Mitochondrial DNA Phylogeny and the Reconstruction of the Population History of a Species: The Case of the European Anchovy (*Engraulis encrasicolus*)

Antonios Magoulas,* Nikolaos Tsimenides,*† and Eleftherios Zouros*‡

*Institute of Marine Biology of Crete, †Department of Biology, University of Crete, and ‡Department of Biology, Dalhousie University

Analysis of mitochondrial DNA restriction fragment length polymorphism in European anchovy (*Engraulis encrasicolus*) revealed a large number of mitotypes that form two distinct clusters (phylads). Phylad A consists of one common mitotype and many rare secondary mitotypes that are one mutational step removed from the main type. Nucleotide diversity and number of homoplasious changes are low. Phylad B has a complex pattern of mitotype connectedness, high nucleotide diversity, and a large number of homoplasious changes. It is suggested that the two phylads evolved in isolation from each other and that present coexistence is the result of a secondary contact. Moreover, phylad A has a "star" phylogeny, which suggests that it has evolved in a population that experienced a drastic bottleneck followed by an explosion of size. Phylad A is practically the only phylad present in the Black Sea, with its frequency dropping to 85% in the northern Aegean, and to 40% in the rest of Mediterranean and the Bay of Biscay. The Black Sea is, therefore, the most likely place of origin of phylad A. Molecular data are consistent with a population bottleneck in the Black Sea during the last glaciation event and a subsequent exit of phylad A with the outflow into the Aegean following the ice melting. Phylogenetic analysis of anchovy mtDNA provides a reconstruction of population history in the Mediterranean, which is consistent with the geological information.

Introduction

"In due course they found themselves entering the narrowest part of the winding straits. Rugged cliffs hemmed them in on either side, and Argo as she advanced began to feel a swirling undercurrent. They moved ahead in fear, for now the clash of the Colliding Rocks and the thunder of the surf on the shores fell ceaselessly on their ears" (Apollonius of Rhodes, 3rd Century BC).

According to Greek Mythology, the Colliding Rocks at the northernmost end of the Bosporus Straits were immobilized when the Argonauts first went through them, so that man may, henceforth, move freely between the Aegean and the Black Seas. For several marine species, however, this remains a one-way passage. This is particularly so for species with passive dispersal, such as planktonic species or species which disperse mainly during their pelagic immature stages. According to Moraitou-Apostolopoulou (1985) the transportation of planktonic animals from the Mediterranean to the Black Sea presents several difficulties and this explains why planktonic forms which are abundant in the Aegean have not been found in the Black Sea. The fauna of the Black Sea is much poorer than that of the adjacent Aegean. Of the approximately 6,000 species known to exist in the Aegean today, only 1,500 are also found in the Black Sea (Tolmazin 1985). Economidis and Bauchot (1976) recorded 114 fish species in the Aegean, of which only 58 are known to occur in the Black Sea. Though many of the Aegean species may enter the Black Sea but fail to establish themselves there, others simply cannot traverse the sea-way that joins the two seas.

Because of its narrow opening to the Atlantic Ocean through the Straits of Gibraltar and its division into several basins with distinct geologic, hydrographic, and biotic characteristics, the Mediterranean has been called a "sea of seas" (e.g., Pastor 1991, p. 20). At the beginning of the Messinian (approximately 5.5 million years ago), communication with the Atlantic was re-established (Por and Dimmentman 1985; Selli 1985). The great bulk of present-day Mediterranean species entered the region during the early Pliocene, slightly less than 5 million years ago, when communication with the Atlantic was re-established (Por and Dimmentman 1985; Sara 1985; Por 1989). At the end of Pliocene, a connection was established between the Black and Aegean Seas through the Dardanelles and Bosporus straits (Dermitzakis 1984; Bacescu 1985; Tortonese 1985). Mediterranean waters intruded into the Black Sea and de-
stroyed the largest part of its brackish paratethyan fauna (Ekman 1968, p. 96). In each of the glaciation events that followed, the communication between the two seas was interrupted during lowering of the sea level at glacial maxima and re-established during the flooding of the interglacial periods (e.g., Por 1978, p. 16; Sara 1985). It is generally believed that the biogeographic consequences of glaciations were much more dramatic in the Mediterranean than in open ocean systems (Por 1978, p. 4). For these reasons the Mediterranean appears to provide one of the best natural scenes to study how colonization events, population bottlenecks, long term isolation, and subsequent mixing have affected the lineage structure and geographical differentiation of present-day populations of marine species.

