Aquaculture 621 (2026) 743980

Contents lists available at ScienceDirect

Aquaculture

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

Application of whole-eDNA sequencing for monitoring broodstock genetic diversity in breeding groups of gilthead seabream (*Sparus aurata*)

Francesca Bertolini[a,*], Valeria Taurisano[a], Samuele Bovo[a], Jacopo Vegni[a], Riccardo Napolitano[b], Marco Martinoli[b], Fabrizio Capoccioni[b], Luca Fontanesi[a]

[a] *Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale G. Fanin 46, 40127 Bologna, Italy*
[b] *Centro di Ricerca Zootecnia e Acquacoltura, Consiglio per la ricerca in agricoltura e l'analisi Dell'economia Agraria (CREA), 00198 Roma, Italy*

ARTICLE INFO

Keywords:
Aquaculture genomics
Environmental DNA (eDNA)
Non-invasive monitoring
Pool-Seq
Population genomics
Whole-genome sequencing

ABSTRACT

Environmental DNA (eDNA) analysis represents a non-invasive approach for assessing biodiversity and its use is expanding in several fields. However, its potential to recover population genomic information remains largely unexplored. In this study we evaluated the feasibility of whole eDNA sequencing to infer broodstock genetic diversity in a breeding aquaculture population, using the gilthead seabream (*Sparus aurata*) as a model species. Water samples were collected from four broodstock tanks and filtered through three pore sizes (0.2, 1.2, and 5 μm), along with tissue samples from the fish in each tank. DNA extracted from the filters (eDNA) and pooled from the fish in each of the four tanks was sequenced using next-generation sequencing and reads were mapped to the gilthead seabream reference genome. Q20 mapped reads were used for variant calling and single nucleotide polymorphisms were utilised to calculate heterozygosity-based indices (F_{IS}-like) and Fixation index (F_{ST}) within and across DNA pools derived from the four tanks and corresponding eDNA samples. Mapping rates confirmed the expected predominance of non-target DNA in eDNA samples across filters, with the 5 μm filters showing the highest rate of mapped reads (from 34% to 51% in three out of 4 tanks), resulting in a genome coverage that ranges from 7.05 to 12.89× of Q20 mapped reads. Moreover, F_{IS}-like and F_{ST} values were similar between DNA pools and eDNA datasets, with low differentiation observed across tanks and filters ($\Delta F_{IS\leq}0.01$ for the depth-balanced dataset). These results indicate that whole eDNA sequencing can recover genome-wide estimates of genetic diversity comparable to those obtained from fish tissue samples, representing a first proof-of-concept for its application in breeding aquaculture populations.

1. Introduction

Genetic diversity refers to the variation in the genetic makeup among individuals within a population or species. Genetic diversity is influenced by numerous natural and human-driven processes, and it is crucial for adaptation to environmental changes to ensure the overall survival of a species or population (Hughes et al., 2008; Ellegren and Galtier, 2016). Monitoring genetic diversity is essential in aquaculture to maintain long-term population health, productivity, and ensure the sustainability of the system. Effective management of genetic variation within stocks provides the foundation for selective breeding aimed at genetic improvement, while also preserving the ability to adjust breeding goals in response to new challenges or needs. This helps prevent undesired and uncontrolled declines in performance and economic loss (Lind et al., 2012; Houston et al., 2020). In fact, erosion of genetic diversity is a major contributor to inbreeding depression, accompanied by an increase in the frequency of homozygous negative and recessive alleles. This can lead to detrimental effects on growth, fertility, survival, disease resistance, and an increase in deformities (Tave, 1999; Charlesworth and Willis, 2009; Gao et al., 2015; Bertolini et al., 2020; Lyu et al., 2025). Although advanced selective breeding programs offer an effective way to improve performance and manage inbreeding, they are only available for a few aquaculture species and commercial operations globally (Gjedrem et al., 2012; Chavanne et al., 2016; Sonesson et al., 2023). However, in most cultivated species, where these programs are not yet established due to high costs, limited broodstock resources, logistical constraints, and/or a fragmented industry structure, routine monitoring of genetic diversity is still challenging despite its relevance.

The estimation of genetic diversity generally requires sampling and genotyping of individual fish, which can be invasive, time-consuming,

---

* Corresponding author.
E-mail address: francesca.bertolini3@unibo.it (F. Bertolini).

https://doi.org/10.1016/j.aquaculture.2026.743980
Received 12 December 2025; Received in revised form 28 March 2026; Accepted 5 April 2026

and costly. These limitations may restrict how often and at what scale genetic diversity can be monitored in aquaculture settings. In this context, environmental DNA (eDNA) emerges as a promising tool for non-invasive and cost-effective genetic surveillance. Environmental DNA extracted from the water has been successfully used to detect and monitor aquatic species across a wide range of taxa and environments, enabling biodiversity assessments and early detection of invasive and endangered species or pathogens (e.g., Westgaard et al., 2024; Thomsen and Willerslev, 2015; Taberlet et al., 2018; Rees et al., 2014; Deiner et al., 2017; Gomes et al., 2017; Benedicenti et al., 2024). Environmental DNA has also been successfully applied to estimate species abundance and biomass through correlations between DNA concentration and organismal density (e.g., Takahara et al., 2012; Rourke et al., 2022; Kasmi et al., 2025; Doi et al., 2017). Recently, eDNA approaches have also been utilised to infer population genetic structures or intraspecific variation using mitochondrial or a limited number of targeted nuclear genome markers (Sigsgaard et al., 2019; Sigsgaard et al., 2016; Andres et al., 2021; Weitemier et al., 2021). However, to our knowledge, no studies have explored the use of eDNA to assess whole genomic diversity in any finfish farmed species.

The gilthead seabream (*Sparus aurata* L.) is the main species reared in Mediterranean countries, reaching 323,544 t (34.5% of the total production of both fish and shellfish) in 2023 (FAO, 2025). Despite its economic relevance, the majority of commercial production of this species relies on a limited number of hatcheries that do not always apply selective breeding schemes or genetic monitoring of broodstock diversity. This lack may increase the risk derived from excessive inbreeding in the breeding stocks, accompanied by an uncontrolled erosion of genetic variability over successive generations (Cossu et al., 2019; Villanueva et al., 2022). Recently, several genomic resources have become available for gilthead seabream, providing an opportunity to perform genome-wide analyses to detect loci linked with genetic defects and other economically relevant traits (Bertolini et al., 2020; Bertolini et al., 2021; Aslam et al., 2020; Yu et al., 2024). This also enables genome-wide assessments of population structure and connectivity among wild and farmed stocks (Maroso et al., 2021; Peñaloza et al., 2021).

In this study, we evaluated the feasibility of using untargeted whole-eDNA sequencing to estimate broodstock genetic diversity in the gilthead seabream in a commercial hatchery located in Italy, one of the main Mediterranean gilthead seabream producers. Broodstock management at this facility is designed to minimise inbreeding through a cohort-segregation strategy. Despite this rigorous demographic control, no prior molecular characterization was available for these stocks. We assessed the sequencing performance of eDNA collected from various tanks with different types of filters as well as their ability to assess genetic diversity, through comparison with whole-genome sequencing data of DNA pools from tissues sampled directly from the same fish included in the eDNA sampled tanks. This comparison through *F*-based statistics indirectly identified the most suitable filter type for capturing representative genomic information from gilthead seabream. Additionally, we estimate genetic diversity between tanks, using both DNA-pool and eDNA sequencing data, and explored, as a by-product, the richness of non-fish derived reads from eDNA samples.

2. Material and methods

2.1. Ethics statement

The tissue samplings were performed exclusively on fish reared in a commercial aquaculture facility, within the framework of the farm's routine management and authorised practices. No experimental procedures beyond those ordinarily permitted by the farm were carried out. Therefore, no specific ethical approval was required, as the sampling activities were part of normal husbandry operations and complied with all applicable animal welfare standards.

2.2. Dataset description, sampling, and DNA extraction

The overall experimental design is summarised in Fig. 1. Genetic diversity was quantified from a total of 245 individual broodstock, distributed into four tanks (indicated as T1, T2, T3, and T4, respectively). Environmental DNA was collected by sampling water from the same tanks using sterile polyvinyl chloride tubes and subsequently filtering it through filters with different pore sizes. More details on samples and sampling procedures are reported below.

2.2.1. Individual fish tissue sampling

The animals were maintained in a recirculating aquaculture system (RAS) comprising six circular tanks, each measuring 3 × 1 m of diameter and containing approximately 7000 L of water maintained at 20 °C. Four target tanks housed between 57 and 67 gilthead seabream (*Sparus aurata*) individuals that were uniquely tagged, with an approximate male-to-female ratio of 1:2 and an average body weight ranging from 2.64 to 4.03 kg (Table 1). Following light anesthesia, a ∼ 1 cm^2 section of caudal fin was collected from each specimen using sterilised scissors and immediately stored on dry ice during transport to the laboratory, where samples were preserved at −20 °C. Sampling was carried out by hatchery staff during the annual broodstock tank cleaning to minimise handling stress. No mortality or injuries occurred during tissue collection.

