



OPEN

Genetic patterns in *Mugil cephalus* and implications for fisheries and aquaculture management

Piero Cossu^{1,2}✉, Laura Mura³, Fabio Scarpa^{1,2}, Tiziana Lai¹, Daria Sanna^{1,4}, Ilenia Azzena^{2,4}, Nicola Fois³ & Marco Casu^{1,2}

Exploitation of fisheries and aquaculture practices are exposing marine fish populations to increasing genetic risks. Therefore, the integration of genetic information into fisheries and aquaculture management is becoming crucial to ensure species' long-term persistence. The raising commercial value of grey mullet (*Mugil cephalus*) and its roe represents a growing challenge to the sustainable management of this economically important fishery resource. Here, microsatellites were used to investigate patterns of genetic variation in a Mediterranean area that harbor flourishing fisheries and practice semi-intensive farming of grey mullet. Genetic diversity within populations is smaller than values reported in previous studies as a result of the lower polymorphism displayed by the new microsatellite loci. Lack of genetic structuring points to the existence of a unique genetic stock, which is consistent with the species' high dispersal capabilities. Nonetheless, differences in local population effective size as well as the excess of related individuals do not completely fit the picture of a large panmictic population. Baseline genetic information here gathered will allow to set up the genetic monitoring of regional fish stocks, which is needed to assess the impact of both harvesting and aquaculture on the genetic integrity of *Mugil cephalus* wild populations.

The increase of fish consumption during the last decades raised concerns about the overexploitation of marine resources¹. The need to avoid the excessive erosion of fish stocks boosted research in aquaculture, whose production nowadays is comparable to overall captures². Aquaculture fulfils different goals, one of which aims to replenish or just increase the biomass of wild stocks by sea ranching, stock restoration and stock enhancement³. The supplementation of wild stocks with early-generation captive fish may induce a reduction of total population effective size in the supplemented wild population if they are the offspring of a handful of breeders⁴. This artificial bottleneck, known as the Ryman–Laikre effect⁵ may thus lead to the loss of genetic diversity in the combined captive-wild system by increasing inbreeding and random genetic drift⁶. The negative effects of supplementation will be enhanced in marine species subject to variance in reproductive success (recruitment sweepstake effect⁷): in these cases, even a limited reduction of population size can result in the loss of low-frequency alleles that can be important for the adaptation to environmental changes³. Therefore, protection of genetic diversity within- and amongst-population should deserve high priority in the planning and implementation of supplementation programmes to preserve species' adaptive and evolutionary potential and thus their long-term persistence³.

The grey flathead mullet (*Mugil cephalus* L.), also commonly referred to as the striped mullet (and henceforth the grey mullet), is a world-wide distributed coastal fish species that spawns offshore and uses estuarine and brackish-water environments as nursery habitats⁸. Early larval stages passively disperse by drifting in ocean currents and then move onshore at the post-flexion larval stage, where they inhabit the surf zone as schools of early juveniles⁹. After about a month spent in the sea, fry colonises coastal ponds, lagoons, estuaries and sometimes the adjoining river catchments¹⁰. Adult individuals migrate back to sea, where they show highly variable dispersal capabilities, whose extent may range from 32 to 700 km⁹. The species is an important economical resource worldwide and in Mediterranean waters⁹, where it is captured during the spawning migration to harvest the egg roe, which is salted and dried to be sold as a delicatessen¹¹. In many parts of the World, the increased demand for mullet roe raised the species' commercial value, which has been called “the grey gold” by fishermen⁹. Therefore, the grey mullet is often stocked in brackish coastal lagoons to improve fish yield, raised in commercial freshwater ponds, and it has also been introduced into inland freshwater lakes and reservoirs to develop new fisheries⁹.

¹Department of Sciences for Nature and Environmental Resources, University of Sassari, Sassari, Italy. ²Department of Veterinary Medicine, University of Sassari, Sassari, SS, Italy. ³Dipartimento per la Ricerca Nelle Produzioni Animali, Agris Sardegna, Olmedo, SS, Italy. ⁴Department of Biomedical Sciences, University of Sassari, Sassari, Italy. ✉email: picossu@uniss.it

Notwithstanding the growing interest in fisheries and aquaculture of *M. cephalus*, few studies investigated the distribution of genetic diversity at the spatial and temporal scales that are relevant to the management of this natural resource. Perhaps due to its high dispersal potential, genetic studies focussed mainly on investigating large-scale genetic structuring and assessing whether *M. cephalus* is a single, cosmopolitan species or a complex of cryptic species (e.g.^{12–15}). Consistent with this picture, an overall lack of genetic structuring was reported at regional and basin-wide spatial scales^{16,17}. However, the use of more variable markers as microsatellites evidenced genetic structuring within the Mediterranean Sea, roughly matching well-known biogeographical barriers to dispersal, as well as an isolation by distance pattern¹⁸.

Therefore, one knowledge gap to be filled concerns the distribution of genetic diversity on local and regional scales in those areas where the grey mullet is, or can be, a commercially valuable resource for fisheries and aquaculture. This is one of the main goals envisioned in the responsible genetic approach to fisheries and aquaculture practises: the protection of the genetic integrity of wild populations³. The distribution of genetic diversity in wild populations represents the baseline information to plan and implement management strategies that minimise genetic risks for natural populations, thus preserving the long-term species persistence¹⁹. Here, we focus our attention on several populations of *M. cephalus* from Sardinia island (Fig. 1), where the species represents an important economic resource: in 2017, 156 out of 401 tonnes of total captures from Italy were indeed from Sardinian coastal waters². Furthermore, to increase or maintain yields over years, mullets are cultured in semi-intensive, extensive systems either to be marketed for direct human consumption or to harvest mullet roe¹¹. However, in recent years the amount of fish captured by fishermen covered only a small part of the growing market demand for mullet roe, forcing Sardinian manufacturers to purchase an increasing amount of frozen egg roe from fishing areas other than the Mediterranean Sea²⁰.

The aim of this study was to investigate patterns of genetic diversity in exploited populations of *M. cephalus*, which will provide the baseline information needed to enforce a responsible genetic approach for this commercially valuable resource. Genetic data will help to plan a more effective fishery management, as well as sea ranching or restocking programmes that minimise potential genetic risks arising from the inter-breeding amongst wild and farmed individuals. To achieve this goal, we set up multiplex PCR reactions based on available microsatellite loci for this species: these markers are a suitable tool to carry out regional- and fine-scale population genetic studies in the grey mullet²¹.

