Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular Phylogenetics and Evolution 39 (2006) 734–746
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (*Engraulis encrasicolus*)

A. Magoulas [a,*,1], R. Castilho [b,1], S. Caetano [b], S. Marcato [c], T. Patarnello [c]

[a] *Institute of Marine Biology and Genetics, Hellenic Centre for Marine Research, P.O. Box 2214, GR 710 03 Heraklion, Greece*
[b] *Center for Marine Sciences (CCMAR), University of Algarve, Campus de Gambelas, 8005-139 FARO, Portugal*
[c] *University of Padova, Department of Biology, Section of Genetics, via Trieste 75 35 121 Padova, Italy*

Received 20 September 2005; revised 9 January 2006; accepted 11 January 2006
Available online 3 March 2006

Abstract

This study extends the geographic coverage of a previous study of mitochondrial DNA restriction fragment length polymorphism in European anchovy. Both studies together include 24 samples representing 17 localities extending from the Black Sea, through the Mediterranean Sea to the eastern Atlantic as far south as Dakar, Senegal. Eighty-eight haplotypes define two clades (A and B) separated by 3.2% sequence divergence. Clade A has a star-like genealogy indicative of a recent population expansion. Clade B has a more complex genealogy, consisting of several haplotypes at intermediate frequencies. The distributions of these clades consist of a mosaic with abrupt changes between some areas and gradients between other areas. Clade A predominates the Black and Aegean seas, but is present throughout the Mediterranean. Unexpectedly, new data show that clade A is also at a high frequency in the Atlantic, from Portugal to at least Senegal. Overall, the level of genetic differentiation among populations is high ($F_{ST} = 0.148$, $p < 0.0001$), with the greatest differences between basins. AMOVA reveals four main geographical groups: Atlantic, central Mediterranean, Aegean Sea, and Black Sea. Mismatch distribution clearly indicates historical bottleneck and population expansion for clade A, while for clade B such evidence is equivocal. This difference may reflect a range expansion for both clades, but with higher gene flow (*Nm* values) between demes for clade A. Both contemporary and historical processes are important in shaping the complex genetic population structure of European anchovy.

Keywords: Anchovy; Mitochondrial DNA; Phylogeography; Mediterranean; Atlantic

1. Introduction

The general lack of barriers in marine waters can facilitate high levels of contemporary gene flow between populations, especially in species with pelagic eggs or larvae, or in species with highly migratory adults. Although phylo geographic studies show that marine species generally exhibit less population structuring compared to the freshwater species (e.g., Hauser and Ward, 1998), population structure has been detected in some marine species, especially species with limited dispersal capabilities (Baus et al., 2005; Doherty et al., 1995; Planes, 1998). Other studies show that under some circumstances, phylogeographic structure can appear in marine pelagic species (Bargelloni et al., 2005; Magoulas et al., 1996; Papetti et al., 2005; Zardoya et al., 2004).

Ocean-climate cycles during the Pleistocene ice ages have had an important influence on marine species. Shifts in temperature and lowered sea levels have favored population displacements, vicariances, extinctions, and colonizations. Genetic imprints of these events are evident in many marine pelagic species (Bremer et al., 2005; Lecomte et al., 2004; Magoulas et al., 1996; Viñas et al., 2004). On shorter time scales, decadal climate shifts can

* Corresponding author. Fax: +30 2810 337822.
E-mail address: magoulas@her.hcmr.gr (A. Magoulas).
[1] These authors have contributed equally to this work.

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.01.016

lead to population declines and local extinctions (Grant and Bowen, 1998).

The focus of the present study is on European anchovy (anchovy hereafter), which is distributed in the Black Sea, throughout the Mediterranean and in the Atlantic off West Africa and Europe (Whitehead et al., 1988). Geographical variation in morphological characters indicates that the Mediterranean may include several subspecies or races. The taxonomic status of some of these groups remains doubtful (Spanakis et al., 1989). Other groups may include cryptic species. Borsa (2002) and Borsa et al. (2004) have described an additional anchovy species, *Engraulis albidus*, that seems to be limited to estuaries in the Mediterranean coast of south France and in the northern Adriatic. In addition to this diversity, anchovy appears to have an exceptional degree of population structure among marine species. Studies of both allozyme and mitochondrial (mt) DNA population markers show that anchovy populations are subdivided in the eastern Mediterranean (Bembo et al., 1995, 1996; Spanakis et al., 1989), but less so in the western Mediterranean (Tudela et al., 1999).

In a previous study, Magoulas et al. (1996) used RFLPs to study the mtDNA variation of anchovy populations in the Mediterranean and adjoining seas. In total, 46 haplotypes defined two divergent clades (phylads A and B), which differed greatly in frequency among samples. Clade A was nearly fixed in the Black Sea, whereas its frequency was about 0.85 in the northern Aegean, and about 0.40 in the rest of Mediterranean and the Bay of Biscay. The lowest frequency of clade A was 0.14 in northern Adriatic. The authors postulated that clade A originated in the Black Sea after isolation from the Mediterranean in the late Pliocene-early Pleistocene. The presence of Clade A haplotypes in the Mediterranean and the Bay of Biscay could be attributed to an outflow of anchovy from the Black Sea, possibly during the last deglaciation, ca. 10 thousand years ago when the Bosporus Strait reopened. It should be noted that the two mtDNA clades do no seem to have been sorted out in the two anchovy species described by Borsa et al. (2004), as they are present in both *Engraulis encrasicolus* and *E. albidus*, even if in different frequencies.

Here, we extended the sample base to include additional localities in the Mediterranean, Black Sea and Atlantic, as far south as Dakar, Senegal. These samples, together with those in Magoulas et al. (1996), cover most of the distribution of the anchovy in the Mediterranean and eastern North Atlantic. The aim of the present article is to provide a deeper understanding of the historical and contemporary components of anchovy population structure. Mismatch analysis is used to estimate historical demographic variables. Population upheavals were expected to be a prominent feature of anchovy biology over the last few hundreds of thousands of years because of strong geological and ocean-climate changes. Egg and larval movement in currents and adult dispersal is expected to most influence the contemporary component of population structure.

2. Materials and methods

For the present work, 14 samples of fish were collected over a large portion of the species' distribution, including the Black Sea, the Mediterranean, and the northeastern Atlantic (Table 1, Fig. 1). Samples were obtained using fishing vessels and were immediately transported on ice to the laboratory, where they were frozen at −70 °C, until DNA extraction. In the laboratories, DNA was extracted from muscle of individual fish following standard methods (Sambrook et al., 1989).

Samples were subject to RFLP analysis of mtDNA following the DIG-non-radioactive method described in Magoulas et al. (1996). Five restriction endonucleases were used, and each individual was assigned a five-letter code, which corresponded to the restriction profile of each enzyme, with the order *Bgl*I, *Bgl*II, *Hin*dIII, *Bam*HI, and *Eco*RI. The analysis of the samples from Bari, Bay of Biscay (Bis2), and Dakar encountered technical difficulties and therefore these samples were only scored by using a PCR-based clade test. This assay was developed for the easy scoring of individuals to identify clade origin. This test has over 95% accuracy in correct characterization of the clades (unpublished data, the primer sequences and technical details are available on request). However, in the haplotype-based analysis (see results) it was not possible to use this information.

The RFLP data for these 14 samples were combined with those in the 10 samples of Magoulas et al. (1996) for further analysis. The digestion profiles were used to produce a matrix of presence/absence of restriction sites that was used to estimate the evolutionary distances between haplotypes. The program DA in REAP (McElroy et al., 1991) was used to estimate haplotype (h) and nucleotide diversity (π) and nucleotide divergence (d_{xy}) among populations (Nei and Tajima, 1981). F_{ST} pairwise values were estimated by Arlequin version 2 (Schneider et al., 2000) and genetic heterogeneity among samples was estimated by the exact test of sample differentiation based on haplotype frequencies (Raymond and Rousset, 1995) also implemented in Arlequin version 2 (Schneider et al., 2000).

To examine hierarchical population structure as well as the geographical pattern of population subdivision, we used analysis of molecular variance (AMOVA; Excoffier et al., 1992). This procedure estimates the amount of genetic structuring at different hierarchical levels by quantifying the inter- and intra-group component of total variance by F statistics, the significance of which was tested by 10,000 permutations.

