**Sardines at a junction: seascape genomics reveals ecological and oceanographic drivers of variation in the NW Mediterranean Sea**

Aglaia Antoniou1, Tereza Manousaki1, Francisco Ramírez2, Alessia Cariani3, Rita Cannas4, Panagiotis Kasapidis1, Antonios Magoulas1, Marta Albo-Puigserver2,5, Elena Lloret-Lloret2, Jose Maria Bellido6, Maria Grazia Pennino7, Maria Cristina Follesa4, Antonio Esteban6, Claire Saraux8, Mario Sbrana9, Maria Teresa Spedicato10, Marta Coll2, Costas S. Tsigenopoulos1\*

1: Institute of Marine Biology, Biotechnology and Aquaculture, Hellenic Centre for Marine Research, Gournes Pediados, Irakleio, P.O. Box 2214, 71003 Crete, Greece

2: Institute of Marine Science (ICM-CSIC), Passeig Marítim de la Barceloneta, nº 37-49, 08003, Barcelona, Spain.

3: Dept. Biological, Geological & Environmental Sciences (BiGeA) Alma Mater Studiorum – Università di Bologna Via Sant’Alberto 163, 48123 Ravenna, Italy - Consorzio Nazionale Interuniversitario per le Scienze del Mare (CoNISMa), Rome, Italy.

4: Department of Life and Environmental Sciences, University of Cagliari, via T. Fiorelli 1-09126 Cagliari, Italy- Consorzio Nazionale Interuniversitario per le Scienze del Mare (CoNISMa), Rome, Italy.

5: Centro de Ciências do Mar, Universidade do Algarve (CCMAR-UAlg), Campus de Gambelas, 8005-139, Faro, Portugal.

6: Instituto Español de Oceanografía, Centro Oceanográfico de Murcia, C/ Varadero 1, 30740 San Pedro del Pinatar (Murcia), Spain.

7: Instituto Español de Oceanografía. Centro Oceanográfico de Vigo. Subida a Radio Faro 50-52, 36390 Vigo, Pontevedra, Spain.

8: MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Sète, France.

9: Consorzio per il Centro Interuniversitario di Biologia Marina ed Ecologia Applicata G. Bacci. Viale Nazario Sauro 4 - 57128 Livorno- Italy.

10: COISPA Tecnologia & Ricerca. Via dei trulli 18. 70126 Bari-Italy.

\* Corresponding author: Costas S. Tsigenopoulos, e-mail: tsigeno@hcmr.gr

**Abstract**

By evaluating genetic variation across the entire genome, one can address existing questions in a novel way while new can be asked. Such questions include how different local environments influence adaptive and neutral genomic variation within and among populations, providing insights into local adaptation of natural populations, their responses to global change and exploitation-induced evolution. Here, under a seascape genomic approach, ddRAD data were used along with environmental information to detect signals of adaptive divergence across environmental gradients with gene flow in European sardines (*Sardina pilchardus*) of the Western Mediterranean and adjacent Atlantic waters. The studied sardines constitute two clusters, a pattern confirmed even when only adaptive (outlier) loci were in use, highlighting the importance of local adaptation. The trend in the number of days with sea surface temperature (SST) above 19oC, a critical threshold for successful sardine spawning, was crucial at all levels of population structuring with implications on species’ key biological processes. Our findings provide evidence for a dynamic equilibrium where population structure is maintained by physical and biological factors under the opposing influences of migration and selection. This dynamic in nature system, postulates a continuous monitoring under a seascape genomic approach that can benefit by incorporating a temporal as well as a more detailed spatial dimension. Our results may contribute further studies aimed at providing deeper insights into the mechanistic processes underlying population structuring; key for the understanding and prediction of future changes and responses of this highly exploited species in the face of climate change.

**Keywords**

Climate change, ddRAD, outliers, population genomics, SNPs

**Introduction**

Identifying genomic signatures of selection provide insights not only into local adaptation of natural populations but also into their adaptive responses to global change and the exploitation-induced evolution through harvesting (Nielsen et al., 2009). This can be better achieved through population genomic studies, by separating the effects of neutral processes (drift-migration) from those of selection, where frequency shifts point to rapid changes in environmental conditions (Hoffmann & Willi, 2008).

Modern genomic tools provide the means to assess the population structure of species in high resolution and get a detailed view of their status (Greenbaum, Rubin, Templeton, & Rosenberg, 2019) providing new means for examining the genetic basis of local adaptation (Price, Lopez, Platts, & Lasky, 2020).

Combining state-of-the-art genomic approaches with spatially-explicit information on the main habitat features allows to link genotypes to environmental variables, within the field of seascape genomics (Riginos, Crandall, Liggins, Bongaerts, & Treml, 2016), enhancing our understanding on the processes affecting the inhabitant populations, including adaptive responses to environmental fluctuations reflecting either environmental heterogeneity or climate change (Selkoe et al., 2010). This can provide novel insights into the observed patterns of spatial genetic structure (Galindo, Olson, & Palumbi, 2006). but also valuable clues on how these populations cope with environmental change, in the face of ocean warming (Liggins, Treml, & Riginos, 2019).

The Mediterranean Sea has experienced a long history of anthropogenic perturbations (Coll et al., 2012) strongly affected by marine debris (García-Rivera, Lizaso, & Millán, 2017; UNEP/MAP, 2015), with one of the world’s highest percentage of stocks fished at unsustainable levels (FAO, 2020). In addition to the human-induced loss of habitat and biodiversity through direct exploitation (Coll, Steenbeek, Ben Rais Lasram, Mouillot, & Cury, 2015), indirect effects such as climate change and Lessepsian migration/invasive species, pose additional threats that will likely increase in severity in the future ( Pennino et al., 2020; Ramírez, Afán, Davis, & Chiaradia, 2017). Climate change and fisheries combined are currently threatening Mediterranean fish stocks ( Peristeraki et al., 2019; Ramírez et al., 2021).

A shift towards sustainable fisheries in the Mediterranean Sea is more urgent than ever and requires robust fisheries and stock assessments, which need that biological and management units are aligned (Vasilakopoulos, Maravelias, & Tserpes, 2014) to preserve stocks with adaptive diversity that can contribute to the long-term maintenance of productive fisheries and ecosystems (Antoniou & Magoulas, 2013).

Because of their commercial value and ecological role within marine communities, small pelagic fish are important species in the Mediterranean Sea (Coll et al., 2019; FAO, 2018; Palomera et al., 2007). Overall, declines of commercially important small pelagic fish populations, particularly of the European sardine (*Sardina pilchardus*, hereafter sardine), have been observed in the Mediterranean Sea mostly attributed to the combined effects of climate impacts and fishing pressure (Coll et al., 2019; Pennino et al., 2020; Ramírez, Coll, Navarro, Bustamante, & Green, 2018; Ramírez et al., 2021; Saraux et al., 2019).

Sardines as other clupeids, although highly mobile, display localized spawning behaviour and migratory patterns that may result in restricted gene flow between populations and subsequent divergence (Bacha, Jemaa, Hamitouche, Rabhi, & Amara, 2014). This has already been demonstrated within the Mediterranean and to a lesser extend in the eastern Atlantic, mainly reflecting the levels of environmental heterogeneity of the two regions (Caballero-Huertas, Frigola-Tepe, Coll, Muñoz, & Viñas, 2022 and references therein). The prevailing conditions are distinct in the different parts of the area in terms of productivity, circulation patterns, wind events, biogeochemical and biological components but also topographic features. Genetic studies conducted on sardines in the Mediterranean employing different types of markers have mostly reported shallow phylogeographic structure, low genetic differentiation and weak or absent population structure, with signs of late Pleistocene expansion (Kasapidis, 2014 and references therein). Significant genetic structure was only revealed in one particular locus and a microsatellite probably linked to it. The pattern observed coincided with a genetic cline tightly connected to the local thermal and hydrodynamic characteristics of the Atlantic-Mediterranean junction, (Chlaida et al., 2009; Kasapidis, Silva, Zampicinini, & Magoulas, 2012; Laurent, Caneco, Magoulas, & Planes, 2007). Within the Mediterranean, the Almeria-Oran oceanographic front (AOF) is a key oceanographic feature, situated at the North-eastern Alboran Sea that separates the Mediterranean waters from inflowing Atlantic waters.Its hydrological characteristics are closer to those of the North-Eastern Atlantic than to the Western Mediterranean (Bacha et al., 2014), indicative of enhanced gene flow between the Alboran Sea and the Atlantic (Northeast and Moroccan Atlantic Ocean, Baibai et al., 2012). AOF constitutes a barrier to sardines’ dispersal that could be held responsible for founder effects, followed by genetic drift with selection acting on the local scale driving adaptations mostly related to minimum sea surface temperature (Baltazar-Soares, Lima, & Silva, 2021).

Despite the wealth of previous studies, the role of multiple stressors in shaping sardines’ stock structure within the Mediterranean Sea remains elusive, thus prohibiting inferences of local adaptive variation and the potential of populations to respond to environmental changes with implications for the conservation and sustainable management of this declining fish stock.

The observed spatial heterogeneity is expected to generate adaptive divergence by imposing challenges to the organisms inhabiting the area further enhanced by the changing environmental conditions (Grummer et al., 2019; Sandoval-Castillo & Beheregaray, 2020). The latter has been held responsible for the recent decline of sardine’s recruitment (Santos et al., 2018) which in conjunction with spawning and larval behavior constitute the main drivers of recruitment variability (Lowerre-Barbieri et al., 2017 and references therein).

Considering all the above, and given the fragmented information on sardines biology, ecology and evolution (Coll & Bellido, 2019), with this study we aim to provide a better understanding of the processes affecting sardine populations and to define environmental variables with high impact in shaping the species genomic patterns. Given that populations may show evident signatures of local adaptation regardless of gene flow (Diopere et al., 2018; Moody et al., 2015) and that adaptive variation is expected to be maintained in cases where the selection of locally adapted alleles is stronger than the influx of non-locally adapted alleles (Bal et al., 2021), we conducted an exploratory analysis aimed at finding relevant variables to sardine’s adaptive potential as well as to its susceptibility to climate change. Analyzing genome wide SNPs within a seascape genomics approach, we evaluated the influence of climatological ocean conditions and their trends, as well as landscape characteristics, biological components and exploitation based on different gear types on the spatial patterns of population genetic structure across the environmentally heterogeneous region of Central – Western Mediterranean and adjacent Atlantic waters. This set of variables include both well-established as significant for the species according to literature as well as variables relevant to the species biology that were never tested before. Our approach enabled the assessment of the species population structure as well as genome-environment associations by disentangling the patterns of adaptive and neutral genetic variation. We further attempted to explain whether observed patterns in size and body condition i.e. the North/South gradient in size and body condition observed in the Western Mediterranean (Albo-Puigserver et al., 2021; Bachiller et al., 2020; Brosset et al., 2017; Coll & Bellido, 2019) reflect random genetic processes and demography, or selection imposed by environmental forcing (both the biotic and abiotic attributes of the areas under study), or their combined effect.

The methodology and findings of this study can contribute to better understand processes shaping population differentiation and to evaluate the adaptive potential of species in response to environmental change. It is also anticipated to shed light on how the interplay between natural selection and gene flow affects local adaptation in this broadcast spawner, especially at times of increasing anthropogenic influence on the natural world (Flanagan, Forester, Latch, Aitken, & Hoban, 2018). This study can offer valuable information for sustainable species management providing important clues on the definition and delineation of stocks and their potential to adapt. Finally, it may assist the prediction of the evolutionary responses to the ongoing and future climate change scenarios.

