Evidence for microsatellite hitchhiking selection in European sardine (*Sardina pilchardus*) and implications in inferring stock structure

PANAGIOTIS KASAPIDIS [1], ALEXANDRA SILVA [2], GIANPAOLO ZAMPICININI [1,3]
and ANTONIOS MAGOULAS [1]

[1] Hellenic Centre for Marine Research, Institute of Marine Biology and Genetics, P.O. Box 2214, 71003 Heraklion, Greece. E-mail: kasapidi@her.hcmr.gr
[2] Instituto de Investigação das Pescas e do Mar (INRB-IPIMAR), Av. Brasilia s/n, 1449-006 Lisboa, Portugal.
[3] Dipartimento di Biologia Animale e dell'Uomo, Universita di Torino, Italy.

SUMMARY: The genetic structure of the European sardine (*Sardina pilchardus*) was assessed throughout its geographic range using five microsatellite loci. One of the loci seemed to be under hitchhiking selection and exhibited a latitudinal cline along the eastern Atlantic, with abrupt change in allele frequencies from the Alboran Sea to the western Mediterranean and from the east Atlantic coast to the Azores and Madeira. This pattern was very similar to that previously described for the allozymic locus SOD* and these 2 loci could be linked. A Bayesian analysis of environmental factors with the genetic data indicated temperature as a potential selection factor. Selection pressure may be stronger at the southern limit of sardine distribution, because heterozygosity of the non-neutral locus was much lower there. The abrupt change in allele frequencies of the non-neutral locus in certain regions seem to be related more to strong barriers to gene flow, which were not evident for neutral loci, than to abrupt changes in selection pressure. These areas of discontinuity provide a guideline to define and delineate genetic stocks and are generally consistent with areas of phenotypic change in sardine, but they are not in concordance with the currently recognized morphological subspecies.

Keywords: sardine, genetic structure, microsatellite DNA, selection, Atlantic, Mediterranean.

RESUMEN: EVIDENCIA PARA SELECCIÓN HITCHHIKING DE MICROSATÉLITE EN SARDINA EUROPEA (*SARDINA PILCHARDUS*) Y LAS IMPLICACIONES PARA INFERIR LA ESTRUCTURA DEL STOCK. – La estructura genética de la sardina europea se evaluó a lo largo de su área de distribución geográfica, con cinco loci de microsatélites. Uno de los loci parecía estar bajo selección por *hitchhiking* y exhibió una clina latitudinal a lo largo del Atlántico oriental, con el cambio abrupto en las frecuencias alélicas del Mar de Alborán al Mediterráneo Occidental y de la costa Oriental del Atlántico hacia las Azores y Madeira. Este patrón es muy similar al descrito previamente para el locus alozímico SOD*, por lo que estos dos loci podrían estar ligados. Un análisis bayesiano de factores ambientales con los datos genéticos indicó que la temperatura puede ser el factor de selección. La presión de selección puede ser más fuerte en el límite sur de distribución de la sardina, dado que la heterozigosis del locus no neutro fue allí mucho más baja. El abrupto cambio en las frecuencias alélicas de locus no neutro parece estar relacionado más con barreras fuertes al flujo de genes, que no se evidenciaron en el caso de los loci neutrales, que a cambios en la presión de selección. Estas áreas de discontinuidad proporcionan una guía para definir y delinear *stocks* genéticos y son generalmente consistentes con las áreas de cambio fenotípico de la sardina, pero no están conformes con las subespecies morfológicas actualmente reconocidas.

Palabras clave: sardina, estructura genetica, ADN microsatelite, selección, Atlántico, Mediterráneo.

INTRODUCTION

A large number of genetic population studies in marine fish have demonstrated a general pattern of weak intraspecific genetic structure (Waples 1998). This pattern has been attributed mainly to the high gene flow, facilitated by the high dispersal abilities in larval and adult stages of many marine fish and the ap-

parent lack of physical barriers in the marine environment. Moreover, the effect of genetic drift in promoting differentiation is diminishing due to the very large population sizes (Allendorf and Phelps 1981). Hence, the very low or statistically non-significant values of F_{ST} usually found in marine fish may be inconclusive about the recent history or present levels of gene flow (Carvalho and Hauser 1998, Ward 2000). This is a major problem in fisheries management since lack of genetic differentiation may indicate the presence of a single *genetic stock* (i.e. a unit that is more or less reproductively isolated from another, and thus may react independently to exploitation) (Ovenden 1990), but does not exclude the presence of more than one *harvest stock* (i.e locally accessible fish resources in which fishing pressure on one resource has no effect on the abundance of fish in another contiguous resource) (Gauldie 1988). The latter units are currently of interest to fisheries managers, but this concept does not imply any genetic or phenotypic differences between stocks.