Tajima's D (Tajima, 1989) was used to test for neutrality. Tajima's D is based on the difference between estimates of Θ from the average number of pairwise restriction site differences (π) and from the number of segregating sites (S). Under the infinite mutation model, significant negative values indicate an excess of low-frequency haplotypes that can arise from selection or rapid population growth. Estimation of significance was performed by bootstrap resampling with 10,000 replicates of Tajima's D.

Table 1
List of samples location, codes, date of collection, sample size (N) and measures of genetic diversity, number of haplotypes (n_h), number of sample specific haplotypes (n_{hs}), the ratio of haplotypes to the number of individuals sampled (n_h/N), haplotype diversity with standard deviation (h ± SD) and nucleotide diversity (π) within each of the sampling sites and also within the groups indicated by AMOVA results

Sampling site	Code	Date	N	n_h	n_{hs}	n_h/N	h (±SD)	π
◆Dakar, Senegal, Atlantic	Dak	Mar. 99	34					
Canary Islands, Spain, Atlantic	Can	Mar. 99	48	13	3	0.27	0.5293 ± 0.08842	0.007021
Tangier, Atlantic coast of Morocco	Tan	Dec. 97	62	4	1	0.06	0.2110 ± 0.06753	0.002432
Olhão, South coast of Portugal	Olh	Jul. 97	57	10	1	0.18	0.4292 ± 0.08199	0.007342
Aveiro, West coast of Portugal	Ave	Mar. 98	70	19	5	0.27	0.5843 ± 0.07045	0.007364
✦Bay of Biscay, Northern Atlantic coast of Spain	Bis1	Mar. 93	47	10	0	0.20	0.7937 ± 0.03449	0.016911
◆Bay of Biscay, Northern Atlantic coast of Spain	Bis2	Apr. 01	50					
Malaga, Alboran Sea	Mal	Nov. 98	47	13	2	0.28	0.5902 ± 0.08424	0.008190
✦Gulf of Lions, Northeastern Mediterranean	Lio	Dec. 92	50	11	2	0.22	0.7959 ± 0.03656	0.017335
Livorno, Ligurian Sea	Liv	May. 98	55	19	2	0.35	0.8929 ± 0.02586	0.015895
Otranto, Italy, Southern Adriatic	Otr	Aug. 98	71	14	2	0.20	0.8298 ± 0.02951	0.012402
✦Chioggia, Italy, Northern Adriatic	Chi1	Nov. 93	57					
Chioggia, Italy, Northern Adriatic	Chi2	Jul. 97	57					
Chioggia, Italy, Northern Adriatic	Chi	Pooled	114	22	8	0.39	0.8070 ± 0.02349	0.009692
◆Bari, Italy (Southern Adriatic)	Bar	Jul. 97	70					
✦Patraikos Gulf, East Ionian Sea	Pat1	Aug. 89	121					
✦Patraikos Gulf, East Ionian Sea	Pat2	Oct. 89	55					
✦Patraikos Gulf, East Ionian Sea	Pat	Pooled	176	22	4	0.13	0.8173 ± 0.01474	0.016659
✦Saronikos Gulf, South Aegean Sea	Sar	Jun. 93	59	16	4	0.27	0.8393 ± 0.02981	0.016923
✦Pagasitikos Gulf, Central Aegean Sea	Pag	Sep. 92	20	7	0	0.35	0.5211 ± 0.13464	0.009812
✦Gulf of Kavala, North Aegean Sea	Kav1	May. 89	57					
✦Gulf of Kavala, North Aegean Sea	Kav2	Oct. 89	144					
✦Gulf of Kavala, North Aegean Sea	Kav	Pooled	201	23	5	0.11	0.5595 ± 0.04174	0.010253
✦Varna, Bulgaria, Western Black Sea	Var1	Oct. 92	70					
Varna, Bulgaria, Western Black Sea	Var2	Oct. 97	45					
Varna, Bulgaria, Western Black Sea	Var	Pooled	115	10	3	0.09	0.2580 ± 0.05413	0.001721
Crimea, Ukraine, Northern Black Sea	Ukr	Nov. 97	43	6	2	0.14	0.3732 ± 0.09234	0.001731
Batumi, Georgia, Eastern Black Sea	Geo	Feb. 99	50	8	4	0.16	0.3894 ± 0.08638	0.002150
Regions								
Atlantic (Mal included)			284	38	12	0.13	0.4717 ± 0.03771	0.006371
Mediterranean			525	44	13	0.08	0.8315 ± 0.01146	0.016750
Aegean Sea			221	24	5	0.11	0.5554 ± 0.04016	0.010177
Black Sea			208	17	10	0.08	0.3144 ± 0.04240	0.001838
Total			1238	88		0.07	0.6012 ± 0.04240	0.009637

◆, samples analysed only by clade test, i.e., only clade frequency data available; ✦, samples from previous work (Magoulas et al., 1996).

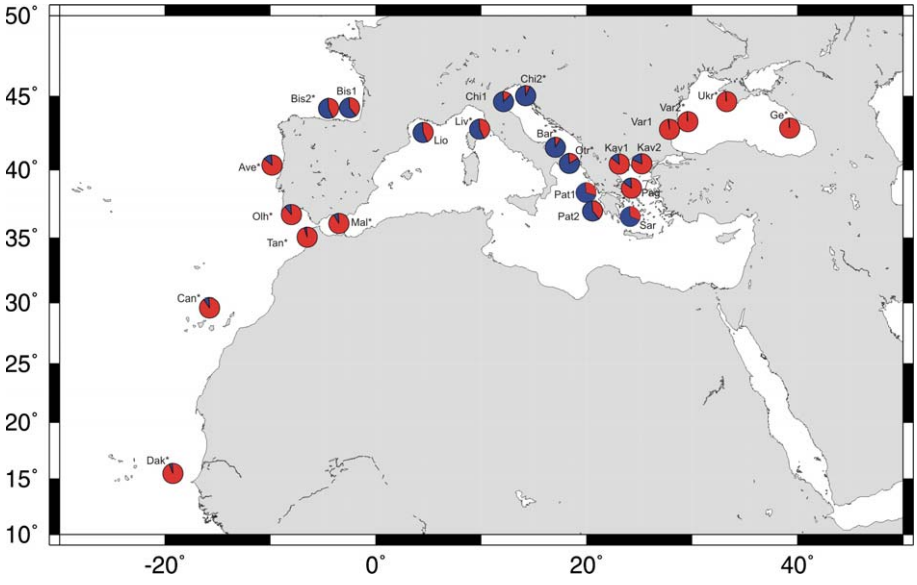

Fig. 1. Sampling sites (present work samples denoted by asterisks in the map and with the codes given in Table 1, other samples from Magoulas et al., 1996) and mtDNA clade frequencies (in red clade A, in blue clade B).

The distributions of the number of pairwise restriction sites differences between individuals (mismatch analyses) were used to estimate historical demographic variables (Rogers and Harpending, 1992). The shape of a mismatch distribution is affected by historical demographic changes of the population: sudden expansion after a bottleneck generates a unimodal mismatch distribution, while stable or slowly declining populations produce a variety of multimodal distributions. The mismatch distribution of an expanding population can provide estimations of the parameters $\Theta_0 = 2Nf_0u$, $\Theta_1 = 2Nf_1u$, and $\tau = 2ut$, where Nf_0 is the initial female population size, Nf_1 the female population size after growth, u the mutation rate of the DNA region under study, and t the time of growth in generations. The mutation rate (u) can be estimated from $\hat{u} = 2\mu k$, where μ is the mutation rate per nucleotide. In this case, 2μ was considered to be 2 or 4% per million years (Bowen and Grant, 1997) and $k = 240$, the number of nucleotide sites that are covered by the restriction sites in the data.

AMOVA, Tajima's D and mismatch analyses were performed using the software Arlequin version 2.0 (Schneider et al., 2000). Both Tajima's D and mismatch analyses were performed for each clade separately, since the apparent mixing of the two divergent clades (see Section 3 and Section 4) in the samples does not allow analyses of the two clades together.

A phylogenetic network of the 88 haplotypes was constructed from the Median Joining algorithm implemented in Network (Bandelt et al., 1999). Ambiguous connections between haplotypes were resolved according to Pfenninger and Posada (2002).

3. Results

In total, 745 new individuals were genotyped in this study. For some individuals it was not possible to score all five enzymes, but nevertheless clade type could confidently be inferred from the enzymes that were successfully scored. In total, 42 new haplotypes were found, which were added to the 46 haplotypes in the study of Magoulas et al. (1996), defined by the same five restriction enzymes. Haplotype and clade frequencies of all the samples appear in Table 2. In total 41 sites were assayed by the battery of five restriction enzymes used in the study and of them only four were not variable.