**Materials & Methods**

*Sampling, DNA extraction, library construction and genotyping*

Sardines were collected from 12 different locations across the Western and Central Mediterranean Sea, and the adjacent Atlantic waters (Figure 1), during scientific surveys (Mediterranean International bottom Trawl Survey, MEDITS and MEDiteranean International Acoustic Survey, MEDIAS) or commercial hauls in the period November 2017 – June 2018. A total of 398 fish (166 females, 220 males, and 6 immatures and 6 n/a) were collected. Samples from two neighboring GSA sites showed differences in respect to fish total length and weight (Figure 1, Table S1).

Genomic DNA was extracted from tissue samples and libraries were prepared following the double-digest restriction site-associated DNA sequencing approach of Kess et al. (2015). Sequenced reads were analyzed using STACKS v.2.4 pipeline (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013), to quality control the reads, identify the genomic loci sequenced, genotype each individual, and conduct basic population genetic analyses against the sardines’ genome (Louro et al., 2019). A detailed description on sampling, library construction and genotyping is provided at the Supporting Information.

*Population genetic analyses*

Population genetic structure was assessed by both Bayesian and multivariate ordination methods. First a model-based clustering was performed with STRUCTURE 2.3.4 (Pritchard, Stephens, & Donnelly, 2000) under the correlated allele frequency model allowing admixture, without location prior and a varying number of clusters (K = 1-8) with 10 replicate runs at each K. The inference of K was evaluated with different methods and replicate runs were averaged to identify sets of highly similar runs, as well as distinct modes in the space of possible solutions with CLUMPAK server (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015). In cases where K > 1, samples allocated to clusters with high membership coefficient (q ≥ 0.9) were further analyzed with STRUCTURE for each cluster and with the same settings as above, in an attempt to examine, based on non-admixed individuals, whether further sub-structuring occurs.

Second, a Discriminant Analysis of Principal Components (DAPC) was performed with Adegenet v. 2.1.1 package (Jombart, 2008), in the R environment (R Core Team, 2020) as means to infer population subdivision of the samples under study, with an independent of population genetics model method.

To determine the number of expected genetic clusters (K) present in the dataset, without any *a priori* population definition, the *find.clusters* function included in adegenet was used to run successive numbers of K-means clusters of the individuals, across a range of K = 1–8. We identified the best supported number of clusters through comparison of the Bayesian Information Criterion (BIC) for the different values of K.

Genetic diversity between the 12 sampling localities and the identified clusters was then compared in terms of observed (Ho) and expected heterozygosity (He), as means of genetic differentiation and levels of gene flow using STACKS population program. Furthermore, levels of differentiation among localities and clusters were assessed by the fixation index FST, statistical significance was evaluated through 10,000 permutations with the software strataG R package (Archer, Adams, & Schneiders, 2017), and P values were adjusted for multiple testing through false discovery rate (FDR) correction (Benjamini & Hochberg, 1995) using the ‘p.adjust’ function in R v.3.6.0 (R Core Team, 2020). A detailed description of the followed procedure so far is given at Supporting Information.

*Seascape genomic analyses and outlier detection*

*The datasets*

Four types of genomic datasets were analyzed as to detect whether the different levels of population structure and the observed contrast in physical condition are driven by distinct environmental drivers. The first one included all studied samples (“all samples” i.e. a dataset that displayed population structure, see Results). The second and third datasets corresponded to sardines’ population clusters discovered in the studied area i.e. “Atlantic” [ATL] and “Mediterranean” [MED] clusters, respectively (see Results ). Finally, the fourth dataset contained samples from northern and southern sampling sites of the Western Mediterranean Sea (“northern vs southern sites” i.e. GSA07a, GSA07b, GSA06a versus GSA06c) that displayed remarkable differences in their physical condition i.e. fish length and weight. This difference had a North/South gradient with larger and heavier sardines found in the southernmost areas of the Western Mediterranean (Table S1, Figure 1). Although those differences might be attributed to the sampling strategy itself (both time and gear employed for sampling), as well as the inclusion of six immature samples at the northern sites, this seems less likely given that sampling was conducted during the same oceanographic surveys, and the consistency of such differences with recent studies reporting a similar trend in maximum size and body condition of sardines in the area (Albo-Puigserver et al., 2021; Bachiller et al., 2020; Brosset et al., 2017).

*Outlier detection with gINLAnd, PCAdapt and RDA*

Significant genotype–environment associations were inferred with gINLAnd (Guillot, Vitalis, Rouzic, & Gautier, 2014), which implements a spatially explicit generalized linear mixed model to evaluate the correlation between allele frequencies and environmental variables using linear or logistic regression. To minimize the occurrence of false positives, loci with a log10 Bayes factor (logBF) of ≥ 3 were interpreted as having an outstanding statistical dependence with a certain environmental variable and therefore likely to belong to a genomic region under selection.

Environmental variables were selected considering evolutionary history, climate variability or change and fishing activity as the dominant processes affecting small pelagic fish. Given that the impacts of climate change are a rather complex issue that refers to more than one parameter (e.g. an increase in temperature), we tested whether there are any environmental variables related to topography, hydrodynamics, biochemical and biological components at different layers (i.e. surface, water column, bottom), different types of exploitation tools and food availability acting as selective agents in sardines of the North-Western Mediterranean. Variables were selected as to include all previously reported variables and their trends with significant impact on sardines such as temperature (Garrido et al., 2017; Gordó-Vilaseca, Pennino, Albo-Puigserver, Wolff, & Coll, 2021), salinity (Santos, Re, Dos Santos, & Peliz, 2006) and hydrodynamics (for a review see Caballero-Huertas et al., 2022), as well as variables related to the biology of the species whose impact has never been assessed before. The latter include topography related variables associated to sardine’s spawning and feeding behaviour, human pressure through fishing effort, nutrients as a proxy to food availability and quality but also to habitat conditions. Finally, regionalizations based on climatological and/or biological components of the Mediterranean ecosystems were also used as they delimit provinces within which physical conditions, chemical properties, and biological communities are reasonably homogeneous. These were conducted under an exploratory framework as to pinpoint significant factors shaping sardine’s genomic patterns, that have previously been neglected with subsequent implications on future directions of global change. Environmental variables employed, comprise direct observations (i.e. satellite data) as well as model derived products. The relevant raw and derived abiotic and biotic variables from five different sources (detailed information is provided at the Supporting Information) were incorporated in the analysis (Table 1 and Table S2), permitting tests of genotype–environment correlation and the detection of putative loci under selection (outlier loci). For the last dataset, i.e., dataset of northern vs southern sites, four extra environmental variables were used to better decipher the observed differences in individuals’ length and body mass related to the climate impact on sardine biomass and spawning and to fishing pressure by trawlers and purse-seiners (the main fishing gears targeting sardines in the Western Mediterranean Sea, details in Table S2). Raster values from all layers at the sampling sites were collected with the Point Sampling Tool plugin of QGIS v. 3.4.15 (QGIS.org, 2021, QGIS Geographic Information System, QGIS Association <http://www.qgis.org>).

The association patterns of outlier loci detected with gINLAnd across environmental variables were further examined with *superheat* R package (Barter & Yu, 2018). This package enabled the visualization of patterns of shared associations with the use of paired dendrograms and heatmaps.

Outlier detection was also conducted by a Principal Component Analysis (PCA) based approach on individual genotype data using the *pcadapt* R package (Luu, 2017). The most appropriate number of clusters i.e. K, was selected following Cattell’s rule (Jackson, 1993) while q-values were used to account for false discovery rate, and SNPs were considered as outliers at significance level of α ≤ 0.05 following a Bonferroni adjustment.

Redundancy analysis (RDA) implemented in the R package VEGAN was employed as a multivariate method to detect environmental variables relevant for sardines as well as outlier loci (Oksanen et al., 2012). The RDA was conducted in two datasets with different levels of population structure given the not well known performance of RDA in systems with increased levels of population structure or metapopulation dynamics (Forester, Lasky, Wagner, & Urban, 2018). The two datasets included “all samples” and MED while for the remaining two it was not possible due to the low number of sites that resulted in collinearity of all the environmental variables. Environmental covariates were selected as to minimize collinearity among them using Variance Inflation Factors (VIF < 10). Because RDA requires complete data frames, we imputed missing values by replacing them with the most common genotype across individuals. Significant constrained axes were identified using 999 permutations of the response data and a p-value threshold of 0.05. We identified candidate adaptive loci as SNPs loading ±3 SD from the mean loading of these significant RDA axes. We then identified the covariate most strongly correlated with each candidate SNP (i.e., highest correlation coefficient), to group candidates by potential driving environmental variables. Detailed information for this section is provided at the Supporting Information.

*Neutral and putatively adaptive genomic variation*

To minimize the detection of false positives, only SNPs that were selected by gINLAnd and *pcadapt*,were considered as outliers, given that the number of SNPs selected by all methods (gINLAnd, PCAdapt, RDA) was extremely low (five) as it is usually the case among the different approaches (e.g. Forester et al., 2018). Similarly, a putative neutral SNPs set included only SNPs that were not highlighted as outliers by any of the methods employed. This yielded two different sets of loci, i.e., putatively under selection and putatively neutral. The two sets of loci were analyzed with STRUCTURE 2.3.4 (Pritchard et al., 2000), using the same settings as above, in order to define whether different signals of population structure could be revealed.

*Outlier loci functional annotation*

The two groups of outlier loci were further analyzed as to identify which of these loci were located within known genic regions (i.e., from the start up to the end of the genes including introns), and for those found in intergenic regions, which is the closest gene. To perform this step, we downloaded the genome annotation files (gff) from the sardine genome portal (<https://bioinformatics.psb.ugent.be/gdb/Spil/>) and compared the outlier marker locations with the genes start and end positions using the function ‘closest’ from bedtools (Quinlan & Hall, 2010). Then, the protein sequences of the genes recovered were used in a blastp similarity search (e-value threshold 10-8) against *swissprot* database. To identify potential biological functions involved in the studied populations response to the environmental variables, a functional annotation analysis followed. Gene Ontology (GO) terms were retrieved per gene and were summarized using the tool WEGO v2.00 (<https://wego.genomics.cn/>, Ye et al., 2018).

All computational intensive analyses were conducted at IMBBC HPC facility (Zafeiropoulos et al., 2021).

**Results**

*ddRAD sequencing and data analysis*

We sequenced one billion paired reads from 398 sardine individuals with an average of 2.5 million reads/individual ranging from 713,308 up to 4,359,594. Mapping rate per individual ranged from 565,056 up to 3,638,348, resulting in a comparable performance across samples and a total of 755,128,044 properly mapped reads. From the mapped reads, a catalogue with 275,864 loci was built. After applying all filters to the SNP dataset and selecting one SNP per locus, 4,609 SNPs were retained and used for downstream analyses.

