Vol. 229: 279–289, 2002 | MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser | Published March 20

Population analysis of the sand smelt *Atherina boyeri* (Teleostei Atherinidae), from Italian coastal lagoons by random amplified polymorphic DNA

Leonardo Congiu*, Remigio Rossi, Giuseppe Colombo

Dipartimento di Biologia, Sezione di Biologia Evolutiva, Università di Ferrara, Via Borsari 46, 44100 Ferrara, Italy

ABSTRACT: *Atherina boyeri* Risso, 1810 is a euryhaline species of ecological and economic relevance, highly diffused in most Mediterranean coastal lagoons and characterized by high morphological variability among populations. To study the genetic relationships among *A. boyeri* populations, 11 Italian lagoons and 2 freshwater lakes were sampled and a random amplified polymorphic DNA analysis was performed. Six arbitrarily designed primers were used. Of the 43 scored markers, 39 were polymorphic. No population-specific markers were found. Statistical analysis based on inferred allele frequencies (using Wright's F_{st} statistic) and on presence or absence of bands (using analysis of molecular variance) revealed significant population structure. Unweighted pair group method with arithmetic mean cluster analysis based on Nei's genetic distances showed a geographic clustering of populations. Mantel tests confirmed a high correlation between genetic and geographic distances along coasts. These results suggest the occurrence of coastal gene flows among populations, probably because the anadromic behavior of this species may be only partially phylopatric. However, these migratory movements have not blurred the existing differences among populations.

KEY WORDS: *Atherina boyeri* · Sand smelt · Random amplified polymorphic DNA · Coastal fishes · Population genetic structure · Coastal lagoons

INTRODUCTION

Brackish water ecosystems are often exposed to wide variations of environmental parameters such as temperature and salinity, which may cause strong selective pressures on organisms. These pressures, in association with geographical discontinuity, can play a relevant role in separating species inhabiting these environments into different populations. The analysis of the geographical distribution of genetic variability of brackish water species may be especially relevant to study the evolutionary role of these habitats.

Many fish species are closely tied to these environments, for at least a part of their life cycle, and their genetic structure is the result of 2 opposing factors: the first is the tendency to isolation among different populations because of considerable habitat fragmentation; the second is the homogenizing effects of gene flow among populations, which depends on the dispersal capability of each species (Waples 1987, Ward et al. 1994). Typically, species with high dispersal capabilities show low levels of population structure while species with little dispersal capabilities or with physical or behavioral barriers to movement present high differences among populations (Waldman & Wirgin 1994). Furthermore, some fish species have anadromic behavior and move to brackish water for reproduction; for these species, population structuring is mediated by the level of homing fidelity of spawning adults to their natal lagoons (Stabile et al. 1996).

*E-mail: leo@dns.unife.it

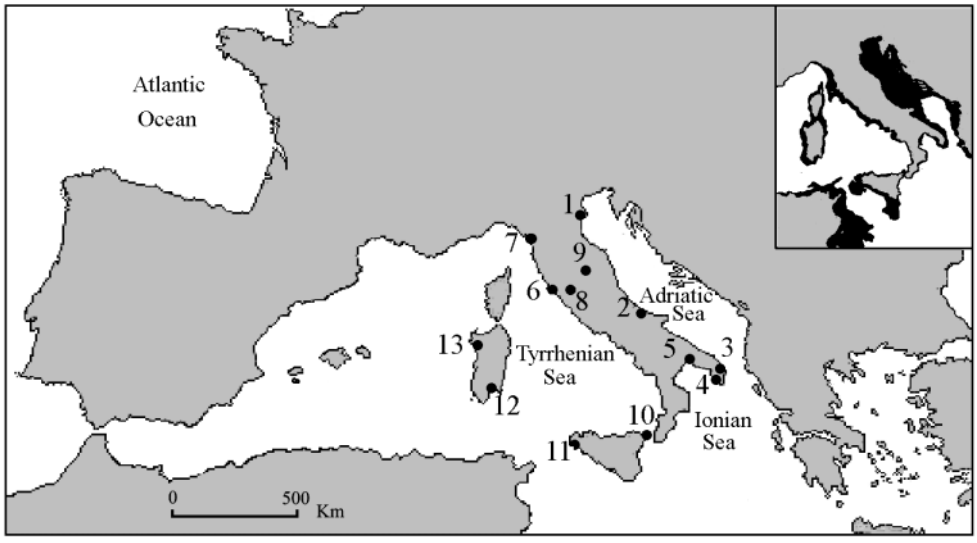

Fig. 1. *Atherina boyeri.* Map of the sampling sites for populations: Scardovari (1), Lesina (2), Acquatina (3), Ugento (4), Taranto (5), Orbetello (6), Massaciuccoli (7), Bolsena Lake (8), Trasimeno Lake (9), Faro (10), Marsala (11), Cagliari (12) and Alghero (13). The inset shows the land extension during the last glaciation (18 000 years ago)

Along the over 8000 km of Italian coastline (including Sicily and Sardinia), there are many lagoons displaying different ranges of temperature, salinity, oxygen concentration and surface area (Sacchi 1979). One of the most common fish species of these environments is the sand smelt *Atherina boyeri* Risso, 1810 (Teleostei Atherinidae), which is a consistent part of fish and bird predator catch and of fisheries (Moretti et al. 1959, Boscolo 1970, Froglia & Orel 1979). This species is euryhaline: the adults migrate to sea in autumn and enter the lagoons in spring for reproduction. They spawn in inshore shallow waters and attach the eggs by filaments to seaweeds and rocks (Jorné-Safriel & Shaw 1966, Bamber & Henderson 1985). Larvae and juveniles grow there (Boscolo 1970). *A. boyeri* may also spend its whole life cycle in freshwater lakes, where it is accidentally introduced (Moretti et al. 1959). The *A. boyeri* populations from different lagoons vary in morphological, morphometric and meristic characters (Marfin 1981, 1982, Mistri & Colombo 1988, Creech 1991, Trabelsi et al. 1994, 1995), and this variability caused controversies about the systematics of the genus (Kiener & Spillmann 1969, 1972).

The morphological and morphometric differences among populations from different lagoons may reflect not only phenotypic plasticity but also the effect of isolation, selection and genetic drift. Indeed, the fragmentation of the habitat could be the cause of a genetic differentiation among populations, as already found in another common lagoon species, *Aphanius fasciatus* (Maltagliati 1999). On the other hand, annual migrations of sand smelt adults from lagoons toward the coastal sea may be sufficient to guarantee an exchange of individuals among populations with a consequent gene flow among brackish environments.

A random amplified polymorphic DNA (RAPD) analysis (Welsh & McClelland 1990, Williams et al. 1990) was performed on 13 different Italian populations, 11 from coastal lagoons and 2 from freshwater lakes, to describe the level of genetic structure in *Atherina boyeri* populations.

