University of Sassari

Dissertation for the Degree of Doctor of Philosophy in Environmental Biology

presented at Sassari University in 2015

XXVII cycle

IMPACT OF AQUACULTURE ON THE GENETIC STRUCTURE OF ANIMAL MARINE POPULATIONS

PH.D. CANDIDATE: ***Dr. Laura Mura***

DIRECTOR OF THE SCHOOL: ***Prof. Marco Curini Galletti***

SUPERVISOR: ***Prof. Marco Casu***

CO-SUPERVISOR: ***Prof. Marco Curini Galletti***

University of Sassari

Dissertation for the Degree of Doctor of Philosophy in Environmental Biology

presented at Sassari University in 2015

XXVII cycle

IMPACT OF AQUACULTURE ON THE GENETIC STRUCTURE OF ANIMAL MARINE POPULATIONS

PH.D. CANDIDATE: ***Dr. Laura Mura***

DIRECTOR OF THE SCHOOL: ***Prof. Marco Curini Galletti***

SUPERVISOR: ***Prof. Marco Casu***

CO-SUPERVISOR: ***Prof. Marco Curini Galletti***

INDEX

Abstract

The development of marine aquaculture systems has raised concerns about the impact that intentional (restocking) or unintentional (escapes) releases may have on natural populations. On this regard the genetic integrity of wild populations is threatened by: 1) biological invasions, 2) loss of genetic variation, 3) change of genetic architecture, and 4) loss of local adaptation. My study concerns the genetic characterization of both wild populations and (un)intentionally introduced populations as recommended by the Code of Conduct for Responsible Fisheries. This aim has been pursued focusing on three main issues. The first, on the gilthead sea bream, highlighted that farmed populations were genetically divergent from wild ones. The other two issues are related to restocking of potentially exploited natural stocks. To this end we investigated the Mediterranean basin-scale genetic structure of the grooved carpet shell; our results evidenced an overall lack of genetic structure throughout the western Mediterranean which may be used in planning future management of this natural resource. In addition, we analysed also populations of its allochthonous counterpart, the manila clam, that show relatively high levels of genetic variation suggesting that this species may cause problem to grooved carpet shell. With regard to the grey flathead mullet, results show levels of genetic differentiation as small as 1% and suggested the occurrence of a unique genetic stock along Sardinian coasts.

Key words: Maricolture; Genetic diversity; *Sparus aurata*; *Ruditapes decussatus*; *Mugil cephalus.*

Riassunto

Lo sviluppo dei sistemi di acquacoltura causa preoccupazioni sull'impatto che il rilascio intenzionale (ripopolamento) o accidentale (fughe) di individui può avere sulle popolazioni naturali. È noto che l'integrità genetica delle popolazioni naturali sia minacciata da: 1) invasioni biologiche, 2) perdita di variabilità genetica, 3) cambiamenti nell'architettura genetica, e 4) perdita di adattamento locale. A tal proposito, nel mio progetto sono state caratterizzate geneticamente popolazioni selvatiche e non, focalizzando l'attenzione su tre casi principali. Il primo, volto ad analizzare i pattern di variabilità genetica in poplazioni selvatiche e allevate di orata, ha evidenziato che le popolazioni allevate sono geneticamente divergenti, presentando una minore variabilità genetica. Gli altri due temi sono connessi al ripopolamento di stock naturali potenzialmente sfruttati. A tal fine è stata investigata la vongola verace che ha mostrato una generale mancanza di strutturazione genetica in tutto il bacino del Mediterraneo occidentale. Parallelamente è stata analizzata la sua controparte alloctona, la vongola filippina, che ha mostrato livelli relativamente alti di variabilità genetica, suggerendo che questa specie possa competere con successo a scapito della nostra vongola verace. Infine, è stato effettuato uno studio volto all'identificazione degli stock genetici di muggine, che ha suggerito la presenza di un unico stock lungo le coste della Sardegna.

Parole chiave: Maricultura; Variabilità genetica; *Sparus aurata*; *Ruditapes decussatus*; *Mugil cephalus.*

Introduction

Background

The term "aquaculture" refers to the technique for farming aquatic organisms such as fish, crustaceans, mollusks, etc. Aquaculture involves cultivating freshwater and saltwater populations under human-controlled conditions, and it can be contrasted with commercial fishing, which is the harvesting of wild populations/stocks. Aquaculture has been observed since ancient times in the economy of different civilizations, such as Phoenicians, Etruscans, Romans, Egyptians and Chinese. Nowadays, it represents a very conspicuous share of the economic sector of food production. For example, in 2003 fish aquaculture products accounted for about 31% (41.9 million tons) out of a total of approximately 132 million tons of fish caught. Its contribution to world supplies of aquatic products has increased from one million tons per year in the early 1950s, to a production of 55.1 million tons in 2009 (FAO, 2006) (Grigorakis and Rigos 2011). Due to greater demand for seafood products, aquaculture growth has observed to be very rapid in the world: for many species the increase account for more than 10% per annum, while on the contrary the contribution of traditional fishing has remained steady or declining in the last decade (FAO, 2006). Indeed, the pessimistic forecasts for fisheries production represent one of the mainly reason for aquaculture spreading, which doubtless is the fastest growing animal-food production fuelled by governmental support.

Over time, aquaculture has been diversified in several methods, such as intensive, iperintensive, extensive and mariculture. The three first methods refer to landlocked tanks, whilst mariculture refers to aquaculture practiced in marine environments and in

underwater habitats. The latter one comes from a specialized branch that involve the cultivation of marine organisms for food and other products in the open ocean, an enclosed section of the ocean, or in tanks, ponds or raceways which are filled with seawater.

Also in Mediterranean, mariculture has followed the global trend during the last three decades, with a very rapid growth. Mediterranean mariculture began in the 80's, mainly by the culture of European sea bass (*Dicentrarchus labrax*), the gilthead sea bream (*Sparus aurata*) and diverse shellfish. During the '90s, the Atlantic bluefin tuna (*Thunnus thynnus*) rearing and the farming of other sparids (Sparidae, Teleostei) were introduced. More recently other species, mainly from Scienidae and Caragidae families entered the industry with the most promising including the meagre (*Argyrosomous regious*) and the greater amberjack (*Seriola dumerilii*).

The expansion of the use of mariculture techniques has been favored by several factors, such as new technologies, improvements in formulated feeds, and a not negligible greater biological understanding of farmed species. Indeed, production has grown from 90,000 tons in 1985 to 436,401 in 2007 only for marine fish, and to 174,385, only for mollusk (CAQ-GFCM, 2008, 2009; FAO, 2008; Barazi-Yeroulanos, 2010; Grigorakis and Rigos 2011). With regard to shellfish, the production is mainly represented by the European mussel (*Mytilus galloprovincialis*), the Japanese carpet shell (*Ruditapes philippinarum*), the European flat oyster (*Ostrea edulis*) and the Pacific oyster (*Crassostrea gigas*) (CAQ-GFCM, 2008; Grigorakis and Rigos 2011).

Maricultured fish are also perceived to be of higher quality than fish raised in ponds or tanks, and offer more diverse choice of species. In addition, mariculture promises

economic and environmental benefits, indeed it can produce fish at lower cost than industrial fishing, leading to better human diets and the gradual elimination of unsustainable fisheries (Katavic, 1999; Brugere et al., 1999). However, despite the undeniable advantages of this technique, it presents also several problems. In particular, a crucial issue is represented by the entities of the impact of the "escapees", that are linked to several factors.

Introduction of exotic species and the occurrence of the so-called genetic pollution, represent the most dangerous problem given by escapes of farmed species; indeed, together with habitat destruction, physical and chemical pollution, they have been identified as two of the most negative factor that affect marine biodiversity (Cognetti et al., 2006). For what concern genetic pollution, it is meant as an alteration of the natural genetic architecture and microevolutionary processes of wild populations due to the gene flow from farmed conspecifics that have escaped from aquaculture plants (Cognetti et al., 2006). On this regard, one of the main important consideration to keep in mind is that that farmed species can descend by ancestors or relatives genetically structured in different ways along a continuum, ranging from a situation of panmixia to an array of highly isolated local populations. Natural patterns of species genetic structure are multifarious and the amount of interpopulation connectivity (gene flow) may vary greatly across different species. Furthermore, within a given species, gene flow may vary in both spatial and temporal sense. To date, thanks to the availability of several classes of molecular markers, it is possible to check and monitor genetic diversity of both farmed and wild populations and to assess the degree of genetic variability lost in cultured populations. Indeed, in recent years, a number of studies

addressing the problem of genetic variability in farmed marine species in the Mediterranean have been carried out (Cognetti et al., 2006), and all of them evidenced that more and more researches with these aims are stressed. Indeed, although a better maintenance and servicing of tanks can reduce escaped, probably genetic interaction between cultured and wild species cannot be precluded at all, because even sexually mature males and females are known to release gametes or fertilized eggs in the surrounding water.

For what concern exotic species, it is meant as species introduced into an area outside of its native range and it has a tendency to spread, which is believed to cause damage to the environment or to native species. Escape of exotic farmed species is perceived as a threat to natural biodiversity in marine waters. Indeed, they may cause undesirable ecological effects in native populations through interbreeding, competition for food or habitats, as well as transfer of pathogens to wild fish or other farmed stocks (Thorstad et al., 2010). For instance, autochthonous species may suffer from competition with the introduced one (see e.g. Chessa et al., 2005). This competition is due to very often the introduced species can progressively occupy the ecological niche of the indigenous species, and sometimes this can lead to the extinction of local populations (see e.g. Jensen et al., 2004).

Aims of PhD thesis

On the basis of what described in the previous paragraph, the main purpose of my PhD project was the enlargement of the knowledge on genetic structure of several species of particular commercial interest in the Mediterranean, and, particularly, in Sardinia. This objective was achieved through three main issues:

1) To investigate the patterns of genetic diversity in wild and farmed populations of the gilthead sea bream, *Sparus aurata* (Teleosta: Sparidae) along the Sardinian coasts, and to compare the genetic variation of these populations with that of wild and farmed populations out of this region;

2) To study the genetic variability of the autochthonous carpet shell clam, *Ruditapes decussatus* (Bivalvia: Veneridae) from Sardinian and compare it with that present in the rest of the Mediterranean and Atlantic, in order to evaluate how population transplantations from a given area to another one, with refurbishment purposes. may influence the genetic variation of this species. The occurrence of a new established population of the allochthonous congeneric manila clam, *R. philippinarum* (Bivalvia: Veneridae) in the Gulf of Olbia, also give me the opportunity to study the first phases of colonization, with a genetic focus;

3) To achieve information about the genetic variability of the Sardinian stocks of grey flathead mullet, *Mugil cephalus* (Teleostei: Mugilidae) in order to obtain useful information for the proper management of the resource and the restocking programs, as well as to preserve the genetic diversity of natural populations.

The above issues have been structured in four study cases, which led to the production of six manuscripts/PhD thesis chapters:

1) Genetic diversity of reared and wild populations of *Sparus aurata* in Sardinia.

2) Genetic variation in wild and farmed populations of the gilthead sea bream (*Sparus aurata*) in Sardinia inferred with microsatellite markers.

3) Preliminary data on the genetic variability of the clam *Ruditapes decussatus*.

4) Genetic structuring of the carpet shell clam *Ruditapes decussatus* along Mediterranean and northern Atlantic coast.

5) Genetic variability in the Sardinian population of the manila clam, *Ruditapes philippinarum*.

6) Development of a microsatellite toolkit for the identification of genetic stocks in Sardinian populations of the grey flathead mullet (*Mugil cephalus*).

- *Sparus aurata*

The case study of the gilthead sea bream *Sparus aurata* (Teleostei: Sparidae) (chapters 1, 2) has been performed because this fish is one of most cultured species in Mediterranean. Accordingly, its potentially escapes has raised concerns on the potential threats on the genetic integrity of natural populations. Although the impact of escapes remains to be quantified, though some authors emphasized that aquaculture may have already contributed to current genetic patterns. However, to our knowledge, no studies have compared patterns of genetic diversity among farmed and wild populations of

gilthead sea bream along the Italian coasts, as done for instance by Alarcón et al. (2004) at large spatial scales, or at smaller geographic scales in Greece (Loukovitis et al. 2012) and Croatia (Šegvić-Bubić et al. 2011). For this reason, my work represents the first study on this regard, focused into bridging this gap, in which I have investigated patterns of genetic variation in farmed and wild populations of gilthead sea bream, centering mainly the Sardinian region, where farming took place from more than a decade ago. For this purpose I chose microsatellite markers, because for this kind of study they represent the first choice. Moreover, other researches on the genetic structure of wild population used microsatellite (see e.g. De Inocentiis et al. 2004, Franchini et al. 2012 and references therein), and thus I could choose best primers among the literature. Here, analyses were based on a set of eighteen microsatellite markers.

- *Ruditapes decussatus*

The second study case (chapters 3, 4) refers to the grooved carpet shell clam, *Ruditapes decussatus* (Bivalvia: Veneridae), an autochthonous bivalve of the Mediterranean. The rationale for this research is given by the fact that the grooved carpet shell clam, considered as one of the most important for Mediterranean aquaculture, experimented a reduction of its population as possible consequence of an unregulated commercial exploitation. On this regard, knowledge on its genetic structure is matter of considerable importance when you want to consider the possibility of future restocking. Accordingly, in order to provide management plans devoted to preserve the genetic variability of this species, it is of great relevance knowing the present status of Mediterranean *R. decussatus* genetic variation on a geographic scale as large as it is possible. This study

represents an effort in this sense, as I analyzed mitochondrial COI (Cytochrome c Oxidase subunit I) gene relationships among samples from the eastern and western Mediterranean and Atlantic.

- *Ruditapes philippinarum*

The third study case (chapter 5) is linked to the above described one. Indeed, it refers to the allochthonous manila clam *Ruditapes philippinarum* (Bivalvia: Veneridae). This species has been introduced in Europe since to 1970, and it represents the allochthonous counterpart of our autochthonous grooved carpet shell clam (*R. decussatus*). Since, in the Adriatic the manila clam progressively occupied the ecological niche of grooved carpet shell, which led to the extinction of local populations (Jensen et al., 2004), this study was designed to analyze the genetic variation among *R. philippinarum* specimens collected at different sites in the Gulf of Olbia, with the twofold aim of detecting whether (a) genetic substructuring is present within this area and whether (b) a recent founder effect occurred in the Gulf of Olbia. To this end, we compared estimates of within-population genetic variability of Sardinian samples with those of a formerly established, farmed population from the North Adriatic. These purposes were achieved by means of microsatellite markers.

- *Mugil cephalus*

The last study case is about the grey flathead mullet, *Mugil cephalus* (Teleostei: Mugilidae), also commonly referred to as the striped mullet, a world-wide distributed

coastal fish species that uses estuarine environment as nurseries habitat and then migrates offshore for spawning (Thomson, 1955; Bacheler et al., 2005). Although nowadays it is not a *sensu strict* farmed species, the grey flathead mullet caught my attention because of its high commercial value in many parts of the world. This condition makes it a noteworthy species to be inserted in my PhD project. On this regard, this study case was focused on the assessment of genetic variation and in carrying out the so-called genetic stock identification (GSI) in populations of *M. cephalus*. Also in this case, these information may be helpful tools in the proper management of the resource and the restocking programs, as well as to preserve the genetic diversity of natural populations.

References

Alarcón J.A., Magoulas A., Georgakopoulos T., Zouros E., Alvarez M.C. (2004) Genetic comparison of wild and cultivated European populations of the gilthead sea bream (*Sparus aurata*). Aquaculture 230: 65–80.

Bacheler N.M., Wong R.A., Buckel J.A. (2005) Movements and mortality rates of striped mullet in North Carolina. North American Journal of Fisheries Management 25: 361–373.

Barazi-Yeroulanos L. (2010). Synthesis of Mediterranean Marine Finfish Aquaculture – A Marketing and Promotion Strategy. Studies and Reviews. General Fisheries Commission for the Mediterranean. FAO No. 88, Rome, p. 198.

Brugere C.M.J.F., Young J.A. (1999). Green grow the fishes-Oh? Environmental attributes in marketing aquaculture products. Aquculture Economics & Management 3(1): 7-17.

CAQ-GFCM (2008) Committee on Aquaculture General Fisheries Commission for the Mediterranean Trends and Issues of Marine and Brackish Mediterranean Aquaculture. Trends and Issues of Marine and Brackish Mediterranean Aquaculture, Tirana, Albania, 17–19 September, 2008.

CAQ-GFCM (2009) Committee on Aquaculture, General Fisheries Commission for the Mediterranean. Developing of a Strategy for Marketing and Promotion of Mediterranean Aquaculture, Tangier, Morocco, 27–28 October, 2009.

Chessa L.A., Paesanti F., Pais A., Scardi M., Serra S., Vitale L. (2005) Perspectives for development of low impact aquaculture in a Western Mediterranean lagoon: the case of the carpet clam Tapes decussatus. Aquacult. Int. 13: 147–155.

Cognetti G., Maltagliati F., Saroglia M. (2006) The risk of ‘‘genetic pollution’’ in Mediterranean fish populations related to aquaculture activities. Marine Pollution Bulletin 52: 1321–1323.

De Innocentiis S., Lesti A., Livi S., Rossi A. R., Crosetti D., Sola L. (2004) Microsatellite markers reveal population structure in gilthead sea bream *Sparus auratus* from the Atlantic Ocean and Mediterranean Sea. Fisheries Science 70: 852–859.

FAO (2006) Food and Agriculture Organization of the United Nations. The state of world Fisheries and Aquaculture 2006, FAO Fisheries and Aquaculture Department, Rome, 2007.

FAO (2008) Food and Agriculture Organization of the United Nations. The state of world Fisheries and Aquaculture 2008, FAO Fisheries and Aquaculture Department, Rome, 2009.

Franchini P., Sola L., Crosetti D., Milana V., Rossi A.R. (2012) Low levels of population genetic structure in the gilthead sea bream, *Sparus aurata*, along the coast of Italy. ICES Journal of Marine Science 69: 41–50.

Grigorakis K., Rigos G. (2011) Aquaculture effects on environmental and public welfare – The case of Mediterranean mariculture. Chemosphere 855: 899–919

Jensen A.C., Humphreys J., Caldow R.W.G., Grisley C., Dyrynda P.E.J. (2004). Naturalization of the Manila clam (*Tapes philippinarum*), an alien species, and establishment of a clam fishery within Poole Harbour, Dorset. J. Mar. Biol. Assoc. UK 84: 1069–1073.

Katavic I. (1999). Mariculture in the New Millennium. Agriculture Conspectus Scientificus 64(3): 223-229.

Loukovitis D., Sarropoulou E., Vogiatzi E., Tsigenopoulos C.S., Kotoulas G., Magoulas A., Chatziplis D. (2012) Genetic variation in farmed populations of the gilthead sea bream *Sparus aurata* in Greece using microsatellite DNA markers. Aquaculture Research 43: 239–246.

Šegvić-Bubić T., Lepen I., Trumbić Ž., Ljubković J., Sutlović D., Matić-Skoko S., Grubišić L., Glamuzina B., Mladineo I. (2011) Population genetic structure of reared and wild gilthead sea bream (*Sparus aurata*) in the Adriatic Sea inferred with microsatellite loci. Aquaculture 318: 309–315.

Thomson J.M. (1955) The movements and migrations of mullet (*Mugil cephalus* L.). Aust J Mar Freshw Res 6: 328–347.

Thorstad E.B., Whoriskey F., Rikardsen A.H., Aarestrup K. (2008) Aquatic Nomads: The Life and Migrations of the Atlantic Salmon, in Atlantic Salmon Ecology (eds Ø. Aas, S. Einum, A. Klemetsen and J. Skurdal), Wiley-Blackwell, Oxford, UK. doi: 10.1002/9781444327755.ch1.

Chapter 1

Genetic diversity of reared and wild populations of *Sparus aurata* in Sardinia

Laura Mura, Gian Luca Dedola, Tiziana Lai, Fabio Scarpa, Daria Sanna, Marco Curini-Galletti, Marco Casu, Nicola Fois, Piero Cossu

Pubblished in *Biologia Marina Mediterranea*, 20 (1): 106-107 (2013)

Biol. Mar. Mediterr. (2013), 20 (1): 106-107

L. Mura, G.L. Dedola*, T. Lai*, F. Scarpa*, D. Sanna*, M. Curini-Galletti*, M. Casu*, N. Fois, P. Cossu*

Agris Sardegna, Località Bonassai, Olmedo (SS), Italy.
lmura@agrisricerca.it
*Dip. di Scienze della Natura e del Territorio, Università di Sassari, Sassari, Italy.

GENETIC DIVERSITY OF REARED AND WILD POPULATIONS OF *SPARUS AURATA* IN SARDINIA

DIVERSITÁ GENETICA NELLE POPOLAZIONI ALLEVATE E SELVATICHE DI SPARUS AURATA *IN SARDEGNA*

Abstract *- This study aims at comparing patterns of genetic diversity in farmed and wild populations of the gilthead sea bream,* Sparus aurata *Linnaeus, 1758 (Sparidae) along the Sardinian shores. We aimed at assessing if farmed populations are genetically divergent and have lower genetic variability than wild populations. As a by-product we investigated genetic structure among wild populations.*

Key-words: Sparus aurata*, genetic diversity, microsatellites, farm-cage escapes.*

Introduction - The gilthead sea bream, *Sparus aurata*, is one of the most important sparids reared in the Mediterranean Sea. The fact that the species is usually farmed in off-shore sea-cages has been raising concern on the consequences that escapes of farmed individuals may have on the local wild stocks. Introduction of genotypes not adapted to local environments and/or genetic erosion may decrease the fitness of wild populations, leading to the decline of local stocks (Fraser *et al*., 2010). This study investigates patterns of genetic variation in wild and farmed populations of *Sparus aurata* along the Sardinian shores, where farming dates back to over 20 years ago. Our goal is to assess whether farmed populations are genetically divergent and displayed lower genetic variability than wild populations.

Materials and methods - A total of 158 wild specimens were collected at four locations (Gulf of Asinara, Alghero, Porto Pino, Tortolì), and 164 individuals from four fish farms (Alghero, Orosei, Golfo Aranci, Calasetta). Genomic DNA was obtained from fin clips using the salting-out extraction method (Wasko *et al*., 2003). Multiplex PCR amplification was performed using eight microsatellite markers (PbMS2, SaGT32, SaI12, SauDI82, SauG46, SauH98, SauI47, SauK140) following Navarro *et al*. (2008). Electrophoresis runs were performed in a 3130xl Genetic Analyzer Capillary (Applied Biosystem); microsatellites were screened and scored using GeneMapper 4.0 software. Departure from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were tested using GENEPOP on the web (Rousset, 2008). The presence of null alleles, stuttering, and large allele dropouts was tested using MICROCHECKER 2.2.3 program (Van Oosterhout *et al*., 2004). Population summary statistics were estimated using GENALEX 6.5 (Peakall and Smouse, 2012). Number of effective alleles (Ne), expected (*He*) and observed (*Ho*) heterozygosity were used as estimates of within-population genetic variability. Genetic structure was estimated using F-statistics; Population pairwise F_{st} values were estimated among each population pair, and significance of population differentiation was tested by 1,000 permutations of individuals among populations. Where necessary, the False Discovery Rate method (FDR) was applied to correct probability values due to multiple testing (Benjamini and Yekutieli, 2001).

Results - Significant departure from HWE was observed at locus SauG46 for most populations; as the excess of homozygotes was due to a high frequency of

null alleles this locus was removed from further analysis. Observed and expected Heterozygosity ranged from *Ho* =0.686±0.094 (Alghero farmed population) to *Ho* =0.742±0.060 (Golfo Aranci), and *He* =0.695±0.118 (Orosei) to *He* =0.746±0.082 (Alghero wild population). Pairwise t-tests carried out on He, and pairwise Wilcoxon sign-rank tests carried out on Ne, did not evidence significant differences between the wild and farmed samples. We found a significant genetic structure (*F*st =0.020, *P* =0.001). Significant genetic differentiation was found between all farmed populations. Among these, Calasetta and Orosei are also genetically divergent from all the wild samples, whereas Alghero and Golfo Aranci are not, with the exception of the pair Golfo Aranci-Porto Pino.

Conclusions - Levels of genetic within-population genetic variability and genetic structure are comparable with those found in other studies (Šegvić-Bubić *et al*., 2011 and references therein). Our data do not evidence a lower genetic variation in farmed populations than the wild populations, which may prevent genetic erosion of native stocks due to farm escapes. Nonetheless, the presence of farmed populations that are genetically divergent from native wild stocks may represent an issue. Indeed, if native populations are adapted to local environments, the introduction of alien genotypes by farm-escaped individuals may reduce the fitness and survival of wild populations, as shown in the Atlantic *Salmo salar* (Fraser *et al*., 2010).

References

BENJAMINI Y., YEKUTIELI D. (2001) - The control of the false discovery rate in multiple testing under dependency. *Ann. Statist*., **29**: 1165-1188.

FRASER D.J., HOUDE A.L.S., DEBES P.V., O'REILLY P., EDDINGTON J.D., HUTCHINGS J.A. (2010) - Consequences of farmed-wild hybridization across divergent wild populations and multipletraits in salmon. *Ecol. Appl*., **20**: 935-953.

NAVARRO A., BADILLA R., ZAMORANO M.J., PASAMONTES V., HILDEBRANDT S., SÁNCHEZ J.J., AFONSO J.M. (2008) - Development of two new microsatellite multiplex PCRs for three sparid species: Gilthead seabream (*Sparus auratus* L.), red porgy (*Pagrus pagrus* L.) and redbanded seabream (*P. auriga*, Valenciennes, 1843) and their application to paternity studies. *Aquaculture*, **285**: 30-37.

PEAKALL R., SMOUSE P.E. (2012) - GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, **28**: 2537-2539.

ROUSSET F. (2008) - Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol. Ecol. Res*., **8**: 103-106.

ŠEGVIĆ-BUBIĆ T., LEPEN I., TRUMBIĆ Ž., LJUBKOVIĆ J., SUTLOVIĆ D., MATIĆ-SKOKO S., GRUBIŠIĆ L., GLAMUZINA B., MLADINEO I. (2011) - Population genetic structure of reared and wild gilthead sea bream (*Sparus aurata*) in the Adriatic Sea inferred with microsatellite loci. *Aquaculture*, **318**: 309-315.

VAN OOSTERHOUT C., HUTCHINSON W.F., WILLS D.P.M., SHIPLEY P. (2004) - Microchecker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes*, **4**: 535-538.

WASKO A.P., MARTINS C., OLIVEIRA C., FORESTI F. (2003) - Non-destructive genetic sampling in fish. An improved method for DNA extraction from fish fins and scales. *Hereditas*, **138**: 161-165.