While considerable attention has been given to the effects of historical and present hydrographic complexity on the species composition of the various basins of the Mediterranean Sea (e.g., Tortonese 1964; Quignard 1978), there has been little attempt to utilize information about hydrographic complexity and the geological history of the Mediterranean to address issues of intraspecific phylogenies and geographical population structure. This is now possible because the introduction of molecular tools in population genetics allows one to combine genetic information with information from the geological, paleontological, and climatic records to reconstruct the population history of a species. This molecular probing into the biogeographical past of a species has, indeed, been used in a number of cases with remarkable success (for a recent review see Avise 1994).

We have attempted to do this in the European anchovy (*Engraulis encrasicolus*) (anchovy hereafter for simplicity), by using restriction fragment length polymorphism of mitochondrial DNA (mtDNA). The anchovy is a small pelagic fish whose distribution includes the Mediterranean and Black Seas, and the Atlantic coasts of north Africa and Europe as far north as the southern North Sea and the coasts of the British Isles (Whitehead 1984). In addition, the anchovy is a euryhaline species persisting in waters from low (5‰) to high (40‰) salinities (Demir 1963). In the Black Sea, where it is the most abundant fish species, it tolerates salinities of approximately 17‰ and, when water temperatures are suitable, it enters the Sea of Azof, where salinity varies from 0 to 17‰ (Ivanov and Beverson 1985). It is a multiple spawner with pelagic eggs and larva. On the basis of variation in morphological characters several authors have suggested the existence of different subspecies or races in the Mediterranean, but the taxonomic status of these divisions remains doubtful (Spanakis, Tsimenides, and Zouros 1989 and references therein).

**Materials and Methods**

**Materials**

The study includes samples of anchovy from eight different localities (fig. 1). Samples from Kavala and Patraiko Guls were taken twice in order to test the stability of the genetic composition in time. Thus, 10 samples were analysed in total.

The Greek samples were either obtained by the Research Vessel “Philia” of the Institute of Marine Biology of Crete or bought fresh from the wharf. In the former case the fish were frozen immediately after capture and brought to the laboratory in -30°C. In the latter case the fish were air-shipped to the laboratory frozen in portable refrigerators. The Black Sea sample was obtained from the Institute of Fisheries of Varna (Bulgaria) and air-shipped to our laboratory frozen in portable refrigerators. Livers of individual fish from Adriatic, Gulf of Lions, and Bay of Biscay were shipped in 70% ethanol. All samples were kept in -30°C until DNA extraction, except the ones preserved in ethanol, which were kept in 4°C until processed.

**Methods**

We employed a restriction site analysis in which fragments were visualized after Southern blotting and hybridization. The hybridization probes were obtained
as follows. Enriched mtDNA was mass-extracted from mature oocytes following the protocol of Lansman et al. (1981), but without the step of purification with cesium chloride gradient centrifugation. For further purification of mtDNA from nuclear DNA, the extracted DNA was digested with restriction enzymes found to cut the mtDNA molecule in few fragments (two such enzymes are BamHI or BgIII). Electrophoresis of the digest produced clear mtDNA bands and a smear of nuclear DNA. The mtDNA bands were recovered from the gel by constructing a well and lining one of its walls by a dialysis membrane. The DNA was blocked onto the membrane and then recovered from the well after a short reversion of the polarity (Magoulas, in preparation). In addition to using it as a probe, this mtDNA was digested with 15 restriction six-cutter endonucleases (AvaI, BamHI, BglII, BglIII, DraI, EcoRI, EcoRV, HindIII, KpnI, PvuII, SacII, SalI, SphI, XmaI, XhoI) with the goal of selecting endonucleases that produce easily and unambiguously detected polymorphisms. The enzymes BglII (cutting site GCCNNN/GGNGC), BglIII (A/GATCT), HindIII (A/AGCTT), BamHI (G/GATCC), and EcoRI (G/AATTC) were finally chosen. Total DNA was extracted from the liver of scored individuals according to Harrison, Rand, and Wheeler (1985). Extraction from the ethanol-preserved livers followed a standard proteinase K/phenol/chloroform method (see Sambrook, Fritsch, and Maniatis 1989), after vacuum desiccation of the samples. Approximately 1/20 of the total volume of each preparation was digested with each restriction endonuclease and the fragments were separated on 0.7% agarose gels at 1.0 V/cm for 10–15 hours. DNA was transferred to nylon membranes (Sambrook, Fritsch, and Maniatis 1989) and hybridized to approximately 100 ng of probe DNA. As alternative probe (with no difference in the scoring power) we used cloned fragments of mtDNA of the American shad (Alosa sapidissima). In the earlier stages of the study labelling of the probe was done with \( ^{32}P \) using the random priming technique (Feinberg and Vogelstein 1983), and the mtDNA fragments were visualised by autoradiography. Subsequently, the probe was labelled with digoxigenin and the fragments detected according to the instructions of the supplier (Boehringer and Mannheim Cat. No. 1093657). HindIII digests of lambda DNA were added in the gels as molecular weight markers. Each distinct restriction profile produced by any of the five endonucleases used for routine scoring (in the order BglII, BglIII, HindIII, BamHI, EcoRI) was assigned an upper-case letter code in alphabetical order of detection (A, B, etc.). Thus, each individual was finally assigned a five-letter composite mitotype.

