Biological Journal of the Linnean Society, 2002, **75**, 261–269. With 5 figures

Allozyme, mitochondrial-DNA, and morphometric variability indicate cryptic species of anchovy (*Engraulis encrasicolus*)

PHILIPPE BORSA*

Institut de Recherche pour le Développement (UR081), BPA5, Nouméa, New Caledonia

Received 22 April 2001; accepted for publication 2 November 2001

Previous surveys of population structure in the Atlantic-Mediterranean anchovy *Engraulis encrasicolus* L. have reported heterogeneity in morphology, allozyme frequencies, and mitochondrial DNA haplotype frequencies at a regional scale. In particular, two stocks of anchovy have been detected in the Adriatic Sea. In this paper, the available data is reviewed with the aim to relate genetic variation to geography at the widest possible geographical scale, for investigating the evolutionary mechanisms underlying stock structure in anchovy. Correspondence analysis of allozyme frequencies (24 samples, three polymorphic loci) compiled from the literature indicates three distinct entities in the Mediterranean Sea. Open-sea or oceanic anchovy populations are genetically different from inshore-water populations within a region (Nei's $\hat{G}_{ST}$ = 0.035–0.067), while broadscale geographical variation is weak for each of these two habitat-specific forms ($\hat{G}_{ST}$ = 0.005–0.006). Mitochondrial-DNA haplotype frequencies support the distinction between an inshore form and an oceanic form ($\hat{G}_{ST}$ = 0.067–0.107), with virtually no genetic differences among oceanic populations across the Gulf of Biscay, the western Mediterranean and the Ionian Sea ($\hat{G}_{ST}$ = –0.001). If natural selection on marker loci is unimportant, these results indicate the occurrence of two parapatric, genetically distinct, habitat-specific forms that are widely distributed throughout the Mediterranean Sea. Persistent allele and haplotype-frequency differences between these forms indicate reproductive isolation and the presence of an *E. encrasicolus* species complex in the Mediterranean. *Biological Journal of the Linnean Society* **75**: 261–269.

ADDITIONAL KEYWORDS: correspondence analysis – habitat specificity – isolation by distance – Mediterranean Sea – systematics.

INTRODUCTION

Because of the economic value of the anchovy, *Engraulis encrasicolus* L., and the need for efficient management of its fisheries, several genetic surveys designed to investigate stock structure have been carried out on this species, mostly in the Mediterranean Sea. Pasteur & Berrebi (1985) detected allozyme-frequency differences between anchovy populations in different habitats (open sea, brackish lagoon) in the Golfe-du-Lion. Spanakis *et al.* (1989) detected significant differences between populations in the Aegean Sea and the Ionian Sea. Bembo *et al.* (1996a) found significant differences between a putative south-central stock and a putative northern stock in the Adriatic Sea. In contrast, Tudela *et al.* (1999) reported no significant genetic structure (Nei's G_{ST} = 0.003) among oceanic populations over an area, 1000 km wide, in the northern part of the western Mediterranean. These genetic differences are associated with morphological differences in the eastern Mediterranean (Spanakis *et al.*, 1989; Bembo *et al.*, 1996a). Tudela (1999), however, reported morphological differences against a background of genetic homogeneity among anchovy samples from the northern part of the western Mediterranean and suggested that the environment was the main determinant of morphological variation among anchovy populations. Mitochondrial (mt) DNA studies have also been conducted to investigate stock structure. Bembo *et al.*

*E-mail: Philippe.Borsa@noumea.ird.nc

(1995) detected heterogeneity in mtDNA haplotype frequency among anchovy samples around the Italian Peninsula. In another study, Magoulas *et al.* (1996) reported that European anchovy mtDNA belonged to one of two distinct haplogroups or phylads (*A, B*), distinguished from each other by 3.7% nucleotide divergence. Magoulas *et al*. (1996) suggested that the two phylads evolved in geographical isolation from each other and that their present coexistence in the Mediterranean Sea resulted from secondary contact. Significant heterogeneity is found in the distribution of the two haplogroups across the Mediterranean (Magoulas *et al*., 1996).

Such levels of genetic variation among populations are rare in the pelagic realm where, as a rule, only small genetic differences are detected across oceans (e.g. see Grant & Utter, 1984; Grant, 1984; Ward, 1995; Gold & Richardson, 1998; Grant & Bowen, 1998; Graves, 1998). The case of *E. encrasicolus,* therefore, is of potential interest to study the patterns and processes of genetic differentiation in pelagic fishes. However, so far no clear geographical picture of genetic variation is available in Atlantic-Mediterranean anchovy.

In this paper, the available morphological, allozyme, and mtDNA data for anchovy to: (1) further explore the patterns of genetic variation at a supra-regional scale, and (2) examine its systematic status are re-examined. The possibility of geographical patterns in genetic differentiation among populations was tested, and morphological and mtDNA correlates of allozyme-frequency variation searched for.

METHODS

Allozyme frequency data on *E. encrasicolus* were compiled from several sources in the literature (Pasteur & Berrebi, 1985; Spanakis *et al*., 1989; Bembo *et al*., 1996a,b; Tudela *et al*., 1999). The total data set was heterogeneous, with numbers of polymorphic loci scored (*r*) varying from two (Bembo *et al*., 1996a) to eight (Bembo *et al*., 1996b; Tudela *et al*., 1999). In this study, the locus is considered to be polymorphic when the frequency of the most common electromorph in the total sample was <0.95 in at least one sample. The geographical coverage of any study was generally restricted to a particular region: the Golfe-du-Lion (Pasteur & Berrebi, 1985; $r = 6$, $N = 5$, where N = number of samples); the Ionian and Aegean Seas (Spanakis *et al*., 1989; $r = 4$, $N = 8$); the Adriatic Sea (Bembo *et al*., 1996a; $r = 2$, $N = 36$); and the northern part of the western Mediterranean (Tudela *et al*., 1999; $r = 8$, $N = 6$). Only Bembo *et al*. (1996b; $r = 8$, $N = 13$) considered a range of samples, from the Gulf of Biscay in the north-eastern Atlantic and across the Mediterranean, but with only one sample from the western Mediterranean. The largest-possible combined data set, in terms of both geographical range ($N = 24$; Fig. 1) and number of polymorphic loci scored in common ($r = 3$; Table 1) combines data from Pasteur & Berrebi (1985), Bembo *et al*. (1996b) and Tudela *et al*. (1999).