In general, large haplotype and nucleotide diversities were found in the samples, characteristic of many marine fishes (Grant and Bowen, 1998). Average nucleotide diversity (π) over all samples was ≈1%. The highest nucleotide diversity of pooled samples per region was found in the Mediterranean (1.7%), intermediate values were found in the Atlantic Ocean (0.6%) and Adriatic and Aegean Seas (around 1%), and the lowest in the Black Sea (0.2%) (Table 1). The average haplotype diversity (h) overall populations was 0.60, and the individual haplotype diversities followed the same geographical pattern as that for π (Table 1). Out of character were samples from the Bay of Biscay, which had higher nucleotide and haplotype diversities than those in other Atlantic samples, and Malaga, which had lower nucleotide and haplotype diversities than the average values in the Mediterranean. Genetic variability does not conform to an isolation by distance model, as several pairs of the most distant populations have nucleotide divergences of zero (Table 3).

Replicate sampling carried out in some areas (Var 1–2, Kav 1–2, Par 1–2, and Chi 1–2) indicates that the gene pools are temporally stable and genetically homogeneous with the exact tests for differentiation in haplotypic frequencies (results not shown), therefore these replicate samples from Var, Kav, Par, and Chi, respectively, were pooled by location for the following analyses.

A total of 71 in 136 pairwise exact tests of F_{ST} are significant (Table 3). When the Atlantic region includes Malaga, but excludes Bay of Biscay, only one of the significant comparisons, Chi-Liv, involves an intra-region population comparison. F_{ST} values between samples from the Atlantic and the Mediterranean and from the Mediterranean and the Black Sea were generally significant, but were not significant between the Atlantic and the Black Sea.

There is an overwhelming presence of clade A in the Black Sea (227 of 228 individuals) (Fig. 1). Clade B has a frequency of ≈15% in the northern Aegean and increases to a frequency of ≈65% in the southern Aegean and Ionian seas. The Adriatic is characterized by a high frequency of clade B (>82%), with a slight north to south cline in clade frequencies. Samples from the Ligurian Sea and the Gulf of Lions have clade B frequencies of ≈56%. The most salient finding of the present study is that clade A ranged from 85 to 97% in samples from Malaga in the Alboran Sea and from Tangier, Olhão, Aveiro, the Canaries and Dakar in the Atlantic. These high frequencies stood in sharp contrast to those in neighbouring areas on both sides, namely the Gulf of Lions (42%) and the Bay of Biscay (45%). Samples from the Bay of Biscay differed from other Atlantic samples, exhibiting a Mediterranean clade distribution type, whereas the sample from Malaga has an Atlantic clade profile (Fig. 1).

The haplotype network (Fig. 2) shows the two clades, separated by six mutational steps between their central haplotypes. Clade A is characterized by a star-like genealogy centered on one geographically widespread haplotype (No. 1 in Fig. 1), connected to numerous less-frequent haplotypes differing from one to 6 mutation steps. Clade B has several haplotypes at intermediate frequencies, and has a more reticulated network. Interclade nucleotide divergence was 3.2%.

On a finer scale, analysis of differentiation between samples was based on the frequencies of haplotypes rather than on clade frequencies, to take into account all the variability detected. One gene pool test by AMOVA returned significant levels of genetic structure among the populations ($\Phi_{ST} = 0.1475$; $p < 0.0001$) indicating the existence of two or more groups of populations (Table 4). Additional AMOVA

Table 2
Codes and frequencies of the 88 composite mtDNA haplotypes observed among *Engraulis encrasicolus*, and percentage of clade groups in each sample

Haplotypes	Nb	Total	Atlantic Ocean							West Mediterranean			Adriatic Sea		Ionian Sea				Aegean Sea						Black Sea				
			Dak*	Can*	Tan*	Olh*	Ave*	Bis1	Bis2*	Mal*	Lio	Liv*	Otr*	Bari*	Chi1	Chi2*	Chi	Pat1	Pat2	Pat	Sar	Pag	Kav1	Kav2	Kav	Var1	Var2*	Var	Ukr*	Geo*
Clade A																														
AAAAA	1	652		33	55	43	45	16		30	17	14	10		5	1	6	30	18	48	17	14	92	40	132	58	41	99	34	39
AAACA	2	21		2						4													5	2	7	1		1	3	4
AAABA	3	20		1			5	1				1			1	1	2						4	1	5	1		1	2	2
AAAAL	4	13		2											1		1	3		3	1		4	2	6					
ABAAA	5	12			4	2						2				1	1		2	2			1		1					
AAAAH	6	10		1							1											1	2		2	5		5			
AABAA	7	8		1			1					2							1	1			3		3					
AAAAG	8	7						1				1											2		2	1	1	2	1	
AAHAA	9	6		1			1			1								1		1							2	2		
CAAAA	10	6								1									1	1			3		3					1
AAGAA	11	5					1	1		1													2		2					
AJAAA	12	5				1	2			1											1									
AAAAK	13	4								2		1	1																	
AAEAA	14	4								1												1	1		1		1	1		
AAFAA	15	4					1				1							1		1			1		1					
BAAAA	16	4				1	1						2																	
ABBBB	58	4				2						2																		
DAAAA	17	3																1		1		1	1		1					
AAAAO	18	2																											2	
AACAA	19	2					1					1																		
AGAAG	20	2																								2		2		
AIAAA	21	2				1															1									
ABCBB	71	2										1	1																	
IBABC	72	1								1																				
AAAAB	22	1										1																		
AAAAC	23	1										1																		
AAAAI	24	1									1																			
AAAAJ	25	1																						1	1					
AAAAP	26	1		1																										1
AAABL	27	1																												1
AAADA	28	1									1																			
AAAEA	29	1																												
AABBA	30	1														1	1													
AADAG	31	1				1																								
AAGBA	32	1					1																							
AAIAA	33	1																											1	
AAKAA	34	1					1																							
AALAA	35	1		1																										
ABAAL	36	1					1																							
ABBAA	37	1					1																							
ABCAA	38	1					1																							
ABCAC	39	1														1	1													
AACBC	40	1											1																	
AHAAA	41	1			1																									
BAABA	42	1														1	1													

CBAAA	43	1								1																				
DAAAG	44	1		1																										
DAABA	45	1																												1
EAAAA	46	1																				1		1						
EAADA	47	1																				1		1						
GAABA	48	1																							1		1			
IAAAA	49	1																												1
Clade B																														
BBBBB	50	168		1	2	4	3	11			14	9	24		21	13	34	30	10	40	15	1	8	2	10					
BBCBB	51	126		2				2		2	6	7	12		16	18	34	31	9	40	6		9	6	15					
BBBBC	52	52		1		1	1	10			2	2	2		9	5	14	8	3	11	6	1	1		1					
BBCBC	53	31						1			4		5		2	3	5	6	5	11	2		1	2	3					
BABBB	54	18						2		1	2	3			2	1	3	4	1	5	1					1		1		
BBBAB	55	9											7			2	2													
BBBBA	56	6										3									3									
BBABB	57	4				1	1					2																		
FBBBC	59	4						2											1	1	1									
BCBBB	60	3								1								1		1	1									
BBBAC	61	3											3																	
BDCBF	62	2																1	1	2										
BBCBK	63	2																1	1	2										
BBCAB	64	2											1			1	1													
BBBBE	65	2									1											1								
BBBBD	66	2														1	1		1	1										
BBABA	67	2										1				1	1													
BBAAB	68	2					1									1	1													
BABBC	69	2										1												1	1					
BAABB	70	2					1						1																	
BHBBC	73	1																			1									
BFBBB	74	1																	1	1										
BEBBC	75	1																			1									
BEBBB	76	1																1		1										
BCDBB	77	1																					1		1					
BBKBC	78	1														1	1													
BBJBB	79	1														1	1													
BBEBB	80	1														1	1													
BBCBN	81	1																			1									
BBCBM	82	1																1		1										
BBCBF	83	1																					1		1					
BBCBA	84	1																			1									
BBCAC	85	1														1	1													
BBBBG	86	1														1	1													
BBABC	87	1																1		1										
BABAB	88	1											1																	
Total		1285	34	48	62	57	70	47	50	47	50	55	71	70	57	57	114	121	55	176	59	20	144	57	201	70	45	115	43	50
Clade A %		64.4	94.0	91.7	96.8	89.5	90.0	40.4	42.0	91.5	42.0	49.1	21.1	10.9	12.3	10.5	11.4	29.8	40.0	33.0	33.9	85.0	85.4	80.7	84.1	98.6	100.0	99.1	100.0	100.0
Clade B %		35.6	6.0	8.3	3.2	10.5	10.0	59.6	58.0	8.5	58.0	50.9	78.9	89.1	87.7	89.5	88.6	70.2	60.0	67.0	66.1	15.0	14.6	19.3	15.9	1.4	0.0	0.9	0.0	0.0

Asterisk represent location sampled in this work; other samples are from Magoulas et al. (1996). Columns in grey represent samples for which only clade data is available.

