# Population structure and genetic variability in wild and farmed Mediterranean populations of gilthead seabream and European seabass inferred from a 60K combined species SNP array

Aquaculture Reports 24 (2022) 101145

Contents lists available at ScienceDirect

Aquaculture Reports

journal homepage: www.elsevier.com/locate/aqrep

# Population structure and genetic variability in wild and farmed Mediterranean populations of gilthead seabream and European seabass inferred from a 60K combined species SNP array

Beatriz Villanueva [a,*], Almudena Fernández [a], Ramón Peiró-Pastor [a], Carolina Peñaloza [b], Ross D. Houston [b,1], Anna K. Sonesson [c], Costas S. Tsigenopoulos [d], Luca Bargelloni [e], Kutsal Gamsız [f], Bilge Karahan [f], Emel Ö. Gökçek [f], Jesús Fernández [a,2], María Saura [a,2]

[a] Departamento de Mejora Genética Animal, INIA-CSIC, Ctra. de La Coruña, Km 7.5, 28040 Madrid, Spain
[b] The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian EH25 9RG, UK
[c] NOFIMA, Ås, Norway
[d] Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC), Hellenic Centre for Marine Research (HCMR), Thalassocosmos Gournes Pediados, 71500 Heraklion, Crete, Greece
[e] Universitá degli Studi di Padova, Via 8 Febbraio 1848, 2, 35122 Padova PD, Italy
[f] Department of Aquaculture, Faculty of Fisheries, University of Ege, TR-35100 Izmir, Turkey

ARTICLE INFO

*Keywords:*
Gilthead seabream
European seabass
Population structure
Genetic variability
Effective population size
SNP array

ABSTRACT

Knowledge of population structure and genetic diversity within and between wild and farmed populations of gilthead sea bream (*Sparus aurata*) and European seabass (*Dicentrarchus labrax*) is important to achieve sustainable aquaculture production of these species and to assess the risk of genetic impacts of fish escaped from farms. Previous population genetic studies on these species have been based on a limited number of genetic markers and samples. In this study, these features were assessed using samples from 24 seabream and 25 seabass populations distributed throughout the Mediterranean Sea, and 3 wild seabream Atlantic populations. Samples were genotyped with a newly developed combined species SNP array that includes ~60K SNPs. Data from sequencing pools of individual DNA from the same populations were also used. Different approaches were employed for identifying the extent of population stratification within species. The effective population size (a parameter inversely related to the rate at which genetic variability is lost) was estimated for each population based on linkage disequilibrium. Population structure results revealed a clear differentiation between wild and farmed populations in both species. Wild populations showed a low degree of differentiation, particularly in seabream. Despite this, a slight differentiation was observed between Atlantic and Mediterranean seabream populations and between western and eastern Mediterranean seabass populations. However, farmed populations were quite heterogeneous and showed a high degree of differentiation. Some farmed populations of both species showed a genetic makeup similar to that found in wild populations. In general, the effective population size was large (> 1000) for wild and small (< 100) for farmed populations of both species. About 40% of the seabream and 80% of the seabass farmed populations had estimates of effective population size smaller than 50 highlighting the need of applying measures to control the rate at which genetic variability is lost.

## 1. Introduction

Knowledge of population structure and genetic diversity within and between wild and farmed fish populations is of paramount importance to develop optimal strategies for the conservation of native fish and to achieve sustainable aquaculture production. The success of any breeding program is critically dependent on the way in which the base population of breeders is built, as the genetic variability initially available in the

* Corresponding author.
*E-mail address:* villanueva.beatriz@inia.csic.es (B. Villanueva).
[1] Present address: Benchmark Genetics, 1 Pioneer Building, Edinburgh Technopole, Milton Bridge, Penicuik EH26 0GB, UK.
[2] These authors contributed equally

https://doi.org/10.1016/j.aqrep.2022.101145
Received 14 January 2022; Received in revised form 19 April 2022; Accepted 27 April 2022
Available online 9 May 2022

founders will determine the genetic progress achieved in the subsequent generations of selection. These base populations can be established from wild and farmed populations and it is thus necessary to determine the relatedness between them in order to optimise their creation (Gjedrem and Baranski, 2010; Fernández et al., 2014). Also, once genetic improvement programs are operational, the control of inbreeding and loss of genetic variability is needed to ensure their long-term sustainability. In particular, given the high fecundity of fish species, a large number of individuals can be produced from a reduced number of broodstock with the consequent decrease in effective population size (a parameter inversely related to the loss of genetic variability) and increase in inbreeding. The problem may be exacerbated in marine species that exhibit mass-spawning behaviour as a large variance in individual parental contribution to offspring may be expected (e.g. Brown et al., 2005).

Gilthead seabream (*Sparus aurata*) and European seabass (*Dicentrarchus labrax*) are two of the main cultured finfish species in Europe. They are highly traded, particularly towards prosperous markets (FAO, 2018) and farmed almost entirely in the Mediterranean where they have become the most important marine farmed fish (Žužul et al., 2019), ranking third (seabass) and fourth (seabream) in European aquaculture production (Eurostat, 2019). Although still less developed than in salmonids, selective breeding is expected to play an increasingly important role in the commercial production of both species. Different companies have initiated breeding programs in recent years and the proportion of genetically improved stocks has increased (Janssen et al., 2017). There are about four to five breeding programs in operation for each species and the most recent available information (Janssen et al., 2017) indicates that the number of selected generations varied between two and eight in seabass and between one and five in seabream.

Escapees from aquaculture facilities are perceived as a major threat to natural biodiversity. It is clear that they pose an ecological risk of transferring diseases to wild fish and may cause undesirable genetic effects in native populations due to interbreeding. Little is known about escapees from seabream and seabass farms (Polovina et al., 2020) although some studies have estimated rates as high as 10–15% (Šegvić-Bubić et al., 2011; Brown et al., 2015). Knowledge on the relatedness between wild and farmed populations will help to assess the level of risk of genetic impacts of any interbreeding between wild fish and fish escaped from farms (Žužul et al., 2019; Polovina et al., 2020).

Most studies investigating population structure in gilthead seabream (Alarcón et al., 2004; De Innocentiis et al., 2004; Šegvić-Bubić et al., 2011; Coscia et al., 2012; Franchini et al., 2012; Loukovitis et al., 2012; Žužul et al., 2019; Polovina et al., 2020) and European seabass (Allegrucci et al., 1997; Caccone et al., 1997; Naciri et al., 1999; Bahri-Sfar et al., 2000; Quéré et al., 2012; Souche et al., 2015; Bodur et al., 2017; Polovina et al., 2020) have been based on a limited number of genetic markers (mostly microsatellites) and often on a limited number of samples. Results on the degree of differentiation among populations from these studies have been inconclusive, particularly for Mediterranean populations.

Despite the importance of effective population size, estimates of this parameter for seabream and seabass are scarce. Until very recently the few published estimates for both wild and farmed populations were obtained from demographic data (Brown et al., 2005; Borrell et al., 2011) or from linkage disequilibrium measures based on a very limited number of microsatellite markers (Loukovitis et al., 2012; Šegvić-Bubić et al., 2017). Recently, Saura et al. (2021) provided estimates obtained using a large number of SNP genotypes (> 8000) derived from RAD-seq but only for a limited number of farmed populations (two seabream and one seabass populations).

Key genomic resources have been developed for both species including high-quality reference genomes (Tine et al., 2014, Pauletto et al., 2018), which have recently been improved (see GenBank: GCA_905237075.1 and GCA_900880675.2), and SNP arrays. SNP arrays are high-throughput genotyping platforms that are potentially more powerful tools for assessing population structure and genetic variability. In particular, Griot et al. (2021) report the development of two ThermoFisher Axiom™ SNP arrays, one 57K for seabass (DlabChip) and one 60K for seabream (SaurChip) while a combined ~60K SNP array for both species (the 'MedFish', also a ThermoFisher Axiom™ SNP array) has recently been developed. Details on SNPs shared by platforms are given in Peñaloza et al. (2021). The MedFish array, a result of collaboration between the EU projects MedAID (http://www.medaid-h2020.eu/) and PerformFISH (http://www.performfish.eu/), is described in Peñaloza et al. (2021), and it includes ~30K markers for seabream and ~30K markers for seabass. The array was designed using samples from 24 populations of European seabass and 27 populations of gilthead seabream that were distributed throughout the species range.

The objective of this study was to make use of the potentially powerful combined MedFish SNP array for gilthead seabream and European seabass to i) identify the extent of wild and farmed population stratification within species, considering the geographic origin of individuals; and ii) estimate genetic variability, inbreeding and effective population size in populations of both species. A large number of wild and farmed populations distributed throughout the Mediterranean Sea were analyzed for both species, including also three wild seabream populations of Atlantic origin.

## 2. Material and methods

### 2.1. Data description

#### 2.1.1. Samples

Samples available came from 11 wild and 13 farmed seabream populations and from 9 wild and 16 farmed seabass populations, distributed from the Western to the Eastern Mediterranean Sea. Additionally, three wild seabream populations sampled in the Atlantic Ocean were also included in the analyses. Most samples were the same as those described in Peñaloza et al. (2021). A brief description of the samples available for the study is given in Table 1. Populations were ordered according to their geographic location, from West to East. Sampling locations for wild populations are shown in Fig. 1.

#### 2.1.2. SNP array data

The genotype datasets analyzed in this study were basically those described in Peñaloza et al. (2021). SNP genotypes were available for 462 seabream and 516 seabass individuals belonging to 26 seabream and 24 seabass populations (see Table 1). From the initial number of informative SNPs in the array, monomorphic SNPs and those without observed heterozygote genotypes were removed from the dataset. After this filtering, the total number of SNPs available for analysis was 25,319 for seabream and 22,507 for seabass. For seabream, the number of SNPs on chromosome-level scaffolds (Table S1 in Supplementary Material) ranged from 816 (chromosome 24) to 1188 (chromosome 6) and the SNP density (number of SNPs per Mb) ranged from 40.30 (chromosome 1) to 51.17 (chromosome 24). For seabass, the number of SNPs per linkage group (LG) (Table S1 in Supplementary Material) ranged from 566 (LG3) to 1095 (LG5) and the SNP density ranged from 33.59 (LG5) to 45.57 (LG24). The distributions of the MAF (minimum allele frequency) for the SNPs used in the analyses of both species are shown in Fig. S1 (Supplementary Material).