Results

Marker validation. All microsatellite loci were polymorphic at the 5% level across all populations (Supplementary Table S3), with the number of alleles ranging from 3 at locus Mce11 to 32 at locus Mcs2DM (Supplementary Table S4). Overall, 7 out of 14 loci departed from HWP: 5 loci showed a heterozygote deficit in at least one population (Mce22, Mce6, Mce24, Mce25 and Mce3), whereas Mcs16DM and Mcs2FH displayed a heterozygote excess at MIS and PPI, respectively. Based on the LM method, two loci (Mce25 and Mce3) displayed a significant departure from HWP ($P < 0.05$), which cannot be explained by chance alone. The presence of null alleles was the most likely reason for the heterozygote deficit observed at Mce25 and Mce3 (Supplementary Table S5): as their frequency exceeded the threshold of 0.08 in both detection methods, these loci were excluded from downstream analyses. Another locus, Mce22 showed null allele frequencies > 0.08 at MIS and was thus removed from the final dataset.

Overall, 96 out of 364 locus pairs showed LD, which largely exceeded those expected by chance, based on a cumulative binomial distribution (27 locus pairs, $P = 0.05$). The locus pairs showing LD were unevenly distributed across populations: 8, 70, 10 and 8 pairs at CAB, MIS, PPI and TOR, respectively. After applying the FDR correction, the number of locus pairs showing LD was nearly halved (58), but still 54 pairs were found in MIS versus one or two pairs in the other populations, thus suggesting that population level processes such as tiny population size might be responsible for high levels of apparent LD in this population. Overall, 15 locus pairs showed LD in at least two populations, evidencing significant departures according to the LM method ($P < 0.05$). However, most of such locus pairs involved loci that did not show LD in former studies^{18,21}; therefore, it seems unlikely that LD was due to true non-random associations amongst the same loci in the present study. Nevertheless, to be as conservative as possible, one further locus (Mcs16DM) that showed LD with locus Mcs1EH in all populations was cautiously excluded from downstream analyses. Finally, no locus was detected as a candidate for selection by all the outlier detection methods simultaneously (Supplementary Figs. S1, S2 and S3). The locus Mce24 was detected as a potential outlier for balancing selection by the Bayesian method but not the other two methods (Supplementary Fig. S1). In contrast, the *LnRH* statistic detected the loci Mcs2DM and Mcs15CM as potential candidates for balancing selection at 95% level but they lie within the expected distribution for neutral alleles after applying the FDR correction (Supplementary Fig. S2). Instead, no locus was detected as an outlier for either divergent or balancing selection by the *FDIST2* approach (Supplementary Fig. S3). Therefore, considering loci detected by only one method as false positives, all loci were subsequently deemed as selectively neutral. Based on these results, 10 out of 14 loci were retained for downstream analyses.

Genetic diversity and population effective size. The number of alleles (A), the allelic richness (A_R), the mean observed and expected heterozygosity (H_O and H_E , respectively), and the inbreeding coefficient (F_{IS}) are summarised in Table 1. Values of H_E were similar across samples with the highest value found in PPI (mean $H_E = 0.81 \pm 0.02$); this population also showed the largest allelic richness ($A_R = 13.66 \pm 0.87$). Nonetheless, pairwise Wilcoxon rank-sum tests did not evidence differences amongst any population pair at both metrics of genetic diversity within populations ($P > 0.05$).

The heterozygosity excess test did not evidence signatures of recent population declines (Wilcoxon sign-rank test, $P > 0.05$): results were not affected by varying model parameters, thus showing to be robust to different model

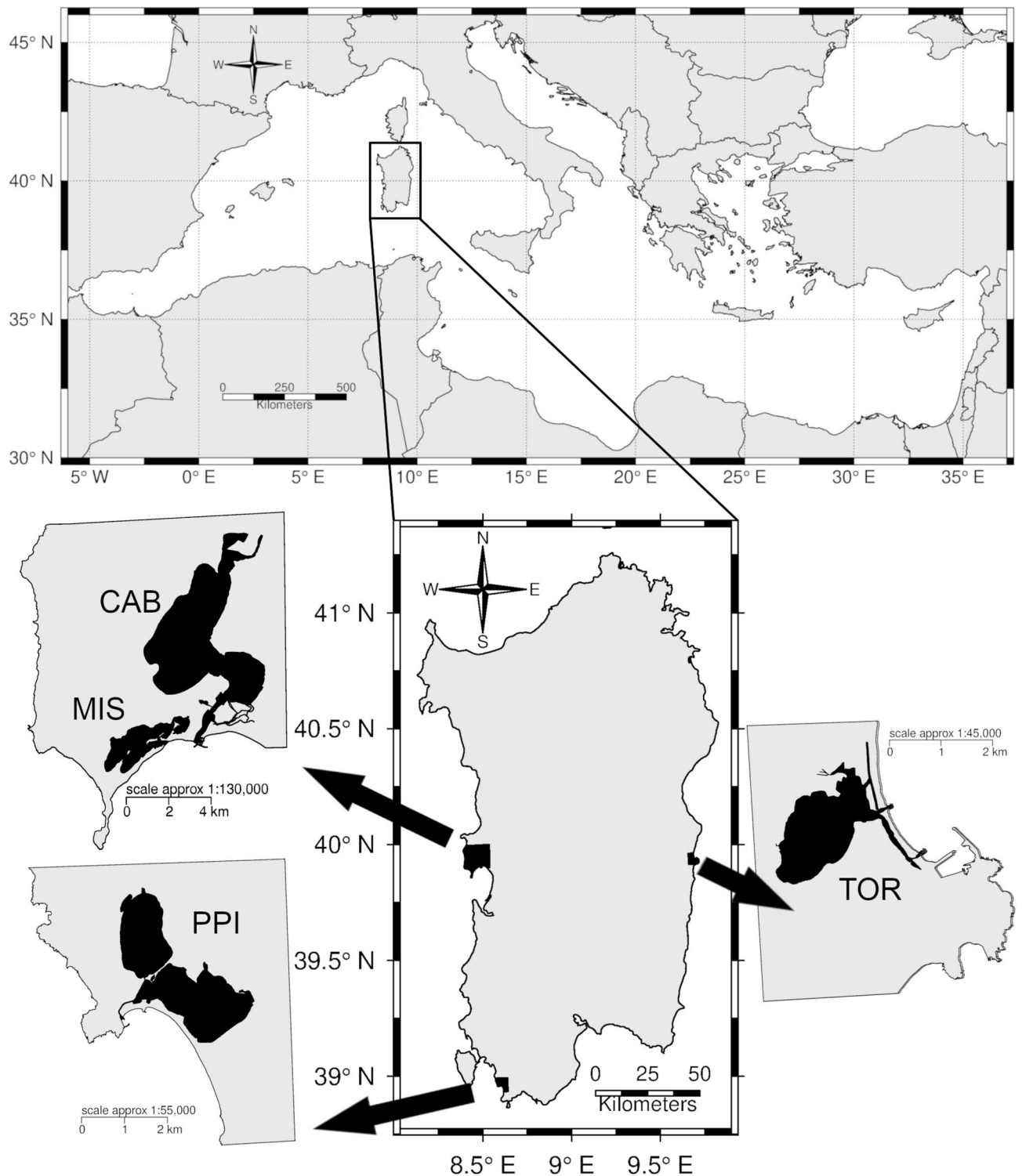


Figure 1. Distribution map of the sampling areas. Note the different scale among coastal ponds. CAB Cabras pond, MIS Mistras pond, PPI Porto Pino ponds, TOR Tortoli pond. This map was created using R 3.5.2 (<https://www.r-project.org/>) and then modified using the free software GIMP 2.10.2 (<https://www.gimp.org/>).

assumptions (Table 2). All populations displayed finite mean estimates of contemporary population effective size, albeit only one population (MIS) also showed a finite N_e estimate for the upper bound of the 95% CI (Table 2). The latter population was characterised by the smallest N_e estimate and the 95% CI of population effective size did not overlap with those of the other populations. Remarkably, CAB, which is the geographically closest population to MIS showed the largest mean population effective size ($N_e = 756$) with a lower bound ($N_e = 131$) that is more than three times larger than the upper bound estimated for the latter ($N_e = 35$).

