

# Mitochondrial variability of sand smelt *Atherina boyeri* populations from north Mediterranean coastal lagoons

L. Astolfi<sup>1</sup>, I. Dupanloup<sup>2</sup>, R. Rossi<sup>1</sup>, P. M. Bisol<sup>3</sup>, E. Faure<sup>4</sup>, L. Congiu<sup>1,\*</sup>

<sup>1</sup>Department of Biology, University of Ferrara, Via Borsari 46, 44100 Ferrara, Italy

<sup>2</sup>Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland

<sup>3</sup>Department of Biology, University of Padua, Via Colombo 3, 35121 Padua, Italy

<sup>4</sup>EA Biodiversité, Université de Provence, 13331 Marseille cedex 3, France

**ABSTRACT:** The role of transition environments, such as lagoons or deltas, in shaping the geographical pattern of genetic variability may be investigated through the analysis of population structure in the species living in these habitats. Several fish species are strictly bound to deltas or coastal lagoons for at least part of their life cycle. This paper reports the analysis of mitochondrial DNA variability in several populations of *Atherina boyeri* (Risso, 1810) sampled along the north Mediterranean coasts from the Atlantic Ocean to the Black Sea. The results show a high genetic structure among populations, probably due to the deep fragmentation of the lagoon environments in which this species spends most of its life cycle. On the other hand, a clear phylogeographic pattern and a significant positive correlation between genetic and geographical distances support the hypothesis of gene flow among populations, probably due to an imprecise anadromic behaviour from the lagoons to the close coastal sea, which allows for an exchange of individuals between contiguous populations. Our results are in good agreement with those previously obtained on the same species by RAPD techniques. Our analysis among different *Atherina* genera are also compatible with the most recent classification of these species based on morphological and genetic characters.

**KEY WORDS:** D-loop · *Atherina boyeri* · Coastal lagoons · Population genetics · Phylogeography

— Resale or republication not permitted without written consent of the publisher —

## INTRODUCTION

The role of physical and geographical factors in shaping the distribution of genetic variability is especially relevant for species closely bound to discontinuous habitats such as coastal lagoons (Sinclair & Iles 1988, 1989). Besides geographical fragmentation, these environments are characterised by relevant variability of chemical and physical parameters: the species inhabiting them are exposed to strong evolutionary pressures (Bamber & Henderson 1988). All these factors may result in a high degree of population structure, and the genetic analysis of lagoon species may help to describe the phylogeographic effects of these environments.

Many fish species are closely linked to lagoon environments for at least part of their life cycle. Their subdivision in genetically distinct populations is the result of 2 opposite influences: the tendency to isolation from habitat fragmentation and the tendency to homogeneity due to migrations, which are related to the dispersal ability of the species (Ward et al. 1994). A relevant role may also be played by oceanographic features, such as physical-chemical parameters or currents, which may ease or obstruct migratory movements, thus contributing to shape phylogeographic patterns.

The sand smelt *Atherina boyeri* is a very common species in the Mediterranean Sea. Individuals of this species spend a relevant part of their life cycle within brackish coastal environments (Bamber & Henderson

1988) where they represent a relevant resource for predators and, mainly in the Adriatic Sea, for fisheries. Adults spawn in inshore shallow waters, and the eggs are attached to seaweed or rocks to avoid dispersion. During larval and juvenile stages the fishes reside within lagoons. Only adults, during cold months, leave the inshore waters and move toward the sea until the next reproductive season (Henderson & Bamber 1987, Henderson et al. 1988, Mistri & Colombo 1988). Due to the low dispersal capability of these small fishes, the percentage of animals re-entering the native lagoons is probably very high, as is also confirmed by morphological studies (Berrebi & Britton-Davidian 1980); this may have the same demographic effect of a homing behaviour, limiting gene flow among sites. A clear genetic structure among *A. boyeri* populations from different lagoons was indeed recently observed along the Italian coasts using random amplified polymorphic DNA (RAPD) markers (Congiu et al. 2002). However, a significant correlation between genetic and geographical distances among the sampled populations was found, thus suggesting the presence of some migratory movements along the coastal line. Occasional exchanges of individuals among populations could be favoured by the annual migration of adults toward the coastal sea during the cold season, establishing a pattern of isolation by distance along the coasts.

This observation contrasts with the results obtained for another atherinid species (*Odontesthes argentinensis*) inhabiting the Atlantic coasts of Brazil and Argentina, for which no significant correlation between genetic and geographical distances was observed, when estuarine populations were compared (Beheregaray & Sunnucks 2001). On the other hand, marine populations, which are considered to have recently colonised the estuaries, showed a clear geographical structure of genetic variants (Beheregaray & Sunnucks 2001). This pattern was considered the product of the rapid adaptive divergence and consequent reproductive isolation of the estuarine populations in the face of potential gene flow along the coastal line (Beheregaray & Sunnucks 2001).

The presence of an isolation-by-distance pattern in the Italian populations of *Atherina boyeri* (Congiu et al. 2002) seems to indicate that the homogenising effect of gene flow counteracts the reproductive isolation as a consequence of adaptive divergence in this region. However, the reduced geographical scale of this study (Congiu et al. 2002), as well as the large size of the sampled populations, may have affected our conclusions.

To investigate the genetic relationships among *Atherina boyeri* populations at a larger geographical scale, we sampled 12 lagoons along the north Mediterranean coasts from the Atlantic Ocean to the Black Sea. We also typed a different genetic marker in these samples,

i.e. sequences of the mtDNA control region, to determine whether the results previously obtained with RAPD markers could find additional support (Congiu et al. 2002). The application of RAPD analysis to population-genetic studies has often been criticised because of limited data reproducibility and because of the 'dominant' nature of its markers (MacPherson et al. 1993, Micheli et al. 1994, Rieseberg 1996). The comparison of the results obtained by mitochondrial markers with those previously yielded by RAPD analysis may provide interesting information about the usefulness, reliability and consistency of the 2 approaches.