2.2.2. Sampling eDNA from tank water

Two months after the tissue sampling, water was filtered from the tanks. The tanks contained the same fish at the time of tissue sampling, except 11 fish (4 in T1, 1 in T2, 5 in T3 and 1 in T4) that died prior to water sampling and were therefore excluded from DNA pooling procedure. The water of the first two tanks (i.e., T1 and T2) was filtered during water flow/circulation of the RAS system, while the water of T3 and T4 was filtered when the flow was absent. For each tank, 7 L of water was collected with a 1 m sterile polyvinyl chloride tube (one for each tank) that was positioned 15 cm below the water surface, then filtered with three types of filters with different pore sizes: a 0.2 μm (15 cm, Ø90 mm) polyethersulfone (PES) filter, 1.2 μm and 5 μm (Ø28 mm) Minisart® surfactant-free cellulose acetate (SFCA) filters (Sartorius, Goettingen, Germany). A total of 3 L was filtered mechanically with a vacuum pump with a 0.2 μm filter. Four L of water were then manually filtered with a sterile 100 mL syringe for the 1.2 μm and 5 μm filters until reaching 2 L of filtered water in each filter. After filtration, each 0.2 μm filter was immediately removed from its housing using sterilised forceps, while 1.2 μm and 5 μm filters were extracted from their solid envelopes using a rotary tool. Extracted filters were then placed in labelled sterile tubes, immediately stored in dry ice, and for long-term at −20 °C (Fig. S1).

2.2.3. DNA extraction from fin clips and construction of DNA pools

Total genomic DNA was isolated from a small portion of each fin clip (2 × 2 mm) using a cetyltrimethylammonium bromide (CTAB) modified protocol: 500 μL of CTAB buffer extraction buffer [2% (*w/v*) cetyltrimethylammoniumbromide; 1.4 M NaCl; 100 mM Tris-HCl; 20 mM EDTA pH 8.0] were added to a 1.5 mL tube containing fin clip tissue. After the addition of 10 μL of proteinase K (20 mg/mL), the tubes were incubated at 56 °C for three hours with gentle mixing. After the incubation, the samples were treated with 5 μL of RNase A at 37 °C for 15 min. Then tubes were centrifuged at 15000*g* at room temperature for 5 min, the supernatant was transferred into tubes containing 500 μL of chloroform:isoamyl alcohol (24:1), vortexed for 30 s, and centrifuged at 15000*g* for 15 min at room temperature. The upper aqueous phase was transferred to a new 1.5 mL tube containing 500 μL of cold isopropanol. Tubes were carefully mixed by inversion and kept at −20 °C for 30 min. DNA was pelleted by centrifugation at 15000*g* for 10 min at room temperature. The DNA pellets were washed by the addition of 500 μL of 70% ethanol and centrifuged for 10 min at 15000*g*. Ethanol was

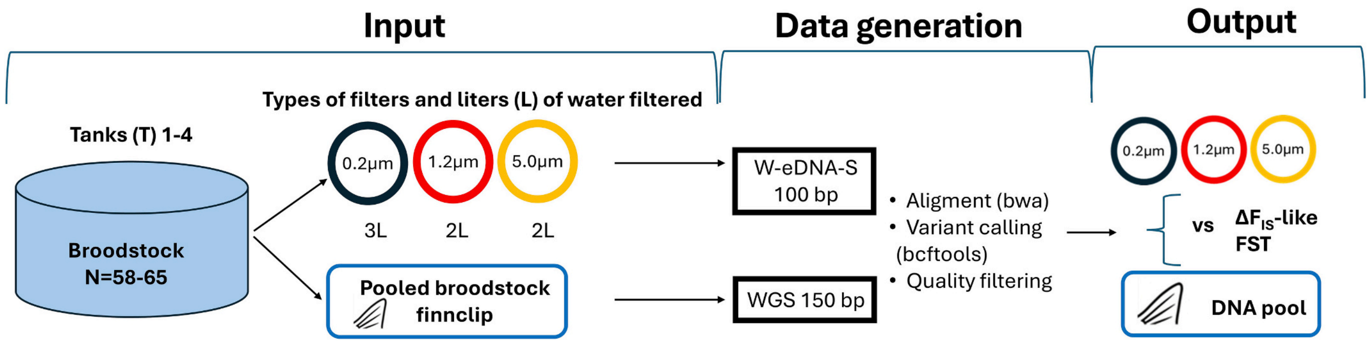

Fig. 1. Experimental design and summary of the main analyses performed. *Legend: WGS: whole genome sequencing; W-eDNA-seq: whole-eDNA sequencing.*

Table 1
Number of gilthead seabream (*Sparus aurata*) broodstock individuals and mean body weight (kg). The table reports the number of females (F) and males (M) sampled in each tank (T1–T4) at the time of tissue sampling. ± sd) per tank and sex.

Tank name	Fish (n.)	Sex (n.)	Mean (kg)	± s.d.
T1	67	F (43)	3.58	0.51
		M (24)	2.64	0.34
T2	58	F (37)	4.17	0.66
		M (21)	3.55	0.57
T3	63	F (35)	4.03	0.51
		M (28)	3.54	0.67
T4	57	F (35)	3.84	0.51
		M (22)	2.89	0.52

removed, and the tubes were left to dry at room temperature. Pellets were rehydrated with 30 µL of sterile H_2O and stored at −20 °C. Genomic DNA quality was individually assessed using a NanoPhotometer (Implen, Munich, Germany), with 260/280 ratios >1.7 and 260/230 ratios >1.9. DNA integrity was evaluated by electrophoresis on a 0.8% agarose gel, and concentrations were measured using a Qubit 4 fluorometer (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Based on Qubit quantification, samples were normalised to equal concentrations prior to pooling.

2.2.4. Environmental DNA extraction from filters

Environmental DNA present in the tank water was extracted from the filters using the following protocol. Filters were soaked in 700 µL of CTAB extraction buffer with the addition of 5 µL of RNase A, incubated at 60 °C for 10 min with gentle mixing. After this initial incubation, 30 µL of proteinase K (20 mg/mL) was added to each sample, and a second incubation was performed for 90 min at 65 °C. All tubes were cooled to room temperature, then centrifuged at 16000*g* for 10 min. Next, tubes were centrifuged at 15000*g* at room temperature for 5 min. The supernatant was transferred into tubes containing 700 µL of chloroform:isoamyl alcohol (24:1), vortexed for 30 s, and centrifuged at 15000*g* for 15 min at room temperature. The supernatant was then transferred to a new 1.5 mL tube, and the DNA was precipitated in two steps using isopropanol and 70% ethanol. The DNA pellets were rehydrated with 30 µL of sterile H_2O.

2.3. High-throughput DNA sequencing, read mapping, and variant calling

The four fish DNA pools, constructed from the tissue collected from the fish contained in each of the four tanks, and 12 eDNA samples, three for each tank (from the three different types of filters, i.e., 0.2, 1.2, and 5 µm), were subjected to high-throughput DNA sequencing, using the DNBSEQ sequencing platform (BGI Genomics Co., Ltd., Shenzhen, China). From the tissue DNA pools, whole genome paired-end sequencing was based at 150 bp DNA fragment size. Environmental DNA samples were sequenced with the same pipeline and chemistry but with 100 bp read length to account for possible higher fragmented DNA, according to the metagenomic pipeline. All sequenced reads were mapped against the gilthead seabream reference genome sequence (NCBI: fSpaAur1.2, GCF_900880675.2) using bwa-0.7.17 mem (Li, 2013) with standard options. Subsequently, reads were filtered retaining only those with mapping quality >10 (Q10) and > 20 (Q20) with Samtools 1.2 (Li et al., 2009), for an evaluation of the mapping success at different quality thresholds. Variant calling was then performed on each of the Q20 filtered data with bcftools 1.2 mpileup (Danecek et al., 2021), considering only biallelic single nucleotide polymorphisms (SNPs) with variant quality >30 (Q30) and a depth from 10 to 150 ×, defined after calculating overall depth distribution of called SNPs across tanks and tissue/filter origin, as shown in Fig. S2.

2.4. Comparison of sequencing data derived from eDNA samples and fish DNA pools

Shared SNPs among each eDNA sample and the respective DNA pools derived from the same tanks were considered for pairwise comparisons. In these analyses, we included only SNPs that shared the same alternative allele among the filters and the DNA pool derived from the same tank. These comparisons were performed with R 4.4.2 (R Core Team R, 2024) and with the R package poolfstat (Hivert et al., 2018).

2.4.1. Depth balanced analyses

To correct for uneven sequencing depth among samples derived from the same tanks (eDNA filters and fish DNA pools), we aimed to obtain comparable per-tank information content and avoid distortions in allele frequency estimates due to unequal coverage (Tilk et al., 2019). To minimise biases associated with extreme depth differences, the dataset with higher coverage (i.e., the tissue-derived DNA pool, which serves as a benchmark and baseline for comparison) was matched to the shallower one (i.e., the filter-derived eDNA) by binomial down-sampling of both reference and alternate read counts, adapting a random thinning approach previously described (Hivert et al., 2018; Kofler et al., 2011; Bhatia et al., 2013).