Tudela *et al*. (1999) sampled one region, the Tyrrhenian Sea, in common with Bembo *et al*. (1996b), and the Golfe-du-Lion, in common with Pasteur & Berrebi (1985). Identities of the most common electro-

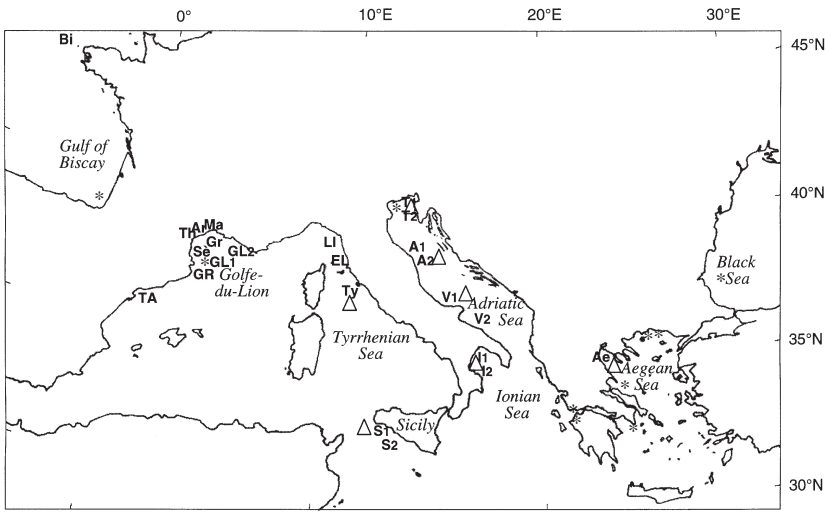

Figure 1. Map of the Gulf of Biscay and Mediterranean Sea with sampling locations for European anchovy, *Engraulis encrasicolus*. Abbreviations (the same as in Table 1) designate samples analysed for allozyme variation. Δ: samples also analysed for mtDNA variation by Bembo *et al*. (1995); *: additional samples analysed for mtDNA variation by Magoulas *et al*. (1996).

morphs across these three studies were established on the basis of similarity in relative electrophoretic mobility and regional frequencies. Identities of rarer electromorphs were uncertain, all faster (respectively, slower) migrating electromorphs at a locus were collapsed into a single '*F*' (resp.'*S*') class (Table 1). No diagnostic locus was found for any paired sample, either from Table 1, or from any of the allozyme surveys considered in this review.

Correspondence analysis (CA; Benzécri, 1982) is an ordination technique based on eigenanalysis of a contingency table, that differs from principal component analysis by deriving both row and column-axes simultaneously, and by using the χ^2-distance metric instead of the Euclidean distance. CA was performed on matrices of allozyme frequencies per sample using the AFC procedure implemented in BIOMECO (Lebreton *et al*., 1990). Guinand (1996) showed that in this case the eigenvalues of each CA's axis can be used as estimates of genetic differentiation between populations. CA allows the graphical representation of multiple–locus associations among samples through clustering in hyperspace. The CA algorithm allows the use of 'supplementary elements'. This are either rows or columns of the data matrix, which are not taken into account for the calculation of the eigenvalues of an axis (unlike 'active elements'), but which nevertheless can be placed in the graphical output. Supplementary elements are therefore useful as reference marks.

CA was also used for determining whether the

Table 1. Electromorph frequencies at three polymorphic allozyme loci in European anchovy, *Engraulis encrasicolus*. Compilation of data from several sources: Pasteur & Berrebi (1985) (PB85); Bembo *et al*., 1996b (BE96); Tudela *et al*., 1999 (TU99). Data obtained by cross comparisons of electromorph mobilities and frequencies in Pasteur & Berrebi (1985) (PB85), Bembo *et al*., 1996b (BE96), and Tudela *et al*. (1999) (TU99) at loci *IDHP-2*, *LDH-1*, and *GPI* (respectively, *Idh-1*, *Ldh-2*, and *Pgi* of PB85). Abbreviations for samples: *Ar*, Arnel lagoon, Golfe-du-Lion; *Gr*, Grau-du-Roi, Golfe-du-Lion; *Ma*, Mauguio lagoon, Golfe-du-Lion; *Sè*, Sète, Golfe-du-Lion; *Th*, Thau lagoon, Golfe-du-Lion (Pasteur & Berrebi, 1985); *Ae*, Aegean Sea; *A1*, *A2*, Ancona, Adriatic sea; *Bi*, Gulf of Biscay; *I1*, *I2*, Ionian Sea; *S1*, *S2*, Sicily; *Ty*, Tyrrhenian Sea; *T1*, *T2*, Trieste, Adriatic Sea; *V1*, *V2*, Vieste, Adriatic Sea (Bembo *et al*., 1996b); *EL*, Elba Island, Tyrrhenian Sea; *GL1*, *GL2*, Golfe-du-Lion; *GR*, Golfo-de-Roses; *LI*, Ligurian Sea; *TA*, Tarragona (Tudela *et al*., 1999). *N*, average sample size per locus