	N	A ± SE	A _R ± SE	H _E ± SE	H _O ± SE	F _{IS} (95% CI)
CAB	52	15.60 ± 0.85	12.64 ± 0.70	0.79 ± 0.02	0.78 ± 0.02	-0.006 (-0.039, 0.028)
MIS	34	14.10 ± 1.02	12.14 ± 0.85	0.79 ± 0.03	0.81 ± 0.03	-0.026 (-0.075, 0.022)
PPI	48	16.90 ± 1.09	13.66 ± 0.87	0.81 ± 0.02	0.83 ± 0.02	-0.031 (-0.050, -0.006)
TOR	50	15.60 ± 0.89	12.83 ± 0.73	0.79 ± 0.03	0.79 ± 0.03	0.003 (-0.031, 0.037)

Table 1. Summary statistics of within population genetic variation averaged over loci for each population. CAB, Cabras pond; MIS, Mistras pond; PPI, Porto Pino ponds; TOR, Tortoli pond; N, sample size; A, number of alleles; A_R, allelic richness; H_E, expected heterozygosity; H_O, observed heterozygosity; F_{IS}, inbreeding coefficient; SE, Standard Error; 95% CI, 95% Confidence Interval.

	Population effective size	Heterozygosity excess test		
	N _e LD	P _{TPM70}	P _{TPM80}	P _{TPM90}
CAB	756 (131-∞)	0.91	0.95	0.99
MIS	17 (9-35)	0.96	0.99	0.99
PPI	289 (93-∞)	0.84	0.93	0.98
TOR	554 (158-∞)	0.69	0.88	0.90

Table 2. Mean current effective population size and Heterozygosity excess test for population size decline. Contemporary population effective size was estimated using the Linkage Disequilibrium method (N_e^{LD}); 95% confidence interval (CI) of N_e based on Jackknifing over loci is reported within brackets. Recent population bottlenecks were assessed testing for the probability of heterozygosity excess related to the heterozygosity expected at mutation-drift equilibrium (H_E > H_{EQ}). Equilibrium heterozygosity was estimated by a Two-Phase Mutation Model with increasing proportion of single-step mutations: 70% (P_{TPM70}), 80% (P_{TPM80}) and 90% (P_{TPM90}). Populations are abbreviated as in Table 1.

	CAB	MIS	PPI	TOR
CAB	-	(-0.0083, 0.0065)	(-0.0048, 0.0066)	(-0.0047, 0.0073)
MIS	-0.0021	-	(-0.0074, 0.0079)	(-0.0067, 0.0095)
PPI	0.0003	-0.0007	-	(-0.0057, 0.0071)
TOR	0.0006	0.0003	-0.0001	-

Table 3. Pairwise population differentiation estimated using Weir and Cockerham's Θ. Observed values are reported below the diagonal. Lower and upper limits of 95% confidence intervals based on 10,000 bootstraps are reported within brackets above the diagonal. Values outlined in bold indicate population pairs showing significant genetic divergence based on the Fisher's exact test for genetic homogeneity of genotype frequencies. Due to multiple testing, probability values were adjusted using the False Discovery Rate (FDR) method⁵². Populations are abbreviated as in Table 1.

Genetic structure and kinship analysis. Since results based on D_{EST} values (Supplementary Table S6) mirrored those resulting from F_{ST}, only the latter are reported hereafter. Global F_{ST} did not show significant genetic differentiation amongst populations as its confidence interval straddled zero (F_{ST} = -0.0002; CI₉₅ = -0.004, 0.004). Consistent with this picture, the exact G test of population differentiation did not evidence heterogeneous genotypic frequencies amongst samples ($\chi^2 = 29.06$ with 20 degrees of freedom, P = 0.086). Nor pairwise F_{ST} values nor exact G tests after correction for multiple testing (FDR method) evidenced genetic divergence amongst any population pair (Table 3). The simulation of pseudo-datasets with different levels of population differentiation indicated that the actual dataset reached almost 95% power to detect genetic divergence when F_{ST} = 0.0025 (Supplementary Fig. S4). When F_{ST} ≥ 0.005, the power of detecting population differentiation was perfect; moreover, dataset was unlikely to infer genetic heterogeneity when it is not true nor was too conservative (α -error = 0.024).

Bayesian clustering clearly supported K = 1 as the best clustering solution (Supplementary Table S7). Given that Bayesian clustering methods perform poorly when F_{ST} ≤ 0.01 regardless the number of markers used²², we cannot rule out that this outcome might reflect lack of resolving power rather than true panmixia. Using ponds as predefined groups in the Discriminant Analysis of Principal Components (DAPC), 70 principal components were retained based on the cross-validation procedure. The ordination plot of the first two discriminant functions did not highlight a stark separation among individuals from different ponds (Supplementary Fig. S5). Notwithstanding DAPC try to minimize differences within groups, individuals sampled in the same pond were as scattered as, or more scattered than, those from different ponds.

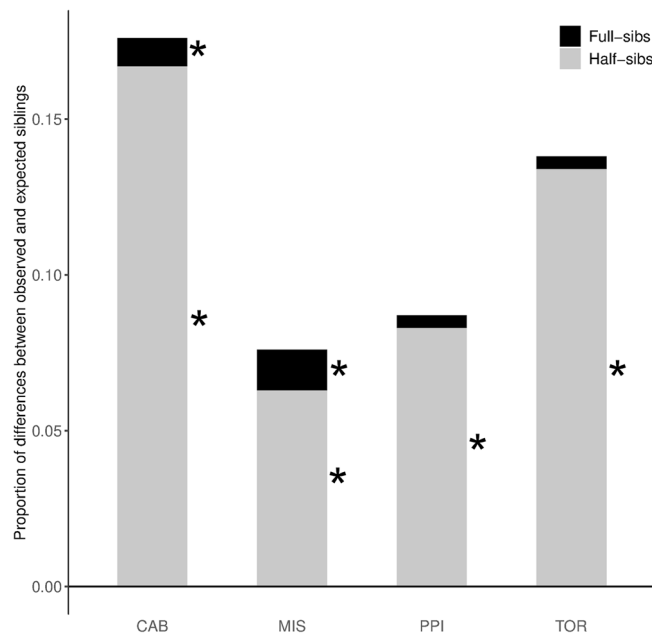


Figure 2. Difference between expected and observed proportions of full- and half-siblings, represented by black and grey bars, respectively. Asterisks mark differences that are significantly larger than those expected by chance. This figure was created using R 3.5.2 (<https://www.r-project.org/>) and then modified using the free software GIMP 2.10.2 (<https://www.gimp.org/>).