For each SNP in the deeper DNA pool (i.e., the tissue-derived DNA pool):

$$R'_{s,\mathrm{deep}} \sim \mathrm{Binom}\left(R_{s,\mathrm{deep}}, f\right), A'_{s,\mathrm{deep}} \sim \mathrm{Binom}\left(A_{s,\mathrm{deep}}, f\right),$$

Where:

$R_{s,\mathrm{deep}}=$ number of reference allele reads observed at SNP s in the deeper DNA pool,

$A_{s,\mathrm{deep}}=$ number of alternate allele reads observed at the same SNP,

$R'_{s,\mathrm{deep}}$ and $A'_{s,\mathrm{deep}}=$ down-sampled read counts retained after thinning,

$f=$ thinning factor applied equally to both alleles, defined as:

$$f = \min\left(1, \frac{\text{median coverage of shallow pool}}{\text{median coverage of deep pool}}\right).$$

defines the thinning factor applied to both allele counts. A fixed random seed (set.seed(123)) was used to ensure reproducibility. After depth normalisation, only SNPs where all DNA pools exhibited coverage within the range of 10–150 reads (inclusive) were retained, with no limit on the minor allele counts.

2.4.2. F_{IS}-like and F_{ST} analyses

Two pairwise comparisons of filters against tissue DNA pools were performed for both depth-balanced and unbalanced datasets, yielding $\Delta\mathrm{F_{IS}}$-like and $\mathrm{F_{ST}}$ estimates. The $\mathrm{F_{IS}}$-like index represents a genome-wide, within-pool heterozygosity deficit relative to Hardy–Weinberg expectations (analogous to the inbreeding coefficient $\mathrm{F_{IS;}}$ Wright, 1951; Weir and Cockerham, 1984; O'Reilly et al., 2024), while $\mathrm{F_{ST}}$ quantifies genetic differentiation between the two pools (Weir and Cockerham, 1984; Kofler et al., 2011; Bhatia et al., 2013). These complementary metrics allow the evaluation of whether eDNA-derived allele frequencies reproduce the levels of heterozygosity and population structure observed from tissue DNA:

$\Delta\, F_{IS}^{(\mathrm{like})}$: for each DNA pool, a F (like) statistic was calculated with the formula

$$F_{IS}^{(\mathrm{like})} = \frac{He - Hw}{\mathrm{He}} = 1 - \frac{Hw}{\mathrm{He}}$$

where:

H_w = observed heterozygosity within the DNA pool, estimated as:

$$H_w = 1 - \sum_i p_i^2$$

with p_i representing the frequency of the i-th allele (for biallelic SNPs, $p_1 = p, p_2 = 1 - p$); equivalently, $H_w = 2p(1-p)$.

H_e = expected heterozygosity under Hardy–Weinberg equilibrium, computed directly from the allele frequency estimates derived from read counts:

$$H_e = 2p(1-p)$$

with = frequency of the reference allele in each DNA pool, calculated as:

$$\mathrm{p} = \frac{\mathrm{R} + \mathrm{A}}{\mathrm{A.}}$$

where R and A are the reference and alternate allele read counts, respectively.

To obtain uncertainty while accounting for linkage, we performed a leave-one-chromosome-out (LOCO) jackknife. For each chromosome b we recomputed H_W and H_E after excluding all SNPs on b, then derived $F_{IS}^{(\mathrm{like})}$. With B chromosomes, the jackknife means, and standard error (SE) for any statistic θ (were)

$$(\overline{\theta}) = \frac{1}{B}\sum_{b=1}^{B}\theta_{(-b)}$$

$$\mathrm{SE}(\overline{\theta}) = \sqrt{\frac{B-1}{B} * \frac{1}{B}\sum_{b-1}^{B}\left(\theta_{(-b)} - \overline{\theta}\right)^2}$$

$$CI_{95\%} = \overline{\theta} \pm 1.96 \times SE$$

Finally, the difference in heterozygosity deficit between the eDNA and tissue-derived DNA pools from the same tank was computed as follows:

$$\Delta F_{IS}^{(\mathrm{like})} = F_{IS}^{(\mathrm{like})}(pool) - F_{IS}^{(\mathrm{like})}(eDNA)$$

SE was obtained from the LOCO jackknife applied to $F_{IS}^{(\mathrm{like})}$ across chromosomes. We also calculated the 95% CI for $\Delta F_{IS}^{(\mathrm{like})}$.

2.4.3. Within tank pairwise F_{ST}

Genetic differentiation between each eDNA sample and the corresponding fish DNA pool from the same tank was quantified by the unbiased F_{ST} estimator implemented in *poolfstat* (Hivert et al., 2018), which extends the Weir and Cockerham (1984) formulation to DNA pool-sequencing data. For each locus:

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

where H_S is the average expected heterozygosity within DNA pools and H_T the total heterozygosity across all DNA pools. Genome-wide estimates were obtained by averaging across SNPs, and uncertainty was again assessed via the LOCO jackknife across chromosomes.

Pairwise $\mathrm{F_{ST}}$ of the filters with the highest depth and coverage was estimated against the corresponding fish DNA pool using a sliding-window approach. For each chromosome, SNPs were partitioned into consecutive windows of 1 Mb (or another specified size). Windows containing fewer than 20 SNPs were excluded from downstream analyses. For every window, the corresponding subset of SNPs was extracted, and $\mathrm{F_{ST}}$ was calculated using the multilocus estimator implemented in the compute.fstats function (poolfstat), without block-jackknife resampling.

2.5. Between tanks pairwise F_{ST}

Pairwise genetic differentiation among tanks was calculated separately for tissue-derived DNA pools and for each eDNA filter type (0.2 μm, 1.2 μm, and 5 μm), using the same unbiased $\mathrm{F_{ST}}$ estimator and LOCO jackknife procedure described above. In this analysis, $\mathrm{F_{ST}}$ values quantify the genome-wide differentiation among tanks within each sampling matrix, providing a measure of the consistency between eDNA- and tissue-based estimates of population structure. Since sequencing depth was comparable across filters and pools, no binomial down-sampling was conducted before $\mathrm{F_{ST}}$ estimation. Genome-wide pairwise $\mathrm{F_{ST}}$ values were averaged across all SNPs for each tank pair (1 vs 2, 1 vs 3, 1 vs 4, 2 vs 3, 2 vs 4, 3 vs 4), and standard errors were calculated from the jackknife variance among chromosomes. These estimates were used to compare the magnitude and rank order of genetic differentiation across filter types and to evaluate whether whole-eDNA sequencing replicates the same between-tank relationships observed in individual tissue-derived DNA pools.

2.6. Analyses of reads not mapped to the gilthead seabream reference genome derived from eDNA samples

Approximately 3 M unmapped reads randomly selected for each filter and each tank, were analysed using blastx with the software NCBI-blast-2.12.0 (Camacho et al., 2009), utilising the uniprot_swiss_2025 database. Here, taxa IDs with the highest score were retained and converted into taxonomic classifications using the taxize R package (Chamberlain and Szöcs, 2013) with additional manual verification based on the NCBI taxonomy database. For each tank, read counts were aggregated by phylum, and the total number of reads per phylum was used to compute relative abundances (i.e., % of total annotated reads per phylum). Differences in phylum-level proportions among filters were expressed as percentage-point changes and tested for statistical significance using the *prop.test* function in R, to control for multiple testing, p-values were adjusted using the Bonferroni correction. Significance was assessed at α = 0.05 for corrected p-values.

3. Results

3.1. Overall sequencing quality, SNP calling, and genome distribution

High-throughput sequencing produced an average of 744.2 (±13.05) M reads for the DNA pools, and 261.58 (±45.43) M reads for the eDNA samples (Table 2). The majority of the DNA-pool derived reads mapped against the gilthead seabream reference genome (>99%). For the sequencing data obtained from eDNA samples, the 5 μm filters produced the highest fraction of reads mapping to the gilthead seabream reference genome (34–51%), except in the case of Tank 1. In Tanks 3 and 4, mapping rates increased steadily as filter pore size increased. The proportion of reads mapped with high quality against the gilthead seabream reference genome showed consistent patterns across the four tanks (T1–T4) and between mapping quality thresholds Q10 and Q20 (Fig. 2). As expected, increasing the mapping quality threshold from Q10 to Q20 resulted in a slight decrease in the percentage of mapped reads, with differences generally below 2%. The most pronounced variation in mapping efficiency was observed in T2, where the proportion of mapped reads reached approximately 10–11% of difference between Q10 and Q20, respectively. Overall, the decrease in mapped reads associated with Q20 quality filtering was modest compared with Q10 and did not affect the general ranking among eDNA samples derived from the two filter types, supporting the use of Q20 as a conservative threshold for downstream variant calling. Sequencing depth and genome coverage at Q20 varied markedly between eDNA samples and DNA pools (Table 3). As expected, DNA pools showed the highest sequencing depth, averaging approximately 120× across tanks and covering more than 96% of the gilthead seabream reference genome at depth > 10. In contrast, eDNA samples exhibited substantially lower depth (2.7–12.9×) and more variable genome coverage (0.2–82.7%), with the highest performance obtained from the eDNA data derived from the filter with 5 μm pores. The same pattern was observed when assessing the number of high-quality SNPs detected. Sequencing data derived by DNA pools captured from ~6.4 to 6.9 M SNPs. The closest eDNA sample reaching that value was that obtained with a 5 μm pore filter from T3 and T4, with 4,157,156 and 2,925,454, respectively, while T1 and T2 reached 16,839 and 206,792 SNPs. The chromosomal distribution of high-quality SNPs obtained from eDNA samples is shown in Fig. 3. Across all tanks (T1–T4), SNPs were uniformly distributed along the 24 chromosomes of the gilthead seabream genome, with no evidence of chromosomal bias.