Table 3
Pairwise population subdivision F_{ST} values (above diagonal) (significant exact test values $p < 0.05$ with 10,000 permutations in bold, after Bonferroni correction for 136 simultaneous exact tests) and population pairwise nucleotide divergence (below diagonal) among 17 populations of *Engraulis encrasicolus*

	Can	Tan	Olh	Ave	Bis1	Mal	Lio	Liv	Otr	Chi	Pat	Sar	Pag	Kav	Var	Ukr	Geo
Can		0.0543	0.0010	−0.0039	**0.1235**	−0.0078	**0.1209**	**0.1244**	**0.2268**	**0.2728**	**0.1376**	**0.1300**	−0.0187	−0.0075	0.0409	0.0024	0.0007
Tan	0.000		0.0196	0.0691	**0.2912**	0.0837	**0.2840**	**0.2884**	**0.3878**	**0.4193**	**0.2605**	**0.2901**	0.0555	**0.0559**	0.0068	0.0248	0.0275
Olh	0.000	0.000		0.0073	**0.1591**	0.0171	**0.1561**	**0.1671**	**0.2682**	**0.3175**	**0.1730**	**0.1672**	−0.0136	0.0105	0.0241	0.0053	0.0054
Ave	0.000	0.000	0.000		**0.1046**	0.0021	**0.1053**	**0.1092**	**0.2089**	**0.2601**	**0.1298**	**0.1144**	−0.0118	0.0029	0.0628	0.0169	0.0163
Bis1	0.007	0.009	0.006	0.006		**0.1124**	0.0086	0.0178	0.0538	**0.0808**	0.0270	−0.0028	0.0985	**0.1230**	**0.3082**	**0.1970**	**0.1956**
Mal	0.000	0.000	0.000	0.000	0.007		**0.1075**	**0.1047**	**0.2067**	**0.2538**	**0.1248**	**0.1157**	−0.0080	−0.0010	0.0711	0.0151	0.0106
Lio	0.007	0.008	0.005	0.006	0.000	0.006		0.0050	0.0210	0.0606	0.0011	−0.0080	0.1016	**0.1117**	**0.3027**	**0.1954**	**0.1935**
Liv	0.006	0.007	0.004	0.005	0.000	0.005	0.000		0.0279	**0.0496**	0.0065	−0.0016	**0.1104**	**0.1244**	**0.3099**	**0.1968**	**0.1958**
Otr	0.014	0.015	0.012	0.012	0.001	0.013	0.001	0.002		0.0176	0.0200	0.0194	**0.2092**	**0.2257**	**0.4144**	**0.3021**	**0.3010**
Chi	0.018	0.020	0.016	0.017	0.003	0.018	0.003	0.003	0.000		0.0315	0.0478	**0.2622**	**0.2729**	**0.4426**	**0.3457**	**0.3445**
Pat	0.010	0.011	0.008	0.009	0.000	0.009	0.000	0.000	0.000	0.001		0.0036	**0.1308**	**0.1308**	**0.2751**	**0.2021**	**0.1999**
Sar	0.009	0.010	0.007	0.008	0.000	0.008	0.000	0.000	0.000	0.002	0.000		0.1110	**0.1279**	**0.3117**	**0.2035**	**0.2019**
Pag	0.000	0.000	0.000	0.000	0.005	0.000	0.005	0.004	0.011	0.016	0.008	0.000		−0.0083	0.0351	0.0033	0.0018
Kav	0.000	0.000	0.000	0.000	0.005	0.000	0.005	0.004	0.011	0.015	0.001	0.007	0.000		**0.0510**	0.0165	0.0149
Var	0.000	0.000	0.000	0.000	0.010	0.000	0.009	0.008	0.017	0.022	0.013	0.012	0.000	0.001		0.0108	0.0151
Ukr	0.000	0.000	0.000	0.000	0.010	0.000	0.009	0.008	0.017	0.022	0.013	0.012	0.007	0.001	0.000		−0.0146
Geo	0.000	0.000	0.000	0.000	0.010	0.000	0.009	0.008	0.017	0.022	0.013	0.012	0.000	0.001	0.000	0.000	

tests were therefore performed by pooling populations into alternative groups trying to maximize Φ_{CT} values (Table 4). The maximum Φ_{CT} values for partitions within Atlantic and the Mediterranean corresponded respectively to the Atlantic divided into two sets (Canaries + Tangier + Olhão + Aveiro in one set and Bay of Biscay alone), and the Mediterranean into three sets (Malaga aside; Gulf of Lions + Livorno + Otranto + Chioggia + Patras + Saronikos ; Pagasitikos + Kavala). All of these groupings were significant ($p < 0.05$). The three samples from the Black Sea constitute a single group. Other arrangements based on geographical proximity and genetic similarity of Malaga with the Atlantic, yield the highest Φ_{CT} when Malaga is grouped with Canaries, Tangier, Olhão and Aveiro and when Bay of Biscay is joined to the Mediterranean. This in our view justifies the inclusion of Malaga in the Atlantic region group for the subsequent calculations. We did not pool Biscay into any region, because there is no geographical sense in including it in the Mediterranean group.

The results of tests for population expansion for clades A and B as a whole gave large significantly negative values of Tajima's D (Table 5), consistent with demographic expansion. However, the D value of clade A (-2.180, $p = 0$) was considerably lower than that of clade B -1.544, $p = 0.034$. Mismatch statistics for clade A yield a pattern with a mode of zero, and a mean number of mismatches of 0.564 (Fig. 3, Table 5). The parametric bootstrapping tests (SSD and Raggedness) did not reject fit of the observed distribution to an expansion model (Table 5). Therefore, the population parameters could be estimated, and values of $\tau = 0.811$, $\theta_0 = 0$ and $\theta_1 = 0.794$ were obtained. With $2\mu = 0.02$ per million years, τ yielded 84,000 years since population growth started, while this estimate drops to 42,000 years ago, if $2\mu = 0.04$ per million years. The estimate of θ_1 yielded a recent female population size of 0.083×10^6.

Mismatch distribution for clade B, revealed a smooth unimodal pattern with a mode of one, and a mean number of mismatches of 1.197 (Table 5, Fig. 3). SSD and raggedness tests did not support a fit to the expansion model (Table 5), however, mismatch distribution analysis is a rather conservative approach and some evidence for population expansion can be found in Tajima's D value, as stated above. If one estimates the mismatch parameters, values of $\tau = 1.328$, $\theta_0 = 0$ and $\theta_1 = 1892.2$ are obtained. The τ value corresponds to a population growth that started about 140,000 years ago for $2\mu = 0.02$ per million years, or 70,000 years ago for $2\mu = 0.04$ per million years. The estimate of θ_1 indicates a recent female population size of 197×10^6.

4. Discussion

European anchovy shows a remarkable degree of genetic population subdivision and phylogeographic complexity, in comparison to other coastal pelagic fishes. Populations of other species of anchovy that have been surveyed with molecular population markers, including populations off southern Africa ('*E. capensis*,' Grant, 1985a) and California-Mexico (*E. mordax*, Hedgecock et al., 1989; Lecomte et al., 2004) show virtually no genetic population subdivision among areas. Similar undivided population structures have also been observed for populations of sardines, which often co-occur with anchovies. These include populations of *Sardinops sagax* along the coasts of California-Mexico (Lecomte et al., 2004) and southern Africa (Grant, 1985b) and *Sardinella aurita* in the western Atlantic (Tringali and Wilson, 1993), although differences between Atlantic and Mediterranean populations were reported for the latter species (Chikhi et al., 1997). The contrast between European anchovy and other pelagic fishes can be seen in the differences in values of F_{ST} for the two groups. European anchovy have an overall F_{ST} of about 0.15, whereas most other coastal pelagic species have a F_{ST} of 0.01 or less, often much less.

