44] algorithms were applied. While sharing the same theoretical basis, i.e., selective pressure for each site is constant along the entire phylogeny; algorithms vary on the methodology to infer dN and dS substitution rates. Specifically, SLAC uses the maximum-likelihood and counting approach, FEL uses a maximum-likelihood approach and FUBAR relies on a Bayesian framework. Significance was assessed accordingly to the default value of each specific test, namely, with the level of significance set to $\alpha = 0.1$ (rounded-up to the decimal) in likelihood models SLAC and FEL, and with a posterior probabilities threshold of >0.9 (rounded-up to the decimal), for $(\alpha > \beta)$ in FUBAR (default).

2.3. Microsatellite Amplification and Diversity Estimates

Nuclear genetic variation was investigated with five microsatellites previously described in other sardine studies—SAB07, SAR09, SR15, SAR112, SAR218, and SP17 [45,46]. Amplification and genotyping protocols can be found in Appendix A.1. Nei's unbiased heterozygosity (H_e), observed heterozygosity (H_o), Wright's inbreeding coefficient (FIS) and Garza-Williams' index were computed in Arlequin v 3.5.2.2 [34]. Rarefied allelic richness (A_r) and private rarefied allelic richness (pA_r) were estimated for both each location and averaged over loci in HP-Rare 1.1 [47]. The presence of null alleles was inspected with Micro-Checker [48].

2.3.1. Estimates of Genetic Differentiation and Structure among Sampled Locations

We first assessed the likelihood of each locus to be in the Hardy-Weinberg equilibrium (HWE), as consistent deviations could compromise assessment of neutral differentiation. HWE was computed per loci and population in Arlequin v 3.5.2.2. The possibility of population structure was then investigated through multiple approaches. First, we utilized a Discriminant Analysis of Principal Components (DAPC) to explore the number of clusters (K) that most likely explains the observed distribution of nuclear diversity implemented in the R package adegenet [49]. This was achieved through the application of the function `find.clusters` with `max.n.clusters = 5`, identification of most likely K through the inflection point of the Bayesian Inference Criteria (BIC) curve and subsequent calculation of discriminant components. Lastly, we plotted the 95% confidence interval ellipses of the identified K groups overlaid on the original distribution. For that, we utilized the two major linear discriminants (LD1 and LD2) extracted from the discriminant analysis as graphical coordinates. We further estimated the pairwise F_{ST} among locations, derived from allelic frequencies in Arlequin v3.5.2.2 (10,000 bootstrap). Isolation by distance was also inspected and analyses were performed similarly to those described in the previous Section 2.2.2.

2.3.2. Inferences on Deviations from Evolutionary Neutrality

Additional locus-specific analyses were performed because one of the utilized markers (SAR112) was reported to be the subject of clinal selection along the Mediterranean [46]. In this context, first, the frequency of HWE deviation across locations for each locus was explored. Then, the locus-specific allele frequency distribution among populations was analyzed by comparing the average F_{ST} obtained from respective pairwise comparisons. Frequencies were compared with ANOVA and Tukey's post-hoc tests being applied to investigate the significance of pairwise relationships.

3. Results

3.1. Mitochondrial DNA Sequencing and Variant Calling

The 96 individuals sequenced for the four targeted genes produced a total of 912,609 usable reads. On average, 2223.84 (SE \pm 128.96) paired-end reads per individual were

utilized for the ATP6, 2599.17 (SE \pm 115.99) for the COI, 2693.29 (SE \pm 140.73) for the CYTB and 1990.04 (SE \pm 97.33) for the ND5. Average coverage per individual varied from 304.07 (SE \pm 18.80) on ATP6 to 208.19 (SE \pm 9.22) at CYTB. The position of variant sites identified with *Stacks* was inspected visually and utilized to construct haplotypes from the reference mitogenome. As genes were analyzed as fully independently, fragments were trimmed to remove uninformative invariant sites at the edges of the DNA string prior to population genomic analysis. Sizes of the fragments utilized varied across genes, specifically ATP6 (381 bp), COI (291 bp), CYTB (201 bp) and ND5 (330 bp).

3.2. Estimates of Mitochondrial Genetic Diversity and Population Structure

Genetic diversity indices expectedly varied among genes. Here, CYTB exhibited the higher indices of diversity ($\mu_S = 3.6$, $SE_S = 0.219$; $\mu_{nHap} = 4.6$, $SE_{nHap} = 0.219$; $\mu_{Hd} = 0.717$, $SE_{Hd} = 0.02$; $\mu_{\pi} = 6 \times 10^{-3}$, and $SE_{\pi} = 3 \times 10^{-4}$), while the COI showed a high conservation rate ($\mu_S = 1.4$, $SE_S = 0.219$; $\mu_{nHap} = 2.4$, $SE_{nHap} = 0.219$; $\mu_{Hd} = 0.164$, $SE_{Hd} = 0.02$; and $\mu_{\pi} = 5 \times 10^{-4}$, $SE_{\pi} = 8 \times 10^{-5}$). Among the sampled locations, genetic diversity was only pronouncedly different for the ATP6, as Tarragona and the Gulf of Lions reported a single haplotype in contrast to the Bay of Biscay and Sesimbra, where we observed three different haplotypes (Table 1). Estimates of pairwise F_{ST} only revealed two significant differentiation values, specifically for ATP6 haplotype frequencies between Sesimbra and the Gulf of Lions ($F_{ST} = 0.045$, $p = 0.042$) and ND5 between Tarragona and Olhão ($F_{ST} = 0.114$, $p = 0.033$). Investigating the overall number of haplotypes shared among locations revealed that closer sampled locations shared more haplotypes, suggesting that geographic distance might play a role in connectivity, despite the absence of clearly structured populations (Figure 2) and non-significant Mantel tests across genes (Table S3).

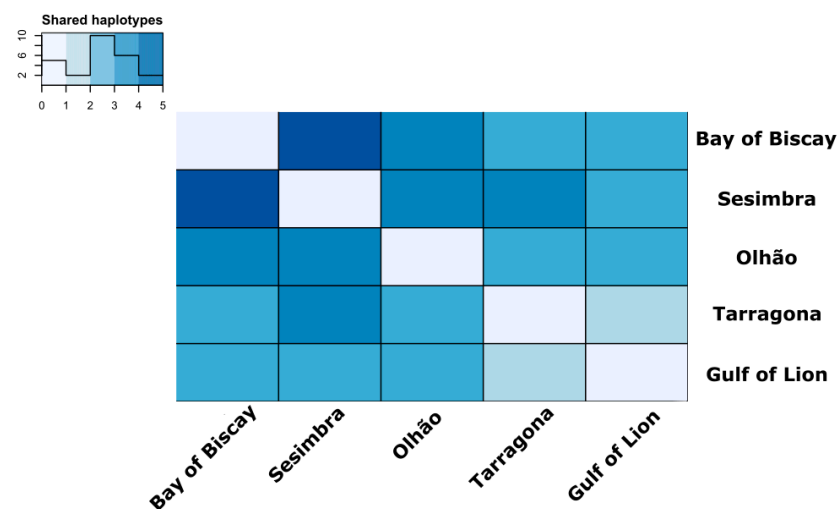


Figure 2. Heatmap of shared haplotypes. Quantification of the number of haplotypes shared by sampling locations. The legend refers to absolute numbers shared among locations, with the histogram line representing the frequency distribution.