*Assessing the population genetic structure*

According to STRUCTUREs‘ clustering analysis, the optimal number of clusters that best fit the data was K = 2 (with all 10 replicates providing identical clustering solutions). Samples with membership coefficients (q-values) ≥ 0.9 were assigned to either cluster with high confidence while samples with intermediate values were considered as admixed i.e. having a mixed ancestry from the two clusters (Figure 2A). The first cluster, named “ATL” hereafter, included samples from only two sites, the GoC and GSA01, while the second cluster included samples from all sampling sites except GoC, thus named “MED” hereafter. Individuals of mixed ancestry were observed in all sampling sites with the highest numbers in sites GSA06a, GSA06b, GSA06c, GSA07a and GSA07b (Figure 2B). In the ATL cluster, 43 SNPs were monomorphic. The fixed alleles of each of the 43 SNPs in the ATL cluster were also present in the MED cluster in high frequency ranging from 0.7547 to 0.8966. No monomorphic SNPs were observed in the MED cluster. No further sub-structuring (within clusters) was detected.

The DAPC analysis indicated the occurrence of two clusters, grouping sardine individuals into two well differentiated and partially overlapping genetic clusters (Figure 3). Based on the cross-validation estimation, 100 PCs were retained while the discriminant function had an eigenvalue of 2441.312 and explained all the variance in the data.

The allocation of samples to two clusters was consistent between the two employed methods (i.e., STRUCTURE and DAPC). In the DAPC analysis all samples were fully allocated (posterior membership probability, pp. ≥ 0.9) to either cluster (DAPC\_ATL or DAPC\_MED) except one sample from GSA06a (sample id: 6a\_20, allocated to DAPC\_ATL with pp. 0.72). The allocation of samples to the two groups in DAPC analysis was identical to that of STRUCTURE with DAPC detecting less admixed individuals.

In order to detect adaptations of the two differentiated gene pools, as a strict measure of defining as pure clusters as possible, seascape genomic analyses were conducted based on STRUCTURE allocation to clusters i.e., on the ATL and MED clusters, without taking under consideration the admixed samples.

Similar levels of observed and expected heterozygosity estimates over all loci were found across sampling sites (Ho = 0.266–0.307 and He = 0.266–0.296). The FST values among sites varied between 0 and 0.0485 (Table S3) with the highest values observed in comparisons involving GoC site (0.0106-0.0485, all statistically significant) as well as GSA01 site comparisons (0.0105-0.0182, all statistically significant) with the rest of the sites. Furthermore, GoC site appeared to be closer, in terms of FST values, with GSA01 site (FST = 0.0106, statistically significant). The rest of the Mediterranean sampling sites (i.e., all Mediterranean sites except GSA01) had very low but significant FST values ranging from 0.0016 to 0.0043, while only among a few sites FST was zero (i.e., involving comparisons among GSA05, GSA06a, GSA06b, GSA06c, GSA07a, GSA07b, GSA09, GSA10, GSA19). Following FDR correction, all (56) except one pairwise comparisons remained significant (i.e. GSA07a vs GSA09, FDR = 1 / (1 + 56)= 0.017). Finally, the observed and expected heterozygosity estimates over all loci for the two clusters detected by STRUCTURE were Ho 0.279 and He 0.304 for the ATL cluster and Ho 0.282 and He 0.316 for the MED cluster. The FST value between the two clusters was 0.0703, statistically significant (p=0.0001) and the highest estimated.

*Seascape genomic analyses*

The total number of outlier loci that were highlighted either by gINLAnd or PCAdapt in any of the datasets summed up to 1478. Of these 196 were common between the two approaches (i.e., gINLAnd and PCAdapt), with most of them highlighted in the dataset containing all samples, along with any of the other three datasets (n=186) and only few of them exclusively in the ATL, MED (n=3 in each dataset) and the “northern vs southern” dataset (n=4). Across datasets with all levels of population structure (i.e., “all samples” dataset that displayed population structure and ATL and MED datasets at the lower level of population structure), there were no outlier loci in common when gINLAnd was used. On the contrary, when PCAdapt was used, a core of 33 loci was shared among the three datasets.

In the “all samples” dataset, gINLAnd detected 262 outlier loci with logBF > 3, indicative of strong evidence of selection (Kass & Raftery, 1995). These outlier loci were associated to 39 environmental variables with none of them related to topography or fishing effort. Of those outliers, 53 had logBF values above 10 in at least two and up to 30 environmental variables. Three main groups of loci were revealed by clustering analysis, two of which exhibited shared patterns of association across multiple environmental variables, while the last one that contained the largest number of loci exhibited shared patterns across fewer variables (Figure 4) and lacked variables related to salinity. The highest number of loci were invoked by three variables i.e., surcurrent, nitrate and sst\_m\_sl (see Table S4). Environmental variables were also clustered in three main groups (Figure 4). For the same dataset, PCAdapt detected 462 outlier loci, with 186 of those being detected by both approaches (common outliers).

Analyzing the MED dataset with gINLAnd yielded 68 outliers with logBF > 3 (and up to 6.5) in 18 environmental variables. The highlighted variables (as previously) included all categories except the ones related to topography and fishing effort. Three main groups of loci were revealed with one including a single locus while the other two were almost evenly populated by similar number of loci (Figure 5). Shared patterns of association across variables involved only few loci while four environmental variables invoked the majority of the loci in four distinct clusters respectively (Figure 5). Those were bo\_pH, surcurrent, sst\_max\_sl and tot\_impact. For the same dataset, PCAdapt detected 209 loci of which 8 where common with gINLAnd’s loci.

In the ATL dataset, gINLAnd detected 37 outlier loci with logBF > 3 (and up to 5.66) in 24 environmental variables. This is the only dataset where variables related to topography co-variate with SNP data. Hierarchical clustering analysis revealed two groups of loci with one exhibiting shared patterns of association across the majority of environmental variables including the ones related to topography (Figure 6). The variable with the highest number of co-variate SNPs was biogeo03, while with the exception of parmean that correlated with one SNP, all the other variables correlated with a high number of SNPs ranging from nine to 23. Environmental variables were clustered in two main groups (Figure 6). PCAdapt analysis detected 957 loci for the same dataset, of which 5 were common with gINLAnd’s list of outlier loci.

The “northern vs southern” dataset yielded 45 outlier loci with logBF > 3 according to gINLAnd analysis (and up to 6.34), in seven environmental variables. One variable i.e., sst\_max\_sl displayed correlation with most of the highlighted SNPs (23 SNPs). The next most important variables in terms of the number of highlighted loci were i) cum\_impact† with 11 loci, and ii) sst\_tr19 and tot\_impact, each with seven loci (Figure 7). The above-mentioned variables were the ones that grouped the loci in two main clusters, two of which were further subdivided. One SNP was found to co-variate with clim\_impact†. PCAdapt for the same dataset detected 218 outliers of which 3 were common with gINLAnds’ outliers.

The RDA model for the “all samples” dataset was significant (p=0.001), performed with 9 environmental variables (sst\_tr19, Currents, surcurrent, nitrate, sst\_max\_sl, sst\_tr12, FrontiersS, fishing\_ef, biogeo03) all being statistically significant while explaining only 3.35% of the observed genomic variance (adjusted R2: 0.011). Nitrate (20.6) and sst\_tr19 (20.6) explained most of the variation while the remaining parameters explained similar percentages of variation (13.2-18). There were four significant RDA axes, which returned 125 unique candidate loci that loaded ±3 SD from the mean loading on each axis: 92 SNPs detected on RDA axis 1, 14 on RDA axis 2, 3 on RDA axis 3 and 16 on RDA axis 4. The detected SNPs displayed shared patterns of association across multiple environmental variables. Individual genotypes from the ATL and GSA01 sites are positively related to nitrate, plan curvature and Photosynthetically Active Radiation (parmean) respectively. Furthermore, a distinction of northern (GSA07a, GSA07b, GSA06a) versus southern sites (GDS06c) is evident with a positive relationship with the number of days with SST < 12oC (sst\_tr12) and with the maximum annual SST (sst\_max\_sl) respectively (Figure 8A, B).

For the MED dataset, RDA was significant (p=0.002) performed using 11 environmental variables (sst\_tr19, Currents, surcurrent, nitrate, sst\_max\_sl, sst\_tr12, salinity, BioReg, FrontiersS, fishing\_ef, biogeo03) with four of them being statistically significant (sst\_tr19, nitrate, sst\_tr12 and salinity, explaining similar levels of variation 31.5-32.4) while explaining only 6.64% of the observed genomic variance (adjusted R2: 0.0024). Individual genotypes from GSA07b were positively related to the number of days with SST > 19oC threshold (sst\_tr19), from GSA11 positively related with sst\_tr12, while genotypes from GSA10 were negatively related to nitrate. Finally on the third RDA axis a positive relationship of GSA05 and GSA19 haplotypes with salinity was observed (Figure 8C, D). The single significant RDA axis in the ordination of the MED dataset, returned 23 unique candidate loci that loaded ±3 SD from the mean loading on axis. The detected SNPs displayed shared patterns of association across multiple environmental variables. The outlier loci detected by the RDA in the two datasets revealed overlapping as well as unique loci with 14 in common among the two datasets and 111 and 9 unique in “all samples” and MED datasets respectively. Given that RDAs were conducted on a different set of environmental variables than the one used in gINLAnd, a direct comparison of the relevant for the European sardines’ environmental variables as well as the outlier loci detected by the two methods is not possible. Nevertheless, although examining two out of the four datasets when RDA was employed, five outlier loci were in common through all approaches. Furthermore, the environmental variables highlighted as significant by the two approaches were highly consistent.

Overall, from all examined datasets and all employed methods (i.e., gINLAnd, RDA and PCAdapt), 1607 SNPs were highlighted as potential outliers . Fifty out of 54 environmental variables used in this study correlated with at least one SNP and up to 156 (see Table S4). Four variables that did not correlate with any of the SNPs in any of the datasets were: bathymetry, biogeo01, biogeo13, and Berline’s regionalization. The sst\_tr19 variable was the only one correlated with SNPs in all four studied datasets.

Regarding gINLAnd results, eight variables correlated with SNPs in all datasets except the one comparing northern vs southern sites. Those variables are related to currents (Currents, surcurrent), and nutrients (nitrate, bphosphate, bo2utilize, bo2dissolv, parmean, bo\_pH). The variables shared among the dataset that contained all samples and MED dataset were mostly related to temperature (tot\_impact, sst\_tr26, sst\_max\_sl, sst\_tr12, sst\_min\_sl), as well as salinity (salinity), Reygondeaus’ epipelagic biogeochemical regions and FrontiersS. Water column temperature (btemp) was a shared environmental variable among MED and ATL dataset though none of the correlated SNPs were in common.

Some variables were exclusively highlighted in particular datasets when gINLAnd was used. For the dataset including all samples those were mostly related to temperature (sst\_m\_slwi, sst\_m\_sl, biogeo17, biogeo16, biogeo15, biogeo14), biology (PLD\_60, Mesopelag and Bathypelag, EcoReg), climatology (Raw\_CluCli, ClusterCli), salinity (biogeo10, biogeo08), primary productivity (primpod) and nutrients (calcite). When RDA was conducted with all samples, two environmental variables, not indicated for the same dataset by gINLAnd, were highlighted i.e. fishing effort (fishing\_ef) and plan curvature (biogeo03). For the MED dataset RDA environmental variables overlapped with those indicated by gINLAnd. ATL dataset exclusively contained topography related variables (biogeo02, biogeo03, biogeo06, biogeo05, biogeo07, biogeo04) as well as two variables related to salinity (biogeo12, biogeo11). Finally, the dataset comparing northern vs southern sites exclusively correlated with SSS of the freshest month (biogeo09).

Population structure analyses on only gINLAnd and PCAdapt common outliers (n=196) estimated K = 2 clusters and allocated the samples to the MED and ATL clusters as with the complete loci dataset. On the contrary, when using only the neutral set of loci i.e. 3,002 loci (deducting 1,607 outliers from the total 4,609 loci), no signal of population structure was detected.