Most genetic studies on fish population structure are typically based on markers that are considered neutral, and the underlying genetic theory is based on the interplay between gene flow, drift and mutation. However, like all other marine organisms, fishes are permanently exposed to external physical factors such as temperature, salinity, and other environmental conditions and to anthropogenic factors such as fishing pressure. These factors may promote divergent selection by causing differential survival or mortality to animals with different genotypes in different environments (Guinand *et al.* 2004). Divergent selection on a locus will cause its adaptive divergence to a degree that often reflects a balance between the strength of selection and rates of gene flow (see Nosil *et al.* 2009 and references therein). Furthermore, selection on one locus can also strongly affect the allele frequency at physically close or "tightly linked" loci, even when the latter are selectively neutral—a process known as genetic hitchhiking (Maynard Smith and Haigh 1974).

There are studies in marine fish in which adaptive divergence has been detected at the molecular level (see Guinand *et al.* 2004 and Nielsen *et al.* 2009 for reviews). Well-known cases are the pantophysin gene *Pan*I in Atlantic cod (Pogson 2001), the heat-shock cognate protein gene *Hsc*70 in European flounder (Hemmer-Hansen *et al.* 2007) and the *Ldh*-B gene in killifish (Schulte *et al.* 2000). More recently, genome scans using several genetic markers of different types and a variety of methods have identified several candidate loci for selection (e.g. Mäkinen *et al.* 2008, Moen *et al.* 2008). However, in most cases the selection agent and the underlying mechanism have not been demonstrated.

European sardine is a small pelagic clupeoid fish inhabiting the Mediterranean Sea and part of the eastern Atlantic from the North Sea to Senegal, with peripheral populations around the Azores, Madeira and the Canary islands (Parrish *et al.* 1989). Two subspecies of sardine have been proposed, based on phenotypic variation mainly in gill raker counts and head length: *S. p. pilchardus* in the eastern Atlantic from the North Sea to southern Portugal, and *S. p. sardina* in the Mediterranean Sea and off the northwest African coast (Andreu 1969, Parrish *et al.* 1989). Sardine supports important fisheries in the northeast Atlantic and in the Mediterranean Sea, with approximately 130000 tons landed on the European coast, 660000 tons landed on the African coast and 80000 tons landed in the Mediterranean area (GFCM 2006, FAO 2008, ICES 2009). Despite the importance of the sardine fishery, stock delineation for management purposes is still a matter of debate throughout the distribution range of the species (ICES 2006, FAO 2008).

There have been a number of studies addressing the genetic structure of sardine in different parts of its global distribution, using different molecular markers such as allozymes (Spanakis *et al. 1*989, Chlaida *et al. 2*006, Laurent *et al.* 2007, Chlaida *et al.* 2009), mitochondrial DNA (Tinti *et al.* 2002, Atarhouch *et al.* 2006) and microsatellites (Gonzalez and Zardoya 2007a). The results of these studies, though not completely congruent, suggest a very weak genetic structure, with the exception of the allozymic locus SOD*, which could be subject to selection (Chlaida *et al.* 2006; Laurent *et al.* 2007).

In this study, we assessed genetic differentiation in sardine using 5 microsatellite loci. One of the loci (Sp22) appeared to be under selection and exhibited a geographic pattern that closely resembles that of allozymic locus SOD* (Laurent *et al.* 2007). We further checked for correlation between genetic variation and environmental factors (latitude, longitude, temperature and salinity). Patterns of neutral and non-neutral variation from this and previous studies are compared and the implication of these results together with published non-genetic evidence in inferring sardine stock structure are discussed.