#### 2.1.3. Pool-sequencing data

For the same seabass and seabream populations, data obtained from sequencing pools of individual DNA (pool-seq data) used for developing the array (Peñaloza et al., 2021) were also available. Compared with SNP array genotyping, pool-seq has the advantage that it generates information for millions of SNPs, which allows high precision when estimating allele frequencies at the population level. The information available for pool-seq was therefore used to perform some population structure analyses based on allele frequencies (see below). As described

**Table 1**
Summary of seabream and seabass populations sampled, including origin, region, country, population ID, number of fish per pool (*N*/pool) and number of fish genotyped in each population ($N_g$).

| | Seabream | | | | | Seabass | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Origin | Region | Country | ID | *N*/pool | $N_g$ | Region | Country | ID | *N*/pool | $N_g$ |
| Wild | Atlantic | Spain | W-SPA-1 | 25 | 12 | Mediterranean | Morocco | W-MOR | 25 | 22 |
| | | Spain | W-SPA-2 | 25 | 12 | | Spain | W-SPA | 25 | 10 |
| | | France | W-FRA | 25 | 4 | | France | W-FRA | 25 | 24 |
| | | | | | | | Italy | W-ITA | 25 | 20 |
| | Mediterranean | Spain | W-SPA-3 | 25 | 19 | | Croatia | W-CRO | 25 | 11 |
| | | Spain | W-SPA-4 | 25 | 10 | | Greece | W-GRE-1 | 25 | 24 |
| | | Tunisia | W-TUN | 25 | 19 | | Greece | W-GRE-2 | 25 | 25 |
| | | Italy | W-ITA-1 | 25 | 23 | | Turkey | W-TUR | 25 | 24 |
| | | Italy | W-ITA-2 | 25 | 24 | | Cyprus | W-CYP | 15 | 10 |
| | | Greece | W-GRE-1 | 25 | 15 | | | | | |
| | | Greece | W-GRE-2 | 25 | 21 | | | | | |
| | | Greece | W-GRE-3 | 25 | 12 | | | | | |
| | | Greece | W-GRE-4 | 25 | 16 | | | | | |
| | | Greece | W-GRE-5 | 25 | 22 | | | | | |
| | | Turkey | W-TUR | 25 | 25 | | | | | |
| Farmed | Mediterranean | Spain | F-SPA-1 | 25 | 18 | Mediterranean | Spain | F-SPA-1 | 25 | 22 |
| | | Spain | F-SPA-2 | 25 | 25 | | Spain | F-SPA-2 | 25 | 25 |
| | | France | F-FRA | 25 | 24 | | Italy | F-ITA | 25 | 24 |
| | | Italy | F-ITA | 25 | – | | Croatia | F-CRO-1 | 25 | – |
| | | Croatia | F-CRO | 25 | 19 | | Croatia | F-CRO-2 | 25 | 24 |
| | | Greece | F-GRE-1 | 14 | 12 | | Greece | F-GRE-1 | 25 | 21 |
| | | Greece | F-GRE-2 | 13 | 13 | | Greece | F-GRE-2 | 25 | 21 |
| | | Greece | F-GRE-3 | 25 | 21 | | Greece | F-GRE-3 | 25 | 21 |
| | | Greece | F-GRE-4 | 25 | 24 | | Greece | F-GRE-4 | 25 | 23 |
| | | Greece | F-GRE-5 | 25 | 20 | | Greece | F-GRE-5 | 12 | 12 |
| | | Turkey | F-TUR | 25 | 25 | | Greece | F-GRE-6 | 25 | 21 |
| | | Egypt | F-EGY | 15 | 14 | | Greece | F-GRE-7 | 23 | 22 |
| | | Israel | F-ISR | 25 | 13 | | Turkey | F-TUR-1 | 25 | 22 |
| | | | | | | | Turkey | F-TUR-2 | – | 55 |
| | | | | | | | Egypt | F-EGY | 15 | 12 |
| | | | | | | | Cyprus | F-CYP | 25 | 21 |

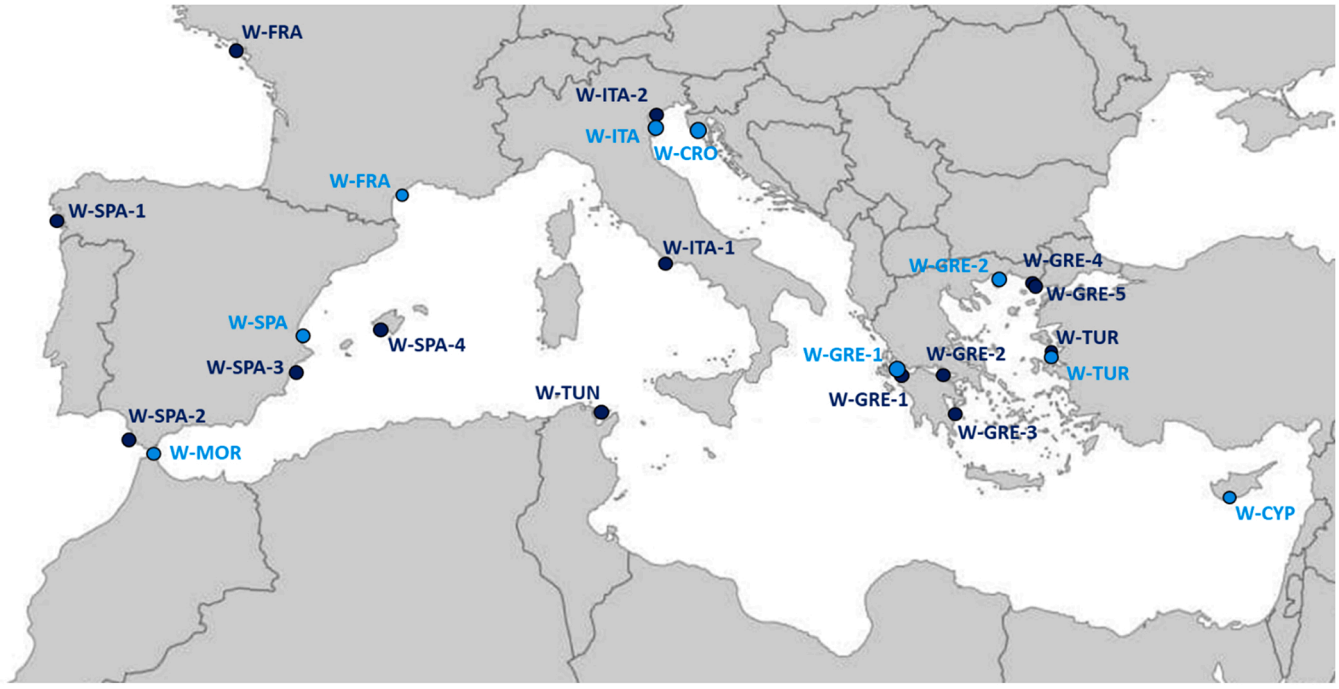


**Fig. 1.** Map of sampling locations for wild seabream (dark blue) and seabass (light blue).

in Peñaloza et al. (2021), most pools were prepared by combining DNA of 25 individual fish (see Table 1) and most population pools were prepared in duplicate; i.e. the DNA of the same 25 individuals was used to prepare two separate pools, which were sequenced independently on an Illumina platform. For a small number of populations, the number of individuals per pool was less than 25. These pools did not have a technical replicate. Paired-end reads from each population pool were aligned to earlier versions of the European seabass and gilthead seabream

genomes (Tine et al., 2014, Pauletto et al., 2018). Read alignment, post-alignment quality control (QC), SNP calling and filtering process are described in Peñaloza et al. (2021). Allele frequencies were extracted for both fish species from the list of high-quality variants called from the pool-seq data. For seabream, the dataset contained frequency information for 1,059,883 SNPs genotyped across 27 population pools (14 wild and 13 farmed). For seabass, the dataset represented 1,064,510 SNPs genotyped across 24 population pools (9 wild and 15 farmed).

Data files were processed using an in-house script to deal with SNPs having a single or both alleles unknown due to QC filters. We first assigned a tag to missing genotype values. Then, we removed SNPs showing more than 40% missing values (in either wild or farmed sets). In the remaining, the missing values were replaced by the median of the frequency values in the rest of populations of the same type (wild or farmed) separately for each SNP. Finally, SNPs with all values equal to zero were removed. Due to this clean up, the number of SNPs was reduced to 1,052,447 (99.3% of the original SNPs) for seabream and to 420,173 (39.5% of the original SNPs) for seabass. These cleansed datasets were the starting point for further analyses. For populations with two replicates, the results presented are based on the average MAF between the two replicates.

### 2.2. Population structure analyses

Different approaches were used for identifying the extent of population stratification within species. These approaches included i) Principal Component Analysis (PCA), using pool-seq and SNP array data; ii) distance-based clustering methods, using pool-seq data; iii) model-based clustering methods, using SNP array data; and iv) Wrightś fixation index ($F_{ST}$), using SNP array data. For the approaches using SNP array data, we pruned the SNP dataset based on pairwise linkage disequilibrium to produce a reduced set of more independent markers, using PLINK (Purcell et al., 2007). Pairwise $r^2$ (the squared correlation between pairs of SNPs; Hill and Robertson, 1968) were computed using a window approach (window size of 50 SNPs, and shift size of 10 SNPs), and the threshold was set to 0.1. A total of 5851 (out of the 25,319) and 7036 (out of the 22,507) SNPs were pruned out for seabream and seabass, respectively.

#### 2.2.1. Principal Component Analysis

For pool-seq data, PCA values were calculated using the 'prcomp' function of the R software. For SNP array data, PCA values were calculated using PLINK. Graphical representations for both datasets were performed using R software. Wild and farmed populations were analyzed jointly.

#### 2.2.2. Distance-based clustering methods

Different clustering methods were used for inferring population structure. In this case, pool-seq rather than SNP array data were used to produce a clearer picture. In an initial stage, hierarchical and non-hierarchical clustering methods were compared using several metrics computed with the 'clValid' function of the 'clValid' R package (Brock et al., 2008): connectivity, Dunn index and Silhouette index. According to these metrics, the optimal clustering method was hierarchical with two groups. Four different hierarchical methods were then used, including three agglomerative clustering methods (complete-linkage, single-linkage and Ward's) and a divisive clustering method (Rokach and Maimon, 2005). To compare the accuracy of the four methods of hierarchical clustering, two metrics were computed: the cophenetic correlation index and the Gower's distance. Wild and farmed populations were analyzed jointly.

#### 2.2.3. Model-based clustering method

Inference of genetic ancestry was performed using the SNP array data and the software Admixture version 1.3.0 (Alexander et al., 2009; Alexander and Lange, 2011; Liu et al., 2013). This software applies a model-based algorithm that simultaneously estimates population allele frequencies along with ancestry proportions. The analysis is based on maximum likelihood estimation of individual ancestries from SNP genotypes, providing the best proportions of admixing components (clusters) for any hypothetical number ($K$) of genetic groups. In order to identify the value of $K$ for which the model has the best predictive accuracy, the Admixture software uses a cross-validation procedure. The $K$ value chosen was that with the lowest standard error for the cross-validation. Wild and farmed populations within each species were analyzed jointly and also separately.

#### 2.2.4. Wrightś fixation index $F_{ST}$

Pairwise $F_{ST}$ coefficients, that measure the degree of differentiation between populations, were obtained from allele frequencies, correcting for sample size following Nei and Chesser (1983) and using the software Metapop2 (López-Cortegano et al., 2019). The fixation index ranges from zero (no differentiation between populations) to one (fixation of different alleles in the different populations). Thus, with low rates of gene flow among populations, they genetically diverge and $F_{ST}$ increases.

### 2.3. Genetic diversity

Genetic diversity within each population was measured as the expected heterozygosity ($H_E$, also called gene diversity), using the SNP array genotypes. It was computed as $\left[2\sum_{k=1}^{L} p_k(1-p_k)\right]\Big/L$, where $L$ is the number of SNPs and $p_k$ is the frequency of the minor allele for SNP $k$. Note that $H_E$ equals $1-f$, where $f$ is the average coancestry. The observed heterozygosity ($H_O$) which equals $1-F$ (where $F$ is the average molecular inbreeding coefficient), was computed for each population as the number of heterozygotes for SNP markers, divided by the number of fish in the population and averaged over all SNPs. Molecular inbreeding coefficients based on SNP genotypes were also calculated as the deviations of the observed frequency of homozygotes from the expected frequency in the base population under Hardy–Weinberg proportions (Li and Horvitz, 1953). Specifically, the inbreeding coefficient of Li and Horvitz ($F_{L\&H}$) was computed as $[(1-H_O)-(1-H_E)]/[1-(1-H_E)]$.

### 2.4. Effective population size

Estimates of current effective population size ($N_e$) were obtained for all populations from linkage disequilibrium between independent SNPs mapped in different chromosomes or linkage groups (Table S1 in Supplementary Material), following Waples (2006) as implemented in the software NeEstimator (Do et al., 2014). Linkage disequilibrium was measured as $r^2$, the squared correlation between pairs of loci (Hill and Robertson, 1968).