All the populations considered in the present study showed an overabundance of half-siblings (Fig. 2). The largest difference between observed and expected half-siblings was observed at CAB, which also displayed a proportion of full-siblings larger than expected. Though the excess of half-siblings at MIS was smaller than that observed in other populations, nevertheless the former showed a larger excess of full-siblings than CAB (Fig. 2). Nevertheless, the genotype sharing method showed an overabundance of related individuals compared with those observed using the maximum likelihood (ML) pairwise estimator (Supplementary Table S8). Whilst the proportion of full-siblings detected by both methods was similar in all populations but CAB, the genotype sharing method detected 4–5 times more half-siblings than the ML method.

Discussion

In the present study, microsatellite data were used to gather the baseline genetic information that is necessary for applying a responsible genetic approach³ to fisheries and aquaculture-based management of *Mugil cephalus*. This approach will help preserve the genetic integrity of natural populations according to both FAO²³ and the Convention on Biological Diversity (Aichi 2010, <http://www.cbd.int/sp/>) guidelines. It will also complement management strategies aimed at ensuring the sustainable exploitation of this commercially valuable resource.

Overall, patterns of genetic diversity are consistent with former studies^{17,18}, even though observed heterozygosity is on average smaller in the present study (Table 1). Such discrepancy reflects the lower polymorphism of some microsatellite loci used in the present study, which are less variable than those developed by Miggiano et al.²¹, rather than small population size or recent population declines (Table 2). In fact, if we consider only the latter markers (Supplementary Table S4), levels of Genetic diversity are as high as those found by the previous studies. By and large, levels of genetic diversity here reported are within the range observed in other species of mullets^{14,24,25}.

Based on simulations, power of detecting genetic differentiation is very close to 95% when $F_{ST} = 0.0025$ (Supplementary Fig. S4), which is four times smaller than the threshold that is usually adopted for genetic stock identification ($F_{ST} = 0.01$,^{26,27}). Therefore, the panel of microsatellites is suitable to detect population units relevant to fishery management. Accordingly, the lack of genetic differentiation observed amongst the four populations (Table 3), the Bayesian clustering and the DAPC outcomes (Supplementary Table S7 and Figure S5) indicate that the four samples analysed in the present study may belong to a unique genetic stock. This result is consistent with Durand et al.'s¹⁸ findings: in the Mediterranean, genetic divergence was found across major oceanographic fronts, whereas populations were genetically homogeneous within basins. These results are also consistent with the species' high dispersal capabilities: migrations of more than 200 km have been observed and adult grey mullets may disperse as far as 700 km⁹. Such migrations may ensure high levels of gene flow amongst locations, leading to genetic homogenisation at regional scales, as observed along the Tunisian coasts¹⁷.

Nevertheless, some outcomes of the present study do not fulfil completely the picture of a unique large pan-mictic population. In this case, local effective population sizes are expected to be roughly of the same magnitude as they shall tend to converge on the global metapopulation size²⁸. In contrast, the four ponds here studied show different local effective population sizes (Table 2), albeit the large and overlapping confidence intervals do not

rule out a similar, large contemporary N_e in three of them. Mistras lagoon is a noteworthy exception to this trend, as it shows the lowest effective population size amongst all populations, even when accounting for uncertainty. Remarkably, populations from Mistras and Cabras show the largest difference in N_e , albeit the channels connecting the ponds to the sea are only one km apart. Thus, it seems unlikely that the low N_e recorded at Mistras might reflect recruitment from a larval pool other than Cabras. Small N_e may depend on recent population declines, which may occur in unpredictable and highly variable environments as brackish-water habitats. Nevertheless, the excess heterozygosity test does not evidence any recent population size decline in Mistras as well as any other population (Table 2). Perhaps, this result depends on the fact that changes in trophic conditions likely affect recruitment of mullets from the Sea²⁹ rather than determining post-recruitment decreases of population abundance. Consistent with this picture, Mistras does not show a reduction of genetic diversity relative to other populations, notwithstanding its small N_e . The loss of genetic diversity because of small local effective population sizes may be counterbalanced by dispersal and gene flow: the offshore spawning migrations of adult mullets may increase the global effective population size, thus preventing inbreeding and genetic drift on local scales in the short- and the long-term, respectively³⁰.

We may advance two hypotheses that might explain the small local N_e observed at Mistras compared to the other populations. Genotype-based habitat selection may drive recruitment in fish that use brackish-water environments as nursery habitats: for instance, the white sea bream (*Diplodus sargus*) and the gilthead sea bream (*Sparus aurata*) migrate into those coastal lagoons where juveniles growth is maximised^{31–33}. Likewise, larvae and juveniles of *M. cephalus* prefer to settle in oligohaline and mesohaline brackish waters, which favour their growth³⁴. Therefore, Cabras likely represents a more suitable habitat than Mistras for the settlement of grey mullet recruits: at the onset of winter season, when juveniles migrate into lagoons, Cabras pond is characterised by mesohaline conditions, whereas Mistras lagoon remains hyperhaline throughout the entire year. We may speculate that early life stages of grey mullet could be mixtures of individuals with different phenotypic traits, each of which maximises growth in different habitats. In this case, different local population sizes will depend on genotype-based habitat selection and the frequency of those genotypes-phenotypes in the global metapopulation rather than sweepstakes recruitment events. Otolith microchemistry provides a further line of evidence that may support this hypothesis: in the Mediterranean Sea, grey mullets show different environmental migratory patterns, spending either part or their entire lifetime in fresh- or brackish-water habitats, as well as preferring seawater and high salinity habitats³⁵.

Alternatively, local N_e might be influenced by the semi-intensive farming of grey mullets, which is a common practise across Mediterranean coastal ponds and in many Sardinian wetlands¹¹. The release of such early-generation hatchery fish may trigger the Ryman–Laikre effect in the supplemented wild population if the captive fishes are the offspring of a limited number of breeders^{4,8}. In addition to decrease N_e , this process should also result in larger than expected relatedness as significant kinship structure should not occur when effective population size is large³¹. Unexpectedly, all four samples show an excess of half-and/or full-siblings as detected by the genotype sharing method (Fig. 2), albeit this result should be interpreted with caution because of the difficulty to assess relationships among individuals using pairwise relatedness³⁶. Indeed, this method estimated a larger proportion of related individuals than that obtained using a maximum likelihood approach and both methods mainly depart in the number of detected half-siblings (Table S8), which likely correspond to the increasing variance of estimators as relatedness decreases³⁷. Disregarding the reliability and precision of each method, the proportion of related individuals is similar across all populations (Table S8), hence downplaying the possibility that the small N_e recorded in Mistras might depend on the Ryman–Laikre effect.