One relevant result obtained by RAPD analyses of Italian sand smelt populations was the high genetic distance which separates the Marsala population (southern Sicily) from all the other populations (Congiu et al. 2002). A possible explanation for this finding could be the presence of migratory movement across the Sicily channel from Tunisia toward southern Sicilian coasts. This hypothesis is supported by the results observed in another lagoon fish species, *Aphanius fasciatus*, in which the population from the Marsala lagoon was found to be more similar to Tunisian populations than to Italian ones (Maltagliati 1999). Since these 2 species exhibit low dispersal capabilities, these migrations should be prompted by marine currents. To test the above hypothesis, a few *Atherina boyeri* individuals from a Tunisian lagoon were included in some analyses. If the genetic similarity between the north African and the Marsala populations found in *A. fasciatus* is observed also in the sand smelt, this may confirm the usefulness of the interspecific approach. Indeed, the same phylogeographic pattern observed in different species should represent a relevant insight into the factors that shape the distribution of genetic diversity across the Mediterranean Sea.

The taxonomic relationships within genus *Atherina* are controversial due to the high morphological variability which characterises these fishes (Kiener & Spillmann 1969, 1972, Marfin 1982, Bamber & Henderson 1988, Creech 1991, 1992). Phenotypic polymorphism is especially high in *A. boyeri*. Before 1969 nearly 20 distinct species were described in the literature until Kiener & Spillman (1969, 1972) considered them all synonymous with *A. boyeri*. The other 2 *Atherina* spp. known to be present in the Mediterranean Sea are *A. hepsetus* and *A. presbyter*. Due to the high morphological variability of the species belonging to the genus *Atherina*, a wrong classification of the samples could not be excluded. For this reason, in order to verify the monophyly of the *A. boyeri* samples studied here, some individuals from different populations of *A. hepsetus* and *A. presbyter* were also analysed. Besides the above congeneric species, well documented and formally recognised, 2 additional species, until now considered morphological vari-

ants of *A. boyeri*, have recently been proposed (Klossa-Kilia et al. 2002, Trabelsi et al. 2002a, 2002b). According to these studies, based on both morphological and molecular data, the species known as *A. boyeri* should be considered as a complex of 3 different species (2 marine and 1 living in lagoons), which should be renamed according to their distribution and to the presence of dark spots along the lateral line. The 1st species (*A. boyeri*) would be composed of non-punctuated (non-spotted) marine individuals, the 2nd species (*A. punctata*) of punctuated marine individuals and the 3rd species (*A. lagunae*) of lagoon individuals.

This paper focuses on the comparison of lagoon samples which, if the above revision of the genus is accepted, should be considered *Atherina lagunae*. Until then, though, we should consider these samples as *A. boyeri* lagoon populations. The 2 newly proposed marine species, included in our dataset with only 1 specimen each, kindly provided by Dr. M. Trabelsi, University of Tunis, will be indicated as non-punctuated and punctuated forms. The analysis of these 2 samples and of the non-*A. boyeri* ones should not be regarded as an insight into the phylogenetic relationships among the *Atherina* spp.; for this purpose, a more extended sample collection should be performed. Our goal was to verify the presence of a clear genetic differentiation among our lagoon samples and the other species (either recognised or putative) through the analysis of the mitochondrial control region. This could, in turn, support the reliability of this marker as a tool for taxonomical investigations within genus *Atherina*.

## MATERIALS AND METHODS

**Sample collection.** Twelve sand smelt populations were sampled in 2 Atlantic, 9 Mediterranean and 1 Black Sea lagoons (Fig. 1 shows 12 sites). The 1st Atlantic population was sampled in the Tago River estuary in Portugal and the 2nd one in the Zonar lagoon, connected to the Guadalquivir River in Spain. The other lagoons analysed were Or and Fucecchio along the continental Tyrrhenian coast; Cagliari on Sardinia island; Marsala and Faro on Sicily island; Acquatina, Scardovari and Spalato along the Adriatic Sea coast; and the Danube Delta in the Romanian Black Sea coast. For each of the above populations, 10 individuals were sampled.

The hypothesis of genetic contact between the coasts of Tunisia and Sicily was tested by typing 6 specimens from a Tunis lagoon. Unfortunately, it was not possible to obtain more individuals from this locality; therefore, the Tunis population was excluded from some data analyses because of the low sample size. Phylogenetic analyses of

the sampled individuals were completed by the inclusion of specimens of the punctuated and non-punctuated forms of *Atherina boyeri* sampled in Corsica, as well as several specimens of 3 populations of *A. hepsetus* sp. (Cagliari, Spalato and Lampedusa) and 2 populations of *A. presbyter* sp. (Arcachon and Swansea) (Fig. 1).

Within a few hours of capture, the fishes were cut into 3 to 4 pieces, separately preserved in ethanol and stored at  $-20^{\circ}\text{C}$  until DNA extraction.

**DNA extraction.** Approximately 200 mg of head tissue, cleaned from skin and scales, were frozen in liquid nitrogen, ground and resuspended in an extraction buffer (100 mM Tris HCl, 1% SDS, 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 3 M Na acetate, washed in 70% ethanol, dried and resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). Tunisian samples and the punctuated and non-punctuated forms were obtained as already purified DNA.

**DNA amplifications.** Approximately 400 bp of the mitochondrial control region, including the first hypervariable region (HVR1), were amplified using the following primers: Lpro 2 (AACTCCCACCACTAACTCCCAAAGC) and Hdl 1 (CCTGAAGTAGGAACCA-GATGCCAG). Reactions were performed in a total volume of 25  $\mu\text{l}$  with the following final concentrations: 10 mM Tris (pH 9), 50 mM KCl, 1.5 mM  $\text{MgCl}_2$ , 200  $\mu\text{M}$  for each dNTP, and 10 pmoles for each primer. For each reaction, 1 U of *Taq* DNA polymerase (Amersham Biosciences) and about 100 ng of template were used.

A thermal cycler Gene Amp PCR System 9700 (Perkin Elmer) was used for amplifications with the following parameters: 2 min of denaturation at  $94^{\circ}\text{C}$ , followed by 25 cycles of 30 s at  $94^{\circ}\text{C}$ , 45 s at  $55^{\circ}\text{C}$  and 45 s at  $72^{\circ}\text{C}$ . The last cycle was followed by 5 min at  $72^{\circ}\text{C}$ . In each thermal cycling a negative control (water instead of template) was included to rule out amplification products due to external contamination. The amplified fragments were purified by the ExoSap-IT kit (Amersham Biosciences). The reaction products were directly sequenced for both forward and reverse strands with the amplification primers using an automated DNA sequencer ABI PRISM 3730XL (Applied Biosystems) (sequencing service: BMR Bio Molecular Research, CRIBI, Padova, Italy). Multiple sequence alignment was performed using the Clustal X program (Thompson et al. 1997) and visually verified on a chromatogram. (GenBank accession number: DQ102774 to DQ102851)