## 3. Results

### 3.1. Population structure

#### 3.1.1. Principal Component Analysis

The PCA plots using the two first principal components and computed with pool-seq and SNP array data are shown for both species in Fig. 2. Using pool-seq allele frequency data, the two first principal components explained about 14% and 25% of the total variance for seabream and seabass, respectively (Fig. 2a and c). Using SNP array genotype data, these percentages increased to 22% and 28% (Fig. 2b and d).

PCA separated wild from farmed seabream populations (with the exception of F-EGY and F-ISR that grouped with wild populations) (Fig. 2a and b). Also, wild populations of Atlantic origin (W-SPA-1, W-SPA-2 and W-FRA) grouped together and separated from most wild Mediterranean populations, particularly when using pool-seq data

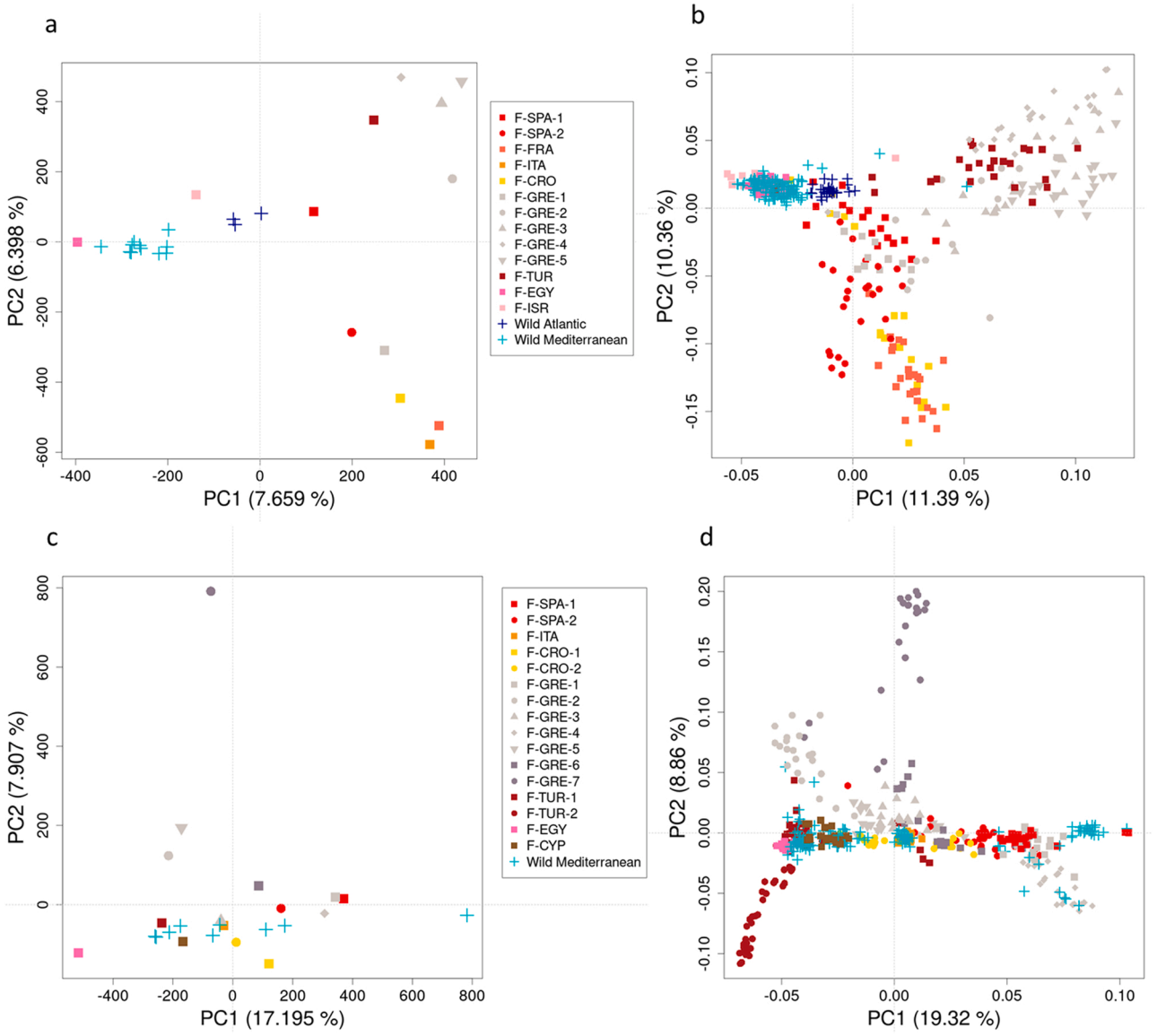


**Fig. 2.** Genetic structure of wild and farmed seabream (a and b) and seabass (c and d) populations obtained from Principal Component Analysis using pool-seq (a and c) and SNP array (b and d) data.

(Fig. 2a). Farmed populations also formed groups that loosely follow their hatchery location. However, seabass results from PCA using pool-seq or SNP array data indicated that there is no clear differentiation between wild and farmed populations (Fig. 2c and d). One farmed population (F-GRE-7) clearly differentiated from other farmed populations and one wild population (W-MOR) differentiate from other wild populations. In general, the dispersion within populations was high in both species (Fig. 2b and d).

### 3.1.2. *Distance-based clustering methods*

For both species, the most accurate hierarchical distance-based clustering methods were the single-linkage (lowest Gowerś distance) and the divisive (highest cophenetic correlation coefficient) methods. Both clustering methods led to equivalent results and only results from the single-linkage method, using pool-seq data, are presented (Fig. 3). These methods clearly differentiated wild and farmed populations of both species. For seabream, clustering analyses revealed that wild populations grouped together and separated from farmed populations that also grouped together (Fig. 3a). Similar results were observed for seabass (Fig. 3b) but there were some exceptions. In particular, two farmed populations (F-ITA and F-TUR-1) were grouped with wild populations and two wild populations (W-MOR and W-CYP) were grouped with farmed populations.

### 3.1.3. *Model-based clustering method*

When analyzing farmed and wild populations jointly, the admixture model results revealed that the number of ancestral populations (*K*) with the lowest cross-validation error was 7 for gilthead seabream and 15 for European seabass (Figs. S2a and b in Supplementary Material). Admixture results also showed a clear differentiation between wild and farmed populations (Fig. 4). In general, there was very little differentiation among wild seabream populations although a slight differentiation was observed between Atlantic and Mediterranean populations (Fig. 4a). However, farmed seabream populations were heterogeneous and showed different admixture proportions although, confirming the PCA results, some populations geographically close showed more similar patterns; i.e. the Spanish, French, Croatian and eastern Greek populations (Fig. 4b). It is worth noting that the F-EGY population showed admixture proportions similar to those found in wild populations. It is also relevant the observation of some farmed populations

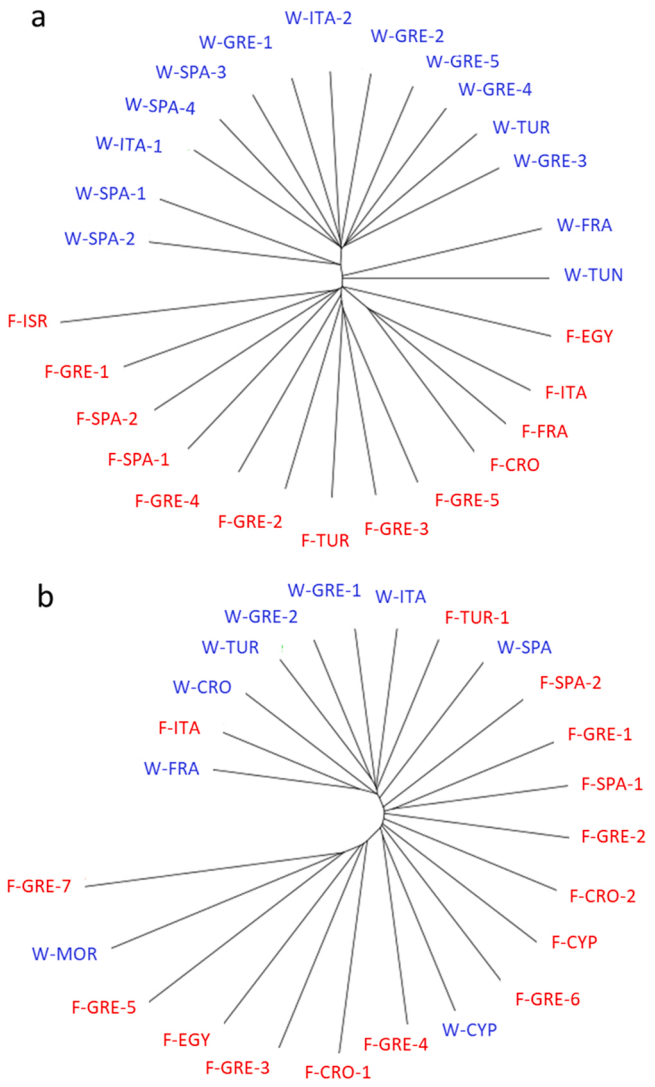


**Fig. 3.** Single-linkage clustering constructed from allele frequencies estimated from pool-seq data for seabream (a) and seabass (b) populations. Blue labels are for wild populations and red labels are for farmed populations.

(e.g. F-SPA-2) presented two clearly different groups, probably reflecting a very recent mixture of populations.

Within wild seabass populations, the Moroccan population (W-MOR) was very different from the rest (Fig. 4c). Also, the Cyprus population (W-CYP) had a different genetic makeup. For the rest of the wild seabass populations, a gradual change was observed in the proportions of different origins and this change was associated with the geographic locations as we move from West to East. It was also observed the presence of individuals within a particular population that clearly differed from others in the same population (e.g. in populations W-SPA and W-CRO). As with seabream, farmed populations were very heterogeneous and there was no direct relationship between the genetic structure and the location of the populations. It is interesting to note that also for this species, some farmed populations (F-ITA, F-TUR-1 and F-EGY) showed admixture proportions similar to those found in wild populations. In general, the Admixture results agreed well with the PCA and distance-based clustering results.

Admixture model results obtained when analyzing wild and farmed populations separately are shown in Fig. 5. In this case, the number of ancestral populations ($K$) with the lowest cross-validation error was one for wild seabream populations (Fig. S2c in Supplementary Material) indicating a complete lack of population structure. Using $K = 7$ (the number of ancestral populations with the lowest cross-validation error for farmed populations; Fig. S2e) showed again very little differentiation among seabream wild populations. However, contrary to what occurred when analyzing wild and farmed populations jointly, one population (W-TUN) appeared to be much more heterogeneous than the rest. Given the way that fish were sampled in this Tunisian population, it is possible that a subset of the fish had farm origin. It is also interesting to note that two wild populations (W-SPA-2 and W-GRE-2) contained individuals with a different makeup than the rest of individuals within the population.

Just as it happened when wild and farmed populations were analyzed jointly, wild seabass populations W-MOR and W-CYP were clearly different from the rest (Fig. 5c) both with $K = 2$ (the optimal number of clusters for the wild seabass analysis; Fig. S2d) and $K = 13$ (the optimal number of clusters for the farmed seabass analysis; Fig. S2f). Also, the Italian wild population (W-ITA) had a large number of individuals (about a third) with a genetic composition very different to that observed for wild populations. For both species, patterns for farmed populations were similar to those observed when analyzing them jointly with wild populations; i.e. a great heterogeneity was observed within and between populations.