Table 1. Diversity estimates obtained for each mitochondrial gene region coding and respective complex of the electron transport chain.

Gene	Population	<i>n</i>	S	nHap	Hd	π	Tajima's-D
ATP6 Respiratory Complex V	Bay of Biscay	15	3	4	0.371	0.00135	−1.31654
	Gulf of Lion	15	0	1	0	0	na
	Olhão	18	2	2	0.11	0.00058	−1.50776
	Sesimbra	14	3	4	0.396	0.0012	−1.67053
	Tarragona	17	0	1	0	0	na
COI Respiratory Complex IV	Bay of Biscay	17	2	3	0.228	0.00081	−1.50358
	Gulf of Lion	20	1	2	0.1	0.00034	−1.16439
	Olhão	15	1	2	0.133	0.00046	−1.15945
	Sesimbra	14	1	2	0.143	0.00049	−1.15524
	Tarragona	18	2	3	0.216	0.00076	−1.50776
CYTB Respiratory Complex III	Bay of Biscay	16	3	4	0.642	0.0051	0.38767
	Gulf of Lion	17	3	4	0.676	0.00563	0.78185
	Olhão	16	4	5	0.717	0.00589	−0.05743
	Sesimbra	14	4	5	0.769	0.00716	0.476
	Tarragona	15	4	5	0.781	0.00625	0.07027
ND5 Respiratory Complex I	Bay of Biscay	19	4	5	0.696	0.00284	−0.53717
	Gulf of Lion	20	1	2	0.521	0.00158	1.53133
	Olhão	19	2	3	0.526	0.0017	−0.04521
	Sesimbra	13	2	3	0.615	0.0021	0.2084
	Tarragona	18	4	5	0.66	0.00273	−0.67309

n = number of individuals per population; S = segregation sites; nHap = number of haplotypes; Hd = Haplotype diversity; π = nucleotide diversity.

3.3. Relating Haplotype Frequencies and Environmental Variation

The low number of different haplotypes in relation to explanatory variables prevented us building a single, full-scale model to incorporate maximum, minimum and averages of both dissolved oxygen and sea surface temperature. However, to preserve the fundamental objective of exploring and disentangling the role of those variables in sardine's distribution, we built four RDA models for each specific set of minima, maxima, mean and range. Only the model that incorporated minima was revealed to be statistically significant (ANOVA: $F = 2.84$, d.f. = 2, $p = 0.05$), and with an adjust $R^2 = 0.48$ (Figure 3A) (Table 2). CAP1 was revealed to significantly explain 43% of observed variability (ANOVA: $F = 3.343$, $p = 0.050$) and the marginal effect of each environmental variable showed minimum sea surface temperature (sst_min) to be significant (ANOVA: $F = 3.109$, $p = 0.033$) (Table 3).

We further identified haplotype 1 of ATP6 and haplotype 3 of CYTB to present a significant negative correlation with (minimum) sea surface temperature (Figure 3b). Models with average and maxima dissolved temperature and sea surface temperature did not significantly explain the haplotype frequencies distribution (Table 2).