Chapter 2

Genetic variation in wild and farmed populations of the gilthead sea bream (*Sparus aurata*) in Sardinia inferred with microsatellite markers

Piero Cossu, Laura Mura, Fabio Scarpa, Tiziana Lai, Daria Sanna, Marco Curini-Galletti, Nicola Fois, Marco Casu

manuscript

Genetic variation in wild and farmed populations of the gilthead sea bream (*Sparus aurata*) in Sardinia inferred with microsatellite markers

P. Cossu[1], L. Mura[2], F. Scarpa[1], T. Lai[1],D. Sanna[1], M. Curini-Galletti[1], N. Fois[2], M. Casu[1]

[1] Dipartimento di Scienze della Natura e del Territorio, Università di Sassari, Via F. Muroni 25, 07100 Sassari (SS), Italy

[2] Agris Sardegna, Agenzia per la Ricerca in Agricoltura. Dipartimento per la Ricerca nelle Produzioni Animali, Località Bonassai, 07040 Olmedo (SS), Italy

Corresponding author:

Piero Cossu

Dipartimento di Zoologia e Genetica Evoluzionistica

Università di Sassari

Via F. Muroni 25

07100 Sassari (SS), Italy

Phone: +39 079 228924

Fax: +39 079 228665

picossu@uniss.it

Abstract

The intensive farming of fish in ponds or floating cage systems has raised concerns on the potential threats that escapes from farms might have on the genetic integrity of natural populations. Concerning the gilthead sea bream (*Sparus aurata* Linaeus 1759), the impact of escapes remains to be quantified, though some authors emphasized that aquaculture may have already contributed to current genetic patterns. Here, we have investigated patterns of genetic variation in farmed and wild populations of gilthead sea bream, focusing mainly on the Sardinian region, where farming took place more than a decade ago. The main results of this study, based on a set of eighteen microsatellite markers highlight that i) genetic variation was greater in wild than farmed populations; ii) farmed populations were genetically divergent from their wild counterparts; and iii) wild and farmed populations shared part of their genetic makeup. Patterns of genetic variation observed at wild and farmed populations might reflect the interplay of two main factors: on one hand reared stocks were originated from a heterogeneous mixture of fry from non local broodstocks. On the other side small effective population size of reared stocks point to increased random genetic drift. Finally, the uncertainty observed in individual assignment may underpin the ongoing homogenization of wild populations due to past escapes from finfish farms. The subsequent admixture between farmed and wild individuals might have resulted in a more or less constant gene flow from farmed to natural populations over time.

Keywords: Microsatellites, Sparidae, genetic diversity, aquaculture, genetic pollution

Introduction

The increased production of fish due to the development of marine aquaculture systems based on floating cage technology has raised concerns about the impact of escaped farmed specimens on Mediterranean natural fish populations; indeed, escapes due either to culture system failure, accidents or carelessness, cannot be completely prevented and inevitably lead to gene flow between cultured and wild stocks (Dempster et al. 2007). Particularly, in floating cages accidents (e.g. extreme sea-floods) might result in the escape of reared fish that may reach up to several tens of thousands (De Inocentiis et al. 2005, Miggiano et al. 2005). As a consequence, a precautionary approach for a sustainable and responsible use of aquatic resources is recommended in order to protect local biodiversity and minimize the risk of genetic pollution and its harmful effects in case of escapes from fish farms (FAO 1995). The potential threats to the genetic integrity of wild populations include 1) loss of genetic variation (usually farmed individuals are the offspring of a small number of breeders), 2) change of population genetic structure and composition, and 3) loss of local adaptation, the latter two being more relevant when hatchery juveniles are from non-local broodstocks (Laikre et al. 2010 and references therein).Thus, according to Code of Conduct for Responsible Fisheries (FAO 1995), any genetic change has to be monitored through the genetic characterization of both wild populations and (un)intentionally introduced populations.

The gilthead sea bream (*Sparus aurata* Linaeus 1759) is a euryhaline and eurythermal perciform fish that, together with the European sea bass (*Dicentrarchus labrax*), is considered as the most important finfish aquaculture product in the

Mediterranean Sea (http://fao.org). Increased effort into developing intensive production practices started in the early 1980s, after which large fry production spreads from France and Italy to the rest of the Mediterranean countries, reaching an industrial level of fish production in the 1990s (Loukovitis et al. 2012, Šegvić-Bubić et al. 2012). Currently, the production of farmed gilthead sea breams in Europe reached 151'787 t in 2013; Italy with 8'400 t has been the fourth largest producer after Greece, Turkey and Spain (FEAP report 2014: www.aquamedia.org). With regard to this species escapees from sea cages have been sporadically reported (Dempster et al. 2002, 2005, Boyra et al. 2004, Tuya et al. 2005, 2006, Valle et al. 2007, Fernandez-Jover et al. 2008). Furthermore, the simulated escape of farmed individuals tagged with acoustic transmitters or external tags showed that a significant fraction of tagged individuals survived in natural habitats as they switch to natural food resources within a short time after release (Arechavala-Lopez et al. 2012). This finding raises concerns given the common practice of breeding gilthead sea breams from Atlantic broodstocks in Mediterranean hatcheries, as the shape and the growth performance of Atlantic form are particularly appreciated (e.g. Miggiano et al. 2005).

Despite Italy has been since 1980s one of the pioneering countries in farming of gilthead sea breams, studies that investigated intra- and inter-population genetic variation in reared populations and their potential impact on wild populations have been scarce. For instance, De Innocentiiis et al. (2005) investigated the geographic origin of individuals belonging to two commercial Italian hatchery broodstocks. Miggiano et al. (2005) investigated the genetic impact of escaped farmed individuals on a natural population by simulating a cage breakdown experiment. To our knowledge, however,

no studies have compared patterns of genetic diversity among farmed and wild populations of gilthead sea bream along the Italian coasts, as done for instance by Alarcón et al. (2004) at large spatial scales, or at smaller geographic scales in Greece (Loukovitis et al. 2012) and Croatia (Šegvić-Bubić et al. 2011).

Therefore, this study represents a first step into bridging this gap, as it focus mostly on analyzing patterns of genetic variation in farmed and wild populations along the Sardinian coast using microsatellite markers. The rationale for this choice is twofold; first, Sardinia harbours several finfish farms that have been working since at least 10 years, several of which (at least those used in this study), import the fry from other Italian finfish farm facilities (see Materials and Methods section). Second, two studies that analyzed genetic variation in wild populations of gilthead sea breams along Italian coasts collected several samples along Sardinian coasts (De Inocentiis et al. 2004, Franchini et al. 2012). With regard to Sardinian region these studies yielded contradictory results; De Innocentiis et al. (2004) using four microsatellite markers found that the three Sardinian populations (Sardinian sea, Sardinian channel and central Thyrrenian sea) were genetically differentiated. Conversely, results of Franchini et al. (2012), based on a set of ten microsatellite markers did not support the findings of the former authors. Overall genetic differentiation was lower, and in particular, Sardinian populations were gathered into a genetically homogeneous group. This lead to hypothesize that such differences might reflect the increased power of resolution of their dataset, nonetheless they also emphasize the chance that current genetic patterns might reflect population homogenization due to aquaculture practices.

The aim of this study was to investigate the patterns of genetic diversity in wild and farmed populations of gilthead sea bream along the Sardinian coasts, and to compare the genetic variation of these populations with that of wild and farmed populations out of this region. In particular we want to assess i) if genetic variability within farmed population is lower than that within wild populations; ii) quantify the degree of genetic differentiation among wild and farmed populations and within these groups, and iii) use assignment tests to identify genetic stocks and detect the presence of wild individuals with a genetic makeup similar to that of farmed individuals.

Materials and methods

Sampling and DNA extraction

Fin-clips of wild and farmed gilthead sea breams were collected throughout the period 2009-2013 at six locations from western Mediterranean and six finfish farms (Table 1and Fig.1) for wild (n = 223) and farmed specimens (n = 227), respectively. Wild individuals sampled along Sardinian coast were located more or less in the vicinity of a farm facility (Fig.1). In addition to the wild Sardinian samples, as a comparison we collected wild individuals from Marina di Massa (MMW) and Mauguio lagoon (LMW) in the north-western Mediterranean (Fig. 1). The four Sardinian fin-fish farms we picked up use off-shore sea cages to rear gilthead sea breams since 10 years at least. Since then accidental escapes have occurred frequently due to system failure (damages to the cages), particularly during the first years, whereas decreased in the last years due to a technological improvement of sea-cages, as reported by the fin-fish farmers. More

rarely, escapes occurred during the harvest of fish or were due to extreme sea-floods. None of the surveyed farms breeds its own fry from local broodstocks, buying fingerlings from other Italian hatcheries (Table S1).

Dorsal Fin-clips were preserved in absolute ethanol and stored at -80°C until DNA isolation. Genomic DNA was purified using the salting-out extraction method (Wasko *et al*., 2003), and then stored in TE buffer. DNA quantity and quality was assessed using a fluorimeter (Nanodrop 2000), and diluted if necessary.

PCR protocols

A series of Multiplex PCR reaction protocols were setup to amplify 18 microsatellite markers (Table S2) starting from the protocols developed by Navarro et al. (2008) and Porta et al. (2010). PCR reactions were prepared in a final volume of 25 µl, containing 20–30 ng of genomic DNA, 1X reaction buffer (Euroclone), 2 mM MgCl2, 0.250 µM of each dNTP, 0.03 - 0.09 µM of each primer, 1.25 U of EuroTaq DNA polymerase (Euroclone). PCR reactions were carried out in 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, 58°C for 30 s, 72°C for 45 s; a final extension of 72°C for 7 min. After checking 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 Analyzer (Applied Biosystems). Microsatellite loci were screened and scored by resulting electropherograms using GeneMapper v4.0 software package (Applied Biosystems).

Statistical data analysis

Departure from Hardy-Weinberg equilibrium (HWE) at each locus and population, as well as linkage disequilibrium (LD) among each pair of loci were tested using GENEPOP 4.0 (Rousset 2008). The probability of departures was estimated using the Markov chain method (10'000 dememorization steps, 100 batches of 10'000 iterations each), and applying the False Discovery Rate (FDR) method of correction for multiple comparisons (Benjamini & Yekutieli 2001). The presence of null alleles, stuttering, and large allele dropouts was tested using MICROCHECKER 2.2.3 (Van Oosterhout et al. 2004). Since two loci (PbMS2 and SaI21) showed signs of null alleles due to homozygote excess in some but not all populations, we used the software FreeNA (Chapuis & Estoup 2007) to calculate null allele frequencies and estimate global and pairwise F_{ST} values. This software implements the excluding null alleles (ENA) method, which correct efficiently for the bias induced by null alleles on F_{ST} estimation (Chapuis & Estoup 2007). Then, the magnitude of the bias due to null alleles was assessed carrying out a Mantel test between uncorrected and corrected pairwise F_{ST} values with 10,000 random permutations to test whether they were significantly correlated.

We also carried out two outlier tests based on the levels of inter-population differentiation for identifying loci under possible selection.Both methods use a coalescent-based simulation approach to identify outlier loci that display unusually high and low values of F_{ST} by comparing observed F_{ST} values with those expected under neutrality (Beaumont & Nichols 1996). The first method is implemented in the Lositan Selection Workbench (Antao et al. 2008) and uses the FDIST2 outlier detection method

based on the finite island model of migration (Beaumont and Balding 2004). We run 100,000 simulations with 100 demes under the infinite allele model (IAM) and the stepwise mutation model (SMM), enabling "neutral mean F_{ST}" and "force mean F_{ST}" options to increase reliability of neutral mean F_{ST} estimates. These options perform a preliminary run to detect loci lying outside the 99.5% confidence interval (CI) and exclude them before the estimation of mean neutral F_{ST}. This refined estimate is then used as the mean neutral F_{ST} for the final run, and the probability that an outlier locus lying outside the 99.5% CI is under selection is corrected setting the FDR rate threshold at 0.05.

The second approach, implemented in ARLEQUIN 3.5 (Excoffier & Lischer 2010), overcomes possible false positives due to strong hierarchical population structure (Excoffier et al. 2009) using the FSDIST2 approach with a hierarchical island model instead of a finite island model of migration. Samples were split into seven groups based on pairwise F_{ST} values and Bayesian model-based clustering (see results: wild samples are grouped together). The program was run assuming 100 demes per 10 groups, and 100,000 permutations. Due to multiple comparisons probabilities were corrected using the FDR method (Benjamini & Yekutieli 2001).

We used MSA 4.05 (Dieringer & Schlötterer 2003) to calculate summary statistics for the mean number of alleles (N_A), gene diversity (H_E), and observed heterozigosity (H_O). We also estimated the allelic richness (A_R) and the private allelic richness (A_{PR}) to account for differences in sample size. These summary statistics are the mean number of alleles and the mean number of private alleles corrected for the smallest sample size in the dataset, respectively, using the rarefaction method developed by Kalinowski (2004),

and implemented in the software HP-Rare (Kalinowski 2005). A non-parametric Mann-Whitney U test was then used to assess if wild populations showed values of H_E, A_R, and P_{AR} greater than farmed populations. We used the framework developed by Hothorn et al. (2006) implemented in the package Coin (Hothorn et al. 2008), a library working within the R 3.0 statistical environment (R Core team 2013). The test statistics was computed using a permutation test (10'000 Montecarlo replicates); moreover, we setup a randomized block design in which each microsatellite locus is a block and permutations are allowed only within blocks. We used H_E as a measure of genetic variability within populations as it is robust to the presence of null alleles (Chapuis et al. 2009), whereas A_R is a more sensitive measure than the former to short and recent bottlenecks such as founder events (Norris et al. 1999), as allelic loss may occur faster than the loss of heterozigosity (Maruyama and Fuerst 1985).

Genetic differentiation was assessed calculating overall and population pairwise F_{ST} statistics (Weir & Cockerham 1984) with MSA 4.05. A permutation test with 10'000 replicates was used to assess whether or not F_{ST} was significantly greater than zero; for pairwise F_{ST} values the FDR correction method (Benjamini & Yekutieli 2001) was applied to account for type I errors due to multiple comparisons. We also tested for the influence of mutation rate on the estimate of population differentiation using the microsatellite allele size permutations test (Hardy et al. 2003) implemented in SPAGEDI 1.4 (Hardy & Vekemans 2002). This test permutes allele size among populations and then computes R_{ST} (an F_{ST} analogous based on microsatellite allele sizes, Slatkin 1995). The test statistics computes the probability the observed R_{ST} is significantly greater than the value expected under the null hypothesis that mutation rate

has no effect on the population structure. Here we carried out 1'000 permutations to compute the one-tail probability that mutation rate was not negligible and thus assess whether R_{ST} should be used instead of F_{ST} to estimate population differentiation (Whitlock 2011).

The more likely population structure was also investigated using the Bayesian model-based clustering implemented in the software Structure 2.3.4 (Pritchard et al. 2000). We run several simulations varying the number of clusters (K) from 1 to 12 and using the admixture model with correlated allelic frequencies (Falush et al. 2003). Moreover, we also carried out a second set of simulations implementing the "LOCPRIOR" model (Hubisz et al. 2009), which uses prior information about sampling locations and may be helpful when there is weak genetic structure. For each simulation we performed 10 independent runs, consisting of 100,000 iterations following a burn-in period of equal length. To detect the more likely number of populations we used a one-tail two-sample paired Wilcoxon test to determine if there was a significant difference in the likelihood among successive values of K (Rosenberg et al. 2001). To check the consistency of results across replicate runs, we used the software CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007); the average pairwise similarity (*H*) among replicate runs (Rosenberg et al. 2002) for a given *K* was estimated by a Greedy search algorithm with 10,000 random input orders. Replicate runs were merged to obtain the average membership of all individuals to each cluster and the results were displayed using the software DISTRUCT 1.1 (Rosenberg 2004).

For each population, the current effective population size (N_e) was estimated using the linkage disequilibrium method implemented in Neestimator V2 (Do et al. 2014).

The minor allele frequency allowed was 0.02 when sample size was greater than 25 individuals, and 0.03 otherwise (Waples & Do 2010). Additionally, we used the software Bottleneck 1.2 (Piry et al. 1999) to assess if populations showed the signature of a recent expansion following a bottleneck. The program tests for an excess of heterozygosity, assuming that after a bottleneck the gene diversity will be greater than the heterozigosity expected at equilibrium ($H_E > H_{eq}$), because the observed number of alleles (from which H_{eq} is calculated) will be reduced faster than gene diversity (Cornuet and Luikart 1996). The program was run with 1'000 iterations assuming a two-phase model of microsatellite mutation (TPM), with the proportion of mutations following the stepwise mutation model set at 80% and variance at 12% as in Coscia et al. (2011). The significance of heterozigosity excess against expected heterozigosity at equilibrium over all loci was assessed using a one-tailed Wilcoxon sign-rank test, as recommended for less than 20 loci (Piry et al. 1999). Furthermore we also used the graphical test to detect shifts from the L-shaped distribution of allele frequencies expected at equilibrium (Luikart et al. 1998).

Finally, we used two individual assignment tests to determine the extent to which individuals could be assigned to their sampling population (genetic stock identification, GSI). The first method of self-assignment based upon the leave-one-out cross-validation method developed by Anderson et al. (2008) and is implemented in the software ONCOR (Kalinowski 2007) which is based on a maximum likelihood approach to assign individuals.We also used Bayesian method of Rannala and Mountain (1997) implemented in Geneclass2 (Piry et al. 2004). In the latter method the probability of the more likely genetic source for a given individual was computed using the Montecarlo

resampling method of Paetkau et al. (2004) with 10'000 iterations and a threshold for the probability of exclusion (PE) set at 0.01. Individuals from farmed populations were included in the analysis as well; being of known origin these specimens may serve as a cross-check to evaluate the reliability of individual self-assignment (Griffiths et al. 2010).

Results

All microsatellite loci were polymorphic at the 5% level across all populations, with the exception of the locus SauH98 that was monomorphic in the MOB population. Overall, we scored 385 alleles over 18 loci (minimum 8 alleles for SauD182; maximum 57 alleles for SaI21). After adjusting probabilities for multiple comparisons significant LD was found at 78 out of 1'796 loci pairs; however physical linkage is unlikely as LD among pairs of loci was not consistent across populations (the same pair showed signs of LD at no more than three populations). Furthermore, among the seven populations at which LD was found, five were farmed populations (CAB, GAB, MOB, GRB and TUB), thus suggesting that founder events and/or limited number of progenitors in the sample collection rather than gametic associations may explain the results.

Considering each wild and farmed sample as an individual population, all samples showed departure from HWE at least for one locus (25 out of 216 tests were significant after adjustment for multiple comparisons). Overall 9 out of 18 loci showed signature of Hardy-Weinberg disequilibria, but only three loci (PbMS2, SaI21 and SaI10) deviated from HWE across several populations. Nonetheless, in 20 out of 25 significant tests HWE departures may be due to the presence of null alleles at all but 5 loci: SaI12, SaI19

and SaI10 in MAB, SauK140 in CAB and SaI15 in PPW. Among loci affected by the presence of null alleles, only PbMS2 and SaI21 exhibited null allele frequencies > 0.2, the threshold below which the bias induced by null alleles for population structure estimates is negligible (Dakin & Avise 2004). However the effect of null alleles on global F_{ST} estimates seems negligible (F_{ST} = 0.0212 and 0.0215, with and without applying the ENA method, respectively). Furthermore, the strong correlation among corrected and uncorrected pairwise F_{ST} values (R^2 = 0.999, P < 0.0001) suggests that null alleles should not bias the results (Vandewoetijne & Dyck 2010).

The two outlier tests used in this study yield contradictory results; the FDIST method based on the finite island migration model indicated SaI21 as a candidate target of divergent selection using the 99.5% CI. Conversely, the FDIST accounting for hierarchical genetic structure another locus (SaI15) as a candidate outlier for balancing selection. Furthermore, as any locus was detected as an outlier by both methods after false discovery rate correction (FDR = 0.05), these results prompt that loci used in this study are selectively neutral. Therefore, all the subsequent results are based on the entire set of microsatellites, as even removing loci departing from HWE, with null allele frequencies > 0.2, or potentially under selection did not affect the outcomes (data not shown).

The number of alleles (A), the mean observed and expected heterozygosity (H_O and H_E, respectively), as well as the mean allelic richness (A_R) and the mean allelic private richness (A_{PR}) are summarized in Table 2. Gene diversity (H_E), ranging from 0.79 to 0.81, in wild populations was significantly greater than values found in farmed populations, that ranged between 0.76 and 0.79, (one-tail Mann-Whitney U test, Z = -

5.801, P < 0.0001). Allelic richness over all loci, varying from 11.2 to 12.2 and from 8.4 to 10.7 in wild and farmed populations, respectively, was significantly greater than expected by chance as well (one-tail Mann-Whitney U test, Z = -7.289, P < 0.0001). Private alleles were found in both wild (32 private alleles) and farmed populations (16) with the highest frequency encountered in MAB (locus SaI14, allele 227: f = 0.042); once again wild populations showed a greater private allelic richness than their farmed counterparts (one-tail Mann-Whitney U test, Z = -5.922, P < 0.0001) with average values ranging between 0.7 - 1.1, and 0.3 – 0.6, respectively.

The global F_{ST} value (F_{ST} = 0.021 ± 0.007, P < 0.001) showed significant genetic variation among the populations of gilthead sea bream included in the present study. According to pairwise multilocus F_{ST} estimates (Table 3), wild populations were not genetically differentiated (F_{ST} range 0 - 0.007, P > 0.05 after FDR correction). Conversely, most genetic differentiation was distributed among farmed populations (F_{ST} range 0.010 - 0.063). All farmed samples were genetically differentiated from the wild ones, with the exception of the pair LMW-MAB (F_{ST} 0.006, P > 0.05). interestingly, MAB is the farmed population showing the lowest genetic differentiation from the wild samples, as compared to the other farmed populations, whereas MOB is that displaying on average the highest genetic differentiation (Table 3). Overall, pairwise F_{ST} values among the farmed and the wild samples were intermediate between those observed within wild and farmed groups, respectively (F_{ST} range 0.006 - 0.046). The non-metric multidimensional scaling ordination carried out on the pairwise F_{ST} values (Fig. 2) showed a good fit to the data (stress = 0.014), highlighting a clear separation of MOB from all other samples along the first dimension, as well as TUB and GRB on the

second dimension. Overall, wild and farmed populations are not starkly separated into two distinct groups; noteworthy, wild populations show a lesser degree of scattering with respect to farmed populations, which suggest greater genetic similarity among the former ones.

Due to weak genetic structuring, only the Bayesian clustering model using prior information on the sampling populations was able to detect genetic structure, hence only results based on the LOCPRIOR model are shown. The one-tailed two-sample paired Wilcoxon sign-rank statistic indicated K = 6 as the maximum number of genetic clusters that can be detected by the Bayesian clustering on the data at hand (Fig. S1). Indeed, the difference of the likelihood between models with K > 6 was not significant (P > 0.05 after FDR correction), suggesting that gain of information is negligible assuming more than six clusters (Garnier et al. 2004). The average pairwise similarity among the 10 runs for this model (H = 0.76) was below the threshold indicated by Rosenberg et al. (2002) for highly similar clustering results (H = 0.85). Nonetheless the consistence among runs was good as evidenced by Fig. 3 which compare the best run (Fig. 3a) with the result averaged over the 10 runs (Fig. 3b). All individuals from wild samples showed a very similar genetic makeup, being assigned to the same cluster (Fig. 3); in most of wild samples the membership into this cluster reached the threshold that, according to Vähä et al. (2007) indicate pure ancestry (Q > 0.80, supplementary material table S3 and S4). Greek and Turk farmed specimens showed pure ancestry into two distinct and homogeneous clusters as well (Averaged Q = 0.83 and 0.84, respectively). None of the Sardinian farmed populations showed pure ancestry into a distinct cluster (Q < 0.80, Table S3 and S4); within each population specimens were

admixed (showing membership in more than one cluster) or displayed different genetic makeups (individuals assigned to different clusters). With the exception of MOB, a not trivial portion of the genome of Sardinian farmed populations was assigned to the same cluster as the wild specimens (Q > 0.20, Tables S2 and S3). Nonetheless the majority of individuals at CAB, GAB, and MOB had a similar genetic makeup, respectively, which allow to identify roughly genetic clusters corresponding to each population (Fig. 3). Conversely, it was not possible to recognize a genetic cluster roughly matching MAB, and all individuals from this farmed population were fairly admixed, being assigned mostly to the same clusters as wild and CAB specimens (Averaged Q = 0.46 and 0.27, respectively).

The estimated N_e values for wild populations were higher than those estimated for farmed populations (Table 4). Though the mean estimated N_e was not finite for ASW, TOW and LMW, the lower 95% CI bounds were finite and from about 3 up to 17 times greater than the highest lower bound value found across farmed populations. Pooling the wild populations according to F-statistic and Bayesian clustering, and thus a single panmictic unit, yielded a finite N_e estimate of 2'391 individuals (95% CI: 1'483 – 5'916). Farmed populations exhibited finite N_e Estimates ranging from 23 (MOB, 95% CI: 21 – 26) to 145 individuals (MAB, 95% CI: 83 - 483). None of the populations examined went through a recent bottleneck event, under both assumed mutation models as evidenced by the Wilcoxon sign-rank test for heterozygosity excess (P > 0.05, see Table 4 for details), and the distribution of mutation shifts that was L-shaped as expected at equilibrium.

Results of individual self-classification performed with ONCOR and Geneclass2 the Bayesian method are summarized in Tables 5 and 6. Using the approach of Anderson et al. (2007) ONCOR correctly assigned 191 out of 437 individuals to their sample collection (43.7%), whereas with the method of Rannala and Mountain (1997) implemented in Geneclass only 123 individuals were correctly assigned (28.1%). However, this fraction is even lower if we deemed as assigned to a given genetic source only individuals to which the resampling method of Paetkau et al. (2004) give a probability >50% of being from that population (data not shown). The two methods showed also contrasting results in the assignment of wild individuals to farmed populations; ONCOR allocated 27 wild specimens to farmed populations, whereas geneclass2 only two individuals. With the exception of MAB both methods reached the highest accuracy of self-assignment in farmed populations, as it may be expected for groups of individuals of known origin. Once again ONCOR outperformed Geneclass2. Assuming these misclassifications were errors, as all farmed individuals are of known origin, ONCOR showed higher accuracy than geneclass2 (20.3% and 62.1% of misclassified farmed individuals, respectively).

Discussion

The present study aimed at giving a first insight on the genetic impact that farmed populations of *Sparus aurata* may exert on natural populations in the western Mediterranean. Indeed, it moves another step toward the need for the genetic tagging and monitoring of farmed gilthead sea breams in the Mediterranean (Miggiano et al. 2005, Loukovitis et al. 2012) and the distinction between micro-evolutive processes and

domestication effects (Alarcón et al. 2004, Franchini et al. 2012). Ideally these studies should be carried out preferentially at the onset of farmed populations or restocking programs (Laikre et al. 2010); however, this is not possible in the gilthead sea bream as rearing has been going since 1980s along Italian coasts. Likely, past rearing without an appropriate control of broodstock origin and history, the extensive exchange of breeders and fry between hatcheries, restocking with hatchery fry in natural lagoons, accidental releases from sea cages, and the release of gametes by farmed gilthead sea breams into the marine systems in the past (Dimitriou et al. 2007) may already have caused a mixture of different gene pools (Franchini et al. 2012). Thus, as emphasized by these authors aquaculture may already have partially contributed to wild population homogenization, being actually the factor that best explains current population patterns.