#### 3.1.4. *Wright's fixation index $F_{ST}$*

Pairwise $F_{ST}$ coefficients (Tables 2 and 3) were generally low but indicated a higher differentiation between seabass than between seabream populations. Also, in general, the differentiation was higher among farmed than among wild populations in both fish species. For seabream, $F_{ST}$ ranged from 0.01 to 0.04 in wild populations and from 0.02 to 0.06 in farmed populations. For seabass, if we exclude the Moroccan population, $F_{ST}$ ranged from 0.01 to 0.05 in wild populations and from 0.02 to 0.09 in farmed populations. For seabream, the highest $F_{ST}$ values were observed for comparisons involving the Israeli population (F-ISR). For seabass, the highest $F_{ST}$ values were observed for comparisons involving the Moroccan population (W-MOR) which showed values up to 0.07 with other wild populations. This agrees with the results from the distance-based and model-based (Admixture) clustering analyses.

### 3.2. *Genetic diversity*

Genetic diversity, measured as $H_E$ obtained from the SNP array data within each population, is shown in Table 4. For both species, the average $H_E$ across wild populations (0.379 for seabream and 0.382 for seabass) was only slightly higher than the average $H_E$ across farmed populations (0.372 for seabream and 0.376 for seabass). Consequently, the average coancestry was slightly higher for farmed (0.628 for seabream and 0.624 for seabass) than for wild populations (0.621 for seabream and 0.618 for seabass). In all populations $H_O$ was higher than $H_E$.

In agreement with the results for $H_E$ and $H_O$, most inbreeding coefficients ($F_{L\&H}$) were below zero in both species (Fig. 6), reflecting that the observed homozygosity was lower than the expected. The $F_{L\&H}$ values observed for wild populations were closer to zero than those found for farmed populations although differences in $F_{L\&H}$ between both types of populations were small. The average $F_{L\&H}$ across seabream populations was −0.038 for wild and −0.048 for farmed populations. Equivalent figures for seabass were −0.014 and −0.068. The variation in $F_{L\&H}$ across individuals was clearly higher in farmed than in wild populations for both species (Fig. 6).

### 3.3. *Effective population size*

In general, estimates of $N_e$ were very high for seabream wild populations ($N_e > 1000$ fish) and relatively low for farmed populations (Table 4). However, there were some exceptions. In particular, estimates for wild seabream populations W-SPA-2, W-TUN and W-GRE-2 were relatively low ($N_e < 70$), and the estimate for the farmed Egyptian population (F-EGY) was high ($N_e > 400$). In seabass, estimates of $N_e$

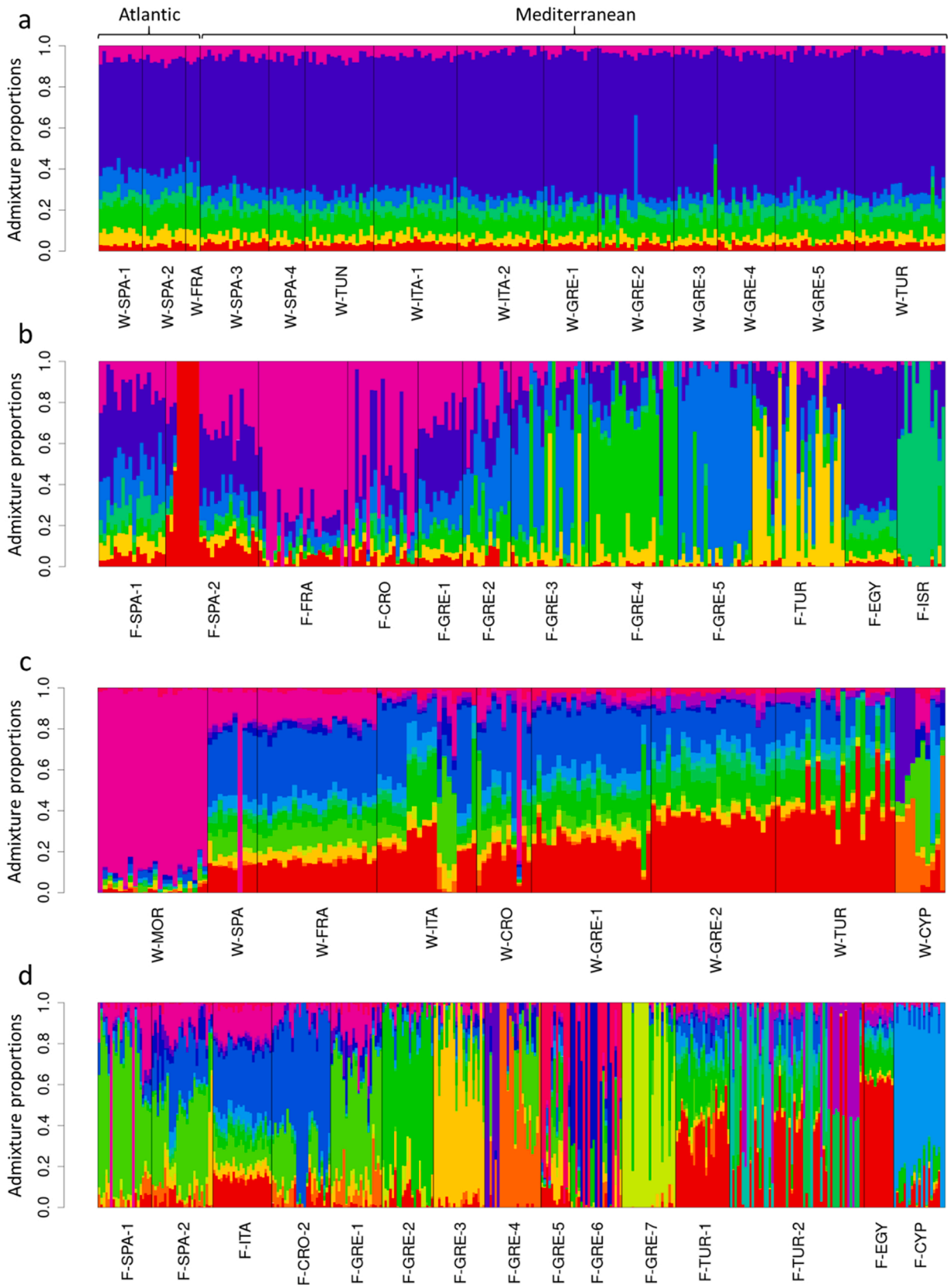


**Fig. 4.** Admixture plots showing the ancestry components of seabream (a and b) and seabass (c and d) populations computed with the SNP array data and analyzing wild and farmed populations jointly. Models assumed 7 (seabream) and 15 (seabass) ancestral components (*K*) which were those with the lowest cross-validation error. Population codes are represented on the x-axis. Populations are ordered according to their geographic location, from West to East. Each individual is represented by a thin vertical line partitioned into coloured segments whose lengths are proportional to the genetic contributions of the ancestral components to the genome of the individual. Seabream wild populations of Atlantic and Mediterranean origin are also indicated.

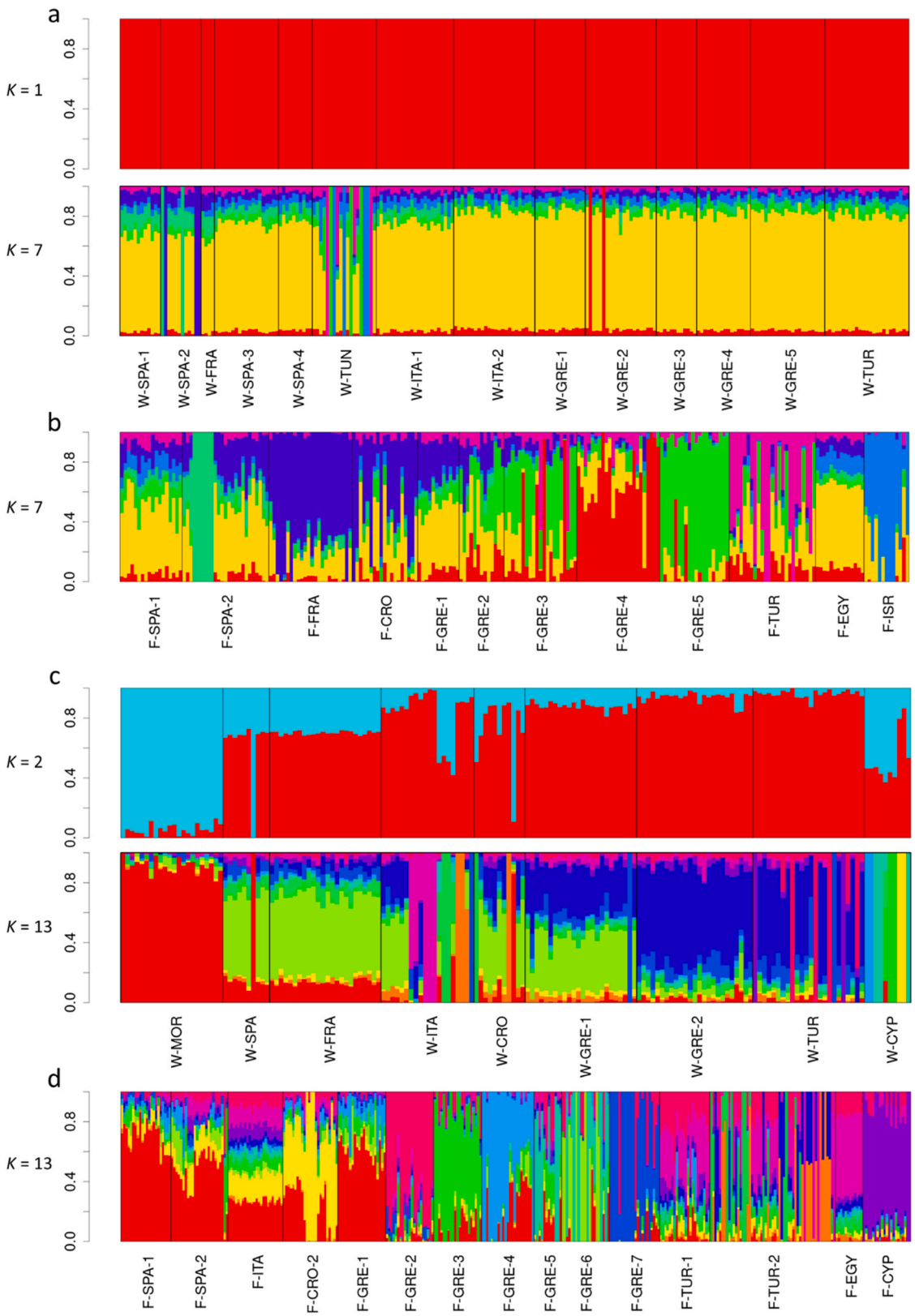


**Fig. 5.** Admixture plots showing the ancestry components of seabream (a and b) and seabass (c and d) populations computed with the SNP array data and analyzing wild and farmed populations separately. Models assumed 1 and 7 ancestral components (*K*) for seabream and 2 and 13 ancestral components for seabass. The optimal *K* value (that with the lowest cross-validation error) was 1 for wild seabream, 7 for farmed seabream, 2 for wild seabass and 13 for farmed seabass. Note that for wild scenarios, plots are also represented for the optimal *K* value obtained for the corresponding farmed populations. Population codes are represented on the x-axis. Populations are numbered in accordance to the geographic sampling location from West to East. Each individual is represented by a vertical line partitioned into coloured segments whose lengths are proportional to the genetic contributions of the ancestral components to the genome of the individual. Fig. 5a for $K = 1$ is included for consistency.