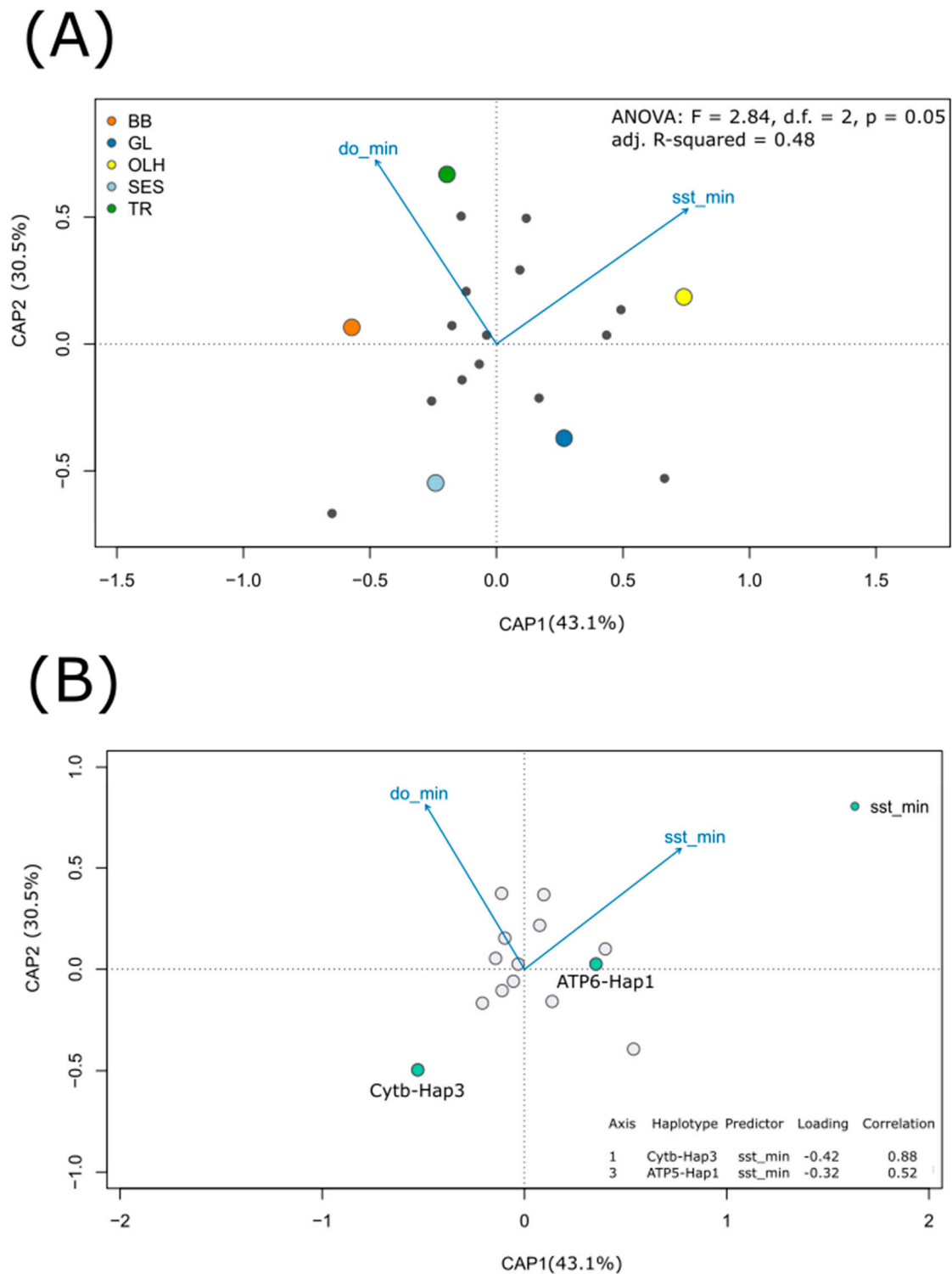


Figure 3. Presentation of redundancy analysis ordination with minimum values of environmental variables. (A) Ordination plot with haplotypes (response variables) as grey dots, populations in coloured dots and environmental variables (explanatory variables) as vectors. BB = Bay of Biscay, GL = Gulf of Lion, TR = Tarragona, OLH = Olhão, SES = Sesimbra; sst_min = minimum sea surface temperature, do_min = minimum dissolved oxygen. (B) Ordination plot highlighting haplotypes, Cytb-Hap3 and ATP6-Hap1, whose frequency is predicted to vary with minimum sea surface temperature (Cytb-Hap3: $r = 0.88$, $p < 0.05$; ATP6-Hap1: $r = 0.52$, $p < 0.05$), also showing respective axis and correlation coefficients.

Table 2. Model results obtained with the *anova.cca* function for each pair of parameter sea surface temperature (T °C) and dissolved oxygen (mol·m⁻³). Significant models are highlighted in bold.

Maximum Temperature and Dissolved Molecular Oxygen				
	df	SS	F	p (>F)
model	2	0.366	1.858	0.117
Residuals	2	0.197	-	-
Minimum Temperature and Dissolved Molecular Oxygen				
	df	SS	F	p (>F)
model	2	0.416	2.841	0.050
Residuals	2	0.146	-	-
Range of Temperature and Dissolved Molecular Oxygen				
	df	SS	F	p (>F)
model	2	0.150	1.087	0.483
Residuals	2	0.413	-	-
Mean of Temperature and Dissolved Molecular Oxygen				
	df	SS	F	p (>F)
model	2	0.202	0.560	0.867
Residuals	2	0.361	-	-

Model formula: Haplotype frequencies ~ SST+ DO. In bold, values significant for $p \leq 0.05$.

Table 3. ANOVA on the minimum SST and DO model to test the marginal effect of terms with 500 permutations to access.

	df	SS	F	p (>F)
CAP1	1	0.244	3.343	0.050
CAP2	1	0.171	2.339	0.150
Residual	2	0.146		
	df	SS	F	p (>F)
sst_min	1	0.230	3.109	0.033
do_min	1	0.202	2.756	0.058
Residual	2	0.146		

In bold, values significant for $p \leq 0.05$.

3.4. Inferences of Episodic Selection across the Phylogeny of Sampled Specimens

Investigating the occurrence of pervasive selection across the entire gene phylogenies revealed sites deviating from a neutral evolution background on all our target genes. SLAC detected six sites on CYTB, three of which revealed signatures of positive or directional selection and three others on negative or purifying selection. FEL and FUBAR, whose statistical power is more robust than SLACs, both identified episodes of pervasive selection on genes ATP6 and ND5. For ATP6, FEL detected site 53 ($p < 0.1$), while FUBAR detected sites 53, 49, 24, 36, and 52 (posterior probability of $\alpha > \beta$, ≥ 0.9). For ND5, FEL detected site 58 ($p = 0.05$), while FUBAR detected sites 58, 64, and 90 (posterior probability of $\alpha > \beta$, ≥ 0.9). The Bayesian approach further identified four sites on the COI, namely, 8, 44, 53, and 57 (posterior probability of $\alpha > \beta$, ≥ 0.9) (Table 4). All these signatures relate to negative/purifying selection.

Table 4. Phylogenetic inferences on episodic or pervasive historical events of selection and adopting three different algorithms.

Method	Gene	Codon	Site	<i>p</i> Value	pp [$\alpha > \beta$]
SLAC	CYTB	14	42	0.1	n/a
		15	45	0.1	n/a
		16	48	0.1	n/a
		26	78	0.1	n/a
		40	120	0.1	n/a
		48	144	0.1	n/a
		53	159	0.1	n/a
FEL	ATP6	53	159	0.1	n/a
	ND5	58	174	0.1	n/a
FUBAR	ATP6	24	72	n/a	0.9
		49	147	n/a	0.9
		36	108	n/a	0.9
		52	156	n/a	0.9
		53	159	n/a	0.9
	ND5	58	174	n/a	0.9
		64	192	n/a	0.9
		90	270	n/a	0.9
	COI	8	24	n/a	0.9
44		132	n/a	0.9	
53		159	n/a	0.9	
57		171	n/a	0.9	

Only sites whose deviations from the null hypothesis from neutral evolution were reported and are shown. Significance was determined with threshold for $p \leq 0.1$, in likelihood models SLAC and FEL, and with posterior probabilities in FUBAR. Only sites denoting significant values (in bold) are reported. Note that the *p*-value on SLAC relates to $P(dN/Ds < 1)$. All sites are reporting synonymous substitutions. n/a denotes statistic not available in the specific test, as defined in the Methods Section.

3.5. Microsatellite Amplification and Diversity Estimates

In general, we did not detect prominent variation regarding population specific diversity. Observed heterozygosity (H_o) values ranged between 0.65 ($SD \pm 0.27$) in Sesimbra and 0.74 ($SD \pm 0.16$) for the Gulf of Lion. Rarefied allelic richness over all loci varied between Sesimbra (12) and the Bay of Biscay (13.44). Private allelic richness, and indicator of population-specific alleles, was also similar from one location to another, being the extremes Tarragona (1.18) and the Bay of Biscay (2.24). Garza-Williams index, which estimates the effects of potential bottlenecks from observed heterozygosity, also indicated a striking similarity across populations (Table 5). Observed and expected heterozygosity (H_e) were revealed to significantly differ across sampled locations (ANOVA: $F_{1,4} = 6.02$, $p = 0.01$) (Figure S1A).

Table 5. Diversity estimates regarding the five microsatellite markers.