The aim of this work was to test whether the high morphological diversity reported in the literature is associated with high genetic diversity among populations. We also investigated possible evidence of gene flow along Italian coasts. This information, together with the amount of genetic structure among populations, could allow the evaluation of the specificity of homing behavior during the reproductive season and to contribute to a more general knowledge on population dynamics of coastal Mediterranean fish species.

MATERIALS AND METHODS

Sample collection. Thirteen sand smelt populations were sampled in 11 Italian coastal lagoons and in 2 freshwater lakes of central Italy (Fig. 1). The lagoons were Scardovari (1), Lesina (2) and Acquatina (3) along the Adriatic shore; Ugento (4) and Taranto (5) along the Ionian shore; Orbetello (6) and Massaciuccoli (7) near the Tyrrhenian sea; Faro (10) and Marsala (11) in Sicily; and Cagliari (12) and Alghero (13) in Sardinia. The lakes were Trasimeno (9) and Bolsena (8). The presence of sand smelts in these closed basins is probably accidental, due to introduction within restocking programs of Mugilidae.

Within a few hours after capture the fishes were cut into 3 to 4 pieces, separately preserved in about 40 ml of 95% ethanol and then stored at –20°C until DNA

extraction. Genomic DNA was extracted from 10 individuals from each population analyzed.

DNA extraction. Classical extraction procedures including a digestion step with proteinase-K were unsuccessful: the DNA was completely degraded into approximately 200 base pair fragments (data not shown). Most likely, DNA was degraded during the proteinase-K digestion step. We had to resort to an alternative extraction method originally developed for plants (Roy et al. 1992, Wolff & Peters-Van Rijn 1993). Approximately 200 mg of head tissue, cleaned from skin and scale, was frozen in liquid nitrogen, ground and then resuspended in an extraction buffer (100 mM Tris HCl, 1% sodium dodecyl sulfate, 100 mM NaCl, 10 mM EDTA, pH 8). Three extraction steps were performed with equal volumes of phenol, phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform-isoamyl alcohol (24:1), respectively. DNA was precipitated with 2 volumes of absolute ethanol and 1/10 volume of 3M Na acetate, washed in 70% ethanol, dried and resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). The DNA concentration of samples was measured spectrophotometrically.

DNA amplifications (RAPD). Reactions were performed in a total volume of 25 µl with the following final concentrations: 10 mM Tris-HCl (pH 9), 50 mM KCl, 2 mM MgCl, 0.1% Triton, 0.1 mM for each deoxy (d) -ATP, dTTP, dCTP and dGTP, and 0.3 µM primer (a decamer oligonucleotide). For each reaction 1 U of *Taq* DNA polymerase (Promega Corporation) and 20 ng of template were used.

A thermal cycler PTC-100 (MJ Research) was used for amplifications with the following parameters: 2 min of denaturation at 94°C, followed by 45 cycles of 10 s at 94°C, 1 min at 35°C and 2 min at 72°C. The last cycle was followed by 5 min at 72°C. In each thermal cycling a negative control (water instead of template) was included to rule out amplification products due to external contamination.

Primer screening, band scoring and reproducibility. The primer choice was based on the screening of 30 RAPD decamer oligonucleotides by Operon Technologies. The primers tested were the OPA series from 1 to 20 and the OPE series from 1 to 10. The 6 primers that gave the best amplification products in terms of band resolution were selected (Table 1).

Amplified DNA was electrophoresed on 1.4% agarose gel (Fig. 2) and stained with ethidium bromide. Gels were photographed with 667 Polaroid film on a UV transilluminator.

Table 1. Primer sequences and corresponding loci used as markers

Name	Sequences	Progressive numeration of loci
OPA 11	CAATCGCCGT	1 – 7
OPA 12	TCGGCGATAG	8 – 16
OPA 16	AGCCAGCGAA	17 – 27
OPA 19	CAAACGTCGG	28 – 32
OPA 20	GTTGCGATCC	33 – 36
OPE 01	CCCAAGGTCC	37 – 43

Amplifications of 3 populations were loaded on each gel. One of the 3 populations was always reamplified and loaded in the following run. In this way each gel shared one population with that following, with 2 advantages: easier comparison of different gels during scoring and control of amplification reproducibility. For each amplified population the DNA of the first 5 individuals was pooled and reamplified simultaneously with the others, providing further reproducibility control within the same thermal cycling (Fig. 2). Because of the reduced concentration of individual DNA in pools, the scoring performed on gel photographs does not always allow detection of a given band present in a single individual within pools; however, on the gel these bands are generally visible, even if faint.

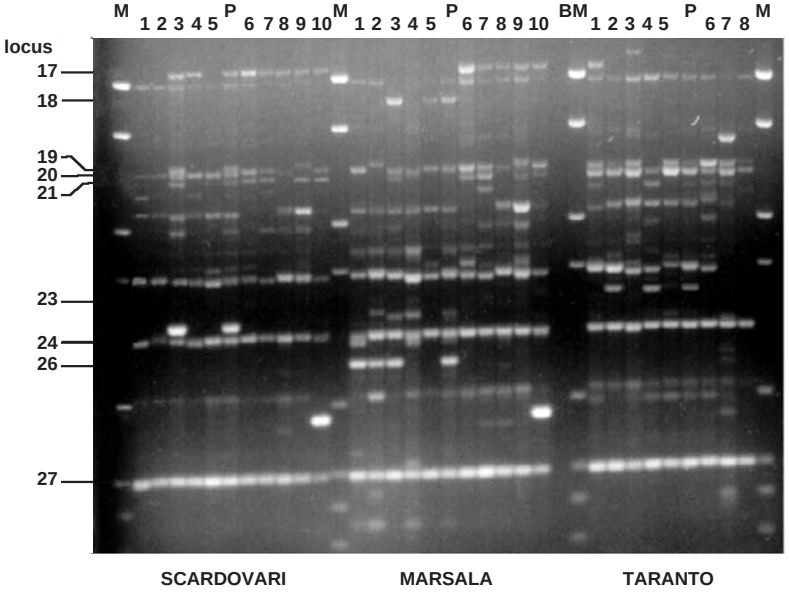

Fig. 2. *Atherina boyeri.* Examples of random amplified polymorphic DNA profiles obtained after amplification with primer OPA16, on genomic DNA from 3 populations. B: blank; M: molecular weight marker VI (Boehringer Mannheim); P: pooled DNA of individuals from 1 to 5. The locus marker bands are emphasized by the lines on the left

The selection of locus marker bands resulted from comparison of 2 scorings, independently performed by 2 people and based on band resolution and reproducibility. If scoring of a given band was ambiguous in one population, that band was excluded from the dataset of all populations.

For each primer amplification, each locus was scored in all individuals. The presence of a given band was recorded as 1 and its absence as 0. In this way all individuals were represented by a Boolean vector with as many 0 and 1 components as the locus marker bands.