			Electromorph								
Source	Sample	*N*	*IDHP-2*^{F}	*IDHP-2*^{100}	*IDHP-2*^{S}	*LDH-1*^{F}	*LDH-1*^{100}	*LDH-1*^{S}	*GPI*^{F}	*GPI*^{100}	*GPI*^{S}
PB85	Sè	32.3	0.05	0.85	0.10	0.23	0.77	0.00	0.09	0.91	0.00
	Gr	38.7	0.12	0.87	0.01	0.25	0.75	0.00	0.00	1.00	0.00
	Th	27.7	0.46	0.52	0.02	0.33	0.67	0.00	0.00	1.00	0.00
	Ar	27.0	0.46	0.54	0.00	0.39	0.61	0.00	0.00	1.00	0.00
	Ma	24.7	0.53	0.47	0.00	0.28	0.72	0.00	0.00	1.00	0.00
BE96	T1	50.0	0.47	0.52	0.01	0.24	0.76	0.00	0.03	0.97	0.00
	T2	48.7	0.42	0.54	0.04	0.24	0.76	0.00	0.00	1.00	0.00
	A1	50.0	0.22	0.73	0.05	0.16	0.84	0.00	0.01	0.99	0.00
	A2	50.0	0.22	0.76	0.02	0.22	0.78	0.00	0.01	0.97	0.02
	V1	50.0	0.17	0.82	0.01	0.23	0.77	0.00	0.00	1.00	0.00
	V2	49.3	0.17	0.78	0.05	0.30	0.70	0.00	0.00	1.00	0.00
	I1	50.0	0.12	0.83	0.05	0.22	0.78	0.00	0.10	0.90	0.00
	I2	48.7	0.23	0.77	0.00	0.14	0.86	0.00	0.00	1.00	0.00
	S1	50.0	0.18	0.79	0.03	0.24	0.76	0.00	0.20	0.80	0.00
	S2	49.7	0.14	0.82	0.04	0.29	0.71	0.00	0.18	0.81	0.01
	Ty	50.0	0.08	0.88	0.04	0.22	0.78	0.00	0.00	1.00	0.00
	Ae	44.0	0.36	0.62	0.02	0.18	0.81	0.01	0.14	0.85	0.01
	Bi	39.3	0.13	0.81	0.06	0.00	1.00	0.00	0.03	0.97	0.00
TU99	TA	100.0	0.12	0.84	0.04	0.31	0.68	0.01	0.04	0.96	0.00
	GR	99.3	0.15	0.82	0.03	0.26	0.74	0.00	0.04	0.96	0.00
	GL1	100.0	0.12	0.85	0.03	0.36	0.64	0.00	0.01	0.98	0.01
	GL2	100.0	0.14	0.81	0.05	0.26	0.74	0.00	0.01	0.98	0.01
	EL	99.7	0.10	0.87	0.03	0.26	0.74	0.00	0.02	0.98	0.00
	LI	100.0	0.08	0.90	0.02	0.27	0.73	0.00	0.02	0.97	0.01

[100]Most common electromorph at a locus; [F]faster and [S]slower compound electromorphs.

sample of mtDNA haplotypes analysed by Bembo *et al*. (1995) was heterogeneous, to eventually allow comparisons of the mtDNA data of these authors with those of Magoulas *et al*. (1996). Bembo *et al*. (1995) characterized individual mtDNA haplotypes by amplifying a 2.5-kb fragment of the ND5/6 gene using polymerase chain reaction (PCR) and cutting it using restriction enzymes. Table 1 in their study presents 53 composite haplotypes (actually 51, because two of these appear to have been counted twice) defined by restriction profiles for six enzymes. For each enzyme, each restriction profile was assigned a value of one (observed) or zero (not observed). The matrix of haplotypes × restriction profiles was then subjected to CA using the AFC procedure of BIOMECO.

G_{ST}, Nei's (1973) parameter for the apportion of genetic diversity within and across subpopulations (k = 1 to s) was estimated over polymorphic loci as $\hat{G}_{ST} = 1-(\hat{H}_S)/\hat{H}_T$, where $\hat{H}_S$ equals the average gene diversity across loci (j = 1 to r) in subpopulation k ($\hat{H}_S = 1/r.\Sigma_j\hat{h}_{jk}$), and $\hat{H}_T$ equals average gene diversity across loci in the total population ($\hat{H}_T = 1/r.\Sigma_j\hat{h}_{Tj}$). The gene diversities at a locus were estimated from the electromorph frequencies (x_{ijk}) as: $\hat{h}_{jk} = 2N_k/(2N_k-1).(1-\Sigma_i x_{ijk}^2)$ and $\hat{h}_{Tj} = 2\Sigma_k N_k/(2\Sigma_k N_k-1).(1-1/\Sigma_k N_k.\Sigma_i\Sigma_k N_k x_{ijk}^2)$, where N_k = sample size. Sample-pairwise genetic distance estimates [$\hat{G}_{ST}/(1-\hat{G}_{ST})$; Rousset, 1997] were plotted against geographical distances to test isolation by distance. The lack of significant correlation between genetic distance and geographical distance would lead to the rejection of the isolation-by-distance model.