*Functional Annotation*

Functional annotation analyses of the 196 outlier loci highlighted in any of the datasets by gINLAnd and PCAdapt approaches revealed that 156 are located within scaffolds that contain predicted genes. Few of them (i.e. 12) were found within genic regions while for the rest the closest gene was identified (top blast hit per gene is given in Table S5). GO terms were retrieved for 140 of these genes (Figure S1).

The most frequent terms on the ‘Biological Process’ category were ‘cellular process’, ’metabolic process’, ‘biological regulation’, ‘regulation of biological process’ and ’signaling’. Out of the five outlier loci in common through all approaches, four were found in scaffolds with genes in their vicinity. Those were: Heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1), G protein-activated inward rectifier potassium channel 2 (GIRK-2), Myosin regulatory light chain 2, ventricular/cardiac muscle isoform (MLC-2) and Transmembrane protein 178B.

**Discussion**

*Understanding sardines’ population structure*

The present study indicates the existence of two clearly differentiated gene pools of sardines that meet in the Western Mediterranean i.e., ATL and MED clusters. The two clusters are geographically well defined, in agreement to the geographical distribution of the two sardine subspecies recognized by morphological and molecular characters (Atarhouch, Rami, Naciri, & Dakkak, 2007; Fonseca et al., submitted; Parrish, Serra, & Grant, 1989). Admixed individuals are found throughout the Mediterranean sampling sites, with AOF having a transition role between the two regions, in agreement to previous studies (e.g. Atarhouch et al., 2007).

Gene flow likely occurs from the Atlantic to the Mediterranean since in GoC only ATL and admixed individuals (not MED) were found, while ATL individuals were also found within the westernmost MED site, i.e., GSA01 in the Alboran Sea. This was also supported by the FST index, with the highest values observed among the Atlantic and the Mediterranean sites except GSA01, with the latter being closer to GoC and displaying intermediate values when the Mediterranean sites were considered. Additionally, the number of admixed individuals decreased towards the edges of the east-west axis i.e., towards central Mediterranean and the Atlantic respectively (Figure 2B).

Our results suggest that AOF acts as the actual barrier separating the Atlantic and the Mediterranean sardines allowing for some interaction although being spatially separated as reported in previous studies (Santos et al., 2018 and references therein), also coinciding with those of other fish species (e.g. Borsa et al., 1997, Bargelloni et al., 2003*,* Abaunza et al., 2008)(Naciri, Lemaire, Borsa, & Bonhomme, 1999).

The two clusters could be detected even with the use of only 196 outlier SNPs linking candidate SNPs to the region’s environmental heterogeneity as seen in other sardine species (Teske et al., 2021). In contrast, when only putatively neutral SNPs were used (i.e. 3,002 SNPs), detection of any pattern of population structure failed, suggesting that natural selection and local adaptation play a key role in driving genetic change (Metivier, Kim, & Addison, 2017) among the Atlantic and the Mediterranean Sea. This was further corroborated by the fact that different outlier SNPs were highlighted in the datasets where population structure was present versus the unstructured datasets (i.e., all versus ATL, MED clusters) and a slight overlap across datasets was detected by PCAdapt (33 SNPs) and RDA (14 SNPs).

*Seascape genomics provides insights on the main environmental conditions affecting sardine’s evolution*

Water temperature, and derived environmental products informing on changing temperature conditions were highlighted as extremely important variables driving natural selection and local adaptation of sardines in the study area. The trend in the number of days with SST above 19oC was the sole environmental variable highlighted in all datasets. This is indicative of the extreme conditions experienced by the species within the past few years (as in European sea bass, Anastasiadi, Shao, Chen, & Piferrer, 2021), impacting its reproduction (Palomera et al., 2007) and spawning success potential (Gordó-Vilaseca et al., 2021). This variable is critical for the reproduction of sardines in the Western Mediterranean Sea, where spawning starts when water temperature falls below 19ºC and salinity is over a threshold (Palomera et al., 2007). Furthermore, temperature could be a proxy of other variables, namely of coastal upwelling strength (e.g. Zelle, Appeldoorn, Burges, & van Oldenborgh, 2004), and therefore larval dispersal (Gordó-Vilaseca et al., 2021). Analyses of the relationship between sardine recruitment and several abiotic factors suggested that SST was a top-ranking candidate in determining recruitment success (Garrido et al., 2017; Gordó-Vilaseca et al., 2021; Solari et al., 2010). The impact of SST on sardine’s abundance has been demonstrated in California as well, debunking previous hypotheses of competitive exclusion with anchovies as the driving factor in population size for both species (Sugihara et al., 2012).

Apart from SST, nutrient availability, sea currents and temperature velocity of change were important variables for the differentiation and local adaptation of Atlantic and Mediterranean clusters. This was evident given that those variables were the most important in all datasets except the one comparing northern vs southern sampling sites. These variables are expected to majorly affect early life stages, which are particularly vulnerable to environmental variability, and thus the strength of recruitment. In sardines, such environmental forcing acting upon the egg and larval stages has already been demonstrated (Garrido et al., 2017), where recruitment strength is affected by differences of the prevailing environmental conditions in the studied areas. In the main recruitment hotspots of the species in Atlantic-Iberian waters, high recruitment years have been associated with high chlorophyll-*a* concentration (Chla, as an indicator of food availability) and low SST (reflecting the optimal range for larval development) during periods of sardine larvae development (Garrido et al., 2017; Gordó-Vilaseca et al., 2021). Temperature and food availability also affect the intensity of the reproduction and quality of the eggs produced (Garrido & van der Lingen, 2014).

employed

*The genomic profile of the outlier loci*

Functional annotation analyses of the outliers-associated geneset revealed that the genes are linked to key biological processes associated to environmental pressures (Figure S1). It is worth noting that the retrieved top GO categories (e.g. cellular process, metabolism, biological regulation and response to stimulus) have been reported in other studies that measure transcriptomic responses in experimentally induced heat stress. For example, the same processes were associated with heat stress in turbot (Huang et al., 2020) and in other fish species (e.g. Huang, Li, Liu, Kang, & Wang, 2018; P. Li et al., 2021). Those findings reflect a similar functional profile of the mechanisms selected during sardines’ adaptation to sea warming.

Reproduction and sex determination are main processes greatly impacted by temperature in fish (Vandeputte et al., 2020). Amongst others, our gene dataset included the proteins Gametogenetin-binding protein 2, AT-rich interactive domain-containing protein 4B, Myosin-7 and Thioredoxin reductase 3 all involved in the process of spermatogenesis and Cilia- and flagella-associated protein 52 involved in sperm mobility, reflecting a possible impact of the significant environmental parameters of our analysis to sardines’ reproduction. Another key process that seems to be highly impacted is the muscle formation which has been described to be affected by temperature (Balbuena-Pecino et al., 2019; Georgakopoulou, Katharios, Divanach, & Koumoundouros, 2010), including one of the genes found as outlier in all analyses conducted herein, i.e. Myosin regulatory light chain 2 as well as others (e.g. paxilin involved in response to muscle stretch, triadin involved in regulation of cardiac muscle cell etc.)

Finally, our geneset included genes involved in the response to nitrate availability which were Carbamoyl-phosphate synthetase I involved in nitrogen compound metabolic process and Atrial natriuretic peptide receptor 3 involved in positive regulation of nitric-oxide synthase activity. Interestingly, the latter is also involved in regulation of cold-induced thermogenesis suggesting a link among the two processes.

Overall, the genomic profile of the dataset of the 196 outlier loci suggests that selective pressures are acting on multiple fronts with emphasis on “binding”, impacting a plethora of biological processes that involve interaction with the environment.

*Sardines from the Mediterranean and the Atlantic respond differently to environmental pressures*

Τwo datasets, i.e., all samples and ATL cluster datasets displayed shared patterns of association of individual SNPs with the majority of environmental variables. Loci showing multiple environmental associations were found in the same loci cluster using hierarchical clustering, with these clusters representing proxies for functional modules (Kess et al., 2020; Lotterhos, Yeaman, Degner, Aitken, & Hodgins, 2017). Such modular genetic architectures are favored when adapting to climate on a landscape (Kliebenstein, 2011) where complex spatial and temporal environments with both abiotic and biotic challenges occur at different spatial scales (Le Nagard, Chao, & Tenaillon, 2011) as well as when multiple traits are under a combination of directional (among populations) and stabilizing (within populations) selection (Griswold, 2006; Le Corre & Kremer, 2003; Wagner & Altenberg, 1996). Such modularity can be beneficial not only for facilitating evolvability in response to a variable environment (Edwards & Weinig, 2011), but also by allowing the system to function as a set of independent groups of genes, allowing a species to better sample its potential phenotypic and genotypic diversity (Leroi, 2000). Sardines undoubtedly face spatially and temporarily complex environments within and between the two areas i.e., Mediterranean Sea and Atlantic Ocean. Whether sardine possesses phenotypic and genomic characteristics that could be indicative of directional and stabilizing selection among and within populations respectively, demands more data and testing. Current data and analyses are in favor of a hypothesis of among populations directional selection in sardines i.e., having as indication the 43 SNPs, which appear monomorphic in the Atlantic cluster, and with high frequency of the fixed allele in the Mediterranean cluster. However, given that our sampling scheme includes only one site from the Atlantic, no firm conclusion can be drawn.

For the Mediterranean cluster, high temperature related variables are of primary importance, probably reflecting the pressures of climate change and/or variability acting within the Western Mediterranean that directly affect sardines’ body condition, reproduction and larvae survival (Albo-Puigserver et al., 2021; Garrido et al., 2016). The same holds for the northern versus southern sites dataset in the Western Mediterranean (Albo-Puigserver et al., 2021; Coll & Bellido, 2019; Gordó-Vilaseca et al., 2021; Pennino et al., 2020), while the correlation of a locus to the SSS of the freshest month might be indicative of the influence of the lower salinity levels to the species reproduction especially in areas of river runoffs as GSA07 (Palomera et al., 2007).

The Atlantic cluster seems to be mostly affected by the seascape and especially plan curvature, as indicated by both gINLAnd and RDA analyses, probably reflecting adaptations to its habitat which is different from that within the Mediterranean e.g., wandering at such an extensive body of water as the Eastern Atlantic which displays fewer physical barriers associated with coastal relief (Santos et al., 2018), and with its bottom topography impacting its oceanographic elements such as currents, gyres, upwelling, and oceanic fronts (Castelao & Luo, 2018). A significant relationship between the extent of the potential sardine habitat, larval food availability (defined by depth, SST and Chla) and sardine recruitment has already been found in the Atlantic coasts off Morocco (Machu et al., 2009). It was suggested that years of low recruitment occurred when the optimal spawning period in terms of temperature and salinity did not coincide with periods of high food availability. However, further investigations over wider geographical areas in the Atlantic Ocean are necessary to draw firm conclusions on the main physical, biological or environmental drivers of the genetic variability in the Atlantic cluster.