**Table 2**
Pairwise $F_{ST}$ between seabream populations. Colour codes indicate $F_{ST}$ values > 0.03 (pale yellow) and > 0.04 (intense yellow).

| | W-SPA-2 | W-FRA | W-SPA-3 | W-SPA-4 | W-TUN | W-ITA-1 | W-ITA-2 | W-GRE-1 | W-GRE-2 | W-GRE-3 | W-GRE-4 | W-GRE-5 | W-TUR | F-SPA-1 | F-SPA-2 | F-FRA | F-CRO | F-GRE-1 | F-GRE-2 | F-GRE-3 | F-GRE-4 | F-GRE-5 | F-TUR | F-EGY | F-ISR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **W-SPA-1** | 0.024 | 0.033 | 0.018 | 0.025 | 0.022 | 0.016 | 0.017 | 0.022 | 0.019 | 0.024 | 0.021 | 0.018 | 0.016 | 0.025 | 0.025 | 0.029 | 0.026 | 0.029 | 0.027 | 0.024 | 0.028 | 0.033 | 0.022 | 0.025 | 0.050 |
| **W-SPA-2** | | 0.036 | 0.020 | 0.028 | 0.025 | 0.018 | 0.019 | 0.025 | 0.021 | 0.026 | 0.023 | 0.020 | 0.018 | 0.027 | 0.027 | 0.031 | 0.028 | 0.031 | 0.029 | 0.026 | 0.030 | 0.036 | 0.024 | 0.027 | 0.052 |
| **W-FRA** | | | 0.025 | 0.041 | 0.027 | 0.021 | 0.021 | 0.031 | 0.024 | 0.037 | 0.030 | 0.023 | 0.020 | 0.030 | 0.025 | 0.028 | 0.029 | 0.040 | 0.036 | 0.027 | 0.027 | 0.033 | 0.024 | 0.034 | 0.055 |
| **W-SPA-3** | | | | 0.018 | 0.018 | 0.012 | 0.013 | 0.016 | 0.014 | 0.018 | 0.016 | 0.014 | 0.013 | 0.022 | 0.024 | 0.028 | 0.023 | 0.024 | 0.023 | 0.023 | 0.029 | 0.033 | 0.022 | 0.019 | 0.043 |
| **W-SPA-4** | | | | | 0.022 | 0.016 | 0.016 | 0.022 | 0.018 | 0.025 | 0.021 | 0.017 | 0.015 | 0.026 | 0.025 | 0.029 | 0.026 | 0.032 | 0.030 | 0.026 | 0.030 | 0.035 | 0.024 | 0.025 | 0.051 |
| **W-TUN** | | | | | | 0.017 | 0.018 | 0.021 | 0.019 | 0.022 | 0.021 | 0.018 | 0.017 | 0.027 | 0.028 | 0.032 | 0.027 | 0.029 | 0.027 | 0.028 | 0.033 | 0.037 | 0.027 | 0.024 | 0.048 |
| **W-ITA-1** | | | | | | | 0.012 | 0.014 | 0.013 | 0.016 | 0.014 | 0.013 | 0.012 | 0.021 | 0.023 | 0.028 | 0.022 | 0.022 | 0.021 | 0.022 | 0.028 | 0.032 | 0.022 | 0.018 | 0.040 |
| **W-ITA-2** | | | | | | | | 0.013 | 0.012 | 0.015 | 0.014 | 0.012 | 0.012 | 0.022 | 0.025 | 0.029 | 0.023 | 0.023 | 0.022 | 0.024 | 0.029 | 0.033 | 0.023 | 0.017 | 0.040 |
| **W-GRE-1** | | | | | | | | | 0.015 | 0.020 | 0.018 | 0.015 | 0.014 | 0.025 | 0.026 | 0.031 | 0.026 | 0.029 | 0.027 | 0.027 | 0.031 | 0.036 | 0.025 | 0.022 | 0.047 |
| **W-GRE-2** | | | | | | | | | | 0.017 | 0.016 | 0.014 | 0.013 | 0.023 | 0.025 | 0.030 | 0.024 | 0.025 | 0.023 | 0.024 | 0.030 | 0.033 | 0.024 | 0.019 | 0.042 |
| **W-GRE-3** | | | | | | | | | | | 0.018 | 0.015 | 0.014 | 0.026 | 0.026 | 0.030 | 0.027 | 0.031 | 0.028 | 0.026 | 0.029 | 0.035 | 0.024 | 0.023 | 0.049 |
| **W-GRE-4** | | | | | | | | | | | | 0.014 | 0.012 | 0.025 | 0.026 | 0.030 | 0.025 | 0.028 | 0.026 | 0.026 | 0.031 | 0.035 | 0.024 | 0.020 | 0.046 |
| **W-GRE-5** | | | | | | | | | | | | | 0.011 | 0.022 | 0.024 | 0.029 | 0.023 | 0.024 | 0.022 | 0.024 | 0.029 | 0.033 | 0.023 | 0.017 | 0.040 |
| **W-TUR** | | | | | | | | | | | | | | 0.021 | 0.023 | 0.028 | 0.022 | 0.022 | 0.021 | 0.023 | 0.028 | 0.032 | 0.022 | 0.015 | 0.038 |
| **F-SPA-1** | | | | | | | | | | | | | | | 0.028 | 0.032 | 0.027 | 0.030 | 0.027 | 0.027 | 0.034 | 0.035 | 0.026 | 0.028 | 0.047 |
| **F-SPA-2** | | | | | | | | | | | | | | | | 0.028 | 0.025 | 0.027 | 0.025 | 0.030 | 0.037 | 0.038 | 0.030 | 0.028 | 0.049 |
| **F-FRA** | | | | | | | | | | | | | | | | | 0.018 | 0.027 | 0.027 | 0.032 | 0.041 | 0.042 | 0.034 | 0.032 | 0.053 |
| **F-CRO** | | | | | | | | | | | | | | | | | | 0.025 | 0.024 | 0.028 | 0.035 | 0.036 | 0.028 | 0.028 | 0.051 |
| **F-GRE-1** | | | | | | | | | | | | | | | | | | | 0.029 | 0.028 | 0.033 | 0.037 | 0.027 | 0.031 | 0.056 |
| **F-GRE-2** | | | | | | | | | | | | | | | | | | | | 0.019 | 0.027 | 0.027 | 0.022 | 0.029 | 0.053 |
| **F-GRE-3** | | | | | | | | | | | | | | | | | | | | | 0.025 | 0.023 | 0.020 | 0.029 | 0.050 |
| **F-GRE-4** | | | | | | | | | | | | | | | | | | | | | | 0.037 | 0.027 | 0.033 | 0.053 |
| **F-GRE-5** | | | | | | | | | | | | | | | | | | | | | | | 0.029 | 0.038 | 0.059 |
| **F-TUR** | | | | | | | | | | | | | | | | | | | | | | | | 0.027 | 0.047 |
| **F-EGY** | | | | | | | | | | | | | | | | | | | | | | | | | 0.048 |

were also very high ($N_e$ > 1000) for about half of the wild populations and relatively low for farmed populations (Table 4). Wild populations W-SPA, W-ITA, W-CRO and W-TUR had a relatively low $N_e$ (< 100) and population W-CYP showed an extremely low $N_e$. The estimate of $N_e$ for two seabass farmed populations (F-ITA and F-EGY) was unexpectedly high ($N_e$ > 1000). It is important to note that five out of the 12 farmed seabream populations and 12 out of the 15 farmed seabass populations had a $N_e$ lower than 50 fish.

## 4. Discussion

This study has shown that the recently developed MedFish SNP array is a very useful tool for determining the genetic structure across gilthead seabream and European seabass populations and the diversity that exists within populations. The population genetic analysis has been carried out with a very high number of SNP markers and samples and thus, our results provide new high-resolution population genetic data to inform future breeding and stock management in these key European aquaculture species.

We applied a range of approaches to investigate population structure in seabream and seabass, using both SNP array and pool-seq data. The two first principal components explained a substantial percentage of the total variance in both species (22% and 28% for seabream and seabass, respectively, using SNP array genotype data), and the PCA results showed a differentiation between wild and farmed populations for seabream but not for seabass. However, hierarchical clustering and admixture model results showed more resolution than PCA and revealed a clear differentiation between farmed and wild populations in both species.

A slight differentiation was observed between wild Atlantic and Mediterranean seabream populations and between wild Western and Eastern Mediterranean seabass populations but, in general, the degree of differentiation among wild populations was low, particularly in seabream. This indicates that considerable gene flow exists. These results, in addition to the general high $N_e$ found, also indicate that there are no specific requirements for conservation of wild populations (i.e. there are no specific target populations to be conserved).

On the contrary, farmed populations of both fish species were quite heterogeneous, which may indicate that they were generated from several origins, possibly including some fish from the Atlantic region. In fact, there is anecdotal evidence that broodstock was originally made up using fish of different geographic origin across and within hatcheries. Our results also showed a high degree of differentiation, pointing to different sources of fish for different companies. This was observed when analyzing them both jointly or separately from wild populations. In any case, Peñaloza et al. (2021) showed that there is some haplotype sharing among Mediterranean farmed populations in both species. Also, they detected two groups of seabream farms between which haplotype sharing was reduced. One group included farms from France, Spain, Croatia and Greece and another group included farms exclusively based in Greece. This observation agrees with our PCA and admixture results (Figs. 2a, b and 4b). A few farmed populations of both species showed genetic compositions similar to those found in wild populations which may indicate that their selective breeding programs are still in their beginnings or that broodstock has been recently renewed using wild fish as a means of mitigating inbreeding.

Given the clear genetic differentiation between wild and farmed populations of seabass and seabream, precautions must be taken to avoid escapes that could have undesirable genetic effects in native populations. The admixture results showed that, within a particular population, some wild individuals have a genetic composition very different from the rest of individuals in the population. These

**Table 3**
Pairwise $F_{ST}$ between seabass populations. Colour codes indicate $F_{ST}$ values > 0.03 (pale yellow), > 0.04 (intense yellow) and > 0.07 (orange).