RESULTS

Two main clusters of samples appeared in the correspondence analysis of compound electromorph frequencies (Table 1) of *E. encrasicolus* samples from the Gulf of Biscay, the northern part of the Western Mediterranean, the Tyrrhenian Sea, Sicily, the Adriatic Sea, the Ionian sea and the Aegean Sea (Fig. 2). One cluster, 'Group I', included all samples from coastal habitats, from either brackish lagoons in the Golfe-du-Lion (samples *Ar*, *Ma*, and *Th*) or the extreme north of the Adriatic Sea (samples *T1*, *T2*). A second cluster ('Group II') comprised samples exclusively from the open-sea habitat (i.e. samples *Bi*, *Gr*, *A1*, *A2*, *I1*, *Ty*, *V1*, *V2*, *EL*, *GL1*, *GL2*, *GR*, *LI*, *TA* of Fig. 2). We included two samples from Sicily (*S1*, *S2*) together with sample *Ae* from the northern Aegean Sea into a third ensemble ('Group III'). It was not clear whether samples *Sè* (Golfe-du-Lion) and *I1* (Ionian Sea) belonged to Group II, or were slightly intermediate between Group II and Group III. Genetically intermediate samples may consist of mixtures of individuals from each group, or of hybrids. This could

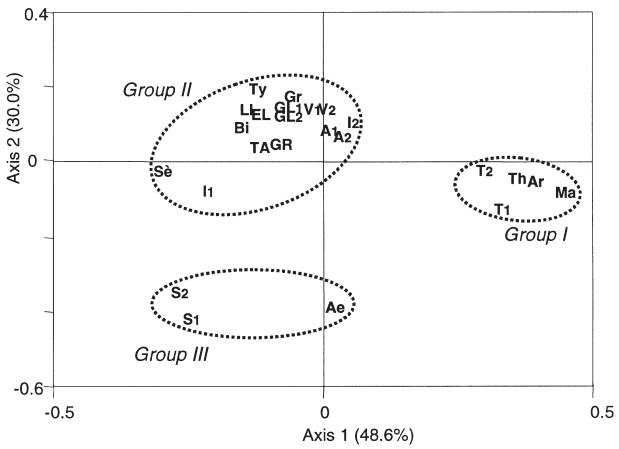

Figure 2. Correspondence analysis (*CA1*): projection on plane defined by axis-1 and axis-2 (percentage of total inertia carried by an axis in brackets) of 24 *Engraulis encrasicolus* samples from the north-eastern Atlantic ocean and the Mediterranean Sea, characterized by their electromorph frequencies at three polymorphic allozyme loci. Abbreviations for samples as in Table 1.

be tested from genotype counts: an excess of heterozygotes would suggest the presence of hybrids, and heterozygote deficiency would suggest Wahlund's effect. Unfortunately, individual genotype data were not provided in any of the source articles to test these hypotheses. It is also possible that these samples belong to different genetic demes, or that the relative dispersion of samples in Group II reflect stochastic sampling variation. To check for internal stability of the clusters, CA was rerun on a portion of the data, as advised by Greenacre (1984). Thus considering a part of the data set [for instance, either all data of Pasteur & Berrebi (1985), or those of Tudela *et al*. (1999)] not as active, but as supplementary samples in the CA did not alter its outcome, only slightly diminishing the inertias of the two first axes (not shown). Importantly, Groups I-III defined as above remained.

Sample-pairwise genetic distances were plotted against the corresponding geographical distances (Fig. 3). Even though Mantel's randomization test on the total matrix of pairwise data indicated a slight increase of genetic distance with geographical distance ($P < 0.05$), suggesting an effect of isolation-by-distance, pairwise comparisons of samples from different groups (open circles in Fig. 3) formed a cluster distinct from paired comparisons within Groups I, II, or III (filled circles). Hence, most of the genetic variability shown in Figure 2 cannot be explained by isolation by distance. The following results and discussion will focus mainly on the distinction between Group I and Group II anchovies.

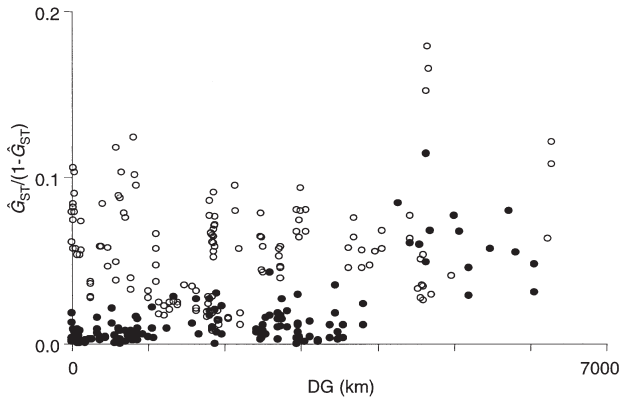

Figure 3. *Engraulis encrasicolus*. Regression of pairwise genetic distances, $\hat{G}_{ST}/(1-\hat{G}_{ST})$ (Rousset, 1997), on geographical distances [*DG*, measured using a map measurer (Wedo, Germany) by following the continental shelf on a 1:3000000 Delisle equidistant conic projection map; Anonymous (1990)] among the samples of Figure 1. A correlation between the two pairwise distance matrices was observed [Mantel's test (implemented in GENETIX; Belkhir *et al.* 1996), $R^2 = 0.069$; $P = 0.036$], but heterogeneous clustering shows that this primarily was not due to isolation by distance. Source data identical to those used for Figure 2. Full (resp. open) circles designate pairs of samples from a single (resp. two different) Group(s).