3.2. Genetic diversity patterns from eDNA samples and DNA pools

Comparisons between estimates of genetic diversity from eDNA samples and DNA pools revealed consistent trends across tanks and filters (Table 4 and Table S1). When using the unbalanced dataset, where allele counts were based on raw read depth, eDNA samples exhibited a systematic excess of homozygosity compared to the corresponding DNA pools, resulting in ΔF_{IS} values ranging from 0.07 to 0.1 across tanks and filters. After depth balancing by binomial down-sampling, ΔF_{IS} values decreased in magnitude, reaching values from 3.91E-03 to 0.01. This reduction indicates that the initial heterozygosity deficit observed in eDNA samples may be largely due to coverage differences rather than genuine biological divergence. Standard errors for F_{IS}-like and ΔF_{IS} estimates were uniformly small (typically <0.002), and the 95% confidence intervals (CI95%) were narrow and symmetrically distributed around the mean, supporting the high precision and numerical stability of the genome-wide estimates across chromosomes (Table S1). These values were small both before and after depth balancing, indicating that the estimated differences between eDNA and tissue pools are statistically consistent and not driven by sampling noise. Pairwise F_{ST} values between eDNA samples and DNA pools were generally low, with several comparisons approaching zero, suggesting overall genetic concordance between the two sources of DNA once read filtering and depth balancing were applied. Pairwise F_{ST} between eDNA samples and DNA pools remained low in all cases, with several comparisons indistinguishable from zero (Table 4 and Table S1). Moderate values persisted mainly for some 5 μm filters, consistent with a greater contribution of non-gilthead seabream DNA in that fraction. While balanced F_{ST} values were slightly higher than unbalanced in some tanks (e.g., T1–T2), they remained small in absolute terms with an increase of a maximum of 0.01 or 0.02 in T1 with the eDNA obtained with the 0.2 μm filter. Downsampling was applied to ensure comparability among datasets with different sequencing depths; however, this procedure reduces the amount of information and may increase sampling variance, leading to slightly inflated F_{ST} estimates despite negligible biological differentiation. Distribution of F_{ST} values across 1 Mb windows is shown in Fig. 4. Here, the values both in terms of average F_{ST} and in terms of distribution seem to be reduced by increasing the depth of sequencing (higher in T3 and T4), meaning that the overall concordance of allele frequency is also reflected throughout the whole genome.

3.3. Genetic diversity across tanks

Genome-wide pairwise F_{ST} values among tanks were consistently close to zero across all sampling matrices (Table 5), indicating negligible genetic differentiation within the broodstock population. Estimates derived from eDNA filters (0.2, 1.2, and 5 μm) were highly consistent with those obtained from tissue-derived DNA pools, with absolute F_{ST}

Table 2

Sequencing summary for environmental DNA (eDNA) and DNA pools from the four tanks (T1–T4). The table reports, for each DNA source (eDNA from filters with pore sizes 0.2 μm, 1.2 μm, and 5 μm, and DNA pools), the total number of raw reads generated by sequencing, number of total reads mapped to the reference genome, and after applying base-quality thresholds (Q10 and Q20).

TANKS	DNA SOURCE	NO. READS	NO. MAPPED READS	NO. MAPPED READS Q10	NO. MAPPED READS Q20
T1	0.2	360,304,022	28,421,723	23,014,864	22,604,276
	1.2	272,390,730	30,797,614	25,505,383	25,083,921
	5	360,399,269	33,431,060	28,209,211	27,747,359
	pool	739,306,394	735,190,707	692,734,217	686,960,661
T2	0.2	240,179,893	67,587,314	44,170,661	41,944,153
	1.2	221,602,383	50,831,862	35,237,437	33,722,183
	5	240,564,652	83,712,420	61,031,076	58,794,520
	pool	766,296,507	762,212,598	718,148,160	712,246,506
T3	0.2	240,337,713	55,513,205	42,338,576	41,159,805
	1.2	240,580,091	81,774,170	68,009,372	66,576,509
	5	240,786,159	124,906,989	109,488,327	107,467,889
	pool	742,905,382	738,855,255	695,903,368	690,181,326
T4	0.2	240,320,424	63,457,819	46,869,430	45,552,764
	1.2	240,583,792	77,720,911	63,119,854	61,539,960
	5	240,947,512	101,817,257	90,140,409	88,931,069
	pool	731,166,894	727,612,399	683,561,872	677,963,451

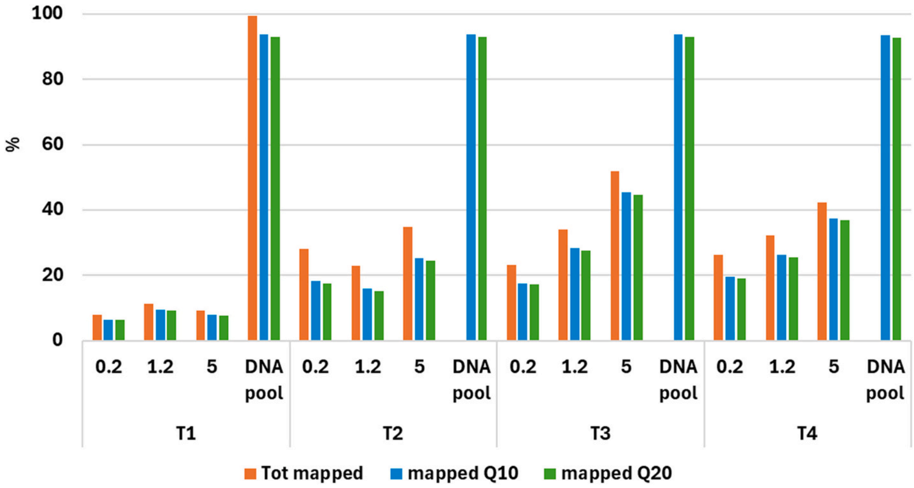

Fig. 2. Percentage of mapped reads against the gilthead seabream reference genome of unfiltered (orange), Q10 (blue), and Q20 (green) filtered reads obtained from eDNA (derived from the three filters: 0.2, 1.2, and 5 μm) and DNA pools from the four tanks (T1-T4). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3

Sequencing depth (Q20), genome coverage (% of bases with depth (d) > 10), and number (n.) of high-quality SNPs identified in each tank (T1–T4) and DNA source (eDNA from filters of 0.2 μm, 1.2 μm, and 5 μm pore size, and DNA pools).

TANKS	SOURCE	DEPTH (X)	COVERAGE D > 10 (%)	HIGH QUALITY SNPS (NO.)
T1	0.2	2.71	0.24	3056
	1.2	3.01	0.49	4862
	5	3.33	1.13	16,839
	Pool	123.62	96.53	6,503,666
T2	0.2	5.03	0.84	4852
	1.2	4.05	0.95	5596
	5	7.05	12.13	206,792
	Pool	128.17	96.55	6,403,778
T3	0.2	4.94	1.20	19,074
	1.2	7.99	30.45	875,526
	5	12.89	82.68	4,157,156
	Pool	124.20	96.53	6,533,946
T4	0.2	5.46	1.57	19,503
	1.2	7.38	18.19	399,621
	5	10.67	66.44	2,925,454
	Pool	122.00	96.33	6,867,322

values below ~0.03 in all comparisons and standard errors typically <0.002. Occasionally, slightly negative estimates (e.g., $F_{ST} = -0.0186 \pm 0.0003$) occurred, a known artifact of unbiased estimators under very low differentiation (Weir and Cockerham, 1984; Bhatia et al., 2013). The number of SNPs retained for each pairwise comparison varied widely across matrices, from approximately 1300 to 6400 loci in the 0.2 μm datasets, 1300–177,000 in the 1.2 μm filters, 8000–2.3 M in the 5 μm filters, and over 5 M in the tissue DNA pools. Despite these strong differences in data volume, genome-wide F_{ST} estimates remained highly concordant, demonstrating that the eDNA-based approach can reproduce the same pattern of genetic diversity observed from tissue sequencing.

Mean F_{ST} across filters ranged from 0.004 (0.2 μm) to 0.02 (1.2 μm), while mean F_{ST} across tissue pools averaged $0.001 \pm 6.7 \times 10^{-5}$. These extremely low FST values indicate negligible genetic differentiation among tanks. In parallel, FIS-like estimates were consistently negative across tanks ($-0.11 <$ FIS-like < -0.09, see Table S1), suggesting a mild excess of heterozygosity relative to Hardy–Weinberg expectations within tanks. Together, these results support the presence of a single, well-mixed broodstock population, consistent with the absence of structured breeding or family separation in the facility.