Sardines widely distribute in highly productive areas of enhanced food availability. In the Atlantic Ocean, these major feeding areas are largely associated to upwelling offshore systems (Checkley, Asch, & Rykaczewski, 2017). In general, the offshore occurrence of sardines in the Atlantic Ocean has been held responsible for their larger size with respect to other small pelagic fishes (i.e., anchovies), but also for their longer migrations and lower levels of population structure (Checkley et al., 2017). In contrast, marine productivity and highly productive marine areas in the Mediterranean Sea are largely driven by river discharges (e.g. Feuilloley et al., 2020). Sardines stay closer to the coastline and feed on areas directly influenced by river runoffs (Costalago, Garrido, & Palomera, 2015). Based on these lines of evidence, we hypothesize that the larger size of the individuals of the ATL cluster might be due to its adaptation to wander in offshore areas. An indication of the distinct selective pressures acting on the different hierarchical clustering levels is provided by the fact that there was almost zero overlap in the lists of loci that were associated to environmental variables across the different datasets (i.e., zero overlap in gINLAnd, 33 loci in PCAdapt and 14 loci in RDA across the datasets).

*Is fishing effort driving sardine’s adaptation?*

Co-variation of SNPs with variables reflecting fishing effort was observed in the cumulative impacts variable i.e. cum\_impact† (Ramírez et al., 2018, 2021), where climate impacts on sardine biomass and spawning is combined with fishing effort footprint, as well as in fishing effort (fishing\_ef) when RDA was employed. Such findings agree with multiple lines of evidence where fishing may alter fluctuations in sardine stocks, but may not be the primary source of change (Checkley et al., 2017; Coll et al., 2019; Saraux et al., 2019) and the fact that environmental parameters (ICES, 2015) and in particular temperature and food are essential factors for sardines in general. Unarguably, fishing effort is affecting sardine stocks and might become detrimental to their sustainability (Ramírez et al., 2018, 2021). It has been the main reason why elderly age classes have been wiped out from particular areas (Albo-Puigserver et al., 2021; Coll & Bellido, 2019), having the maximum age of sardines in the Mediterranean not exceeding five years. Spatially-explicit information on fishing pressure is only available for a relatively short time period, thus preventing from capturing long-term changes in fishing activity (FAO, 2020) but also the selective pressure acting on the stock. In part, this may explain the lack of strong selective pressure of fisheries (when considered alone) on sardine populations. However, from a genomics perspective, whether sardines’ primary control in the Mediterranean is bottom-up (i.e., climate variability) rather than top-down (i.e., fishing) remains to be fully assessed, with our results pointing to a probable combination of factors affecting sardine’s population dynamics.

*North/South gradient is solely affected by the environment or does it have a heritable basis?*

According to our results, the genetic patterns of sardines sampled in the area do not reflect the observed north to south gradient in size and body condition. Furthermore, key variables of the northern vs southern dataset are related to high temperature velocity of change (sst\_max\_sl) and sea surface salinity (SSS) of the freshest month (biogeo09) with its lowest value observed in the northernmost site GSA07a, indicating that reproduction and survival are important forces acting on sardines in the area. Variables reflecting directly or indirectly food availability (e.g., nitrate) failed to justify the poorer physical condition of sardines in northern versus southern sites through genomic adaptations. Nitrate, a significant index of ecosystem production, was invariable in the relevant sites. Primary production is varying in a counter anticipated way to that expected when food availability causes differences in fish condition, i.e., it is lower in the wealthier GSA06c area and higher in the poorer GSA07a, GSA07b and GSA06a areas. The only variables that reflected northern poorness to southern richness were PAR mean (Photosynthetically Active Radiation) and bottom nitrate that have been found to vary on a seasonal basis (Zhao & Costello, 2019). Nevertheless, not strictly looking into the seasonal variation of the environmental variables with time (we only looked on average values of two periods i.e. May-October and November-April approximating the species reproductive season) inevitably disregards any seasonal/temporal changes occurring as well as their spatial extent with the role of seasonality remaining to be further assessed in future studies.

According to Van Beveren et al. (2014), the recent decrease in abundance of small pelagic fishes in the Gulf of Lions was related to slower growth and the disappearance of older individuals rather than to a decrease in the recruitment strength or overfishing. As growth and condition decreased, the authors proposed that the current decline in sardine biomass could be due to qualitative and/or quantitative modifications in the planktonic production (i.e. a bottom-up control) or mass mortalities of adults due to an epidemic disease (Van Beveren et al., 2014). This is further supported by the findings of recent studies where the decrease in body condition was not related to selective processes (Saraux et al., 2019), with spatial-seasonal factors at local scales being pivotal (Lloret-Lloret et al., 2022). It has also been observed that the spatial differences of the feeding ecology of adult sardines might affect their lipid composition, and also the fatty acid reserves transferred to their progeny (Garrido et al., 2007, 2008), with direct implications on their condition. Such phenotypic differences are not always in line with observed genetic divergence, as in the case of sardines off Moroccan coasts (Atarhouch et al., 2006; 2007) that were also attributed to responses to local environmental variables. Unfortunately, the quality of the food was not directly included in our analysis due to the lack of variables to consider and no conclusions on the matter can be drawn.

Overall, observed responses to local environments comply with the environmental conditions prevailing in those areas and their deterioration (Ramírez et al., 2021). According to the marine regionalization of Zhao & Costello (2019), the Mediterranean Sea is occupied by five “ecosystems” (defined as “an enduring, spatially bounded environment where biological and energy interactions are greater within than with other ecosystems”), three of which are present in the study area (Ecosystems 4, 7 and 8). during winter (December-February) and high PAR during summer (June-August) at the southern sites The above, coupled with rapid changes in temperature (Ramírez et al., 2018, 2021), could be held responsible for such differences in physical conditions between the north and the south that might be reinforced by periods of lower connectivity with areas submitted to milder pressures of extreme environmental variability and where healthier fishes are found such as GSA06c sampling site (as in European hake, Hidalgo et al., 2019). These intensively fluctuating conditions constrain species either to tolerate such variation, adjust their activity, growth and reproduction in response to seasons or evolve into distinct biogeographic assemblages within tropical ecosystems (as Ecosystem 7), where ecosystem boundaries are more spatially stable (Zhao & Costello, 2019).

*Conclusions*

Insights into the main environmental parameters that drive local adaptation and adaptive responses to changes were gained under a seascape genomics approach through the separation of signals of neutral and putative adaptive variation. Undoubtedly, the seascape is affecting and driving the evolution of sardine’s key biological processes in Western Mediterranean Sea. The results obtained are not only relevant for the biology of the species, but also for understanding of natural as well as human-induced evolutionary processes and for planning an adapted fisheries management in this area by predicting future trajectories of biodiversity and setting conservation priorities. Despite admixture between the two groups (i.e. MED and ATL), their discrete management will maintain locally adapted groups, enhancing their sustainability. Our findings provide evidence for a dynamic equilibrium where population structure is maintained by physical and/or biological (including predation and behavior) factors under the opposing influences of migration and selection. Drastic changes may occur in such dynamic systems requiring a continuous monitoring under a seascape genomic approach that could benefit by including a temporal as well as a more detailed spatial dimension. Finally, our results may contribute to further studies aimed at providing deeper insights into the mechanistic processes underlying population structuring of sardines; a key step for understanding and predicting future changes and responses by this highly exploited species in the face of climate change.

**Acknowledgements**

The authors acknowledge MEDITS and MEDIAS oceanographic campaigns and the biological sampling program of the Data Collection Framework for collecting samples. D. A. Jadaud is acknowledged for her help collecting samples from MEDITS GSA07 survey (<http://www.medias-project.eu/medias/website/>, <https://doi.org/10.18142/7>). This study was carried out within the European Research Contract SPELMED (EASME/EMFF/2016/032) funded by the Executive Agency for Small- and Medium-sized Enterprises (EASME) or of the European Commission (DGMARE). The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Commission and EASME. Neither EASME, nor the European Commission can guarantee the accuracy of the data included in this publication. Neither EASME, nor the European Commission or any person acting on the EASME’s or on the European Commission’s behalf may be held responsible for the use which may be made of the information contained therein. MC would like to acknowledge partial funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 869300 (FutureMARES). MC and FR acknowledge the institutional support of the ‘Severo Ochoa Centre of Excellence’ accreditation (CEX2019-000928-S). FR was supported by Ramón y Cajal programme (Spanish Ministerio de Ciencia e Innovación, RYC2020-030078-I). This research was supported in part through computational resources provided by IMBBC (Institute of Marine Biology, Biotechnology and Aquaculture) of the HCMR (Hellenic Centre for Marine Research). Funding for establishing the IMBBC HPC has been received by the MARBIGEN (EU Regpot) project, LifeWatchGreece RI and the CMBR (Centre for the study and sustainable exploitation of Marine Biological Resources) RI, implemented under the Action “Reinforcement of the Research and Innovation Infrastructure”, funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund).

**References**

Abaunza, P., Murta, A. G., Campbell, N., Cimmaruta, R., Comesaña, A. S., Dahle, G., … Zimmermann, C. (2008). Stock identity of horse mackerel (Trachurus trachurus) in the Northeast Atlantic and Mediterranean Sea: Integrating the results from different stock identification approaches. *Fisheries Research*. doi: 10.1016/j.fishres.2007.09.022

Albo-Puigserver, M., Pennino, M. G., Bellido, J. M., Colmenero, A. I., Giráldez, A., Hidalgo, M., … Coll, M. (2021). Changes in Life History Traits of Small Pelagic Fish in the Western Mediterranean Sea. *Frontiers in Marine Science*, *8*, 1197. doi: 10.3389/FMARS.2021.570354/BIBTEX

Anastasiadi, D., Shao, C., Chen, S., & Piferrer, F. (2021). Footprints of global change in marine life: Inferring past environment based on DNA methylation and gene expression marks. *Molecular Ecology*, *30*(3), 747–760. doi: 10.1111/mec.15764

Antoniou, A., & Magoulas, A. (2013). Application of Mitochondrial DNA in Stock Identification. In *Stock Identification Methods: Applications in Fishery Science: Second Edition*. doi: 10.1016/B978-0-12-397003-9.00013-8

Archer, F. I., Adams, P. E., & Schneiders, B. B. (2017). stratag: An r package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources*, *17*(1), 5–11. doi: 10.1111/1755-0998.12559

Atarhouch, T., Rami, M., Naciri, M., & Dakkak, A. (2007). Genetic population structure of sardine (Sardina pilchardus) off Morocco detected with intron polymorphism (EPIC-PCR). *Marine Biology*, *150*(3), 521–528. doi: 10.1007/s00227-006-0371-8

Atarhouch, Touriya, Rüber, L., Gonzalez, E. G., Albert, E. M., Rami, M., Dakkak, A., & Zardoya, R. (2006). Signature of an early genetic bottleneck in a population of Moroccan sardines (Sardina pilchardus). *Molecular Phylogenetics and Evolution*, *39*(2), 373–383. doi: 10.1016/j.ympev.2005.08.003

Bacha, M., Jemaa, S., Hamitouche, A., Rabhi, K., & Amara, R. (2014). Ocean : evidence from otolith shape analysis. *Ices Journal of Marine Science*, *71*, 2429–2435.