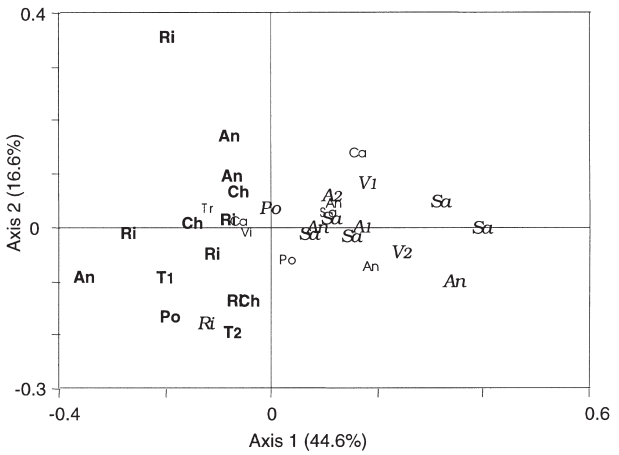

Figure 4. Correspondence analysis (*CA2*): projection on plane defined by Axis 1 and Axis 2 of 35 *Engraulis encrasicolus* samples from the Adriatic sea, characterized by their electromorph frequencies at allozyme loci *G3PDH-2* and *IDHP-2 (*data from Bembo *et al*., 1996a). Canonical correspondence analysis on morphometric data for 27 of these samples (Bembo *et al*., 1996a) distinguished two groups: bold, first canonical variate (CV1) < 0; italic, CV1 > 0; smaller characters, no morphometric data available. Sample locations: *An*, Ancona; *Ca*, Cattolica; *Ch*, Chioggia; *Po*, Porto Garibaldi; *Ri*, Rimini; *Sa*, San Benedetto; *Tr*, Trieste; *Vi*, Vieste. Samples *A1*, *A2*, *T1*, *T2*, *V1*, and *V2* of Figure 2 were treated as supplementary variables in *CA2*, so that *CA2* was independent from *CA1*.

The results of CA inform us about data structure. To make inferences regarding population structure requires a confirmatory analysis. The distinction between Groups I and II was confirmed by a second CA on an independent dataset from the Adriatic Sea, consisting of 29 additional samples characterized at two polymorphic loci (Bembo *et al*., 1996a); six of a total of 35 samples analysed by these authors (*Ancona 5/93* and *9/93*, *Trieste 5/93* and *9/93*, and *Vieste 4/93* and *9/93*) were excluded from the analysis because they appeared to be the same as samples *A1* and *A2*, *T1* and *T2*, and *V1* and *V2*, respectively, in Bembo *et al*. (1996b), which were already used for Figure 2. Figure 4 represents all the Adriatic Sea samples on the plane defined by the two first axes of CA. Two distinct groups appeared: one was defined by negative values along CA's axis-1 and corresponded to Group I defined above, by the position of samples *T1* and *T2* used as supplementary samples; the other was defined by positive values along axis-1 and corresponded to Group II, as indicated by the position of supplementary samples *A1*, *A2*, *V1*, and *V2*. The coordinate of a sample along axis-1 of this CA, which characterized samples genetically, was strongly correlated with the first canonical variate which characterizes them morphometrically (Fig. 5 of Bembo *et al*., 1996a) (Spearman rank correlation coefficient, $r_S = 0.791$; $N = 27$; $P < 0.001$).

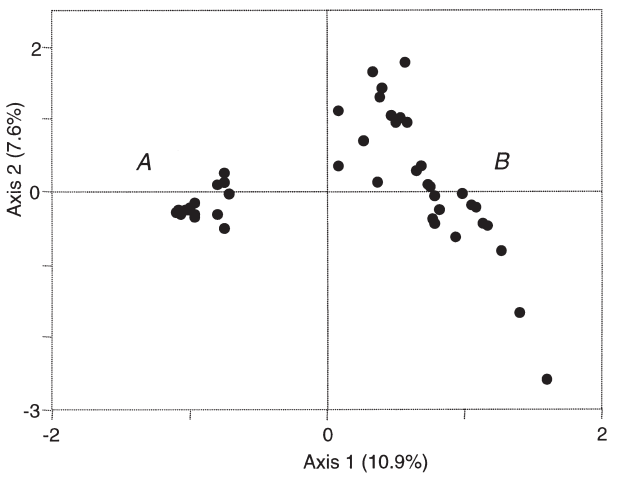

Figure 5. Correspondence analysis: projection on plane defined by axis-1 and axis-2 of the 51 mitochondrial-DNA haplotypes found in seven Mediterranean samples of European anchovy, *Engraulis encrasicolus* (data from Bembo *et al*., 1995) leading to the distinction between *A* and *B* haplogroups. The restriction profile for each of six restriction enzymes was considered as an independent character in the matrix of haplotypes × characters and coded as 1 (present) or 0 (absent).

Focusing on Groups I and II in the Golfe-du-Lion and Adriatic Sea regions, $\hat{G}_{ST}$ intra-Group across regions = 0.005–0.006 ($P = 0.050$); $\hat{G}_{ST}$ inter-Groups in a region = 0.035–0.067 ($P < 0.001$). [Null-hypothesis probabilities of occurrence of $\hat{G}_{ST}$-values larger or equal to the observed value were estimated by permutation tests (GENETIX; Belkhir *et al.*, 1996).]