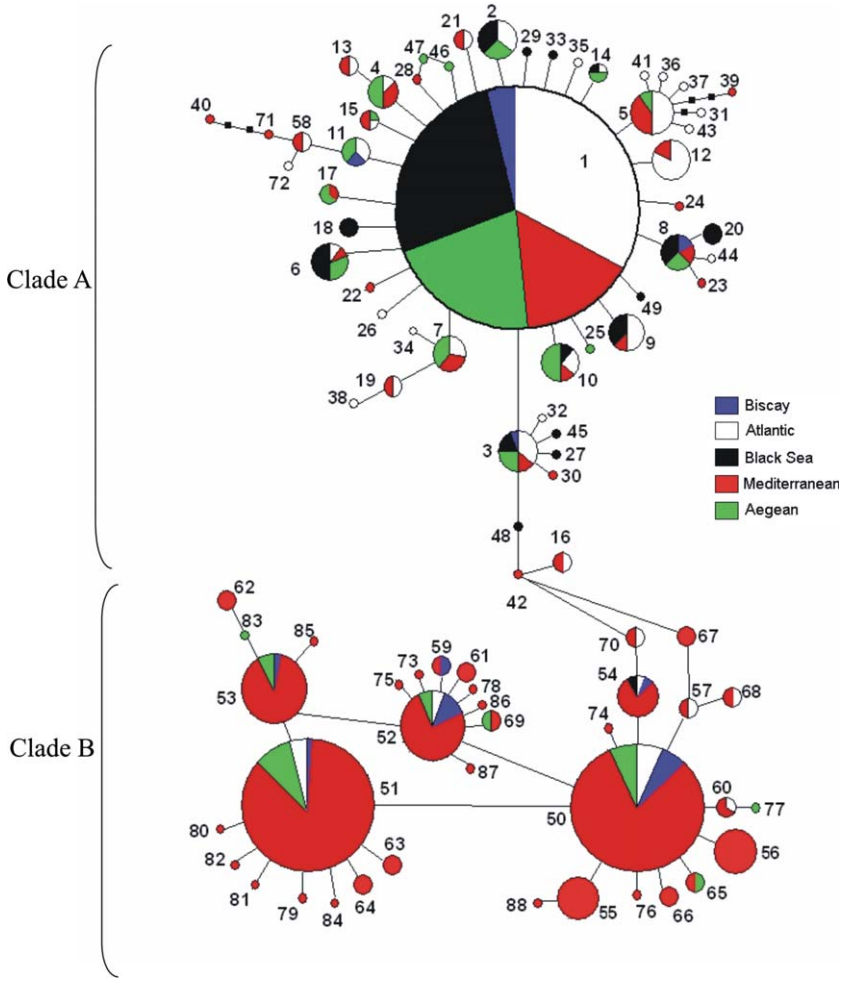

Fig. 2. Median joining tree of the 88 haplotypes detected in *Engraulis encrasicolus*. Haplotypes were divided into seven frequency size classes for a balanced representation. The area of the circles is proportional to the haplotype frequency class. Shading refers to the region in which haplotypes were found. In the case where haplotypes are shared among regions, shading is proportional to the frequency of the haplotype in each region. Black squares represent missing haplotypes. Regions were constituted in agreement with AMOVA results. The Atlantic includes the Malaga sample.

The much higher level of differentiation among populations of European anchovy can be attributed to present-day factors that shape population genetic structure and to historical isolations that have produced the deep divergence between the two clades. It is likely that the two clades have evolved in isolated populations in the past and their coexistence in most of the present distribution range of anchovy is a result of secondary mixing after range expansion.

Clade A revealed a star-like phylogeny, which is a typical signature of a recent population expansion following a population bottleneck (Slatkin and Hudson, 1991). The SSD and Raggedness tests for population expansion, as well as Tajima's *D* test also indicated a relatively recent bottleneck and subsequent population expansion for this clade. In contrast, for clade B the two categories of tests provided contradictory results, SSD and Raggedness tests did not support the population expansion model, while the Tajima's test provided evidence for population growth. The difference between the two clades could be explained if a scenario of range expansion rather than simply population (deme) expansion is considered. It is reasonable to assume that populations of anchovy were restricted in refugia or in isolated basins during the glacial maxima of the Pleistocene and that their range was expanded during the warming of post-glacial eras. Ray et al. (2003) have suggested that the shape of gene genealogy and the overall pattern of diversity between demes included in a range expansion depend not only on the age of the expansion, but also on the level of gene flow between neighboring demes. Star shaped gene genealogies and significant negative values in neutrality tests like Tajima's *D* are produced after a range expansion, only if the demes exchange a high number of migrants, that is if high *Nm* values occur, where *N* is the deme size and *m* is the migration rate between the demes. It is therefore possible that clade A evolved in a population with higher *Nm* values, which expanded spatially with a stronger pace compared to clade B.

If this hypothesis of range expansion is true, the lower genetic variability encountered in the Black Sea can be explained if we consider it as a peripheral deme sensu Ray et al. (2003). The geographic configuration of the Black Sea,

Table 4
Engraulis encrasicolus analysis of molecular variance (AMOVA) with the groupings that maximize Φ_{CT}, with variance, percentage of total variance, Φ statistics (Φ_{SC}, Φ_{ST}, Φ_{CT}) and its significance (p) based on 10,000 replicates

Structure tested	Variance	% total	Φ statistics	p
Observed partition				
One gene pool = no structure (Can; Tan; Olh; Ave; Bis1; Mal; Lio; Liv; Otr; Chi; Pat; Sar; Pag; Kav; Var; Ukr; Geo)				
Among populations	0.053	14.75	$\Phi_{ST} = 0.1475$	<0.0001
Within populations	0.308	85.25		
One gene pool in the Atlantic Ocean (Can; Tan; Olh; Ave; Bis1)				
Among populations	0.024	8.83	$\Phi_{ST} = 0.0883$	<0.0001
Within populations	0.248	91.17		
Atlantic Ocean (Can; Tan; Olh; Ave)(Bis1)				
Among groups	0.055	17.94	$\Phi_{CT} = 0.1794$	0.1985
Among populations/group	0.005	1.56	$\Phi_{SC} = 0.0191$	0.0068
Within populations	0.248	80.5	$\Phi_{ST} = 0.1950$	<0.0001
West-Central-East Mediterranean (Mal) (Lio; Liv; Otr; Chi; Pat; Sar) (Pag; Kav)				
Among groups	0.057	13.21	$\Phi_{CT} = 0.1331$	0.00762
Among populations/group	0.009	1.98	$\Phi_{SC} = 0.0228$	<0.0001
Within populations	0.368	84.81	$\Phi_{ST} = 0.1519$	<0.0001
One gene pool in the Black Sea (Var; Ukr; Geo)				
Among populations	0.001	0.74	$\Phi_{ST} = 0.0074$	0.15832
Within populations	0.157	99.26		
Combination of previous partitions (Can; Tan; Olh; Ave; Mal)(Bis1) (Lio;Liv;Otr; Chi;Pat;Sar)(Pag;Kav)(Var;Ukr;Geo)				
Among groups	0.059	15.91	$\Phi_{CT} = 0.1591$	<0.0001
Among populations/group	0.006	1.67	$\Phi_{SC} = 0.0198$	<0.0001
Within populations	0.308	82.43	$\Phi_{ST} = 0.1757$	<0.0001

Table 5
Neutrality statistics of *Engraulis encrasicolus* within major regions as indicated by AMOVA results

Pooled samples	Clade A	Clade B
Sample size	827	457
Mean number of pairwise differences (d)	0.564	1.197
p (sim. SSD $\geqslant$ obs. SSD)	0.732	0.000
p (sim. Rag. $\geqslant$ obs. Rag.)	0.664	0.000
τ	0.811	1.328
θ	0	0
θ_1	0.794	1892.2
t	84,479	138,833
Nfe_0	0	0
Nfe_1	82,708	197,104,167
Tajima's D	−2.18	−1.544
p	0.000	0.034

p measures the probability of deviation from a hypothesis of sudden population expansion.