A restriction-site map for the above enzymes was constructed for all haplotypes using the method of partial and double digestions (Sambrook, Fritsch, and Maniatis 1989). The map was oriented by hybridizations of single and double digests with characterized clones of shad mtDNA. The minimum number of restriction site changes between the different mitotypes was estimated and used to manually construct a minimum length phylogenetic network. Maximum parsimony phylogenetic analysis of the mitotypes was performed using the program PAUP (Swofford 1985). Mitotypes were entered as taxa and presence/absence of restriction sites as characters.

Evolutionary distances (number of nucleotide substitutions per nucleotide site; Nei 1987, equation 8.5) and Nei's genetic distances (Nei 1972). The samples were clustered on the basis of their genetic distances, according to Fitch-Margoliash method, using the program FITCH of the package PHYLIP (Felsenstein 1993). Nucleotide diversities within populations (Nei 1987, equation 10.19) were also estimated using the program REAP.

**Results and Discussion**

**Intraspecific Phylogeny**

Seven different restriction profiles were found for BglII (A–G), 10 for BglIII (A–J), 8 for HindIII (A–H), 7 for BglII, 8 for HindIII, 4 for BamHI (A–D), and 14 for EcoRI (A–N) in 673 fish for which the composite five-enzyme pattern was obtained. Typical profiles for each of these enzymes were published in Magoulas and Zouros (1993). Mapping revealed a total of 34 restriction sites (5 for BglII, 7 for BglIII, 8 for HindIII, 4 for BamHI, and 10 for EcoRI) shown in figure 2. Five of the 34 sites (three HindIII, one BamHI and one EcoRI) were invariable. Character state (presence/absence) combinations at the 29 variable sites generated 46 mitotypes.

The large number of individuals examined and of mitotypes detected in this study allows us to make inferences about the evolution of mtDNA and invites the application of tests to evaluate the effects of significant historical events on present-day structure of mtDNA lineages. Recent work (Rand, Dorfsman, and Kann 1994; Ballard and Kreitman 1994) has challenged the generally held view that mtDNA variation in natural populations is governed solely by the random processes.
FIG. 2.—Restriction site map of *Engraulis encrasicolus* mtDNA. Total length is approximately 17,200 bp. The orientation of the molecule is shown inside the circle. Invariable sites are given in bold. Sites whose exact locations are not determined are indicated with an asterisk. The evolutionary changes of each site are also indicated. @ denotes a site change in the path connecting the two phylads, A denotes a change in phylad A, and B denotes a change in phylad B. Superscripts + and - denote site gain and loss, respectively.

of neutral evolution. We will present our results on the assumption of neutrality, while recognizing that the nature of our data cannot provide direct support for one or the other view.

Attempts to connect the 46 mitotypes with a single mutational step (loss or gain of a restriction site) revealed the existence of two distinct groups of mitotypes, henceforth called phylad A and phylad B (fig. 3). Within phylad A, mitotype AAAAA is connected to all but 3 of the 22 mitotypes by a single step. The remaining three types are connected to AAAAA by two steps through an existing intermediary. An ambiguity arises with EAADA which can be connected to AAAAA through EAAAA or AAADA. The first alternative was chosen because both EAADA and EAAAA were found in the same population (Black Sea), whereas AAADA was found once in the Gulf of Lions. The high frequency of AAAAA and the fact that it occupies the centre of a star-type phylogeny suggests that AAAAA is the most likely ancestral mitotype in this phylad. Using the infinite allele model under the neutral coalescent theory, Crandall and Templeton (1993) arrived at three criteria for rooting an intra-specific phylogeny. Mitotype AAAAA meets all these criteria:

a) It is by far the most frequent mitotype (mean intraphylad frequency 0.81), b) it has 19 connections in the mutational network, whereas mitotypes with the second degree of connectedness have only two connections, and c) it is located in the midpoint of the cladogram. Thus, under the assumptions of neutral coalescent theory, AAAAA is the most likely ancestral mitotype of phylad A.

Phylad B is more complex. There is no single mitotype with which the other 22 types can be connected by one or two steps and there are many networks that can be produced using the criterion of minimal number of steps for joining all 23 types. The network shown in figure 3 was chosen because it joins with single steps the most common type, BBBBB (intraphylad frequency of 0.378), to the next two most frequent ones, BBCBB and BBCBC (frequencies of 0.274 and 0.135), and, then, joins the fourth most common type, BBCBC (frequency of 0.078), to the second most frequent type (BBCBB), again with a single step. We have examined 26 other equally parsimonious networks. In all of them mitotype BBBBB has the highest connectedness and in all but one occupied a position close to midpoint. Given that it is also the most common mitotype in phylad B, BBBBB appears as the most likely ancestral mitotype in this phylad under the assumptions of neutral coalescent theory.