Locations	<i>n</i>	rAR	pA	H_o	H_e
Bay of Biscay	20	13.44	2.24	0.71	0.86
Gulf of Lion	20	12.41	1.80	0.73	0.83
Olhão	20	12.89	2.10	0.74	0.86
Sesimbra	16	12.00	1.78	0.65	0.84
Tarragona	20	12.18	1.18	0.63	0.80

n = number of individuals; rAR = rarefied allelic richness; pA = private alleles; H_o = Observed heterozygosity and H_e = Expected heterozygosity.

3.6. Estimates of Microsatellite Genetic Differentiation and Structure among Sampled Locations

Inferences on the structure of sardine populations revealed a nearly panmictic scenario. Discriminant component analyses suggested a $K = 2$, though 95% CI ellipses largely

overlap (Figure S2A). Pairwise F_{ST} comparisons further reinforced such a scenario as all but one showed a significant result, for $p = 0.02$ ($F_{ST}^{Bay\ of\ Biscay\ vs\ Gulf\ Lion} = 0.05$). While those locations are at the maximum geographic distance possible within our sampled sites, correlating geographic and genetic distances obtained no significant value. Lastly, STRUCTURE analyses also revealed panmixia (Figure S2B).

Regarding locus-specific analyses, we first detected locus SAR218 to be constantly departing from Hardy-Weinberg equilibrium: it was the only locus whose H_o significantly deviated from H_e in all populations. This is probably due to the presence of null alleles at this locus, which was found to be consistent across all examined populations (Table S4).

Exploring the structure and explaining each locus, we did not observe significant departures from what would be expected from panmixia: CIs of the 95% ellipses largely overlapped in every locus (Figure S3). Pairwise F_{ST} analyses suggested a very similar picture, with only one significant pairwise comparison between Sesimbra and Olhão for the locus SAR112 ($F_{ST} = 0.013$, $p = 0.036$). However, we detected averaged pairwise F_{ST} among comparisons and this varied significantly across loci (ANOVA: $F_{1,5} = 21.385$, $p < 0.01$) (Figure 1B). Post-hoc comparisons further revealed locus SAR218 to be the one originating in significantly higher average F_{ST} than all the other loci, Tukey HSD ($Avg_{SAR218} = 0.023$, $SD_{SAR218} = 0.005$).

4. Discussion

Environmental alterations driven by the ongoing rapid climatic shifts impose intense selective pressures on populations of marine organisms, and particularly on those that are currently exploited. Understanding the response repertoire of those species is crucial to develop effective and sustainable exploitation plans to accommodate potential shifts in the distribution range and overall impacts on the viability of populations. By screening the variation on genes directly involved in metabolic performance, we aimed to infer the adaptive potential of sardine populations to shifts in temperature and dissolved oxygen. We observed a high degree of connectivity among populations and further revealed a putative molecular basis of adaptation to cold temperatures. Under the predicted increase in sea surface temperature, our results indicate that sardines might be able to follow environmental optima and extend their distribution northwards.

4.1. Population Structure Reveals High Connectivity around Iberian Peninsula

Genetic diversity estimates across investigated markers suggested the Bay of Biscay harboring higher diversity among investigated populations. At the mitochondrial level, three out of four genes (CYTB being the exception) exhibited a higher number of haplotypes and/or nucleotide diversity, and microsatellite markers consistently showed a higher rarefied number of alleles and private alleles among sardines collected on the Bay of Biscay. The expected a positive correlation between genetic diversity and effective population size and may simply suggest that the Bay of Biscay assembles a larger number of reproducing individuals in comparison to all others [50]. The results are coherent with observations reporting the Atlantic sardine stocks holding higher abundances than those in the Mediterranean [51–53]. Alternatively, differences among observed diversity estimates may be associated with variable fishing efforts in relation to stock size, as overfished stocks are predicted to exhibit lower genetic diversity than those that are sustainably managed [54]. Population structure analyses revealed a scenario of near panmixia, punctuated by deviations likely associated to geographic distance among sampled sites. Specifically, not only nearby locations shared a higher number of haplotypes, but also the F_{ST} pairwise estimate associated with microsatellite data between the Bay of Biscay and the Gulf of Lion. As Mantel tests did not support a general pattern of isolation by distance, we consider having identified a signature of limited dispersal among sampled locations, possibly due to the natural barrier imposed by the Alboran Sea, as suggested by Ramon and Castro (1997) [55].

Generally, our observations are consensual within the existing studies dedicated to sardine population structure in the region, though the majority suggests a sardine panmictic

stock [22,51,56], and others point towards a soft structure on a specific region of the species distribution [55]. While our study did not aim to exhaustively characterize the population structure of the species, we believe that it reinforces the need to undertake genome-wide approaches and temporal sampling to convincingly resolve the stock structure of the European sardine.

4.2. Temperature-Dependent Distribution of Most Frequently Observed Haplotypes

Exploring the role of specific environmental factors on the distribution of sardine's mitochondrial diversity revealed that only minimum values of temperature and dissolved oxygen across sampled locations significantly explained the gradients of genetic variation. By showing that the distribution of the more frequently observed haplotypes of two of the targeted genes—ATP and CYTB—negatively correlated with minimum surface temperature, we can infer that the sardine stock in the screened geographical area is selected to tolerate colder sea surface temperatures. These results agree with previous works on mitochondrial DNA selection on other marine fishes, particularly on anchovies [27] and on Atlantic salmon [26], probably indicating a functional need to increase metabolic efficiency at lower temperatures [26]. North eastern Atlantic latitudes covered by our work are characterized by a strong seasonal upwelling, that creates oxygen- and nutrient-rich environments nearby coastal locations but originates from lower sea surface temperatures [57]. One of the anticipated effects of ongoing climate change on coastal upwelling systems is an increase in intensity of spring and summer upwelling events [57]. As the Iberian Peninsula is particularly sensitive to increases in upwelling intensity [58], we interpret the relationship identified in our work as a putative genomic signature of the species' response. Interestingly, historical signatures of purifying selection identified amongst the phylogenies of targeted genes were associated with low frequency haplotypes. As all variants are exposed to selection in haploid systems, such observation suggests the evolutionary dynamics of the sardine genetic pool purged specific haplotypes from the population due potentially deleterious effects [59,60]. While synonymous mutations do not alter the sequencing of encoded protein, they are known to impact secondary structures affecting thermodynamics and stability, as observed on mRNA molecules [61], misfolding of proteins [62] or translational selection and codon use bias in both proto and metazoans [63,64]. In addition, and specific to small pelagic fishes, the often large effective population sizes would reduce the efficiency of genetic drift on removing rare variants, but see Cvijović et al. (2018) for alternative theoretical expectations on site-frequency distribution under purifying selection [65]. Lastly, genetic signatures on mitochondrial genes of the OXPHOS complexes are also worth exploring at the light of cyto-nuclear co-evolution [66] or having their functional impact on organismal fitness effectively assessed given the existence of compensatory mechanisms to buffer the expression of deleterious alleles [67].