3.4. *Information from the unmapped reads*

The analyses of a random subset of the unmapped reads of the eDNA samples from the different tanks detected 125 phyla. Among those, differences in the relative abundance among the three filters were significant for 39 phyla in T1, 47 in T2, 55 in T3, and 24 in T4 (Bonferroni-adjusted $p < 0.05$; Table S2–5). Twenty-two of these significant phyla were common to all four tanks (Fig. S3). Similar patterns of filter-dependent abundance were observed in these phyla across all tanks (Table S2–S5). In fact, the most abundant bacterial group, *Pseudomonadota*, showed a progressive decrease in relative abundance from the smallest (0.2 μm) to the largest (5 μm) filter across all tanks, with differences ranging from −1.9 to −6.8 percentage points. Other dominant bacterial phyla, such as *Bacteroidota*, *Bacillota*, and *Mycoplasmatota,* exhibited comparable decreasing trends, suggesting a preferential retention of free-living or smaller bacterial cells in the finer filters. Multicellular eukaryotic phyla such as *Chordata*, *Nematoda*, and *Arthropoda* consistently increased in abundance with larger pore sizes, reflecting their higher representation in the 5 μm fractions. *Streptophyta* (Eukaryotic plants) and *Basidiomycota* (Eukaryotic fungi) displayed similar patterns, with relative abundance differences between eDNA derived from 0.2 μm and 5 μm filters typically exceeding +0.5 percentage points in all tanks.

4. Discussion

Since its first applications in biodiversity and ecological monitoring, eDNA has provided a non-destructive, scalable, and repeatable method for assessing genetic variation (Ficetola et al., 2015; Thomsen and Willerslev, 2015). Therefore, we aimed to evaluate the ability of untargeted whole-eDNA sequencing to approximate broodstock genomic diversity in a controlled aquaculture setting using the gilthead seabream (*Sparus aurata*) as a model species. The main finding of this study is that eDNA-based whole-genome sequencing and DNA-pool sequencing provided highly consistent estimates of genetic diversity and structure across tanks. These findings suggest that eDNA can reliably replicate the genomic information typically obtained from tissue samples. Sequencing and mapping results confirmed that eDNA-based libraries had lower overall mapping rates compared to DNA pools derived from individual fish tissues. This was expected due to the high degree of DNA fragmentation and the presence of non-target environmental taxa. For this reason, in our work we sequenced 100 bp read lengths for eDNA samples compared with the 150 bp of the pools. This

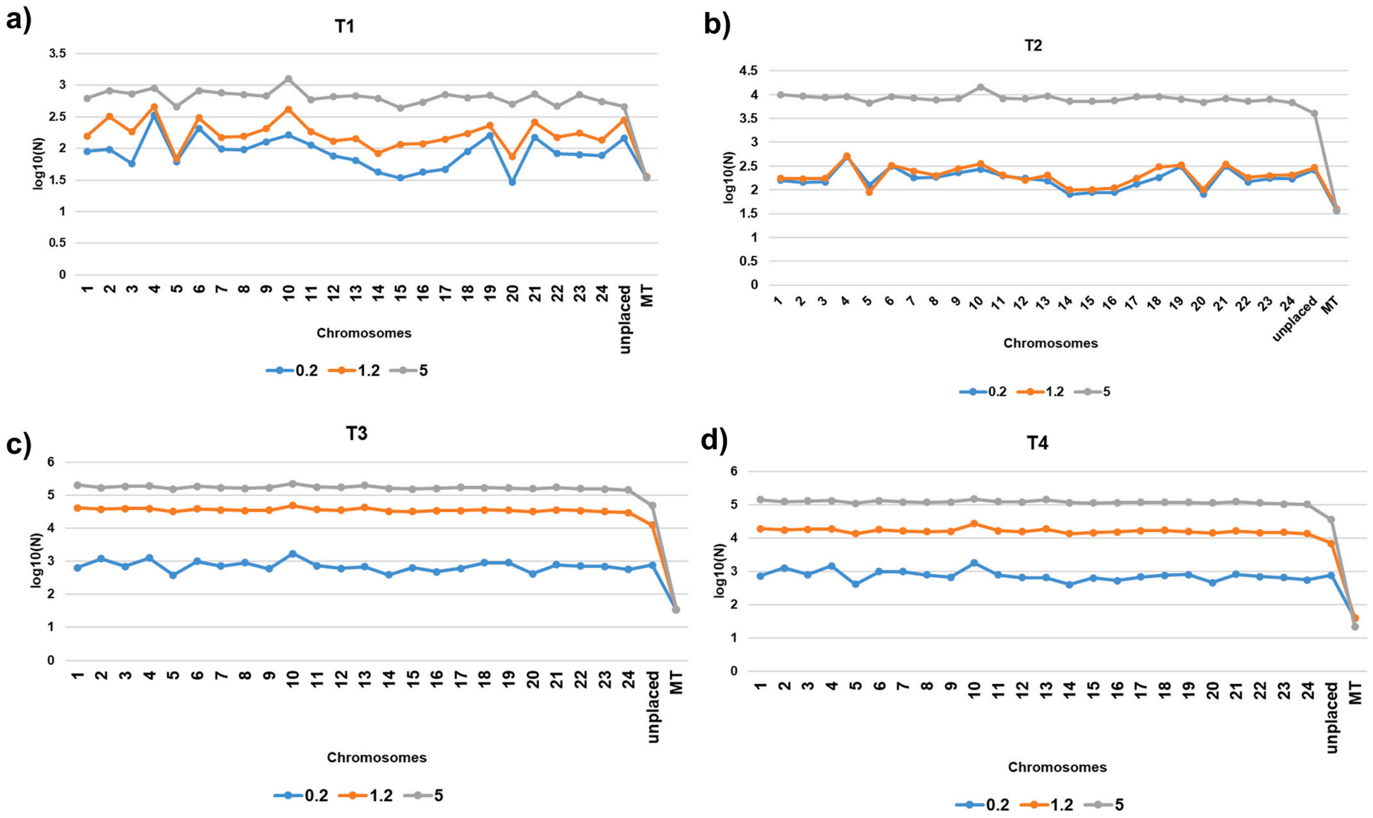

Fig. 3. Chromosomal distribution of high-quality single nucleotide polymorphisms (SNPs) detected from environmental DNA (eDNA) across the four tanks T1-T4 (a–d). Each line represents information derived from eDNA obtained from a different filter pore size (0.2 μm = blue, 1.2 μm = orange, 5 μm = grey). The *x*-axis indicates the 24 chromosomes of the *Sparus aurata* reference genome, with unplaced scaffolds and the mitochondrial genome (MT). The *y*-axis reports the $\log_{10}$ number of SNPs (N). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4

Comparison of within-tank genetic diversity estimates between eDNA samples obtained using different filters (0.2 μm, 1.2 μm, 5 μm) and corresponding DNA pools for each tank (T1–T4). The table reports the number of common SNPs shared with the respective pool, the ΔF_{IS} (difference in F_{IS}-like index between filter and pool) and its standard error (SE), and the corresponding F_{ST} values under both unbalanced and depth-balanced conditions. Confidence intervals (CI95%) refer to the ΔF_{IS} estimates obtained by leave-one-chromosome-out (LOCO) jackknife resampling.

			DEPTH UNBALANCED				DEPTH BALANCED			
TANKS	Filter (μm)	Common SNPs with DNA pool	ΔF_{IS}	SE ΔF_{IS}	ΔF_{IS} Ci95_low, CI95_high	F_{st}	ΔF_{IS}	SE ΔF_{IS}	ΔF_{IS} CI95_low, CI95_high	F_{ST}
	0.2	218	0.09	2.67E-04	-8.68E-02,-8.57E-02	0.23	0.01	3.78E-04	−6.66E-03-5.17E-03	0.26
T1	1.2	1848	0.09	7.72E-05	−9.28E-02,-9.25E-02	0.16	0.01	1.25E-04	−1.42E-02-1.37E-02	0.16
	5	16,061	0.09	5.10E-05	−9.07E-02,-9.05E-02	0.03	0.01	5.78E-05	−1.22E-02-1.20E-02	0.04
	0.2	872	0.08	1.62E-04	−8.46E-02,−8.40E-02	0.31	0.01	2.41E-04	−8.36E-03 − 7.41E-03	0.32
T2	1.2	1452	0.08	1.79E-04	-8.38E-02,-8.31E-02	0.31	0.01	1.86E-04	-7.36E-03-6.63E-03	0.30
	5	22,5875	0.09	8.26E-06	−9.20E-02,-9.20E-02	0.01	0.01	1.02E-05	−1.17E-02-1.17E-02	0.01
	0.2	11,394	0.10	2.75E-05	−9.62E-02,-9.61E-02	0.04	0.01	5.20E-05	−1.32E-02-1.30E-02	0.03
T3	1.2	1,001,514	0.09	3.74E-06	−8.82E-02,-8.81E-02	0.01	0.01	7.03E-06	−7.77E-03-7.74E-03	0.01
	5	3,725,349	0.06	6.89E-06	−6.36E-02,-6.36E-02	0.01	3.91E-03	6.55E-06	−3.93E-03-3.91E-03	0.01
	0.2	12,285	0.09	2.73E-05	−9.35E-02,-9.34E-02	0.07	0.01	4.07E-05	−1.20E-02-1.19E-02	0.06
T4	1.2	493,292	0.09	4.97E-06	−9.08E-02,-9.08E-02	0.01	0.01	6.81E-06	−1.17E-02-1.17E-02	0.01
	5	2,832,887	0.07	1.18E-05	−7.03E-02,-7.02E-02	0.01	1.74E-03	1.41E-05	−1.78E-03 − 1.72E-03	0.01

choice was made to improve recovery of degraded DNA fragments commonly found in aquatic environments, where extracellular DNA is subject to enzymatic degradation, UV radiation, and microbial activity (Barnes and Turner, 2016; Jo et al., 2019; Harrison et al., 2019). However, our findings revealed that, despite potential sources of noise in the eDNA samples, a consistent proportion of reads was aligned to the *S. aurata* reference genome across filters and tanks. This suggests that eDNA retains a detectable genomic signal from the target broodstock population even in complex environmental backgrounds.