Using again the assumption of coalescent theory, Casteloe and Templeton (1994) proposed a heuristic method to estimate the probability that a haplotype is the oldest in a network of haplotypes. When this argument was applied to phylad A, it gave a root probability of 0.280 for mitotype AAAAA, 0.235 for mitotype AAABA, 0.233 for AAAG, and 0.230 for EAAAA. When applied to phylad B, it gave a probability of 0.283 for BBBBB, 0.227 for BBCBB, 0.186 for BBCBC, 0.134 for BCBBB, and 0.122 for BBCBC. These results agree with the characterization of AAAAA and BBBBB as the most likely ancestral mitotypes of phylads A and B respectively. The approach of Casteloe and Templeton (1994) takes into account whether the haplotype lies in the interior or at the tip of the network, but ignores the connectedness of the haplotype. This apparently explains why mitotypes AAABA, AAAG, and EAAAA were assigned probabilities comparable to that of AAAAA, despite the fact that they are much less frequent and have much lower connectedness (2 versus 19). The same applies to mitotype BCBBB in phylad B.

There are several ways to join the two phylads into one composite network. If we assume that types AAAAA and BBBBB are the most likely ancestral types in each phylad, we may join them directly. This option would require a six-step inter-phylad gap and
would consider the intermediate types AABAA, GAA-
BA, BABBB, and BBBBA as homoplasious deriva-
tives of one or the other central type. The other option
is to minimize the gap separating the two phylads, by
assuming that existing rare intermediate mitotypes are
older than one or the other of the central types (see fig.
3). The network joining all 46 mitotypes involves ei-
ther fifty steps (if AAAA is directly joined to
BBBBB) or 47 steps (if AAAA is joined to BBBB
through existing minor mitotypes).

The manual network described above is in agree-
ment with the results of a phylogenetic analysis using
the PAUP program. More than 100 equally parsimo-
nious trees were produced, all of which connected the
mitotypes of phylad A into one group and the mito-
types of phylad B into another. Mitotypes within each
phylad were connected by single steps and the two
phylads were connected to each other by three steps,
as expected from the criterion of maximum parsimony.
The consistency index of these trees was 0.617, which
indicates that 18 of the 47 changes were homoplasious.
Separate phylogenetic analysis for each phylad showed
that four homoplasious changes occurred in phylad B
and one in phylad A.

The UPGMA phenogram of mitotypes derived
from their evolutionary distance matrix is shown in fig-
FIG. 4.—UPGMA phenogram of the mitotypes, based on their evolutionary distance. The two phylads are connected by a node at a distance of approximately 0.037. Mitotypes are numbered as in Appendix 1.

The main feature of this figure is that the interphylad divergence was three times as large as the intraphylad divergence. Specifically, the mean within phylad divergence was 0.012 and 0.016 for phylad A and B respectively, whereas the mean sequence divergence between the two phylads was 0.037, which is among the highest degrees of within-species mtDNA divergence reported in marine fish species (see Billington and Hebert 1991, table 5).

Geographic Variation

Appendix 1 gives the distribution of the 46 mitotypes in the ten samples. The Monte-Carlo simulation method of Roff and Bentzen (1989) showed that the samples are highly heterogeneous. Out of 1,000 trials none produced a chi-square value higher than that produced by the observed distribution. The frequencies of Appendix 1 were used to calculate Nei's pairwise genetic distances and construct a Fitch-Margoliash phenogram among populations (fig. 5A). It can be seen from the phenogram that the Black Sea and the three North Aegean samples are grouped together. Adriatic appears to be distinct from the other samples due to its low frequency of phylad A, but more samples are needed to substantiate this difference.

The recognition of two distinct mitotype phylads raises the question of how these phylads are distributed among samples. Table 1 summarises Appendix 1 in terms of the most common mitotypes in each phylad. For 54 individuals, not included in the appendix, the complete five-fold restriction profile could not be obtained, but the phylad could be safely inferred from the scored enzymes. These individuals are included in table 1 (part B). The frequencies of phylads are shown in figure 1. Nei's distances were also calculated considering the mitotypes of each phylad separately and the resulting Fitch-Margoliash phenograms are shown in figure 5B. It can be seen from figure 5B that there is considerable intra-phylad heterogeneity among samples for phylad B, but practically none for phylad A.