| | W-SPA | W-FRA | W-ITA | W-CRO | W-GRE-1 | W-GRE-2 | W-TUR | W-CYP | F-SPA-1 | F-SPA-2 | F-ITA | F-CRO-2 | F-GRE-1 | F-GRE-2 | F-GRE-3 | F-GRE-4 | F-GRE-5 | F-GRE-6 | F-GRE-7 | F-TUR-1 | F-TUR-2 | F-EGY | F-CYP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **W-MOR** | 0.042 | 0.047 | 0.061 | 0.049 | 0.064 | 0.068 | 0.072 | 0.057 | 0.039 | 0.045 | 0.047 | 0.059 | 0.042 | 0.084 | 0.069 | 0.071 | 0.068 | 0.062 | 0.087 | 0.069 | 0.062 | 0.084 | 0.087 |
| **W-SPA** | | 0.016 | 0.024 | 0.026 | 0.021 | 0.024 | 0.027 | 0.047 | 0.030 | 0.022 | 0.016 | 0.026 | 0.030 | 0.040 | 0.035 | 0.049 | 0.044 | 0.036 | 0.055 | 0.028 | 0.020 | 0.051 | 0.044 |
| **W-FRA** | | | 0.017 | 0.015 | 0.015 | 0.019 | 0.022 | 0.033 | 0.029 | 0.020 | 0.011 | 0.023 | 0.029 | 0.035 | 0.031 | 0.049 | 0.031 | 0.033 | 0.053 | 0.022 | 0.023 | 0.036 | 0.039 |
| **W-ITA** | | | | 0.018 | 0.016 | 0.017 | 0.019 | 0.039 | 0.036 | 0.025 | 0.018 | 0.025 | 0.034 | 0.032 | 0.032 | 0.054 | 0.034 | 0.035 | 0.055 | 0.020 | 0.020 | 0.036 | 0.039 |
| **W-CRO** | | | | | 0.017 | 0.019 | 0.022 | 0.045 | 0.032 | 0.022 | 0.015 | 0.024 | 0.031 | 0.036 | 0.033 | 0.049 | 0.040 | 0.035 | 0.053 | 0.023 | 0.018 | 0.045 | 0.040 |
| **W-GRE-1** | | | | | | 0.014 | 0.016 | 0.038 | 0.038 | 0.027 | 0.016 | 0.025 | 0.037 | 0.030 | 0.032 | 0.056 | 0.032 | 0.036 | 0.053 | 0.017 | 0.019 | 0.031 | 0.036 |
| **W-GRE-2** | | | | | | | 0.012 | 0.040 | 0.042 | 0.030 | 0.019 | 0.028 | 0.041 | 0.030 | 0.034 | 0.060 | 0.032 | 0.038 | 0.055 | 0.014 | 0.015 | 0.026 | 0.037 |
| **W-TUR** | | | | | | | | 0.043 | 0.046 | 0.033 | 0.022 | 0.031 | 0.045 | 0.030 | 0.036 | 0.063 | 0.035 | 0.041 | 0.057 | 0.015 | 0.016 | 0.026 | 0.039 |
| **W-CYP** | | | | | | | | | 0.036 | 0.030 | 0.033 | 0.040 | 0.031 | 0.059 | 0.047 | 0.029 | 0.060 | 0.045 | 0.069 | 0.042 | 0.031 | 0.072 | 0.059 |
| **F-SPA-1** | | | | | | | | | | 0.024 | 0.029 | 0.037 | 0.021 | 0.059 | 0.044 | 0.047 | 0.047 | 0.040 | 0.065 | 0.043 | 0.042 | 0.059 | 0.061 |
| **F-SPA-2** | | | | | | | | | | | 0.020 | 0.029 | 0.022 | 0.046 | 0.034 | 0.043 | 0.036 | 0.033 | 0.056 | 0.032 | 0.034 | 0.046 | 0.049 |
| **F-ITA** | | | | | | | | | | | | 0.023 | 0.029 | 0.036 | 0.031 | 0.049 | 0.032 | 0.033 | 0.053 | 0.022 | 0.024 | 0.036 | 0.039 |
| **F-CRO-2** | | | | | | | | | | | | | 0.036 | 0.046 | 0.040 | 0.056 | 0.040 | 0.042 | 0.063 | 0.031 | 0.031 | 0.045 | 0.049 |
| **F-GRE-1** | | | | | | | | | | | | | | 0.058 | 0.041 | 0.038 | 0.045 | 0.036 | 0.064 | 0.042 | 0.041 | 0.059 | 0.060 |
| **F-GRE-2** | | | | | | | | | | | | | | | 0.045 | 0.078 | 0.046 | 0.051 | 0.052 | 0.029 | 0.029 | 0.046 | 0.053 |
| **F-GRE-3** | | | | | | | | | | | | | | | | 0.063 | 0.045 | 0.046 | 0.065 | 0.035 | 0.034 | 0.051 | 0.054 |
| **F-GRE-4** | | | | | | | | | | | | | | | | | 0.064 | 0.056 | 0.085 | 0.059 | 0.057 | 0.077 | 0.080 |
| **F-GRE-5** | | | | | | | | | | | | | | | | | | 0.037 | 0.064 | 0.035 | 0.026 | 0.058 | 0.054 |
| **F-GRE-6** | | | | | | | | | | | | | | | | | | | 0.066 | 0.039 | 0.037 | 0.055 | 0.059 |
| **F-GRE-7** | | | | | | | | | | | | | | | | | | | | 0.055 | 0.052 | 0.072 | 0.078 |
| **F-TUR-1** | | | | | | | | | | | | | | | | | | | | | 0.018 | 0.028 | 0.039 |
| **F-TUR-2** | | | | | | | | | | | | | | | | | | | | | | 0.020 | 0.036 |
| **F-EGY** | | | | | | | | | | | | | | | | | | | | | | | 0.054 |

**Table 4**
Estimates of genetic diversity measured as expected ($H_E$) and observed ($H_O$) heterozygosity, and effective population size ($N_e$) for seabream and seabass populations.

| Seabream | | | | Seabass | | | |
|---|---|---|---|---|---|---|---|
| Population | $H_E$ | $H_O$ | $N_e$ | Pop. ID | $H_E$ | $H_O$ | $N_e$ |
| W-SPA-1 | 0.378 | 0.397 | > 1000.0 | W-MOR | 0.327 | 0.340 | >1000.0 |
| W-SPA-2 | 0.376 | 0.402 | 59.6 | W-SPA | 0.397 | 0.400 | 87.8 |
| W-FRA | 0.342 | 0.441 | > 1000.0 | W-FRA | 0.403 | 0.411 | >1000.0 |
| W-SPA-3 | 0.386 | 0.398 | > 1000.0 | W-ITA | 0.393 | 0.392 | 37.7 |
| W-SPA-4 | 0.376 | 0.403 | 936.3 | W-CRO | 0.398 | 0.398 | 82.4 |
| W-TUN | 0.378 | 0.397 | 36.0 | W-GRE-1 | 0.386 | 0.397 | >1000.0 |
| W-ITA-1 | 0.387 | 0.396 | > 1000.0 | W-GRE-2 | 0.381 | 0.392 | >1000.0 |
| W-ITA-2 | 0.385 | 0.394 | > 1000.0 | W-TUR | 0.378 | 0.396 | 73.7 |
| W-GRE-1 | 0.380 | 0.395 | > 1000.0 | W-CYP | 0.373 | 0.390 | 4.8 |
| W-GRE-2 | 0.385 | 0.395 | 63.6 | | | | |
| W-GRE-3 | 0.378 | 0.399 | > 1000.0 | | | | |
| W-GRE-4 | 0.381 | 0.394 | > 1000.0 | | | | |
| W-GRE-5 | 0.385 | 0.394 | > 1000.0 | | | | |
| W-TUR | 0.387 | 0.396 | > 1000.0 | | | | |
| | | | | | | | |
| F-SPA-1 | 0.377 | 0.397 | 75.1 | F-SPA-1 | 0.389 | 0.394 | 16.3 |
| F-SPA-2 | 0.376 | 0.392 | 21.5 | F-SPA-2 | 0.402 | 0.418 | 22.1 |
| F-FRA | 0.375 | 0.395 | 107.1 | F-ITA | 0.402 | 0.411 | >1000.0 |
| F-ITA | – | – | – | F-CRO-1 | – | – | – |
| F-CRO | 0.380 | 0.394 | 116.9 | F-CRO-2 | 0.391 | 0.415 | 19.1 |
| F-GRE-1 | 0.372 | 0.402 | 94.4 | F-GRE-1 | 0.391 | 0.414 | 49.1 |
| F-GRE-2 | 0.377 | 0.402 | 99.3 | F-GRE-2 | 0.362 | 0.397 | 30.0 |
| F-GRE-3 | 0.377 | 0.387 | 51.3 | F-GRE-3 | 0.381 | 0.408 | 35.1 |
| F-GRE-4 | 0.368 | 0.388 | 43.6 | F-GRE-4 | 0.363 | 0.381 | 9.3 |
| F-GRE-5 | 0.364 | 0.395 | 49.0 | F-GRE-5 | 0.375 | 0.435 | 8.9 |
| F-TUR | 0.377 | 0.392 | 49.2 | F-GRE-6 | 0.384 | 0.428 | 8.9 |
| F-EGY | 0.376 | 0.395 | 455.0 | F-GRE-7 | 0.356 | 0.454 | 7.0 |
| F-ISR | 0.342 | 0.412 | 15.4 | F-TUR-1 | 0.381 | 0.403 | 82.4 |
| | | | | F-TUR-2 | 0.366 | 0.384 | 17.8 |
| | | | | F-EGY | 0.342 | 0.385 | > 1000.0 |
| | | | | F-CYP | 0.356 | 0.400 | 13.1 |

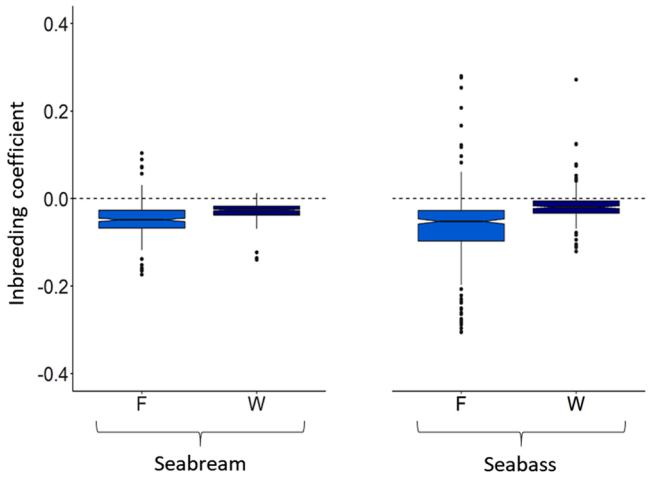


**Fig. 6.** Inbreeding coefficient distributions for wild (W) and farmed (F) seabream and seabass populations.

individuals may represent escapees from local farms (e.g. Brown et al., 2015). Also, there have been initiatives in the past to increase the natural supply of fish by means of restocking (Santos et al., 2006). Given the differentiation between wild and farmed populations, it is recommended that only wild broodstock should be used if fish restocking were to be carried out.

Previous studies that used a much more limited number of markers and samples have also found a differentiation between wild Atlantic and Mediterranean seabream populations (Alarcón et al., 2004; De Innocentiis et al., 2004; Franchini et al., 2012). However, their results in terms of the differentiation among wild Mediterranean populations led to different conclusions. Some studies have suggested a lack of genetic structure when using different types of genetic markers (Alarcón et al., 2004; Loukovitis et al., 2012; Žužul et al., 2019) while other studies have suggested a weak differentiation (De Innocentiis et al., 2004; Šegvić-Bubić et al., 2011). Results have been inconclusive in other studies (e.g. Franchini et al., 2012). For wild seabream, the study coming closest to ours was that of Maroso et al. (2021) which was also based on a large number of SNPs and samples (1159 SNPs and 23 samples from different locations). They found a differentiation between wild Atlantic and Mediterranean and also among Mediterranean populations. In particular, their analyses suggested a weak subdivision of seabream into four major genetic clusters (Atlantic, West Mediterranean, Ionian/Adriatic seas and the Aegean Sea). While our results confirm the differentiation between Atlantic and Mediterranean populations, they do not support a differentiation among Mediterranean populations.

Most previous studies on wild seabass have showed evidence of some genetic differentiation among Mediterranean populations (Allegrucci et al., 1997; Caccone et al., 1997; García De León et al., 1997; Naciri et al., 1999; Bahri-Sfar et al., 2000; Quéré et al., 2012; Souche et al., 2015; Bodur et al., 2017). Although in general, this genetic differentiation was low, some studies (Bahri-Sfar et al., 2000; Souche et al., 2015) have described differentiated clusters rather than the gradual change in the genetic composition of populations when moving from West to East found here using a very large number of SNP markers (see also Vandeputte et al., 2019).