**Data analyses.** Genetic relationships among the haplotypes observed in the 12 lagoons were sum-



Fig. 1. Map of sampling locations. ●: *Atherina boyeri* samples, ☆: *A. presbyter* samples, □: *A. hepsetus* samples, and O: specimens of the newly proposed species, here indicated as (1) punctuated and (2) non-punctuated forms, respectively

marised by a neighbour-joining tree (Saitou & Nei 1987) based on Kimura's 2 parameter distances (Kimura 1980) using the PHYLIP package (Felsenstein 1993). To avoid sample classification errors and to test the usefulness of the mitochondrial control region in solving taxonomical issues of this systematic group (Trabelsi et al. 2002a,b), we included in this tree the haplotypes found in punctuated and non-punctuated specimens, originally thought to be *Atherina boyeri* (Trabelsi et al. 2002a,b), as well as those observed in several individuals from *A. hepsetus* and *A. presbyter*. Finally, the 6 individuals from the Tunisian population were also analysed to test the presence of a gene flow from African coasts to Sicily (Congiu et al. 2002). Sequences of 4 *Atherinomorpha* spp. were used as outgroups: 1 (*Exocoetus volitans*) belonging to the family Exocoetidae and 3 (*Atherinosoma elongate*, *Leptatherina presbyteroides* and *Leptatherina wallacei*) belonging to the family Atherinidae (GenBank accession numbers NC003184, AY313135, AY325980 and AY326789 respectively). Bootstrap resampling of the original haplotypes (1000 replicates) was per-

formed to evaluate the robustness of the tree topology (Felsenstein 1985).

The reliability of the neighbour-joining method in representing the genetic relationships among haplotypes was checked by a maximum-likelihood method, also implemented in the PHYLIP package (Felsenstein 1993). To infer haplotype relationships, statistical parsimony (Templeton et al. 1992) was also applied using TCS software (Clement et al. 2000).

For 11 lagoons (Tunis excluded), we computed some measures of within-population variability (gene diversity,  $H$ ) and mean number of pairwise differences ( $\pi$ ) using Arlequin software v.2.0 (Schneider et al. 2000). We attempted to gain some insight into the past demographic history of the Adriatic–Ionian region by calculating the mismatch distribution on the pooled populations. Irregular and multimodal distribution of pairwise differences between sequences is indeed expected in stationary or shrinking populations, whereas a smooth, unimodal shape is typical of expanding populations (Rogers & Harpending 1992, Rogers et al. 1996, Excoffier & Schneider 1999).

Departures from mutation-drift or mutation-selection equilibrium were tested on pooled Adriatic and Faro populations using Tajima's  $D$  (Tajima 1989) and Fu's  $F_s$  (Fu 1997) tests. These 2 statistics are sensitive to demographic processes:  $D$  and  $F_s$  show negative values when the population expands and positive values when it shrinks. The significance of  $D$  and  $F_s$  was tested by randomisation as implemented in the Arlequin package (Schneider et al. 2000). For both of these 2 indices and for mismatch distribution analysis, the criterion adopted to pool the populations was the non-significance of pairwise  $\phi_{ST}$ . These analyses were not individually performed on each population because of the small sample size.

Genetic divergence between populations was analysed using several different methods. The hierarchical components of mtDNA variation were computed under the analysis of molecular variance (AMOVA) framework (Excoffier et al. 1992, Schneider et al. 2000). Pairwise comparisons between populations were performed by computing  $\phi_{ST}$  distances based on Kimura's 2 parameter distances among haplotypes. For both the AMOVA analyses and the computations of  $\phi_{ST}$  distances between lagoons, the heterogeneity of mutation rates across mtDNA sites was taken into account by assuming a gamma distribution of rates (Hasegawa et al. 1993) with parameter  $\alpha = 0.12$ . The value for  $\alpha$  was estimated by a maximum-likelihood method with PAML software, v.3.11 (Yang & Kumar 1996), from the total set of haplotypes observed in the samples.

A multidimensional scaling was reconstructed using the matrix of  $\phi_{ST}$  distances between populations. Mantel tests were performed to measure the correlation between  $\phi_{ST}$  distances and geographic distance matrices using the Arlequin package (Schneider et al. 2000). The significance of the correlation was assessed by performing 1000 permutations of original distance matrices. Geographic distances were computed along the coast on a 1:20 000 000 scale map of Europe.

We used the SAMOVA approach (Dupanloup et al. 2002) to define groups of lagoon populations which are geographically and genetically homogeneous and maximally differentiated

from each other. The method is based on a simulated annealing procedure which aims at maximising the proportion of total genetic variance due to differences among groups of populations.

## RESULTS

### Differences among mtDNA haplotypes

The numbered list of haplotypes is reported in Table 1. The neighbour-joining tree representing the nucleotide variations among the control region

Table 1. *Atherina* spp. Numbered haplotypes with the corresponding sampling localities. The number of individuals found with each haplotype is given in brackets

Haplotype	Sampling location	Haplotype	Sampling location
01	Acquatina (3)	40	Scardovari (1)
02	Danube (1)	41	Fucecchio (1)
03	Tago (3)	42	Fucecchio (1)
04	Danube (1)	43	Or (1)
05	Marsala (4)	44	Scardovari (1)
06	Acquatina (1)	45	Scardovari (1)
07	Acquatina (1)	46	Acquatina (2)
08	Acquatina (2)	47	Fucecchio (1)
09	Danube (2)	48	Or (1)
10	Danube (1)	49	Spalato (1)
11	Tago (7)	50	Danube (1)
12	Danube (1)	51	Danube (1)
13	Spalato (1)	52	Danube (1)
14	Spalato (5)	53	Faro (2), Scardovari (1)
15	Spalato (1)	54	Spalato (1)
16	Zonar (9)	55	Spalato (1)
17	Or (1)	56	Zonar (1)
18	Cagliari (1)	57	Danube (1)
19	Cagliari (6)	58	Faro (1)
20	Fucecchio (3)	59	Faro (1)
21	Scardovari (1)	60	Faro (1)
22	Scardovari (1)	61	Faro (1)
23	Cagliari (2)	62	Faro (1)
24	Marsala (4)	63	Faro (2)
25	Or (2)	64	Faro (1)
26	Or (1)	65	Tunis (4)
27	Scardovari (1)	66	Tunis (1)
28	Fucecchio (2)	67	Tunis (1)
29	Marsala (1)	68	<i>A. non punctata</i> (1)
30	Or (1)	69	<i>A. punctata</i> (1)
31	Scardovari (1)	70	<i>A. hepsetus</i> Spalato (2)
32	Fucecchio (1)	71	<i>A. hepsetus</i> Spalato (1)
33	Fucecchio (1)	72	<i>A. hepsetus</i> Cagliari (2)
34	Marsala (1)	73	<i>A. hepsetus</i> Cagliari (1)
35	Or (1)	74	<i>A. hepsetus</i> Lampedusa (2)
36	Acquatina (1), Scardovari (2)	75	<i>A. presbyter</i> Swansea (2)
37	Cagliari (1)	76	<i>A. presbyter</i> Swansea (1)
38	Or (1)	77	<i>A. presbyter</i> Arcachon (1)
39	Or (1)	78	<i>A. presbyter</i> Arcachon (1)