The Evolutionary History of Anchovy mtDNA

The marked genetic differentiation among geographic populations of anchovy that we report here is not common among marine fishes with pelagic eggs and larvae (Avise, Reeb, and Saunders 1987). Also un-
usual is the deep division of mitotypes in two phylads and the heterogeneous geographic distribution of these phylads. Avise et al. (1987) have distinguished five phylogeographic categories on the basis of molecular divergence of mitotypes and their geographic distribution. Anchovy fits category II, which is characterized by the co-existence of discontinuous and highly diverged mitotypes in several geographic regions. This pattern implies an allopatric evolution of the diverged mitotypes and a secondary contact within the areas of coexistence. The evidence summarized in table 2 strongly suggests that phylad A has evolved in a population that maintained a much lower effective size than was the case for phylad B. The present-day distribution of mitotypes and the geologic record of the region converge on the hypothesis that the Black Sea is the place of origin of phylad A. The strongest evidence comes from the geographical isolation of the Black Sea from the Mediterranean and the fact that the present population in this sea consists almost exclusively of individuals of phylad A. The mean mtDNA sequence divergence (Nei 1987) among individuals of the Black Sea population (excluding the single type B animal) is 0.0017. If we use the conventional mtDNA evolutionary rate of 2% per million years (e.g., Moritz, Dowling, and Brown 1987), this value translates into an expected average time to common ancestry for the mitotypes of ca. 80,000 years. Since anchovies have

Table 2

| Values of Various Measurements for Phylad A, Phylad B, and the Whole Collection of Mitotypes |
|------------------------------------------|-----------|-----------|
| Phylad A                                 | Phylad B  | Total     |
| Number of haplotypes                     | 23        | 23        | 46        |
| Branching pattern                        | Radial    | Complex   | —         |
| Number of homoplasies                    | 1         | 4         | 21        |
| Number of steps the most distant mitotype is removed from the presumed ancestral mitotype | 2         | 4         | —         |
| Mitotype diversity                       | 0.347     | 0.756     | 0.748     |
| Nucleotide diversity                     | 0.0017    | 0.0051    | 0.0171    |
| Tajima’s D                              | -2.203 ($P < 0.01$)* | -1.289 (NS)* | 0.012 (NS) |
| Tajima’s D corrected for homoplasies     | -2.194 ($P < 0.01$) | -1.106 (NS) | 0.074 (NS) |

* $P =$ probability that the value is nonsignificant.
* NS = nonsignificant.

Notes.—For the calculation of mitotype and nucleotide diversities for phylads all individuals belonging to the same phylad were pooled into one hypothetical population, assuming that the collection of our samples is representative of the global distribution of mitotypes within phylads.
roughly one generation per year (see Magoulas and Zouros 1993), the approximate female effective population size can be estimated at 80,000 (the mean number of generations to common ancestry; see Avise 1989). If the rate of evolution is only a fifth of the above (Rand 1994, and references therein), the effective population size will be 400,000. Even the latter estimate is incompatible with the present huge sizes of anchovy populations in the Black Sea (the usual annual catch is approximately 500,000 mt [GFCM statistical bulletin no. 8, FAO, Rome, 1991, p. 2011, which corresponds roughly to $20 \times 10^9$ individuals). A population bottleneck in the Black Sea could explain this reduction of the effective population size. The large number of mitotypes in phylad A, all of which are recent derivatives of AAAAA, suggests that this was the mitotype that survived the bottleneck and gave rise to other mitotypes, perhaps during an explosive recovery of the population following the bottleneck.

Support for this hypothesis comes from the application of Tajima’s test for neutrality (Tajima 1989), which compares the observed average number of pairwise nucleotide differences ($k$) between haplotypes in a sample to the number ($M$) expected from the number of segregating sites. Under the infinite site mutation model the quantity $D$, defined as $k - M$, is zero. A positive $D$ indicates balancing selection or population subdivision. A negative $D$ indicates directional selection, recent population bottleneck, or background selection of slightly deleterious alleles (see also Rand, Dorfsman, and Kann 1994). The results from the application of Tajima’s test are shown in table 2. For phylad B the test gives a negative but not significant $D$. For phylad A it gives a negative and significant $D$. The significance remains after correction for the incidence of homoplasy (by estimating the number of restriction site differences between the mitotypes from the network of fig. 3 rather than from the matrix of presence/absence of restriction sites). The negative value of phylad A provides further support for a recent bottleneck in the Black Sea, although it could also be explained by a recent selective sweep.