MATERIALS AND METHODS

Sampling

Twenty-one sardine samples of 40-88 individuals each (73 individuals per sample on average, 1540 individuals in total) were used in the present study (Table 1, Fig. 1). The sampling scheme covered almost the entire distribution range of sardine. The samples had been collected within the framework of SARDYN and other international projects during the period 1999-2004. Seventeen samples had been collected during the spawning season and 4 during the feeding season. Some of the samples had been collected from the same region in different years and/ or in different seasons, in order to check the temporal stability of the genetic structure and to account for random sampling errors.

TABLE 1. – Information on sardine samples used in the present study. N, sample size; Ho and He, observed and expected heterozygosity per sample for locus Sp22, and average for loci Sp2, Sp7, Sp8 and SpI5; S, salinity in psu; T, mean spawning temperature.

Sample code	Sampling locality	Sampling time	Geographical coordinates	N	H_o/H_e for Sp22	H_o/H_e for Sp2,7,8,I5	S	T(°C)
1-NSea04	North Sea	Feb 2004	54°47.0'N, 01°17.4'E	84	0.71/0.77	0.94/0.93	34.5	13.62
2-Channel03	Western Channel	Jan 2003	50°01.0'N, 05°22.0'W	82	0.74/0.75	0.93/0.92	35.4	13.76
3-Celtic00	Celtic Sea	Mar 2000	48°30.0'N, 05°26.0'W	67	0.67/0.73	0.94/0.93	35.4	13.76
4-FraAtl03	Bay of Biscay	Jun 2003	45°01.7'N, 01°53.2'W	61	0.64/0.67	0.94/0.93	35.5	15.73
5-Cant00	Bay of Biscay	Apr 2000	43°40.3'N, 06°30.1'W	76	0.66/0.71	0.94/0.93	35.5	13.99
6-NPort00	North Portugal	Mar 2000	40°59.8'N, 08°42.8'W	88	0.60/0.64	0.93/0.93	35.7	15.59
7-NPort03	North Portugal	Feb 2003	41°05.0'N, 08°59.9'W	85	0.72/0.71	0.93/0.93	35.8	15.59
8-FNPort03	North Portugal	Aug 2003	Matosinhos port	76	0.46/0.50	0.93/0.93	n/a	n/a
9-SPort03	South Portugal	Mar 2003	38°32.0'N, 09°13.6'W	69	0.67/0.68	0.93/0.92	36.1	15.59
10-Cadiz00	Gulf of Cadiz	Mar 2000	36°33.3'N, 06°23.6'W	67	0.73/0.70	0.93/0.92	36.3	17.00
11-Cadiz03	Gulf of Cadiz	Mar 2003	36°30.2'N, 06°30.9' W	83	0.64/0.67	0.94/0.93	36.3	17.00
12-FCadiz03	Gulf of Cadiz	Aug 2003	Cadiz port	87	0.69/0.71	0.93/0.93	n/a	n/a
13-AlbEsp99	Alboran Sea	Dec 1999	36°43.0'N, 03°04.6'W	82	0.66/0.72	0.93/0.92	38.2	15.45
14-AlbMor04	Medit. Morocco	Oct 2004	n/a	40	0.75/0.74	0.92/0.91	38.2	16.29
15-MorAtl00	Morocco, Casablanca	Jan 2000	n/a	86	0.57/0.59	0.93/0.92	36.4	17.00
16-Barc99	Barcelona	Dec 1999	40°50.2'N, 01°12.0'E	48	0.63/0.77	0.93/0.92	37.9	13.80
17-Aegean03	Aegean Sea	Jun 2003	40°50.8' N, 24°23.9'E	76	0.66/0.70	0.93/0.93	39.1	13.29
18-Azores00	Azores, St Miguel Isl.	May 2000	n/a	78	0.69/0.76	0.93/0.92	n/a	n/a
19-Azores03	Azores	Nov 2003	n/a	80	0.66/0.73	0.93/0.92	n/a	n/a
20-Madeira03	Madeira	Mar 2003	n/a	47	0.66/0.68	0.93/0.92	n/a	n/a
21-Maur03	Mauritania	Dec 2003	20°38.0'N, 17°12.0'W	78	0.21/0.22	0.93/0.92	36.1	17.61
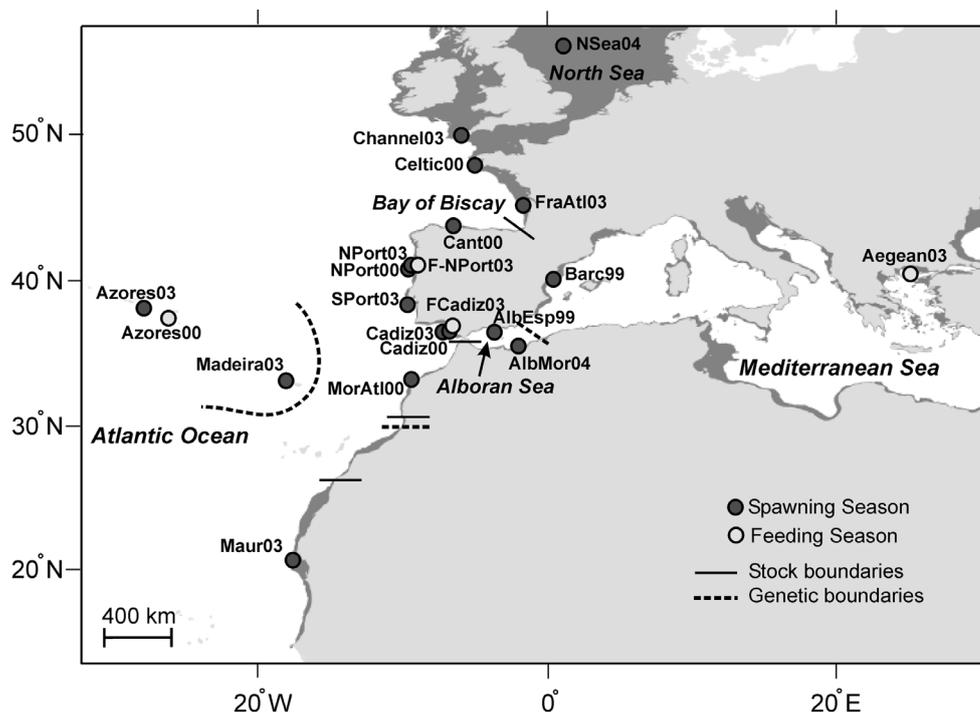
FIG. 1 – Map of sampling localities with the geographic distribution of sardine from FAO (http://www.fao.org/figis/geoserver/factsheets/species.html). Sample abbreviations are as in Table 1. Stock boundaries in the Atlantic are indicated, as are the genetic boundaries identified in this study and in Chlaida et al. (2006).