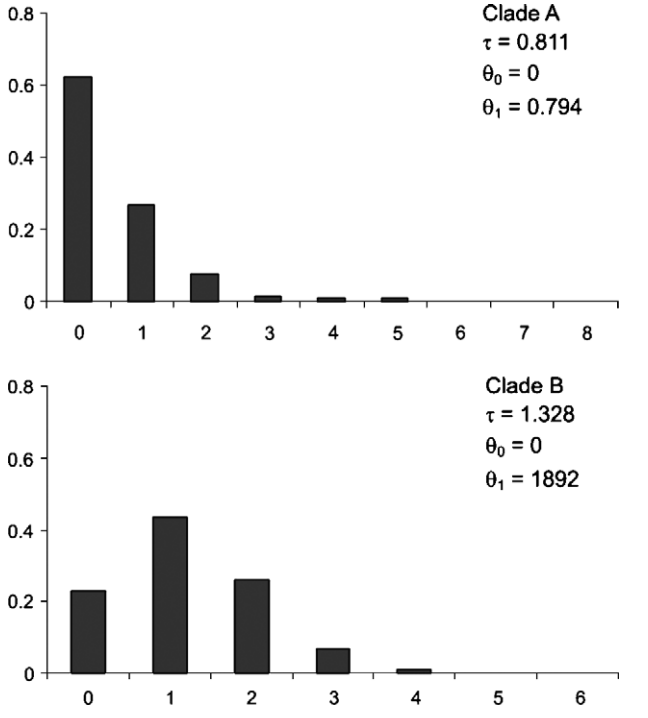

Fig. 3. Number of restriction site pairwise mismatch distributions for clades A and B.

which is an almost completely isolated basin, as well as the hydrography in the Bosporus straits, which further hamper the gene flow to it, further contribute to the reduced level of genetic diversity in this sea.

4.1. Population genetic structure

European anchovy inhabit spatially complex coastal areas, and this complexity tends to isolate populations by reducing levels of gene flow between regions. One fundamental difference between European anchovy and other coastal pelagic species, which show little genetic population subdivision, is that European anchovy inhabit coastal seas isolated from one another by peninsulas and narrow straits. For example, peninsulas partially isolate the Adriatic and Aegean seas from the rest of the Mediterranean, shallow sills in the Bosporus Strait isolate the Black sea from the Mediterranean, and the constriction of the Strait of Gibraltar partially isolates the Mediterranean from the eastern

North Atlantic. Many of the haplotype frequency discontinuities occur across these potential barriers to gene flow.

In addition to physical barriers imposed by complex shorelines, the Mediterranean is characterized by local upwelling, current patterns and isotherms that reinforce isolations among populations. High abundances of anchovy are generally associated with areas of upwelling, which support high levels of productivity (Agostini and Bakun, 2002). However, high levels of productivity by themselves do not ensure the development of fish stocks. Additional requirements are ocean fronts that concentrate larvae food and current gyres that retain eggs and larvae for recruitment into the local population (Agostini and Bakun, 2002). Ocean conditions within the Mediterranean favorable to anchovy populations were identified in the Aegean Sea, Adriatic Sea, Alboran Sea, Gulf of Lions and Catalan coast, and the Strait of Sicily. Upwelling and high levels of productivity in the Canary Current off Northwest Africa (Huntsman and Barber, 1977) also supports large anchovy populations (Cury and Roy, 1989). Notably, these oceanic domains also define to a large extent groups of populations that are genetically distinct from neighboring population groups in other ocean domains (Carlsson et al., 2004; Peijnenburg et al., 2004).

The large-scale patterns of differentiation provide evidence to two previously unrecognized barriers to gene flow between anchovy populations. One occurs within the Mediterranean between the Alboran Sea and the Gulf of Lions and appears to coincide with the oceanic boundary at the Almeria-Oran front (Bargelloni et al., 2003, 2005; Cimmaruta et al., 2005). Another barrier appears in the Atlantic, most likely off the north coast of Spain, and separates Portuguese coastal areas from the Bay of Biscay. The geographically isolated populations of anchovies in the Bay of Biscay have a Mediterranean genetic profile, and populations in the Alboran Sea in the western Mediterranean have a genetic profile typical of Atlantic populations in the Canary Current. The intrusion of Canary Current anchovy into the Alboran Sea may be due to the hydrographic patterns in the Gibraltar straits and to ecological similarities between the two areas, as also suggested for other marine species (Papetti et al., 2005). The Almeria-Oran front is a long recognized barrier to gene flow (Bargelloni et al., 2003, 2005; Perez-Losada et al., 2002; and references therein) apparently acting by preventing the eastward dispersal of planktonic stages. The abrupt genetic change between anchovy in the Bay of Biscay and the Canary Current is most likely due to a combination of historical events and present day hydrographic or ecological patterns. It is likely that the anchovy population in the Bay of Biscay became extinct during the last (Würm) glaciation and was reestablished after its termination, possibly by migrants from the Mediterranean. Present-day hydrography and ecological factors in the area of north Spanish coast likely prevent homogenization of populations of the Bay of Biscay and the Canary Current. During the recent past it has been noted that the anchovy stock has nearly disappeared from the north coast of Spain, while at the same time its distribution has expanded northwards. Anchovy presently is found in the North Sea and extending into the Skagerak and the Kattegatt (ICES, 2005). This fact supports the distinctness of the anchovy population in the Bay of Biscay compared with those of the rest of the Atlantic.

4.2. Historical isolations and dispersals

While the isolating effects of complex shorelines, local upwelling and ocean gyres may explain many of the shallow genetic differences among populations, they cannot account for the deep divergence between the two major clades. Nor can present-day oceanic conditions explain the abrupt shifts in frequencies between the clades in some areas. Explanations for both observations are most likely based mainly on historical events rooted in the paleaoceanography of the Mediterranean and its adjacent seas. Over the past few million years, numerous ocean-climate events undoubtedly precipitated range displacements, population subdivisions, population declines and local extinctions. The Messinian salinity crisis about 5.3 mya, which lowered the level of the Mediterranean Sea and greatly increased salinity, probably predates the events producing the deep haplotype divergences in anchovy. Anchovy populations in the Mediterranean are unlikely to have survived the extreme conditions (Por and Dimentman, 1985).

Except for altered shorelines due to lowered sea levels, the shape of the Mediterranean basin and adjoining seas has not change significantly following the Messinian salinity crisis. The last two million years have been characterized by at least ten ice ages, each with numerous climatic subcycles on scales of thousands and tens of thousands of years (Cronin, 1999). These cycles led to ocean-climate changes (Bond et al., 1993) and lowered sea levels (Thiede, 1978) that greatly influence the distribution and abundance of marine populations, including anchovy populations, in the Mediterranean and northeastern Atlantic.

Magoulas et al. (1996) proposed that the origin of the deep divergence between clades A and B in European anchovy resulted from geographical isolations brought about by climate cycles during the late Pliocene and throughout the Pleistocene. The ancient basin that existed in the broader geographical area of the modern Black and Caspian Seas, lost its connection to the world ocean in the late Pliocene-early Pleistocene, some 1.5–3 mya, and did not regain it until the Riss-Würm interglacial, 100,000–150,000 years ago. At that time the Dardanelles opened for the first time since the formation of the Tethys Sea (Zaitsev and Mamaev, 1997). The connection was lost again during the last glacial maximum (Würm), when lowered sea level cutoff the potential exchange of water between the Black Sea and the Mediterranean across the sill (36 m) in the Bosporus Strait. At the end of the last glaciation, a torrent of ice-melted water gushed through the Bosporus Strait from the Black Sea into the Mediterranean (Herman, 1988). If

anchovy populations in the Black Sea survived the low temperatures and the strong drops in salinity that accompanied the glacial periods, they were possibly isolated from populations in the Mediterranean during most of the Pleistocene. It was also postulated that previously isolated populations of Black Sea anchovy moved into the Mediterranean with the massive water outflow from the Black Sea at the end of the last glaciation. This massive exit is compatible with the hypothesis of a range expansion with high levels of gene flow.

The new information provided by sampling additional areas in the Atlantic, show that clade A is also at a high frequency in populations extending from Portugal to Senegal, but also in populations in the Alboran Sea. This finding calls for a re-consideration of the above-described scenario. The similarity in haplotypic composition of clade A in the two basins dominated by it (Canary current area and Black Sea), as well as in the other regions (Aegean and Mediterranean) is demonstrated by the lack of difference in F_{ST} tests. However, the Atlantic is characterized by higher clade-A haplotype diversity compared to the Black Sea, and this indicates that the Black Sea was not actually the place of origin of clade A, but rather a recipient area, the source population being the Atlantic. In this case, the assumed population bottleneck in the Black Sea could actually be a founder event.