5. Conclusions and Future

Under and intense fishing pressure, the capacity for the European sardines to positively respond to ongoing climatic shifts must be resolved before developing practices of sustainable exploitation. Here, we focused on the European sardine stocks around the Iberian Peninsula and provided additional evidence for the existence of an apparent near panmictic population with a high degree of connectivity. This connectivity might be important to maintain frequency distribution of key mitochondrial haplotypes across the species distribution range, ensuring not only the species adaptive potential, but also an overall capacity to migrate to follow environmental optima: eggs and larvae of European sardine have been reported as North as the Baltic Sea [68]. Strong northward shifts are nevertheless expected for temperate-cold water species such as European sardines, according to species distribution models under future climatic conditions [69].

Shifts in thermal conditions are particularly stressful for ectotherms. Additionally, there are certainly other mechanisms involved in a putative adaptive response to thermal

variation. For instances, adaptive and non-adaptive phenotypic plasticity are known to play a decisive role in assisting organisms to cope with shifting temperatures [70,71]. Epigenetic variation is also another possible route towards adaptation, and evidence of such keeps on accumulating among both natural populations and under controlled experiments [72–74]. Though we interpreted the relationship between haplotype frequency and minimum sea surface temperature as a signature of selection, we cannot disregard the potential effect of small sample sizes or the fact that we did not screen the edges of specie’s distribution. Still, with the seascape genetics approach we undertook, we are confident that our study provides a robust glimpse of sardines’ adaptive potential and is an important contribution to the growing body of evidence supporting adaptation among populations of small pelagic fishes.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2073-4425/12/1/91/s1>, Figure S1: Comparison of diversity and differentiation estimates obtained with microsatellite loci. (A) Expected (H_e) and observed heterozygosity (H_o); (B) Averaged F_{ST} per locus considering scores of the five populations, Figure S2: Population differentiation analyses produced with microsatellite loci. (A) DAPC plots considering all loci—ellipses are at 95% C.I.; (B) STRUCTURE assignment plots for $K = 5$, Figure S3: DAPC plots considering each locus separately—ellipses are at 95% C.I, Table S1: NCBI references sequences utilized to design mitochondrial primers, Table S2: Primers utilized to amplify mitochondrial genes, Table S3: Mantel tests for each mitochondrial gene and overall microsatellite loci, Table S4: Micro-checker reports for each loci, per population. In bold, loci where null alleles were detected. Text S1: sequences in *fasta* format of mitochondrial’s amplified regions obtained in this work.

Author Contributions: Conceptualization, M.B.-S. and G.S.; methodology, M.B.-S. and G.S.; validation, M.B.-S., G.S. and A.R.d.A.L.; formal analysis, M.B.-S. and A.R.d.A.L.; investigation, M.B.-S., G.S. and A.R.d.A.L.; resources, G.S.; data curation, M.B.-S.; writing—original draft preparation, M.B.-S.; writing—review and editing, M.B.-S., G.S. and A.R.d.A.L.; visualization, M.B.-S. and A.R.d.A.L.; supervision, G.S.; project administration, G.S.; funding acquisition, G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study had the support of Fundação para a Ciência e Tecnologia (FCT), through the strategic project UIDB/04292/2020 granted to MARE, and the project PTDC/BIA-BMA/32209/2017, co-financed by FCT through national funds and by the European Union through FEDER (European Regional Development Fund).

Data Availability Statement: Sequence data is available as supplementary file. Sequences are deposited in NCBI under the accession numbers MW437948–MW438277.

Acknowledgments: The authors would like to acknowledge to Naiara Rodriguez-Ezpeleta, Jordi Viñas, Marta Frigola, Xènia Tepe, Joana Boavida, Emílie Villar, David Piló, Laurie Casalot, Didier Aurelle and Fauvelle Vincent for their availability, and their collaboration on collecting and shipping tissue samples. The authors also thank Flying Sharks for transporting samples collected on the Gulf of Lions and Susana Costa Lopes (CIBIO) for sequencing and library preparation.

Conflicts of Interest: The authors declare no conflict of interests.

Appendix A

Appendix A.1. Library Preparation and Amplicon Sequencing Protocol

Multiplexed Illumina MiSeq amplicon libraries for the four mitochondrial regions were amplified using a two-stage PCR approach [75]. First-stage PCR conditions differed among genes. The ATP6 fragment was amplified in a 10 ML reaction with 5 ML of MyTaq® Mix (BioLine, UK), 0.4 mM of each primer with overhangs as in McInnes et al. [76] and 10 ng of DNA template. Thermal cycling was performed in a T100 Thermal Cycler® (BioRad) consisting of an initial step at 95 °C for 10', followed by 40 cycles at 95 °C for 30'', 58 °C for 30'' and 72 °C for 30'' and a final extension of 72 °C for 10'. Amplification of the remaining genes was performed under the following conditions: an initial step at 95 °C for 10', followed by 95 °C for 30'', 68 °C –0.5 °C/cycle for 30'', 72 °C for 20'',

followed by 31 cycles at 95 °C for 30", 64 °C for 30", 72 °C for 20", and a final extension of 72 °C for 10". A template-free PCR was included in each amplification to control for potential contamination. PCR products were purified using reversible immobilization (SPRI) paramagnetic beads[®] (Agencourt AMPure XP, Beckman Coulter, CA, USA) with 0.8× of beads per microliter of PCR product.

Individual tags from McInnes et al. (2017) [76] were added to each sample PCR products obtained from first-stage amplifications. Second-stage PCRs were prepared with 12.5 ML of KAPA HiFi HotStart ReadyMixPCR Kit[®] (Kapa Biosystems, Wilmington, MA, USA), 2.5 ML of 1 mM of each tag primer and 2.5 ML of 1:10 diluted PCR product in 25 ML final volume. Thermal cycler conditions were as following: 95 °C for 3', followed by 10 cycles of 95 °C for 30", 55 °C for 30", 72 °C for 30", and a final extension of 72 °C for 5'. PCR products were purified using reversible immobilization (SPRI) paramagnetic beads.