sequences observed in 12 *Atherina boyeri*, 3 *A. hepsetus* and 2 *A. presbyter* samples and 2 punctuated and non-punctuated specimens (cf. *A. boyeri*) is reported in Fig. 2. These results are in complete agreement with those obtained by maximum likelihood (data not shown). The total number of haplotypes (78 in 131 individuals) are grouped in 4 clusters, strongly supported by bootstrap values. Three of these clusters correspond to the 3 species known to be present in the Mediterranean Sea (*A. presbyter*, *A. hepsetus* and *A. boyeri*); the 4th one includes the 2 sequences observed in the punctuated and non-punctuated specimens, provisionally supporting the hypothesis of a division of *A. boyeri* into distinct species (Trabelsi et al. 2002a,b).

The haplotypes found in the 12 sampled lagoons are distributed in the tree according to geographical location, as also confirmed by the network shown in Fig. 3. However, the absence of shared haplotypes among lagoons, except 2 (Acquatina-Scardovari and Faro-Scardovari), can be ascribed to non-extensive gene flow among populations.

The clustering of Marsala haplotypes with the Tunisian sequences is compatible with 2 non-mutually exclusive hypotheses: (1) the Mediterranean Sea between the Tunisian and Sicilian coasts does not constitute a strong barrier to migration and/or (2) the common origin of Tunisian and Sicilian populations is relatively recent. Two specimens sampled in Cagliari fall within the Marsala-Tunisian cluster, suggesting genetic contact among Tunisia, Sicily and Sardinia.

The Danube haplotypes constitute a distinct genetic group, well supported by bootstrap resampling (1000/1000), in agreement with the peripheral position of this sampling site in our study area. Only haplotype 15, from the Spalato lagoon, falls within this cluster.

A high bootstrap value (976/1000) strongly supports monophyly of the remaining haplotypes, compatible with

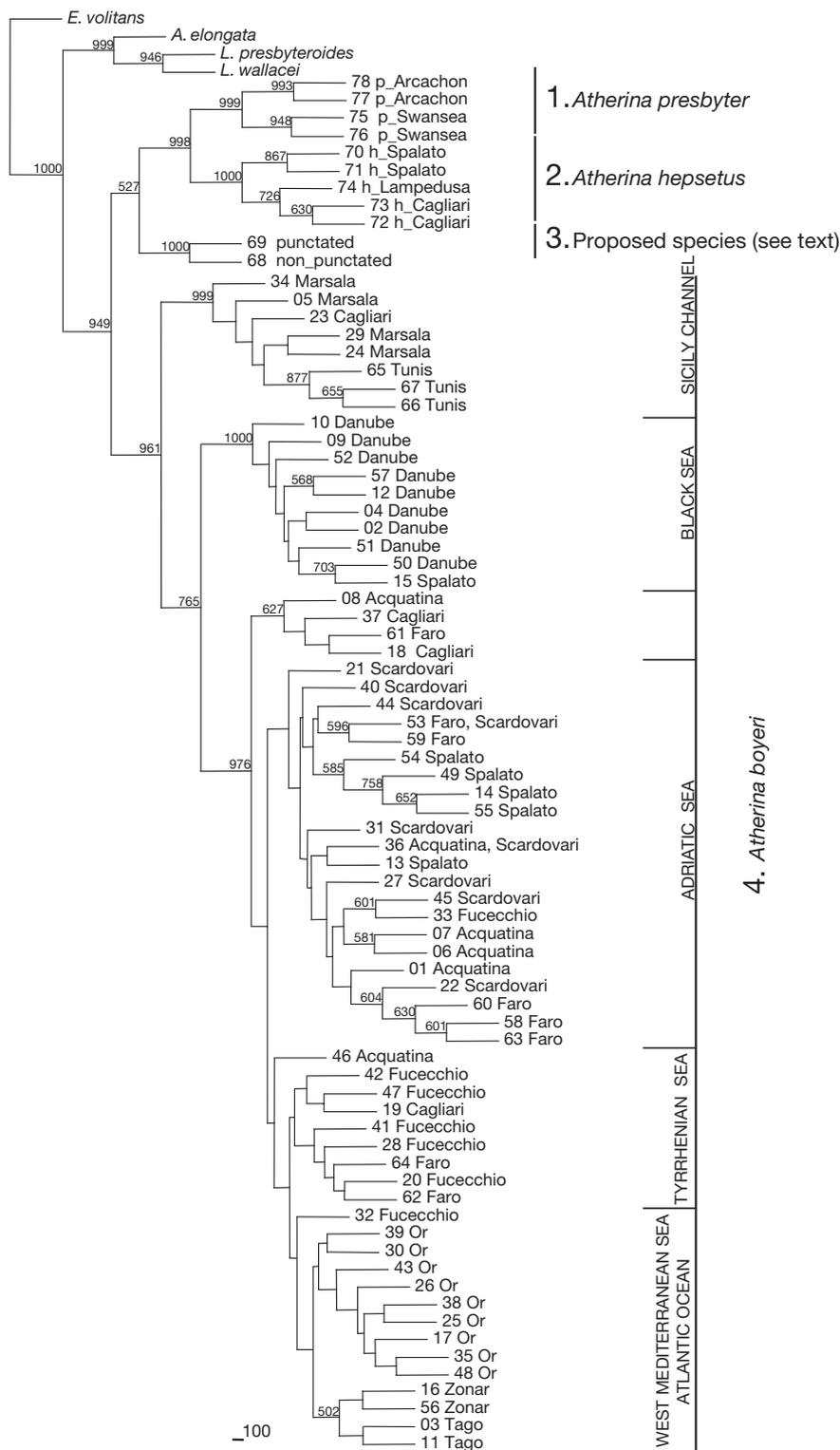


Fig. 2. *Atherina boyeri*. Neighbour-joining tree based on control region haplotypes found in *A. boyeri* and specimens from closely related species (see text). The number of each haplotype is followed by the sampling locality. Bootstrap values (1000 replicates) larger than 500 are reported. Bold numbers 1, 2, 3 and 4 indicate the 4 main clusters discussed in the text. Within cluster 4 the main geographical regions are indicated by vertical superscripts