Data analysis. Linkage disequilibrium (LD) between pairs of loci was estimated according to Apostol et al. (1996) to detect possible physical association of 2 markers, for instance due to nested priming sites: a mutation in one site could cause the loss of both fragments and the same information would be used twice. If an apparent LD is generated by the above artifacts, that LD is expected to occur for the same markers in all populations and is easily detected. For LD estimation we used the program RAPDLD (Black & Krafsur 1985).

Since RAPD markers are dominant, it is not possible to ascertain whether a given band is due to the amplification of a homozygote or a heterozygote genotype. However, both genotypes can be distinguished from the *absence/absence* homozygote, which does not yield any band. Therefore, the allele frequencies can only be estimated assuming the Hardy-Weinberg equilibrium for all population samples. The frequencies of the alleles were calculated by the program RAPDPLOT (Black 1993), whose output is directly readable by BIOSYS-1.7 (Swofford & Selander 1989) for subsequent elaborations. From allele frequencies, Nei's (1972, 1978) genetic distances between populations were estimated and used for cluster analysis by the unweighted pair group method with arithmetic mean (UPGMA) (Sneath & Sokal 1973). On the original data set of allele frequencies, a bootstrap analysis was performed as follows: 1000 new data sets were produced by randomly resampling the original allele frequencies and on each new data set genetic distances between populations were estimated and cluster analysis performed. Monophyletic groups occurring at the highest possible frequency were used to generate a consensus tree. For these procedures the SEQBOOT, GENDIST, NEIGHBOR and CONSENSE programs of PHYLIP (Felsenstein 1989) were used.

Multidimenisonal scaling was also performed on a genetic distances matrix.

Similarity between individuals was assessed by performing a correspondence analysis on individual band profiles (using the NTSYS program, Exeter Software). Homogeneity tests of allele frequency distribution among samples were performed through a Monte Carlo sampling process. One thousand independent randomizing processes of the data set of allele frequencies for each locus were performed and the corresponding χ^2 values estimated. The probability of encountering an χ^2 value as large as that calculated for the original matrix was determined. For this analysis the REAP software was used (McElroy et al. 1992).

Population structure was described by 2 approaches. The first was Wright's F_{st} (Wright 1951) based on allele frequencies (hereafter referred to as the genotypic approach). For this analysis we used the program RAPDFST (Black 1995), which computes F_{st} values over all loci assuming the Hardy-Weinberg equilibrium. A second approach, analysis of molecular variance (AMOVA), was based directly on the presence or absence of bands. This information, although obtained from direct DNA observation, could be considered as phenotypic since no assumption is made about genotype and allele frequencies (Haig et al. 1994). The AMOVA is a nested analysis based on a matrix of pairwise distances between all individuals. Here the distances were represented by the number of differences between individual band patterns, and the WINAMOVA 1.55 package (Excoffier et al. 1992) was used. Variance was partitioned among 3 nested levels: among individuals within populations, among populations within groups and among groups. Grouping of populations was based on the results of cluster analysis (obtained from allele frequencies). The significance of each variance component was tested with permutational procedures. Note that direct analysis of the RAPD banding pattern by AMOVA allows one to cross-check the results obtained by the genotypic approach, avoiding the untestable assumption of the Hardy-Weinberg equilibrium, which is indeed one of the drawbacks of RAPD-based population studies (Haig et al. 1994, Travis et al. 1996).

By WINAMOVA 1.55 we also calculated the Φ_{st} value, without grouping populations. This value is defined as an F_{st} analogue for molecular data (Excoffier et al. 1992) but, in our case, it is an index of phenotypic variability since it is not based on allele frequencies.

Nei's (1972) genetic distance matrix was correlated with a matrix of geographic distances between sampling locations, measured along the coast on a 1:1 000 000 scale map of Italy. Populations of freshwater lakes were excluded from the correlation test because of their artificial introduction in the environment and because of their forced isolation. The Sardinian populations were also excluded from this analysis due to the impossibility of estimating the distances along the coasts to other populations. The Messinian Strict was not considered as a barrier and the Sicilian populations were consequently included in the analysis. The Mantel test (Mantel 1967) computed by the program NTSYS was used for correlations.

RESULTS

After the scoring procedures, ambiguous bands were excluded and 43 bands (numbered from 1 to 43) were chosen as markers, with an average of 7.2 markers per primer. Table 1 shows the markers aligned with the primers that generated them. Thirty-nine (90.7%) were polymorphic across the populations analyzed (Table 2). No marker exclusive of a given population is ever found in all its individuals.

After LD analysis no pair of loci showed a significant LD in more than 3 populations (data not shown); therefore, a physical association between loci appeared unlikely and no locus was excluded from the data set.

Nei's (1972) genetic distances (Table 3) were used for constructing a dendrogram by UPGMA (Fig. 3). The dendrogram reveals a geographic clustering of populations. Even if the bootstrap values are rather low, the topology of the consensus tree was identical to that of the UPGMA.

Table 2. *Atherina boyeri*. Frequencies of 'presence' allele at the 43 loci in the 13 analyzed populations