In this study, all samples of wild seabass came from the Mediterranean Sea but two populations (W-CYP and W-MOR) appeared genetically very different to the rest of wild populations when using the different approaches to identify genetic structure (Figs. 2c, d, 3b, 4c and 5c and Table 3). In the case of the Cypriot population (W-CYP), two locations were included in the samples and these locations were very close to fish farms. Thus, it is expected that they are a mix of farmed and wild fish (see also Brown et al., 2015). Although no geographically speaking pure Atlantic wild seabass populations were sampled, the Moroccan population (W-MOR) was located in the Alboran Sea where the Atlantic intersects the Mediterranean lineages. This, together with the fact that it seems that the introgression is asymmetric mostly from the Atlantic to the Mediterranean lineage (Tine et al., 2014; Duranton et al., 2018; Vandeputte et al., 2019) can lead us to consider the W-MOC population as an Atlantic population.

Our results comparing wild and farmed populations agree with previous results showing a genetic differentiation both in seabream (Alarcón et al., 2004; Šegvić-Bubić et al., 2011; Loukovitis et al., 2012; Žužul et al., 2019) and in seabass (Brown et al., 2015; Polovina et al., 2020). The knowledge acquired will be very useful when base populations need to be created in an optimal way to start new breeding programs and also to detect and correct introgression arising from escapes from farmed populations.

In all populations, $H_O$ was higher than $H_E$. As a consequence, the inbreeding coefficient $F_{L\&H}$ was negative. Note that $F_{L\&H}$ depends on the allele frequencies in the base population (Villanueva et al., 2021) that in our case is the current population. In this situation, $F_{L\&H}$ would be expected to be zero. However, it is not exactly zero but takes a negative value close to zero due to sampling of gametes. In a finite population, it is expected that there will be random differences between the allele frequencies between both sexes and this generates an excess of observed heterozygotes ($H_O$) with respect to those expected with Hardy-Weinberg equilibrium ($H_E$) (Caballero, 2020). In any case, the heterozygosities (both observed and expected) given here for the different populations are likely to be somewhat overestimated due to the ascertainment bias resulting from the selection of SNPs to be included in the array. It is common practice to develop SNP arrays with an underrepresentation of SNPs with extreme allele frequencies. (e.g. Geibel et al., 2021).

The average expected heterozygosity ($H_E$) across wild populations was only slightly higher than that across farmed populations, for both species. There are two reasons why $H_E$ was relatively high in farmed populations: i) if the time in which farmed populations were established was relatively recent, the founder effect would led to a faster loss of rare alleles than in heterozygosity because rare allele have little effect on $H_E$ (i.e. many alleles can be lost without much reduction in $H_E$) in the short term (e.g. Luikart and Cornuet, 1998); and ii) the high heterogeneity of farmed populations observed here suggests that they were generated from many different origins which leads to a high $H_E$. However, as expected, in general, estimates of $N_e$ (the parameter related to the rate at which genetic variability is lost) were much higher for wild than for farmed populations in both species (Table 4). This general result was also observed by Žužul et al. (2019) in seabream and by Šegvić-Bubić et al. (2017) in seabass, using microsatellite markers. In our study, most wild seabream populations showed estimates higher than 1000 fish and most farmed populations showed estimates lower than 100 fish. Exceptions included three wild populations (W-SPA-2, W-TUN and W-GRE-2) with relatively low $N_e$. Given that there was no clear differentiation among wild populations, it is likely that these low estimates are due to the fact that some of the fish included in these samples are escapees as most samples were collected by local fishermen and they often fish close to the sea cages. In fact, when analyzing wild and farmed populations separately, results show that for these three populations, there are individuals with a genetic composition different from the rest of individuals within the particular population (Fig. 5a with $K = 7$). It is known that for instance, in Tunisia, fish farms import juveniles from different hatcheries in Greece, France, Spain and Italy, and this increases the risk that escapees come from divergent genetic backgrounds. The low $N_e$ found for some wild populations could be also due to a reduced number of families included in the samples.

On the other hand, the unexpected high $N_e$ estimated for the Egyptian seabream farmed population (F-EGY) agrees with the results of the Admixture analysis in which this population showed admixture proportions very similar to those found in wild populations. This may indicate that their selective breeding program is recent and that fish in the farm are still closely related to wild fish.

More exceptions to the general rule of high $N_e$ in wild and low $N_e$ in farmed populations were found for seabass. About half the wild populations had high $N_e$ estimates ($N_e > 1000$) as expected. However, estimates for populations W-SPA, W-ITA, W-CRO and W-TUR were relatively low ($N_e < 100$) and that for W-CYP was extremely low ($N_e \sim 5$). These low estimates can be also due to the way sampling was done (i.e. escapees included in the samples) or to a reduced number of families sampled. As described above, the Cypriot population (W-CYP) is expected to be composed by a mix of farmed and wild fish. The very high $N_e$ estimate ($> 1000$) for two farmed seabass populations (F-ITA and F-EGY) could indicate that selection is still in its beginnings or that broodstock is renewed using wild fish, a common practice in Mediterranean aquaculture, particularly in seabass.

Estimates of $N_e$ below 50 fish were found in many farmed populations (5 seabream and 12 seabass populations). These estimates are lower than the critical size of 50 needed to avoid inbreeding depression and retain fitness in the short-term (FAO, 2013). Estimates of this magnitude have been found in many other farmed fish populations of different species (Eknath and Doyle, 1990; Su et al., 1996; Pante et al., 2001; Gallardo et al., 2004; Brown et al., 2005; Borrell et al., 2011; Yáñez et al., 2014; Šegvić-Bubić et al., 2017; Barría et al., 2018; Barría et al., 2019; D'Ambrosio et al., 2019; Saura et al., 2021). This highlights the need of applying selection and mating approaches designed to control inbreeding and loss of genetic variability in order to ensure the sustainability of the breeding programs.

## 5. Conclusions

The MedFish SNP array recently developed is a powerful tool for discriminating between wild and farmed populations of gilthead seabream and European seabass and for determining the genetic structure among populations and the genetic diversity that exists within populations. Within a species, farmed populations present marked differences among them whereas wild populations show low levels of differentiation, particularly in seabream. Nonetheless, there was a slight differentiation between Atlantic and Mediterranean seabream populations and a gradual change in seabass in the proportions of different origins that was associated with the geographic locations as we move from West to East. There was a clear differentiation between wild and farmed populations of both species which indicates that care must be taken to avoid escapes that could have undesirable genetic effects in native populations. The low estimates of effective population size found in several farmed populations highlight the need to apply selection and mating approaches designed to control inbreeding and loss of genetic variability in order to ensure the sustainability of the breeding programs.

## Funding

This work was supported by the European Commission Horizon 2020 (H2020) Framework Programme through grant agreement no 727315 MedAID project (Mediterranean Aquaculture Integrated Development) and by MCIN/ AEI /10.13039/501100011033 (Project PID2020-114426GB-C2).

## CRediT authorship contribution statement

**Beatriz Villanueva:** Conceptualization, Writing – original draft, Visualization, Supervision, Project administration, Funding acquisition. **Almudena Fernández:** Formal analysis, Data curation. **Ramón Peiró-Pastor:** Formal analysis. **Carolina Peñaloza:** Investigation, Resources, Writing – review & editing. **Ross D. Houston:** Writing – review & editing, Supervision, Funding acquisition. **Anna K. Sonesson:** Writing – review & editing, Project administration, Funding acquisition. **Costas S. Tsigenopoulos:** Resources, Writing – review & editing. **Luca Bargelloni:** Resources, Writing – review & editing. **Kutsal Gamsız:** Resource. **Bilge Karahan:** Resources, Writing – review & editing. **Emel Ö. Gökçek:** Resources, Writing – review & editing. **Jesús Fernández:** Conceptualization, Formal analysis, Writing – review & editing, Visualization. **María Saura:** Conceptualization, Formal analysis, Writing – review & editing, Visualization.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We are grateful to the coordinators (A. López-Francos, B. Basurco, D. Furones, M. Izquierdo and K. Moutou) and partners of MedAID and PerformFish (H2020 grant agreement No 727610) as well as collaborating institutions and companies that contributed with the samples. We also thank the 'Centro de Supercomputación de Galicia' (Cesga) for their computing support.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.aqrep.2022.101145.

## References

Alarcón, J.A., Magoulas, A., Georgakopoulos, T., Zouros, E., Alvarez, M.C., 2004. Genetic comparison of wild and cultivated European populations of the gilthead sea bream (*Sparus aurata*). Aquaculture 230, 65–80. https://doi.org/10.1016/S0044-8486(03)00434-4.

Alexander, D.H., Lange, K., 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. BMC Bioinformatics 12, 246. https://doi.org/10.1186/1471-2105-12-246.

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19, 1655–1664. https://doi.org/10.1101/gr.094052.109.

Allegrucci, G., Fortunato, C., Sbordoni, V., 1997. Genetic structure and allozyme variation of sea bass (*Dicentrarchus labrax* and *D. punctatus*) in the Mediterranean Sea. Mar. Biol. 128, 347–358. https://doi.org/10.1007/S002270050100.

Bahri-Sfar, L., Lemaire, C., Ben Hassine, O.K., Bonhomme, F., 2000. Fragmentation of sea bass populations in the Western and Eastern Mediterranean as revealed by microsatellite polymorphism. Proc. Royal Soc. B 267, 929–935. https://doi.org/10.1098/rspb.2000.1092.

Barría, A., López, M.E., Yoshida, G., Carvalheiro, R., Lhorente, J.P., Yáñez, J.M., 2018. Population genomic structure and genome-wide linkage disequilibrium in farmed Atlantic salmon (*Salmo salar L*) using dense SNP genotypes. Front. Genet. 9, 649. https://doi.org/10.3389/fgene.2018.00649.

Barría, A., Christensen, K.A., Yoshida, G., Jedlicki, A., Leong, J.S., et al., 2019. Whole genome linkage disequilibrium and effective population size in a coho salmon (*Oncorhynchus kisutch*) breeding population using a high-density SNP array. Front. Genet. 10, 498. https://doi.org/10.3389/fgene.2019.00498.

Bodur, T., Tsigenopoulos, C., Cagatay, I.T., 2017. Genetic structure of wild European sea bass (Dicentrarchus labrax L, 1758) populations in Aegean and Levantine Sea using microsatellite markers. Turkish. J. Fish. Aquat. Sci. 1, 1–11. https://doi.org/10.4194/1303-2712-v17_1_02.

Borrell, V., Gallego, V., García-Fernández, C., Mazzeo, I., Pérez, L., et al., 2011. Assessment of parental contributions to fast- and slow-growing progenies in the sea bream *Sparus aurata* L using a new multiplex PCR. Aquaculture 314, 58–65. https://doi.org/10.1016/j.aquaculture.2011.01.028.

Brock, G., Pihur, V., Datta, S., Datta, S., 2008. clValid: An R package for cluster validation. J. Stat. Softw. 25. ⟨https://www.jstatsoft.org/v25/i04⟩.

Brown, C., Woolliams, J.A., McAndrew, B.J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. Aquaculture 247, 219–225. https://doi.org/10.1016/j.aquaculture.2005.02.002.

Brown, C., Miltiadou, D., Tsigenopoulos, C.S., 2015. Prevalence and survival of escaped European seabass *Dicentrarchus labrax* in Cyprus identified using genetic markers. Aquacult. Environ. Interact 7, 49–59. https://doi.org/10.3354/aei00135.

Caballero, A., 2020. Quantitative Genetics. Cambridge University Press, Cambridge, UK.

Caccone, A., Allegrucci, G., Fortunato, C., Sbordoni, V., 1997. Genetic differentiation within the European sea bass (*D. labrax*) as revealed by RAPD-PCR assays. J. Hered 88, 316–324. https://doi.org/10.1093/oxfordjournals.jhered.a023109.