The Population History of the Anchovy in the Mediterranean

The existence of two discontinuous mitotype phylads and the pattern of geographic distribution of mitotypes in anchovy (figs. 1, 3, and 5) can be understood in terms of past geological events and current hydrographic patterns in the Mediterranean. Specifically, we will provide evidence to support the following points: (a) The anchovy entered the Mediterranean about 5 million years ago, when the present opening of Gibraltar was established. (b) Phylad A originated in the Black Sea, which was colonized through the Mediterranean at the end of the Pliocene. (c) The current high frequency of phylad A in the Mediterranean and the Atlantic is due to a massive outflow from the Black Sea. (d) Current hydrographic forces contribute to the dynamic maintenance of the geographical structure of mitotypic diversity that originated from these geologic events.

Colonization of the Mediterranean and the Black Sea

Anchovy is not included in the list of fish species cited by Por and Dimentman (1985) as possible Messinian relics in the Mediterranean, i.e., species that survived the Messinian salinity crisis in situ. It must have entered the Mediterranean from the Atlantic during the Trubi transgression, at the beginning of the Pliocene, when the communication between Atlantic and Mediterranean was opened again and marine conditions were re-established permanently in the Mediterranean.

It is known that 80% of the Black Sea fish species are of Mediterranean origin, the rest tracing their origin from Paratethys, the body of water which during Oligocene and Miocene covered present-day central Europe and extended beyond the Black and Caspian Seas (Ekman 1968; Quignard 1978). The view that the Black Sea anchovy is a relic of the Tertiary Period was entertained in the past, but is no longer accepted (Zenkovich 1963, p. 435). This view is also supported by the fact that the implied long separation of the Atlanto-Mediterranean and Paratethyan stocks (as far back as 15-20 million years) would have caused a higher divergence between phylads A and B, even under the assumption of a slow rate of mtDNA evolution for fish. We conclude that anchovy has been one of the many species that invaded the Black Sea during the late Pliocene, when this sea was connected to the Mediterranean.

As mentioned, several lines of evidence suggest that the Black Sea is the place of origin of phylad A and that a population bottleneck resulted in the "star" phylogeny and low diversity of this phylad. An explanation for this presumptive bottleneck can be found in the glaciation history of the Black Sea in the Pleistocene. During glacial maxima, the anchovy populations must have suffered drastic reductions in the isolated Black Sea, as this species cannot reproduce in temperatures below 13°C (Demir 1963; Palomera 1992). The most recent bottleneck in the Black Sea may have occurred 20,000–30,000 years ago, during the height of the last Glacial (Por 1989, pp. 24 and 27).
Massive Exit of Anchovy from the Black Sea into the Mediterranean

Phylad A is now present in all localities sampled. Several observations suggest that this is the result of a recent exit of anchovy from the Black Sea into the Eastern Mediterranean rather than of gradual diffusion via the Aegean Sea. These include the relatively young age of phylad A, the high homogeneity of intra-phylad A composition of samples (fig. 5R), the fact that type AAAAA is by far the most common mitotype of phylad A in all samples, and that the minor A types found in all samples are recent derivatives of this type. The sharp drop of the phylad’s frequency over the relatively small distance separating the North from the Central Aegean suggests that, under the present regime of hydrographic conditions in the Aegean, gene flow from the North Aegean into the main Mediterranean is not free.

Massive water spills from the Black Sea into the Mediterranean during the interglacial periods and especially the deglaciation events (or terminals) are accepted facts in physical oceanography. According to Herman (1988), “following warming and subsequent massive deglaciation, sea level rose; when the sea stand reached the Bosporus sill (~36 m) the connection between the Mediterranean and the Black Sea was re-established and the low salinity Black Sea water spilled over into the Mediterranean”, resulting in the flooding of the eastern Mediterranean by large volumes of low salinity water. Other authors (Por 1989, and references therein) write that by the end of the last Glacial and especially during its termination event (9,000–12,000 years ago), the Levantine basin was hydricographically separated from the Ionian basin by a tongue of diluted and cold (13°C) water penetrating from the direction of the Aegean Sea and reaching the African coast during the winter. We note that the population recovery from the bottleneck in the Black Sea and the outflow into the Mediterranean are not very distant in terms of geologic time, as both must have occurred during the warming period after the last glacial maximum.