Sample size	10	10	10	10	8	10	10	10	10	10	10	10	10
Populations	1	2	3	4	5	6	7	8	9	10	11	12	13
Locus													
1	0.000	0.000	0.000	0.000	0.209	0.105	0.000	0.163	0.000	0.000	0.683	1.000	0.225
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0.292	0.367	0.552	0.051	0.387	0.105	0.163	0.105	0.000	0.367	0.552	0.292	0.051
4	0.225	0.163	0.000	0.051	0.064	0.105	0.292	0.163	0.051	0.000	0.163	0.163	0.051
5	0.163	0.225	0.105	0.105	0.292	0.000	0.000	0.292	0.225	0.000	0.000	0.000	0.051
6	0.000	0.000	0.105	0.367	0.000	0.000	0.000	0.000	0.000	0.000	0.051	0.000	0.000
7	1.000	0.683	0.552	0.683	1.000	1.000	1.000	0.552	0.552	1.000	0.105	0.292	0.105
8	0.367	0.452	0.367	0.552	0.292	0.105	0.051	0.051	0.000	0.367	0.000	0.000	0.000
9	0.000	0.225	0.105	0.163	0.064	0.000	0.000	0.000	0.000	0.051	0.000	0.000	0.000
10	0.105	0.367	0.163	0.163	0.209	0.000	0.000	0.000	0.000	0.051	0.000	0.000	0.000
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.552	0.051	1.000	1.000
12	0.163	0.051	0.105	0.225	0.209	0.163	1.000	0.105	0.000	0.051	0.163	0.163	0.225
13	1.000	1.000	0.683	0.292	0.646	0.225	0.163	0.163	0.452	0.367	0.225	0.225	0.367
14	0.000	0.000	0.000	0.000	0.000	0.552	0.552	0.552	0.225	0.105	0.000	0.000	0.225
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.683	1.000	0.683	0.225	1.000	1.000
16	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.051	0.000	0.000
17	0.452	0.225	0.105	0.225	0.064	0.000	0.000	0.000	0.000	0.051	0.292	0.000	0.000
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.105	0.163	0.000	0.105	0.163	0.000
19	0.105	0.163	0.105	0.225	0.000	0.000	0.000	0.000	0.367	0.051	0.105	0.000	0.000
20	1.000	1.000	1.000	0.552	0.646	1.000	0.683	1.000	0.452	0.225	0.683	1.000	1.000
21	0.292	0.452	0.552	0.552	1.000	0.292	0.683	0.552	0.000	1.000	0.292	0.292	0.000
22	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.367
23	0.000	0.000	0.000	0.000	0.133	0.292	0.452	0.105	0.683	0.163	0.000	0.292	0.225
24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.163	0.000	0.000	0.000	0.000	0.000
26	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.452	0.000	0.000
27	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
28	0.225	0.367	0.051	0.000	0.000	0.225	0.367	0.225	0.292	0.105	0.000	0.225	0.000
29	0.105	0.163	0.292	0.163	0.064	0.683	1.000	0.683	0.452	0.452	0.105	0.452	0.292
30	0.000	0.000	0.000	0.000	0.000	0.051	0.000	0.000	0.000	0.000	0.683	0.000	0.000
31	0.292	0.163	0.051	0.000	0.064	0.000	0.105	0.000	0.000	0.000	1.000	1.000	0.000
32	0.000	0.000	0.000	0.292	0.064	0.552	0.225	0.225	1.000	0.105	0.051	0.367	0.292
33	0.105	0.292	0.051	0.000	0.209	0.225	0.225	0.163	0.452	0.225	0.105	0.000	0.000
34	0.000	0.000	0.225	0.292	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
35	0.552	0.552	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.683	0.683	1.000	1.000
36	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.225	0.000	0.683	0.163
37	0.292	0.163	0.292	0.367	0.209	0.000	0.000	0.000	0.000	0.452	0.163	0.000	0.000
38	0.105	0.225	0.292	0.367	0.292	1.000	0.683	0.452	0.452	0.552	0.367	1.000	0.552
39	1.000	1.000	1.000	1.000	1.000	0.683	0.225	1.000	1.000	1.000	0.163	0.292	0.105
40	0.367	0.683	0.552	0.552	1.000	1.000	1.000	0.552	1.000	0.292	1.000	0.683	0.105
41	0.292	0.292	0.105	0.163	0.064	0.163	0.105	0.225	0.225	0.000	0.000	0.000	0.225
42	0.105	0.051	0.225	0.105	0.133	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
43	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. *Atherina boyeri.* Nei's (1972) genetic distances (below diagonal) and geographic distances (kilometers) measured along the coasts in a geographic map (above diagonal). Populations from Sardinia and freshwater lakes were not considered (NC) in estimating geographic distances

	1	2	3	4	5	6	7	8	9	10	11	12	13
(1) Scardovari	–	421	757	892	999	2208	2396	NC	NC	1436	1845	NC	NC
(2) Lesina	0.016	–	336	471	578	1787	1975	NC	NC	1015	1387	NC	NC
(3) Acquatina	0.032	0.024	–	135	242	1560	1748	NC	NC	788	1019	NC	NC
(4) Ugento	0.053	0.051	0.027	–	134	1316	1504	NC	NC	544	918	NC	NC
(5) Taranto	0.053	0.039	0.029	0.036	–	1209	1397	NC	NC	436	805	NC	NC
(6) Orbetello	0.095	0.084	0.068	0.067	0.066	–	188	NC	NC	772	1111	NC	NC
(7) Massaciuccoli	0.152	0.143	0.119	0.113	0.100	0.041	–	NC	NC	960	1295	NC	NC
(8) Bolsena	0.080	0.070	0.047	0.055	0.061	0.028	0.071	–	NC	NC	NC	NC	NC
(9) Trasimeno	0.121	0.101	0.099	0.081	0.097	0.045	0.112	0.068	–	NC	NC	NC	NC
(10) Faro	0.099	0.099	0.077	0.067	0.065	0.105	0.132	0.087	0.138	–	326	NC	NC
(11) Marsala	0.221	0.216	0.197	0.218	0.206	0.217	0.246	0.203	0.252	0.179	–	NC	NC
(12) Cagliari	0.184	0.178	0.149	0.167	0.163	0.121	0.153	0.130	0.164	0.200	0.144	–	NC
(13) Alghero	0.127	0.129	0.098	0.108	0.148	0.096	0.134	0.085	0.122	0.128	0.153	0.094	–

Four main clusters can be identified: the first comprises all populations of the Adriatic-Ionian coasts and the Faro population (1, 2, 3, 4, 5 and 10), the second comprises the populations of the Tyrrhenian coasts and lakes (6, 7, 8 and 9), the third comprises the Sardinian populations (12 and 13), and the fourth consists only of the Marsala population (11) from the western Sicilian coast. When the same analysis was repeated using Nei's (1978) genetic distances, especially suitable for small size samples such as ours, the results yielded a dendrogram identical to Fig. 3.

Results of multidimensional scaling analysis appear consistent with that of cluster analysis, with the exception only of the Faro population (Fig. 4). Peculiarities of north Adriatic and Sicilian populations will be discussed below.

Results of correspondence analysis are shown in Fig. 5, where the first 2 axes explain 23.10% of the total genetic variance. This approach yielded information consistent with that of cluster analysis. Two main groups are visible: the first is composed of the individuals from Marsala (11) and the second by all the others. In this latter group, individuals are located according to their geographical position and it is possible to distinguish 3 main subgroups: the first is formed by the 2 close but well-separated populations from Sardinia (Cagliari and Alghero), the second is composed of individuals from the Tyrrhenian coasts (and lakes), and the third is composed of the individuals from the Adriatic-Ionic coasts. The position of individuals from Faro (10) is intermediate between the Tyrrhenian and the Adriatic-Ionian populations. Individuals from the Adriatic-Ionian coasts and individuals from the Tyrrhenian coasts form 2 separate and internally homogeneous groups, in which no population can be distinguished. For this reason we chose to label in Fig. 5 all Adriatic-Ionian and all Tyrrhenian individuals with the same symbols.

As a first step in analyzing the genetic structure of populations, homogeneity tests of allele frequency distribution were performed. Of 39 polymorphic loci tested, 2 were not significantly heterogeneous between populations. After Bonferroni correction for multiple testing (Sokal & Rohlf 1995), an additional locus did not show significant heterogeneity. As a consequence, for 36 loci there seems to be sufficient variation to justify an analysis of geographic structure.