Notwithstanding such potential caveats this study highlights some interesting patterns. First, wild samples exhibited a within-population genetic diversity greater than that observed in farmed populations, as evidenced by Mann-Whitney U tests ($P < 0.05$). These results are consistent with those reported in previous studies carried out at both large (Alarcón et al. 2004) and small spatial scales (Šegvić-Bubić et al. 2011, Loukovitis et al. 2012). The lower genetic diversity observed in farmed populations might reflect the increased genetic drift as fingerlings are usually the offspring of a rather small number of progenitors (Loukovitis et al. 2012). Overall, estimates of genetic variation were similar, though those found in our study, and in particular allelic richness tend to be greater than values reported in the studies of Loukovitis et al. (2012) and Šegvić-Bubić et al. (2011). Maybe our results were affected by highly variable loci; the mean number of alleles per locus (21.4) was more than twice the corresponding

values reported in the above mentioned studies (see Table S2 for details). However, several clues point to disprove this possibility, at least partly. First, the allelic richness and gene diversity of GRB are within the range observed in the samples of Loukovitis et al. (2012); second, estimates of allelic richness in our samples are still greater than those of Šegvić-Bubić et al.'s populations, even if we used the subset of loci corresponding to their dataset (data not shown). We may advance the hypothesis that the greater genetic variation found in Sardinian reared stocks may be due to the fact that fingerlings are mixtures of different broodstocks (Table S1). Interestingly the Sardinian farmed population showing the lowest values of A_R and H_E is MOB, whose fry is from a single supplier. Finally, the estimates of gene diversity observed in wild populations (mean H_E = 0.79 – 0.81) were lower than values previously reported for the western Mediterranean region (De Innocentiis et al. 2004, Franchini et al. 2012). Nevertheless these results likely reflect the presence of loci with low mutation rate and heterozygosity (e.g. SauH98); after the removal of this locus he estimates were similar to those reported previously (data not shown). Furthermore, gene diversity range at loci SaGT32 (H_E = 0.89 – 0.91), SauK140 (H_E = 0.83 – 0.85) and SaGT26 (H_E = 0.81 – 0.88) were very similar to those reported by De Innocentiis et al. (2004) and Franchini et al. (2102) for the same loci, thus suggesting that genetic diversity seems to be stable over time.

The pattern of genetic structure evidenced by F_{ST} values was similar to those reported by former studies. As in Loukovitis et al. (2012) we found significant genetic differentiation between the wild and farmed populations and among all the farmed populations (Table 3). In particular, genetic differentiation was greater among farmed

populations as evidenced by the higher degree of scatter in the nMDS ordination of pairwise F_{ST} values (Fig. 2). Concerning the wild populations, our results did not evidence any significant genetic structuring, supporting partly the findings of Franchini et al. (2012). These authors pointed out the occurrence of a weak genetic structuring between Sardinian and Ligurian Sea regions. Conversely our results based either on F_{ST} values and Bayesian clustering (Fig. 3) did not find a genetic subdivision between Sardinian and north-western wild samples, suggesting the presence of a single panmictic unit. These results likely reflect a high level gene flow among the western Mediterranean populations, consistent with the marine circulation pattern in this region (Franchini et al. 2012). In addition to the different geographic origin of the broodstocks of farmed samples with respect to the wild specimens, patterns of genetic structuring might be due to the lower effective population sizes of farmed populations rather than the use of a small number of founders (Table 4). Our results suggest that random genetic drift due to small N_e may be as or more important than geographic origin as a factor explaining genetic divergence in farmed populations. For instance the most genetically divergent population was that showing the smallest effective population size (MOB, mean N_e = 23). Conversely, the farmed populations with greater Ne (MAB and GAB) were those less genetically divergent from either wild and farmed populations as well.

The most striking result of this study is at odds with our expectation of asymmetric gene flow from farmed to wild populations due to escapes from finfish farms. This expectation is grounded on the fact that a significant fraction of escaped farmed individuals survive in natural habitats where compete for food resources (Arechavala-Lopez et al. 2012). As escapes may lead over time to admixture between wild and

farmed samples we may expect to find in natural populations individuals with a genetic makeup overall or partly similar to that observed in farmed populations. Instead we observed the opposite pattern as a not trivial fraction of the genetic makeup of many farmed individuals was similar to the genetic makeup of wild specimens (Fig. 3). In particular, all individuals from MAB were actually admixed, showing a genetic makeup partly similar to that of wild individuals and partly similar to the genome of other farmed individuals. These results were also mirrored by individual self-assignment, as farmed individuals were more often assigned to wild populations than the opposite (Tables 5 and 6). As we can rule out the use of local broodstocks in the western Mediterranean finfish farms, we may advance the hypothesis that this result might underpin the consequences of past escapes from farms. All the four Sardinian farms sampled in this study reported escape events since the onset of the gilthead sea bream farming (Table S1). More than a decade of frequent escapes, especially during the first years, as well as cage break events due to sea-floods may have already altered the genetic composition of natural populations as suggested by Franchini et al. (2012). To this end it should be noted that at least 16 finfish farms have been rearing gilthead sea breams along Sardinian coasts during these years (www.acquacoltorisardi.com). The picture may be further complicated by the fact that release of male and female gametes from farmed individuals may have occurred in the past (Dimitriou et al. 2007) resulting in extensive admixture among wild and farmed individuals.

In conclusion, the results of this study further highlight the need for a more thorough genetic monitoring of reared stocks and their potential impact on natural populations. However this task in the western Mediterranean may have been yet hampered by the

escapes occurred over time in the past leading to the homogenization of wild populations. Future studies should investigate the genetic variation of broodstocks supplying fry to finfish farms; such information can be used either as a baseline to trace back the genome fraction of wild populations originated from reared stocks, either to simulate the future impact of farm escapes into wild populations.

Acknowledgements

We are indebted to the four finfish farms that kindly allow to collect fin-clip samples of and give us information about their activity and the origin of their stock: Maricoltura Alghero Srl (Alghero), Marina 2000 SNC (Calasetta), Compagnie ittiche Riunite – CIR Srl (Golfo Aranci), and Sardegna Maricoltura Srl (Orosei). We wish also thank Dr. ferruccio Maltagliati and Prof. Pierre Quignard that collect fin-clip samples from Marina di Massa (MMW) and Mauguio lagoon (LMW), respectively. This work was part of the LM's Ph.D. thesis.

References

Alarcón J.A., Magoulas A., Georgakopoulos. Zourosb, T.,E, Alvareza M.C. (2004) Genetic comparison of wild and cultivatedEuropean populations of the gilthead sea bream (*Sparus aurata*). Aquaculture 230: 65 – 80

Anderson E.C., Waples R.S., Kalinowski S.T. (2008) An improved method for estimating the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences 65: 1475-1486

Antao T., Lopes A., Lopes R.J., Beja-Pereira A., Luikart G. (2008) LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics 9:323

Arechavala-Lopez P., Uglem I., Fernandez-Jover D., Bayle-Sempere J.T., Sanchez-Jerez P. (2012) Post-escape dispersion of farmed sea bream (*Sparus aurata* L.) and recaptures by Local fisheries in the Western Mediterranean Sea. Fisheries Research 121-122: 126–135

Beaumont M.A., Nichols R.A. (1996) Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society of London, Series B: Biological Sciences 263: 1619–1626

Beaumont M.A., Balding D.J. (2004) Identifying adaptive genetic divergence among populations from genome scans. Molecular Ecology 13: 969–980

Benjamini Y., Yekutieli D. (2001) The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics 29: 1165–1188

Boyra A., Sanchez-Jerez P., Tuya F., Espino F., Haroun R. (2004) Attraction of wild coastal fishes to Atlantic subtropical cage fish farms, Gran Canaria, Canary Islands. Environmental Biology of Fishes 70: 393–401

Chapuis M.-P., Estoup A. (2007) Microsatellite null alleles and estimation of population differentiation. Molecular and Biological Evolution 24: 621–631

Chapuis M.-P., Loiseau A., Michalakis Y., Lecoq M., Franc A., Estoup A. (2009) Outbreaks, gene flow and effective population size in the migratory locust, *Locusta migratoria*: a regional-scale comparative survey. Molecular Ecology 18: 792–800

Cornuet J.M., Luikart G. (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001–2014

Coscia I., Vogiatzi E., Kotoulas G., Tsigenopoulos T.S., Mariani S. (2011) Exploring neutral and adaptive processes in expanding populations of gilthead sea bream, *Sparus aurata* L., in the North-East Atlantic. Heredity 108: 537 - 546

Dakin E.E., Avise J.C. (2004) Microsatellite null alleles in parentage analysis. Heredity 93: 504–509

De Innocentiis S., Lesti A., Livi S., Rossi A. R., Crosetti D., Sola L. (2004) Microsatellite markers reveal population structure in gilthead sea bream *Sparus auratus* from the Atlantic Ocean and Mediterranean Sea. Fisheries Science 70: 852–859

De Innocentiis S., Miggiano E., Ungaro A., Livi S., Sola L., Crosetti D. (2005) Geographical origin of individual breeders from gilthead sea bream (*Sparus auratus*) hatchery broodstocks inferred by microsatellite profiles. Aquaculture 247: 227– 232

Dempster T., Sanchez-Jerez P., Bayle-Sempere J.T., Giménez-Casalduero F., Valle C. (2002) Attraction of wild fish to sea-cage fish farms in the south-western Mediterranean Sea: spatial and short-term temporal variability. Marine Ecology Progress Series 242: 237–252

Dempster T., Fernandez-Jover D., Sanchez-Jerez P., Tuya F., Bayle-Sempere J., Boyra A., Haroun R.J. (2005) Vertical variability of wild fish assemblages around sea-cage fish farms: implications for management. Marine Ecology Progress Series 304: 15–29

Dempster T., Moe H., Fredheim A., Jensen Q., Sanchez-Jerez P. (2007) Escapes of marine fish from sea-cage aquaculture in the Mediterranean Sea: status and prevention. In: Briand, F. (Ed.), Impact of Mariculture on Coastal Ecosystems, Lisboa, 21–24 February, CIESM Workshop Monographs. No. 32, 86 p.

Dieringer D., Schlötterer C. (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Molecular Ecology Notes: 3: 167–169

Dimitriou E., Katselis G., Moutopoulos D.K., Akovitiotis C., Koutsikopoulos C. (2007) Possible influence of reared gilthead sea bream (*Sparus aurata*, L.) on wild stocks in the area of the Messolonghi lagoon (Ionian Sea, Greece). Aquaculture Research 38: 398–408

Do C., Waples R.S., Peel D., Macbeth G.M., Tillett B.J., Ovenden J.R. (2014) NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Molecular Ecology Resources 14: 209-214

Efron B. (1983) Estimating the error rate of a prediction rule: improvement on cross-validation. Journal of the American Statistical Association 78: 316–331

Excoffier L., Estoup A., Cornuet J.M. (2005) Bayesian analysis of an admixture model with

mutations and arbitrarily linked markers. Genetics 169: 1727–1738

Excoffier L., Hofer T., Foll M. (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103: 285–298

Excoffier L., Lischer H.E.L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564-567

Falush D., Stephens M., Pritchard J.K.(2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

FAO (1995) Code of conduct for responsible fisheries. FAO, Rome, 41 p.

Faulks L.K., Gilligan D.M., Beheregaray L.B. (2010) The role of anthropogenic vs natural in stream structures in determining connectivity and genetic diversity in an endangered freshwater fish, Macquarie perch (*Macquaria australasica*). Evolutionary Applications 4: 589–601

Fernandez-Jover D., Sanchez-Jerez P., Bayle-Sempere J.T., Valle C., Dempster T. (2008) Seasonal patterns and diets of wild fish assemblages associated to Mediterranean coastal fish farms. ICES Journal of Marine Science 65: 1153–1160

Franchini P., Sola L., Crosetti D., Milana V., Rossi A.R. (2012) Low levels of population genetic structure in the gilthead sea bream, *Sparus aurata*, along the coast of Italy. ICES Journal of Marine Science 69: 41–50

Garnier S., Alibert P., Audiot P., Prieur B., Rasplus J.Y. (2004) Isolation by distance and sharp discontinuities in gene frequencies: implications for the phylogeography of an alpine insect species, *Carabus solieri*.Molecular Ecology 13: 1883-1897

Garza J.C., Williamson E.G. (2001) Detection of reduction in population size using data from microsatellite loci. Molecular Ecology 10: 305–318

Griffiths A. M., Machado-Schiaffino G., Dillane E., Coughlan J., Horreo J.L., Bowkett

A. E., Minting P., Toms S., Roche W., Gargan P., McGinnity P., Cross T., Bright D., Garcia-Vazquez E., Stevens, J. R. (2010) Genetic stock identification of Atlantic salmon (*Salmo salar*) populations in the southern part of the European range. BMC Genetics 11, 31

Hardy O. J., Vekemans X. (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2: 618-620

Hardy O. J., Charbonnel N., Fréville H., Heuertz M. (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. Genetics 163: 1467-1482

Hothorn T., Hornik K., Van de Wiel M.A., Zeileis A. (2006). A Lego System for Conditional Inference. The American Statistician 60: 257-263

Hothorn T., Hornik K., Van de Wiel M.A., Zeileis A. (2006). Implementing a Class of Permutation Tests: The coin Package. Journal of Statistical Software 28(8), 1-23. URL http://www.jstatsoft.org/v28/i08/

Hubisz M.J., Falush D., Stephens M., Pritchard J. (2009) Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9: 1322-1332

Jakobsson M., Rosenberg N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics23: 1801-1806

Kalinowski S.T. (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling designs. Conservation Genetics 5: 539-543

Kalinowski S.T. (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Molecular Ecology Notes 5: 187-189

Kalinowski S.T., Manlove K.R., Taper M.L. (2007) ONCOR: a computer program for genetic stock identification. http://www.montana.edu/kalinowski/Software/ONCOR.htm

Laikre L., Schwartz M.K., Waples R.S., Ryman N. (2010) Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. Trends in Ecology and Evolution 25: 520–529

Loukovitis D., Sarropoulou E., Vogiatzi E., Tsigenopoulos C.S., Kotoulas G., Magoulas A., Chatziplis D. (2012) Genetic variation in farmed populations of the gilthead sea bream *Sparus aurata* in Greece using microsatellite DNA markers. Aquaculture Research 43: 239 – 246

Maruyama T., Fuerst P.A. (1985) Population bottlenecks and non-equilibrium models in population genetics. Genetics 111: 675 – 689

Miggiano E., De Innocentiis S., Ungaro A., Sola L., Crosetti D. (2005) AFLP and microsatellites as genetic tags to identify cultured gilthead sea bream escapees: data from a simulated floating cage breaking event. Aquaculture International 13: 137–146

Luikart G., Allendorf F.W., Cornuet J.M., Sherwin W.B. (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. Journal of Heredity 89: 238–247

Navarro A., Badilla R., Zamorano M.J., Pasamontes V., Hildebrandt S., Sánchez J.J., Afonso J.M. (2008) Development of two new microsatellite multiplex PCRs for three sparid species: Gilthead sea bream (*Sparus auratus* L.), red porgy (*Pagrus pagrus* L.) and redbanded sea bream (*P. auriga*, Valenciennes, 1843) and their application to paternity studies. Aquaculture 285: 30–37

Norris A.T., Bradley D.G., Cunningham E.P. (1999) Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. Aquaculture 180: 247 – 264

Paetkau D., Slade R., Burden M., Estoup A. (2004) Direct, real-time estimation of migration rate using assignment methods: a simulation-based exploration of accuracy and power. Molecular Ecology13: 55-65

Piry S., Luikart G., Cornuet J.M.(1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. The Journal of Heredity: 90: 502-503

Piry S., Alapetite A., Cornuet J.-M., Paetkau D., Baudouin L., Estoup A. (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95: 536 – 539

Porta J., Porta J. M., Béjar J., Del Carmen Álvarez M.(2010) Development of a microsatellite multiplex genotypingtool for the fish Gilthead sea bream (*Sparus aurata*): applicability in population genetics and pedigree analysis. Aquaculture Research 41: 1514-1522

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

Rannala B., Mountain J.L., (1997) Detecting immigration by using multilocus genotypes.

Proceedings of the National Academy of Sciences U.S.A. 94: 9197–9201

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Rosenberg N.A., Burke T., Elo K., Feldman M.W., Freidlin P.J., Groenen M.A.M., Hillel J., Mäki-Tanila A., Tixier-Boichard M., Vignal A., Wimmers K., Weigend S. (2001) Empirical evaluation of genetic clustering methods from 20 chicken breeds. Genetics 159: 699 – 713

Rosenberg N.A., Pritchard J.K., Weber J.L., Cann H.M., Kidd K.K., Zhivotovsky L.A., Feldman M.W. (2002) Genetic structure of human populations. Science 298: 2381-2385

Rosenberg N.A. (2004) Distruct: a program for the graphical display of structure results. Molecular Ecology Notes 4: 137-138

Rousset F. (2008) Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. Molecular Ecology Resources 8: 103–106

Šegvić-Bubić T., Lepen I., Trumbić Ž., Ljubković J., Sutlović D., Matić-Skoko S., Grubišić L., Glamuzina B., Mladineo I. (2011) Population genetic structure of reared and wild gilthead sea bream (*Sparus aurata*) in the Adriatic Sea inferred with microsatellite loci. Aquaculture 318: 309–315

Tuya F., Boyra A., Sanchez-Jerez P., Haroun R. (2005) Non-metric multivariate analysis of the demersal ichthyofauna along soft bottoms of the Eastern Atlantic:

comparison between unvegetated substrates, seagrass meadows and sandy bottoms under the influence of sea-cage fish farms. Marine Biology 147: 1229–1237

Tuya F., Sanchez-Jerez P., Dempster T., Boyra A., Haroun R. (2006) Changes in demersal wild fish aggregations beneath a sea-cage fish farm after the cessation of farming. Journal of Fish Biology 69: 682–697

Vähä J.-P., Erkinaro J., Niemela E., Primmer C.R. (2007) Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. Molecular Ecology 16: 2638-2654

Vandewoetijne S., Van Dyck H. (2010) Population genetic differences along a latitudinal cline between original and recently colonized habitat in a butterfly. PLoS ONE, 5: e13810

Valle C., Bayle-Sempere J.T., Dempster T., Sanchez-Jerez P., Gimenez-Casualdero F. (2007) Temporal variability of wild fish assemblages associateith a sea cage fish farm in the southwestern Mediterranean Sea. Estuarine Coastal and Shelf Science 72: 299–307

Van Oosterhout C., Hutchinson W.F., Wills D.P.M., Shipley P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538

Waples R. S., Do C. (2010) Linkage disequilibrium estimates of contemporary *Ne* using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. Evolutionary Applications 3: 244-262

Wasko A.P., Martins C., Oliveira C., Foresti F. (2003) Non-destructive genetic sampling in fish. An improved method for DNA extraction from f sh fi ns and scales. Hereditas 138: 161-165

Weir B.S., Cockerham C.C. (1984) Estimating F-Statistics for the Analysis of Population Structure, Evolution 38: 1358-1370

Whitlock M.C. (2011) G'_{ST} do not replace F_{ST}. Molecular Ecology 20: 1083 – 1091

Figures and Tables

Fig. 1 Sampling map of *Sparus aurata*. Grey dots indicate the location of the Sardinian farms from which farmed specimens were collected. As we do not know the exact location of farmed samples from Greece and Turkey, we shaded the countries of origin in light and dark gray, respectively. Boxes outline approximately the fishing grounds in which wild gilthead sea breams were sampled.

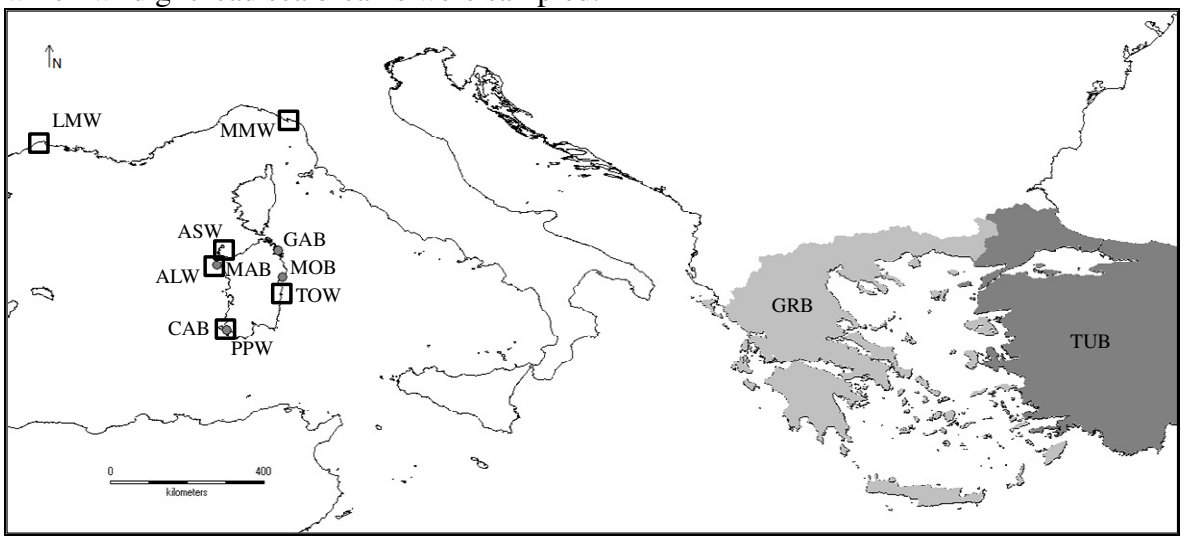

Fig. 2 Non-Metric Multidimensional scaling (nMDS) ordination of wild and farmed samples of Sparus aurata. The plot depicts the first two dimensions of the nMDS based on the observed pairwise FST values. The stress value indicates a good fit of the ordination to the original data.

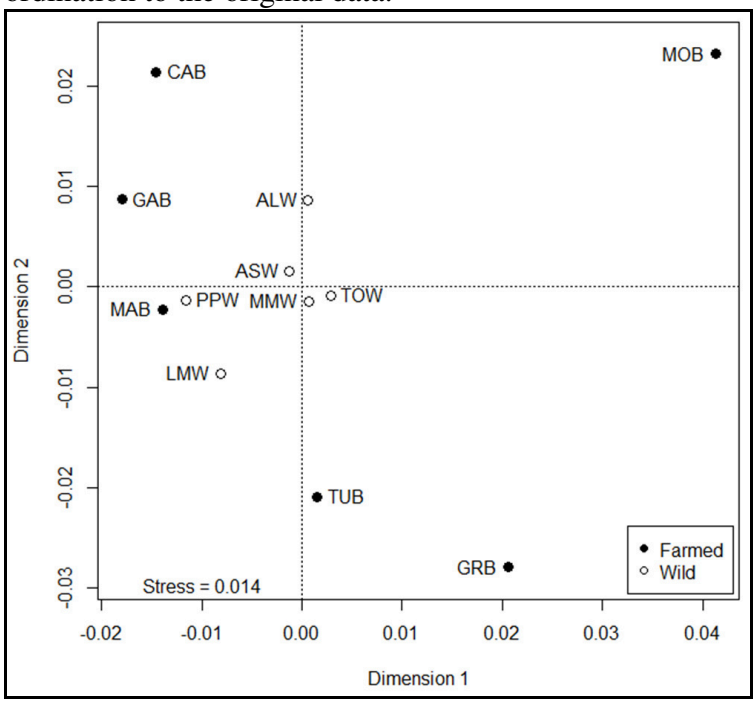

Fig. 3 Bayesian model-based clustering barplot. results correspond to the number of groups (K) for which the additive information due to clusters was maximum (K = 6). Individuals are represented by vertical segmented vertical bars; the color of each segment corresponds to a different cluster and the height to the membership coefficient into that cluster (Q). a) barplot corresponding to the best likelihood run at K = 6; b) barplot showing the individual memebership to each cluster averaged over 10 runs, using the large K Greedy algorhithm implemented in CLumpp (Jakobson & Rosenberg 2007).

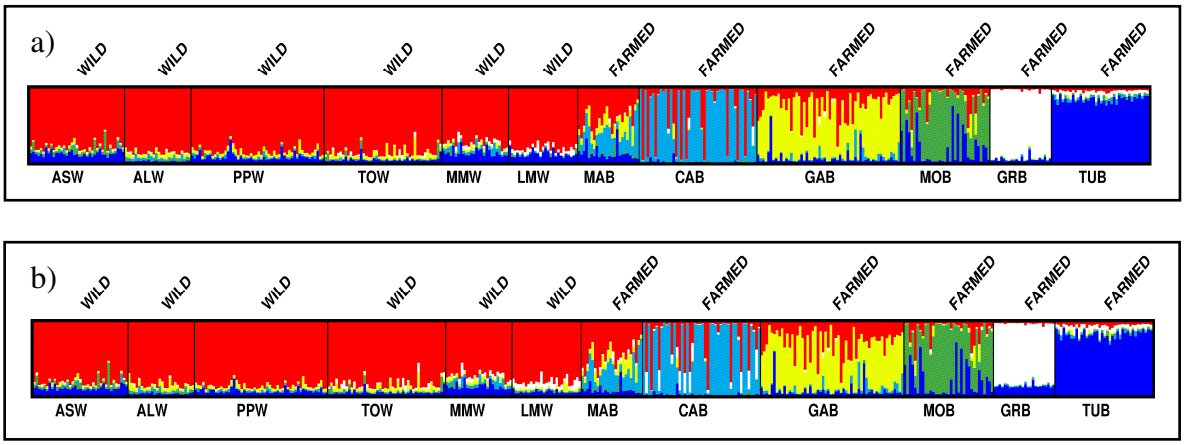

Table 1 Details of gilthead sea bream (*Sparus aurata*) samples analyzed in this study: ID = identification label; N = sample size.

ID[1]	Type	Locality	Origin	Year	N
ASW	Wild	Gulf of Asinara	Italy	2011 – 2012	37
ALW	Wild	Bay of Alghero	Italy	2011 – 2013	26
PPW	Wild	Porto Pino	Italy	2012	52
TOW	Wild	Tortolì	Italy	2011	46
MMW	Wild	Marina di Massa	Italy	2012	26
LMW	Wild	Mauguiolagoon	France	2013	27
MAB	Farmed	Bay of Alghero	Italy	2009	24
CAB	Farmed	Calasetta	Italy	2012	46
GAB	Farmed	Golfo Aranci	Italy	2012	56
MOB	Farmed	Marina di Orosei	Italy	2012	35
GRB	Farmed	Unknown	Greece	2011	24
TUB	Farmed	Unknown	Turkey	2011	38

[1] the first two letters of the identification label indicate the locality or the farm where individuals are collected, and the third letter whether samples are of wild or farmed origin

Table 2 Summary statistics of within population genetic variation averaged over loci for each gilthead sea bream (*Sparusaurata*) sample. Populations are abbreviated as in Table 1; A = number of alleles, A_R = allelic richness, A_{PR} = private allelic richness, H_E = expected heterozigosity, H_O = observed heterozigosity, F_{IS} = fixation index.