The results outlined above may help complement fishery and aquaculture management in several ways. In the first place, the presence of a unique genetic stock indicates that *M. cephalus* could be considered as a single management unit at regional level, based on the concept that genetically distinct stocks need to be managed as separate units³⁸. Second, regional genetic homogeneity will ease the enforcement of a responsible approach to aquaculture-based fisheries management³⁹, which may help local small fisheries to withstand yield reductions triggered by environmental fluctuations. Our results suggest that early-generation captive individuals from other ponds could be used to supplement depleted populations without affecting their genetic integrity, unless they are the offspring of a small number of breeders. In this context, the genetic information here gathered is also a crucial prerequisite to select and determine the number of individuals that are needed to raise broodstocks that maintain levels of genetic diversity as high as those of wild populations.

Our results may also provide some clue in on the assessment of demographic independence, which is a goal of great relevance in fisheries management. Although it is difficult to delineate demographic units with genetic data^{27,40}, kinship analyses may help reach this goal as these metrics are suitable to assess ongoing dispersal and thus local recruitment patterns⁴¹. Nevertheless, caution is needed to avoid a biased estimation of these demographic processes as the present study further evidences the difficulty of estimating relatedness among individuals³⁷. Based on our results (Supplementary Table S8), considering only high-level relationships might reduce the uncertainty in the estimation of relatedness, which is analogous to increase the threshold above which individuals are deemed as related⁴². Alternatively, only relationships that overlap across two or more methods could be deemed as reliable; in this case, however, recruitment patterns will be determined by the most conservative method or that underestimating relatedness.

In conclusion, our results further confirm the importance of using several lines of evidence to draw inferences on species' fine-scale genetic patterns. This framework suits the needs of fishery managers, who usually adopt a 'best available science' approach in which information from several sources and the associated uncertainty are considered before making management decisions⁴⁰. For instance, local population effective size clues in on a factor that may hamper the supplementation of wild populations. Habitat selection might dampen the success of aquaculture-based Fisheries management programmes if sub-optimal environmental conditions affect the growth and the survival of the introduced captive individuals. This process is difficult to unravel with genetic

metrics as it does not affect neutral genetic variation, unless genetic markers directly or indirectly involved in the selection process are used^{17,31,32}. Finally, the Baseline genetic information here gathered may help improve the sustainable management of *Mugil cephalus* fisheries by setting the stage for the genetic monitoring of wild populations. Monitoring genetic metrics over time is fundamental to evaluate the impact of supplementation programmes on wild populations, as well as the health, exploitation, recruitment and connectivity patterns of natural stocks³⁹.

Methods

Ethical approval. The study did not involve endangered or protected species. No specific permissions were required for locations and activities. Capture, non-lethal sampling and experimental protocols followed the principles of laboratory animal care and regulations on animal welfare enforced by national laws (D. Lgs 116/1992 and D. Lgs. 26/2014) and EU Directive 2010/63/EU. No approval was needed by an institutional ethics committee: fin-clips were obtained from fishes that had to be sold in local markets or used to produce roe mullet and were kindly provided by local fishermen.

Study area. Sardinia island is located in the Western Mediterranean (Fig. 1) and approximately 150 km² are covered by wetlands, which are exploited for fishing and semi-intensive aquaculture of euryhaline fish¹¹. In the present study, a small portion of caudal fin was clipped from individuals in four coastal ponds where either semi-intensive aquaculture for sea ranching or traditional fisheries of grey mullet are common practises.

Cabras, Mistras, Porto Pino and Tortoli ponds (Fig. 1) are confined, shallow, non-tidal systems, which differ for extent and physico-chemical characteristics (Supplementary Table S1). Cabras pond (CAB) is a hypertrophic system, in which salinity may drop to < 10 PSU (Practical Salinity Units) in winter and raise to > 30 PSU in summer, depending on the rainfall⁴³. Mistras pond (MIS) and Porto Pino salty ponds (PPI) are oligotrophic and hyperhaline systems in which salinity may increase above 40 PSU^{44–46}. Tortoli coastal pond (TOR) is a euhaline basin characterised by low eutrophication levels because of a good water exchange with the sea⁴⁷. Salinity ranges between 31 and 38 PSU in winter and summer, respectively⁴⁴.

Sampling, DNA extraction and PCR protocols. Caudal fin-clips of *M. cephalus* ($N=200$) were collected during 2013, preserved in absolute ethanol and stored at -80°C until DNA isolation. Genomic DNA was purified using the salting-out extraction method⁴⁸, and then stored in TE buffer. DNA quantity and quality were assessed using a fluorimeter (Nanodrop 2000) and diluted if necessary.

Multiplex Polymerase Chain reaction (PCR) protocols were setup to amplify 14 microsatellite markers, which were marked on the 5' end of the forward primer with the 6-FAM, VIC, NED and PET fluorochromes (Supplementary Table S2). PCR reactions contained 20–30 ng of genomic DNA, 1X reaction buffer (Euroclone), 2 mM MgCl₂, 0.250 μM of each dNTP, 0.06–0.14 μM of each primer, 1.25 U of EuroTaq DNA polymerase (Euroclone). PCR reactions were carried out on a MJ DNA Engine PTC-100 thermal cycler under the following conditions: an initial denaturation step at 94 °C for 5 min; 30 cycles of 94 °C for 30 s, annealing temperature of each multiplex for 30 s, 72 °C for 30 s; a final extension of 72 °C for 7 min. After chequing for successful amplicons by electrophoresis on a 2% Agarose gel stained with Ethidium Bromide, 1 μl of PCR product mixed with 9.90 μl of Formamide and 0.10 μl of GeneScan 500 (-250) LIZ size standard (Applied Biosystems) was run on an ABI PRISM 3130xl Genetic Analyser (Applied Biosystems). Microsatellite alleles were scored and binned using GeneMapper v4.0 software package (Applied Biosystems).

Population genetics analysis. Departure from Hardy–Weinberg proportions (HWP), and linkage disequilibrium (LD) were tested following⁴⁹. First, the Markov chain method (10,000 dememorization steps, 100 batches of 10,000 iterations each) implemented in Genepop 4.7⁵⁰ was used to compute the probability of HWP departures for either heterozygote deficit or heterozygote excess and LD. The binomial likelihood method (LM) was used to combine probabilities across individual tests as it is not affected by small probability values as the Fisher's exact test is⁵¹. The joint probability of departure from HWP was computed combining probabilities of single tests by locus or by populations. For LD, single tests were grouped by locus pair or population. If the joint probability within each group was smaller than 0.05, the B–Y method of correction for multiple tests was used to adjust the probability values of single tests⁵². The procedures above were automated running two customised scripts in the R 3.5.2 statistical environment⁵³, one of which was designed for carrying out the LM method, while the other was used in Cossu et al.⁵⁴.