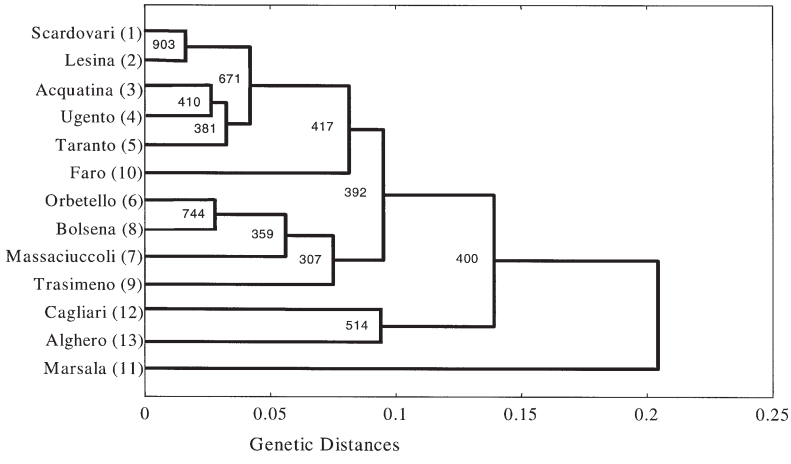

Fig. 3. *Atherina boyeri.* Unweighted pair group method with arithmetic mean cluster analysis of Nei's (1972) genetic distances on 13 populations. The bootstrap values after 1000 permutations are also reported

Table 4a reports the average F_{st} calculated for individual loci. Overall, the F_{st} values suggest a clear genetic structure of populations (F_{st} = 0.32, p < 0.0001). Performing the same analysis separately on the 3 main clusters shown in Fig. 3 (the first composed of populations 1, 2, 3, 4, 5 and 10; the second of 6, 7, 8 and 9; and the third of 12 and 13), the F_{st} values decreased as expected (0.18, 0.19 and 0.19, respectively), indicating a high degree of differentiation also within the individual clusters. The Nm values, estimated from F_{st} according to Wright (1951), are 0.5 for overall populations and 1.1, 1.0 and 1.0 for populations of the 3 main clusters. The choice to group populations on the basis of cluster analysis results is justified by the fact that the simple geographical position does not take into account environmental factors, such as oceanographic features, which, more than distance, may act as barriers to genetic flow. However, additional F_{st} analysis was also performed on 3 groups of populations, chosen on the basis of geographical proximity, excluding lake populations because of their forced isolation (Table 4b). For the 3 groups analyzed (1, 2, 3, 4 and 5; 6 and 7; 10 and 11) the F_{st} values were, respectively, 0.13, 0.12 and 0.2, while the corresponding Nm values were 1.7, 1.8 and 1.

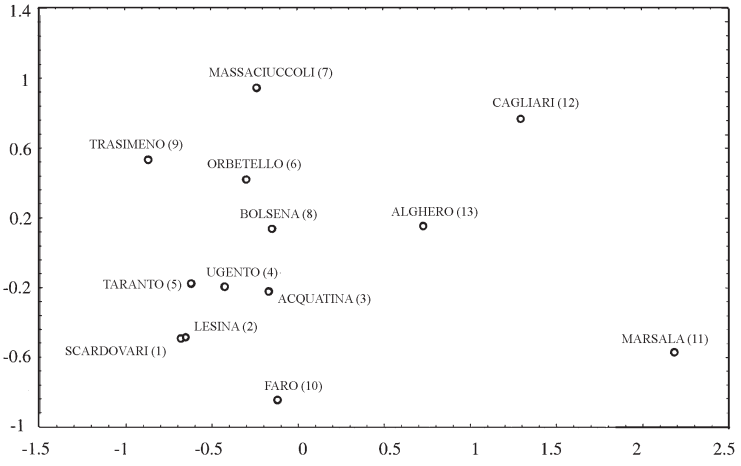

Fig. 4. *Atherina boyeri.* Multidimensional scaling analysis of Nei's (1972) genetic distances on 13 populations. Stress value = 0.07

The AMOVA results are reported in Table 5 and show that a high percentage of genetic variability (25.20%) is accounted for by the differences among population groups. However, up to 11.03% of total variability can be ascribed to the differences among populations of the same group. The same table reports the Φ_{st} values obtained by AMOVA without grouping the 13 populations. This value (F_{st} = 0.33) is very close to the Φ_{st} estimated for all the populations analyzed.

The correlation analysis, by Mantel test, between matrices of genetic and geographic distances was first performed considering all the coastal populations of continental Italy and Sicily, and yielded the following result: r = 0.42, p = 0.026, n = 36. The distribution of the population pairwise comparison, according to their genetic and geographic distance, is shown in Fig. 6: all pairs that include the Marsala population have higher than average genetic distances and appear to be outliers.

Accordingly, when the population of Marsala was excluded from the analysis, the correlation coefficient in-

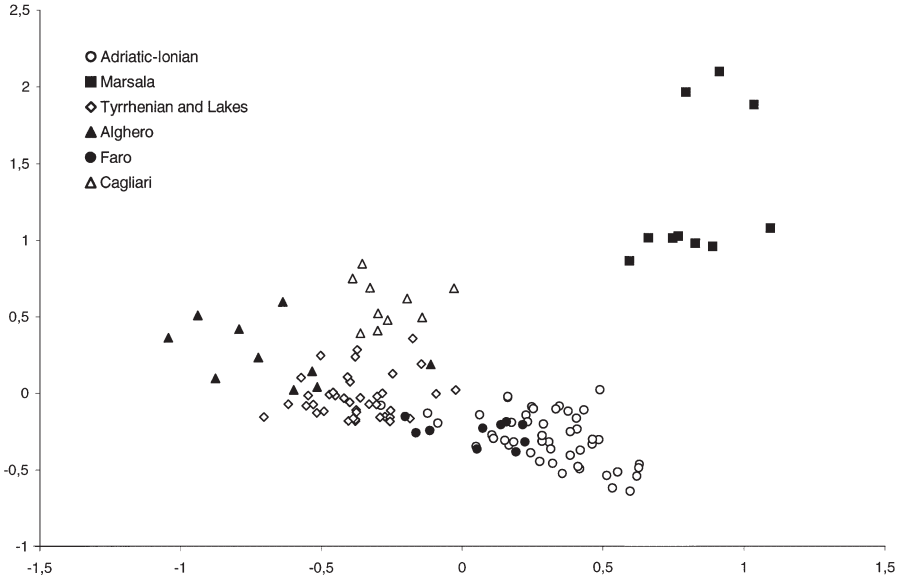

Fig. 5. *Atherina boyeri.* Genetic differentiation among individual multilocus profiles, based on correspondence analysis. All Adriatic-Ionian and all Tyrrhenian and lake individuals are labelled with the same symbols

Table 4. *Atherina boyeri.* Average Wright F_{st} calculated for individual loci. F_{st} values were estimated for different groups of populations, selected on the basis of (a) cluster analysis results or (b) their geographical proximity. The corresponding Nm values are also reported

Populations	F_{st}	p	Nm
All 13 populations	0.32	<0.0001	0.5
(a) Populations grouped on basis of cluster analysis results			
1, 2, 3, 4, 5, 10 (Adriatic-Ionian + Faro)	0.18	<0.0001	1.1
6, 7, 8, 9 (Tyrrhenian and lakes)	0.19	<0.0001	1.0
12, 13 (Sardinian)	0.19	<0.0001	1.0
(b) Populations grouped on basis of geographical proximitiy			
1, 2, 3, 4, 5 (Adriatic-Ionian)	0.13	<0.0001	1.7
6, 7 (Tyrrhenian)	0.12	<0.0001	1.8
10, 11 (Sicilian)	0.2	<0.0001	1

creased (r = 0.78, p = 0.001, n = 28), showing that genetic distances between populations of continental coasts largely reflect the geographic distances between sampling sites.