It is reasonable to assume that clade A originated in a refugium, in which anchovy was possibly restricted during glacial periods. If the hypothesis of Atlantic origin of this clade is correct, a possible location for this refugium is the west African coast, but other possibilities could not be excluded. At a certain time later, approximately 100,000 years ago, under favorable climatic conditions, clade-A anchovy should have migrated to the north east Atlantic up to the west Portuguese coast, while it also entered the Mediterranean and colonized the Black Sea. As mentioned above, this range expansion should have taken place under conditions that allowed high levels of gene flow. There are two explanations for the fact that clade A is today under-represented in the analyzed samples from the northern coasts of the Mediterranean. Either the upheavals of the Pleistocene have dissipitated the traces of this movement or the route of migration could have been along the African Mediterranean rim. It should be noted that anchovy populations do presently exist along the Mediterranean African coasts, but unfortunately we lack information on their genetic composition.

This study provided evidence for a long-lasting presence of clade B in the Mediterranean. A fit to the expansion model was not consistently in evidence for this clade, and this may suggest that the population in which it was evolved experienced a spatial expansion, but with a low number of migrants (Ray et al., 2003). It is likely that this population was restricted in a Mediterranean refugium during the last glacial maximum, likely the easternmost Mediterranean, from which it expanded during the deglaciation period in the rest of the Mediterranean.

The episodic population upheavals and dispersals caused by geological and ocean-climate changes during glaciations may have produced the mosaic pattern of haplotype distributions revealed in this study. The oldness, complexity and repeatedness of these phenomena, together with the demographic instability of small pelagic fish, make it very difficult to reconstruct completely the population phenomena that accompanied them. At this stage only speculations can be made based on the pieces of evidence available. Furthermore, when these disturbances ended, the present hydrographic conditions erected barriers to gene flow.

A pattern of mtDNA differentiation strikingly similar to that of the anchovy was recently found in Mediterranean populations of bonito (*Sarda sarda*) by Viñas et al. (2004). In this species two highly divergent clades were also found co-existing in the Mediterranean, one of which displayed a star-like phylogeny and the other a more complex network. To explain this pattern these authors also considered allopatric isolation during the Pleistocene, population bottleneck followed by a sudden expansion that gave origin to the clade with a star-like phylogeny, and later secondary contact of the two previously isolated populations. It is likely that the same historical factors have shaped the phylogeographic patterns in both species.

Recently, Borsa (2002) and Borsa et al. (2004), proposed the existence of a new inshore anchovy species, *E. albidus*, which is less common than the oceanic species *E. encrasicolus*. The recognition of this new species is based on the occurrence of morphologic and genetic differences revealed after a re-examination of previous allozyme (Bembo et al., 1995, 1996; Pasteur and Berrebi, 1985; Tudela et al., 1999) and mtDNA results (Bembo et al., 1995; Magoulas et al., 1996). The relative proportions of the mtDNA clades A and B seem to be different between the two proposed species, but do not represent fixed characters defining each one (Borsa, 2002). Thus, these species bear ancestral polymorphisms that have not completely sorted between species. As described above, these polymorphisms appear to have originated through isolations during ice-age cycles. There are several examples where genetic divergence between populations was suggested to have originated during the early or mid-Pleistocene and were apparently not dissipated afterwards (Bargelloni et al., 2003; Borsa et al., 1997; Brunner et al., 2001; Gysels et al., 2004; Luttikhuizen et al., 2003; Quesada et al., 1995).

It is difficult to correlate the results of the present study with the existence of the *E. albidus*. This species seems to have a restricted distribution in lagoonal areas, and in our opinion represents locally adapted populations in clade-B dominated regions. It is likely that none of our samples belonged to *E. albidus*, since they were in general collected from open oceanic areas. The only exception to this could be the samples collected in the Chioggia (northern Adriatic), which belong to the area of distribution suggested for *E. albidus*. However, these samples did not differ significantly in their haplotype composition from other samples

from the central and south Adriatic (Bari and Otranto), which should normally belong to *E. encrasicolus* species. We consider that more work is needed regarding the distribution and genetic constitution of the newly described species, before sound conclusions can be drawn.

5. Conclusion

European anchovy populations are genetically structured to a greater extent than are populations in other species of anchovies and other species of small pelagic fishes. This structure appears to arise from the interaction between the biology of anchovy and the complex geography and hydrography of the Mediterranean. The pattern of genetic variability among populations has also been greatly influenced by the history of the Mediterranean basin. The high demographic instability makes it difficult to disentangle the effects of contemporary gene flow with those of historic population extinctions, expansions and colonisations, which lead to the present-day puzzling population structure of the European anchovy. However, the combined use of nuclear and mitochondrial markers, as well as the use of other biological data (e.g., characterization of spawning areas and periods) could help to shed light on the role of these various factors on shaping the present-day population structure of anchovy. This would yield important information on both scientific and fisheries-management grounds for such an important and overexploited commercial species.

Acknowledgments

The authors wish to thank everyone who kindly provided samples, namely: Dr. Alain Hervé (IRD, Senegal) (Dakar), Dr. Maria Teresa Garcia Santa Maria (Centro Oceanografico de Canarias, Spain) (Canaries), Dr. Pablo Abaunza (Centro Oceanografico de Santander, Spain) (Santander), Prof. Margarida Castro and Prof. Karim Erzini (CCMAR, Portugal) (Morocco); Mr. Urbi (IOE, Málaga, Spain) (Málaga); Mestre Alfredo Cabrita (Olhão); Mr. Aníbal Guerra (Mútua de Pescadores de Aveiro, Portugal) (Aveiro); Dr. K. Mikhailov (Fisheries Research Institute, Bulgaria) (Varna); Dr. Alexei Petrov (Institute of Biology of the Southern Seas, Ukraine) (Sevastopol). Thanks are also due to Mrs. Vasso Terzoglou and Miss Céu Carvalho for their technical assistance, and Dr. Andy Banks (HCMR) for proofreading. We thank Stewart Grant and two anonymous reviewers for constructive criticism and helpful comments, This work has been funded by the EU grant FAIR-CT96-1899.

References

Agostini, V.N., Bakun, A., 2002. 'Ocean triads' in the Mediterranean Sea: physical mechanisms potentially structuring reproductive habitat suitability (with example application to European anchovy, *Engraulis encrasicolus*). Fish. Oceanogr. 11, 129–142.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Bargelloni, L., Alarcon, J.A., Alvarez, M.C., et al., 2003. Discord in the family Sparidae (Teleostei): divergent phylogeographical patterns across the Atlantic-Mediterranean divide. J. Evol. Biol. 16, 1149–1158.

Bargelloni, L., Alarcon, J.A., Alvarez, M.C., et al., 2005. The Atlantic-Mediterranean transition: discordant genetic patterns in two seabream species, *Diplodus puntazzo* (Cetti) and *Diplodus sargus* (L.). Mol. Phylogenet. Evol. 36, 523–535.

Baus, E., Darrock, D.J., Bruford, M.W., 2005. Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*. Mol. Ecol. 14, 3373–3382.

Bembo, D.G., Carvalho, G.R., Cingolani, N., et al., 1996. Allozymic and morphometric evidence for two stocks of the European anchovy *Engraulis encrasicolus* in Adriatic waters. Mar. Biol. 126, 529–538.

Bembo, D.G., Carvalho, G.R., Snow, M., Cingolani, N., Pitcher, T., 1995. Stock discrimination among European anchovies *Engraulis encrasicolus*, by means of PCR-amplified mitochondrial DNA analysis. Fish. Bull. 94, 31–40.

Bond, G., Broecker, W., Johnson, S., McManus, J., Labeyrie, L., Jouzel, J., Bonani, G., 1993. Correlations between climate records from North Atlantic sediments and Greenland ice. Nature 365, 143–147.

Borsa, P., 2002. Allozyme, nitochondrial DNA, and morphometric variability indicate cryptic species of anchovy (*Engraulis encrasicolus*). Biol. J. Linn. Soc. 75, 261–269.

Borsa, P., Blanquer, A., Berrebi, P., 1997. Genetic structure of the flounders *Platichthys flesus* and *P. stellatus* at different geographic scales. Mar. Biol. 129, 233–246.

Borsa, P., Duran, J.-D., Collet, A., 2004. Nuclear–DNA markers confirm the presence of two anchovy species in the Mediterranean. C. R. Seances Acad. Sci. D 327, 1113–1123.

Bowen, B.W., Grant, W.S., 1997. Phylogeography of the sardines (*Sardinops* spp.): assessing biogeographic models and population histories in temperate upwelling zones. Evolution 51, 1510–1601.