The presence of null alleles, stuttering, and large allele dropouts was tested using Micro-checker 2.2.3⁵⁵. The frequency of null alleles was also estimated using FreeNA⁵⁶ to minimise the rate of both false positives and false negatives: only null alleles that reached a frequency > 8% by both methods were deemed as true⁵⁷.

Three methods were used to detect outlier loci. BayeScan 2.1⁵⁸ compares a neutral model with a model that include selection. Setting higher prior odds for the former (threshold = 10), proposal distributions were adjusted using 20 pilot runs of 5000 iterations each and then a simulation was run for 150,000 iterations. Records were sampled every 20 iterations after discarding the first 50,000 iterations. The FDR (False Discovery Rate) was set at 5% to correct for multiple testing. The *LnRH* statistic⁵⁹ assumes reduced levels of genetic diversity within populations for outlier loci. Under neutrality, 95% of neutral loci are expected to show standardised *LnRH* values (mean = 0, standard deviation = 1) between -1.96 and $+1.96$, and between -2.87 and $+2.87$ after correction with the B–Y method. Finally, the FDI_{ST2} approach with a finite island model of migration implemented in Arlequin 3.5.2⁶⁰ was run assuming 100 demes and 100,000 permutations. Probability values were corrected using the B–Y method. Only loci detected as outliers by all methods were deemed as true candidates for selection⁶¹.

Within population summary statistics, population size reductions and contemporary effective population sizes were assessed following Cossu et al.⁵⁴. The number of alleles (N_A), allelic richness (A_R), expected and observed heterozygosity (H_E and H_O) and the inbreeding coefficient (F_{IS}) were computed using *diveRsity* 1.9.9⁶². Non-parametric pairwise Wilcoxon rank-sum tests were used to assess if H_E and A_R were different between population pairs. Signatures of recent population declines were assessed using the heterozygosity excess test implemented in *BOTTLENECK* 1.2⁶³. The contemporary effective population size (N_e) was estimated using the linkage disequilibrium (LD) method implemented in *NeEstimator* V2.1⁶⁴, setting the minor allele frequency to 0.02⁶⁵.

Genetic differentiation amongst populations was assessed using *diveRsity* to compute the Weir and Cockerham's F_{ST} estimator θ^{66} and Jost's D_{EST} ⁶⁷ index: means and confidence intervals were assessed carrying out 10,000 bootstrap replicates. Exact G tests implemented in *Genepop* 4.7 were used to compute the probability values of population differentiation. Probabilities of pairwise multiple tests were adjusted applying the B–Y method⁵².

POWSIM V1.2⁶⁸ was used to estimate the statistical power of detecting population differentiation. Pseudo-datasets with the same number of populations (N), loci, alleles and population sampling size (S) as the real dataset were created. Different levels of genetic differentiation ($F_{ST} = 0.00–0.01$) were simulated assuming constant population effective size ($N_e = 2000$), no migration and varying the time since divergence. The statistical power was evaluated computing the fraction of both Chi-square and Fisher's exact tests that successfully detected population differentiation out of 1000 replicates for each simulation.

Kinship analyses were carried out using the method of Blouin et al.⁶⁹, implemented in *DEmerelate* 0.9.3⁷⁰. The method tests for an overabundance of closely related individuals by comparing the proportions of full- and half-siblings observed in empirical populations with those expected in randomised reference populations. Furthermore, to assess the uncertainty in sibship reconstruction³⁶, relationships among individuals were cross-checked using *ML-Relate*⁷¹.

Genetic structure was investigated using the Bayesian model-based clustering implemented in *structure* 2.3.4⁷². Only genetic information was used to group individuals into clusters that minimise Hardy–Weinberg–Linkage disequilibria. Simulations were run using the admixture model with correlated allele frequencies⁷³ and varying the number of clusters ($K = 1–4$). For each K , 10 independent runs were performed, each consisting of 100,000 iterations following a burn-in period of equal length. The mean posterior probability ($lnP(D)$) and its standard deviation (SD) were computed using the *Pophelper* 1.0.10 web app⁷⁴. In addition to model-based Bayesian clustering, a Discriminant Analysis of Principal Components (DAPC), which does not rely upon a population genetic model⁷⁵, was carried out in the *adegenet* R package⁷⁶. DAPC aimed at visualizing differences among ponds as prior groups. The number of retained principal components was optimised using a cross-validation procedure to avoid overfitting of the data.

Data availability

The dataset generated and analysed during the current study is available in the figshare repository, <https://doi.org/10.6084/m9.figshare.12594374>.

Code availability

Custom R codes used for data analysis are available in the figshare repository, <https://doi.org/10.6084/m9.figshare.12594374> and in the DRYAD repository, <https://doi.org/10.5061/dryad.dm97908>.

Received: 3 July 2020; Accepted: 18 January 2021

Published online: 03 February 2021

References

- Garibaldi, L. The FAO global capture production database: A six-decade effort to catch the trend. *Mar. Pol.* **36**, 760–768 (2012).
- FAO (Food and Agriculture Organization). The State of World Fisheries and Aquaculture 2018. in *Meeting The Sustainable Development Goals*. (FAO, Rome, 2018).
- Grant, W. S., Jasper, J., Bekkevold, D. & Adkison, M. Responsible genetic approach to stock restoration, sea ranching and stock enhancement of marine fishes and invertebrates. *Rev. Fish Biol. Fish.* **27**, 615–649 (2017).
- Christie, M. R., Marine, M. L., French, R. A., Waples, R. S. & Blouin, M. S. Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. *Heredity* **109**, 254–260 (2012).
- Ryman, N. & Laikre, L. Effects of supportive breeding on the genetically effective population size. *Conserv. Biol.* **5**, 325–329 (1991).
- Waples, R. S., Hindar, K., Karlsson, S. & Hard, J. J. Evaluating the Ryman–Laikre effect for marine stock enhancement and aquaculture. *Curr. Zool.* **62**, 617–627 (2016).
- Sun, X. & Hedgecock, D. Temporal genetic change in North American Pacific oyster populations suggests caution in seascape genetics analyses of high gene-flow species. *Mar. Ecol. Prog. Ser.* **565**, 79–93 (2017).
- Bacheler, N. M., Wong, R. A. & Buckel, J. A. Movements and mortality rates of striped mullet in North Carolina. *N. Am. J. Fish. Manage.* **25**, 361–373 (2005).
- Whitfield, A. K., Panfili, J. & Durand, J.-D. A global review of the cosmopolitan flathead mullet *Mugil cephalus* Linnaeus 1758 (Teleostei: Mugilidae), with emphasis on the biology, genetics, ecology and fisheries aspects of this apparent species complex. *Rev. Fish Biol. Fish.* **22**, 641–681 (2012).
- Hsu, C.-C., Chang, C.-W., Iizuka, Y. & Tzeng, W.-N. A growth check deposited at estuarine arrival in otoliths of juvenile flathead mullet (*Mugil cephalus* L.). *Zool. Stud.* **48**(3), 315–323 (2009).
- Antuofermo, E. et al. First evidence of intersex condition in extensively reared mullets from Sardinian lagoons (central-western Mediterranean, Italy). *Ital. J. Anim. Sci.* **16**, 283–291 (2017).
- Heras, S., Roldán, M. I. & Castro, M. G. Molecular phylogeny of Mugilidae fishes revised. *Rev. Fish Biol. Fish.* **19**, 217–231 (2009).
- Heras, S., Maltagliati, F., Fernández, M. V. & Roldán, M. I. Shaken not stirred: A molecular contribution to the systematics of genus *Mugil* (Teleostei, Mugilidae). *Integr. Zool.* **11**, 263–281 (2016).
- Shen, K.-N., Jamandre, B. W., Hsu, C.-C., Tzeng, W.-N. & Durand, J.-D. Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet *Mugil cephalus*. *BMC Evol. Biol.* **11**, 83. <https://doi.org/10.1186/1471-2148-11-83> (2011).