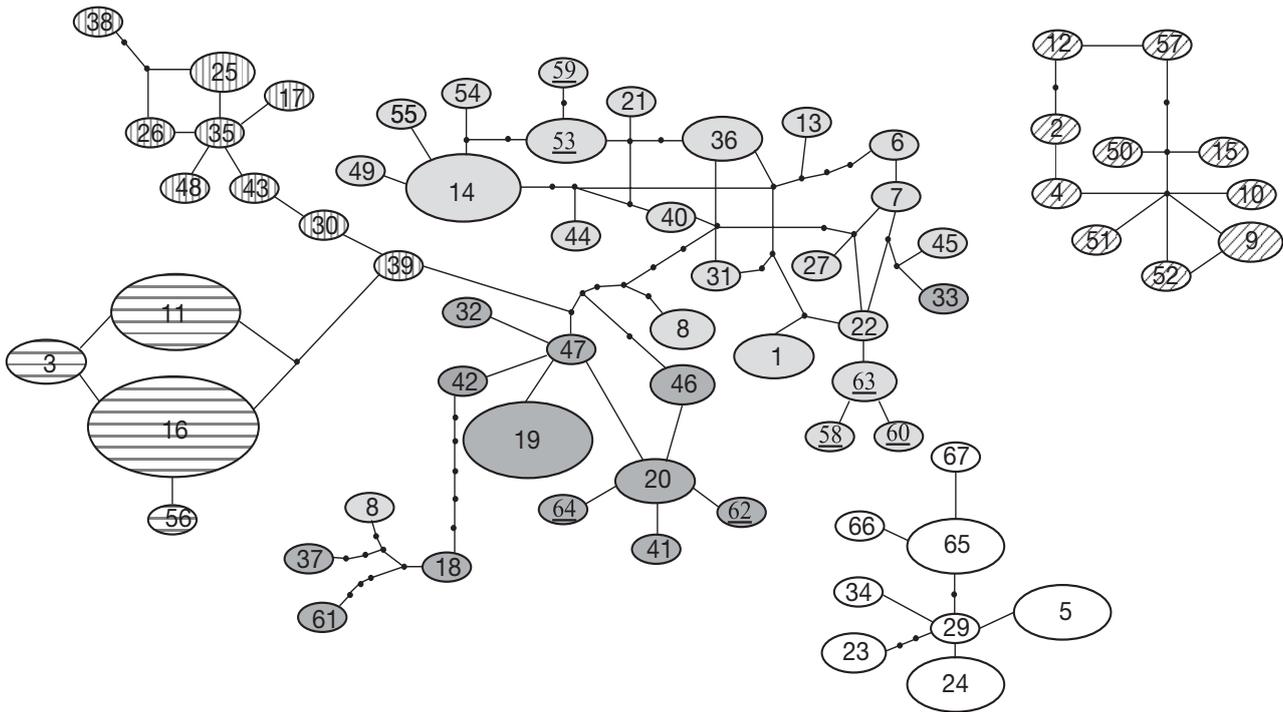


Fig. 3. *Atherina boyeri*. Network of lagoon *A. boyeri* haplotypes (numbered as in Table 1). Circles are sized proportionally to the number of individuals found with the corresponding haplotype. Small black dots = missing haplotypes. Circles are marked according to the geographical origin of the corresponding haplotypes: vertical bars = Western Mediterranean Sea, horizontal bars = Atlantic Ocean, light grey = Adriatic Sea, dark grey = Tyrrhenian Sea, white circles = Sicily Channel, and diagonal bars = Black Sea. Underlined haplotype numbers = Faro population

gene flow among adjacent lagoons and/or the common and recent origin of these gene pools. In this large cluster, sequences found in specimens from Italian lagoons are randomly distributed but distantly related to western Mediterranean and Atlantic haplotypes.

Table 2. *Atherina boyeri*. Genetic diversity in 11 sand smelt populations (Tunis excluded); n: sample size, k: number of different haplotypes, S: number of polymorphic sites, H: gene diversity,  $\bar{\pi}$ : average number of pairwise differences between sequences, D: Tajima's D statistic, Fs: Fu's Fs statistic. ns: non-significant

Locality	n	k	S	H (SD)	$\bar{\pi}$ (SD)
Tago	10	2	1	0.467 (0.132)	0.467
Zonar	10	2	1	0.200 (0.154)	0.200
Or	10	9	8	0.978 (0.054)	2.556
Fucecchio	10	7	13	0.911 (0.077)	3.244
Cagliari	10	4	42	0.644 (0.152)	15.356
Marsala	10	4	3	0.733 (0.101)	1.267
Faro	10	8	24	0.956 (0.059)	8.644
Acquatina	10	6	15	0.889 (0.075)	6.733
Scardovari	10	9	11	0.978 (0.054)	4.089
Danube	10	9	8	0.978 (0.054)	2.644
Adriatic Sea + Faro					
D:				-0.290 (ns)	
Fs:				-12.402 (p < 0.001)	

**Population analysis**

**Genetic diversity.** In the lagoon samples, gene diversity (H) ranges between 0.200 (Zonar) and 0.978 (Or, Danube, Scardovari) (Table 2). Genetic variability is lower in the Atlantic Ocean (only 2 haplotypes in 10 individuals are found in the Tago and Zonar samples) than in the Mediterranean and Black Sea. The mean number of pairwise differences ( $\bar{\pi}$ ) is higher in Spalato, Faro and Cagliari, in agreement with the haplotype tree (Fig. 2), showing extensive divergence between the HVR1 sequences found in these lagoons.

**Mismatch distributions and neutrality tests.** A bimodal mismatch distribution, as well as negative and significant Fu's Fs values (Table 2) is observed in the Adriatic populations, as well as in the Faro population, if all 40 corresponding individuals are included in the analyses. When Haplotype 15 from the Spalato population is excluded, the mismatch distribution becomes unimodal (Fig. 4). Note that this individual is included within the Danube haplotype cluster in the neighbour-joining tree shown in Fig. 2 and in the haplotype network (Fig. 3).