Significant heterogeneity in mtDNA haplotype frequency among anchovy populations across the Mediterranean has been reported (Bembo *et al*., 1995; Magoulas *et al*., 1996). Five (*Trieste*, *Tyrrhenian*, *Ionian*, *Sicily*, *Aegean*) of the seven samples (adding *Ancona* and *Vieste*) of Bembo *et al*. (1995) are useful to the present work as each can be assigned with confidence to one of the populations whose samples (respectively, *T1* or *T2*, *Ty*, *I1* or *I2*, *S1* or *S2*, and *Ae*) were characterized at allozyme loci by Bembo *et al*. (1996b), hence ascribed to either Group I, II, or III. Samples *Ancona* and *Vieste* could both belong to either Group I or Group II depending on the sampling date (Bembo *et al*., 1996a; this Fig. 4). As sampling dates have not been given in Bembo *et al*. (1995), these two samples cannot be used for comparisons with allozymes. The 51 composite mtDNA haplotypes detected in 140 individuals from all seven Mediterranean samples clustered into two distinct clusters when subjected to CA (Fig. 5). One cluster spanned a small area on the plane defined by factorial axes one and two and comprised 21 haplotypes. This contributed 95% of the haplotypes found in the northern Aegean Sea. The other cluster consisted of a diffuse ensemble of 30 haplotypes, to which belonged 80% of the haplotypes found in the northern Adriatic (*Trieste*). The wide dispersion of the points in this cluster means that the molecular relationship of the haplotypes was less than within the other cluster. From these observations it can be deduced that the first cluster corresponds to phylad *A* of Magoulas *et al*. (1996) that is dominant in the Black Sea (99%) and the Aegean Sea (85%) and also exhibits a star-like phylogeny. The second cluster corresponds to phylad *B* that is dominant in the northern Adriatic Sea (86%) and also exhibits a more diffuse phylogeny. Three groups of samples can be distinguished in the study of Bembo *et al*. (1995), as in Magoulas *et al*. (1996), according to the respective proportions of *A* and *B* haplotypes. One group included all three Adriatic Sea samples (*Trieste*, *Ancona*, *Vieste*) where the frequency of *A* was ≤ 25%. A second group included samples from the Tyrrhenian Sea, the Ionian Sea, and Sicily, with 50%, 45%, and 45% *A* haplotypes, respectively. A third group corresponded to the northern Aegean Sea (95% *A*). From the restricted information available, it appears that Group I (sample *Trieste*) harbours a majority of haplogroup-*B* mitochondria, while approximately 50% of mitochondria in Group II (samples *Tyrrhenian*, *Ionian*) are of the *A* type. Group III samples were variable, with *Sicily* having 45% *A* mitochondria, while *Aegean* reached 95%.

Combining the results of Magoulas *et al*. (1996) with those of Bembo *et al*. (1995; present re-analysis) shows that Group II anchovy have similar haplotype frequencies in the Gulf of Biscay (40% *A*; $N = 48$), the Golfe-du-Lion (42% *A*; $N = 50$), the Tyrrhenian sea (50% *A*; $N = 20$), and the Ionian Sea (45% *A*; $N = 20$). This translates into $\hat{G}_{ST} = -0.001$. Using either the northern Adriatic sample of Magoulas *et al*. (1996) (14% *A*; $N = 65$), or the *Trieste* sample of Bembo *et al*. (1995) (20% *A*; $N = 20$) as representing Group I anchovy yielded $\hat{G}_{ST} = 0.067–0.107$ between the two groups. This result appeared to closely match allozymes.

DISCUSSION

Local variation in morphology has been documented in Clupeoids, including *E. encrasicolus* (Lee & Juge, 1965; Quignard *et al.*, 1973; Spanakis *et al*., 1989, and references therein; Bembo *et al*., 1996a; Tudela, 1999). Although various subspecies of *E. encrasicolus* has been distinguished on the basis of morphology, they have not been retained in the current taxonomy of anchovies (Whitehead *et al*., 1984; Spanakis *et al*., 1989). Population genetic surveys have also yielded evidence of unusually high levels of genetic heterogeneity in this pelagic fish (Pasteur & Berrebi, 1985; Spanakis *et al*., 1989; Bembo *et al*., 1995, 1996a; 1996b; Magoulas *et al*., 1996; Grant *et al*. in Grant & Bowen, 1998). None of these studies, however, have raised the possibility that *E. encrasicolus* may consist of genetically distinct forms that occupy different habitats.

Allozyme frequency differences between anchovy samples collected at regular intervals over a 2-year period throughout the Adriatic yielded evidence for separate, temporally stable 'stocks' (Bembo *et al*., 1996a; Carvalho & Hauser, 1998). The geographical limit between stocks, as delineated by a drop in electromorph frequency for locus *IDHP-2*, coincided with a change in hydrology between shallow (<50 m) waters of the northern Adriatic and the deep sea waters of the southern-central Adriatic (Bembo *et al*., 1996a). Remarkably, this boundary also corresponded to another delineation based on *G3PDH-2* electromorph frequencies (Bembo *et al*., 1996a). Although the northern Adriatic region is in geographical continuity with the open sea, hydrological data show that this region is a distinct, wide shallow area of brackish water. The northern Adriatic region is strongly affected by variation in freshwater discharge from rivers and is subjected to variable climatic forcing, both related to strong spatial and temporal variability of trophic conditions (Fonda-Umani *et al*., 1992; Vichi *et al.*,

1998). Such conditions which are characteristic of brackish-water ecosystems, are encountered in the Golfo-di-Trieste where some of the Group I anchovy samples were collected.

The present analysis shows that the south-central Adriatic stock is genetically similar to geographically distant populations (western Mediterranean, Gulf of Biscay), whereas the northern Adriatic stock is genetically similar to anchovies in the brackish lagoons of the Golfe-du-Lion. The hydrological regime in the Adriatic Sea does not physically separate local populations, but rather delineates a major habitat that harbours a widely distributed form of anchovy, specifically adapted to inshore areas.

The heterogeneous arrangement of the points on Figure 3 departs from the expectations under the neutral model of isolation by distance. Figure 3 could be explained by the inclusion of ecologically and genetically distinct populations in the analysis. Assuming that the habitat-specific forms (Group I, Group II) are conspecific requires selection. Bembo *et al*. (1996a) dismissed the selection hypothesis on the grounds that too high mortalities would be necessary to counter the effect of migration, albeit small, between 'stocks'. Considering that the two loci showing significant differences between Group I and Group II anchovies in the Adriatic Sea are independent, selection should be invoked simultaneously for two loci, which these authors deemed as highly unlikely. An additional argument for ruling out selection of particular allozyme genes arises from the evidence uncovered in this review. It seems very unlikely that the equilibrium between selection and migration would yield the same allelic frequency differences between Groups I and II in the Adriatic Sea and in the Golfe-du-Lion: this would amount to admitting that the very same selective coefficients and the very same migration rates between groups are at work in these geographically distant regions.