DISCUSSION

The size of samples available for this study is certainly limited. As a consequence, some possible effects of sampling errors cannot be ruled out. However, RAPD was chosen because it allows simultaneous typing of many markers. In this way, the effect of the small sample size is somewhat balanced by the comparatively large number of loci analyzed. For a general justification of this choice, see Bowcock et al. (1994).

The F_{st} values estimated under the uncertain assumption of the Hardy-Weinberg equilibrium (Table 4) and the Φ_{st} values estimated from the presence or absence of bands (Table 5) agree in indicating significant population structuring. F_{st} is 0.32 over the Italian range of the species, whereas it is close to 0.20 within each of the 3 population clusters that can be recognized in Fig. 3. In a broad-scale comparison of genetic diversity across 57 marine and 49 freshwater fish species, Ward et al. (1994) found an average G_{st} (which can be regarded as an F_{st} analogue for multiallelic loci) (Nei 1973, 1977, Takahata & Nei 1984, Hartl & Clark 1989) of 0.06 in marine species, 0.11 in anadromous species and of 0.22 in freshwater species. Thus, the values estimated in this study indicate a high level of genetic diversification compared with that of fish species with populations relatively unfettered by physical barriers. If, as seems likely, genetic differentiation among freshwater and anadromous fish samples is largely due to a limited gene flow occurring among populations (Ward et al. 1994), then *Atherina boyeri* appears to have evolved under similar evolutionary constraints. Both biological and geographic factors probably contribute to reducing the gene flow among *A. boyeri* populations. Besides the anadromic behavior, the small size of the species also probably plays a relevant role in determining the observed genetic structure. The small size is indeed often related to short-range dispersal (Munday & Jones 1998). Moreover, the planktonic larval stages develop within their natal lagoons, avoiding dispersal by coastal marine currents. Concerning the geographical constraints, most important is probably the limited number of environments suitable for reproduction; this decreases the chance for adults to migrate to different lagoons during the reproductive season.

Notwithstanding the high level of genetic structure observed, the close correlation between genetic and geographic distances measured along the coast suggests the presence of migratory movements following the coastal line. The movements are also probably related to sand smelt reproductive strategy and migratory behavior. Occasional exchanges of individuals between populations could be favored by the annual migration of adults toward the coastal sea during the cold season, establishing a sort of isolation by distance along the coasts.

Open sea tracts may have acted as geographic barriers enhancing the genetic drift. If this is the case, it is not surprising that the dendrogram in Fig. 3, according to correspondence analysis results, allows the identification of 4 main clusters, corresponding, respectively, to the Sardinian, Tyrrhenian, Adriatic-Ionian and Marsala populations. The bootstrap values are rather low but the topology of the consensus tree is identical to that of the tree obtained from the original dataset, supporting its reliability. The low bootstrap values in a population genetic study may be explained by evolutionary or demographic processes: indeed, a recent separation or the presence of genetic flow among populations may result

Table 5. *Atherina boyeri.* Nested analysis of molecular variance (AMOVA) for 128 individuals of 13 populations. The populations were grouped following the main clusters of the cluster analysis. The null distribution for each variance component is obtained by 1000 permutational runs. Φ_{st} was estimated without grouping populations for a comparison with Wright's F_{st}

Variance component	df	Variance	% of total	p-value
Among groups	3	1.58	24.94	<0.001
Among populations, within groups	9	0.81	12.78	<0.001
Within populations	115	3.96	62.28	<0.001
Φ_{st} without grouping populations			0.33	

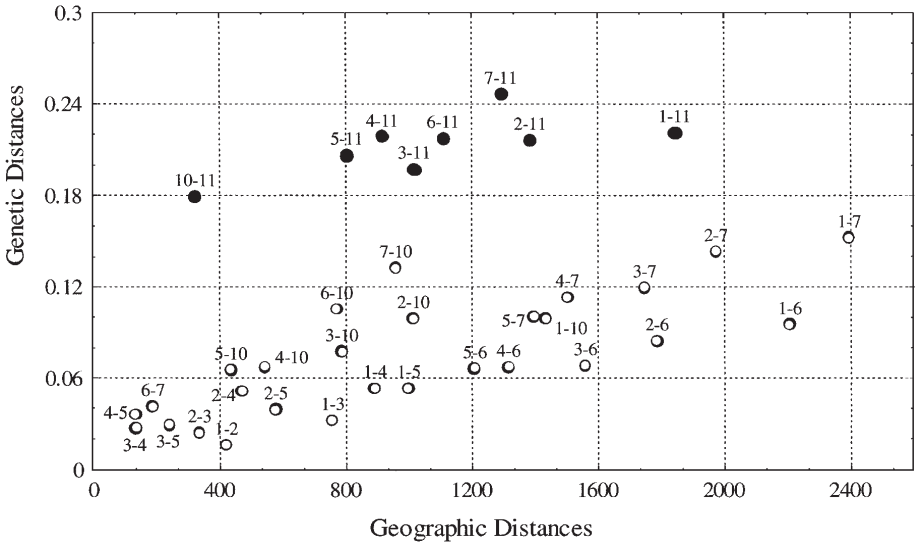

Fig. 6. *Atherina boyeri.* Pairwise population comparisons according to the genetic and geographic distances. The numbers refer to the populations described in Fig. 1. The black dots represent the pairs that include the Marsala population. The Sardinian populations are not plotted

in similar allelic frequencies at many loci with a consequent decrease of bootstrap nodal values. The main clusters will be individually discussed.

Sardinian populations show a relevant genetic distance not only from the continental ones (probably because of an isolation effect caused by the geographic barrier of the Tyrrhenian Sea) but also from each other. Cagliari (12) and Alghero (13) lagoons are, however, connected to seas with different oceanographic characteristics, and this fact could explain a certain degree of isolation and, consequently, a high genetic distance between the 2 populations.