Previous studies have shown that moderate changes in filter pore size often have limited effects on overall DNA yield once extraction efficiency is standardised, yet, the type of biological material and information captured by the filter may vary (Li et al., 2018; Majaneva et al., 2018; Putri et al., 2024; Liu et al., 2024). Among the tested filter pore sizes utilised in our study, the 5 μm filter consistently produced the highest number of mapped reads and recovered the largest proportion of gilthead seabream-derived DNA. This suggests that larger pore sizes may be more effective in capturing intact cells and cellular debris rather than free extracellular DNA. Shotgun metagenomic datasets are particularly sensitive to the ratio of target and non-target DNA: a higher proportion

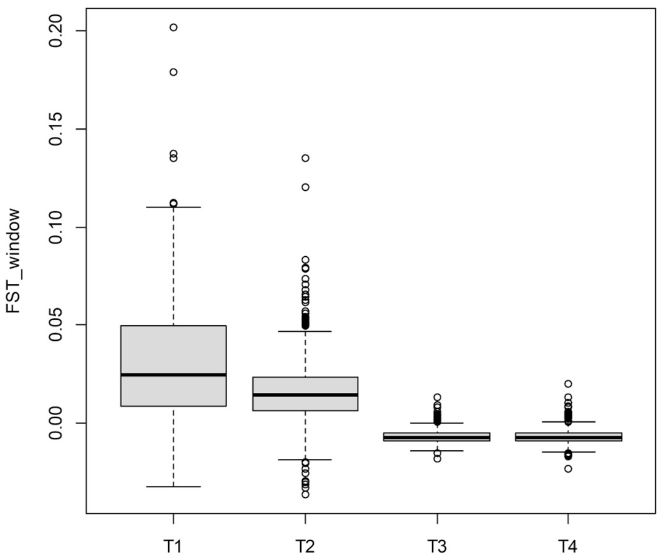

Fig. 4. Window-based F_{ST} for each tank showing distribution of F_{ST} values of the SNPs detected with the 5 μm filter against the DNA pools.

of non-target content reduces the effective depth on the host genome and can bias SNP discovery and allele-frequency estimation. Our findings align with broader metagenomic research indicating that increasing host DNA levels can positively impact sensitivity and quantitative performance (McArdle and Kaforou, 2020). Although the 5 μm filter generally produced the highest recovery of host DNA, variability in host read proportions was observed in all filters among tanks, with differences between T1-T2 and T3-T4. Tank dimensions and water conditions were identical across tanks; however, water sampling for T3 and T4 was conducted while water flow was temporarily stopped. This may have promoted settling of organic particles, such as residual feed, thereby reducing suspended host-derived eDNA available for filter capture, as eDNA often associates with cellular debris or suspended particles. Reduced water mixing may promote the settling of particle-associated eDNA and influence its distribution in the water column, potentially leading to differences in detected eDNA patterns (Sassoubre et al., 2016; Nevers et al., 2020).

Genome-wide heterozygosity and F-statistics were consistent across tanks but varied with pore size. F_{IS}-like estimates from 5 μm eDNA were most similar to those from DNA pools, while 0.2 and 1.2 μm fractions showed greater divergence from that of the pools, especially when non-target DNA was more prevalent. ΔF_{IS}-like and F_{ST} values between eDNA from various filters and DNA pools were generally small, indicating that eDNA captured a consistent genomic signature from the broodstock, with greatest similarity seen as depth and coverage increased. The similarity between eDNA derived from filters and DNA pools also increased in terms of ΔF_{IS}-like when adjusting for sequencing depth. Despite minor increases in some F_{ST} estimates after depth balancing, absolute values remained extremely low in all comparisons. This may be due to down-sampling reducing per-site coverage, leading to inflated binomial sampling variance in allele frequency estimates. Considering that F_{ST} quantifies the variance in allele frequencies among populations, this inflation can cause small positive shifts in F_{ST} even when true differentiation is zero (Bhatia et al., 2013; Tilk et al., 2019). Moreover, balancing coverage also helps remove depth-related biases that could mask subtle differences among samples, making minor stochastic variation more apparent. The filtering process itself may impact the retained SNP set by excluding loci with insufficient coverage or minor allele counts after balancing, potentially enriching for slightly more variable sites. However, this effect likely affects both balanced and unbalanced datasets, contributing to the observed $F_{ST} \neq 0$. At very low differentiation levels, F_{ST} estimates are known to fluctuate symmetrically around zero due to finite-sample noise (Miller et al., 2008). Similar artificially inflated differentiation patterns at low coverage have been reported in other DNA-pooled or low-depth sequencing datasets (Tilk et al., 2019; Gautier et al., 2024).

The between-tank comparisons showed that genome-wide genetic differentiation among broodstock groups was negligible, with pairwise F_{ST} estimates consistently close to zero across all datasets (tissue pools and all eDNA filters). This indicates that the four tanks represent a single, well-mixed broodstock population, with no detectable substructuring or drift among groups. This finding is consistent with the hatchery management practices, which do not enforce family segregation, and supports the expectation that the observed genomic uniformity is biologically real rather than a methodological artifact.

When analysing the composition of non-gilthead seabream reads, it was found that eDNA from the 0.2 μm and 1.2 μm filters contained a significantly higher proportion of bacterial and archaeal DNA (e.g., *Pseudomonadota*, *Bacteroidota*, *Mycoplasmatota*, *Thermoproteota*). This finding was expected due to their higher ability to capture finer particles and cell-free nucleic acids. The higher proportion of eukaryotic species detected in the eDNA from 5 μm filters (e.g., *Chordata*, *Ascomycota*, *Streptophyta*, *Arthropoda*) may have originated from the ambient ecosystem, tank biofilm, detritus, or organisms present in the water or recirculating system. Additionally, residual uneaten food was visible in the tank, even though the sampling date did not coincide with a feeding day (Izquierdo et al., 2015; Rivera et al., 2022). This effect reflects differences in the composition of the material retained by filters of different pore sizes, rather than the ability of larger filters to retain smaller particles such as bacteria. These findings highlight the

Table 5
Pairwise genome-wide F_{ST} estimates among tanks based on DNA pools and eDNA from filters (0.2, 1.2, and 5 μm). For each comparison (tank pair), the number of SNPs retained after filtering (n.SNPs), the mean F_{ST} value (F_{ST}), and its standard error (SE) obtained by leave-one-chromosome-out (LOCO) jackknife are reported. Negative F_{ST} estimates may occur due to stochastic sampling variance in cases of very low differentiation.

	eDNA(filter 0.2 μm)			eDNA (filter 1.2 μm)			eDNA (filter 5 μm)			DNA pools		
Tank comparisons	No. of SNPs	F_{ST}	SE	No. of SNPs	F_{ST}	SE	No. of SNPs	F_{ST}	SE	No. of SNPs	F_{ST}	SE
1 vs 2	1332	4.28E-03	9.64E-04	1315	6.81E-03	1.29E-03	8451	4.12E-03	5.56E-04	5,408,896	-1.57E-04	6.66E-05
1 vs 3	1595	2.04E-02	8.80E-04	3547	3.21E-02	1.49E-03	17,898	2.25E-02	3.73E-04	5,487,792	2.25E-04	6.94E-05
1 vs 4	1690	1.01E-02	9.23E-04	3742	5.40E-03	9.35E-04	10,876	1.81E-02	4.07E-04	5,603,013	2.08E-03	6.11E-05
2 vs 3	2682	−7.52E-03	6.22E-04	3214	3.27E-02	1.47E-03	241,376	5.70E-03	1.69E-04	5,420,344	−5.70E-04	5.66E-05
2 vs 4	2539	−7.50E-03	6.59E-04	2823	1.12E-02	1.31E-03	177,487	6.53E-03	1.66E-04	5,551,984	5.37E-05	7.03E-05
3 vs 4	6451	−1.87E-02	3.37E-04	177,643	−1.77E-02	1.31E-04	2,267,032	−8.54E-03	1.31E-04	5,637,907	5.40E-04	8.25E-05

importance of considering water quality in tanks when studying fish population genomic parameters using eDNA, particularly with respect to non-target organic components. At the same time, this feature may represent an advantage when adopting a more holistic approach, as eDNA can capture broader biological information from the system. From these observations, it appears evident that filtering 2 L of water from a tank requires substantially less human effort in the field and in the laboratory and induces less stress to fish than individual fin-clip sampling. Moreover, as demonstrated in this study, eDNA-based approaches may provide information not only on the fish present in the tanks but also on the broader biological environment.