	A	A_R	A_{PR}	H_E	H_O	F_{IS}
ASW	13.6 ± 1.4	11.3 ± 1.1	0.8 ± 0.2	0.80 ± 0.04	0.77 ± 0.04	0.025 ± 0.035
ALW	12.0 ± 1.3	11.2 ± 1.1	0.7 ± 0.2	0.81 ± 0.03	0.77 ± 0.04	0.053 ± 0.032
PPW	15.1 ± 1.6	11.6 ± 1.2	1.0 ± 0.2	0.80 ± 0.04	0.77 ± 0.04	0.034 ± 0.030
TOW	14.6 ± 1.8	11.7 ± 1.3	0.8 ± 0.2	0.81 ± 0.04	0.77 ± 0.04	0.056 ± 0.023
MMW	12.4 ± 1.5	11.5 ± 1.3	0.9 ± 0.2	0.80 ± 0.04	0.78 ± 0.04	0.028 ± 0.029
LMW	13.4 ± 1.6	12.2 ± 1.4	1.1 ± 0.3	0.79 ± 0.04	0.78 ± 0.04	0.015 ± 0.022
MAB	11.2 ± 1.9	10.7 ± 1.0	0.6 ± 0.1	0.78 ± 0.04	0.75 ± 0.04	0.020 ± 0.040
CAB	10.8 ± 1.1	9.0 ± 0.8	0.3 ± 0.1	0.77 ± 0.04	0.73 ± 0.05	0.053 ± 0.033
GAB	13.6 ± 1.4	10.4 ± 1.0	0.6 ± 0.1	0.79 ± 0.03	0.73 ± 0.04	0.068 ± 0.032
MOB	9.8 ± 0.9	8.7 ± 0.7	0.3 ± 0.1	0.76 ± 0.05	0.73 ± 0.05	0.019 ± 0.044
GRB	8.6 ± 0.8	8.4 ± 0.7	0.3 ± 0.1	0.76 ± 0.04	0.77 ± 0.04	-0.028 ± 0.037
TUB	10.8 ± 1.1	9.3 ± 0. 9	0.4 ± 0. 1	0.76 ± 0.04	0.73 ± 0.04	0.025 ± 0.022

Table 3 Pairwise FST values (below diagonal) and the probability each value is significantly different from zero (above diagonal). Due to multiple testing probability values have been corrected using the False Discovery Rate (FDR) method suggested by Benjamini and Yekutieli (2001).

	ASW	ALW	PPW	TOW	MMW	LMW	MAB	CAB	GAB	MOB	GRB	TUB
ASW		0.164	1	0.275	1	0.144	**0.02**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**
ALW	0.005		0.18	1	1	0.059	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**
PPW	0.001	0.004		0.14	1	0.247	**0.034**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**
TOW	0.003	0.001	0.003		1	0.185	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**
MMW	-0.003	0.000	0.001	-0.001		1	**0.024**	**0.001**	**0.005**	**0.001**	**0.001**	**0.001**
LMW	0.004	0.007	0.003	0.004	0.002		0.164	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**
MAB	0.007	0.016	0.007	0.011	0.009	0.006		**0.019**	**0.004**	**0.001**	**0.001**	**0.001**
CAB	0.025	0.025	0.023	0.026	0.022	0.026	0.013		**0.001**	**0.001**	**0.001**	**0.001**
GAB	0.010	0.013	0.010	0.012	0.010	0.014	0.010	0.028		**0.001**	**0.001**	**0.001**
MOB	0.031	0.037	0.046	0.036	0.032	0.044	0.046	0.063	0.052		**0.001**	**0.001**
GRB	0.030	0.036	0.034	0.028	0.021	0.027	0.034	0.049	0.040	0.060		**0.001**
TUB	0.018	0.032	0.021	0.023	0.015	0.019	0.019	0.045	0.025	0.053	0.039	

Table 4 Estimates of the mean current effective population size (Ne) and probability of heterozygosity excess (He>Heq). 95% confidence intervals (CI) for Ne estimates are based on jackknifing over loci.Probability that populations underwent a recent bottleneck based on heterozygosity excess were estimated under the Stepwise Mutation Model (SMM) and the Two-Phase Mutation Model (TPM).

	$N_e^{\wedge}$ with 95% CI	Probability heterozigosity excess under SMM	Probability for heterozigosity excess under TPM
ASW	∞ (633 – ∞)	0.998	0.838
ALW	206 (119 – 692)	0.946	0.620
PPW	161 (127 – 214)	1.000	0.838
TOW	∞ (1452 – ∞)	0.988	0.551
MMW	386 (151 – ∞)	0.997	0.779
LMW	∞ (232 – ∞)	0.997	0.868
MAB	145 (85 – 438)	0.990	0.725
CAB	30 (26 – 33)	0.995	0.584
GAB	97 (82 – 118)	0.998	0.893
MOB	23 (21 – 26)	0.756	0.103
GRB	48 (35 – 75)	0.868	0.114
TUB	95 (69 – 146)	0.998	0.837

Table 5 Results of Individual assignment tests using the maximum-likelihood method implemented in ONCOR (Kalinowski et al. 2007): self-assignment of individuals of *Sparus aurata* to the more likely population was performed with the leave-one-out method. Individuals correctly assigned are outlined in bold.

	Assigned to											
Populations	ASW	ALW	PPW	TOW	MMW	LMW	MAB	CAB	GAB	MOB	GRB	TUB
ASW	**10**	2	7	6	5	2	1		1	1		
ALW	2	**1**	4	7	1		1		2			
PPW	5	4	**19**	4	5	7	2	1	4			1
TOW	6	4	2	**17**	1	9			4			1
MMW	6	2	5	4	**1**	2	1	2	3			
LMW	1	2	5	7	2	**1**			2			
MAB	3		2	2	1		**5**	3	3			2
CAB	1		2	1		1	2	**33**	2			
GAB	6	2	3	3	2	1	1		**35**			2
MOB	1	1		2	2				2	**26**		
GRB				1	1						**17**	1
TUB	3		1	1	3				1			**26**
Correctly assigned (%)	28.6	5.6	36.5	38.6	3.8	5.0	23.8	78.6	63.6	76.5	85.0	74.3

Table 6 Results of Individual assignment tests using the bayesian method of Rannala and Mountain (1997) implemented in Geneclass2 (Piry et al. 2004): the probability of self-assignment was estimated simulating 10'000 individuals with the Montecarlo resampling method (Paetkau et al. 2004). Individuals correctly assigned are outlined in bold.

	Assigned to											
Populations	ASW	ALW	PPW	TOW	MMW	LMW	MAB	CAB	GAB	MOB	GRB	TUB
ASW	**5**	2	7	10	8	4	1					
ALW	3	**5**	3	11	3	1						
PPW	4	5	**13**	7	4	19						
TOW	5	9	2	**15**	3	12						
MMW	4	6	5	4	**2**	4	1					
LMW	1	5	6	4	4	**7**						
MAB	2	4	1	5	1	9	**2**					
CAB	3	3	3	4	4	8	4	**17**				
GAB	8	5	6	8	6	5	1		**17**			
MOB	5	3		5	4					**18**		
GRB				3	2	7					**12**	
TUB	3		7	3	7	7			1			**10**
Correctly assigned (%)	13.5	19.2	25.0	32.0	7.7	25.9	8.3	37.0	30.4	51.4	50.0	26.3

Supplementary Material

Figure S1 Bayesian model-based clustering. Plot of the natural logarithm of posterior probabilities of a model given the data (lnP(D)). Error bars represent the standard deviation of lnP(D) over 10 independent runs. P = probability that a model is significantly different from the former computed using a one-tail paired two-sample Wilcoxon sign-rank test is plotted up to the first non-significant test. Due to multiple comparisons probability values wre corrected with FDR method (Benjamini and Yekutieli 2001).

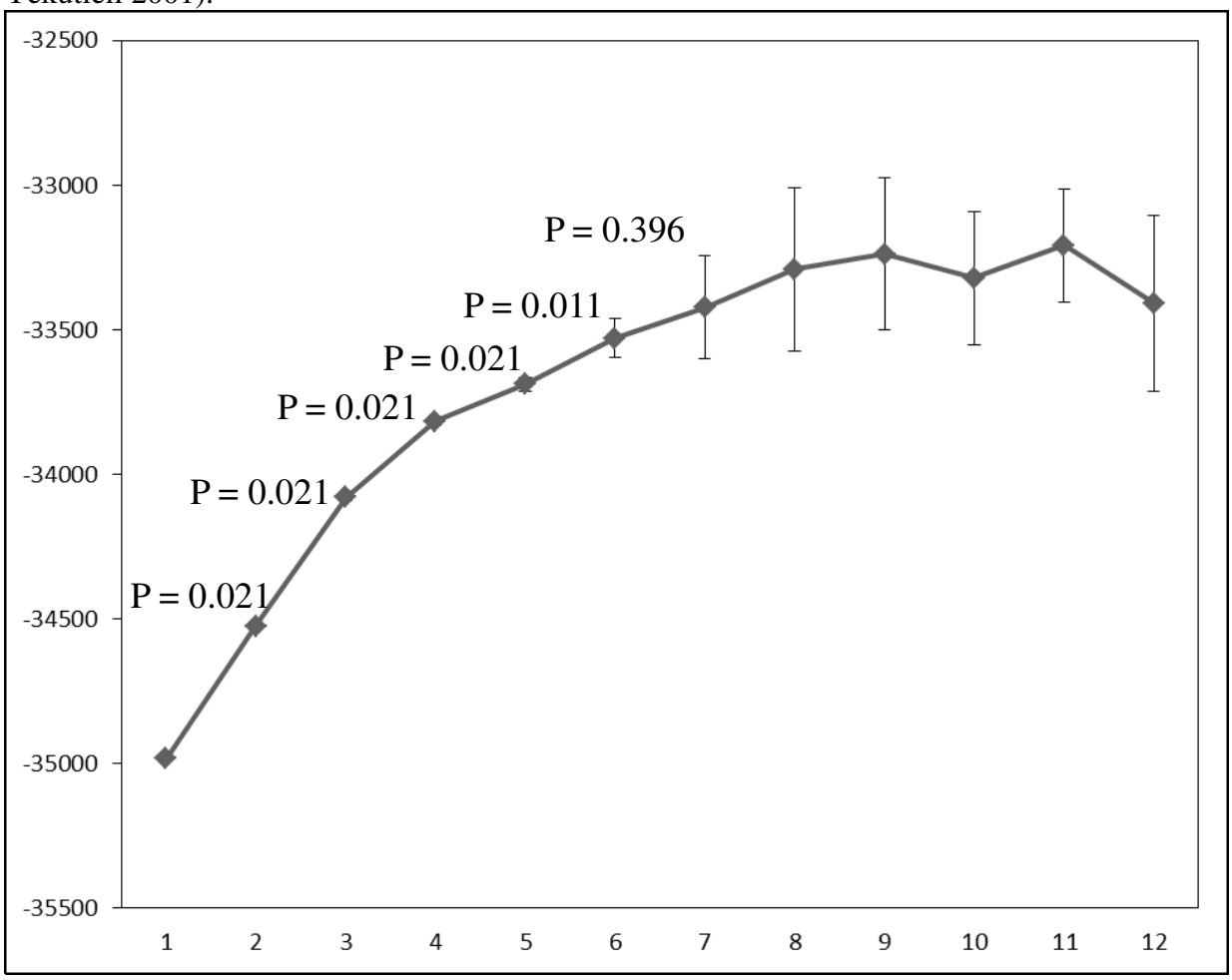

Table S1 Information on the fin-fish farms sampled in this study. Year represents the time since a farm started the rearing of gilthead seabreams. Origin of fry reports the hatcheries supplying fingerlings to the finfish farms.

Code	Name	Locality	Year	Escape causes	Origin of fry
MAB	MaricolturaAlgheroSrl	Alghero	1993	Cage breakdown Other system failure due to extreme sea-floods Operational accidents	Valle Cà Zuliani[a] PanitticaPugliese[b]
CAB	Marina 2000 SNC	Calasetta	2002	Cage breakdown Other system failure	Valle Cà Zuliani[a] PanitticaPugliese[b] Acquacoltura Jonica[c] Ittica Caldoli[d] Acqua Azzurra[e]
GAB	Compagnie Ittiche Riunite Srl	Golfo Aranci	2000	Holes in the cage net Operational accidents	Agroittica Toscana[f]
MOB	Sardegna MaricolturaSrl	Orosei	2000	Hole in the net Extreme sea-floods	PanitticaPugliese[b]
GRB	Unknown	Unknown	Unknown	Unknown	Unknown
TUB	Unknown	Unknown	Unknown	Unknown	Unknown

[a]ValleCàZuliani Soc. Agricola SrlMonfalcone (GO) is located in the North-Adriatic sea and started its activity since the 1980s (www.vallecazuliani.it).

[b]PanitticaPuglieseFasano (BR) is located in the South-Adriatic sea (www.panaquaculture.com).

[c]AcquacolturaJonicaSrl (Taranto), located in the Ionian sea, started its activity since 2006 and uses broodstocks either of Mediterranean either of Atlantic origin (www.acquacolturajonica.it).

[d]ItticaCaldoliSrlLesina(FG) is located in the South Adriatic sea.

[e]AcquaAzzurraS.p.A. Pachino (SR), located in the channel of Sicily between Pachino and Porto Palo di Capo Passero started its activity since 1988 (www.acquaazzurra.it).

[f]Agroittica Toscana Soc. Agr. SrlPiombino (LI), located in the Thyrrenian Sea started its activity since 1994 (www.agroitticatoscana.it). This farm hatches fingerlings that trace back to Valle CàZuliani and Ferme marine de Douhet (Ile D'Oleron, France). The latter hatchery uses broodstocks of both Atlantic and Mediterranean origin to produce its fry (www.douhet.com).

Table S2 Microsatellite loci used in this study.

Locus	Repeat motif	Number of alleles	Allelesize range	GenBank accession number
PbMS2	$(CTT)_6CTC(CTT)_3$	14	122 – 164	AF209085
SaGT32	$(CA)_{32}$	19	139 – 177	Y17264
SaI12	$(GT)_{30}$	20	222 – 266	AY322108
SauD182	$(CA)_9 + (CA)_8$	8	258 – 272	AY173034
SauH98	$(CA)_{17} + (TG)_6$	14	100 – 128	AY172039
SauI47	$(TG)_{19}$	10	93 – 111	AY172041
SauK140	$(CA)_{23}$	17	67 – 103	AY172042
PaGA2A	$(AG)_{16}$	19	106 – 148	AF202885
Sal14	$(GT)_{27}$	34	193 – 263	AY322109
Sal19	$(GT)_{25}$	25	226 – 300	AY322111
Sal21	$(GT)_{41}$	57	119 – 253	AY322112
SauE82	$(CA)_{12}AA(CA)_7$	13	149 – 183	AY172035
SauAN	$(TG)_{18}$	9	150 – 166	AY173032
Sal10	$(GT)_{37}$	23	170 – 230	AY322107
Sal15	$(GT)_{26}$	34	95 – 163	AY322110
pSaGT26	$(CA)_8CT(GT)_{31}$	21	214 – 260	Y17266
SaGT41b	$(AC)_{13}$	26	139 – 195	Y17262
SauE97	$(CA)_{30}$	22	163 - 265	

Table S3 Bayesian clustering results based on the run with the best likelihood at K = 6 showing the cluster membership coefficient (Q) for each population. Highest membership coefficient are outlined in bold. N = number of individuals sampled in each population.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	N
ASW	0.119	0.044	0.019	0.026	0.011	**0.781**	37
ALW	0.017	0.019	0.033	0.050	0.017	**0.864**	26
PPW	0.071	0.005	0.022	0.048	0.005	**0.848**	52
TOW	0.035	0.023	0.006	0.055	0.022	**0.859**	46
MMW	0.131	0.013	0.054	0.036	0.037	**0.730**	26
LMW	0.111	0.020	0.005	0.000	0.060	**0.805**	27
MAB	0.120	0.035	0.269	0.150	0.014	**0.412**	24
CAB	0.009	0.003	**0.681**	0.009	0.013	0.286	46
GAB	0.057	0.002	0.040	**0.665**	0.010	0.226	56
MOB	0.185	**0.647**	0.002	0.043	0.003	0.120	35
GRB	0.041	0.004	0.009	0.008	**0.925**	0.014	24
TUB	**0.854**	0.011	0.033	0.014	0.059	0.030	38

Table S4 Bayesian clustering results averaged on the ten runs at K = 6 showing the cluster membership coefficient (Q) for each population. Highest membership coefficient are outlined in bold. N = number of individuals sampled in each population.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	N
ASW	0.1013	0.0452	0.0191	0.0265	0.0134	**0.7944**	37
ALW	0.0212	0.0161	0.0317	0.0746	0.0170	**0.8394**	26
PPW	0.0547	0.0050	0.0225	0.0486	0.0077	**0.8615**	52
TOW	0.0365	0.0242	0.0067	0.0583	0.0385	**0.8358**	46
MMW	0.1222	0.0149	0.0562	0.0431	0.0450	**0.7185**	26
LMW	0.0440	0.0125	0.0095	0.0363	0.0922	**0.8054**	27
MAB	0.0703	0.0289	0.2732	0.1282	0.0365	**0.4628**	24
CAB	0.0109	0.0026	**0.6847**	0.0070	0.0715	0.2234	46
GAB	0.0603	0.0033	0.0456	**0.6102**	0.0076	0.2730	56
MOB	0.1873	**0.6595**	0.0047	0.0372	0.0091	0.1023	35
GRB	0.1310	0.0037	0.0147	0.0059	**0.8294**	0.0153	24
TUB	**0.8408**	0.0090	0.0272	0.0213	0.0642	0.0375	38

Chapter 3

Preliminary data on the genetic variability

of the clam *Ruditapes decussatus*

Benedetto Cristo, Laura Mura, Gian Luca Dedola, Fabio Scarpa, Tiziana Lai, Piero Cossu, Nicola Fois, Marco Casu, Daria Sanna

Pubblished in *Biologia Marina Mediterranea*, 20 (1): 234-235 (2013)

Biol. Mar. Mediterr. (2013), 20 (1): 234-235

B. Cristo, L. Mura*, G.L. Dedola, F. Scarpa, T. Lai, P. Cossu,
N. Fois*, M. Casu, D. Sanna

Dip. di Scienze della Natura e del Territorio, Sez. di Zoologia, Archeozoologia e Genetica,
Università di Sassari,
Via F. Muroni, 25 - 07100 Sassari, Italy.
darsanna@uniss.it
*Agris Sardegna, Località Bonassai, Olmedo (SS), Italy.

PRELIMINARY DATA ON THE GENETIC VARIABILITY OF THE CLAM *RUDITAPES DECUSSATUS*

DATI PRELIMINARI SULLA VARIABILITÀ GENETICA DEL BIVALVE RUDITAPES DECUSSATUS

Abstract *- The Mediterranean populations of* Ruditapes decussatus *(Linnaeus, 1758) (Veneridae) experienced a decline in the last decades as consequence of commercial exploitation. Any increasing of genetic knowledge, on large geographic-scale, of present populations is essential to address adequate conservation plans for this species. We analysed mtDNA COI region in individuals from Mediterranean and Atlantic. Results pointed out a possible genetic break between eastern and western Mediterranean basins and a molecular affinity between western Mediterranean and Atlantic populations. Further analysis on an enlarged dataset of samples could shed new light on the molecular trends here evidenced.*

Key-words: *grooved carpet shell clam, genetics, population dynamics.*

Introduction - The grooved carpet shell clam, *Ruditapes decussatus* is a bivalve considered as one of the most important for Mediterranean aquaculture, which experimented a reduction of its population as possible consequence of an unregulated commercial exploitation. Furthermore, the introduction for commercial purposes in the Mediterranean of the Manila clam, *R. philippinarum,* native to the Pacific coast of Asia, favoured progressive decline of indigenous populations of *R. decussatus.* In order to provide management plans devoted to preserve the genetic variability of this species, it is of great relevance knowing the present status of Mediterranean *R. decussatus* genetic variation on a geographic scale as large as it is possible. Our study represents an effort in this sense, as we analysed mitochondrial COI gene relationships among samples from the eastern and western Mediterranean and Atlantic.

Materials and methods - Specimens of *R. decussatus* from three sites from Italy - Olbia (Sardinia), Orbetello (Tuscany), and Messina (Sicily) - were sampled. A portion of mitochondrial Cytochrome c Oxidase subunit I gene was amplified and then sequenced using specific primers designed by the authors according to the methodologies previous described by Sanna and colleagues (Sanna *et al*., 2012 and references therein). All COI sequences available on GenBank were added to our dataset obtaining a 541 bp long alignment including individuals of 6 Mediterranean (from France, Italy, Tunisia, and Turkey) and 1 Atlantic (from Portugal) sites. Estimates of genetic diversity were obtained using DnaSP 5.10 (Rozas and Rozas, 1999). Genetic relationships among individuals were inferred using the program Network 4.6.1.1 (www.fluxus-engineering.com). The occurrence of genetic structuring among samples was investigated by Baps 5.4 (Corander and Tang, 2007). Historical population dynamics were estimated by mismatch distribution analysis using DnaSP 5.10.

Results - On 23 individuals analysed, 13 haplotypes, identified by 19 polymorphic sites, were found. The most common haplotype spread in the 48% of individuals including samples from France, Italy, and Tunisia, 10 haplotypes were present in a single individual, while samples from Turkey and Portugal did not share haplotypes

with remaining individuals. Noteworthy, high levels of genetic diversity occurred among samples (h=0.850 and π=0.006). Network analysis pointed out a genetic divergence (represented by 7 mutations) between samples from Turkey and the others. Among western Mediterranean and Atlantic individuals, a star-like fashion evidenced the occurrence of a main haplotype diverging from derived ones for 1-2 mutations only, according to a founder effect model. Consistently, the mismatch distribution analysis showed a bi-modal shape. However, Bayesian analysis evidenced the presence of a single group of haplotypes as possible consequence of the low number of mutations which differentiated Turkish samples from the remaining.

Conclusions - Results obtained suggest the presence of a moderate level of genetic structuring between *R. decussatus* populations from eastern and western Mediterranean. This finding is consistent with the occurrence of a biogeographic boundary, usually located at the Siculo-Tunisian Strait, which reduces gene flow between the two basins (Bianchi, 2007). Our data also support the possible occurrence of a bottleneck among western Mediterranean sites (France, Italy and Tunisia) as it was previously observed in other populations of *R. decussatus* (Gharbi *et al.*, 2010; Manunza *et al.*, 2010). Noteworthy, samples from western Mediterranean show a genetic similarity with samples from Portugal, as possible consequence of an efficient larval dispersal across Gibraltar Strait. A similar trend was found in *Parcentrotus lividus* (Maltagliati *et al.*, 2010). However we should take into account that the low number of mutations separating eastern from western Mediterranean could not be conclusive to infer on the possible occurrence of a genetic structuring between the two basins that must be confirmed by means of further analysis on a larger number of samples.

References

BIANCHI CN. (2007) - Biodiversity issues for the forthcoming tropical Mediterranean Sea. *Hydrobiologia*, **580**: 7-21.

CORANDER J., TANG J. (2007) - Bayesian analysis of population structure based on linked molecular information. *Math. Biosci.*, **205**: 19-31.

GHARBI A., CHATTI N., SAID K., VAN WORMHOUDT A. (2010) - Genetic variation and population structure of the carpet shell clam *Ruditapes decussatus* along the Tunisian coast inferred from mtDNA and ITS1 sequence analysis. *Biologia*, **65** (4): 688-696.

MALTAGLIATI F., DI GIUSEPPE G., BARBIERI M., CASTELLI A., DINI F. (2010) - Phylogeography and genetic structure of the edible sea urchin *Paracentrotus lividus* (Echinodermata: Echinoidea) inferred from the mitochondrial cytochrome b gene. *Biol. J. Linnean Soc.*, **100**: 910-923.

MANUNZA A., MASCIA D., DEMURTAS R., DEIANA AM., CANNAS R. (2010) - On the genetic structure of *Ruditapes decussatus* (Mollusca, Bivalvia) inferred by nuclear and mitochondrial genetic markers. *Rapp. Comm. int. Mer Médit.*, **39**: 384.

ROZAS J., ROZAS R. (1999) - DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**: 174-175.

SANNA D., DEDOLA G.L., LAI T., CURINI-GALLETTI M., CASU M. (2012) - PCR-RFLP: a practical method for the identification of specimens of *Patella ulyssiponensis s.l.* (Gastropoda: Patellidae). *Ital. J. Zool.*, **79**: 50-59.

Chapter 4

Genetic structuring of the carpet shell clam *Ruditapes decussatus* along Mediterranean and northern Atlantic coast

Daria Sanna, Laura Mura, Fabio Scarpa, Gian Luca Dedola, Piero Cossu, Tiziana Lai, Benedetto Cristo, Marcella Carcupino, Marco Curini-Galletti, Nicola Fois, Marco Casu

manuscript

Genetic structuring of the carpet shell clam *Ruditapes decussatus* along Mediterranean and northern Atlantic coast

D. Sanna[1], L. Mura[2], F. Scarpa[1], G.L. Dedola[1], P. Cossu[1],T. Lai[1], B. Cristo[1], M. Carcupino[1], M. Curini-Galletti[1], N. Fois[2], M. Casu[1]

[1] Dipartimento di Scienze della Natura e del Territorio – Sezione di Zoologia, Archeozoologia e Genetica, Università di Sassari, Via F. Muroni 25, 07100 Sassari (SS), Italy

[2] Agris Sardegna, Agenzia per la Ricerca in Agricoltura. Dipartimento per la Ricerca nelle Produzioni Animali, Località Bonassai, 07040 Olmedo (SS), Italy

Corresponding author:

Marco Casu

Dipartimento di Scienze della Natura e del Territorio

Università di Sassari

Via F. Muroni 25

07100 Sassari (SS), Italy

Phone: +39 079 228924

Fax: +39 079 228665

marcasu@uniss.it

ABSTRACT

The grooved carpet shell, *Ruditapes decussatus* is a commercially valuable bivalve mollusk authoctnous in Mediterranean. Due to the introduction of the allochthonous *R. philippinarum* into Europe during the '70s and '80s for fishery purposes and the intense exploitation for human consumption, this species has lead to serious risk of genetic erosion, till to local extinction. Here, we aimed to investigate the level of genetic variability and the occurrence of genetic structuring of *R. decussatus* in Atlantic and Mediterranean a geographic areas, with particular reference to the central-west Mediterranean, analysing sequences of mitochondrial gene for the subunit I of the Cytochrome c Oxidase (COI). The genetic variability we detected is homogenously distributed throughout the area analysed. The obtained network shows a star-like configuration, with many derived, unique haplotypes separated by one point-mutation from one geographically widespread haplotype. This suggests that *R. decussatus* experienced in this area a severe decline of effective population size, likely followed by a recent period of rapid population growth. This condition may be explaned by a contractions of its populations occurred in the past as a consequence of the repeated sea-level fluctuations during Pleistocene glaciation and/or a more recent reduction of population sizes may has been induced, at least locally, by overexploitation.

KEY WORDS: mtDNA, COI, sequencing, genetic variation

1. INTRODUCTION

The grooved carpet shell, *Ruditapes decussatus* (Linnaeus 1758) is a commercially valuable bivalve mollusk. The Food and Agriculture Organisation of the United Nations (FAO 2006) recorded for 2006 a production of the commercially-called "clams, cockles and arkshells" (in which *R. decussatus* is included) of about 740 tons. *Ruditapes decussatus* is naturally distributed along the Mediterranean coasts, and from the coasts of Senegal to Norway in the Atlantic. It is a highly tolerant, euryhaline species, which is present in shallow-water marine and brackish habitats, the adult sedentarily living in muddy-sand sediments. *Ruditapes decussatus* is a gonochoric species with external fertilization and planktotrophic larval stages. Reproduction occurs during spring and summer seasons. Hatchery experiments have shown that the PLD (planktonic larval duration) is about 10–20 days depending on temperature.