Ruling out selection for the particular genes examined, one is left with the hypothesis that Groups I and II, which are each genetically homogeneous across geographically distant regions and which occur in parapatry (or even, perhaps, sympatry) within a region, form distinct gene pools. The report of differences in mtDNA-haplogroup frequency between Groups I and II, though preliminary, suggests some correlation with nuclear-DNA markers. The two genomes do not recombine randomly. Group I and Group II *E. encrasicolus* appear to fit the genotypic-cluster definition of species (Mallet, 1995). High levels of gene flow and possibly stabilizing selection are involved to maintain homogeneity within a group across geographically distant regions. Gene flow between Groups I and II in a region, if any, is comparatively highly restricted. This implies some degree of reproductive isolation between the two groups, which, therefore, correspond to biological species. The maintenance of Groups I and II as distinct species may not be due to reproductive traits *per se*, but perhaps to ecological adaptations that select disruptively against intermediates or hybrids (Mallet, 1995) on either side of the hydrographic boundary between inshore waters and oceanic waters.

Group I (inshore) and Group II (oceanic) anchovies can be distinguished by their morphology (Bembo *et al*., 1996a; Fig. 4), and also have different colour patterns, different counts of vertebrae, different maximal lengths, and different growth rates as inferred from otoliths (Lee & Juge, 1965; Quignard, 1978; Levi *et al*., 1994; Bembo *et al*., 1996a). The morphological distinction between a lagoonal form and an offshore form has also been noted in Tunisia (Quignard *et al*., 1973) but no sample from this region has ever been analysed using allozymes. Currently, it is not possible to determine the genetic affinities of Tunisian anchovies to Groups I-III defined from Figure 1.

The mtDNA data available to date are not sufficient to investigate whether haplotype frequencies are geographically homogeneous for each habitat-specific form. This was suggested, however, for Group II anchovies when combining the results of Magoulas *et al*. (1996) with those of Bembo *et al*. (1995). MtDNA data also confirm that northern Aegean Sea and Sicily anchovy belong to different populations. The coexistence of three differentiated groups of populations (according to both allozyme and mtDNA haplotype frequencies) but only two mitochondrial clades suggests that the evolution of mitochondrial lineages did not parallel that of the nuclear gene pools. Such a complex pattern raises the possibility that previously geographically isolated populations have undergone secondary contact and introgressed (Magoulas *et al*., 1996), but only partly. Alternatively, one may hypothesize that mtDNA diversity reflects ancestral variation that was apportioned differently between demes. Further surveys of genetic variation in European anchovy should be designed to specifically address these issues.

In conclusion, the present review highlights the usefulness of broadscale geographical surveys of genetic variation for interpreting patterns of genetic structure at a regional scale. The emerging result was that *E. encrasicolus* appears to consist of at least two biological species (Group I, Group II) that are both present in the Golfe-du-Lion and in the Adriatic Sea. Group I anchovy occupy the inshore habitat of the two regions, while Group II anchovy are found offshore in the oceanic waters of the Biscay Gulf, the western Mediterranean, the central and southern Adriatic Sea, and the Ionian Sea. I speculate that the form sampled in the northern Aegean Sea and in Sicily (Group III)

might belong to a third species, as preliminarily suggested from morphometric, allozymic and mtDNA data (Spanakis *et al*., 1989; this review). Further population studies of anchovy warrant further insight into the questions of geographical differentiation and genetic architecture in pelagic fishes in relation to habitat specialization and historic events, as extensively debated in recent years (e.g. Bowen & Avise, 1990; Miya & Nishida, 1997; Grant & Bowen, 1998; Graves, 1998; Turan *et al.*, 1998; Nesbø *et al*., 2000). From an applied perspective, updated taxonomy and reassessment of population structure within each species will be indispensable tools for the management of anchovy fisheries in the Mediterranean Sea and perhaps elsewhere in the eastern Atlantic.

ACKNOWLEDGEMENTS

Thanks to F. Bonhomme, V. Castric, W. S. Grant, D. Ponton and J.-P. Quignard who made insightful comments and suggestions for this article. W. S. Grant also offered his help in editing a former manuscript. This is a contribution of Laboratoire Génome Populations Interactions (UMR 5000 CNRS) and UR081/DRV ('Génome / Population / Environnement chez les poissons tropicaux') of Institut de Recherche pour le Développement.

REFERENCES

Anonymous. 1990. *Grand Atlas du Monde*. Paris: Editions Atlas.

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F. 1996. *GENETIX, Version 3.0. Logiciel sous WINDOWS™ pour la Génétique des Populations.* Montpellier : Université Montpellier 2.

Bembo DG, Carvalho GR, Cingolani N, Arneri E, Giannetti G, Pitcher TJ. 1996a. Allozymic and morphometric evidence for two stocks of the European anchovy *Engraulis encrasicolus* in Adriatic waters. *Marine Biology* **126:** 529–538.

Bembo DG, Carvalho GR, Cingolani N, Pitcher TJ. 1996b. Electrophoretic analysis of stock structure in northern Mediterranean anchovies, *Engraulis encrasicolus*. *ICES Journal of Marine Science* **53:** 115–128.

Bembo DG, Carvalho GR, Snow M, Cingolani N, Pitcher TJ. 1995. Stock discrimination among European anchovies, *Engraulis encrasicolus*, by means of PCR-amplified mitochondrial DNA analysis. *Fishery Bulletin* **94:** 31–40.