Coscia, I., Vogiatzi, E., Kotoulas, G., Tsigenopoulos, C.S., Mariani, S., 2012. Exploring neutral and adaptive processes in expanding populations of gilthead sea bream, Sparus aurata L., in the North-East Atlantic. Heredity 108, 537–546. https://doi.org/10.1038/hdy.2011.120.

D'Ambrosio, J., Phocas, F., Haffray, P., Bestin, A., Brard-Fudulea, S., et al., 2019. Genome-wide estimates of genetic diversity, inbreeding and effective size of experimental and commercial rainbow trout lines undergoing selective breeding. Genet. Sel. Evol. 51, 26. https://doi.org/10.1186/s12711-019-0468-4.

De Innocentiis, S., Lesti, A., Livi, S., Rossi, A.R., Crosetti, D., Sola, L., 2004. Microsatellite markers reveal population structure in gilthead sea bream *Sparus auratus* from the Atlantic Ocean and Mediterranean Sea. Fish. Sci. 70, 852–859. https://doi.org/10.1111/j.1444-2906.2004.00879.x.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol. Ecol. Resour. 14, 209–214. https://doi.org/10.1111/1755-0998.12157.

Duranton, M., Allal, F., Fraïsse, C., Bierne, N., Bonhomme, F., Gagnaire, P.A., 2018. The origin and remolding of genomic islands of differentiation in the European sea bass. Nat. Commun. 9, 2518. https://doi.org/10.1038/s41467-018-04963-6.

Eknath, A., Doyle, R.W., 1990. Effective population size and rate of inbreeding in aquaculture of Indian major carps. Aquaculture 85, 293–305. https://doi.org/10.1016/0044-8486(90)90028-L.

Eurostat, 2019. Agriculture, Forestry and Fishery Statistics, two thousand nineteenth ed. Statistical books. https://doi.org/10.2785/798761.

FAO, 2013. In vivo conservation of animal genetic resources. FAO Animal Production and Health Guidelines. No. 14. Rome. https://www.fao.org/3/i3327e/i3327e.pdf.

FAO, 2018. The State of World Fisheries and Aquaculture 2018 - Meeting the sustainable development goals. Rome. Licence: CC BY-NC-SA 3.0 IGO. ⟨https://www.fao.org/3/i9540en/i9540en.pdf⟩.

Fernández, J., Toro, M.A., Sonesson, A.K., Villanueva, B., 2014. Optimizing the creation of base populations for aquaculture breeding programs using phenotypic and genomic data and its consequences on genetic progress. Front. Genet. 5, 414. https://doi.org/10.3389/fgene.2014.00414.

Franchini, P., Sola, L., Crosetti, D., Milana, V., Rossi, A.R., 2012. Low levels of population genetic structure in the gilthead sea bream, Sparus aurata, along the coast of Italy. ICES J. Mar. Sci. 69, 41–50. https://doi.org/10.1093/icesjms/fsr175.

Gallardo, J.A., García, X., Lhorente, J.P., Neira, R., 2004. Inbreeding and inbreeding depression of female reproductive traits in two populations of Coho salmon selected using BLUP predictors of breeding values. Aquaculture 234, 111–122. https://doi.org/10.1016/j.aquaculture.2004.01.009.

García De León, F.J., Chikhi, L., Bonhomme, F., 1997. Microsatellite polymorphism and population subdivision in natural populations of European sea bass *Dicentrarchus labrax* (Linnaeus, 1758). Mol. Ecol. 6, 51–62. https://doi.org/10.1046/j.1365-294X.1997.t01-1-00151.x.

Geibel, J., Reimer, C., Weigend, S., Weigend3, A., Pook, T., Simianer, H., 2021. How array design creates SNP ascertainment bias. PLoS ONE 16 (3), e0245178. https://doi.org/10.1371/journal.pone.0245178.

Gjedrem, T., Baranski, M., 2010. Selective Breeding in Aquaculture: An Introduction. Springer Science & Business Media. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-90-481-2773-3.

Griot, R., Allal, F., Phocas, F., Brard-Fudulea, S., Morvezen, R., et al., 2021. Genome-wide association studies for resistance to viral nervous necrosis in three populations of European sea bass (*Dicentrarchus labrax*) using a novel 57k SNP array DlabChip. Aquaculture 530, 735930. https://doi.org/10.1016/j.aquaculture.2020.735930.

Hill, W.G., Robertson, A., 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38, 226–231. https://doi.org/10.1007/BF01245622.

Janssen, K., Chavanne, H., Berentsen, P., Komen, H., 2017. Impact of selective breeding on European aquaculture. Aquaculture 472, 8–16. ⟨https://www.researchgate.net/publication/297678584⟩.

Li, C.C., Horvitz, D.G., 1953. Some methods of estimating the inbreeding coefficient. Am. J. Hum. Genet. 5, 107–117.

Liu, Y., Nyunoya, T., Leng, S., Belinsky, S.A., Tesfaigzi, Y., Bruse, S., 2013. Softwares and methods for estimating genetic ancestry in human populations. Hum. Genomics 7, 1. https://doi.org/10.1186/1479-7364-7-1.

López-Cortegano, E., Pérez-Figueroa, A., Caballero, A., 2019. Metapop2: re-implementation of software for the analysis and management of subdivided populations using gene and allelic diversity. Mol. Ecol. Resour. 19, 1095–1100. https://doi.org/10.1111/1755-0998.13015.

Loukovitis, D., Sarropoulou, E., Vogiatzi, E., Tsigenopoulos, C.S., Kotoulas, G., et al., 2012. Genetic variation in farmed populations of the gilthead sea bream *Sparus aurata* in Greece using microsatellite DNA markers. Aquac. Res 43, 239–246. https://doi.org/10.1111/j.1365-2109.2011.02821.x.

Luikart, G., Cornuet, J.M., 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conserv. Biol. 12, 228–237. ⟨https://www.jstor.org/stable/2387479⟩.

Maroso, F., Gkagkavouzis, K., De Innocentiis, S., Hillen, J., do Prado, F., et al., 2021. Genome-wide analysis clarifies the population genetic structure of wild gilthead sea bream (*Sparus aurata*). PLoS ONE 16 (1), e0236230. https://doi.org/10.1371/journal.pone.0236230.

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

Nei, M., Chesser, R.K., 1983. Estimation of fixation indices and gene diversities. Ann. Hum. Genet. 47, 253–259. https://doi.org/10.1111/j.1469-1809.1983.tb00993.x.

Pante, M.A., Gjerde, B., McMillan, I., 2001. Effect of inbreeding on body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. Aquaculture 192, 201–211. https://doi.org/10.1016/S0044-8486(00)00467-1.

Pauletto, M., Manousaki, T., Ferraresso, S., Babbucci, M., Tsakogiannis, A., et al., 2018. Genomic analysis of *Sparus aurata* reveals the evolutionary dynamics of sex-biased genes in a sequential hermaphrodite fish. Commun. Biol 1, 119. https://doi.org/10.1038/s42003-018-0122-7.

Peñaloza, C., Manousaki, T., Franch, R., Tsakogiannis, A., Sonesson, A., et al., 2021. Development and testing of a combined species SNP array for the European seabass (*Dicentrarchus labrax*) and gilthead seabream (*Sparus aurata*). Genomics 113, 2096–2107. https://doi.org/10.1016/j.ygeno.2021.04.038.

Polovina, E.S., Kourkouni, E., Tsigenopoulos, C.S., Sanchez-Jerez, P., Ladoukakis, E.D., 2020. Genetic structuring in farmed and wild Gilthead seabream and European seabass in the Mediterranean Sea: implementations for detection of escapees. Aquat. Living Resour. 33, 7. https://doi.org/10.1051/alr/2020007.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. https://doi.org/10.1086/519795.

Quéré, N., Desmarais, E., Tsigenopoulos, C.S., Belkhir, K., Bonhomme, F., Guinand, B., 2012. Gene flow at major transitional areas in sea bass (*Dicentrarchus labrax*) and the possible emergence of a hybrid swarm. Ecol. Evol. 2, 3061–3078. https://doi.org/10.1002/ece3.406.

Rokach, L., Maimon, O., 2005. Clustering methods. In: Maimon, O., Rokach, L. (Eds.), Data Mining and Knowledge Discovery Handbook. Springer, Boston, MA. https://doi.org/10.1007/0-387-25465-X_15.

Santos, M.N., Lino, P.G., Pousão-Ferreira, P., Monteiro, C.C., 2006. Preliminary results of hatchery-reared seabreams released at artificial reefs off the Algarve coast (southern Portugal): a pilot study. Bull. Mar. Sci. 78, 177–184. ⟨https://www.researchgate.net/publication/228918743⟩.

Saura, M., Caballero, A., Santiago, E., Fernández, A., Morales-González, E., et al., 2021. Estimates of recent and historical effective population size in turbot, seabream, seabass and carp selective breeding programmes. Genet. Sel. Evol. 53, 85. https://doi.org/10.1186/s12711-021-00680-9.

Šegvić-Bubić, T., Lepen, I., Trumbić, Ž., Ljubković, J., Sutlović, D., et al., 2011. Population genetic structure of reared and wild gilthead sea bream (*Sparus aurata*) in the Adriatic Sea inferred with microsatellite loci. Aquaculture 318, 309–315. https://doi.org/10.1016/j.aquaculture.2011.06.007.

Šegvić-Bubić, T., Grubišić, L., Trumbić, Ž., Stanić, R., Ljubković, J., et al., 2017. Genetic characterization of wild and farmed European seabass in the Adriatic sea: assessment of farmed escapees using a Bayesian approach. ICES J. Mar. Sci. 74, 369–378. https://doi.org/10.1093/icesjms/fsw155.

Souche, E., Hellemans, B., Babbucci, M., Mac Aoidh, E., Guinand, B., et al., 2015. Range-wide population structure of European sea bass *Dicentrarchus labrax*. Biol. J. Linn. Soc. 116, 86–105. https://doi.org/10.1111/bij.12572.

Su, G.-S., Liljedahl, L.-E., Gall, G.A.E., 1996. Effects of inbreeding on growth and reproductive traits in rainbow trout (*Oncorhynchus mykiss*). Aquaculture 142, 139–148. https://doi.org/10.1016/0044-8486(96)01255-0.

Tine, M., Kuhl, H., Gagnaire, P.A., Louro, B., Desmarais, E., et al., 2014. European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. Nat. Commun. 5, 5770. https://doi.org/10.1038/ncomms6770.

Vandeputte, M., Gagnaire, P.A., Allal, F., 2019. The European sea bass: a key marine fish model in the wild and in aquaculture. Anim. Genet. 50, 195–206. https://doi.org/10.1111/age.12779.

Villanueva, B., Fernández, A., Saura, M., Caballero, A., Fernández, J., et al., 2021. The value of genomic relationship matrices for estimating inbreeding. Genet. Sel. Evol. 53, 42. https://doi.org/10.1186/s12711-021-00635-0.

Waples, R., 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conserv. Genet. 7, 167–184. https://doi.org/10.1007/s10592-005-9100-y.

Yáñez, J.M., Bassini, L.N., Filp, M., Lhorente, J.P., Ponzoni, R., Neira, R., 2014. Inbreeding and effective population size in a coho salmon (*Oncorhynchus kisutch*) breeding nucleus in Chile. Aquaculture 420–421, S15–S19. https://doi.org/10.1016/j.aquaculture.2013.05.028.

Žužul, I., Šegvić-Bubić, T., Talijančić, I., Džoić, T., Lepen Pleić, I., et al., 2019. Spatial connectivity pattern of expanding gilthead seabream populations and its interactions with aquaculture sites: a combined population genetic and physical modelling approach. Sci. Rep. 9, 14718. https://doi.org/10.1038/s41598-019-51256-z.