Bachiller, E., Albo-Puigserver, M., Giménez, J., Pennino, M. G., Marí-Mena, N., Esteban, A., … Coll, M. (2020). A trophic latitudinal gradient revealed in anchovy and sardine from the Western Mediterranean Sea using a multi-proxy approach. *Scientific Reports*, *10*(1), 1–18. doi: 10.1038/s41598-020-74602-y

Baibai, T., Oukhattar, L., Quinteiro, J. V., Mesfioui, A., Rey-Mendez, M., & soukri, A. (2012, March 24). 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. *Reviews in Fish Biology and Fisheries*, Vol. 22, pp. 63–80. Springer. doi: 10.1007/s11160-011-9223-9

Bal, T. M. P., Llanos‐garrido, A., Chaturvedi, A., Verdonck, I., Hellemans, B., & Raeymaekers, J. A. M. (2021). Adaptive divergence under gene flow along an environmental gradient in two coexisting stickleback species. *Genes*, *12*(3). doi: 10.3390/genes12030435

Balbuena-Pecino, S., Riera-Heredia, N., Vélez, E. J., Gutiérrez, J., Navarro, I., Riera-Codina, M., & Capilla, E. (2019). Temperature Affects Musculoskeletal Development and Muscle Lipid Metabolism of Gilthead Sea Bream (Sparus aurata). *Frontiers in Endocrinology*, *10*(MAR), 173. doi: 10.3389/fendo.2019.00173

Baltazar-Soares, M., Lima, A. R. de A., & Silva, G. (2021). Targeted sequencing of mitochondrial genes reveals signatures of molecular adaptation in a nearly panmictic small pelagic fish species. *Genes*, *12*(1), 1–17. doi: 10.3390/genes12010091

Bargelloni, L., Alarcon, J. A., Alvarez, M. C., Penzo, E., Magoulas, A., Reis, C., & Patarnello, T. (2003). Discord in the family Sparidae (Teleostei): divergent phylogeographical patterns across the Atlantic-Mediterranean divide. *Journal of Evolutionary Biology*, *16*, 1149–1158.

Barter, R. L., & Yu, B. (2018). Superheat: An R Package for Creating Beautiful and Extendable Heatmaps for Visualizing Complex Data. *Journal of Computational and Graphical Statistics*, *27*(4), 910–922. doi: 10.1080/10618600.2018.1473780

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Source: Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Borsa, P., Naciri, M., Bahri, L., Chikhi, L., Garcia de Leon, F. . J., Kotoulas, G., & Bonhomme, F. (1997). Infraspecific zoogeography of the Mediterranean: population genetic analysis on sixteen Atlanto-Mediterranean species (fishes and invertebrates). *Zoogeographie Infra-Specifique de La Mer Mediterranee: Analyse Des Donnees Genetiques Populationnelles Sur Seize Especes Atlanto-Mediterraneennes (Poissons et Invertebres)Vie et Milieu*, *47*(4), 295–305.

Bradbury, I. R., Gardiner, K., Snelgrove, P. V., Campana, S. E., Bentzen, P., & Guan, L. (2006). Larval transport, vertical distribution, and localized recruitment in anadromous rainbow smelt (Osmerus mordax). *Canadian Journal of Fisheries and Aquatic Sciences*, *63*(12), 2822–2836. doi: 10.1139/f06-164

Brosset, P., Fromentin, J. M., Van Beveren, E., Lloret, J., Marques, V., Basilone, G., … Saraux, C. (2017). Spatio-temporal patterns and environmental controls of small pelagic fish body condition from contrasted Mediterranean areas. *Progress in Oceanography*, *151*, 149–162. doi: 10.1016/j.pocean.2016.12.002

Caballero-Huertas, M., Frigola-Tepe, X., Coll, M., Muñoz, M., & Viñas, J. (2022). The current knowledge status of the genetic population structure of the European sardine (Sardina pilchardus): uncertainties to be solved for an appropriate fishery management. *Reviews in Fish Biology and Fisheries*, (0123456789). doi: 10.1007/s11160-022-09704-z

Castelao, R. M., & Luo, H. (2018). Upwelling jet separation in the California Current System. *Scientific Reports*, *8*(1), 1–8. doi: 10.1038/s41598-018-34401-y

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

Checkley, D. M., Asch, R. G., & Rykaczewski, R. R. (2017). Climate, Anchovy, and Sardine. *Annual Review of Marine Science*, *9*(1), 469–493. doi: 10.1146/annurev-marine-122414-033819

Chhatre, E. V. (2018). *Distruct v2.3, A modified cluster membership plotting script.* Retrieved from http://distruct2.popgen.org

Chlaida, M., Laurent, V., Kifani, S., Benazzou, T., Jaziri, H., & Planes, S. (2009). Evidence of a genetic cline for Sardina pilchardus along the Northwest African coast. *ICES Journal of Marine Science*, *66*(2), 264–271. doi: 10.1093/icesjms/fsn206

Coll, M., Albo-Puigserver, M., Navarro, J., Palomera, I., & Dambacher, J. M. (2019). Who is to blame? Plausible pressures on small pelagic fish population changes in the northwestern Mediterranean Sea. *Marine Ecology Progress Series*, *617*–*618*, 277–294. doi: 10.3354/meps12591

Coll, M., & Bellido, J. (2019). *Evaluation of the population status and specific management alternatives for the small pelagic fish stocks in the Northwestern Mediterranean Sea (SPELMED) - Final Report SC NR 02 - TENDER EASME/EMFF/2016/32 – SPELMED*. Retrieved from https://op.europa.eu/en/publication-detail/-/publication/f1bd2c63-084e-11eb-a511-01aa75ed71a1

Coll, M., Piroddi, C., Albouy, C., Ben Rais Lasram, F., Cheung, W. W. L., Christensen, V., … Pauly, D. (2012). The Mediterranean Sea under siege: Spatial overlap between marine biodiversity, cumulative threats and marine reserves. *Global Ecology and Biogeography*, *21*(4), 465–480. doi: 10.1111/j.1466-8238.2011.00697.x

Coll, M., Steenbeek, J., Ben Rais Lasram, F., Mouillot, D., & Cury, P. (2015). ‘Low-hanging fruit’ for conservation of marine vertebrate species at risk in the Mediterranean Sea. *Global Ecology and Biogeography*, *24*(2), 226–239. doi: 10.1111/geb.12250

Costalago, D., Garrido, S., & Palomera, I. (2015). Comparison of the feeding apparatus and diet of European sardines Sardina pilchardus of Atlantic and Mediterranean waters: Ecological implications. *Journal of Fish Biology*, *86*(4), 1348–1362. doi: 10.1111/jfb.12645

Diopere, E., Vandamme, S. G., Hablützel, P. I., Cariani, A., Van Houdt, J., Rijnsdorp, A., … Maes, G. E. (2018). Seascape genetics of a flatfish reveals local selection under high levels of gene flow. *ICES Journal of Marine Science*, *75*(2), 675–689. doi: 10.1093/icesjms/fsx160

Edwards, C. E., & Weinig, C. (2011). The quantitative-genetic and QTL architecture of trait integration and modularity in Brassica rapa across simulated seasonal settings. *Heredity*, *106*(4), 661–677. doi: 10.1038/hdy.2010.103

FAO. (2018). *The state of Mediterranean and Black Sea fisheries.* Rome, Italy.

FAO. (2020). The State of Mediterranean and Black Sea Fisheries 2020. In *The State of Mediterranean and Black Sea Fisheries 2020*. Rome, Italy. doi: 10.4060/cb2429en

Feuilloley, G., Fromentin, J.-M., Stemmann, L., Demarcq, H., Estournel, C., & Saraux, C. (2020). *Concomitant changes in the Environment and small pelagic fish community of the Gulf of Lions*. doi: 10.1016/j.pocean.2020.102375ï

Flanagan, S. P., Forester, B. R., Latch, E. K., Aitken, S. N., & Hoban, S. (2018). Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications*, *11*(7), 1035–1052. doi: 10.1111/EVA.12569

Fonseca, R. da, Campos, P., Iglesia, A. R. de la, Barroso, G., Bergeron, L., Nande, M., … Castro, L. F. C. (2021). Low coverage whole genome sequencing reveals the underlying structure of European sardine populations. *Authorea Preprints*, 1–32. Retrieved from https://www.authorea.com/users/402855/articles/514589-low-coverage-whole-genome-sequencing-reveals-the-underlying-structure-of-european-sardine-populations

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, *27*(9), 2215–2233. doi: 10.1111/mec.14584

Galindo, H. M., Olson, D. B., & Palumbi, S. R. (2006). Seascape genetics: A coupled oceanographic-genetic model predicts population structure of Caribbean corals. *Current Biology*, *16*(16), 1622–1626.

García-Rivera, S., Lizaso, J. L. S., & Millán, J. M. B. (2017). Composition, spatial distribution and sources of macro-marine litter on the Gulf of Alicante seafloor (Spanish Mediterranean). *Marine Pollution Bulletin*, *121*(1–2), 249–259. doi: 10.1016/j.marpolbul.2017.06.022

Garrido, S., Cristóvão, A., Caldeira, C., Ben-Hamadou, R., Baylina, N., Batista, H., … Santos, A. M. P. (2016). Effect of temperature on the growth, survival, development and foraging behaviour of Sardina pilchardus larvae. *Marine Ecology Progress Series*, *559*, 131–145. doi: 10.3354/meps11881

Garrido, S., Rosa, R., Ben-Hamadou, R., Cunha, M. E., Chícharo, M. A., & van der Lingen, C. D. (2007). Effect of maternal fat reserves on the fatty acid composition of sardine (Sardina pilchardus) oocytes. *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology*, *148*(4), 398–409. doi: 10.1016/j.cbpb.2007.07.008

Garrido, S., Rosa, R., Ben-Hamadou, R., Cunha, M. E., Chícharo, M. A., & Van Der Lingen, C. D. (2008). Spatio-temporal variability in fatty acid trophic biomarkers in stomach contents and muscle of Iberian sardine (Sardina pilchardus) and its relationship with spawning. *Marine Biology*, *154*(6), 1053–1065. doi: 10.1007/s00227-008-0999-7

Garrido, S., Silva, A., Marques, V., Figueiredo, I., Bryère, P., Mangin, A., & Santos, A. M. P. (2017). Temperature and food-mediated variability of European Atlantic sardine recruitment. *Progress in Oceanography*, *159*(October), 267–275. doi: 10.1016/j.pocean.2017.10.006

Garrido, S., & van der Lingen, C. D. (2014). Feeding Biology and Ecology. In *Biology and Ecology of Sardines and Anchovies* (pp. 135–202). CRC Press. doi: 10.1201/b16682-10

Georgakopoulou, E., Katharios, P., Divanach, P., & Koumoundouros, G. (2010). Effect of temperature on the development of skeletal deformities in Gilthead seabream (Sparus aurata Linnaeus, 1758). *Aquaculture*, *308*, 13–19. 77956618251&partnerID=40&md5=0207c4c018552596d5ad56ba6b888452

Gordó-Vilaseca, C., Pennino, M. G., Albo-Puigserver, M., Wolff, M., & Coll, M. (2021). Modelling the spatial distribution of Sardina pilchardus and Engraulis encrasicolus spawning habitat in the NW Mediterranean Sea. *Marine Environmental Research*, *169*, 105381–105381. doi: 10.1016/j.marenvres.2021.105381

Greenbaum, G., Rubin, A., Templeton, A. R., & Rosenberg, N. A. (2019). Network-based hierarchical population structure analysis for large genomic data sets. *Genome Research*, *29*(12), 2020–2033. doi: 10.1101/gr.250092.119

Griswold, C. K. (2006). Pleiotropic mutation, modularity and evolvability. *Evolution & Development*, *8*(1), 81–93. doi: 10.1111/j.1525-142X.2006.05077.x