5. Conclusions

This study provides the first demonstration, to our knowledge, that untargeted whole-eDNA sequencing can reliably provide indicators of fish genetic diversity traditionally derived from tissue samples in aquatic species. By comparing eDNA-derived and tissue-derived DNA pools from the same broodstock tanks, we showed that eDNA captures genome-wide patterns of heterozygosity and population structure consistent with those obtained through conventional approaches. This represents a proof-of-concept for the use of whole-eDNA as a non-invasive tool for monitoring the genetic status of aquaculture finfish broodstocks, particularly at the population level, where genome-wide patterns can be effectively captured without requiring individual-based sampling. Among the tested filter types, eDNA derived from larger pores (i.e., the 5 μm filters) proved to be the most informative substrate for recovering gilthead seabream genomic information, likely due to the greater efficiency of these filters in retaining intact cells and cellular debris. Environmental DNA datasets also provided complementary insights into the taxonomic composition of non-fish reads, highlighting the potential of this approach to simultaneously characterise both the genetic status of cultivated fish and the surrounding biological environment within rearing systems. Future work should optimise filtering and sequencing depth, test temporal stability of eDNA-derived estimates, and evaluate performance in other rearing systems and species. Once optimised, eDNA monitoring could become a practical tool for broodstock management, helping hatcheries to monitor genetic diversity at the population level, support management decisions, and enhance the sustainability of aquaculture production.

CRediT authorship contribution statement

Francesca Bertolini: Writing – original draft, Supervision, Resources, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Valeria Taurisano:** Writing – original draft, Methodology, Investigation, Data curation. **Samuele Bovo:** Writing – review & editing, Formal analysis, Data curation. **Jacopo Vegni:** Writing – review & editing, Investigation. **Riccardo Napolitano:** Writing – original draft, Methodology, Data curation. **Marco Martinoli:** Writing – review & editing, Methodology. **Fabrizio Capoccioni:** Writing – review & editing, Resources, Conceptualization. **Luca Fontanesi:** Writing – review & editing, Resources, Conceptualization.

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.

Acknowledgments

We thank Carmelita Lobello for technical assistance. We thank Anisa Ribani and Maria Letizia Calabri (University of Bologna) for laboratory assistance. This work received funding from University of Bologna RFO and from internal resources of the Council for Agricultural Research and Economics (CREA).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aquaculture.2026.743980.

Data availability

Data will be made available on request.

References

Andres, K.J., Sethi, S.A., Lodge, D.M., Andrés, J., 2021. Nuclear eDNA estimates population allele frequencies and abundance in experimental mesocosms and field samples. Mol. Ecol. 30, 685–697. https://doi.org/10.1111/mec.15765.

Aslam, M.L., Carraro, R., Sonesson, A.K., Meuwissen, T., Tsigenopoulos, C.S., Rigos, G., Bargelloni, L., Tzokas, K., 2020. Genetic variation, GWAS and accuracy of prediction for host resistance to sparicotyle chrysophrii in farmed gilthead sea bream (sparus aurata). Front. Genet. 11, 594770. https://doi.org/10.3389/fgene.2020.594770.

Barnes, M.A., Turner, C.R., 2016. The ecology of environmental DNA and implications for conservation genetics. Conserv. Genet. 17, 1–17. https://doi.org/10.1007/s10592-015-077.

Benedicenti, O., Måsøy Amundsen, M., Mohammad, S.N., Vrålstad, T., Strand, D.A., Weli, S.C., Patel, S., Sindre, H., 2024. A refinement to eRNA and eDNA-based detection methods for reliable and cost-efficient screening of pathogens in Atlantic salmon aquaculture. PLoS One 19, e0312337. https://doi.org/10.1371/journal.pone.0312337.

Bertolini, F., Ribani, A., Capoccioni, F., Buttazzoni, L., Utzeri, V.J., Bovo, S., Schiavo, G., Caggiano, M., Fontanesi, L., Rothschild, M.F., 2020. Identification of a major locus determining a pigmentation defect in cultivated gilthead seabream (sparus aurata). Anim. Genet. 51, 319–323. https://doi.org/10.1111/age.12890.

Bertolini, F., Ribani, A., Capoccioni, F., Buttazzoni, L., Utzeri, V.J., Bovo, S., Schiavo, G., Caggiano, M., Rothschild, M.F., Fontanesi, L., 2021. A comparative whole genome sequencing analysis identified a candidate locus for lack of operculum in cultivated gilthead seabream (sparus aurata). Anim. Genet. 52, 365–370. https://doi.org/10.1111/age.13049.

Bhatia, G., Patterson, N., Sankararaman, S., Price, A.L., 2013. Estimating and interpreting FST: the impact of rare variants. Genome Res. 23, 1514–1521. https://doi.org/10.1101/gr.154831.113.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformat. 10, 421. https://doi.org/10.1186/1471-2105-10-421.

Chamberlain, S.A., Szöcs, E., 2013. Taxize: taxonomic search and retrieval in r. F1000Res 2, 191. https://doi.org/10.12688/f1000research.2-191.v2.

Charlesworth, D., Willis, J.H., 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10, 783–796. https://doi.org/10.1038/nrg2664 (PMID: 19834483).

Chavanne, H., Janssen, K., Hofherr, J., Contini, F., Haffray, P., Aquatrace Consortium, Komen H., Eg Nielsen, E., Bargelloni, L., 2016. A comprehensive survey on selective breeding programs and seed market in the european aquaculture fish industry. Aquac. Int. 24, 1287–1307. doi:https://doi.org/10.1007/s10499-016-9985-0.

Cossu, P., Scarpa, F., Sanna, D., Lai, T., Dedola, G.L., Curini-Galletti, M., Mura, L., Fois, N., Casu, M., 2019. Influence of genetic drift on patterns of genetic variation: the footprint of aquaculture practices in sparus aurata (teleostei: Sparidae). Mol. Ecol. 28, 3012–3024. https://doi.org/10.1111/mec.15134.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience 10, giab008. https://doi.org/10.1093/gigascience/giab008.

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N., Pfrender, M.E., Bernatchez, L., 2017. Environmental DNA metabarcoding: transforming how we survey animal and plant communities. Mol. Ecol. 26, 5872–5895. doi:https://doi.org/10.1111/mec.14350.

Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T., Minamoto, T., 2017. Environmental DNA analysis for estimating the abundance and biomass of stream fish. Freshwater Biology 62 (1), 30–39. https://doi.org/10.1111/fwb.12846.

Ellegren, H., Galtier, N., 2016. Determinants of genetic diversity. Nat. Rev. Genet. 17, 422–433. https://doi.org/10.1038/nrg.2016.58.

FAO, 2025. The State of Mediterranean and Black Sea Fisheries 2025. General Fisheries Commission for the Mediterranean. Rome. https://doi.org/10.4060/cd7701en.

Ficetola, G.F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., Gielly, L., Lopes, C.M., Boyer, F., Pompanon, F., Rayé, G., Taberlet, P., 2015. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Mol. Ecol. Resour. 15, 543–556. https://doi.org/10.1111/1755-0998.12338.

Gao, B., Liu, P., Li, J., Wang, Q., Han, Z., 2015. Effect of inbreeding on growth and genetic diversity of portunus trituberculatus based on the full-sibling inbreeding families. Aquac. Int. 23, 1401–1410. https://doi.org/10.1007/s10499-015-9892-9.

Gautier, M., Coronado-Zamora, M., Vitalis, R., 2024. Estimating hierarchical f–statistics from pool–seq data. bioRxiv 2024 (11). https://doi.org/10.1101/2024.11.22.624688.

Gjedrem, T., Robinson, N., Rye, M., 2012. The importance of selective breeding in aquaculture to meet future demands for animal protein: a review. Aquaculture 350, 117–129. https://doi.org/10.1016/j.aquaculture.2012.04.008.

Gomes, G.B., Hutson, K.S., Domingos, J.A., Chung, C., Hayward, S., Miller, T.L., Jerry, D. R., 2017. Use of environmental DNA (eDNA) and water quality data to predict protozoan parasites outbreaks in fish farms. Aquaculture 479, 467–473. https://doi.org/10.1016/j.aquaculture.2017.06.021.

Harrison, J.B., Sunday, J.M., Rogers, S.M., 2019. Predicting the fate of eDNA in the environment and implications for studying biodiversity. Proc. Biol. Sci. 286, 20191409. https://doi.org/10.1098/rspb.2019.1409.

Hivert, V., Leblois, R., Petit, E.J., Gautier, M., Vitalis, R., 2018. Measuring genetic differentiation from pool-seq data. Genetics 210, 315–330. https://doi.org/10.1534/genetics.118.300900.

Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., Jenkins, T.L., Selly, S.L.C., Martin, S.A.M., Stevens, J.R., Santos, E.M., Davie, A., Robledo, D., 2020. Harnessing genomics to fast-track genetic improvement in aquaculture. Nat. Rev. Genet. 21, 389–409. https://doi.org/10.1038/s41576-020-0227-y.

Hughes, A.R., Inouye, B.D., Johnson, M.T., Underwood, N., Vellend, M., 2008. Ecological consequences of genetic diversity. Ecol. Lett. 11, 609–623. https://doi.org/10.1111/j.1461-0248.2008.01179.x.