15. Durand, J.-D. *et al.* Systematics of the grey mullets (Teleostei: Mugiliformes: Mugilidae): Molecular phylogenetic evidence challenges two centuries of morphology-based taxonomy. *Mol. Phylogenet. Evol.* **64**, 73–92 (2012).
16. Rossi, A. R., Capula, M., Crossetti, D., Campton, D. E. & Sola, L. Genetic divergence and phylogenetic inferences in five species of Mugilidae (Pisces: Perciformes). *Mar. Biol.* **131**, 213–218 (1998).
17. Blel, H. *et al.* Selection footprint at the first intron of the Prl gene in natural populations of the flathead mullet (*Mugil cephalus*, L. 1758). *J. Exp. Mar. Biol. Ecol.* **387**, 60–67 (2010).
18. Durand, J., Blel, H., Shen, K., Koutrakis, E. & Guinand, B. Population genetic structure of *Mugil cephalus* in the Mediterranean and Black Seas: A single mitochondrial clade and many nuclear barriers. *Mar. Ecol.-Prog. Ser.* **474**, 243–261 (2013).
19. Šegvić-Bubić, T. *et al.* Range expansion of the non-native oyster *Crassostrea gigas* in the Adriatic Sea. *Acta Adriat.* **57**(2), 321–330 (2016).
20. Piras, P. *et al.* A case study on the labeling of bottarga produced in Sardinia from ovaries of grey mullets (*Mugil cephalus* and *Mugil capurrii*) caught in Eastern Central Atlantic coasts. *Ital. J. Food. Saf.* **7**(1), 6893. <https://doi.org/10.4081/ijfs.2018.6893> (2018).
21. Miggiano, E. *et al.* Isolation and characterization of microsatellite loci in the striped mullet, *Mugil cephalus*. *Mol. Ecol. Notes* **5**, 323–326 (2005).
22. Jinliang, W. A parsimony estimator of the number of populations from a STRUCTURE-like analysis. *Mol. Ecol. Resour.* **19**, 970–981 (2019).
23. FAO (Food and Agriculture Organization). *Code of Conduct for Responsible Fisheries* (FAO, Rome, 1995).
24. Mai, A. C. G. *et al.* Microsatellite variation and genetic structuring in *Mugil liza* (Teleostei: Mugilidae) populations from Argentina and Brazil. *Estuar. Coast. Shelf Sci.* **149**, 80–86 (2014).
25. Pacheco-Almanzar, E., Simons, J., Espinosa-Perez, H., Chiappa-Carrara, X. & Ibanez, A. L. Can the name *Mugil cephalus* (Pisces: Mugilidae) be used for the species occurring in the north western Atlantic?. *Zootaxa* **4109**, 381–390 (2016).
26. Waples, R. S. & Gaggiotti, O. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol. Ecol.* **15**, 1419–1439 (2006).
27. Hauser, L. & Carvalho, G. R. Paradigm shifts in marine fisheries genetics: Ugly hypotheses slain by beautiful facts. *Fish Fish.* **9**, 333–362 (2008).
28. Waples, R. S. & England, P. R. Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics* **189**, 633–644 (2011).
29. Murenu, M., Olita, A., Sabatini, A., Follasa, M. C. & Cau, A. Dystrophy effects on the *Liza ramada* (Risso, 1826) (Pisces, Mugilidae) population in the Cabras lagoon (Central-Western Sardinia). *Chem. Ecol.* **20**, 425–433 (2004).
30. Ryman, N., Laikre, L. & Hössjer, O. Do estimates of contemporary effective population size tell us what we want to know?. *Mol. Ecol.* **28**, 1904–1918 (2019).
31. Guinand, B. *et al.* Candidate gene variation in gilthead sea bream reveals complex spatiotemporal selection patterns between marine and lagoon habitats. *Mar. Ecol. Prog. Ser.* **558**, 115–127 (2016).
32. Chaoui, L. *et al.* Microsatellite length variation in candidate genes correlates with habitat in the gilthead sea bream *Sparus aurata*. *Mol. Ecol.* **21**, 5497–5511 (2012).
33. González-Wangüemert, M. & Pérez-Ruzafa, Á. In two waters: contemporary evolution of lagoonal and marine white seabream (*Diplodus sargus*) populations. *Mar. Ecol.* **33**, 337–349 (2012).
34. Cardona, L. Effects of salinity on the habitat selection and growth performance of Mediterranean flathead grey mullet *Mugil cephalus* (Osteichthyes, Mugilidae). *Estuar. Coast. Shelf Sci.* **50**, 727–737 (2000).
35. Fortunato, R. C., Galán, A. R., Alonso, I. G., Volpedo, A. & Durá, V. B. Environmental migratory patterns and stock identification of *Mugil cephalus* in the Spanish Mediterranean Sea, by means of otolith microchemistry. *Estuar. Coast. Shelf Sci.* **188**, 174–180 (2017).
36. Jones, A. G., Small, C. M., Paczolt, K. A. & Ratterman, N. L. A practical guide to methods of parentage analysis. *Mol. Ecol. Resour.* **10**, 6–30 (2010).
37. Taylor, H. R. The use and abuse of genetic marker-based estimates of relatedness and inbreeding. *Ecol. Evol.* **5**, 3140–3150 (2015).
38. Coppinger, C. R. *et al.* Assessing the genetic diversity of catface grouper *Epinephelus andersoni* in the subtropical Western Indian Ocean. *Fish. Res.* **218**, 186–197 (2019).
39. Cushman, E. L. *et al.* Development of a standardized molecular tool and estimation of genetic measures for responsible aquaculture-based fisheries enhancement of American Shad in North and South Carolina. *Trans. Am. Fish. Soc.* **148**, 148–162 (2019).
40. Waples, R. S., Punt, A. E. & Cope, J. M. Integrating genetic data into management of marine resources: How can we do it better?. *Fish. Fish.* **9**, 423–449 (2008).
41. Iacchei, M. *et al.* Combined analyses of kinship and F_{ST} suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. *Mol. Ecol.* **22**, 3476–3494 (2013).
42. Bernardi, G., Beldade, R., Holbrook, S. J. & Schmitt, R. J. Full-sibs in cohorts of newly settled coral reef fishes. *PLoS ONE* **7**(e44953), 2012. <https://doi.org/10.1371/journal.pone.0044953> (2012).
43. Como, S., van der Velde, G. & Magni, P. Temporal variation in the trophic levels of secondary consumers in a Mediterranean coastal lagoon (Cabras lagoon, Italy). *Estuaries Coasts* **41**, 218–232 (2018).
44. Floris, R., Manca, S. & Fois, N. Microbial ecology of intestinal tract of gilthead sea bream (*Sparus aurata* Linnaeus, 1758) from two coastal lagoons of Sardinia (Italy). *Transit. Waters Bull.* **7**(2), 4–12. <https://doi.org/10.1285/i1825229Xv7n2p4> (2013).
45. Merella, P. & Garippa, G. Metazoan parasites of grey mullets (Teleostea: Mugilidae) from the Mistras lagoon (Sardinia-Western Mediterranean). *Sci. Mar.* **65**, 201–206 (2001).
46. Pitacco, V. *et al.* Spatial patterns of macrobenthic alpha and beta diversity at different scales in Italian transitional waters (Central Mediterranean). *Estuar. Coast. Shelf Sci.* **222**, 126–138 (2019).
47. Cioffi, F. & Gallerano, F. From rooted to floating vegetal species in lagoons as a consequence of the increases of external nutrient load: An analysis by model of the species selection mechanism. *Appl. Math. Model.* **30**, 10–37 (2006).
48. Wasko, A. P., Martins, C., Oliveira, C. & Foresti, F. Non-destructive genetic sampling in fish. An improved method for DNA extraction from fish fins and scales. *Hereditas* **138**, 161–165 (2003).
49. Waples, R. S. Testing for Hardy–Weinberg proportions: Have we lost the plot?. *J. Hered.* **106**, 1–19 (2015).
50. Rousset, F. genepop'007: A complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106 (2008).
51. Kinnison, M. T., Bentzen, P., Unwin, M. J. & Quinn, T. P. Reconstructing recent divergence: Evaluating nonequilibrium population structure in New Zealand chinook salmon. *Mol. Ecol.* **11**, 739–754 (2002).
52. Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**(4), 1165–1188 (2001).
53. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2018).
54. Cossu, P. *et al.* Influence of genetic drift on patterns of genetic variation: The footprint of aquaculture practices in *Sparus aurata* (Teleostei: Sparidae). *Mol. Ecol.* **28**, 3012–3024 (2019).
55. Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535–538 (2004).
56. Chapuis, M.-P. & Estoup, A. Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* **24**, 621–631 (2007).