Although *R. decussatus* is a species with high range of distribution, locally abundant, and frequently harvested for the shellfish market, particularly in Europe, Turkey and north African countries, few studies focused on surveying its genetic variability. Wild populations of *R. decussatus* have been used as model - generally, bivalves, with their sessile adult life and passive planktonic dispersal are a good model for surveying phylogeographic pattern shaped by the water circulation to unravel the role of barriers to gene flow, at both a small (Gharbi et al. 2010) and a large geographic scale (Cordero et al., 2013; and references therein). Considering its high commercial values, other studies focused on the molecular discrimination between this species and the pullet carpet shell, *Venerupsis pullastra* (Montagu, 1803), often marketed together. Indeed, the correct labeling of molluscs, and fishery resources in general, is crucial in order to

unequivocally identify the food products that consumers acquire. Molecular tools were also recently used to verify the possibility and the amount of hybridization between *R. decussatus* and the introduced allochthonous *Ruditapes philippinarum* (Adams and Reeve, 1850). Indeed, the introduction of *R. philippinarum*, native of the West Pacific coasts, from Russia to the Philippines, into Europe during the '70s and '80s for fishery purposes, has led to a dramatic reduction of *R. decussatus*, whose populations suffer from competition with *R. philippinarum*.

This fact, and the contemporary intense exploitation of this species for human consumption (FAO, 2000) in different Mediterranean and Atlantic areas undergoes this species to serious risk of genetic erosion, till to local extinction.

In this paper we aimed to investigate the level of genetic variability and the occurrence of genetic structuring of *R. decussatus* in Atlantic and Mediterranean a geographic areas, with particular reference to the central-west Mediterranean, so far under-sampled, which is characterised by an intense activity of harvesting of the species for commercial purposes (FAO, 2000). In particular, we improved sampling plan on the coasts around the Sardinian-Corsican region. Within this region, the level of genetic variability of wild populations would be peculiar, considering that i) *R. decussatus* represents an important marine fishery resource exploited by the coastal countries (Turolla, 2008; Cannas, 2010; Chessa, 2013), and ii) the allochthonous competitor *R. philippinarum* is not yet widespread in the Sardinian-Corsican region, as in other Mediterranean sites (so far, only one conspicuous population of *R. philippinarum* were found in the Gulf of Olbia, North-East Sardinia; see Mura et al., 2012).

We pursued the twofold goal of i) evaluating the level of among-population genetic variability of the species with an enlarged Mediterranean sampling plan, and ii) surveying whether genetic substructuring, with genetically differentiated wild populations, occurs within this marine ecoregions (sensu Spalding et al. 2007). Indeed, the knowledge on the population genetic structure and the level of connectivity among population represents a pivotal information to be achieved, because it may be a critical point both for the understanding of the biology of a given species in a given area, and for better managing their stocks, particularly those that undergo severe exploitation (Utter, 1991; Thorpe et al., 2000). In such a context, the lack of recognition of population units may lead to local overexploitation, and finally to dramatic reduction/extinction of these populations.

To achieve the goals, we used sequences of the mitochondrial gene for the subunit I of the Cytochrome c Oxidase (COI), since COI has shown to be very useful to depict the processes that contribute to the spatial distribution of genetic diversity in marine species, also in *R. decussatus* (Cordero et al. 2013). The presence in GenBank of a conspicuous number of COI sequences from other Atlanto-Mediterranean sites (Brittany, Portugal, Spain, Tunisia, Adriatic, Aegean Sea, and Marmara Sea) gave us the possibility to compare our estimates of genetic variability from the Sardinian-Corsican region and from other sites of the central-west Mediterranean we sampled (Tuscany, Sicily, France and Tunisia), in order to collocate the genetic pattern of *R. decussatus* from this area within a larger geographic-scale perspective.

2. MATERIALS AND METHODS

Sampling

We sampled 271 specimens of *Ruditapes decussatus* from 15 locations (Table 1), sited in the following Atlantic and Mediterranean marine ecoregions (sensu Spalding et al. 2007): South European Atlantic Shelf and, Western Mediterranean.

DNA extraction and PCR

DNA was isolated using the Qiagen DNeasy tissue kit. Mitochondrial regions were amplified using specific primers for COI (L: 5'-gtaattattcggatagagtt-3' and H: 5'-cgtcgggtcaaaaaa-3') designed by the authors.

Each 25 µl PCR mixture contained approximately 100 ng of total genomic DNA, 0.32 µM of each primer, 2.5 U of EuroTaq DNA Polymerase (Euroclone), 1× reaction buffer and 200 µM of each dNTP. $MgCl_2$ concentration was set at 3 mM, and 25 µg of BSA (5ng/ml) was added to the reaction mixture. PCR amplifications were performed according to the following steps: 1 cycle of 2 min at 94°C, 35 cycles of 1 min at 94°C, 1 min at 50°C and 1 min 30 s at 72°C. A post-treatment of 5 min at 72°C was applied. After electrophoresis on 2% agarose gels, the PCR products were purified using ExoSAP-IT (USB Corporation) and sequenced using an external sequencing core service (Macrogen Inc., The Netherlands). Dual peaks of similar height, which could be interpreted as evidence of mitochondrial pseudogenes in the nucleus (Numts) or heteroplasmy, were not observed in any of the electropherograms. The PCR products did not show any occurrence of aspecificity, excluding the possibility of multiple nuclear mtDNA-like sequences.

Statistical analysis

Sequences were aligned using the program Clustal W (Thompson et al. 1994) and deposited in GenBank (Accession Nos.: xxxxxxx;). Estimates of the number of polymorphic sites (S), the number of haplotypes (H), haplotype diversity (*h*), and nucleotide diversity (π) were obtained using the software package DnaSP 5.10 (Librado & Rozas, 2009).

After having added to our sequences all the COI sequences available in the literature (Cordero et al. 2013; Keskin & Atar 2013), we constructed a dataset including the sequences obtained in the present study along with 351 sequences from the Mediterranean ecoregions South European Atlantic Shelf, Adriatic Sea, Aegean Sea, Tunisian Plateau/Gulf of Sidra, Western Mediterranean, Black Sea. (Spalding et al. 2007) (see Table 1 for details).

The presence of population genetic structure was assayed by the Bayesian model-based clustering algorithm implemented in Baps 5.2 (Corander & Tang 2007).

Clustering was performed using the module for linked molecular data and by applying the codon linkage model, which is appropriate for sequence data. Each analysis was performed ten times with a vector of values (1-22) for *K*, each with six replicates. Haplotypes were organised into haplogroups following the results of Bayesian clustering. For each sampling location, we computed the proportion of a given haplogroup to build interpolation maps of haplogroup frequency in the Mediterranean.

Genetic relationships among haplotypes were investigated by a Median-Joining network using the software package Network 4.5.0.1 (www.fluxus-engineering.com).

Patterns of genetic differentiation at the population level were assessed by Arlequin 3.5.1.3 (Excoffier et al. 1992) using a matrix of Tamura and Nei (1993) genetic

distances with a gamma correction according to the best-fitting model of sequence evolution obtained with JModeltest (Posada, 2008) which was the following 6-rates custom model: TPM2uf+G (gamma shape = 1.0630). First, we calculated pairwise Φ_{ST} values among sampling localities. Samples with less than five individuals were excluded from population level analysis due to a lack of statistical power. We next conducted a hierarchical analysis of molecular variance (AMOVA) (Schneider & Excoffier 1999). Alternative groupings of populations were taken into account. Populations were grouped as regional groups that corresponded to the main marine ecoregions (sensu Spalding et al. 2007): South European Atlantic Shelf, Adriatic Sea, Aegean Sea, Tunisian Plateau/Gulf of Sidra, Western Mediterranean, Black Sea.

3. RESULTS

Specific primers allowed us to amplify an internal portion of COI 348 bp long. Therefore, from coupling of our samples and data from Cordero et al. (2013) and Keskin and Atar (2013), we obtained a dataset from 622 *Ruditapes decussatus* individuals. Fifty two polymorphic site (S) resulted in a total of 48 different haplotypes one of which was the most diffused being reported in the 72% of individuals. The 56% of haplotypes found were private owned by a single individual. Total mean haplotype and nucleotide diversity were h = 0.474 and π = 0.005, respectively. An overall low levels of h and π were found among Atlantic populations from the cost of northern Portugal and France. Estimates of COI genetic diversity for each site are reported in Table 2.

Genetic structuring. The Bayesian analysis identified four distinct haplotype groups (R1, R2, R3 and R4). The first most diffused group (R2) include samples from all the marine ecoregions considered, with the only exception of individuals from the Marmara Sea (Black Sea), with an overall lack of geographic structuring. A similar trend was reported for the second most common group (R3) which spread widely across the Atlantic and Mediterranean ecoregions (occurring in all populations except for the Aegean and northern Portugal ones). On the other hand, Aegean samples all predominantly exhibited R4 group, which exclusively spanned this Mediterranean ecoregions with the sole exception of five individuals from the Sardinian population of Tortolì (Western Mediterranean). The latter group (R1) is characterised by the lowest average frequency, and characterises the population from the Marmara Sea (Black Sea) only.

The Median-Joining network analysis identified a low level of divergence among haplotypes, each differing in one or two point mutations, with the only exceptions of the Aegean and Marmara Sea populations (Figure 1). According to a model of star-like distribution, a single haplotype, present in 449 of 622 individuals, were the most common among all samples without evidence of geographic structuring among regions, with the exception of the Aegean samples and individuals from Marmara Sea that do not share haplotypes with the remaining sites.

The pairwise Φ_{ST} estimates indicate an overall significant genetic differentiation between the Aegean and Marmara Sea samples and the counterparts other samples (Table 3). Tunisian population from Sfax and northen Portugal populations (Tunisian Plateau/Gulf of Sidra and South European Atlantic Shelf respectively) were also

significantly differentiated in most comparisons. The AMOVA (Table 4) maximized the largest differences among groups defined *a priori* when the samples from Aegean Sea and Marmara Sea (sensu Spalding et al. 2007) were considered as two separated groups and the other samples were considered a unique separate group (Φ_{CT} = 0.843, P < 0.001). However, a similar value of molecular variance (Φ_{CT} = 0.766, P < 0.005) was obtained when treating four groups: 1) South European Atlantic Shelf and Western Mediterranean, 2) Adriatic Sea and Tunisian Plateau/Gulf of Sidra; 3) Aegean Sea; 4) Black Sea (sensu Spalding et al. 2007). When alternative groupings of samples were tested, the AMOVA showed a decrease in the proportion of variance (data not shown).

4. DISCUSSION

The present study was conducted by means of COI sequencing, in order to determine the level of genetic variation and unravel the phylogeographic structure of populations of the bivalve *Ruditapes decussatus*, also considering a so far uninvestigated biogeographic area, the central-west Mediterranean, in which different coastal sites are interested by intense exploitation of this species.

Supplementing the sequences from central-west Mediterranean with those obtained for Portugal and Tunisia, and those from previous investigations conducted in other sites of the range of distribution of the species allow us to provide a deep insight into the large-scale patterns of spatial genetic variation of *R. decussatus*.

Estimates of genetic diversity recorded in our central-west Mediterranean samples are lower than those reported for other bivalve species (e.g. Katsares et al., 2008; Sanna et al., 2013), but comparable to those obtained in previous studies (e.g., Gharbi et al.,

2010; Cordero et al., 2013) in populations from Atlantic, Spain, Tunisia, Adriatic, and Aegean Sea.

The genetic variability we detected is homogenously distributed throughout the area analysed.

The low levels of genetic differentiation found in this study are not surprising, and it may be explained in the light of the long PLD of *R. decussatus,* which can ensure a high potential for dispersal, at least at regional geographical scale, as evidenced for other bivalve species. However, the network star-like configuration, with many derived, unique haplotypes separated by one point-mutation from one geographically widespread haplotype, suggest that *R. decussatus* experienced in this area a severe decline of effective population size, likely followed by a recent period of rapid population growth (sudden population expansion). Two different, partly overlapping, explanations may account for this finding: i) the contractions of its populations occurred in the past as a consequence of the repeated sea-level fluctuations during Pleistocene glaciation, as evidenced in other marine species (see, e.g., Sanna et al. 2013); ii) a more recent reduction of population sizes may has been induced, at least locally, by overexploitation. Indeed, in different sites of the area investigated the species is not cultured and the harvesting takes place on natural beds (see, e,g, Turolla, 2008). Notably, Chessa et al. (2013) evidenced that in Sardinia, the natural resource represented by *R. decussatus* is at present insufficient for the local market, and that an overexploitation due to the illegal harvesting is an actual threat.

Interestingly, in the sample from Tortolì (Sardinia), an haplotype unexpectedly well differentiated from the commonest one (8 point mutation), and typical of the Aegean

Sea was found in five individuals. Although the occurrence of a unique haplotype with very low frequencies in the area investigated (but present in the Aegean Sea) prevented any statistically supported inference, we may suggest that a human-mediated introduction, e.g. by planktonic larvae passively transported in the ballast water of commercial ships (the commercial harbor of Arbatax is very close to Tortolì), which is a common finding in marine molluscs invasion, or an 'accidental' release of allochthonous individuals in the lagoon by the fishery operators may account for this outcome.

The low genetic divergence among the haplotypes spread throughout the Atlantic and the Mediterranean is suggestive of a common origin of the *R. decussatus* populations. Interestingly, in a previous study, Cordero et al. (2013) evidenced lack of genetic differentiation at COI across the Strait of Gibraltar, that was explained as suggestive of the absence of physical barrier to larval migration. Our results support this finding, being reinforced by the inclusion of additional samples from the central-west Mediterranean, so far not surveyed. Indeed, Bayesian assignment analysis and AMOVA did not evidenced the occurrence of genetic divergence between the Atlantic and the western Mediterranean, indicating that the different biogeographic boundaries which occur in this area (see Bianchi, 2007), did not limit the gene flow in *R. decussatus* from the Atlantic to the Siculo Tunisian Strait. Lack of genetic differentiation at COI between the Atlantic and western Mediterranean has been found in several marine invertebrates and fish (see, e.g., Patarnello 2007 and references therein).

Interestingly, our study showed the occurrence of a genetically divergent group of *R. decussatus* from the Black Sea, which was separated by three point mutations from the

commonest haplotype. The very enclosed nature of the Black Sea, which represents a well-defined phylogeographic region, may account for the genetic divergence retrieved for *R. decussatus*.

5. REFERENCES

Gharbi A., Chatti N., Said K., Van Wormhoudt, A. (2010). Genetic variation and population structure of the carpet shell clam *Ruditapes decussatus* along the Tunisian coast inferred from mtDNA and ITS1 sequence analysis. *Biologia*, *65*(4), 688-696.

Cordero D., Peña J. B., Saavedra, C. (2014). Phylogeographic analysis of introns and mitochondrial DNA in the clam *Ruditapes decussatus* uncovers the effects of Pleistocene glaciations and endogenous barriers to gene flow. *Molecular phylogenetics and evolution*, *71*, 274-287.

Turolla E. (2008). La venericoltura in Italia. FAO Actas de Pesca y Acuicultura. No. 12. Roma, FAO. pp. 177–188.

Cannas Alessandra (2009-2010). Dinamica di popolazione di *Ruditapes decussatus* (L) e insediamento di *Ruditapes philippinarum* (Adams&Reeve) in Sardegna (Italia). Dottorato di Ricerca. Biologia e biochimica dell'uomo e dell'ambiente - Indirizzo biologia animale Ciclo XXIII

G. Chessa, S. Serra, S. Saba, S. Manca, F. Chessa, M. Trentadue, N. Fois (2013). The floating upwelling system (FLUPSY) for breeding of *Venerupis decussata* (Linnaeus, 1758) juveniles in a coastal lagoon in Sardinia (Italy) *Transitional Waters Bulletin* TWB, 7 n. 2, 53-61.

Spalding M.D., Fox H.E., Allen G.R., Davidson N., Ferdaña Z.A., Finlayson M.A.X., Robertson, J. (2007). Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *BioScience*, 57(7), 573-583.

Thompson J.D., Higgins D.G. and Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research,* 22, 4673-4680.

Librado P. and Rozas J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451-1452.

Keskin E. and Atar H.H. (2013). DNA barcoding commercially important aquatic invertebrates of Turkey. *Mitochondrial DNA*, 24(4), 440-450.

Corander J. and Tang J. (2007). Bayesian analysis of population structure based on linked molecular information. *Mathematical biosciences*, 205: 19-31.

Excoffier L., Smouse P.E., Quattro J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction Data. *Genetics,* 131: 479-491.

Tamura K., Nei M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution,* 10, 512-526.

Posada D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution,* 25, 1253-1256.

Schneider S., Excoffier L. (1999). Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics,* 152: 1079-1089.

Katsares V., Tsiora, A., Galinou-Mitsoudi S., Imsiridou A. (2008). Genetic structure of the endangered species *Pinna nobilis* (Mollusca: Bivalvia) inferred from mtDNA sequences. *Biologia*, 63(3), 412-417.

Sanna D., Cossu P., Dedola G. L., Scarpa F., Maltagliati F., Castelli A., Casu, M. (2013). Mitochondrial DNA reveals genetic structuring of *Pinna nobilis* across the Mediterranean Sea*. PloS one*, 8(6), e67372.

Bianchi C.N. (2007). Biodiversity issues for the forthcoming tropical Mediterranean Sea. *Hydrobiologia,* 580: 7–21.

Patarnello T., Volckaert F. A., Castilho R. (2007). Pillars of Hercules: is the Atlantic–Mediterranean transition a phylogeographical break?. *Molecular Ecology*, 16(21), 4426-4444.

FIGURES AND TABLES

Figure 1

Median-joining network with the haplotypes grouped according to the sampling locations of the individuals in the major marine ecoregions proposed by Spalding et al. (2007): South European Atlantic Shelf (Portugal, France) and, Western Mediterranean (Corsica, Sardinia, Sicily, Tuscany, Tunisia); Adriatic Sea (Veneto); Tunisian Plateau/Gulf of Sidra (Tunisia); Aegean Sea (Greece); Black Sea (Turkey). The small red plots on the nodes show median vectors representing the hypothetical connecting sequences that were calculated using the maximum parsimony method. The numbers of mutations between haplotypes are reported on the network branches.

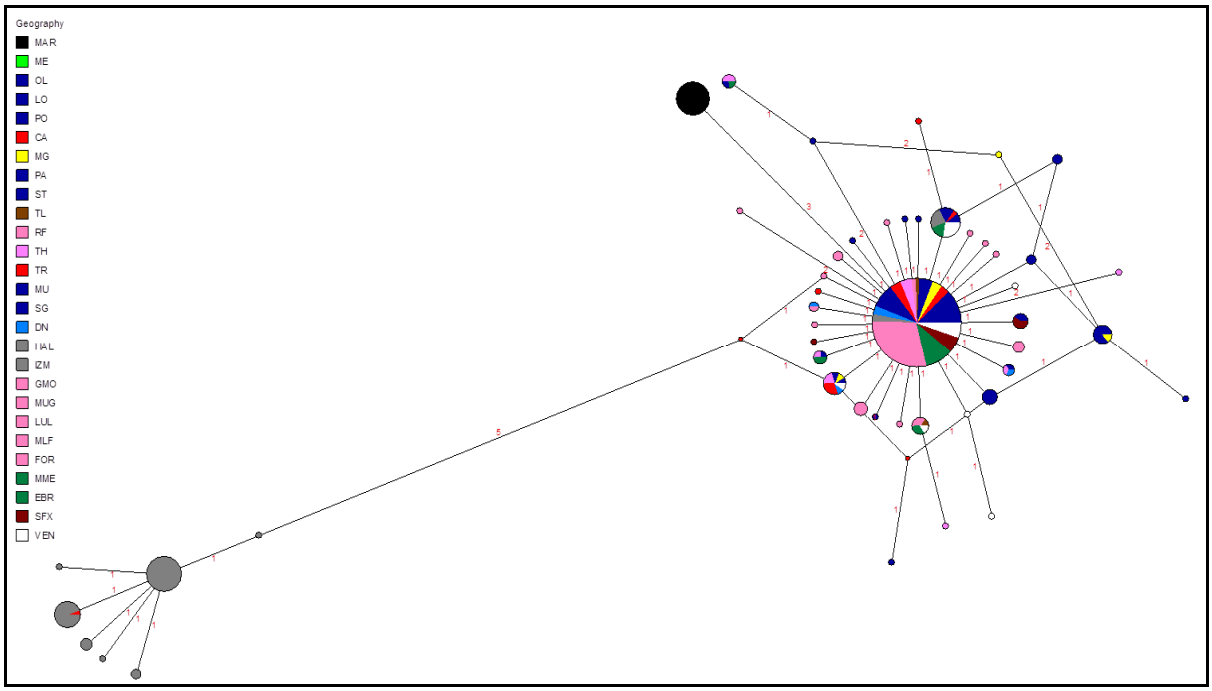

Table 1

Data collection: sampling localities, and sample sizes (N) for individuals of *Ruditapes decussatus* belonging to the marine ecoregion (Spalding et al. 2007) considered.

Marine Ecoregion	Region	Locality	Code	N	Study
Western Mediterranean	Sardinia	Olbia	OL	44	Present
		Lu Canaloni	PA	16	
		PoltuQualtu	PO	19	
		Tortolì	TR	23	
		San Teodoro	ST	11	
		Alghero	CA	17	
		Muravera	MU	20	
		Santa Giusta	SG	20	
	Tuscany	Orbetello	LO	19	
		Elba	ME	3	
	Corsica	Aleria	DN	16	
	Sicily	Messina	MG	19	
	Provence	Thau	TH	26	
	Northern Tunisia	Tunisis North-Lake	TL	6	
	Spain	Ebro delta	EBR	22	Cordero et. al 2013
		Mar Menor	MME	33	
Adriatic Sea	Veneto	Venice	VEN	33	
Tunisian Plateau/Gulf of Sidra	Tunisia	Sfax	SFX	31	
Aegean Sea	Greece	Halkidiki	HAL	33	
	Turkey	Izmir	IZM	30	
Black Sea		Marmara Sea	MAR	23	Keskin&Atar 2013
South European Atlantic Shelf	Portugal	Rio Formosa	RF	12	Present
			FOR	32	Cordero et. al 2013
		Milfontes	MLF	17	
	Spain	Mugardos	MUG	32	
		Lombos do Ulla	LUL	33	
	France	Golfe du Morbihan	GMO	32	

Table 2

Sample sizes and levels of genetic diversity obtained for each Mediterranean marine ecoregions (Spalding *et al.* 2007) considered. N: sample sizes; S: number of polymorphic sites; H: number of haplotypes; *h*: haplotype diversity; *π:* nucleotide diversity. See Table 1 for the sample codes.

Sample	**N**	**S**	**H**	***h***	***π***
OL	44	10	11	0.614	0.00284
PA	16	5	4	0.350	0.00180
PO	19	3	4	0.380	0.00189
ME	3	0	1	0	0
LO	19	6	7	0.608	0.00235
CA	17	2	3	0.228	0.00068
MG	19	5	4	0,298	0.00205
TR	23	11	4	0.383	0.00336
ST	11	2	3	0.345	0.00104
TL	6	1	2	0.333	0.00096
MU	20	0	1	0	0
SG	20	3	2	0.100	0.00086
DN	16	3	4	0.350	0,00108
TH	26	10	7	0.523	0.00302
RF	12	6	6	0.682	0.00287
MAR	23	0	1	0	0
HAL	33	10	5	0.748	0.01306
IZM	30	13	7	0.591	0.00576
GMO	32	5	6	0.343	0.00107
MUG	32	2	3	0.280	0.00143
LUL	33	0	1	0	0
MLF	17	0	1	0	0
FOR	32	3	4	0.236	0.00071
MME	33	5	4	0.280	0.00120
EBR	22	2	3	0.255	0.00076
SFX	31	2	3	0.239	0.00070
VEN	33	6	7	0.502	0.00180
Tot	622	52	48	0.474	0.00532

Chapter 5

Genetic variability in the Sardinian population of the manila clam,

Ruditapes philippinarum

Laura Mura, Piero Cossu, Alessandra Cannas, Fabio Scarpa, Daria Sanna, Gian Luca Dedola, Rossanna Floris, Tiziana Lai, Benedetto Cristo, Marco Curini-Galletti, Nicola Fois, Marco Casu

Published in *Biochemical Systematics and Ecology*, 41: 74–82 (2012)

Biochemical Systematics and Ecology 41 (2012) 74–82

Contents lists available at SciVerse ScienceDirect

Biochemical Systematics and Ecology

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

Genetic variability in the Sardinian population of the manila clam, *Ruditapes philippinarum*

L. Mura [a], P. Cossu [b], A. Cannas [a], F. Scarpa [b], D. Sanna [b], G.L. Dedola [b], R. Floris [a], T. Lai [b], B. Cristo [b], M. Curini-Galletti [b], N. Fois [a], M. Casu [b,*]

[a] Agris Sardegna, Agenzia per la Ricerca in Agricoltura, Dipartimento per la Ricerca nelle Produzioni Animali, Località Bonassai, 07040 Olmedo (SS), Italy
[b] Dipartimento di Zoologia e Genetica Evoluzionistica, Università di Sassari, Via F. Muroni 25, 07100 Sassari (SS), Italy

ARTICLE INFO

Article history:
Received 26 June 2011
Accepted 19 November 2011
Available online 11 January 2012

Keywords:
Biological invasion
Introduced species
Genetic variability
Manila clam

ABSTRACT

The genetic variability of microsatellites was investigated among allochthonous populations of *Ruditapes philippinarum* (Bivalvia, Veneridae) at six sites in the Gulf of Olbia (North-East Sardinia) and one site in the Lagoon of Sacca di Goro (North Adriatic Sea). Our results demonstrate the lack of genetic substructuring of *R. philippinarum* populations within the Gulf of Olbia. Furthermore, we found highly similar estimates of within-population genetic variability without evidence of genetic differentiation between Sardinian and Adriatic populations. The lack of a signature of a founder effect in the Sardinian samples, notwithstanding their recent introduction, indicates that the number of *R. philippinarum* individuals released into the Gulf of Olbia may have been high enough to reduce the loss of genetic variability due to genetic drift. Furthermore, the absence of genetic differentiation observed between samples from the two areas suggests that the two populations belong to the same gene pool. Our results also indicated that *R. philippinarum* is rapidly colonizing the Gulf of Olbia and the neighboring areas, which is likely favored by its high potential for dispersal.

1. Introduction

The introduction of allochthonous species in marine habitats is a traditional practice in aquaculture, particularly when fisheries are depleted due to the overexploitation of native species (see e.g. Breber, 1985; Pellizzato et al., 1989). However, this method can often pose serious risks (e.g., local extinction, hybridization) to indigenous species, particularly when farming is performed without separation of the farmed area from natural biotopes (see Wonham and Carlton, 2005). The introduction of the Manila clam *Ruditapes philippinarum* (Adams and Reeve, 1850) (Bivalvia: Veneridae) to Europe in the 1970s and 1980s (Gosling, 2003) from its native habitat on the West coast of the Pacific from Russia to the Philippines serves as a paradigm for such a phenomenon. Since then, the Manila clam has rapidly spread along European coastlines, becoming the main contributor to local clam fisheries. The Manila clam was released in the North-East Adriatic in the 1980s (Breber, 1985) to counter the dramatic depletion of several native species due to overexploitation and/or pollution (Savini et al., 2010). In the Adriatic, the Manila clam progressively occupied the ecological niche of the indigenous European grooved carpet shell, *Ruditapes decussatus* L., 1758, which led to the extinction of local populations (Jensen et al., 2004).