Microsatellite analysis

Total DNA was extracted from muscle tissue preserved in 95% ethanol, using the salt protocol of Miller et al. (1988). Samples were genotyped for 5 microsatellite loci (Sp2, Sp7, Sp8, SpI5, Sp22). Three of the loci (Sp2, Sp7, Sp8) were isolated from a genomic library screened for $(GT)_n$ microsatellite repeats, following the protocol described in Batargias et al. (1999). Locus

SpI5 was developed from a genomic library enriched for $(AGAT)_5$ repeats, following a modified enrichment protocol (Tsigenopoulos et al. 2003), while locus Sp22 was kindly provided by Dr. F. Tinti, University of Bologna. Characteristics of the loci used are given in Table 2.

Each locus was amplified using the polymerase chain reaction (PCR). PCR reactions were performed in a 10 μL volume and contained about 40 ng of genomic

TABLE 2. – Characterization of the microsatellite markers of sardine used in this study. N_A, number of alleles; T_A, annealing temperature in °C; $MgCl_2$, concentration in mM; H_o and H_e, average observed and expected heterozygosity from the 21 samples of the present study.

Locus	Primers 5'-3'	Repeat motif	Size range	N_A	T_A/MgCl$_2$	H_o/H_e	F_{ST}	F_{IS}	Accession number
Sp2	CGAGGCCTGATAGAAACCC AACCACGGTCAGTTCTCCAG	$(AG)_8/(TG)_{16}$	138-254	58	56/1.0	0.93/0.95	0.0017	0.0362	AJ639616
Sp7	GCACAGGCGCTTACACAC TGTGACACCAGGCAGAGC	$(AC)_{14}$	124-212	41	62/1.5	0.93/0.92	0.0009	0.0000	AJ639618
Sp8	ACGTCACAGTTCCCCACTG ACTGGCTGAGGAGGATGATG	$(CA)_{22}$	109-211	60	62/1.5	0.93/0.94	0.0010	0.0238	AJ639619
SpI5	TGGCTGTGCATGTAAGTCTGT CAATTTTGCCAGTTGTTTAGTC	$(TATC)_8TC$ $(TATC)_2$	135-197	31	49/2.5	0.76/0.88	0.0012	0.1443	HM031962
Sp22	GGTCACGTTGGGTATTTGCTAT GTCTGGAACTGGGTGCTCTC	$(CA)_3GA(CA)_6$ $GA(CA)_8$	194-232	19	49/2.5	0.64/0.67	0.0895	0.0579	AY241284