Benzécri J-P. 1982. *L'Analyse des Données. 2. L'Analyse des Correspondances*. Paris: Dunod.

Bowen BW, Avise JC. 1990. Genetic structure of Atlantic and Gulf of Mexico populations of sea bass, menhaden, and sturgeon: influence of zoogeographic factors and life-history patterns. *Marine Biology* **107:** 371–381.

Carvalho GR, Hauser L. 1998. Advances in the molecular analysis of fish population structure. *Italian Journal of Zoology* **65** (Suppl.)**:** 21–33.

Fonda-Umani S, Franco P, Ghirardelli E, Malej A. 1992. Outline of oceanography and the plankton of the Adriatic Sea. In: Colombo G, ed. *Marine eutrophication and population dynamics*. Fredensborg, Denmark: Olsen and Olsen, 347–365.

Gold JR, Richardson LR. 1998. Mitochondrial DNA diversification and population structure in fishes from the Gulf of Mexico and western Atlantic. *Journal of Heredity* **89:** 404–414.

Grant WS. 1984. Biochemical population genetics of Atlantic herring, *Clupea harengus*. *Copeia* **1984:** 357–364.

Grant WS, Bowen BW. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity* **89:** 415–426.

Grant WS, Utter FM. 1984. Biochemical population genetics of Pacific herring (*Clupea pallasi*). *Canadian Journal of Fisheries and Aquatic Sciences* **41:** 856–864.

Graves JE. 1998. Molecular insights into the population structures of cosmopolitan marine fishes. *Journal of Heredity* **89:** 427–437.

Greenacre MJ. 1984. *Theory and application of correspondence analysis*. London: Academic Press.

Guinand B. 1996. Use of a multivariate model using allele frequency distributions to analyse patterns of genetic differentiation among populations. *Biological Journal of the Linnean Society* **58:** 173–195.

Lebreton J-D, Roux M, Banco G, Bacou A-M. 1990. *BIOMECO (Biométrie-Ecologie), Logiciel d'Ecologie Statistique Pour PC et Compatibles*, Version 3.9. Montpellier: Centre National de la Recherche Scientifique.

Lee J-Y, Juge C. 1965. Observations morphologiques et biologiques sur les anchois (*Engraulis encrasicholus*) du golfe du Lion. *Rapports et Procès-Verbaux des Réunions de la Commission Internationale Pour l'Exploration Scientifique de la Mer Méditerranée* **18:** 221–224.

Levi D, Andreoli MG, Arneri E, Gianetti G, Rizzo P. 1994. Otolith reading as a tool for stock identification. *Fisheries Research* **20:** 97–107.

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*). *Molecular Biology and Evolution* **13:** 178–190.

Mallet J. 1995. A species definition for the Modern Synthesis. *Trends in Ecology and Evolution* **10:** 294–299.

Miya M, Nishida M. 1997. Speciation in the open ocean. *Nature* **389:** 803–804.

Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA* **70:** 3321–3323.

Nesbø CL, Rueness EK, Iversen SA, Skagen DW, Jakobsen KS. 2000. Phylogeography and population history of

Atlantic mackerel (*Scomber scombrus* L.): a genealogical approach reveals genetic structuring among the eastern Atlantic stocks. *Proceedings of the Royal Society of London B* **267:** 281–292.

Pasteur N, Berrebi P. 1985. *Structure Génétique des Populations Naturelles d'Animaux Laguno-Marins sur des Bases de Génétique Biochimique et d'Ecologie* (Contrat no. 83.J.0024 du Ministère de la Recherche, Direction du Développement Scientifique et Technologique et de l'Innovation). Montpellier : Université des Sciences et Techniques du Languedoc.

Quignard J-P. 1978. La Méditerranée creuset ichthyologique. *Bolletin Zoologico* **45** (Suppl. 11)**:** 23–36.

Quignard J-P, Hamdouni T, Zaouali J. 1973. Données préliminaires sur les caractères biométriques des anchois *Engraulis encrasicolus* (Linné, 1758) des côtes de Tunisie et du lac Ichkeul. *Revue Des Travaux de l'Institut Des Pêches Maritimes* **37:** 191–196.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145:** 1219–1228.

Spanakis E, Tsimenides N, Zouros E. 1989. Genetic differences between populations of sardine, *Sardina pilchardus*, and anchovy, *Engraulis encrasicolus*, in the Aegean and Ionian Seas. *Journal of Fish Biology* **35:** 417–437.

Tudela S. 1999. Morphological variability in a Mediterranean, genetically homogeneous population of the European anchovy, *Engraulis encrasicolus*. *Fisheries Research* **42:** 229–243.

Tudela S, García-Marín JL, Pla C. 1999. Genetic structure of the European anchovy, *Engraulis encrasicolus*, in the north-west Mediterranean. *Journal of Experimental Marine Biology and Ecology* **234:** 95–109.

Turan C, Carvalho GR, Mork J. 1998. Molecular genetic analysis of atlanto-scandian herring (*Clupea harengus*) populations using allozymes and mitochondrial DNA markers. *Journal of the Marine Biological Association, UK* **78:** 1–15.

Vichi M, Zavatarelli M, Pinardi N. 1998. Seasonal modulation of microbially mediated carbon fluxes in the northern Adriatic Sea. *Fisheries Oceanography* **7:** 182–190.

Ward RD. 1995. Population genetics of tunas. *Journal of Fish Biology* **47** (Suppl. A)**:** 259–280.

Whitehead PJP, Bauchot M-L, Hureau J-C, Nielsen J, Tortonese E, eds. **1984.** *Fishes of the North-Eastern Atlantic and the Mediterranean, Vol. I.* Paris: Unesco.