All samples were then pooled after normalizing amplified quantities to 10 nM. Amplicon's size distribution was analysed on a TapeStation 2200 using the HS D1000 kit (Agilent Technologies, Santa Clara, CA, USA). Libraries concentration was checked by quantitative PCR on a BIORAD C1000 Real Time Thermo cycler using KAPA Library Quantification kit (Kapa Biosystems, Wilmington, MA, USA). Amplicon libraries were sequenced on a MiSeq genome sequencer (Illumina, San Diego, CA, USA) using MiSeq Reagent Kit Nano v2 (500 cycles) with paired-end reads.

Appendix A.2. Amplification and Genotyping of Microsatellite Loci

Amplification was performed with a Qiagen Multiplex PCR kit (Hilden, Germany) in a single multiplex under the following conditions: initial denaturation step at 95 °C for 15 min, 17 cycles comprising an initial step of at 95 °C for 30 s, an annealing step at 56 °C for 1 min and 30 s (with an increase of 0.5 °C/cycle during the first 13 cycles) and an elongation at 72 °C for 30 s, followed by 27 cycles which included a denaturation at 95 °C for 30 s, annealing at 50 °C for 1 min and an elongation step at 72 °C for 30 s. Final elongation occurred at 60 °C for 30 min. PCR products were separated by capillary electrophoresis on an automatic sequencer ABI3130xl Genetic Analyzer (AB Applied Biosystems, Foster City, CA, USA). Fragments were scored against the GeneScan-500 LIZ Size Standard using the GENEMAPPER 4.1 (Applied Biosystems) and manually checked twice.

References

1. Calosi, P.; De Wit, P.; Thor, P.; Dupont, S. Will life find a way? Evolution of marine species under global change. *Evol. Appl.* **2016**, *9*, 1035–1042. [[CrossRef](#)] [[PubMed](#)]
2. Costa, B.H.E.; Assis, J.; Franco, G.; Erzini, K.; Henriques, M.; Gonçalves, E.; Caselle, J.; Gonçalves, E.J. Tropicalization of fish assemblages in temperate biogeographic transition zones. *Mar. Ecol. Prog. Ser.* **2014**, *504*, 241–252. [[CrossRef](#)]
3. Vergés, A.; Steinberg, P.D.; Hay, M.E.; Poore, A.G.B.; Campbell, A.H.; Ballesteros, E.; Heck, K.L.; Booth, D.J.; Coleman, M.A.; Feary, D.A.; et al. The tropicalization of temperate marine ecosystems: Climate-mediated changes in herbivory and community phase shifts. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20140846. [[CrossRef](#)] [[PubMed](#)]
4. Melzner, F.; Thomsen, J.; Koeve, W.; Oschlies, A.; Gutowska, M.A.; Bange, H.W.; Hansen, H.P.; Körtzinger, A. Future ocean acidification will be amplified by hypoxia in coastal habitats. *Mar. Biol.* **2012**, *160*, 1875–1888. [[CrossRef](#)]
5. Altieri, A.H.; Gedan, K.B. Climate change and dead zones. *Glob. Chang. Biol.* **2014**, *21*, 1395–1406. [[CrossRef](#)] [[PubMed](#)]
6. Eizaguirre, C.; Baltazar-Soares, M. Evolutionary conservation—Evaluating the adaptive potential of species. *Evol. Appl.* **2014**, *7*, 963–967. [[CrossRef](#)]
7. Nielsen, E.E.; Hemmer-Hansen, J.; Larsen, P.F.; Bekkevold, D. Population genomics of marine fishes: Identifying adaptive variation in space and time. *Mol. Ecol.* **2009**, *18*, 3128–3150. [[CrossRef](#)]
8. Grummer, J.A.; Beheregaray, L.B.; Bernatchez, L.; Hand, B.K.; Luikart, G.; Narum, S.R.; Taylor, E.B. Aquatic Landscape Genomics and Environmental Effects on Genetic Variation. *Trends Ecol. Evol.* **2019**, *34*, 641–654. [[CrossRef](#)]
9. Selkoe, K.A.; Henzler, C.M.; Gaines, S.D. Seascape genetics and the spatial ecology of marine populations. *Fish Fish.* **2008**, *9*, 363–377. [[CrossRef](#)]
10. Hutchinson, W.F. The dangers of ignoring stock complexity in fishery management: The case of the North Sea cod. *Biol. Lett.* **2008**, *4*, 693–695. [[CrossRef](#)]
11. Ernande, B.; Dieckmann, U.; Heino, M. Adaptive changes in harvested populations: Plasticity and evolution of age and size at maturation. *Proc. R. Soc. B Biol. Sci.* **2004**, *271*, 415–423. [[CrossRef](#)] [[PubMed](#)]