DNA, 1 X PCR buffer, 1.0-2.5 mM $MgCl_2$ (Table 2), 0.2 mM dNTPs, 0.5 μM of each primer and 0.2 U *Taq* DNA polymerase (Gennaxon). The following PCR profile was used: 2 min at 94°C, 35 cycles consisting of 45 sec at 94°C, 45 sec at annealing temperature (Table 2), 45 sec at 72°C, and a final extension at 72°C for 10 min. Loci Sp2, Sp7 and Sp8 were analyzed on a Vistra automated sequencer and typed manually by determining allele size using allele fragments of known size. For loci SpI5 and Sp22, microsatellite DNA fragments were separated on a Base Station automatic sequencer (MJ Research) and genotypes were scored with the Cartographer software (MJ Research).

Statistical analyses

Microsatellite data were checked for null alleles, stuttering and large allele dropout with the software MICRO-CHECKER 2.2.1. (van Oosterhout *et al.* 2004). Allelic and genotypic frequencies and observed and expected heterozygosity values were computed using GENETIX 4.05 (Belkhir 2000) and FSTAT 2.9.3 (Goudet 1995). Deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium was assessed using GENEPOP 3.4 (Raymond and Rousset 1995). The significance levels were adjusted using sequential Bonferroni corrections (Rice 1989) when multiple tests were applied.

Significance of pairwise differences in allelic and genotypic frequencies was tested using Fisher's exact test as implemented in GENEPOP. Pairwise F_{ST} between samples (Weir and Cockerham 1984) were estimated using ARLEQUIN 3.0 (Excoffier *et al.* 2005); the p values were obtained by performing 1000 permutations of the data.

The hierarchical structure of the samples was investigated using the SAMOVA.1.0 program (Dupanloup *et al.* 2002). This method is an extension of the analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) in that it incorporates geographical coordinates of the samples in the analysis and seeks to define groups of populations that are geographically homogeneous and maximally differentiated from each other. An AMOVA was also performed using ARLEQUIN, to test whether a population structure separating the 2 subspecies of sardine was significant. Samples from southern Portu-

gal to the North Sea were considered as *S. p. pilchardus* and those from the Mediterranean Sea and off the northwest African coast as *S. p. sardina*. A Principal Component Analysis (PCA) on gene frequency data was conducted using PCAGEN software (Goudet. J, http://www2.unil.ch/popgen/softwares/pcagen.htm).

The Mantel test was used to test for correlation between genetic and geographical distances, as implemented in the IBDWS software, v3.15 (Jensen 2005). The logarithm of geographical distances between sample localities (measured in km as the shortest distance along coastal waters of depths less than 200 m, apart from the Azores and Madeira, where the shortest distance from the coast was used) was regressed against the logarithm of Rousset's distances ($F_{ST}/(1-F_{ST})$). Significance was assessed with 1000 permutations.

Marker neutrality

Neutrality of polymorphic loci was tested using 2 approaches: a) the simulation based approach of Beaumont and Nichols (1996) implemented in the program FDIST2 and b) the Bayesian regression approach implemented in BAYESCAN (Foll and Gaggiotti 2008). The FDIST2 program uses coalescent simulations to generate a neutral joint distribution of F_{ST} and heterozygosity; loci with the highest ratios of F_{ST} versus heterozygosity are candidates for having experienced selection. Coalescent simulations were performed using 21 samples and a sample size of 50, assuming an island model (with 100 islands) and an infinite alleles mutational model. The mean F_{ST} value was used in addition to other values close to the mean in order to get half the data points above and half below the median line, as suggested in the software manual.