* Corresponding author. Tel.: +39 (0)79 228924; fax: +39 (0)79 228665.
E-mail address: marcasu@uniss.it (M. Casu).

0305-1978/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bse.2011.11.002

In 2008, several specimens of *R. philippinarum* were found in the Gulf of Olbia (North-East Sardinia). This species has now become abundant (Cannas et al., 2009) and has established the first recorded population of Manila clam in Sardinia. The introduction of this species was likely the result of the 'accidental' release of adult individuals a few years earlier due to careless management by fishery operators that were conducting legal stabulation processes in the town of Olbia (Cannas, pers. comm.; Cristo, pers. comm.).

The amount of initial genetic variability, as well as the potential for dispersal and a high growth rate, may play an important role in determining the spread of an invasive species by favoring its adaptation to new environments (Roman and Darling, 2007; Estoup and Guillemaud, 2010). Furthermore, a high genetic variability may overcome the deleterious effects of inbreeding on population growth (Bakker et al., 2009; Eales et al., 2010). Notably, the evolutionary potential of an invasive species relies more on the initial amount of genetic variability than the number of introduction events (Eales et al., 2010).

Despite the importance of following the evolution of genetic variability of allochthonous species, few studies on the genetic structure of the Manila clam have been performed to date (Kijima et al., 1987; Oniwa et al., 1988; Yokogawa, 1998; Sekine et al., 2006; Vargas et al., 2008; Chiesa et al., 2011). To our knowledge, only two of these studies focused on Mediterranean clam populations: Hurtado et al. (2011) uncovered evidence for hybridization between *R. philippinarum* and *R. decussatus* along North-West Spain, and Chiesa et al. (2011) suggested that patterns of genetic variation in clam populations from the North Adriatic Sea may have been shaped by multiple introduction events.

Therefore, our study was designed to analyze the genetic variation of microsatellites among *R. philippinarum* specimens collected at different sites in the Gulf of Olbia, with the twofold aim of detecting whether (i) genetic substructuring is present within this area and whether (ii) a recent founder effect occurred in the Gulf of Olbia. To this end, we compared estimates of within-population genetic variability of Sardinian samples with those of a formerly established, farmed population from the North Adriatic.

2. Materials and methods

A total of 195 specimens were collected from tidal flats at six different sites within the Gulf of Olbia at an average depth of 5–6 m. The sediment structure at the six sites was mostly characterized by coarse sand (0.5–1.0 mm grain size). Furthermore, 99 specimens were sampled within a farming area in the Lagoon of Sacca di Goro (SGR). For each individual, the total length, width, and height were measured (Table 1, Fig. 1).

Genomic DNA was isolated either from the muscle or from the mantle tissue using a modified alkaline/CTAB/NaCl lysis method and subsequent silicon dioxide treatment (Neudecker and Grimm, 2000). Approximately 50 mg of frozen tissue was incubated for 30 min at 65 °C in 200 µl of extraction buffer (NaOH 0.02 M, NaCl 2.5 M, CTAB 1%). The extract was then neutralized with the addition of 0.02 M Tris–HCl buffer (0.01 M Tris pH 8.0, 0.02 M HCl).

Six microsatellite loci (*Asari16*, *Asari24*, *Asari43*, *Asari55*, *Asari62*, and *Asari64*) (Yasuda et al., 2007) were used in this study. PCR was performed in a 25 µl solution containing 20–40 ng of genomic DNA, 1× reaction buffer, 2 mM $MgCl_2$, 0.2 mM of each dNTP, 1 U of EuroTaq (Euroclone), multiplexed fluorescent dye-labeled primers at different concentrations, and deionized water. PCR was performed in a MJ DNA Engine PTC-100 thermal cycler under the following conditions: an initial cycle at 95 °C for 5 min, followed by 30 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s. Microsatellite loci were visualized on an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems) and analyzed using GENEMAPPER 4.0 software (Applied Biosystems).

Genetic relationships among the sampling sites were visualized by means of a principal coordinate analysis (PCoA). The ordination plot was built using GENALAEX 6.3 software (Peakall and Smouse, 2006) on a matrix of genotypic distances.

Genetic structure was also investigated with an individual-based approach (Mank and Avise, 2004). Genetic substructuring was inferred by the Bayesian model-based clustering algorithm that is included in the STRUCTURE 2.2.3 software

Table 1
Details of the sampling plan. The codes of *Ruditapes philippinarum*, sampling sites, coordinates, number of individuals, and shell size are reported.

Code	Locality	Coordinates	# of individuals	Size range (mm)		
				TL	TW	TH
GOL1	Gulf of Olbia, NE Sardinia	40°55′55.67″N; 09°31′54.08″E	50	13.6–40.0	8.9–27.7	4.9–16.8
GOL2	Gulf of Olbia, NE Sardinia	40°55′16.54″N; 09°30′47.63″E	50	15.1–33.0	10.7–23.0	6.0–16.0
GOL3	Gulf of Olbia, NE Sardinia	40°55′00.59″N; 09°30′46.48″E	25	13.8–42.2	9.4–28.9	5.3–19.7
GOL4	Gulf of Olbia, NE Sardinia	40°55′09.26″N; 09°32′31.60″E	25	10.3–35.0	7.3–25.8	4.2–16.7
GOL5	Gulf of Olbia, NE Sardinia	40°55′09.19″N; 09°33′16.13″E	48	13.2–32.1	8.6–26.6	4.7–14.4
GOL6	Gulf of Olbia, NE Sardinia	40°55′00.23″N; 09°33′19.76″E	25	13.9–38.6	9.4–29.4	5.0–21.0
SGR	Lagoon of Sacca di Goro, N Adriatic	44°48′49.78″N; 12°16′50.63″E	99	31.6–41.5	24–30.1	17.5–22.4

TL: total length; TW: total width; and TH: total height.

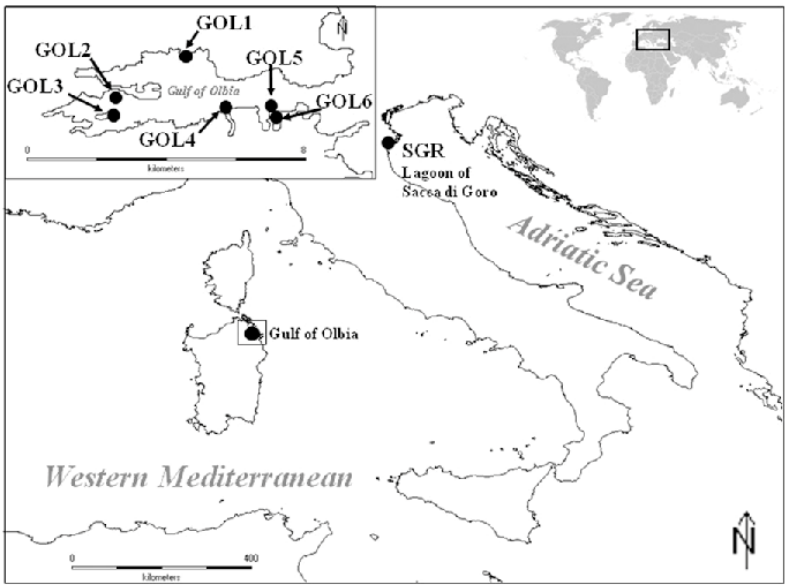

Fig. 1. Sampling sites of *Ruditapes philippinarum*. The samples from the Gulf of Olbia are reported in the inset. For the sample codes, see Table 1.

(Pritchard et al., 2000), applying the admixture model with correlated allelic frequencies (Falush et al., 2003). The most likely population structure was inferred from multilocus individual genotypes by estimating the posterior probability of the data (PPD) according to Pritchard and Wen (2004) and, if necessary, the consistency of results across runs. For each value of K (the number of assumed genetic clusters), ten independent runs were performed, each consisting of 100,000 iterations after a burn-in period of 100,000.

Furthermore, the partitioning of genetic variation within and among sampling sites was estimated by AMOVA (Analysis of Molecular Variance) using ARLEQUIN 3.52 software (Excoffier and Lischer, 2010). The significance of genetic differentiation was tested by a permutation test (10,000 replicates). Overall Φ_{ST} was calculated, followed by a pairwise computation of R_{ST} values (Slatkin, 1995) between all pairs of sampling sites within the Gulf of Olbia and the Adriatic sample. The significance of the random distribution of individuals among sampling sites was evaluated with exact tests that are analogous to Fisher's contingency tables (Raymond and Rousset, 1995). We defined two populations from subsequent genetic analyses based upon the results of genetic structuring: one population was obtained by pooling the samples from the Gulf of Olbia (SPOOL), and the other population represented the sample from the Lagoon of Sacca di Goro (SGR).

Conformity to Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were tested with the Markov chain randomization test (Guo and Thompson, 1992) using GENEPOP 4.0 (Rousset, 2008). Whenever necessary, the sequential Bonferroni correction was applied to P values to avoid type I errors that can occur due to multiple testing (Hochberg, 1988).

Estimates of genetic variability were computed using the GENALEX software. The number and size range of alleles, allelic frequencies, observed (H_o) and expected (H_e) heterozygosity, and inbreeding coefficient (F_{IS}) values were estimated at each microsatellite locus. Moreover, the mean number of alleles (Na), the number of effective alleles (Ne), and the observed (H_o) and expected (H_e) heterozygosity were also calculated for the loci of each inferred population (SPOOL and SGR).

The statistical significance for differences in the expected heterozygosity was tested with a Wilcoxon rank-sum test, using the R 2.11.1 package (R Development Core Team, 2010).

The possibility of a founder effect due to recent colonization or expansion in SPOOL and SGR was tested using the software BOTTLENECK 1.2.02 (Cornuet and Luikart, 1996; Piry et al., 1999). Indeed, after genetic bottlenecks, the number of alleles reduced faster than heterozygosity, leading to excess of gene diversity when compared to a mutation-drift equilibrium. Conversely, an expanding population accumulates new alleles at a high rate, which contribute little to heterozygosity because of their low frequency, resulting in a gene diversity that was lower than expected at mutation-drift equilibrium (Luikart et al., 1998). Because the expected heterozygosity is highly dependent on the selected model of microsatellite evolution (Hawley et al., 2006), the analysis was performed using both the strict stepwise mutation model (SMM) and the two-phase mutation model (TPM) (see Cornuet and Luikart, 1996). Here, the reported results are based on simulations with a TPM model set at 95% and the infinite allele model at 5%. The significance of the tests was assessed

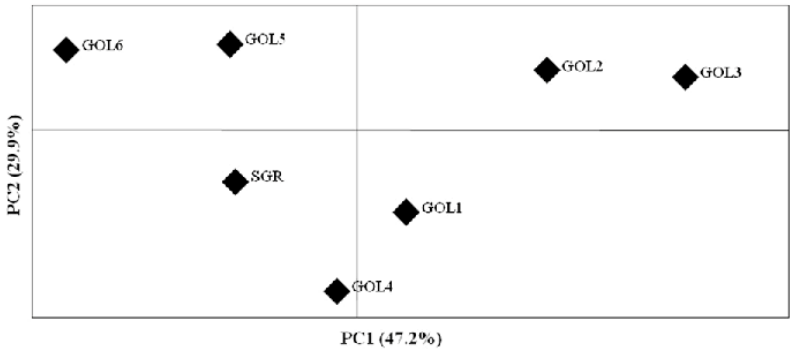

Fig. 2. Principal coordinate analysis (PCoA). The plot depicts the genetic relationships among all sampling sites. For the sample codes, see Table 1.

Table 2
Estimation of the best-fit model of population structure. The number of assumed genetic clusters (K), the likelihood of the posterior probability of the model given the data (LnP(D)), and the posterior probability of the data (PPD) are reported.

K	LnP(D)	PPD
1	−5575.84	1
2	−5757.19	~0
3	−6253.89	~0
4	−7391.75	~0
5	−6212.24	~0

using Wilcoxon rank-sum test, which is the most appropriate test when fewer than 20 microsatellite loci are used (Piry et al., 1999).

3. Results

Screening 294 specimens of *R. philippinarum* yielded a total of 87 alleles (see the Appendix), resulting in 286 unique individual multilocus genotypes (representing approximately 97% of the specimens).

A PCoA plot of the first two principal coordinates (77% of total genetic variation) did not show a clear separation between the Adriatic and Sardinian samples (Fig. 2). However, the ordination of samples from the Gulf of Olbia along the first principal coordinate appeared to confirm their geographic position (Fig. 1).

Nevertheless, the highest posterior probability of the data was obtained using the model with $K = 1$, which was retrieved by STRUCTURE (Table 2), supporting the lack of genetic structuring both within the Gulf of Olbia and between samples from the Gulf of Olbia and the Lagoon of Sacca di Goro. Similar results were obtained using AMOVA ($\Phi_{ST} = 0.004$, $P > 0.05$). Exact tests of population differentiation did not demonstrate divergence (i) among sampling sites within the Gulf of Olbia, and (ii) between these samples and the Adriatic sample (pairwise R_{ST} values are given in Table 3).

Table 3
Pairwise R_{ST} values. The R_{ST} values are shown below the diagonal. The probability values based on an exact test of population differentiation based on a Markov chain with 100,000 steps and 10,000 dememorization steps are shown above the diagonal. For the sample codes, see Table 1.

R_{ST}	GOL1	GOL2	GOL3	GOL4	GOL5	GOL6	SGR
GOL1	–	0.103	0.116	0.110	0.058	0.231	0.107
GOL2	−0.0015	–	1.000	1.000	1.000	1.000	1.000
GOL3	−0.0175	0.0280	–	1.000	1.000	1.000	1.000
GOL4	−0.0035	0.0325	0.0054	–	1.000	1.000	1.000
GOL5	−0.0036	0.0256	−0.0144	−0.0068	–	1.000	1.000
GOL6	0.0047	0.0024	0.0296	0.0597	0.0354	–	1.000
SGR	−0.0049	0.0221	−0.0060	−0.0065	−0.0272	0.0232	–

$^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

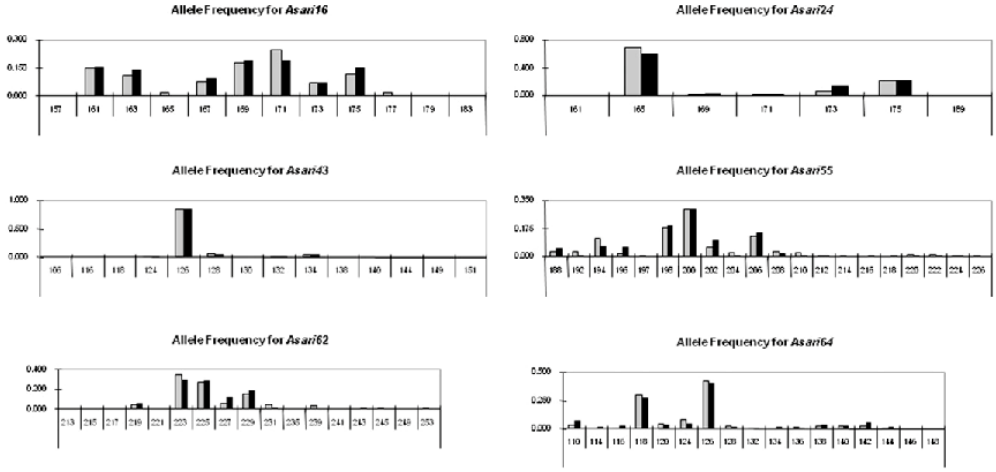

Fig. 3. Histograms depicting the allelic frequencies of each microsatellite locus in the Gulf of Olbia and the Lagoon of Sacca di Goro. Alleles for each locus are reported on the *x*-axis and their frequencies on the *y*-axis. Gray bar: SPOOL (Sardinian pooled sample); black bar: SGR (Adriatic sample).

Table 4
Summary of the microsatellite-based estimates of intrapopulation genetic variability. The locus name, number of alleles, size range, observed heterozygosity (H_O), expected heterozygosity (H_E) and their respective standard errors (SE), and inbreeding coefficient (F_{IS}) are reported.

Locus	# of alleles	Size range (bp)	H_O (±SE)	H_E (±SE)	F_{IS}
Asari16	12	157–183	0.821 ± 0.022	0.836 ± 0.006	0.019
Asari24	7	161–189	0.448 ± 0.025	0.503 ± 0.020	0.109
Asari43	14	106–151	0.319 ± 0.039	0.284 ± 0.031	−0.121
Asari55	20	188–226	0.617 ± 0.049	0.816 ± 0.015	0.245
Asari62	17	213–253	0.746 ± 0.034	0.759 ± 0.013	0.018
Asari64	17	110–148	0.737 ± 0.027	0.708 ± 0.014	−0.041

After pooling the Sardinian samples (SPOOL), we found that the allele frequencies were very similar to the Adriatic population (SGR) (Appendix, Fig. 3). Although population-specific alleles (private alleles) were present in both populations (18 and 10 in SPOOL and SGR, respectively), they occurred at very low frequencies, and 20 out of 28 alleles were found in a single individual. Alleles 165 and 177 at *Asari16*, the private alleles with the highest frequency, were found in seven and six specimens in SPOOL, respectively (Appendix, Fig. 3).

Significant deviations from HWE ($P < 0.05$ after sequential Bonferroni correction for multiple tests) were found at locus *Asari55* in SPOOL and at loci *Asari24* and *Asari55* in SGR. In both cases, the deviations from HWE are caused by a deficit of heteozygotes, as evidenced by the positive F_{IS} values at these loci (Table 4).

Only two out of the 30 pairs of loci showed a significant LD after the sequential Bonferroni correction ($P < 0.05$). Because pairwise comparisons at one or two loci are expected to show genotypic disequilibrium over 30 pairs by chance (specifically 1.5 out of 30, corresponding to the cut-off of 5%), the loci were considered to be unlinked. Moreover, two pairs of loci with a significant LD (*Asari16-Asari55* and *Asari16-Asari62*) were found only in SPOOL.

The mean number and the effective number of alleles were comparable for SPOOL (Na = 12.833 ± 1.797; Ne = 4.074 ± 0.892) and SGR (Na = 11.500 ± 1.607; Ne = 4.151 ± 0.814). In SPOOL we found $H_o = 0.621 \pm 0.077$, and $H_e = 0.662 \pm 0.091$. Comparable values were obtained in SGR ($H_o = 0.626 \pm 0.090$; $H_e = 0.680 \pm 0.087$). Indeed, the expected heterozygosity of SPOOL was not significantly lower than SGR ($P = 1$; Wilcoxon rank-sum test). Table 5 reports the Wilcoxon rank-sum test results for mutation-drift equilibrium under the SMM and TPM models.

Table 5
Wilcoxon rank-sum test of the mutation-drift equilibrium. The stepwise mutation model (SMM) and a two-phase model (TPM) were estimated at six microsatellite loci in SGR (Adriatic sample) and SPOOL (Sardinian pooled sample).

	Heterozygosity deficiency		Heterozygosity excess	
	SMM	TPM	SMM	TPM
SGR	0.016	0.016	0.992	0.992
SPOOL	0.008	0.016	1.000	0.992

4. Discussion and conclusion

This study demonstrates a high degree of genetic variability of *R. philippinarum* at six microsatellite loci. At each locus, the size range of the alleles, the number of alleles, and the number of effective alleles were higher than those reported by Yasuda et al. (2007). The observed and expected heterozygosity displayed values that were comparable in magnitude to those reported by Yasuda et al. (2007) and Chiesa et al. (2011). In particular, although the values of expected heterozygosity were lower in our samples at *Asari24*, *Asari43* and *Asari55* (Table 4), the differences in our values and in those detected by Yasuda et al. (2007) were not significant (Wilcoxon rank-sum test, $P = 0.394$). Two out of six loci, *Asari24* (in SGR) and *Asari55* (in SPOOL and SGR), showed an overall deficit of heterozygosity, a phenomenon that is common in other bivalves (i.e. Hare et al., 1996; Raymond et al., 1997; Vargas et al., 2008).

Chiesa et al. (2011) found moderate levels of genetic differentiation in microsatellites among clam populations from the North Adriatic Sea, hypothesizing that such genetic differentiation may be the result of multiple introductions from different gene pools. In contrast, we detected a lack of genetic substructuring within the Gulf of Olbia (Tables 2 and 3). Our results are suggestive of (i) a single introduction event or (ii) multiple introductions either from the same source population or from genetically undifferentiated populations.

The results provided by model-based clustering for the whole dataset (Table 2) also identified a common origin for the Sardinian and Adriatic samples. Although private alleles were found in SPOOL, the two populations share the same gene pool, likely because not enough time has elapsed to achieve a significant amount of divergence (Hutchison and Templeton 1999).

Notably, while SPOOL showed numerous loci departing from HWE that are lower than SGR, only two pairs of loci displaying LD were found in the Sardinian population. Assuming that LD was not due to chance, the physical linkage of the loci can be ruled out because the same pairs did not display LD in SGR. Furthermore, non-random associations of loci may occur in populations departing from HWE (Liu et al., 2006) and determine the so-called zygotic disequilibrium (Weir, 1996). A substantial LD may also be found in expanding populations that were initially bottlenecked as well as in recently introduced populations (Hansson et al., 2004, and references therein). The latter hypothesis fits with the historical data that point to a recent introduction of the Sardinian population.

A further outcome of the present study was the great similarity in the level of within-population genetic diversity among the Sardinian samples from the Gulf of Olbia, regardless of whether they were pooled, and also in the sample from the Lagoon of Sacca di Goro, as evidenced by Wilcoxon rank-sum test, which was applied to the expected heterozygosity ($P = 1$). In fact, there was little evidence of a significant loss of genetic variability in SPOOL and SGR under the SSM and TPM mutation models (Table 5). Instead, they exhibited a significant deficit of heterozygozity with respect to genetic diversity.

This finding suggests that the 'wild' population of the Gulf of Olbia is currently expanding (see Ribeiro et al., 2010). This hypothesis is supported by the presence of several alleles that occur at low frequencies; for example, 25 out of 77 alleles (~32%) in SPOOL and 29 out of 69 alleles (~28%) in SGR were found at frequencies lower than 1% (Appendix). Thus, we can conclude that the two populations are either derived from the same source, or that the Sardinian population is derived from individuals that were farmed in the Adriatic and released *in situ*.

Regardless of the source of Sardinian samples, the number of individuals of *R. philippinarum* that were scattered in the Gulf of Olbia may have been high enough to avoid the consequence of a founder effect. Indeed, the higher fecundity rates that invasive or introduced species exhibit in new environments (i.a. Grosholz and Ruiz, 2003) often overcome the loss of genetic diversity that is caused by genetic bottlenecks or the founder effect (Andreakis et al., 2009). In this context, *R. philippinarum* may be considered a 'good invader' species due to its effective larval dispersal and high rate of reproduction. Cannas et al. (2009) reported the occurrence of *R. philippinarum* recruits one year after their first recorded release in the Gulf of Olbia.

Interestingly, Delgado and Pérez-Camacho (2007) have reported that under the same experimental conditions, *R. philippinarum* demonstrates superior reproductive performance when compared to *R. decussatus*. This trait, coupled with high levels of genetic diversity, may allow *R. philippinarum* to quickly adapt to the new environmental conditions (see Estoup and Guillemaud, 2010, and the references therein).

The previously recorded hybridization between *R. philippinarum* and *R. decussatus* (Hurtado et al., 2011) represents a further threat, because hybridization may enhance the invasiveness of an allochthonous species by producing novel genotypes that better adapt to local conditions (see Henry et al., 2009, and the references therein).

These factors raise concerns about the conservation and management of the autochthonous *R. decussatus*, whose populations may suffer from competition with *R. philippinarum* (Chessa et al., 2005) as well as the maladaptation that occurs as a result of hybridization between the two species (Hurtado et al., 2011). Finally, the recent introduction of *R. philippinarum* in Sardinia offers the opportunity to follow the evolution of genetic variability in a geographic area where a newly introduced invasive species has begun to spread.

Acknowledgments

This research was financially supported by the Agris Sardegna, Agenzia per la Ricerca in Agricoltura (Sardinia, Italy) and partially by the Comune di Olbia (Sardinia, Italy). We are grateful to the anonymous referee whose comments and suggestions helped improve this manuscript.

Appendix. Allele frequencies

Private alleles in SGR (Adriatic sample) and SPOOL (Sardinian pooled sample) are highlighted in bold. For the sample codes, see Table 1.