12. Dunlop, E.S.; Eikeset, A.M.; Stenseth, N.C. From genes to populations: How fisheries-induced evolution alters stock productivity. *Ecol. Appl.* **2015**, *25*, 1860–1868. [[CrossRef](#)] [[PubMed](#)]
13. Spielman, D.; Brook, B.W.; Frankham, R. Most species are not driven to extinction before genetic factors impact them. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15261–15264. [[CrossRef](#)] [[PubMed](#)]
14. Deyle, E.R.; Fogarty, M.; Hsieh, C.-H.; Kaufman, L.; MacCall, A.D.; Munch, S.B.; Perretti, C.T.; Ye, H.; Sugihara, G. Predicting climate effects on Pacific sardine. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6430–6435. [[CrossRef](#)]
15. Cury, P.; Bakun, A.; Crawford, R.J.M.; Jarre, A.; Quiñones, R.A.; Shannon, L.J.; Verheye, H.M. Small pelagics in upwelling systems: Patterns of interaction and structural changes in “wasp-waist” ecosystems. *ICES J. Mar. Sci.* **2000**, *57*, 603–618. [[CrossRef](#)]
16. Rykaczewski, R.R.; Checkley, D.M. Influence of ocean winds on the pelagic ecosystem in upwelling regions. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1965–1970. [[CrossRef](#)]
17. Koslow, J.A.; Davison, P.; Lara-Lopez, A.; Ohman, M.D. Epipelagic and mesopelagic fishes in the southern California Current System: Ecological interactions and oceanographic influences on their abundance. *J. Mar. Syst.* **2014**, *138*, 20–28. [[CrossRef](#)]
18. Yamaguchi, A.; Matsuno, K.; Abe, Y.; Arima, D.; Imai, I. Latitudinal variations in the abundance, biomass, taxonomic composition and estimated production of epipelagic mesozooplankton along the 155° E longitude in the western North Pacific during spring. *Prog. Oceanogr.* **2017**, *150*, 13–19. [[CrossRef](#)]
19. Duarte, L.O.; Garcia, C. Trophic role of small pelagic fishes in a tropical upwelling ecosystem. *Ecol. Model.* **2004**, *172*, 323–338. [[CrossRef](#)]
20. González-García, S.; Villanueva-Rey, P.; Belo, S.; Vázquez-Rowe, I.; Moreira, M.T.; Feijoo, G.; Arroja, L. Cross-vessel eco-efficiency analysis. A case study for purse seining fishing from North Portugal targeting Euro-pean pilchard. *Int. J. Life Cycle Assess.* **2015**, *20*, 1019–1032. [[CrossRef](#)]
21. Vázquez-Rowe, I.; Villanueva-Rey, P.; Hospido, A.; Moreira, M.T.; Feijoo, G. Life cycle assessment of European pilchard (*Sardina pilchardus*) consumption. A case study for Galicia (NW Spain). *Sci. Total Environ.* **2014**, *475*, 48–60. [[CrossRef](#)] [[PubMed](#)]
22. 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](#)]
23. Antonakakis, K.; Giannoulaki, M.; Machias, A.; Somarakis, S.; Sanchez, S.; Ibaibarriaga, L.; Uriarte, A. Assessment of the sardine (*Sardina pilchardus* Walbaum, 1792) fishery in the eastern Mediterranean basin (North Aegean Sea). *Mediterr. Mar. Sci.* **2011**, *12*, 333–357. [[CrossRef](#)]
24. Letts, J.A.; Fiedorczuk, K.; Sazanov, J.A. The architecture of respiratory supercomplexes. *Nat. Cell Biol.* **2016**, *537*, 644–648. [[CrossRef](#)] [[PubMed](#)]
25. Formosa, L.E.; Ryan, M.T. Mitochondrial OXPHOS complex assembly lines. *Nat. Cell Biol.* **2018**, *20*, 511–513. [[CrossRef](#)] [[PubMed](#)]
26. Consuegra, S.; John, E.; Verspoor, E.; de Leaniz, C.G. Patterns of natural selection acting on the mitochondrial genome of a locally adapted fish species. *Genet. Sel. Evol.* **2015**, *47*, 58. [[CrossRef](#)]
27. Silva, G.; Lima, F.P.; Martel, P.; Castilho, R. Thermal adaptation and clinal mitochondrial DNA variation of European anchovy. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20141093. [[CrossRef](#)]
28. Sebastian, W.; Sukumaran, S.; Zacharia, P.; Muraleedharan, K.; Kumar, P.D.; Gopalakrishnan, A. Signals of selection in the mitogenome provide insights into adaptation mechanisms in heterogeneous habitats in a widely distributed pelagic fish. *Sci. Rep.* **2020**, *10*, 1–14.
29. Hall, T.A. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. In *Nucleic Acids Symposium Series*; Information Retrieval Ltd.: London, UK, 1999; Volume 41, pp. 95–98.
30. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
31. Li, H.; Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **2010**, *26*, 589–595. [[CrossRef](#)]
32. Catchen, J.; Hohenlohe, P.A.; Bassham, S.; Amores, A.; Cresko, W.A. Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **2013**, *22*, 3124–3140. [[CrossRef](#)] [[PubMed](#)]
33. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [[CrossRef](#)]
34. Excoffier, L.; Lischer, H.E.L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **2010**, *10*, 564–567. [[CrossRef](#)] [[PubMed](#)]
35. Goslee, S.C.; Urban, D.L. The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.* **2007**, *22*, 1–19. [[CrossRef](#)]
36. Leigh, J.W.; Bryant, D. POPART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **2015**, *6*, 1110–1116. [[CrossRef](#)]
37. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2013.
38. Assis, J.; Tyberghein, L.; Bosch, S.; Verbruggen, H.; Serrão, E.A.; De Clerck, O. Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Glob. Ecol. Biogeogr.* **2018**, *27*, 277–284. [[CrossRef](#)]
39. Rao, C.R. The use and interpretation of principal component analysis in applied research. *Sankhyā Indian J. Stat. Ser. A* **1964**, *26*, 329–358.