BAYESCAN estimates the probability that a locus is under selection by calculating a Bayes factor, which is simply the ratio of the posterior probabilities of 2 models (selection and neutral) given the data. A Bayes factor above 100 (log10>2) is interpreted as "decisive evidence" of different statistical support for the two models and corresponds to posterior probabilities between 0.99 and 1. We ran 100000 iterations (sample size of 5000 and thinning interval of 20), following 10 pilot runs of 5000 iterations and an additional burn-in of 50000 iterations. A Bayes factor of infinity,

corresponding to a posterior probability of 1, was assigned a log10 value of 5. In order to alleviate potential problems arising from allele frequency correlations between samples (i.e. high rate of false positives), especially when different samples are drawn from the same population (Excoffier *et al.* 2009), we ran a global analysis with samples grouped into 5 populations defined by AMOVA (see Results), as well as different combinations of samples belonging to different genetic groups.

The population structure revealed by locus Sp22 (see Results) had striking similarities to that of the allozymic locus SOD* in Laurent *et al.* (2007), suggesting a possible linkage of these 2 loci. Unfortunately, we could not directly check for linkage disequilibrium because different samples had been used for microsatellite and allozyme genotypying. However, we tested for correlation between allele frequencies of the 2 loci in samples collected from the same geographical areas.

Correlation with environmental parameters

The hierarchical Bayesian method of Foll and Gaggiotti (2006), as implemented in the GESTE software, was used to test for the effect that different environmental factors may have on the genetic structure of sardine. This method estimates F_{ST} values for each local population using the approach first proposed by Balding and Nichols (1995) and relates them to environmental factors using a generalized linear model. The consideration of n factors and their interaction leads to 2^n alternative regression models. The method provides posterior probabilities for each one of them using a reversible jump Markov chain Monte Carlo (MCMC) approach. The model with the highest posterior probability is the one that best explains the data. We used 10 pilot runs of 5000 iterations to obtain the parameters of the proposal distribution used by the MCMC approach, followed by an additional burn-in of 50000 and a thinning interval of 50. All estimates were obtained using a sample size of 20000.

We examined the effect of 4 different factors (latitude, longitude, mean temperature during spawning season and mean salinity) that are related to the physical characteristics of the sampling locations, following the approach of Gaggiotti *et al.* (2009). We chose to use surface temperature at spawning time because it is as close as possible to the most critical time of the sardine life cycle, when the eggs and larvae are found close to the surface and have little option for actively avoiding unfavourable conditions, by moving deeper as the adults do. In order to separate the correlation of temperature and salinity with latitude and longitude which was evident in our data (not shown), we transformed temperature and salinity as the mean absolute difference between the value of the factor at the sampling locality and the average value of all sampling localities. We did the analysis by including all loci and by including only neutral loci in order to check whether there was a link between the outlier locus and any of the environmental factors. These analyses were applied to 16 samples for which temperature data on spawning season existed. The 2 samples collected at the port were excluded, as were samples from the Azores and Madeira, where there are no data for spawning in the literature.

Mean temperature during the spawning season was taken from Coombs *et al.* (2006). For areas for which there were no available estimations, it was calculated from SST monthly data retrieved from the Pathfinder AVHRR dataset for the period 1985-2001 for a 4 x 4 km box around the grid points. Annual mean surface salinity was retrieved from NOAA/NCEP/EMC/CMB/GODAS (Global Ocean Data Assimilation System) for the period 1980-2009 (http://iridl.ldeo.columbia.edu/docfind/databrief/cat-ocean.html).

RESULTS

A total of 1540 individuals were fully genotyped for 5 microsatellite loci, with no missing data. All 5 loci were highly polymorphic with number of alleles ranging from 19 for locus Sp22 to 60 for locus Sp8 (Table 2). Sixteen out of 105 randomization tests for heterozygote deficiency resulted in significant outcomes, in one sample for locus Sp22 and 15 samples for locus SpI5. None of the tests for heterozygote excess was significant. MICRO-CHECKER detected problems that could be attributed to null alleles in 26 out of 105 cases, mainly in locus SpI5 (17 samples) and in locus Sp2 (6 samples). All samples had similar levels of heterozygosity in each locus, with the exception of the sample from Mauritania, which exhibited low heterozygosity in locus Sp22 (Table 1). Overall genetic divergence between samples (F_{ST}) was equal to 0.017. Contribution of each locus to genetic divergence was unequal, with locus Sp22 showing an F_{ST} value of 0.089, while the other 4 loci had F_{ST} values between 0.001 and 0.002 (Table 2).