Locus	Allele	GOL1	GOL2	GOL3	GOL4	GOL5	GOL6	SGR	SPOOL
Asari16	*N*	22	50	25	25	48	25	96	195
	157	0.000	0.000	0.020	0.000	0.000	0.000	0.005	0.001
	161	0.091	0.170	0.180	0.180	0.135	0.160	0.156	0.154
	163	0.091	0.100	0.120	0.120	0.125	0.120	0.135	0.113
	165	0.068	0.000	0.020	0.080	0.000	0.000	0.000	**0.021**
	167	0.114	0.060	0.060	0.020	0.104	0.100	0.094	0.077
	169	0.227	0.250	0.080	0.140	0.156	0.160	0.193	0.177
	171	0.273	0.210	0.340	0.240	0.271	0.140	0.193	0.244
	173	0.000	0.070	0.040	0.060	0.073	0.180	0.073	0.072
	175	0.136	0.090	0.120	0.140	0.115	0.120	0.151	0.115
	177	0.000	0.050	0.000	0.000	0.011	0.020	0.000	**0.018**
	179	0.000	0.000	0.020	0.020	0.000	0.000	0.000	**0.005**
	183	0.000	0.000	0.000	0.000	0.010	0.000	0.000	**0.003**
Asari24	*N*	21	50	25	22	48	25	97	191
	161	0.000	0.000	0.000	0.000	0.000	0.000	**0.005**	0.000
	165	0.690	0.690	0.640	0.636	0.688	0.700	0.598	0.678
	169	0.000	0.020	0.040	0.046	0.031	0.020	0.031	0.026
	171	0.000	0.000	0.000	0.046	0.010	0.000	0.015	0.008
	173	0.048	0.020	0.080	0.136	0.094	0.060	0.134	0.068
	175	0.262	0.270	0.200	0.136	0.177	0.220	0.211	0.215
	189	0.000	0.000	0.040	0.000	0.000	0.000	0.006	0.005
Asari43	*N*	22	50	25	25	45	25	96	192
	106	0.023	0.000	0.000	0.000	0.000	0.000	0.000	**0.003**
	116	0.000	0.000	0.000	0.000	0.011	0.000	0.000	**0.003**
	118	0.000	0.000	0.000	0.000	0.000	0.000	**0.005**	0.000
	124	0.000	0.000	0.000	0.000	0.022	0.100	0.016	0.018
	126	0.818	0.880	0.800	0.920	0.767	0.840	0.839	0.836
	128	0.136	0.050	0.060	0.000	0.089	0.040	0.047	0.063
	130	0.000	0.000	0.020	0.000	0.000	0.000	0.000	**0.003**
	132	0.000	0.010	0.000	0.000	0.033	0.020	0.016	0.013
	134	0.023	0.050	0.120	0.080	0.067	0.000	0.057	0.057
	138	0.000	0.000	0.000	0.000	0.011	0.000	0.005	0.001
	140	0.000	0.010	0.000	0.000	0.000	0.000	0.000	**0.003**
	144	0.000	0.000	0.000	0.000	0.000	0.000	**0.005**	0.000
	149	0.000	0.000	0.000	0.000	0.000	0.000	**0.005**	0.000
	151	0.000	0.000	0.000	0.000	0.000	0.000	**0.005**	0.000
Asari55	*N*	20	50	24	24	46	25	97	189
	188	0.000	0.060	0.019	0.000	0.054	0.000	0.057	0.032
	192	0.100	0.020	0.021	0.000	0.065	0.000	0.010	0.034
	194	0.100	0.110	0.042	0.042	0.196	0.100	0.072	0.111
	196	0.000	0.010	0.063	0.000	0.033	0.000	0.062	0.019
	197	0.000	0.000	0.021	0.000	0.000	0.000	0.000	**0.003**
	198	0.200	0.130	0.292	0.125	0.141	0.260	0.191	0.177
	200	0.350	0.330	0.396	0.333	0.228	0.160	0.294	0.294
	202	0.075	0.090	0.000	0.104	0.054	0.020	0.098	0.061
	204	0.050	0.000	0.021	0.063	0.043	0.000	0.010	0.026
	206	0.075	0.080	0.104	0.188	0.109	0.240	0.144	0.124
	208	0.025	0.070	0.021	0.021	0.000	0.040	0.021	0.032
	210	0.025	0.010	0.000	0.021	0.011	0.120	0.010	0.026
	212	0.000	0.020	0.000	0.021	0.000	0.000	0.006	0.007
	214	0.000	0.010	0.000	0.021	0.022	0.000	0.000	**0.011**
	216	0.000	0.000	0.000	0.000	0.000	0.000	**0.005**	0.000
	218	0.000	0.000	0.000	0.000	0.000	0.000	**0.005**	0.000
	220	0.000	0.010	0.000	0.042	0.011	0.020	0.005	0.013
	222	0.000	0.020	0.000	0.019	0.011	0.040	0.005	0.016
	224	0.000	0.030	0.000	0.000	0.011	0.000	0.005	0.011
	226	0.000	0.000	0.000	0.000	0.011	0.000	0.000	**0.003**

Appendix (*continued*)

Locus	Allele	GOL1	GOL2	GOL3	GOL4	GOL5	GOL6	SGR	SPOOL
Asari62	*N*	21	50	25	25	48	25	96	194
	213	0.000	0.020	0.000	0.000	0.000	0.000	0.005	0.005
	215	0.000	0.010	0.000	0.020	0.000	0.000	0.005	0.005
	217	0.000	0.010	0.000	0.000	0.011	0.000	0.000	**0.005**
	219	0.000	0.060	0.060	0.020	0.042	0.080	0.063	0.046
	221	0.000	0.000	0.000	0.000	0.021	0.000	0.000	**0.005**
	223	0.333	0.380	0.500	0.300	0.292	0.300	0.297	0.348
	225	0.357	0.190	0.200	0.320	0.292	0.320	0.276	0.268
	227	0.048	0.060	0.040	0.040	0.073	0.040	0.120	0.054
	229	0.119	0.140	0.100	0.180	0.177	0.180	0.182	0.152
	231	0.024	0.050	0.040	0.080	0.031	0.040	0.016	0.044
	235	0.000	0.010	0.000	0.000	0.000	0.000	0.000	**0.003**
	239	0.095	0.010	0.020	0.020	0.031	0.040	0.005	0.031
	241	0.000	0.000	0.000	0.000	0.010	0.000	0.000	**0.003**
	243	0.000	0.000	0.000	0.000	0.010	0.000	0.016	0.003
	245	0.024	0.010	0.040	0.020	0.000	0.000	0.000	**0.013**
	249	0.000	0.000	0.000	0.000	0.000	0.000	**0.005**	0.000
	253	0.000	0.050	0.000	0.000	0.010	0.000	0.010	0.015
Asari64	*N*	22	50	25	24	48	25	98	194
	110	0.023	0.040	0.020	0.000	0.052	0.020	0.066	0.031
	114	0.000	0.000	0.000	0.000	0.000	0.000	**0.010**	0.000
	116	0.000	0.010	0.000	0.000	0.000	0.040	0.020	0.008
	118	0.205	0.280	0.360	0.208	0.354	0.360	0.276	0.302
	120	0.136	0.040	0.040	0.021	0.012	0.040	0.031	0.041
	124	0.023	0.140	0.120	0.063	0.052	0.040	0.041	0.080
	126	0.477	0.370	0.400	0.542	0.396	0.420	0.403	0.420
	128	0.000	0.020	0.020	0.042	0.031	0.000	0.010	0.021
	132	0.000	0.010	0.000	0.000	0.000	0.020	0.005	0.005
	134	0.000	0.000	0.000	0.000	0.000	0.000	**0.010**	0.000
	136	0.045	0.000	0.000	0.019	0.031	0.000	0.005	0.015
	138	0.045	0.050	0.020	0.021	0.000	0.000	0.036	0.023
	140	0.023	0.040	0.020	0.021	0.031	0.000	0.026	0.026
	142	0.023	0.000	0.000	0.042	0.021	0.040	0.046	0.018
	144	0.000	0.000	0.000	0.021	0.010	0.000	0.015	0.004
	146	0.000	0.000	0.000	0.000	0.010	0.000	0.000	**0.003**
	148	0.000	0.000	0.000	0.000	0.000	0.020	0.000	**0.003**

References

Andreakis, N., Kooistra, W.H.C.F., Procaccini, G., 2009. High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. Mol. Ecol. 18, 212–226.

Bakker, E.G., Montgomery, B., Nguyen, T., Eide, K., Chang, J., Mockler, T.C., Liston, A., Seabloom, E.W., Borer, E.T., 2009. Strong population structure characterizes weediness gene evolution in the invasive grass species *Brachypodium distachyon*. Mol. Ecol. 18, 2588–2601.

Breber, P., 1985. On-growing of the carpet-shell clam (*Tapes decussatus* (L.)): two years' experience in Venice Lagoon. Aquaculture 44, 51–56.

Cannas, A., Trentadue, M., Manca, S., Mura, L., Fois, N., 2009. Insediamento di *Ruditapes philippinarum* (Adams & Reeve) nel golfo di Olbia (Sardegna, Italia) e distribuzione in rapporto a *Ruditapes decussatuas* (L.). Biol. Mar. Mediterr. 17, 320–321.

Chessa, L.A., Paesanti, F., Pais, A., Scardi, M., Serra, S., Vitale, L., 2005. Perspectives for development of low impact aquaculture in a Western Mediterranean lagoon: the case of the carpet clam *Tapes decussatus*. Aquacult. Int. 13, 147–155.

Chiesa, S., Nonnis Marzano, F., Minervini, G., De Lucrezia, D., Baccarani, G., Bordignon, G., Poli, I., Ravagnan, G., Argese, E., 2011. The invasive Manila clam *Ruditapes philippinarum* (Adams and Reeve, 1850) in Northern Adriatic Sea: population genetics assessed by an integrated molecular approach. Fish. Res. 110, 259–267.

Cornuet, J.M., Luikart, G., 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144, 2001–2014.

Delgado, M., Pérez-Camacho, A., 2007. Comparative study of gonadal development of *Ruditapes philippinarum* (Adams and Reeve) and *Ruditapes decussatus* (L.) (Mollusca: Bivalvia): influence of temperature. Sci. Mar. 71, 471–484.

Eales, J., Thorpe, R.S., Malhotra, A., 2010. Colonization history and genetic diversity: adaptive potential in early stage invasions. Mol. Ecol. 19, 2858–2869.

Estoup, A., Guillemaud, T., 2010. Reconstructing routes of invasion using genetic data: why, how and so what? Mol. Ecol. 19, 4113–4130.

Excoffier, L., Lischer, H.E.L., 2010. 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.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164, 1567–1587.

Gosling, E., 2003. Bivalve Molluscs. Biology, Ecology and Culture. Fishing News Books, Oxford.

Grosholz, E.D., Ruiz, G.M., 2003. Biological invasions drive size increases in marine and estuarine invertebrates. Ecol. Lett. 6, 700–705.

Guo, S.W., Thompson, E.A., 1992. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48, 361–372.

Hansson, B., Westerdahl, H., Hasselquist, D., Åkesson, M., Bensch, S., 2004. Does linkage disequilibrium generate heterozygosity-fitness correlations in great reed warblers? Evolution 58, 870–879.

Hare, M.P., Karl, S.A., Avise, J.C., 1996. Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. Mol. Biol. Evol. 13, 334–345.

Hawley, D.M., Hanley, D., Dhondt, A.A., Lovette, I.J., 2006. Molecular evidence for a founder effect in invasive house finch (*Carpodacus mexicanus*) populations experiencing an emergent disease epidemic. Mol. Ecol. 15, 263–275.

Henry, P., Le Lay, G., Goudet, J., Guisan, A., Jahodová, Š., Besnard, G., 2009. Reduced genetic diversity, increased isolation and multiple introductions of invasive giant hogweed in the western Swiss Alps. Mol. Ecol. 18, 2819–2831.
Hochberg, Y., 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75, 800–802.
Hurtado, N.S., Pérez-García, C., Morán, P., Pasantes, J.J., 2011. Genetic and cytological evidence of hybridization between native *Ruditapes decussatus* and introduced *Ruditapes philippinarum* (Mollusca, Bivalvia, Veneridae) in NW Spain. Aquaculture 311, 123–128.
Hutchison, D.W., Templeton, A.R., 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. Evolution 53, 1898–1914.
Jensen, A.C., Humphreys, J., Caldow, R.W.G., Grisley, C., Dyrynda, P.E.J., 2004. Naturalization of the Manila clam (*Tapes philippinarum*), an alien species, and establishment of a clam fishery within Poole Harbour, Dorset. J. Mar. Biol. Assoc. UK 84, 1069–1073.
Kijima, A., Taniguchi, N., Mori, N., Hagiwara, J., 1987. Genetic variability and breeding structure in *Ruditapes philippinarum* (sic.). Rep. USA Mar. Biol. Inst. Kochi Univ. 9, 173–181.
Liu, T., Todhunter, R.J., Lu, Q., Schoettinger, L., Li, H., Littell, R.C., Burton-Wurster, N., Acland, G.M., Lust, G., Wu, R., 2006. Modelling extent and distribution of zygotic disequilibrium: implications for a multigenerational canine pedigree. Genetics 174, 439–453.
Luikart, G., Allendorf, F.W., Cornuet, J.M., Sherwin, W.B., 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. J. Hered. 89, 238–247.
Mank, J.E., Avise, J.C., 2004. Individuals organisms as units of analysis: Bayesian-clustering alternatives in population genetics. Genet. Res. 84, 135–143.
Neudecker, F., Grimm, S., 2000. High-throughput method for isolating plasmid DNA with reduced lipopolysaccharide content. Biotechniques 28, 107–109.
Oniwa, K., Nakano, M., Fujio, Y., 1988. Heterogeneity within and between geographical populations of the short-necked clam, *Ruditapes philippinarum*. Tohoku J. Agric. Res. 38, 49–60.
Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.
Pellizzato, M., Mattei, N., Renzoni, A., 1989. Allevamento su scala commerciale di *Tapes philippinarum* (Adams and Reeve, 1850) in alcune valli venete. Oebalia XV-2, 735–744.
Piry, S., Luikart, G., Cornuet, J.M., 1999. Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. J. Hered. 90, 502–503.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
Pritchard, J.K., Wen, W., 2004. Documentation for Structure Software: Version 2. University of Chicago Press, Chicago.
R Development Core Team, 2010. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. http://www.R-project.org.
Raymond, M., Rousset, F., 1995. An exact test for population differentiation. Evolution 49, 1280–1283.
Raymond, M., Vääntö, R.L., Thomas, F., Rousset, F., De Meeüs, T., Renaud, F., 1997. Heterozygote deficiency in the mussel *Mytilus edulis* species complex revisited. Mar. Ecol. Prog. Ser. 156, 225–237.
Ribeiro, P.A., Branco, M., Hawkins, S.J., Santos, A.M., 2010. Recent changes in the distribution of a marine gastropod, *Patella rustica*, across the Iberian Atlantic coast did not result in diminished genetic diversity or increased connectivity. J. Biogeogr. 37, 1782–1796.
Roman, J., Darling, J.A., 2007. Paradox lost: genetic diversity and the success of aquatic invasions. Trends Ecol. Evol. 22, 454–464.
Rousset, F., 2008. Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. Mol. Ecol. Resour. 8, 103–106.
Savini, D., Occhipinti-Ambrogi, A., Marchini, A., Tricarico, E., Gherardi, F., Olenin, S., Gollasch, S., 2010. The top 27 animal alien species introduced into Europe for aquaculture and related activities. J. Appl. Ichthyol. 26, 1–7.
Sekine, Y., Yamakawa, H., Takazawa, S., Lin, Y., Toba, M., 2006. Geographic variation of the COX1 gene of the short-neck clam *Ruditapes philippinarum* in coastal regions of Japan and China. Venus 65, 229–240.
Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139, 457–462.
Vargas, K., Asakura, Y., Ikeda, M., Taniguchi, N., Obata, Y., Hamasaki, K., Tsuchiya, K., Kitada, S., 2008. Allozyme variation of littleneck clam *Ruditapes philippinarum* and genetic mixture analysis of foreign clams in Ariake Sea and Shiranui Sea off Kyushu Island. Jpn. Fish. Sci. 74, 533–543.
Weir, B.S., 1996. Genetic Data Analysis II. Sinauer Associates, Sunderland, MA.
Wonham, M.J., Carlton, J.T., 2005. Trends in marine biological invasions at local and regional scales: the Northeast Pacific Ocean as a model system. Biol. Invasions 7, 369–392.
Yasuda, N., Nagai, S., Yamaguchi, S., Lian, C.L., Hamaguchi, M., 2007. Development of microsatellite markers for the Manila clam *Ruditapes philippinarum*. Mol. Ecol. Notes 7, 43–45.
Yokogawa, K., 1998. Morphological variabilities and genetic features in Japanese common clam *Ruditapes philippinarum*. Jpn. J. Malacol. 57, 121–132.

Chapter 6

Development of a microsatellite toolkit for the identification of genetic stocks in Sardinian populations of the grey mullet (*Mugil cephalus*)

Piero Cossu, Laura Mura, Fabio Scarpa, Tiziana Lai, Daria Sanna, Marco Curini-Galletti, Nicola Fois, Marco Casu

manuscript

Development of a microsatellite toolkit for the identification of genetic stocks in Sardinian populations of the grey mullet (*Mugil cephalus*)

P. Cossu[1], L. Mura[2], F. Scarpa[1], T. Lai[1],D. Sanna[1], M. Curini-Galletti[1], N. Fois[2], M. Casu[1]

[1] Dipartimento di Scienze della Natura e del Territorio, Università di Sassari, Via F. Muroni 25, 07100 Sassari (SS), Italy

[2]Agris Sardegna, Agenzia per la Ricerca in Agricoltura. Dipartimento per la Ricerca nelle Produzioni Animali, Località Bonassai, 07040 Olmedo (SS), Italy

Correspondingauthor:

Piero Cossu

Dipartimento di Zoologia e Genetica Evoluzionistica

Università di Sassari

Via F. Muroni 25

07100 Sassari (SS), Italy

Phone: +39 079 228924

Fax: +39 079 228665

picossu@uniss.it

Abstract

The grey flathead mullet is a widely distributed coastal fish species which may colonise marine, brackish-water and freshwater environments. Such versatility has raised interest into aquaculture and restocking projects regarding this species, because of its high commercial value in many parts of the world. Concerning this perspective, the present study copes with the need for genetic monitoring of wild populations at regional and local scale, as it is recommended to preserve local biodiversity and genetic integrity of natural populations. Here we investigated patterns of genetic variation at four coastal ponds in Sardinia, where there is a flourishing production of mullet roe, which is considered a delicatessen on international markets. We tested whether a set of microsatellite loci other than that already used in the Mediterranean can detect genetic structure at such a fine spatial scale and thus its usefulness for genetic stock identification. Seven out of nine microsatellites were polymorphic and showed great variation in the estimate of Nei's gene diversity ranging from 0.291 to 0.911. Albeit showing less genetic variation, these loci had the same power to detect genetic structure than those formerly used in the Mediterranean. Overall, F-statistics and individual assignment testssuggest the existence of a unique genetic stock along Sardinian coasts, consistently with the high dispersal potential of adults. Taking into account outcomes of simulations, increasing the number of sampled individuals and maybe of loci should improve the power of detecting genetic structure at such spatial scales, thus leading to more accurate genetic stock identification.

Keywords: Microsatellites, *Mugil cephalus*, Genetic stock, western Mediterranean

Introduction

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 uses estuarine environment as nurseries habitat and then migrates offshore for spawning (Thomson 1955,Bacheler et al. 2005). The egg and early larval stages passively disperse drifting in ocean currents (Chang et al. 2000), and thenmigrate onshore at the postflexion larval stage which is followed by temporary occupation of the surf zone as early juveniles (Strydom&d'Hotman 2005). After about a month spent in the sea schools of these fry colonise estuaries and sometimes the adjoining river catchments (Hsu et al. 2009). The grey mullet spent the entire juvenile and sub-adult life stages in these habitats before come back to sea as the adult life-stage is reached (Lawson &Jimoh 2010). The extent to which adult grey flathead mullet can migrate may range from 32 to 700 km and vary between geographical areas (Whitfield et al. 2012 and references therein), thus surface marine currents may be an important factor in promoting long distance dispersal and gene flow over a wide geographic scale (Campton and Mahmoudi 1991, Jamandre et al. 2009, Shen et al. 2011).

Perhaps due to its high dispersal potential and its worldwide distribution genetic studies focused mainly on investigating large-scale patterns of genetic differentiation and assessing whether *M. cephalus* is a single, cosmopolitan species or a complex of cryptic species rather than genetic structuring on a regional scale (Whitfield et al. 2012 and references therein). In fact, significant genetic structure and even sub-species as well as cryptic species have been reported among *M. cephalus* populations inhabiting different oceans (Crosetti et al. 1994, Cardona 2000, Rocha-Olivares et al 2000, 2005, Heras et

al., 2009, Shen et al. 2011, Durand et al. 2012) but see (Livi et al. 2011). Conversely, offshore migration of adult to the sea and then larval dispersal towards estuarine areas may constitute important homogenizing events explaining the reported lack of genetic structure among populations inhabiting more restricted areas within the species range such as the Gulf of Mexico (Campton and Mahmoudi1991,Rocha-Olivares et al. 2000) and the Mediterranean Sea (Rossi et al. 1998,Livi et al. 2011). However, the use of more variable markers as microsatellites revealed the occurrence of genetic structure within this region, roughly matching well-known barriers to dispersal, as well an isolation by distance pattern (Durand et al. 2013).

Disentangling the influence of contemporary factors such as dispersal capability and demographyon the species' genetic structure at regional orfine (sub-regional) spatial scaleis also important as the grey mullet is an important resource worldwide and in Mediterranean waters (Nash &Shehadeh 1980), especially in Egypt, Tunisia, and Morocco (FAO 2008), as well as Greece where it is deemed a species of considerable commercial value (Maitland &Herdson 2009). In many regions of the world it is captured during the spawning migration to harvest the egg roe which is salted and dried to be sold as a delicatessen (Livi et al. 2011). The increased demand for mullet roe during last decades has elevated the status of grey mullet, which has been called '' the grey gold'' by fishermen (Hung and Shaw 2006).Therefore, the capability of living in fresh- brackish- and sea-water during its life-cycle made*M. cephalus*an attractive species to use in aquaculture (Tamaru et al. 2005). To this end, the grey mullet is often stocked in brackish coastal lagoons to improve fish yield, raised in commercial freshwater fish-ponds, and it has also been introduced into inland freshwater lakes and

reservoirs to develop new fisheries (Thomson, 1966, Ben Tuvia et al., 1992).These practices may represent a potential threat to the genetic integrity of localwild populations leading to 1) loss of genetic variation, 2) change of population genetic structure and composition, and 3) lossof local adaptation (Laikre et al. 2010 and references therein). Accordingly, in order to preserve local biodiversity and minimize the risk of genetic pollutionthe Code of Conduct for Responsible Fisheries (FAO 1995)recommend that any genetic change has to be monitored through the genetic characterization of both wild populations and (un)intentionally introduced individuals.

In Sardinia (Italy) grey mullet and particularly mullet roe are an important economic resource; however, the amount of fish caught by fisheries is not sufficient to satisfy the increasing demand oninternational markets of these products (Locci et al. 2011). With regard to mullet roe, for instance, an increasing amount of egg roe has being purchased from fishing areas other than the Mediterranean Sea, such as the western central Atlantic (FAO fishing area 31), the eastern central Atlantic (FAO fishing area 34) and the southwest Atlantic (FAO fishing area 41). Currently, in order to meet requests of increasing the amount of fish caught without depleting natural resources, the local government is planning restocking programs of the coastal ponds (Laura Mura Pers. Comm.). In such a context, the genetic tagging of grey mullets from different coastal ponds might be a useful complement to natural restocking programs. Indeed, it may allow to assess the status and viability of natural population(s) and the number of genetic stocks; furthermore such information is necessary to prevent risks related to genetic homogenization of genetically divergent populations or loss of local adaptation. For instance, Blel et al. (2010) evidenced significant adaptive genetic divergence among

populations of grey mullets due to differences in salinity among brackish water and marine environments.

The aim of this study was to develop a ready-to-use microsatellite toolkit and test its efficiency to assess genetic variation and carry out genetic stock identification (GSI) in populations of *M. cephalus*.Such tools may be helpful in the proper management of the resource and the restocking programs, as well as to preserve the genetic diversity of natural populations. To this end we screened nine microsatellite loci available in the literature and used these loci to set up multiplex PCR reactions.Microsatellites are suitable markersfor population genetic studies at regional and fine scales of distribution of *M. cephalus*(Miggiano et al. 2005), as well as for uncoveringthe influence of contemporary factors such as dispersal and demography on population structure (Gonzalez &Zardoya 2007).

Materials and methods

Sampling and DNA extraction

Fin-clips of grey mullets(n = 120) were collected during 2013 at fourpondsalong the coasts of Sardinia(Table 1).We considered three distinct spatial scales, ranging from the gulf of Oristano to different sub-basins such as the Thyrrenian and the Sardinian Sea (Fig.1).Dorsal Fin-clips were preserved in absolute ethanol and stored at -80°C until DNA isolation.Genomic DNA was purified using the salting-out extraction method (Waskoet al. 2003), and then stored in TE buffer. DNA quantity and quality was assessed using a fluorimeter (Nanodrop 2000), and diluted if necessary.

PCR protocols

A series of Multiplex PCR reaction protocols weresetupto amplify the nine microsatellite markers available on Genbank database (Table 2); 5' end of the primers were marked with the 6-FAM, VIC, NED and PET fluorochromes. PCR reactions were prepared in a final volume of 25 µl, containing 20–30 ng of genomic DNA, 1X reaction buffer (Euroclone), 2 mMMgCl2, 0.250 µMof each dNTP, 0.03 - 0.07µM of each primer, 1.25 U of EuroTaq DNA polymerase (Euroclone). PCR reactions werecarried outin a MJ DNA Engine PTC-100 thermal cycler under the following conditions: an initialdenaturation step at 94°C for 5 min; 30 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 30 s; a final extension of 72°C for 7 min. After checking 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 GeneScan500(-250) LIZ size standard (Applied Biosystems) was run on an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems).Microsatellite loci were screened and scored by resulting electropherograms using GeneMapperv4.0 software package (Applied Biosystems).

Statistical data analysis

Departure from Hardy-Weinberg equilibrium (HWE) at each locus and population, as well as linkage disequilibrium (LD) among each pair of loci were tested using GENEPOP 4.0 (Rousset 2008). The probability of departures was estimated using the Markov chain method (10'000 dememorization steps, 100 batches of 10'000 iterations

each), and applying the False Discovery Rate (FDR) Method of correction for multiple comparisons (Benjamini & Yekutieli 2001).The presence of null alleles, stuttering, and large allele dropouts was tested using MICROCHECKER 2.2.3 (Van Oosterhoutet al.2004).

We also carried out an outlier testbased on the levelsof inter-population differentiation for identifying candidate loci for selection.We used the coalescent-based simulation approach that display unusually high and low values of F_{ST} by comparing the observed F_{ST} values with those expected under neutrality(Beaumont & Nichols 1996).We used the FDIST2 approachimplemented in the Lositan Selection Workbench (Antao et al. 2008)based upon the finite island model of migration(Beaumont and Balding 2004). We run 100,000 simulations with 100 demes under the infinite allele model (IAM) and the stepwise mutation model (SMM), enabling "neutral mean F_{ST}" and "force mean F_{ST}" options to increase reliability of neutral mean F_{ST} estimates. These options perform a preliminary run to detect loci lying outside the 99.5% confidence interval (CI) and exclude them before the estimation of mean neutral F_{ST}. This refined estimate is then used as the mean neutral $F_{ST}$for the final run, and the probability that an outlier locus lying outside the 99.5% CI is under selection is corrected setting the FDR rate threshold at 0.05.

We used GenAlex 6.5 (Peakall& Smouse 2012) to calculate summary statistics for the mean number of alleles (N_A), the effective number of alleles (N_E), Nei'sgene diversity (H_E), and observed heterozigosity (H_O). The same software was also used to estimate overall and population pairwise F_{ST} statistics. A permutation test with 10'000 replicates was used to assess whether or not F_{ST} was significantly greater than zero.

Additionally, genetic differentiation was assessed using a Fisher's exact test of population differentiation (Rousset2008) as implemented in GENEPOP 4.0. We tested for the heterogeneity of allelic frequencies among populations using the Markov chain method (10'000 dememorization steps, 100 batches of 10'000 iterations each). Whenever necessary, type I errors due to multiple comparisons were accounted for by using the sequential Bonferroni correction to adjust probability values(Rice 1989).

The simulation program POWSIM V1.2 (Ryman and Palm,2006) was used to estimate the statistical power of the dataset in detecting significant genetic differentiation. The hypothesis of genetic homogeneity can be assessed through optional combinations of the number of samples, sample sizes, and the number of loci, alleles, and allele frequencies for any hypothetical degree of true differentiation (quantified as F_{ST}). The statistical power of the dataset was evaluated using Chi-squared and Fisher's exact tests. Simulations were performed using different combinations of effective populations size (N_e = 2000 and 5000)and the time since divergence in number of generations (t = 5, 10, 25, 50, 100), leading to F_{ST} values in the range 0.001–0.01. Pairwise F_{ST} values are similar to those found in this study (see Results section), whereas $N_e$estimates of this magnitude have been found in most marine fishes(Hauser &Carvalho 2008). We also considered a sample sizes as big and twice as that used in this study to test for the effect of sample size on the power of detecting genetic differentiation. Estimates ofα (type I) error were generated using samples drawn directly from the base population, omitting the drift steps (t = 0) leading to the absence of differentiation (F_{ST}= 0). The estimate of power was reported as the proportion of significant outcomes ($p < 0.05$) after 1000 replicates.