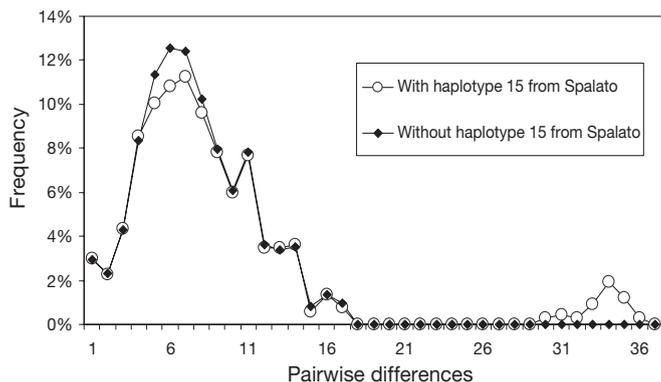


Fig. 4. *Atherina boyeri*. Mismatch distributions in pooled samples from Adriatic and Faro lagoons

**Population differentiation**

The partition of total genetic variance into 2 components (within and among lagoons) confirms the presence of a strong genetic structure in *Atherina boyeri* ( $\phi_{ST} = 0.766$ ,  $p < 0.001$ ). The multidimensional scaling (Fig. 5) confirms the genetic peculiarities of the Danube and Marsala lagoons and the high homogeneity of the Adriatic and Ionian samples. Genetic and geographic distances are significantly correlated ( $r = 0.462$ ,  $p = 0.033$ ), but if we remove the Marsala sample from the analysis, the correlation is stronger ( $r = 0.698$ ,  $p < 0.001$ ).

A strong genetic structure is recognised by the SAMOVA approach, which allows one to identify groups of samples matching those which are best supported by genetic data (Table 3). The Danube and Marsala lagoons are defined as outliers, and 3 groups of samples are defined. The 1st one is composed of the Adriatic and Ionian samples, the 2nd one of Sardinian and Tyrrhenian lagoons and the 3rd one of the Or sample and the Atlantic populations. These results are in good agreement with those obtained by statistical parsimony, represented in the network shown in Fig. 3.

**DISCUSSION**

**Relationships among *Atherina* species**

The choice to analyse *Atherina boyeri* populations together with some samples of the other Mediterranean species of the genus *Atherina* (*A. presbyter* and *A. hepsetus*) was made in order to verify the reliability of the sample identifi-

cations. The extreme phenotypic plasticity and the high similarity among species made species identification difficult. Including individuals of the above groups enabled us to confirm the conspecificity of our samples, excluding the chance of classification errors: our cluster analysis revealed a clear genetic differentiation of the *A. hepsetus* and *A. presbyter* specimens, which thus belonged to independent clades, strongly supported by very high bootstrap values. All *A. boyeri* samples, collected within coastal lagoons, turned out to be monophyletic. According to the systematic revision of genus *Atherina* recently proposed (Klossa-Kilia et al. 2002, Trabelsi et al. 2002a,b), our results showed that the 2 individuals of the punctuated and the non-punctuated forms are separate from the other samples, thus provisionally supporting the newly proposed classification. However, before fully accepting this taxonomical revision, an adequate number of samples should be analysed. The present study showed that the mitochondrial control region may be very useful in clarifying the systematics of genus *Atherina*. We accordingly recommend the use of this marker for this purpose.

**Genetic relationships among *Atherina boyeri* populations**

Our results showed that *A. boyeri* populations are highly structured genetic entities, as previously shown on the basis of nuclear DNA information (Congiu et al. 2002). The overall  $\phi_{ST}$  value ( $\phi_{ST} = 0.766$ ) estimated in the present study using mitochondrial sequences was, as expected, much higher than that previously estimated using RAPD markers ( $\phi_{ST} = 0.33$ ). This should be ascribed to 2 different factors: (1) the 4-fold reduction

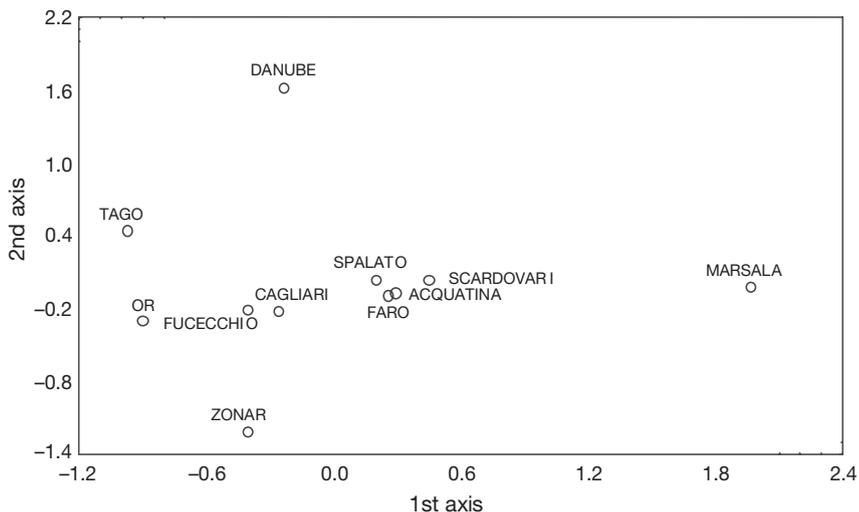


Fig. 5. *Atherina boyeri*. Multidimensional scaling plot based on  $\phi_{ST}$  distances between populations (stress = 0.115)

Table 3. *Atherina boyeri*. Fixation indices corresponding to the 5 groups of populations inferred by SAMOVA algorithms for the 11 sand smelt populations (Tunis excluded) tested for mitochondrial control region sequences. \* $p < 0.001$

Composition of the 5 groups	$F_{SC}$	$F_{ST}$	$F_{CT}$
1. Acquatina, Faro, Scardovari, Spalato	0.278*	0.759*	0.666*
2. Cagliari, Fucecchio			
3. Zonar, Tago, Or			
4. Marsala			
5. Danube			

of effective population size when dealing with mitochondrial markers, with the consequent increase of corresponding  $\phi_{ST}$  values, and (2) the wider geographical scale of the present study compared to the previous one, in which only populations from Italian coasts were analysed. Notwithstanding this high genetic differentiation, a geographic pattern is detectable, as is clearly shown by the cluster analysis of haplotypes (Fig. 2), by the MDS plot (Fig. 5), by SAMOVA results (Table 3) and by statistical parsimony (Fig. 3).