57. Dąbrowski, M. J. *et al.* Reliability assessment of null allele detection: Inconsistencies between and within different methods. *Mol. Ecol. Resour.* **14**, 361–373 (2014).
58. Foll, M. & Gaggiotti, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant Markers: A Bayesian perspective. *Genetics* **180**, 977–993 (2008).
59. Kauer, M. O., Dieringer, D. & Schlötterer, C. A microsatellite variability screen for positive selection associated with the “Out of Africa” habitat expansion of *Drosophila melanogaster*. *Genetics* **165**, 1137–1148 (2003).
60. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
61. Paris, M. *et al.* Genome scan in the mosquito *Aedes rusticus*: population structure and detection of positive selection after insecticide treatment. *Mol. Ecol.* **19**, 325–337 (2010).
62. Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W. & Prodöhl, P. A. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol. Evol.* **4**, 782–788 (2013).
63. Piry, S., Luikart, G. & Cornuet, J.-M. Computer note. BOTTLENECK: A computer program for detecting recent reductions in the effective size using allele frequency data. *J. Hered.* **90**, 502–503 (1999).
64. Do, C. *et al.* NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol. Ecol. Resour.* **14**, 209–214 (2014).
65. Waples, R. S. & Do, C. Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evol. Appl.* **3**, 244–262 (2010).
66. Weir, B. S. & Cockerham, C. C. Estimating F -statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
67. Jost, L. G_{ST} and its relatives do not measure differentiation. *Mol. Ecol.* **17**, 4015–4026 (2008).
68. Ryman, N. & Palm, S. POWSIM: A computer program for assessing statistical power when testing for genetic differentiation. *Mol. Ecol. Notes* **6**, 600–602 (2006).
69. Blouin, M. S., Parsons, M., Lacaille, V. & Lotz, S. Use of microsatellite loci to classify individuals by relatedness. *Mol. Ecol.* **5**, 393–401 (1996).
70. Kraemer, P. & Gerlach, G. Demerelate: Calculating interindividual relatedness for kinship analysis based on codominant diploid genetic markers using R. *Mol. Ecol. Resour.* **17**, 1371–1377 (2017).
71. Kalinowski, S. T., Wagner, A. P. & Taper, M. L. ml-relate: A computer program for maximum likelihood estimation of relatedness and relationship. *Mol. Ecol. Notes* **6**, 576–579 (2006).
72. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
73. Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).
74. Francis, R. M. pophelper: An R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* **17**, 27–32 (2017).
75. Jombart, T., Devillard, S. & Balloux, F. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94. <https://doi.org/10.1186/1471-2156-11-94> (2010).
76. Jombart, T. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).

Acknowledgements

We wish to thank Marco Trentadue for his invaluable help during field sampling. We are indebted to: Nuovo Consorzio Cooperative Pontis (Cabras Pond), Cooperativa Pescatori e Molluschicoltori (Mistras lagoon), Società Cooperativa Pescatori San Giuseppe (Porto Pino ponds) and Cooperativa Pescatori Tortoli (Tortoli pond), which kindly allowed to collect fin-clip samples and supported on field sampling. This work was part of LM’s Ph.D. thesis.

Author contributions

All authors contributed to conceive and design the study. N.F. and L.M. collected the field samples. L.M. carried out laboratory work. P.C. carried out data analysis and interpretation. P.C. wrote the first draft of the manuscript and all authors commented on previous version of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-82515-7>.

Correspondence and requests for materials should be addressed to P.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021