Izquierdo, M.S., Turkmen, S., Montero, D., Zamorano, M.J., Afonso, J.M., Karalazos, V., Fernández-Palacios, H., 2015. Nutritional programming through broodstock diets to improve utilization of very low fishmeal and fish oil diets in gilthead sea bream. Aquaculture 449, 18–26. https://doi.org/10.1016/j.aquaculture.2015.03.032.

Jo, T., Murakami, H., Yamamoto, S., Masuda, R., Minamoto, T., 2019. Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. Ecol. Evol. 9, 1135–1146. https://doi.org/10.1002/ece3.4802.

Kasmi, Y., Nunez-Riboni, I., Blancke, T., Möckel, B., Bernreuther, M., Stransky, C., Hanel, R., 2025. Fish diversity assessment and semi-quantitative biomass estimation through metabarcoding of environmental DNA. Ecol. Indic. 173, 113406. https://doi.org/10.1016/j.ecolind.2025.113406.

Kofler, R., Pandey, R.V., Schlötterer, C., 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27, 3435–3436. https://doi.org/10.1093/bioinformatics/btr589.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasism, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Li, J., Lawson Handley, Read, D.S., Hänfling, B., 2018. The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding. Molecular Ecology Resources 18 (5), 1102–1114. https://doi.org/10.1111/1755-0998.12899.

Lind, C.E., Ponzoni, R.W., Nguyen, N.H., Khaw, H.L., 2012. Selective breeding in fish and conservation of genetic resources for aquaculture. Reprod. Domest. Anim. 47, 255–263. https://doi.org/10.1111/j.1439-0531.2012.02084.x.

Liu, Q., Tan, J., Wang, M., Xin, N., Qi, R., Wang, H., 2024. Optimization of pore size and filter material for better enrichment of environmental DNA. Frontiers in Environmental Science 12, 1422269.

Lyu, D., Fu, Q., Liu, M., Hu, Y., Lyu, G., Wang, W., 2025. Genomic inbreeding evaluation and its effects on phenotype and selection accuracy of resistance to *Edwardsiella tarda* in turbot (*Scophthalmus maximus*). Aquaculture, 743178. https://doi.org/10.1016/j.aquaculture.2025.743178.

Majaneva, M., Diserud, O.H., Eagle, S.H., Boström, E., Hajibabaei, M., Ekrem, T., 2018. Environmental DNA filtration techniques affect recovered biodiversity. Sci. Rep. 8, 4682. https://doi.org/10.1038/s41598-018-23052-8.

Maroso, F., Gkagkavouzis, K., De Innocentiis, S., Hillen, J., do Prado, F., Karaiskou, N., Taggart, J.B., Carr, A., Nielsen, E., Triantafyllidis, A., Bargelloni, L., Aquatrace Consortium, 2021. Genome-wide analysis clarifies the population genetic structure of wild gilthead sea bream (*sparus aurata*). PLoS One 16, e0236230. https://doi.org/10.1371/journal.pone.0236230.

McArdle, A.J., Kaforou, M., 2020. Sensitivity of shotgun metagenomics to host DNA: abundance estimates depend on bioinformatic tools and contamination is the main issue. Access. Microbiol. 2, acmi000104. https://doi.org/10.1099/acmi.0.000104.

Miller, J.R., Wood, B.P., Hamilton, M.B., 2008. F(ST) and Q(ST) under neutrality. Genetics 180, 1023–1037. https://doi.org/10.1534/genetics.108.092031.

Nevers, M.B., Przybyla-Kelly, K., Shively, D., Morris, C.C., Dickey, J., Byappanahalli, M. N., 2020. Influence of sediment and stream transport on detecting a source of environmental DNA. PLoS One 15, e0244086. https://doi.org/10.1371/journal.pone.0244086.

O'Reilly, G.D., Manlik, O., Vardeh, S., Sinclair, J., Cannell, B., Lawler, Z.P., Sherwin, W. B., 2024. A new method for ecologists to estimate heterozygote excess and deficit for multi-locus gene families. Ecol. Evol. 14, e11561. https://doi.org/10.1002/ece3.11561.

Peñaloza, C., Manousaki, T., Franch, R., Tsakogiannis, A., Sonesson, A.K., Aslam, M.L., Allal, F., Bargelloni, L., Houston, R.D., Tsigenopoulos, C.S., 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.

Putri, R.E., Vrouwenvelder, J.S., Farhat, N., 2024. Enhancing the DNA yield intended for microbial sequencing from a low-biomass chlorinated drinking water. Front. Microbiol. 15, 1339844. https://doi.org/10.3389/fmicb.2024.1339844.

R Core Team R, 2024. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.https://www.R-project.org/.

Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R., Gough, K.C., 2014. The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. J. Appl. Ecol. 51, 1450–1459. https://doi.org/10.1111/1365-2664.12306.

Rivera, S.F., Rimet, F., Vasselon, V., Vautier, M., Domaizon, I., Bouchez, A., 2022. Fish eDNA metabarcoding from aquatic biofilm samples: methodological aspects. Mol. Ecol. Resour. 22, 1440–1453. https://doi.org/10.1111/1755-0998.13568.

Rourke, M.L., Fowler, A.M., Hughes, J.M., Broadhurst, M.K., DiBattista, J.D., Fielder, S., Wilkes Walburn, J., Furlan, E.M., 2022. Environmental DNA (eDNA) as a tool for assessing fish biomass: a review of approaches and future considerations for resource surveys. Environ. DNA 4, 9–33. https://doi.org/10.1002/edn3.185.

Sassoubre, L.M., Yamahara, K.M., Gardner, L.D., Block, B.A., Boehm, A.B., 2016. Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. Environ. Sci. Technol. 50, 10456–10464. https://doi.org/10.1021/acs.est.6b03114.

Sigsgaard, E.E., Nielsen, I.B., Bach, S.S., Lorenzen, E.D., Robinson, D.P., Knudsen, S.W., Pedersen, M.W., Jaidah, M.A., Orlando, L., Willerslev, E., Møller, P.R., Thomsen, P. F., 2016. Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. Nat. Ecol. Evol. 1, 4. https://doi.org/10.1038/s41559-016-0004.

Sigsgaard, E.E., Jensen, M.R., Winkelmann, I.E., Møller, P.R., Hansen, M.M., Thomsen, P. F., 2019. Population-level inferences from environmental DNA-current status and future perspectives. Evol. Appl. 13, 245–262. https://doi.org/10.1111/eva.12882.

Sonesson, A.K., Hallerman, E., Humphries, F., Hilsdorf, A.W.S., Leskien, D., Rosendal, Bartley D., Hu, X., Garcia Gomez, R.K., Mair, G.C., 2023. Sustainable management and improvement of genetic resources for aquaculture. J. World Aquacult. Soc. 54, 364–396. https://doi.org/10.1111/jwas.12968.

Taberlet, P., Bonin, A., Zinger, L., Coissac, E., 2018. Environmental DNA: for biodiversity research and monitoring. Oxford University Press. https://doi.org/10.1093/oso/9780198767220.001.0001.

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., Kawabata, Z.I., 2012. Estimation of fish biomass using environmental DNA. PLoS One 7, e35868. https://doi.org/10.1371/journal.pone.0035868.

Tave, D., 1999. Inbreeding and brood stock management, 392. Food & Agriculture Org.

Thomsen, P.F., Willerslev, E., 2015. Environmental DNA–an emerging tool in conservation for monitoring past and present biodiversity. Biol. Conserv. 183, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019.

Tilk, S., Bergland, A., Goodman, A., Schmidt, P., Petrov, D., Greenblum, S., 2019. Accurate allele frequencies from ultra-low coverage pool-seq samples in evolve-and-resequence experiments. G3 Genes|Genomes|Genetics 9, 4159–4168. https://doi.org/10.1534/g3.119.400755.

Villanueva, B., Fernández, A., Peiró-Pastor, R., Peñaloza, C., Houston, R.D., Sonesson, A. K., Tsigenopoulos, C.S., Bargelloni, L., Gamsız, K., Karahan, B., Gökçek, E.Ö., Fernández, J., Saura, M., 2022. 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. Aquacult. Rep. 24, 101145. https://doi.org/10.1016/j.aqrep.2022.101145.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x.

Weitemier, K., Penaluna, B.E., Hauck, L.L., Longway, L.J., Garcia, T., Cronn, R., 2021. Estimating the genetic diversity of pacific salmon and trout using multigene eDNA metabarcoding. Mol. Ecol. 30, 4970–4990. https://doi.org/10.1111/mec.15811.

Westgaard, J.I., Præbel, K., Arneberg, P., Ulaski, B.P., Ingvaldsen, R., Wangensteen, O.S., Johansen, T., 2024. Towards eDNA informed biodiversity studies–comparing water derived molecular taxa with traditional survey methods. Prog. Oceanogr. 222, 103230. https://doi.org/10.1016/j.pocean.2024.103230.

Wright, S., 1951. The genetical structure of populations. Ann. Eugenics 15, 323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.

Yu, X., Bastiaansen, J.W., Gulzari, B., Camara, M., Mulder, H.A., Komen, H., Martien, A. M.G., Megens, H.J., 2024. Genome-wide association analyses reveal genotype-by-environment interactions of growth and organ weights in gilthead seabream (sparus aurata). Aquaculture 589, 740984. https://doi.org/10.1016/j.aquaculture.2024.740984.