Grummer, J. A., Beheregaray, L. B., Bernatchez, L., Hand, B. K., Luikart, G., Narum, S. R., & Taylor, E. B. (2019). Aquatic Landscape Genomics and Environmental Effects on Genetic Variation. *Trends in Ecology and Evolution*, *34*(7), 641–654. doi: 10.1016/j.tree.2019.02.013

Guillot, G., Vitalis, R., Rouzic, A. le, & Gautier, M. (2014). Detecting correlation between allele frequencies and environmental variables as a signature of selection: A fast computational approach for genome-wide studies. *Spatial Statistics*, *8*(C), 145–155. doi: 10.1016/j.spasta.2013.08.001

Hidalgo, M., Rossi, V., Monroy, P., Ser-Giacomi, E., Hernández-García, E., Guijarro, B., … Reglero, P. (2019). Accounting for ocean connectivity and hydroclimate in fish recruitment fluctuations within transboundary metapopulations. *Ecological Applications*, *29*(5), 1–13. doi: 10.1002/eap.1913

Hoffmann, A. A., & Willi, Y. (2008). Detecting genetic responses to environmental change. *Nature Reviews Genetics*, *9*(6), 421–432. doi: 10.1038/nrg2339

Huang, J., Li, Y., Liu, Z., Kang, Y., & Wang, J. (2018). Transcriptomic responses to heat stress in rainbow trout Oncorhynchus mykiss head kidney. *Fish and Shellfish Immunology*, *82*, 32–40. doi: 10.1016/j.fsi.2018.08.002

Huang, Z., Ma, A., Yang, S., Liu, X., Zhao, T., Zhang, J., … Xu, R. (2020). Transcriptome analysis and weighted gene co-expression network reveals potential genes responses to heat stress in turbot Scophthalmus maximus. *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*, *33*, 100632. doi: 10.1016/j.cbd.2019.100632

ICES. (2015). *Report of the Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA), 24-29 June 2015, Lisbon, Portugal*. Lisbon, Portugal. Retrieved from www.ices.dk

Jackson, D. A. (1993). Stopping rules in principal components analysis: A comparison of heuristical and statistical approaches. *Ecology*, *74*(8), 2204–2214. doi: 10.2307/1939574

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

Kasapidis, P, Silva, A., Zampicinini, G., & Magoulas, A. (2012). Evidence for microsatellite hitchhiking selection in European sardine (Sardina pilchardus) and implications in inferring stock structure. *Evidencia Para Selección Hitchhiking de Microsatélite En Sardina Europea (Sardina Pilchardus) y Las Implicaciones Para Inferir La Estructura Del StockScientia Marina*, *76*(1), 123–132.

Kasapidis, Panagiotis. (2014). Phylogeography and Population Genetics. In K. Ganias (Ed.), *Biology and Ecology of Sardines and Anchovies* (pp. 43–75). CRC Press, Taylor and Francis Group. doi: 10.1201/b16682-7

Kass, E. R., & Raftery, E. A. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*, 773–795.

Kess, T., Bentzen, P., Lehnert, S. J., Sylvester, E. V. A., Lien, S., Kent, M. P., … Bradbury, I. R. (2020). Modular chromosome rearrangements reveal parallel and nonparallel adaptation in a marine fish. *Ecology and Evolution*, *10*(2), 638–653. doi: 10.1002/ece3.5828

Kess, T., Gross, J., Harper, F., & Boulding, E. G. (2016). Low-cost ddRAD method of SNP discovery and genotyping applied to the periwinkle Littorina saxatilis. *Journal of Molluscan Studies*, *82*(1), 104–089. doi: 10.1093/mollus/eyv042

Kliebenstein, D. J. (2011). Genetic and functional modularity: How does an organism solve a nearly infinite genetic/environmental problem space. *Heredity*, *106*(6), 909–910. doi: 10.1038/hdy.2010.136

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). <scp>Clumpak</scp> : a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources*, *15*(5), 1179–1191. doi: 10.1111/1755-0998.12387

Laurent, V., Caneco, B., Magoulas, A., & Planes, S. (2007). Isolation by distance and selection effects on genetic structure of sardines Sardina pilchardus Walbaum. *Journal of Fish Biology*, *71*(SUPPL. A), 1–17. doi: 10.1111/j.1095-8649.2007.01450.x

Le Corre, V., & Kremer, A. (2003). Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*, *164*(3), 1205–1219. doi: 10.1093/genetics/164.3.1205

Le Nagard, H., Chao, L., & Tenaillon, O. (2011). The emergence of complexity and restricted pleiotropy in adapting networks. *BMC Evolutionary Biology*, *11*(1). doi: 10.1186/1471-2148-11-326

Leroi, A. M. (2000). The scale independence of evolution. *Evolution and Development*, *2*(2), 67–77. doi: 10.1046/j.1525-142x.2000.00044.x

Li, P., Liu, Q., Li, J., Wang, F., Wen, S., & Li, N. (2021). Transcriptomic responses to heat stress in gill and liver of endangered Brachymystax lenok tsinlingensis. *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*, *38*, 100791. doi: 10.1016/j.cbd.2021.100791

Liggins, L., Treml, E. A., & Riginos, C. (2019). *Seascape Genomics: Contextualizing Adaptive and Neutral Genomic Variation in the Ocean Environment*. doi: 10.1007/13836\_2019\_68

Lloret-Lloret, E., Albo-Puigserver, M., Giménez, J., Navarro, J., Pennino, M. G., Steenbeek, J., … Coll, M. (2022). Small pelagic fish fitness relates to local environmental conditions and trophic variables. *Progress in Oceanography*, *202*, 102745. doi: 10.1016/J.POCEAN.2022.102745

Lotterhos, K. E., Yeaman, S., Degner, J., Aitken, S., & Hodgins, K. A. (2017). Modularity of genes involved in local adaptation to climate despite physical linkage. *BioRxiv*, 1–24. doi: 10.1101/202481

Louro, B., De Moro, G., Garcia, C., Cox, C. J., Veríssimo, A., Sabatino, S. J., … Canário, A. V. M. (2019). A haplotype-resolved draft genome of the European sardine (Sardina pilchardus). *GigaScience*, *8*(5), 1–8. doi: 10.1093/gigascience/giz059

Lowerre-Barbieri, S., DeCelles, G., Pepin, P., Catalán, I. A., Muhling, B., Erisman, B., … Paris, C. B. (2017). Reproductive resilience: a paradigm shift in understanding spawner-recruit systems in exploited marine fish. *Fish and Fisheries*, *18*(2), 285–312. doi: 10.1111/FAF.12180

Luu, K. (2017). *pcadapt : an R package to perform genome scans for selection based on principal component analysis*. *33*, 67–77. doi: 10.1111/1755-0998.12592

Machu, E., Ettahiri, O., Kifani, S., Benazzouz, A., Makaoui, A., & Demarcq, H. (2009). Environmental control of the recruitment of sardines (Sardina pilchardus) over the western saharan shelf between 1995 and 2002: A coupled physical-biogeochemical modelling experiment. *Fisheries Oceanography*, *18*(5), 287–300. doi: 10.1111/j.1365-2419.2009.00511.x

Metivier, S. L., Kim, J. H., & Addison, J. A. (2017). Genotype by sequencing identifies natural selection as a driver of intraspecific divergence in Atlantic populations of the high dispersal marine invertebrate, Macoma petalum. *Ecology and Evolution*, *7*(19), 8058–8072. doi: 10.1002/ece3.3332

Micheli, F., Halpern, B. S., Walbridge, S., Ciriaco, S., Ferretti, F., Fraschetti, S., … Rosenberg, A. A. (2013). Cumulative Human Impacts on Mediterranean and Black Sea Marine Ecosystems: Assessing Current Pressures and Opportunities. *PLoS ONE*, *8*(12), e79889. doi: 10.1371/journal.pone.0079889

Moody, K. N., Hunter, S. N., Childress, M. J., Blob, R. W., Schoenfuss, H. L., Blum, M. J., & Ptacek, M. B. (2015). Local adaptation despite high gene flow in the waterfall-climbing Hawaiian goby, Sicyopterus stimpsoni. *Molecular Ecology*, *24*(3), 545–563. doi: 10.1111/MEC.13016

Naciri, M., Lemaire, C., Borsa, P., & Bonhomme, F. (1999). Genetic study of the Atlantic/Mediterranean transition in sea bass (Dicentrarchus labrax). *Journal of Heredity*, *90*(6), 591–596. doi: 10.1093/jhered/90.6.591

Nielsen, E. E., Hemmer-Hansen, J., Poulsen, N. A., Loeschcke, V., Moen, T., Johansen, T., … Carvalho, G. R. (2009). Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (Gadus morhua). *BMC Evol Biol*, *9*, 276. doi: 1471-2148-9-276 [pii]10.1186/1471-2148-9-276

Oksanen, A. J., Blanchet, F. G., Kindt, R., Legen-, P., Minchin, P. R., Hara, R. B. O., … Stevens, M. H. H. (2012). Community Ecology Package. *… Ecology Package …*, p. 263. Retrieved from http://mirror.bjtu.edu.cn/cran/web/packages/vegan/vegan.pdf

Palomera, I., Olivar, M. P., Salat, J., Sabatés, A., Coll, M., García, A., & Morales-Nin, B. (2007). Small pelagic fish in the NW Mediterranean Sea: An ecological review. *Progress in Oceanography*, *74*(2–3), 377–396. doi: 10.1016/j.pocean.2007.04.012

Parrish, R. H., Serra, R., & Grant, W. S. (1989). The monotypic sardines, Sardina and Sardinops: their taxonomy, distribution, stock structure, and zoogeography. *Canadian Journal of Fisheries and Aquatic Sciences*, *46*(11), 2019–2036. doi: 10.1139/f89-251

Pennino, M. G., Coll, M., Albo-Puigserver, M., Fernández-Corredor, E., Steenbeek, J., Giráldez, A., … Bellido, J. M. (2020). Current and Future Influence of Environmental Factors on Small Pelagic Fish Distributions in the Northwestern Mediterranean Sea. *Frontiers in Marine Science*, *7*, 622. doi: 10.3389/fmars.2020.00622

Peristeraki, P., Bitetto, I., Carbonara, P., Carlucci, R., Certain, G., De Carlo, F., … Tserpes, G. (2019). Investigation of spatiotemporal patterns in mean temperature and mean trophic level of MEDITS survey catches in the Mediterranean Sea. *Scientia Marina*, *83*(S1), 165–174.