Bremer, J.R.A., Viñas, J., Mejuto, J., Ely, B., Pla, C., 2005. Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. Mol. Phylogenet. Evol. 36, 169–187.

Brunner, P.C., Douglas, M.R., Osinov, A., Wilson, C.C., Bernatchez, L., 2001. Holarctic phylogeography of Arctic charr (*Salvelinus alpinus* L.) inferred from mitochondrial DNA sequences. Evolution 55, 573–586.

Carlsson, J., Mcdowell, J.R., Díaz-Jaimes, P., et al., 2004. Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranean Sea. Mol. Ecol. 13, 3345–3356.

Chikhi, L., Agnès, J.F., Bonhomme, F., 1997. Strong differences of mitochondrial DNA between Mediterranean sea and Eastern Atlantic populations of *Sardinella aurita*. C. R. Seances Acad. Sci. D 320, 289–297.

Cimmaruta, R., Bondanelli, P., Nascetti, G., 2005. Genetic structure and environmental heterogeneity in the European hake (*Merluccius merluccius*). Mol. Ecol. 14, 2577–2591.

Cronin, T.M., 1999. Principles of Paleoclimatology. Columbia University Press, New York.

Cury, P., Roy, C., 1989. Optimal environmental window and pelagic fish recruitment success in upwelling areas. Can. J. Fish. Aquat. Sci. 46, 670–680.

Doherty, P.J., Planes, S., Mather, P., 1995. Gene flow and larval duration in seven species of fish from the Great Barrier Reef. Ecology 76, 2373–2391.

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

Grant, W.S., 1985a. Biochemical genetic stock structure of the southern African anchovy, *Engraulis capensis* Gilchrist. J. Fish Biol. 27, 23–29.

Grant, W.S., 1985b. Population genetics of the southern African pilchard, *Sardinops ocellata*, in the Benguela upwelling system. International Symposium on the Upwelling of West Africa. Inst. Invest. Pesqui. Barcelona 1985 (1) pp. 551–562.

Grant, W.S., Bowen, B.W., 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. J. Hered. 89, 415–426.

Gysels, E.S., Hellemans, B., Pampoulie, C., Volckaert, F.A.M., 2004. Phylogeography of the common goby, *Pomatoschistus microps*, with particular emphasis on the colonization of the Mediterranean and the North Sea. Mol. Ecol. 13, 403–417.

Hauser, L., Ward, R.D., 1998. Population identification in Pelagic fish: the limits of molecular markers. Adv. Mol. Ecol., 191–224.

Hedgecock, D., Hutchinson, E.S., Li, G., Sly, F.L., Nelson, K., 1989. Genetic and morphometric variation in the Pacific sardine, *Sardinops sagax caerulea*: comparisons and contrasts with historical data and with variability in the northern anchovy, *Engraulis mordax*. Fish. Bull. USA 87, 653–671.

Herman, Y., 1988. Paleoenvironments and episodic stagnation of the eastern Mediterranean, during the Brunhes epoch. Rapp. Comm. Int. Mer. Medit. 31 (2), 106.

Huntsman, S.A., Barber, R.T., 1977. Primary production off northwest Africa: the relationship to wind and nutrient conditions. Deep-Sea Res. 24, 25–33.

ICES, 2005. Report of the Study Group on regional Scale Ecology of Small Pelagic Fish. ICES CM 2005/G:06.

Lecomte, F., Grant, W.S., Dodson, J.J., Rodríguez-Sánchez, R., Bowen, B.W., 2004. Living with uncertainty: genetic imprints of climate shifts in East Pacific anchovy (*Engraulis mordax*) and sardine (*Sardinops sagax*). Mol. Ecol. 13, 2169–2182.

Luttikhuizen, P.C., Drent, J., Baker, A.J., 2003. Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larval dispersal. Mol. Ecol. 12, 2215–2229.

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 (*Engraulis encrasicolus*). Mol. Biol. Evol. 13, 178–190.

McElroy, D., Moran, P., Bermingham, E., Kornfiled, I., 1991. The Restriction Enzyme Analysis Package. Smithsonian Tropical Research Institute, Balboa, Panama.

Nei, M., Tajima, F., 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97, 145–163.

Papetti, C., Zane, L., Bortolotto, E., Bucklin, A., Patarnello, T., 2005. Genetic differentiation and local temporal stability of population structure in the euphausiid *Meganyctiphanes norvegica*. Mar. Ecol. Prog. Ser. 289, 225–235.

Pasteur, N., Berrebi, P., 1985. Structure génétique des populations naturelles d'animaux laguno-marins sur des bases de génétique biochimique et d'écologie (Ed. 83.J.0024 RdfdCn), Université de Montpellier, Montpellier, p. 63.

Peijnenburg, K.T.C.A., Breeuwer, J.A.J., Pierrot-Bults, A.C., Menken, S.B.J., 2004. Phylogeography of the planktonic chaetognath *Sagitta setosa* reveals isolation in European seas. Evolution 58, 1472–1487.

Perez-Losada, M., Guerra, A., Carvalho, G.R., Sanjuan, A., Shaw, P.W., 2002. Extensive population subdivision of the cuttlefish *Sepia officinalis* (Mollusca: Cephalopoda) around the Iberian Peninsula indicated by microsatellite DNA variation. Heredity 89, 417–424.

Pfenninger, M., Posada, D., 2002. Phylogeographic history of the snail *Candidula unisfasciata* (Helicellinae: Stylommatophora): fragmentation, corridor migration and secondary contact. Evolution 56, 1776–1788.

Planes, S., 1998. Genetic diversity and dispersal capabilities in marine fish. Evol. Biol. 30, 253–298.

Por, F.D., Dimentman, C., 1985. Continuity of Messinian biota in the Mediterranean basin. In: Stanley, D.J., Wezel, F.C. (Eds.), Geological Evolution of the Mediterranean basin. Springer-Verlag, Berlin, pp. 545–557.

Quesada, H., Zapata, C., Alvarez, G., 1995. A multilocus allozyme discontinuity in the mussel Mytilus galloprovincialis: the interaction of ecological and life-history factors. Mar. Ecol. Prog. Ser. 116, 99–115.

Ray, N., Currat, M., Excoffier, L., 2003. Intra-deme molecular diversity in spatially expanding populations. Mol. Biol. Evol. 20 (1), 76–86.

Raymond, M., Rousset, F., 1995. An exact test for population differentiation. Evolution 41, 823–833.

Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9, 552–569.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schneider, S., Roessli, D., Excofier, L., 2000. Arlequin ver. 2.000: A Software for Population Genetic Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Slatkin, M., Hudson, R.R., 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129, 555–562.

Spanakis, E., Tsimenides, N., Zouros, E., 1989. Genetic differences between populations of sardine, *Sardina pilchardus*, and anchovy, *Engraulis encrasicolus*, in the Aegean and Ionean seas. J. Fish Biol. 35, 417–437.

Tajima, F., 1989. Statistical testing for the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Thiede, J., 1978. A glacial Mediterranean. Nature 276, 680–683.

Tringali, M.D., Wilson Jr., R.R., 1993. Differences in haplotype frequencies of mtDNA of the Spanish sardine *Sardinella aurita* between specimens from the eastern Gulf of Mexico and southern Brazil. Fish. Bull. 91, 362–370.

Tudela, S., Garcia-Marin, J.L., Pla, C., 1999. Genetic structure of the European anchovy, *Engraulis encrasicolus*, in the north-west Mediterranean. J. Exp. Mar. Biol. Ecol. 234, 95–109.

Viñas, J., Alvarado Bremer, J., Pla, C., 2004. Phylogeography of the Atlantic bonito (*Sarda sarda*) in the northern Mediterranean: the combined effects of historical vicariance, population expansion, secondary invasion, and isolation by distance. Mol. Phylogenet. Evol. 33, 32–42.

Whitehead, P.J.P., Nelson, G.J., Wongratana, T., 1988. FAO species catalogue. Vol. 7. Clupeoid fishes of the world (Suborder Clupeoidei). An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, shads, anchovies and wolf-herrings. Part 2–Engraulididae. FAO Fish. Synop. 125 (7/2), 305–579.

Zaitsev, Y., Mamaev, V., 1997. Marine Biological Diversity in the Balck Sea. A Study of Change and Decline. United Nations Publications.

Zardoya, R., Castilho, R., Grande, C., et al., 2004. Differential population structuring of two closely related fish species, the mackerel (*Scomber scombrus*) and the chub mackerel (*Scomber japonicus*) in the Mediterranean. Mol. Ecol. 13, 1785–1798.