Current Factors Affecting the Phylogeography of Anchovy in the Mediterranean

Phylad B is practically absent from the Black Sea, while it has a frequency of 15% in the neighbouring North Aegean. This can be explained by postulating either selection against phylad B in the Black Sea or that there is a blockade in the anchovy gene flow from the Aegean to the Black Sea. Everything that we know about the hydrography of the passage joining the two seas and the biology of the species is consistent with a view of unidirectional gene flow. The elongated (ca. 31 km), narrow (no more than 1.5 km wide in most of its length and 550 m at its narrowest point), and shallow (ca. 30 m at its shallowest point) Straits of Bosphorus heavily restricts free water circulation and is to a large extent responsible for the preservation of the physicochemical and biotic differences that distinguish the Black Sea from the Mediterranean Sea. As an example, salinity is as high as 38–39‰ in the Aegean Sea, but only 17–18‰ in the Black Sea. In the Straits of Bosphorus there is a surface current of cold, low-salinity (17‰) water flowing from the Black Sea towards the Marmara Sea and a subsurface current of warmer, high-salinity (38.5‰) water flowing in the opposite direction. In volume, the former is twice as large (Vergnaud-Grazzini 1985). Eggs and larvae of anchovy are primarily distributed in shallow layers, mainly in the upper 10 m (e.g., Demir 1963; Palomera 1991). The anchovy spawns in most parts of the Black Sea, but to a greater extent in the southern regions (Nierman et al. 1994). In the mouth of Bosphorus in the Black Sea the planktonic stages of the anchovy must be almost exclusively at the upper part of the water column that moves swiftly southward. The lower northward-moving current contains no immature stages and, besides, upon entering the northern end of the straits, it sinks precipitously into the hostile oxygen-deprived deep water of the Black Sea (Tulunazin 1985). Deniz (1963) suggests that adult anchovy exit the Black Sea into the Sea of Marmara in the fall and move back in the Black Sea in the spring, but provides no evidence for this claim. In contrast, Ivanov and Beverton (1985) provide detailed information about anchovy migration within the Black Sea but do not include the anchovy among the species that migrate as adults through the Bosphorus straits. Kocatas et al. (1993) list a number of species that move from the Sea of Marmara to the Black Sea, and, again, anchovy is not included among them. It seems, therefore, that the unidirectional flow of immature stages from Black Sea to the Aegean is not counterbalanced by active movement of adults in the opposite direction.

The northern Aegean is the Mediterranean’s portal area of Black Sea waters. Three samples of anchovy from this area, taken from two different stations at two seasons 3 years apart, produced a consistently high frequency (85%) of phylad A (fig. 1). To a large extent this is the result of gene flow from the Black Sea into the Aegean. The large drop of this frequency in the Central Aegean is suggestive of a second barrier to free gene flow, one that geographically coincides with the boundary of the northern and southern basins of the Aegean Sea. The subdivision of the Aegean into two basins has been long recognized. The northern basin is characterized by a continental shelf favouring neritic
and meroplanktonic forms (Moraitou-Apostolopoulou 1985). This basin is under the influence of cool, low-salinity Black Sea waters which are entrained into a cyclonic circulation affecting the northern and western parts of the Aegean and causing an ecological isolation of the northern basin from the southern basin (Theocharis et al. 1993) (see fig. 6). In the southern basin the continental shelf is very limited and the area has the characteristics of a pelagic zone with a fauna that resembles the Western and Eastern Mediterranean (Moraitou-Apostolopoulou 1985).

A similar north-south subdivision is known to exist in the Adriatic Sea. According to Artegiani et al. (1993), the northern part is under the strong influence of fresh water inflow, mainly from the Po River, whereas a topographically controlled cyclonic gyre, situated in the South Adriatic pit, partially isolates the northern Adriatic from the influence of the Mediterranean (fig. 6). This isolation is also reflected by the anchovy distribution within the Adriatic. In the southern Adriatic anchovy is 80 times less abundant than in the northern Adriatic (Casavola et al. 1988). The low frequency of phylad A in the northern Adriatic (14%) can be seen as a result of a historic event that prevented its entrance when this phylad became abundant in the rest of the Mediterranean and of current conditions that prevent the free movement between the South and North Adriatic.

It is known that the Straits of Gibraltar does not represent a zoogeographic barrier for most marine species, with the result that the Mediterranean and the adjacent Atlantic coasts support very similar faunas and form together what is recognized as the Atlanto-Mediterranean province (e.g., Ekman 1968; Tortonese 1964). The homogeneity of the sample from the Bay of Biscay with those from the Mediterranean in terms of phylad A and B proportions is consistent with the free movement of anchovy through the Straits of Gibraltar. The difference of this sample from the others in terms of intra-phylad B composition remains unexplained. This is, however, a minor difference, since all four major B mitotypes of the Mediterranean samples were found in high numbers in the Bay of Biscay and BBBB was also the most common type in this sample.

An electrophoretic study of populations of anchovy from the North Aegean and Ionian Seas (Spanakis, Tsimenides, and Zouros 1989), produced results that varied among loci. All samples were statistically homogeneous for one of the four polymorphic enzyme loci (Est). The samples from the Aegean were highly heterogeneous at two loci (Ldh and Pgm), whereas the samples from the Ionian were not. All the samples from the Aegean were homogeneous for the fourth locus (Pgi) and so were the samples from the Ionian, but the two seas were highly heterogeneous at this locus. These results are consistent with recent suggestions that geographic patterns of allozyme variation may be influenced more by locus specific selection than by forces of gene flow and random drift and may, thus, be less useful in reconstructing the phylogenetic history of a species (Karl and Avise 1992). However, the fact that the only enzyme locus that joined the samples from the same sea into a homogeneous group also recognized the two seas as genetically heterogeneous populations is consistent with our present finding of a large mtDNA differentiation between North Aegean and Ionian samples.