The SAMOVA technique, which allows one to recognise geographically homogeneous and highly differentiated groups of samples, identified 3 groups and 2 outliers in the 11 (Tunis excluded) sampled lagoons. The Marsala population, which showed genetic peculiarities, according to RAPD data (Congiu et al. 2002), is defined as a first mitochondrial outlier. To explain this pattern, Congiu et al. (2002) suggested the presence of gene flow between Sicily and north African lagoons. This exchange of migrants could be favoured by the currents entering the Mediterranean Sea from the Atlantic Ocean (Astraldi & Gasparini 1994). This hypothesis is strongly supported by the results of a study performed on another typical species of Mediterranean lagoons, *Aphanius fasciatus*, in which the Marsala population was found to be genetically more similar to the Tunisian than to the Italian populations (Maltagliati 1999). The clustering of Marsala haplotypes with north African mitochondrial sequences found in Tunisia (Fig. 2) confirmed this hypothesis. However, a common and recent origin of the 2 populations can also be invoked to explain this genetic similarity. No shared haplotypes have been found in the Marsala and Tunis individuals, and the high bootstrap value supporting the separation between Marsala and Tunis haplotypes, may favour this hypothesis.

A first group of samples recognised by the SAMOVA method included populations from the Adriatic and Ionian Seas (Acquatina, Scardovari, Spalato and Faro), which were also characterised by a high genetic homogeneity at the RAPD level. The genetic similarity among these populations is compatible with a recent and common ancestry of these populations followed by

gene flow among them. Migratory movements in this region could indeed have been promoted by the presence of several estuarine systems along the Adriatic coasts and by the occurrence of a constant and significant coastal marine current, flowing anticlockwise in the Adriatic Sea.

A recent and common ancestry of these populations was also suggested by the fact that the Adriatic Sea emerged only after the last glacial

maximum (Thiede 1978). Accordingly, the estuarine and lagoon systems of this area should have been colonised by populations which probably originated in the Ionian Sea. The mismatch distribution (Fig. 4) is indeed compatible with a demographic expansion associated with the colonisation of a new geographical region (Excoffier 2004).

Only Haplotype 15 from the Spalato population did not fit with the general pattern observed in the 4 Adriatic and Ionian populations. In the neighbour-joining tree (Fig. 2), this haplotype was found in the cluster of Danube haplotypes (2nd outlier population according to SAMOVA results). Since sampling errors can be totally excluded, the occurrence in the Adriatic Sea of a haplotype within the range of Danube variability may appear odd, but it can be explained by migration events: the Black Sea is known to have been separated from the Mediterranean basin during the Pleistocene glacial eras, with the consequent isolation of its populations (Magoulas et al. 1996). This allopatric evolution, which is probably the cause of the high genetic difference observed, was followed by a massive flooding of the eastern Mediterranean by large volumes of Black Sea low-salinity waters. This may have favoured flows from the Black Sea genetic pool to more western areas. The analysis of additional samples from the Aegean Sea could be very useful in confirming this hypothesis.

The 2 additional groups of populations identified by the SAMOVA technique include Atlantic and Tyrrhenian populations, the 1st one including Fucecchio and Cagliari and the 2nd one Zonar, Tago and Or. The unexpected grouping of the Or population with the Atlantic ones, notwithstanding the former's closer geographic location to Fucecchio and Cagliari, may be explained by a common origin of these populations. However, the analysis of additional populations along the French and Spanish Mediterranean coasts may help to support this finding. The geographically related pattern described above is well summarised also by the haplotype network shown in Fig. 3, in which, besides the 2 outstanding clusters (Marsala-Tunis and Danube), only 2 haplotypes are connected in unexpected positions (Haplotype 33 sampled in Fucec-

chio and Haplotype 8 sampled in Acquatina). It is interesting to note the dispersed position of haplotypes sampled in the Faro lagoon, which are distributed within the Tyrrhenian and the Adriatic clusters, according to their intermediate geographical locations.

The high genetic structure observed among the 12 analysed populations and the absence of shared haplotypes may support the hypothesis of low genetic flow along the coasts and suggest that the observed phylogeographic pattern is the result of colonisation events followed by the evolution of private haplotypes within the different lagoons. On the other hand, the positive and significant correlation between genetic and geographical distances, as well as the random distribution of sequences within some clusters (such as the Adriatic Sea one), may support the hypothesis of gene flow along the coastal line, not only along the Italian coasts (Congiu et al. 2002), but also on a larger geographical scale. This result should probably be ascribed to occasional exchanges of individuals among populations during the cold season, when the sand smelt stocks migrate from the lagoons to the coastal sea. The observed pattern is probably the result of the combined action of both evolutionary and demographical processes.

The genetic divergence among populations of species inhabiting estuarine environments is the result of the concerted influence of different evolutionary and demographical factors, such as preadaptive plasticity, adaptive selection, genetic flow and genetic drift (Bamber & Henderson 1988, Beheregaray & Sunnucks 2001 and references therein). The resulting geographical pattern of genetic variability, however, may be very different among species. In the South American atherinid *Odontesthes argentinensis*, for which estuarine and marine populations are present, a significant correlation between geographical and genetic distances was observed among the marine populations but absent among the estuarine ones (Beheregaray & Sunnucks 2001). This is probably due to adaptive selection which, acting on the estuarine populations, promoted the divergence of estuarine ecotypes and life histories and their consequent reproductive isolation (Beheregaray & Sunnucks 2001). In *Atherina boyeri*, by contrast, the anadromic behaviour tends to prevent a clear distinction between estuarine and marine populations, and the estuarine gene pools seem to be exposed to the homogenising effect of genetic flow among populations.

Besides life history and related demographical processes, the genetic diversity of *Atherina boyeri* has also been influenced by the geographical history of the Mediterranean Sea. The synergism between paleogeographical and demographic effects is especially clear in the Adriatic Sea zone, where postglacial colonisation occurred over a large area.