The populations from the Bolsena (8) and Trasimeno (9) lakes cluster with those from the Tyrrhenian coast. As previously mentioned, the presence of *Atherina boyeri* in the freshwater lakes could be explained by the accidental introduction together with stocks of mugilids captured in the nearby lagoons of Tyrrhenian coasts. For the Trasimeno lake, only 1 accidental introduction is documented from about 70 years ago (Moretti et al. 1959), while for the Bolsena lake introductions occur almost yearly (P. Franzoi pers. comm.). The Bolsena lake is geographically close to the Orbetello lagoon (6) and the accidentally introduced stocks of *A. boyeri* could originate from this Tyrrhenian lagoon, thus explaining the high genetic similarity between the 2 populations.

Concerning the Adriatic-Ionian cluster, an interesting result is the low genetic distances between these populations, despite their high geographic distances. Two possible explanations can be advanced: the first is related to the oceanographic characteristics of the northern and central Adriatic, a shallow, eutrophic and enclosed sea, with several lagoons and estuarine systems. About one-third of all Mediterranean continental waters flow into the northern and central Adriatic (Bombace 1992) and consequently there is a constant significant current flow along the Italian coasts from north to south. All these factors could contribute to increasing total abundance and gene flow among *Aatherina boyeri* populations. This could explain the genetic similarity between Scardovari (1) and Lesina (2), strongly supported by the perfect overlapping of the two by multidimensional scaling, and allow the hypothesis to be advanced that there is only 1 stock in the northern and central Adriatic.

A second explanation can be found in the recent paleogeographic history of the Mediterranean regions (Thiede 1978, Pirazzoli 1996). During the last glaciation, about 18000 years ago, the sea level was lower and the Adriatic Sea was reduced to its present southern area (Fig. 1, inset). The increased marine level in the post-glacial period brought the sea to its present extension. The Adriatic coasts could have been progressively exposed to a colonization spreading from the Ionian area. In this case, the extant *Atherina boyeri* populations of the Adriatic and Ionian coasts could be descended from the same genetic pool.

Concerning the population from Marsala (11), its relatively high genetic distance from all other populations is not easily explainable. This is also the only population that completely stands out from the overall correlation trend. There are no evident geographic barriers to fish migration among the Marsala population and the others. The high genetic distance between the *Atherina boyeri* population from Marsala and other populations from Sicilian coasts has been recently confirmed by other studies based on allozyme analysis (M. Arculeo pers. comm.). The possibility that the Messina Channel could be a barrier is excluded by the good agreement of the other population from the Sicilian coast (Faro) with the trend of continental ones in the correlation analysis by Mantel test. On the other hand,

the multidimensional scaling analysis shows that the Faro population is not among the Tyrrhenian and Adriatic groups (as expected on the basis of Mantel test results) but appears misplaced toward the Marsala population, leading to the hypothesis of a partial genetic flow between the 2 populations. However, if the Marsala population is excluded from the multidimensional scaling analysis, the Faro population shifts to a completely different position between the Adriatic and Tyrrhenian groups (data not shown).

The geographic position of Marsala is a possible explanation for its high genetic distance from the other population analyzed: its closeness to Tunisian coasts and the currents entering the Mediterranean Sea from the Atlantic Ocean (Astraldi & Gasparini 1994) could indeed favor immigration of individuals from genetically distant populations. To test this hypothesis it would have been necessary to obtain *Atherina boyeri* samples from North African coast, but it was not possible in this study. This hypothesis, however, is strongly supported by a recent study (Maltagliati 1999) based on allozyme markers on another typical species of Mediterranean lagoons, *Aphanius fasciatus*, in which the Marsala population was found genetically more similar to Tunisian populations than to the Italian ones.

In conclusion, the analyzed populations of *Atherina boyeri* are highly structured, leading to the hypothesis that the limited reproduction environments and the anadromic behavior of this species could enhance the effects of the genetic drift. However, the *A. boyeri* populations are not completely isolated from each other and there is some genetic continuity along continental coasts. Although adults do not specifically live at their natal site, the partial isolation observed is probably due to the limited range of movements of this small species. The result of this situation is an isolation by distance pattern. Also, these results are in agreement with the already cited findings on *Aphanius fasciatus*.

It would be interesting to compare *Atherina boyeri* with others species typical of lagoon environments. For example, some mugilids, such as *A. boyeri*, spend part of their life cycle in the coastal lagoons but their larger size allows a wider range of movements. In this case a lower population differentiation could be expected and it would be interesting to verify whether the phylogeographic pattern observed for *A. boyeri* is preserved. On a larger geographical scale (in comparison with the present study), a significant genetic structure was observed among *Mugil cephalus* populations (Crosetti et al. 1994, Rossi et al. 1998, Rocha-Olivares et al. 2000) but no molecular studies on Italian Mugilidae populations are available. A multispecies approach would help to understand the role of oceanographic features or paleoclimatic events in determining phylogeographic patterns (Bernatchez et al. 1998).

Acknowledgements. Many thanks are due to G. Alessio, G. Belmonte, P. Franzoi, M. Lenzi, A. Mazzola, M. Natali, E. Rossetti, G. Scardella, P. Villani, A. Cau, G. Cavallaro and M. Pastore for their kind help in the sample collection; to G. Bertelli for her technical support during the experiments; to M. Chicca and G. Grandi for their valuable comments on the previous version of the manuscript; and to I. Dupanloup for her suggestions and for her kind help. The project was supported by a 40% grant of MURST (Italian Ministry of the University and Scientific and Technological Research) assigned in 1995.

LITERATURE CITED

Apostol BL, Black WC 4th, Reiter P, Miller BR (1996) Population genetics with RAPD-PCR markers: the breeding structure of *Aedes aegypti* in Puerto Rico. Heredity 76:325–334

Astraldi M, Gasparini GP (1994) The seasonal characteristics of the circulation in the Tyrrhenian Sea. In: La Violette PE (ed) Coastal and estuarine studies, book 46. American Geophysical Union, Washington

Bamber RN, Henderson PA (1985) Morphological variation in British Atherinids, and the status of *Atherina presbyter* Cuvier (Pisces: Atherinidae). Biol J Linn Soc 25:61–76

Bernatchez L, Wilson CC (1998) Comparative phylogeography of nearctic and palearctic fishes. Mol Ecol 7:431–452

Black WC 4th (1993) RAPDPLOT-FORTRAN program for analysis of genetic relationships among individuals using RAPD-PCR markers. Colorado State University, Fort Collins

Black WC 4th (1995) RAPDFST 3.0. A FORTRAN program to estimate F_{st} and effective migration rates among subpopulations using RAPD-PCR files. Colorado State University, Fort Collins

Black WC 4th, Krafsur ES (1985) A FORTRAN program for the calculation and analysis of two-locus linkage disequilibrium coefficients. Theor Appl Genet 70:491–496

Bombace G (1992) Fisheries of the Adriatic Sea. In: Colombo G, Ferrari I, Ceccherelli VU, Rossi R (eds) Marine eutrophication and population dynamics. Proc 25th Eur Mar Biol Symp. Olsen & Olsen, Fredensborg, p 379–389