The two anchovy phylads were found to coexist over a large and ecologically diverse area extending from the North Aegean to the Bay of Biscay. This suggests that the phylads A and B are not coincident with geographical races or subspecies, which has been suggested by several authors in the past (Spanakis, Tsimenides, and Zouros 1989). Both phylads were also collected from the same schools in all these localities (except the Black Sea). This apparent lack of ecological or ethological differentiation between animals belonging to different phylads suggests that the molecular discontinuity and the marked geographical differences of the various mitotypes that we have described here are not in themselves evidence for the existence of two incipient taxa. At present, we have no evidence to suggest that the history of the anchovy that resulted in this rich and informative pattern of mitotype diversity has led to the development of any form of reproductive isolation between local populations or between mitotypic variants occurring within the same population.

Acknowledgments

We are grateful to a number of people who helped us with the collection of the samples: Prof. P. Kolarov,
Dr. I. Dobrovolo, K. Michallov, Dr. G. Kotoulas, Dr. T. Patarnello, Dr. G. Goulielmos, Dr. A. Machias, and A. Argyrokastritis. We thank K. Ekonomaki and V. Terzoglou for technical assistance, Dr. G. Kotoulas for discussion and comments during the whole course of the work, and Dr. P. Bentzen for providing the clones of shad mtDNA. A.M. was supported by a graduate scholarship from the Institute of Marine Biology of Crete and the Institute of Molecular Biology and Biotechnology (Foundation for Research and Technology—Hellas). This research was supported by a contract from the Commission of the European Communities.

**APPENDIX 1**

Distribution of Mitotype Frequencies in the Samples. Several Individuals were Heteroplasmic. Heteroplasmy in Anchovy Has Been Discussed Elsewhere (Magoulas and Zouros 1993). For the Purpose of the Present Analysis Only the “Stronger” of the Two Mitotypes Found in Heteroplasmy Was Taken into Account.

<table>
<thead>
<tr>
<th>BLS</th>
<th>KAV1</th>
<th>KAV2</th>
<th>PAG</th>
<th>SAR</th>
<th>PAT1</th>
<th>PAT2</th>
<th>ADR</th>
<th>LION</th>
<th>BISC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>AAAAA</td>
<td>58</td>
<td>90</td>
<td>14</td>
<td>17</td>
<td>30</td>
<td>17</td>
<td>5</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>2.</td>
<td>AAAAG</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3.</td>
<td>AAAAH</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4.</td>
<td>AAAAI</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5.</td>
<td>AAAAJ</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6.</td>
<td>AAAAL</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7.</td>
<td>AAAAB</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8.</td>
<td>AAACA</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9.</td>
<td>AAADA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10.</td>
<td>AAABB</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11.</td>
<td>AAEEA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12.</td>
<td>AAFAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13.</td>
<td>AAAGA</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14.</td>
<td>AHAHA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15.</td>
<td>AABAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16.</td>
<td>AGAAG</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17.</td>
<td>AIAAA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18.</td>
<td>AIJAA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19.</td>
<td>CAAAA</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20.</td>
<td>DAAAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21.</td>
<td>EAAAA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22.</td>
<td>EADDA</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>23.</td>
<td>GAABA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24.</td>
<td>RABB</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25.</td>
<td>BABC</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26.</td>
<td>BBABC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>27.</td>
<td>BBBBB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28.</td>
<td>BBBBA</td>
<td>0</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>15</td>
<td>30</td>
<td>10</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>29.</td>
<td>BBBBC</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>30.</td>
<td>BBBBD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>31.</td>
<td>BBBDE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>32.</td>
<td>BBCBA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>33.</td>
<td>BBCCB</td>
<td>0</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>28</td>
<td>9</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>34.</td>
<td>BCBBC</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>35.</td>
<td>BBBCD</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>36.</td>
<td>BBCBK</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>37.</td>
<td>BBCBM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>38.</td>
<td>BBCDN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>39.</td>
<td>BCBBB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40.</td>
<td>BCDDB</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>41.</td>
<td>BDCBF</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>42.</td>
<td>BBFRR</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>43.</td>
<td>BEBB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>44.</td>
<td>BFBBB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45.</td>
<td>DIDD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>46.</td>
<td>FBBBC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Total ........ 70 141 57 20 59 118 54 57 50 47 673


and between populations from restriction data. Genetics 125:873–879.


RICHARD G. HARRISON, reviewing editor

Accepted September 13, 1995