40. Wickham, H.; Francois, R.; Henry, L.; Müller, K. Vegan: Community Ecology Package. R package version 1.17-4. 2010. Available online: <https://cran.r-project.org/web/packages/vegan/index.html> (accessed on 15 July 2020).
41. Legendre, P.; Oksanen, J.; ter Braak, C.J.F. Testing the significance of canonical axes in redundancy analysis. *Methods Ecol. Evol.* **2011**, *2*, 269–277. [[CrossRef](#)]
42. Weaver, S.; Shank, S.D.; Spielman, S.J.; Li, M.; Muse, S.V.; Kosakovsky Pond, S.L. Datamonkey 2.0: A Mod-ern Web Application for Characterizing Selective and Other Evolutionary Processes. *Mol. Biol. Evol.* **2018**, *35*, 773–777. [[CrossRef](#)]
43. Pond, S.L.K.; Frost, S.D.W. Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. *Mol. Biol. Evol.* **2005**, *22*, 1208–1222. [[CrossRef](#)]
44. Murrell, B.; Moola, S.; Mabona, A.; Weighill, T.; Sheward, D.; Pond, S.L.K.; Scheffler, K. FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection. *Mol. Biol. Evol.* **2013**, *30*, 1196–1205. [[CrossRef](#)] [[PubMed](#)]
45. Ruggeri, P.; Splendiani, A.; Bonanomi, S.; Arneri, E.; Cingolani, N.; Santojanni, A.; Belardinelli, A.; Giovannotti, M.; Barucchi, V.C. 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](#)]
46. 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.* **2011**, *76*, 123–132. [[CrossRef](#)]
47. Kalinowski, S.T. HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Notes* **2005**, *5*, 187–189. [[CrossRef](#)]
48. Van Oosterhout, C.; Hutchinson, W.F.; Wills, D.P.M.; Shipley, P. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **2004**, *4*, 535–538. [[CrossRef](#)]
49. Jombart, T. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [[CrossRef](#)] [[PubMed](#)]
50. Frankham, R.; Briscoe, D.A.; Ballou, J.D. *Introduction to Conservation Genetics*; Cambridge University Press: Cambridge, UK, 2002.
51. Baibai, T.; Oukhattar, L.; Quinteiro, J.V.; Mesfioui, A.; Rey-Mendez, M. First global approach: Morphological and biological variability in a genetically homogeneous population of the European pilchard, *Sardina pilchardus* (Walbaum, 1792) in the North Atlantic coast. *Rev. Fish Biol. Fish.* **2012**, *22*, 63–80. [[CrossRef](#)]
52. Silva, A.; Carrera, P.; Masse, J.; Uriarte, A.; Santos, M.; Oliveira, P.B.; Soares, E.; Porteiro, C.; Stratoudakis, Y. Geographic variability of sardine growth across the northeastern Atlantic and the Mediterranean Sea. *Fish. Res.* **2008**, *90*, 56–69. [[CrossRef](#)]
53. Van Beveren, E.; Fromentin, J.-M.; Rouyer, T.; Bonhommeau, S.; Brosset, P.; Saraux, C. The fisheries history of small pelagics in the Northern Mediterranean. *ICES J. Mar. Sci.* **2016**, *73*, 1474–1484. [[CrossRef](#)]
54. Pinsky, M.L.; Palumbi, S.R. Meta-analysis reveals lower genetic diversity in overfished populations. *Mol. Ecol.* **2014**, *23*, 29–39. [[CrossRef](#)]
55. Ramon, M.; Castro, J. Genetic variation in natural stocks of *Sardina pilchardus* (sardines) from the western Mediterranean Sea. *Heredity* **1997**, *78*, 520–528. [[CrossRef](#)]
56. Correia, A.T.; Hamer, P.; Carocinho, B.; Silva, A. Evidence for meta-population structure of *Sardina pilchardus* in the Atlantic Iberian waters from otolith elemental signatures of a strong cohort. *Fish. Res.* **2014**, *149*, 76–85. [[CrossRef](#)]
57. Bakun, A.; Black, B.A.; Bograd, S.J.; Garcia-Reyes, M.; Miller, A.J.; Rykaczewski, R.R.; Sydeman, W.J. Anticipated effects of climate change on coastal upwelling ecosystems. *Curr. Clim. Chang. Rep.* **2015**, *1*, 85–93. [[CrossRef](#)]
58. Miranda, P.; Alves, J.; Serra, N. Climate change and upwelling: Response of Iberian upwelling to atmospheric forcing in a regional climate scenario. *Clim. Dyn.* **2013**, *40*, 2813–2824. [[CrossRef](#)]
59. Avery, P.J. The population genetics of haplo-diploids and X-linked genes. *Genet. Res.* **1984**, *44*, 321–341. [[CrossRef](#)]
60. Levin, L.; Zhidkov, I.; Gurman, Y.; Hawlena, H.; Mishmar, D. Functional Recurrent Mutations in the Human Mitochondrial Phylogeny: Dual Roles in Evolution and Disease. *Genome Biol. Evol.* **2013**, *5*, 876–890. [[CrossRef](#)] [[PubMed](#)]
61. Chamary, J.V.; Hurst, L.D. Evidence for selection on synonymous mutations affecting stability of mRNA secondary structure in mammals. *Genome Biol.* **2005**, *6*, R75. [[CrossRef](#)]
62. Drummond, D.A.; Wilke, C.O. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* **2008**, *134*, 341–352. [[CrossRef](#)]
63. Castellana, S.; Vicario, S.; Saccone, C. Evolutionary Patterns of the Mitochondrial Genome in Metazoa: Exploring the Role of Mutation and Selection in Mitochondrial Protein-Coding Genes. *Genome Biol. Evol.* **2011**, *3*, 1067–1079. [[CrossRef](#)]
64. Uddin, A.; Chakraborty, S. Synonymous codon usage pattern in mitochondrial CYB gene in pisces, aves, and mammals. *Mitochondrial DNA Part A* **2015**, *28*, 187–196. [[CrossRef](#)]
65. Cvijović, I.; Good, B.H.; Desai, M.M. The Effect of Strong Purifying Selection on Genetic Diversity. *Genetics* **2018**, *209*, 1235–1278. [[CrossRef](#)] [[PubMed](#)]
66. Barreto, F.S.; Watson, E.T.; Lima, T.G.; Willett, C.S.; Edmands, S.; Li, W.; Burton, R.S. Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nat. Ecol. Evol.* **2018**, *2*, 1250–1257. [[CrossRef](#)] [[PubMed](#)]
67. Hill, G.E. Mitonuclear Compensatory Coevolution. *Trends Genet.* **2020**, *36*, 403–414. [[CrossRef](#)] [[PubMed](#)]
68. Alheit, J.; Pohlmann, T.; Casini, M.; Greve, W.; Hinrichs, R.; Mathis, M.; O’Driscoll, K.; Vorberg, R.; Wagner, C. Climate variability drives anchovies and sardines into the North and Baltic Seas. *Prog. Oceanogr.* **2012**, *96*, 128–139. [[CrossRef](#)]
69. Schickele, A.; Goberville, E.; Leroy, B.; Beaugrand, G.; Hattab, T.; Francour, P.; Raybaud, V. European small pelagic fish distribution under global change scenarios. *Fish Fish.* **2021**, *22*, 212–225. [[CrossRef](#)]

70. Schulte, P.M.; Healy, T.M.; Fanguie, N.A. Thermal Performance Curves, Phenotypic Plasticity, and the Time Scales of Temperature Exposure. *Integr. Comp. Biol.* **2011**, *51*, 691–702. [[CrossRef](#)]
71. Yampolsky, L.Y.; Schaer, T.M.M.; Ebert, D. Adaptive phenotypic plasticity and local adaptation for temperature tolerance in freshwater zooplankton. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20132744. [[CrossRef](#)]
72. McCaw, B.A.; Stevenson, T.J.; Lancaster, L.T. Epigenetic Responses to Temperature and Climate. *Integr. Comp. Biol.* **2020**, *60*, 1469–1480. [[CrossRef](#)]
73. Best, C.; Ikert, H.; Kostyniuk, D.J.; Craig, P.M.; Navarro-Martin, L.; Marandel, L.; Mennigen, J.A. Epigenetics in teleost fish: From molecular mechanisms to physiological phenotypes. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **2018**, *224*, 210–244. [[CrossRef](#)]
74. Heckwolf, M.J.; Meyer, B.S.; Häsler, R.; Höppner, M.P.; Eizaguirre, C.; Reusch, T.B. Two different epigenetic information channels in wild three-spined sticklebacks are involved in salinity adaptation. *Sci. Adv.* **2020**, *6*, eaaz1138. [[CrossRef](#)]
75. Binladen, J.; Gilbert, M.T.P.; Bollback, J.P.; Panitz, F.; Bendixen, C.; Nielsen, R.; Willerslev, E. The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing. *PLoS ONE* **2007**, *2*, e197. [[CrossRef](#)] [[PubMed](#)]
76. McInnes, J.C.; Jarman, S.N.; Lea, M.-A.; Raymond, B.; Deagle, B.E.; Phillips, R.A.; Catry, P.; Stanworth, A.; Weimerskirch, H.; Kusch, A. DNA metabarcoding as a marine conservation and management tool: A circumpolar examination of fishery discards in the diet of threatened albatrosses. *Front. Mar. Sci.* **2017**, *4*, 277. [[CrossRef](#)]