Boscolo L (1970) Osservazioni sulla biologia e sulla pesca dell' *Atherina boyeri* Risso 1810 (Osteichthyes, Atherinidae) vivente nelle acque dell'Alto Adriatico. Boll Pesc Piscic Idrobiol 25:61–79 (in Italian)

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457

Creech S (1991) An electrophoretic investigation of populations of *Atherina boyeri* Risso, 1810 and *A. presbyter* Cuvier, 1829 (Teleostei: Atherinidae): genetic evidence in support of the two species. J Fish Biol 39:807–816

Crosetti D, Nelson WS, Avise JC (1994) Pronounced genetic structure of mitochondrial DNA among populations of the circumglobally distributed grey mullet (*Mugil cephalus* Linnaeus). J Fish Biol 44:47–58

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Felsenstein J (1989) PHYLIP – phylogeny inference package, version 3.5. Department of Genetics SK, University of Washington, Seattle

Froglia C, Orel G (1979) Considerazioni sulla pesca a strascico nella fascia costiera delle tre miglia in Adriatico. Atti Soc Toscana Sci Nat Mem Ser B 86:17–25 (in Italian)

Haig SM, Rhymer JM, Heckel DG (1994) Population differentiation in randomly amplified polymorphic DNA of red-cockaded woodpeckers *Picoides borealis.* Mol Ecol 3: 581–595

Hartl DL, Clark AG (1989) Principles of population genetics, 2nd edn. Sinauer Associates, Sunderland

Jorné-Safriel O, Shaw E (1966) The development of schooling in the atherinid fish, *Atherina mochon* (Cuvier). PSZN 35: 76–88

Kiener A, Spillmann CJ (1969) Contribution à l'étude systématique et écologique des Athérines des côtes françaises. Mem Mus Natl Hist Nat Ser A Zool 60:1–74 (in French)

Kiener A, Spillmann CJ (1972) Note complémentaire à l'étude systématique et écologique d' *Atherina boyeri* Risso (Poissons, Cyprinidae) dans sa zone de dispersion actuelle. Bull Mus Natl Hist Nat 3 41:563–580 (in French)

Maltagliati F (1999) Genetic divergence in natural populations of the Mediterranean brackish-water killifish *Aphanius fasciatus.* Mar Ecol Prog Ser 179:155–162

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Marfin JP (1981) Biologie de l'Atherine *Atherina boyeri* Risso, 1810 (poisson-téléostéen) dans trois milieux saumatres du Roussillon. Thèse de doctorat, Accademie de Montpellier (in French)

Marfin JP (1982) Les problèmes liés au polymorphisme de l'éspèce *Atherina boyeri* Risso, 1810. Cybium 6:19–26 (in French)

McElroy D, Moran P, Bermingham E, Kornfield I (1992). REAP: an integrated environment for the manipulation and phylogenic analysis of restriction data. J Hered 83:157–158

Mistri M, Colombo G (1988) Morphometric variability in sandsmelt, *Atherina boyeri* Risso 1810, populations from different Italian sites. Boll Zool 3:129–133

Moretti G, Giannotti FS, Giganti A (1959) Il Latterino (*Atherina mochon* Cuv.) nel Trasimeno. Riv Biol 51:1–38 (in Italian)

Munday PL, Jones GP (1998) The ecological implications of small body size among coral reef fishes. Oceanogr Mar Biol 36:373–441

Nei M (1972) Genetic distances between populations. Am Nat 106:283–292

Nei M (1973) Analysis of gene diversity in subdivided populations. P Natl Acad Sci USA 70:3321–3323

Nei M (1977) F-statistics and analysis of gene diversity in subdivided populations. Ann Hum Genet 41:225–233

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Pirazzoli PA (1996) Sea level changes, the last 20,000 years. John Wiley & Sons, Chichester

Rocha-Olivares A, Garber NM, Stuck KC (2000) High genetic diversity, large inter-oceanic divergence and historical demography of the striped mullet. J Fish Biol 57:1134–1149

Rossi AR, Capula M, Crosetti D, Campton DE, Sola L (1998) Allozyme variation in global populations of striped mullet *Mugil cephalus* (Pisces, Mugilidae). Mar Biol 131:203–212

Roy A, Frascaria N, MacKay J, Bousquet J (1992) Segregating random amplified polymorphic DNAs (RAPDs) in *Betula alleghaniensis*. Theor Appl Genet 85:173–180

Sacchi CF (1979) The coastal lagoon of Italy. In: Jefferies RL, Davy AJ (eds) Ecological processes in coastal environments. Blackwell Scientific Publications, Oxford, p 593–601

Sneath PHA, Sokal RR (1973) Numerical taxonomy. WH Freeman and Co., San Francisco

Sokal RR, Rohlf FJ (1995) Biometry, 3d edn. WH Freeman and Co., New York

Stabile J, Waldman JR, Parauka F, Wirgin I (1996) Stock structure and home fidelity in Gulf of Mexico sturgeon (*A. oxyrinchus desotoi*) based on restriction fragment length polymorphism and sequence analyses of mitochondrial DNA. Genetics 144:767–775

Swofford DL, Selander RB (1989) Biosys-1.7: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. Illinois Natural History Survey, Champaign

Takahata N, Nei M (1984) F_{st} and G_{st} statistics in finite island model. Genetics 107:501–504

Thiede J (1978) A glacial Mediterranean. Nature 276: 680–683

Trabelsi M, Kartas F, Quignard JP (1994) Comparaison du régime alimentaire d'une population marine et d'une population lagunaire d' *Atherina boyeri* des côtes tunisiennes. Vie Milieu 44:117–123 (in French)

Trabelsi M, Focant B, Tomasini JA, Quignard JP (1995) Originalité du complexe *Atherina boyeri* Risso, 1810 (Poissons, Téléostéens, Atherinides) de la reserve naturelles des Îles Lavezzi (Corse). Trav Sci Parc Nat Reg Reserv Nat Corse 56:83–105 (in French)

Travis SE, Maschinski J, Keim P (1996) An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically endangered plant, using AFLP markers. Mol Ecol 5:735–745

Waldman JR, Wirgin I (1994) Use of DNA analyses in the management of natural fish populations. In: Garte SJ (ed) Molecular environmental biology. Lewis Publishers, Boca Raton, FL

Waples RS (1987) A multispecies approach to the analysis of gene flow in marine shore fishes. Evolution 41:385–400

Ward RD, Woodwark M, Skibinski DOF (1994) A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. J Fish Biol 44:213–232

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18: 7213–7218

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Wolff K, Peters-Van Rijn J (1993) Rapid detection of genetic variability in chrysanthemum (*Dendranthema grandiflora* Tzvelev) using random primers. Heredity 71:335–341

Wright S (1951) The genetical structure of populations. Ann Eugen 15:323–354

Editorial responsibility: Simon Archer (Contributing Editor), Guildford, United Kingdom

Submitted: May 1, 2001; Accepted: August 31, 2001
Proofs received from author(s): March 5, 2002