Genetic stock identification (GSI) was performedusing both a maximum-likelihood and a Bayesian approach to test the extent to which individuals could be assigned to their sampling collection, assuming that individuals have originated in that population. The first method of self-assignment uses the leave-one-out cross-validation method developed by Anderson et al. (2008) and is implementedin the software ONCOR (Kalinowski 2007).The second method is implemented in Geneclass2 (Piry et al. 2004) and it has been developedbyRannala and Mountain (1997).However, the probability of the more likely genetic source for a given individual was computed using the Monte Carlo resampling method of Paetkau et al. (2004) with 10'000 iterations and a threshold for the probability of exclusion (PE) set at 0.01.Prior to be assigned, individuals are excluded from the reference dataset according to the leave-one-out method (Efron. 1983).

Results

One locus (MuCE9) was monomorphic in all the samples examined and thus it was not considered further; conversely the locus MuCE55 was removed because all the individuals were heterozygous for the same alleles and hence uninformative. Among the remaining seven loci, the number of alleles ranged between 4 and 20, and Nei's gene diversity between 0.291 and 0.911 (Table 2). Subsequently, 3 out of 7 loci displayed a polymorphism information content (PIC) greater than 0.7, whereas the remaining ones were characterized by a PIC < 0.5 (Table 2).

Significant LD was found at 3 out of 85 loci pairs after controlling for false positives; however physical linkage is unlikely as LD among pairs of loci was not consistent

across populations. Two loci showedsignificant departure from HWE at least in one sample (Mce25 in CAB and POP, Mce3 in CAB).Nevertheless, in the case of CAB deviations from HWE are likely due to the presence of null alleles. None of the seven loci retained for the subsequent analyses showed and F_{ST} greater (divergent selection) or lower (balancing selection) than values expected under the hypothesis of selective neutrality (Fig. 2).

The number of alleles (N_A) and of effective alleles (N_E), the mean observed and expected heterozygosity (H_O and H_E, respectively) are summarized in Table 2. Overall, the mean number of alleles across all populations (N_A) was 7.393 ± 0.777, but the mean effective number of alleles (the number of alleles weighted by the frequency) was less than half (N_E = 3.486 ± 0.513). values of Nei's gene diversity were similar across samples with the highest value fond in POP (mean H_E = 0.591), whereas the lowest was found in CAB (mean H_E = 0.574).The global F_{ST} value did not show significant genetic variation among the samples of grey mullet(F_{ST}=0.015 ± 0.003, P > 0.05). Actually, only one population pair (MIS-POP) showed a significant genetic differentiation according to both pairwise F_{ST}values and the exact test for population differentiation (Table 3).

The simulation of datasets with number of alleles, allele frequencies and sample sizes similar to those characterizing the present study, detected genetic differentiation in at least 95% of cases (χ^2 test) when $F_{ST} \geq 0.01$ (Fig. 3a and 3b). When $F_{ST} \geq 0.005$power decreased to 82.2% for N_e = 2000 (Fig. 3a) and to 68% for N_e = 5000 (Fig. 3b). Doubling the number of sampled individuals increased the power to detect genetic differentiation to >99% when true $F_{ST} \geq 0.005$ and to >70% when true $F_{ST} \geq 0.002$ (Fig.

3a and 3b). the estimate of the type I statistical error (detecting genetic differentiation when genetic homogeneity is true) was smaller than 0.05 in all the simulations, according to the χ^2 test (Fig. 3).

Results of individual assignmentare summarized in Table4. Assuming that specimens have originated nearby the coastal pond from which they have been collected, both the maximum-likelihood and theBayesian approach performed quite poorly. Overall, these methods assigned to the sample collection of origin 25.4 and 25.6 % of individuals, respectively. Among the four samples, TOR showed the lowest self-classification, as only 3 out of 30 individuals collected from TOR were self-assigned to this sample collection, according to both methods. Conversely, maximum likelihood and Bayesian assignment evidenced divergent results with respect to the sample collection showing the highest self-classification. According to the maximum likelihood method POP was the sample collection with the highest proportion of self-assigned individuals (37.5%), whereas MIS (42.8%) ruled over the others according to the Bayesian method.

Discussion

The increasing demand of grey mullets and its most economically valuable product, the mullet roe, has boosted since 2010 aquaculture projects that according to FAO data have reached a global production of 141731 t in 2012 (www.fao.org). Only a small part of such production was from Europe (770 t) and almost half of it was produced in Italy (380 t). In Europe, wild captures still outnumber farmed fish in 2011 (1179 t) though the amount of grey mullet caught has been decreasing since 2005 when fishing reached 2652 t (www.fao.org). Restocking of natural populations or increased production of

farmed grey mullets in Europe might avoid the further depletion of natural resources, particularly if these practices are coupled with the genetic monitoring of natural populations in order to preserve the species' genetic diversity (FAO, 1995).

The present study tested the efficacy of nine microsatellite for assessing patterns of genetic diversity and performing genetic stock identification in grey mullets along the Sardinian coasts.The rationale behind the choice of microsatellite markers other than those which have been successfully used in the Mediterranean Sea (Blel et al 2010, Durand et al. 2013) was twofold. First, while the microsatellites developed by Miggiano et al. (2005) successfully detected genetic structure at basin scale (Durand et al. 2013), the same markers did not evidence genetic structuring on a smaller spatial scale (Blel et al. 2010), whose extent is similar to that adopted in the present study. Hence, increasing the panel of available markers may improve the chance of detecting genetic structure more efficiently than increasing the number of individuals (Kalinowski 2005).

After excluding two non-informative loci (MuCE55 and MuCE9), the remaining seven loci were not significantly affected by Hardy-Weinberg or Linkage disequilibria as well as selection (Fig. 2). Among the three loci departing from HWE deviation was due to the presence of null alleles in one sample; however, at both loci null allele frequency was smaller than 0.2, the threshold below which the bias induced by null alleles for population structure estimates is negligible (Dakin & Avise 2004). Hence, the only locus at which HWE departure did not reflect null alleles, scoring errors or large allele dropout is Mce25 in POP.Such deviationsfrom Hardy–Weinberg equilibrium might reflect non-random mating,population mixing or natural selection (Ward 2006). Nevertheless non-random mating or population admixture (Wahlund effect) are

expected to affect most loci in a population; in the other hand none of the loci used in this study was selected as an outlier for natural selection (Fig. 2). Hence the process responsible for the observed HW disequilibrium remains to be cleared.

Within population genetic variability (Table 2) was lower than estimates that were previously reported in the Mediterranean Sea. For instance, in populations along Tunisian coasts mean observed heterozygosity and number of alleles ranged from 0.854 to 0.895, and from 13.57 to 14.71, respectively (Blel et al. 2010). Similar estimates were reported by Durand et al. (2013) who analysed genetic variation across Mediterranean and the Black Sea. Some loci used in this study were also characterized by lower gene diversity than estimates for the same loci reported by Shen et al. (2011) and Xu et al. (2011) in populations of *M. cephalus* from the north-western Pacific. However it is likely thatgene diversity might have been affected by the mixing of distinct cryptic species at some locations in north-western Pacific region. Indeed, our mean estimate of gene diversity (H_E = 0.577) and number of alleles (N_A = 7.393) are within the range reported for the three main lineages separately by Shen et al. (2011). Hence, it is likely that the lower genetic variation observed at some microsatellite loci in our study is due to a relatively low mutation rate.

This fact is not detrimental *per se*; for instance, microsatellite loci with low to moderate number of alleles may retain phylogeographic signal for a longer time. In fact, high mutation rate may increase homoplasy which in turn hampers the detection of genetic structure (O'Reilly et al. 2004). Another advantage of loci with low to moderate levels of heterzigosity (low mutation rate) concerns the less biased estimates of F_{ST}, and thus of demographic parameters that are important for conservation and management of

natural resources. Because of its negative dependence on heterozygosity the relationship between F_{ST} and true genetic differentiation is not linear at high levels of heterozygosity (Jost 2008). Hence, based on F_{ST} estimates populations may appear weakly differentiated even without sharing any allele (Meirmans& Hedrick 2011). Notwithstanding several F_{ST} analogues have been developed so far, these values are less biased estimates of allelic differentiation and their usefulness in estimating demographic parameters is unclear (Whitlock 2011).

In the present study only one population pair (MIS – POP) displayed a significant genetic differentiation according to both F-statistics and the exact test of population differentiation (Table 3). The simulations carried out to assess whether the observed pattern may reflect a statistical artefact evidenced that microsatellite loci screened in this study had sufficient statistical power to detect a level of genetic structure equal or greater than 0.01 (Fig. 3). This outcome is consistent with the fact that genetic divergence between MIS and POP was greater than these value (F_{ST} = 0.017) and, noteworthy, with results reported in previous studies. For instance all pairwise F_{ST} values among Mediterranean populations reported in Durand et al. (2013) were greater than 0.01 (F_{ST} range = 0.012 – 0.051). In Shen et al. (2011) the only population pair that was genetically divergent within lineages displayed a F_{ST} value of 0.016. In these studies most of sample sizes were similar or lower than sample size considered in the present study. Our results also suggest that doubling the sample sizemight increase the power of detecting genetic differentiation when $F_{ST} \geq 0.05$.

Likely, more loci are needed to detect population structure when $F_{ST} \leq 0.05$, as well as to carry out successfully GSI when genetic differentiation is weak. Based on the

assumption that individuals have originated in the coastal pond where have been collected, both maximum-likelihood and Bayesian self-assignment performed poorly (Table 4). However it should be considered that in addition to passive larval dispersal grey mullets may disperse long distances as adults even against marine currents. Migrations of more than 200 km have been observed and, according to some authors adult grey mullets might disperse as far as 700 km (Whitfield et al. 2012 and references therein). Such migrations may ensure high levels of gene flow among locations, leading to genetic homogenization at regional scales, as observed by Blel et al. (2010) along the Tunisian coasts. The high dispersal potential along shorelines is consistent with the fact that grey mullets sampled at a given coastal pond were more or less uniformly assigned across locations.

In conclusion, the present study tested the usefulness of nine microsatellite loci in detecting patterns of genetic variation in grey mullets collected at four coastal ponds along the Sardinian coast (western Mediterranean). Our results highlighted that loci used in this study had the same power to detect genetic structure than loci already used in the Mediterranean Sea (Blel et al. 2010, Durand et al. 2013). Outcomes based on this molecular dataset point also to the presence of a unique genetic stock along the Sardinian coasts, a scenario fitting the high dispersal capabilities of larvae, juveniles and adults of this species. However, given the limited number of loci used in this study, such scenario need to be confirmed by further studies. Increasing the sample size of individuals and maybe of microsatellite loci will likely improve the power of detect fine scale genetic structure. Future studies might also integrate biophysical models of

dispersal into the analysis framework to get a more detailed picture of population connectivity in *M.cephalus* at regional scale.

Acknowledgements

We are indebted to fishermen that kindly provided fin-clip samples. This work was part of LM's Ph.D. thesis.

References

Anderson E.C., Waples R.S., Kalinowski S.T. (2008) An improved method for estimating the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences 65: 1475-1486

Antao T., Lopes A., Lopes R.J., Beja-Pereira A., Luikart G. (2008)LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics 9:323

Bacheler N.M., Wong R.A., Buckel J.A. (2005) Movements and mortality rates of striped mullet in North Carolina. North American Journal of Fisheries Management 25:361–373

Beaumont M.A., Nichols R.A.(1996) Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society of London, Series B: Biological Sciences 263: 1619–1626

Beaumont M.A., Balding D.J. (2004) Identifying adaptive genetic divergence amongpopulations from genome scans. Molecular Ecology 13: 969–980

Blel H., Panfili J., Guinand B., Berrebi P., Said K., Durand J.D. (2010) Selection footprint at the first intron of the *Prl*genein natural populations of the flathead mullet (*Mugilcephalus*L., 1758). Journal of Experimental Marine Biologyand Ecology 387: 60–67

Ben Tuvia A., Davidoff E.B., Shapiro J., Shefler D. (1992) Biology and management of lake Kinneret fisheries. Israeli Journal of Aquaculture-Bamidgeh 44: 48– 65

Campton D.E., Mahmoudi B. (1991) Allozyme variation and population structure of striped mullet (*Mugilcephalus*) in Florida. Copeia 1991:485–492

Cardona L. (2000) Effects of salinity on the habitat selection and growth performance of Mediterranean flathead grey mullet *Mugilcephalus* (Osteichthyes, Mugilidae). Estuarine Coastal and ShelfScience 50:727–737

Chang C.W., Tzeng W.N., Lee Y.C. (2000) Recruitment and hatching dates of grey mullet (*Mugilcephalus* L.) juveniles in the Tanshui estuary of northwest Taiwan. Zoological Studies 39:99–106

Crosetti D., Nelson W.S., Avise J. (1994) Pronounced genetic structure of mitochondrial DNA among populations of the circumglobally distributed grey mullet (*Mugilcephalus* Linnaeus). Journal of Fish Biology 44:47–58

Dakin E.E.,Avise J.C. (2004) Microsatellite null alleles in parentage analysis. Heredity 93: 504–509

Durand J.-D., Shen K.-N., Chen W.-J., Jamandre B.-W., Blel H., Diop K., Nirchio M., García de León F.J., Whitfield A.K., Chang C.-W., Borsa P. (2012) Systematics of the grey mullet (Teleostei: Mugiliformes: Mugilidae): molecular phylogenetic

evidence challenges two centuries of morphology-based taxonomic studies. Molecular Phylogenetics and Evolution64:73–92

Durand J.D., Blel H., Shen K.N., Koutrakis E.T.,GuinandB. (2013) Population genetic structure of *Mugilcephalus*in the Mediterranean and Black Seas: a singlemitochondrial clade and many nuclear barriers. Marine Ecology Progress Series 474: 243–261

Efron B. (1983) Estimating the error rate of a prediction rule:improvement on cross-validation. Journal of the AmericanStatistical Association 78:316–331

FAO (1995) Code of conduct for responsible fisheries. FAO, Rome, 41 p.

Gonzalez E.G., Zardoya R. (2007) Relative role of life-history traits and historical factors in shaping genetic population structure of sardines *(Sardinapilchardus)*. BMC Evolutionary Biology 7: 197

Hauser L., Carvalho G.R. (2008) Paradigm shifts in marine fisheries genetics: ugly hypothesesslain by beautiful facts. Fish and Fisheries 9: 333–362

Heras S., Roldán M.I., Castro M.G. (2009) Molecular phylogeny of Mugilidae fishes revised. Reviews in Fish Biology and Fisheries 19: 217–231

HsuC.C., ChangC.W., Iisuka Y., TzengW.-N. (2009)Agrowth check deposited at estuarine arrival in otoliths of juvenile flathead mullet (*Mugilcephalus* L.). Zoological Studies 48:315–324

Hung C.-M., Shaw D. (2006) The impact of upstream catch and global warming on the grey mullet fishery in Taiwan: A non-cooperative game analysis. Marine Resource Economics 21:285–300

Jamandre B.W., Durand J.-D., Tzeng W.N. (2009) Phylogeography of the flathead mullet *Mugilcephalus* in the north-west Pacific as inferred from the mtDNA control region. Journal of Fish Biology 75:393–407

Jost L. (2008) GST and its relatives do not measuredifferentiation. Molecular Ecology 17: 4015–4026

Kalinowski S.T. (2005) Do polymorphic loci require large sample sizes to estimate genetic distances? Heredity 94:33-36

Kalinowski S.T. (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Molecular Ecology Notes 5: 187-189

Kalinowski S.T., ManloveK.R., Taper M.L.(2007) ONCOR: acomputer program for genetic stock identification. http://www.montana.edu/kalinowski/Software/ONCOR.htm

Lawson E.O., Jimoh A.A.-A. (2010) Aspects of the biology of grey mullet, *Mugilcephalus*, in Lagos Lagoon, Nigeria. AACL Bioflux 3:181 – 193

Laikre L., Schwartz M.K., Waples R.S., Ryman N. (2010) Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. Trends in Ecology and Evolution 25: 520–529

Livi S., Sola L., CrosettiD.(2011) Phylogeographic relationships among worldwide populations of the cosmopolitan marine species, the striped grey mullet (*Mugilcephalus*), investigated by partial cytochrome b gene sequences. Biochemical Systematics and Ecology 39:121–131

Locci E., Piras C., Mereu S., Marincola F.C., Scano P. (2011) H NMR metabolite fingerprinting and pattern recognition of mullet (*Mugilcephalus*) bottarga. Journal of Agricultural and food Chemistry 59: 9497 – 9505

Maitland P.S., Herdson D. (2009) Key to the marine and freshwater fishes of Britain and Ireland. Environment Agency, Bristol

Meirmans P.G., Hedrick P.W. (2011) Assessing populationstructure: F_{ST} and related measures. Molecular EcologyResources 11: 5– 18

Miggiano E., Lyons R.E., Li Y., Dierens L.M., Crosetti D., Sola L. (2005) Isolation and characterization of microsatellite loci in the striped mullet, *Mugilcephalus*. Molecular Ecology Notes5:323 - 326

Nash E.C., Shehadeh Z.H. (1980) Review of breeding and propagation techniques for grey mullet *Mugilcephalus* L. ICLARM Studies and Reviews 3: 1 – 40

O'Reilly P.T., Canino M.F., Bailey K.M., Bentzen P. (2004) Inverse relationship between

F_{ST} and microsatellite polymorphism in the marine fish, walleye pollock (*Theragrachalcogramma*): implications for resolving weak population structure. Molecular Ecology 13: 1799–1814

Paetkau D., Slade R., Burden M., Estoup A. (2004) Direct, real-time estimation of migration rate using assignment methods: a simulation-based exploration of accuracy and power. Molecular Ecology13:55-65

Peakall R., Smouse P.E. (2012)GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537-2539

Piry S., Alapetite A., Cornuet J.-M., Paetkau D., Baudouin L.,Estoup A. (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95: 536 – 539

Rannala B., Mountain J.L. (1997) Detecting immigration by using multilocus genotypes.

Proceedingsof the National Academy of Sciences U.S.A. 94: 9197–9201

R Core Team (2013) R: A language and environment for statisticalcomputing. R Foundation for StatisticalComputing, Vienna, Austria. URL http://www.R-project.org/

Rice W.R. (1989) Analyzing tables of statistical tests. Evolution 43: 223 – 225

Rocha-Olivares A., Garber N.M., Stuck K.C. (2000) High genetic diversity, large inter-oceanic divergence and historical demography of the striped mullet. Journal of Fish Biology 57: 1134 – 1149

Rocha-Olivares A., Garber N.M., Garber A.F., Stuck K.C. (2005) Structure of the mitochondrial control region and flanking tRNA genes of *Mugilcephalus*. Hydrobiológica 15: 139 – 149

Rossi A.R., Capula M., Crosetti D., Campton D.E., Sola L. (1998) Genetic divergence and phylogenetic inferences in fivespecies of Mugilidae (Pisces: Perciformes). Marine Biology 131:213–218

Rousset F. (2008) Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. Molecular Ecology Resources 8: 103–106

Ryman N., Palm S. (2006) POWSIM: a computer program forassessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6: 600–602

Shen K.-N., Jamandre B.W., Hsu C.–C., Tzeng W.-N., Durand J.-D. (2011) Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet *Mugilcephalus*. BMC Evolutionary Biology 11: 83

Strydom N.A., d'Hotman B.D. (2005) Estuary-dependence oflarval fishes in a non-estuary associated South African surfzone: evidence for continuity of surf assemblages. Estuarine Coastal and Shelf Science 63:101–108

Tamaru C.S., Tamaru C.C.-T., FitzGerald W.J., Ako H., Sato V.T. (2005) Advances in the culture of striped mullet. American Fisheries Society Symposium 46:439–456

Thomson J.M. (1955) The movements and migrations of mullet (*Mugilcephalus* L.). Australian Journal of Marine & Freshwater Research 6: 328 – 347

Thomson J.M. (1966) The grey mullets. Oceanography and Marine Biology - an Annual Review 4:301–355

Van Oosterhout C., Hutchinson W.F., Wills D.P.M., Shipley P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538

Ward R.D. (2006) The importance of identifying spatial population structure inrestocking and stock enhancement programmes. Fisheries Research 80: 9–18

Wasko A.P., Martins C., Oliveira C., Foresti F. (2003) Non-destructive genetic sampling in fish. An improved method for DNA extraction from fish fins and scales. Hereditas 138:161-165

Weir B.S.,Cockerham C.C. (1984) Estimating F-Statistics for the Analysis of Population Structure, Evolution 38: 1358-1370

Whitfield A.K., Panfili J., Durand J.D. (2012) A global review of the cosmopolitan flathead mullet *Mugilcephalus* Linnaeus, 1758 (Teleostei: Mugilidae), with emphasis on the biology, genetics, ecology and fisheries aspects of this apparent species complex. Reviews in Fish Biology and Fisheries 22:641 – 681

Whitlock M.C. (2011) G'_{ST} do not replace F_{ST}. Molecular Ecology 20: 1083 – 1091

Figures and Tables

Fig. 1 Sampling map of *Mugil cephalus*. CAB = Cabras pond; MIS = Mistras pond; POP = Porto Pino pond; TOR = Tortolì pond.

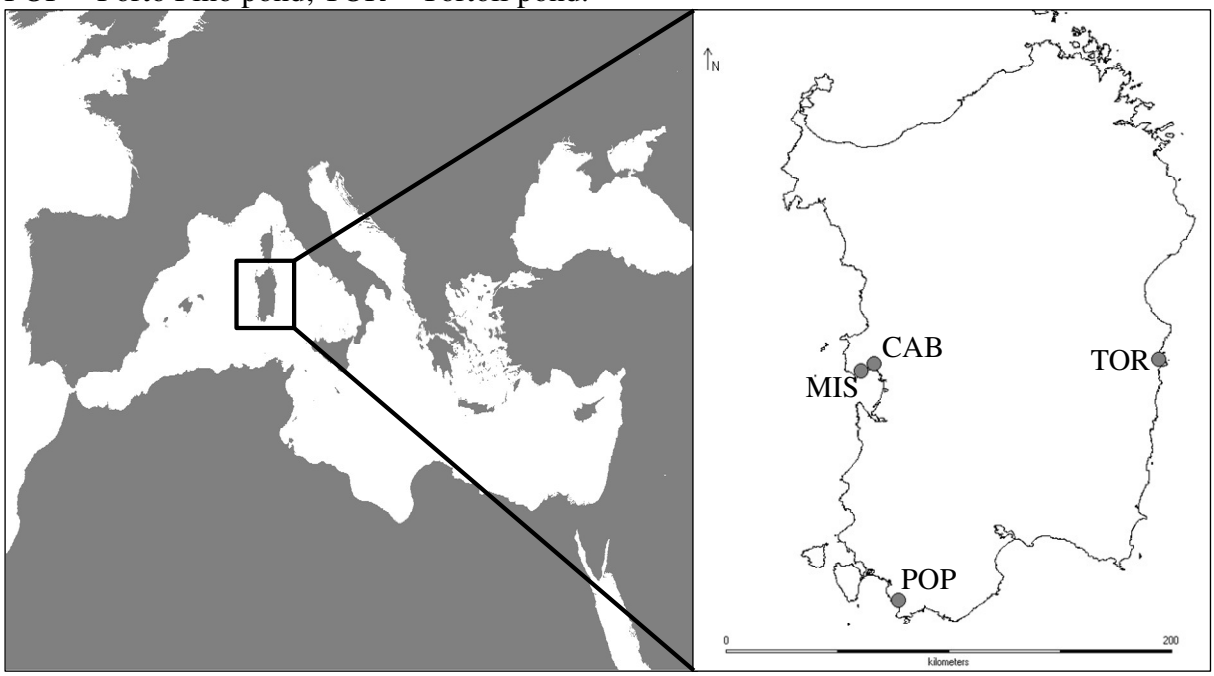

Fig. 2 Results of outlier detection tests. Observed F_{ST} values of loci used in this study (black dots) are plotted against Nei's gene diversity (H_E). The white area delimit the range of neutral simulated F_{ST} expected according to a finite island model of migration. Loci lying in the area shaded with dark grey (F_{ST} greater than the expected null distribution of neutral values) are outliers for divergent selection, whereas those lying in light grey shaded area are outlier for balancing selection (F_{ST} lower than expected based on the null distribution).

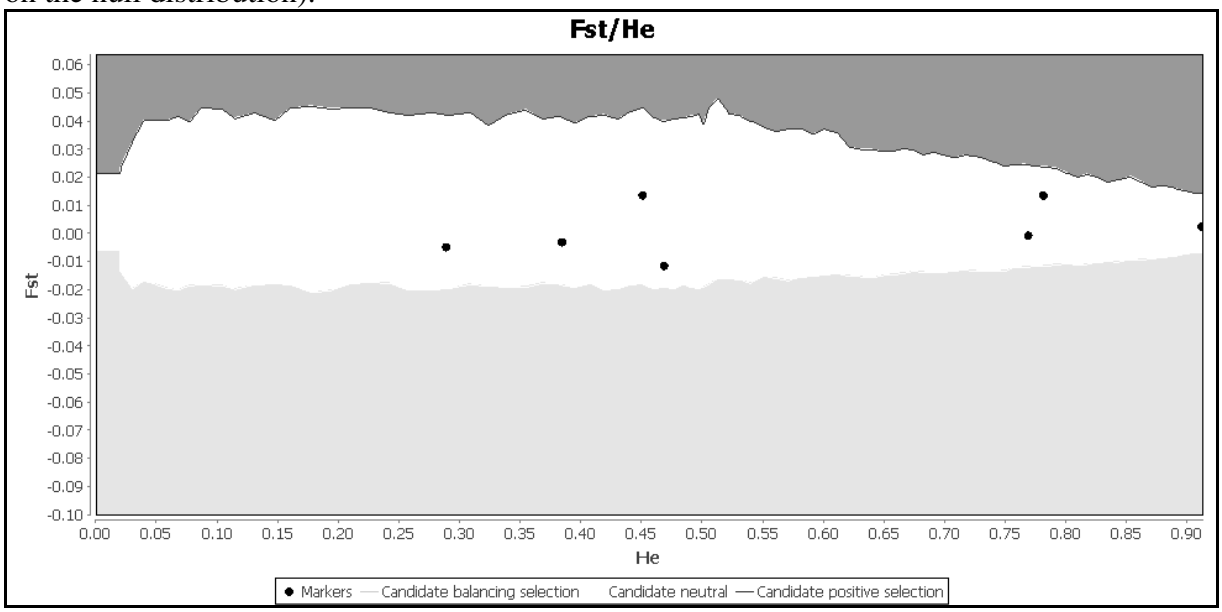

Fig. 3 Results of POWSIM simulations to assess the statistical power of the microsatellite loci to detect genetic differentiation at varying levels of F_{ST}. Targeted F_{ST} values are plotted on the horizontal axis, while the vertical axis displays the proportion of simulationsshowing significant genetic differentiation according to a χ^2 test of genetic heterogeneity. Solid dark grey line plots results obtained in simulation with the same sample size (N = number of individuals) adopted in this study; solid light grey line plots results obtained after doubling the sample size, whereas the dashed black line delimit the critical value for statistical type I error ($\alpha = 0.05$). a) simulations carried out assuming an effective sample size (N_e) of 2000 individual; b) N_e = 5000 individuals.