## General considerations

The study of species strictly bound to transition environments may provide valuable information about the relationships among different lagoons, their relevance for biomass and biodiversity conservation and their role in shaping the distribution of genetic variability. Unfortunately, notwithstanding the high ecological and economic relevance of these environments, very few data on genetic variability of the species inhabiting them are available. Moreover, the discrete distribution of lagoon environments allows precise sample collection and increases the comparability of studies performed on different lagoon species collected in the same sites. A multispecies approach would help to understand the effect of environmental features or paleoclimatic events on the distribution of genetic variability, as shown by the comparison of *Atherina boyeri* and *A. fasciatus* phylogeographic patterns. This approach is therefore strongly recommended.

Once more it is clear that, besides a comparison among different species, a comparison of results obtained by different markers is also recommended. In this study, the reliability of our results gains support by their agreement with those obtained by RAPD analyses (Congiu et al. 2002).

*Acknowledgements.* Many thanks are due to people who provided us with samples, to M. Chicca for her valuable comments on the previous version of the manuscript and to Prof. G. Colombo for his support. The project was supported by a 40% grant from MURST (Italian Ministry of University and Scientific and Technological Research) awarded in 2002.

## LITERATURE CITED

- 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, DC
- Bamber RN, Henderson P (1988) Pre-adaptive plasticity in atherinids and the estuarine seat of teleost evolution. *J Fish Biol* 33(Suppl A):17–23
- Beheregaray L, Sunnucks P (2001) Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish *Odontesthes argentinensis*. *Mol Ecol* 10:2849–2866
- Berrebi P, Britton-Davidian J (1980) Enzymatic survey of four populations of *Atherina boyeri* based on electrophoresis and the occurrence of microsporidiosis. *J Fish Biol* 16: 149–157
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657–1660
- Congiu L, Rossi R, Colombo G (2002) Population analysis of the sand smelt *Atherina boyeri* (Teleostei Atherinidae), from Italian coastal lagoons by random amplified polymorphic DNA. *Mar Ecol Prog Ser* 229:279–289
- 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
- Creech S (1992) A study of the population biology of *Atherina boyeri* Risso, 1810 in Aberthaw Lagoon, on Bristol Channel, in South Wales. *J Fish Biol* 39:807–816
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Mol Ecol* 11:2571–2581
- Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Mol Ecol* 13:853–864
- Excoffier L, Schneider S (1999) Why hunter-gatherer populations do not show signs of Pleistocene demographic expansions. *Proc Natl Acad Sci USA* 96:10597–10602
- 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 (1985) Confidence limit on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791
- Felsenstein J (1993) PHYLIP (Phylogenetic Inference Package) and manual, version 3.5c. Department of Genetics, University of Washington, Seattle
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925
- Hasegawa M, Di Rienzo A, Kocher TD, Wilson AC (1993) Toward a more accurate time scale for the human mitochondrial DNA tree. *J Mol Evol* 37:347–354
- Henderson PA, Bamber RN (1987) On the reproductive biology of the sand smelt *Atherina boyeri* Risso (Pisces: Atherinidae) and its evolutionary potential. *Biol J Linn Soc* 32:395–415
- Henderson PA, Holmes RHA, Bamber RN (1988) Size-selective overwintering mortality in the sand smelt, *Atherina boyeri* Risso, and its role in population regulation. *J Fish Biol* 33:221–233
- Kiener A, Spillmann CJ (1969) Contribution à l'étude systématique et écologique des Athérines des côtes françaises. *Mem Mus Natl Hist Nat (France) (Nouv Ser) (A Zool)* 60: 1–74
- 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* 341:563–580
- Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120
- Klossa-Kilia E, Prassa M, Papisotiropoulos V, Alahiotis S, Kiliass G (2002) Mitochondrial DNA diversity in *Atherina boyeri* populations as determined by RFLP analysis of three mtDNA segments. *Heredity* 89:363–370
- MacPherson JM, Eckstein PE, Scoles GJ, Gajadhar AA (1993) Variability of the random amplified polymorphic DNA assay among thermal cyclers, and effects of primer and DNA concentration. *Mol Cell Probes* 7:293–299
- Magoulas A, Tsimenides N, Zouros E (1996) Mitochondrial DNA phylogeny and the reconstruction of the population history of a species: the case of the European anchovy. *Mol Biol Evol* 13:178–190
- Maltagliati F (1999) Genetic divergence in natural populations of the Mediterranean brackish-water killifish *Aphanius fasciatus*. *Mar Ecol Prog Ser* 179:155–162
- Marfin JP (1982) Les problèmes liés au polymorphisme de l'espèce *Atherina boyeri* Risso, 1810. *Cybium* 6:19–26
- Micheli MR, Bova R, Pascale E, D'Ambrosio E (1994) Reproducible DNA fingerprinting with the random amplified polymorphic DNA (RAPD) method. *Nucleic Acids Res* 22: 1921–1922
- Mistri M, Colombo G (1988) Morphometric variability in sandsmelt, *Atherina boyeri* Risso 1810, population from different Italian sites. *Boll Zool* 3:129–132
- Rieseberg LH (1996) Homology among RAPD fragments in interspecific comparisons. *Mol Ecol* 5:99–105
- Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* 9:552–569
- Rogers AR, Fraley AE, Bamshad MJ, Watkins WS, Jorde LB (1996) Mitochondrial mismatch analysis is insensitive to the mutational process. *Mol Biol Evol* 13:895–902
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva
- Sinclair M, Iles TD (1988) Population richness of marine fish species. *Aquat Liv Resour* 1:71–83
- Sinclair M, Iles TD (1989) Population regulation and speciation in the oceans. *J Cons Int Explor Mer* 45:165–175
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphisms. *Genetics* 123:585–595
- Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132: 619–633
- Thiede J (1978) A glacial Mediterranean. *Nature* 276: 680–683
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 24:4876–4882
- Trabelsi M, Faure E, Quignard JP, Boussaïd M, Focant B, Mâamouri F (2002a) *Atherina punctata* and *Atherina lagunae* (Pisces, Atherinidae), new species in the Mediterranean Sea. 1. Biometric investigations of three Atherinid species. *CR Biol* 325:967–975
- Trabelsi M, Gilles A, Fleury C, Mâamouri F, Quignard JP, Faure E (2002b) *Atherina punctata* and *Atherina lagunae* (Pisces, Atherinidae), new species found in the Mediterranean sea. 2. Molecular investigations of three Atherinid species. *CR Biol* 325:1119–1128
- 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
- Yang Z, Kumar S (1996) Approximate methods for estimating the pattern of nucleotide substitution and the variation of substitution rates among sites. *Mol Biol Evol* 13:650–659