Price, N., Lopez, L., Platts, A. E., & Lasky, J. R. (2020). In the presence of population structure: From genomics to candidate genes underlying local adaptation. *Ecology and Evolution*, *10*(4), 1889–1904. doi: 10.1002/ece3.6002

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *BIOINFORMATICS APPLICATIONS NOTE*, *26*(6), 841–842. doi: 10.1093/bioinformatics/btq033

R Core Team. (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.r-project.org/

Raj, A., Stephens, M., & Pritchard, J. K. (2014). FastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), 573–589. doi: 10.1534/genetics.114.164350

Ramírez, F., Afán, I., Davis, L. S., & Chiaradia, A. (2017). Climate impacts on global hot spots of marine biodiversity. *Science Advances*, *3*(2), e1601198. doi: 10.1126/sciadv.1601198

Ramírez, F., Coll, M., Navarro, J., Bustamante, J., & Green, A. J. (2018). Spatial congruence between multiple stressors in the Mediterranean Sea may reduce its resilience to climate impacts. *Scientific Reports*, *8*(1), 14871. doi: 10.1038/s41598-018-33237-w

Ramírez, F., Pennino, M. G., Albo-Puigserver, M., Steenbeek, J., Bellido, J. M., & Coll, M. (2021). SOS small pelagics: A safe operating space for small pelagic fish in the western Mediterranean Sea. *Science of the Total Environment*, *756*, 144002. doi: 10.1016/j.scitotenv.2020.144002

Ramon, M. M., & Castro, J. A. (1997). Genetic variation in natural stocks of Sardina pilchardus (sardines) from the western Mediterranean Sea. *Heredity*, *78*(5), 520–528. doi: 10.1038/hdy.1997.81

Riginos, C., Crandall, E. D., Liggins, L., Bongaerts, P., & Treml, E. A. (2016). Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies. *Current Zoology*, zow067. doi: 10.1093/cz/zow067

Sandoval-Castillo, J., & Beheregaray, L. B. (2020). Oceanographic heterogeneity influences an ecological radiation in elasmobranchs. *Journal of Biogeography*, *47*(7), 1599–1611. doi: 10.1111/JBI.13865

Santos, A. M. P., Nieblas, A. E., Verley, P., Teles-Machado, A., Bonhommeau, S., Lett, C., … Peliz, A. (2018). Sardine (Sardina pilchardus) larval dispersal in the Iberian upwelling system, using coupled biophysical techniques. *Progress in Oceanography*, *162*(February), 83–97. doi: 10.1016/j.pocean.2018.02.011

Santos, A. M. P., Re, P., Dos Santos, A., & Peliz, Á. (2006). Vertical distribution of the European sardine (Sardina pilchardus) larvae and its implications for their survival. *Journal of Plankton Research*, *28*(5), 523–532. doi: 10.1093/plankt/fbi137

Saraux, C., Van Beveren, E., Brosset, P., Queiros, Q., Bourdeix, J. H., Dutto, G., … Fromentin, J. M. (2019). Small pelagic fish dynamics: A review of mechanisms in the Gulf of Lions. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *159*, 52–61. doi: 10.1016/j.dsr2.2018.02.010

Selkoe, K. A., Watson, J. R., White, C., Ben Horin, T., Iacchei, M., Mitarai, S., … Toonen, R. J. (2010). Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology*, *19*(17), 3708–3726.

Solari, A. P., Santamaría, M. T. G., Borges, M. F., Santos, A. M. P., Mendes, H., Balguerías, E., … Bas, C. (2010). On the dynamics of Sardina pilchardus: Orbits of stability and environmental forcing. *ICES Journal of Marine Science*, *67*(8), 1565–1573. doi: 10.1093/icesjms/fsq107

Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500. doi: 10.1126/science.1227079

Teske, P. R., Emami-Khoyi, A., Golla, T. R., Sandoval-Castillo, J., Lamont, T., Chiazzari, B., … van der Lingen, C. D. (2021). The sardine run in southeastern Africa is a mass migration into an ecological trap. *Science Advances*, *7*(38). doi: 10.1126/sciadv.abf4514

UNEP/MAP. (2015). *Marine Litter Assessment in the*.

Van Beveren, E., Bonhommeau, S., Fromentin, J. M., Bigot, J. L., Bourdeix, J. H., Brosset, P., … Saraux, C. (2014). Rapid changes in growth, condition, size and age of small pelagic fish in the Mediterranean. *Marine Biology*, *161*(8), 1809–1822. doi: 10.1007/s00227-014-2463-1

Vandeputte, M., Clota, F., Sadoul, B., Blanc, M. O., Blondeau-Bidet, E., Bégout, M. L., … Geffroy, B. (2020). Low temperature has opposite effects on sex determination in a marine fish at the larval/postlarval and juvenile stages. *Ecology and Evolution*, *10*(24), 13825–13835. doi: 10.1002/ece3.6972

Vasilakopoulos, P., Maravelias, C. D., & Tserpes, G. (2014). The alarming decline of mediterranean fish stocks. *Current Biology*, *24*(14), 1643–1648. doi: 10.1016/j.cub.2014.05.070

Wagner, G. P., & Altenberg, L. (1996). Perspective: Complex Adaptations and the Evolution of Evolvability. *Evolution*, *50*(3), 967. doi: 10.2307/2410639

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., … Shi, C. (2018). WEGO 2.0: A web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Research*, *46*(W1), W71–W75. doi: 10.1093/nar/gky400

Zafeiropoulos, H., Gioti, A., Ninidakis, S., Potirakis, A., Paragkamian, S., Angelova, N., … Pafilis, E. (2021). 0s and 1s in marine molecular research: a regional HPC perspective. *GigaScience*, *10*(8), 1–12. doi: 10.1093/GIGASCIENCE/GIAB053

Zelle, H., Appeldoorn, G., Burges, G., & van Oldenborgh, G. J. (2004). The relationship between sea surface temperature and thermocline depth in the eastern equatorial Pacific. *Journal of Physical Oceanography*, *34*(3), 643–655. doi: 10.1175/2523.1

Zhao, Q., & Costello, M. J. (2019). Summer and winter ecosystems of the world ocean photic zone. *Ecological Research*, *34*(4), 457–471. doi: 10.1111/1440-1703.12006

**Data Accessibility**

Raw reads have been deposited to ENA under the accession id PRJEB45992.

**Author Contributions**

A. A. performed population genetic and seascape genomic analyses and wrote the paper, T. M. carried out bioinformatic analysis on SNP calling, population genetics and functional annotation and wrote the paper, F.R. produced environmental layers and wrote the paper, A. C., R. C., P. K., A. M., M. A. - P., E. Ll-Ll., J. M. B., M. G. P., M. C. F., A. E., C. S., M. S., M. T. S., provided their expertise and commented on the paper while A. C., R. C., M. A. - P., E. Ll-Ll., J. M. B., M. G. P., M. C. F., A. E., C. S., M. S., M. T. S. were also responsible for samples collection, M. C. and C. S. T. conceived and supervised the study and wrote the paper. All authors approved the final version of the paper. Competing interests: the authors declare no competing interests.

**Figure Captions**

**Figure 1**. Geographic distribution of sampling sites. Colors in pies indicate the proportion of males, females and immatures captured, while pie size is proportional to sardines mean length. Bathymetry of the studied area is also indicated with a gradient of blue colors with darker colors for deeper seas.

**Figure 2**. Population structure of sardines in the Western Mediterranean Sea as inferred by STRUCTURE for K=2. **A.** Bar plot of individual q-values after CLUMPAK analysis (plot created with Distruct v2.3, Chhatre, 2018; Raj, Stephens, & Pritchard, 2014), **B.** Spatial interpolates (Krig interpolation in ArcGIS) of individual counts in the three assignment categories of STRUCTURE (i.e. individuals with q-values ≥ 0.90 in either Atlantic (ATL) or Mediterranean (MED) cluster and individuals with 0.10 ≤ q-value ≤ 0.90 “Admixed”) per sampling site. Sites are identical to the locations and names given in Figure 1.

**Figure 3**. Results of the DAPC analysis: **A.** Bar plot of individual posterior membership probability values for K=2 as indicated by BIC (plot created with Distruct v2.3, Chhatre, 2018; Raj et al., 2014), and **B.** density plot of the distribution of each cluster on the discriminant axis. In blue, the cluster that mostly contains samples from the Atlantic (similar to ATL cluster of STRUCTURE analyses) and in red, the cluster that mostly contains samples from the Mediterranean (similar to MED cluster of STRUCTURE analyses).

**Figure 4:** Correlation of gINLAnd outlier single nucleotide polymorphisms (SNPs) with environmental variables as a signature of selection in “all samples” dataset. Hierarchical clustering of significant (i.e. logBF values ≥ 3) outlier SNPs (in rows) and environmental variables (in columns). Environmental variables in columns correspond to V1: PLD\_60, V2: Mesopelag, V3: Bathypelag, V4: EcoReg, V5: BioReg, V6: primprod, V7: o2saturate, V8: bo\_ph, V9: ClusterCli, V10: windspeed, V11: surcurrent, V12: parmean, V13: Raw\_CluCli, V14: Currents, V15: FrontiersS, V16: nitrate, V17: calcite, V18: bsilicate, V19: bphosphate, V20: bo2utilize, V21: bo2dissolv, V22: bnitrate, V23: biogeo10, V24: biogeo08, V25: salinity, V26: tot\_impact, V27: sst\_tr26, V28: sst\_tr19, V29: sst\_tr12, V30: sst\_min\_sl, V31: sst\_m\_slwi, V32: sst\_m\_sl, V33: sst\_max\_sl, V34: biogeo17, V35: biogeo16, V36: biogeo15, V37: biogeo14, V38: sst\_mayoct, V39: bedtemp.

**Figure 5:** Correlation of gINLAnd outlier single nucleotide polymorphisms (SNPs) with environmental variables as a signature of selection in MED cluster. Hierarchical clustering of significant (i.e. logBF values ≥ 3) outlier SNPs (in rows) and environmental variables (in columns). Numbers in rows indicate the locus that bears the outlier SNP while the environmental variables in columns correspond to V1: BioReg, V2: bo\_ph, V3: surcurrent, V4: parmean, V5: Currents, V6: FrontiersS, V7: nitrate, V8: bphosphate, V9: bo2utilize, V10: bo2dissolv, V11: salinity, V12: tot\_impact, V13: sst\_tr26, V14: sst\_tr19, V15: sst\_tr12, V16: sst\_min\_sl, V17: sst\_max\_sl, V18: btemp.

**Figure 6:** Correlation of gINLAnd outlier single nucleotide polymorphisms (SNPs) with environmental variables as a signature of selection in ATL cluster. Hierarchical clustering of significant (i.e. logBF values ≥ 3) outlier SNPs (in rows) and environmental variables (in columns). Numbers in rows indicate the locus that bears the outlier SNP while the environmental variables in columns correspond to V1: o2saturate, V2: bo\_ph, V3: windspeed, V4: surcurrent, V5: parmean, V6: Currents, V7: nitrate, V8: bsilicate, V9: bphosphate, V10: bo2utilize, V11: bo2dissolv, V12: bnitrate, V13: biogeo12, V14: biogeo11, V15: sst\_tr19, V16: sst\_mayoct, V17: btemp, V18: bedtemp, V19: biogeo02, V20: biogeo03, V21: biogeo06, V22: biogeo05, V23: biogeo07, V24: biogeo04.

**Figure 7:** Correlation of gINLAnd outlier single nucleotide polymorphisms (SNPs) with environmental variables as a signature of selection in the “northern vs southern sites" dataset. Hierarchical clustering of significant (i.e. logBF values ≥ 3) outlier SNPs (in rows) and environmental variables (in columns). Numbers in rows indicate the locus that bears the outlier SNP while the environmental variables in columns correspond to V1: tot\_impact, V2: sst\_tr26, V3: sst\_tr19, V4: sst\_max\_sl, V5: biogeo09, V6: Cli\_impact, V7: Cum\_impact.

**Figure 8:** Triplots of “all samples” dataset for RDA axes 1 and 2 (**A**), and RDA axes 1 and 3 (**B**). Triplots of MED dataset for RDA axes 1 and 2 (**C**), and RDA axes 1 and 3 (**D**). The dark grey cloud of points at the centre of each plot represent the SNPs. The coloured points represent individual sardines with coding by sampling site. Blue vectors represent environmental predictors (see Table 1 for abbreviations).

**Figure S1:** Gene Ontology summary of the GO terms characterizing the genes that include outlier loci discovered in the seascape genomics analysis