Pairwise F_{ST} estimates showed that samples from Mauritania (Maur03), Madeira (Madeira03), the Azores (Azores00 and Azores03), Barcelona (Barc99) and the Aegean (Aegean03) were significantly different from each other and from the rest of the samples, with the following exceptions: Azores03 vs. Madeira03, Barc99 and Aegean03; Barc99 vs. Aegean03, AlbMor04 and NSea04; and FNPort03 vs. Maur03. All the other samples were not significantly different except for FNPort03 vs. AlbEsp99, Nport03, Channel03 and Nsea04. The sample from Mauritania was the most differentiated, with F_{ST} values ranging from 0.006 (Maur03 vs. FNPort03) to 0.114 (Maur03 vs. Madeira03). When locus Sp22 was removed from the analysis, the samples showed very little differentiation, except for the Madeira03 sample being different from all the others and the Azores00 sample being different from all others but the Azores03 sample. Fisher's exact tests for allelic frequency differentiation across
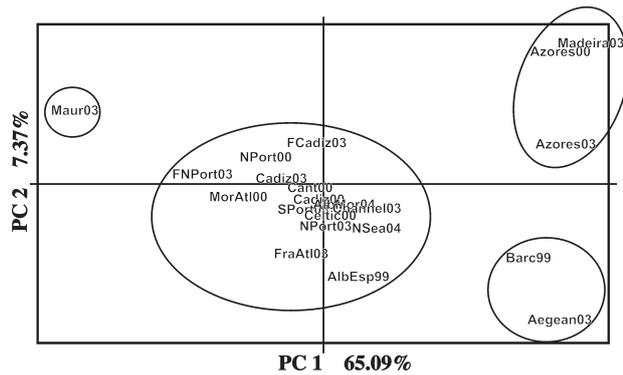
Fig. 2 – Principle component analysis of all 21 samples of sardine used in this study for 5 polymorphic microsatellite loci. Groupings refer to the inferred stocks of sardine. Sample abbreviations are as in Table 1.

all population pairs rejected the null hypothesis of no differentiation for all loci and confirmed the previous findings. No differentiation was observed between samples that had been collected from the same areas in different years, with or without locus Sp22.

SAMOVA analysis was performed using all loci and 19 out of 21 samples and assuming 2 to 9 groups (FNPort03 and FCadiz03 were excluded because they had a dubious origin, having been purchased at the port). The statistically significant configuration with the highest F_{CT} found comprised 5 groups (Barcelona and Aegean, Azores, Madeira, Mauritania, NE Atlantic and Alboran Sea; $F_{CT}=0.0298$, $P=0.00000$). When the same analysis was carried out excluding locus Sp22, the largest F_{CT} value ($F_{CT}=0.0101$, $P=0.0410$) was observed in a 2-group configuration (Madeira vs. all others). The PCA produced similar grouping of the samples to that obtained in the SAMOVA, with the exception of the Azores and Madeira, which seemed to form one group (Fig. 2). A 2-group structure representing the 2 subspecies, tested with AMOVA, was not significant ($F_{CT}=-0.1605$, $P=0.9306$).

A Mantel test showed a significant correlation between genetic and geographic distances ($r=0.4012$, $P<0.0010$) when all loci and the 19 samples described previously were used. After locus Sp22 had been excluded, no correlation ($r=0.0667$, $P=0.3060$) was detected.

A scatter plot of F_{ST} against heterozygosity for each of the 5 microsatellite loci, superimposed on the distribution expected for $F_{ST}=0.005$ (a value that roughly sets half the data points above and half below the median line) was produced, using FDIST2 (Fig. 3A). This plot shows that locus Sp22 is an outlier that falls well above the upper 95% confidence interval, suggesting it is under positive selection. Similarly, in the BAYESCAN approach, the same locus had an infinite probability of being under directional selection, while the other 4 had a posterior probability of less than 0.95 (Fig. 3B). This result was persistent when different combinations of samples belonging to different populations (pooled or not) were tested.
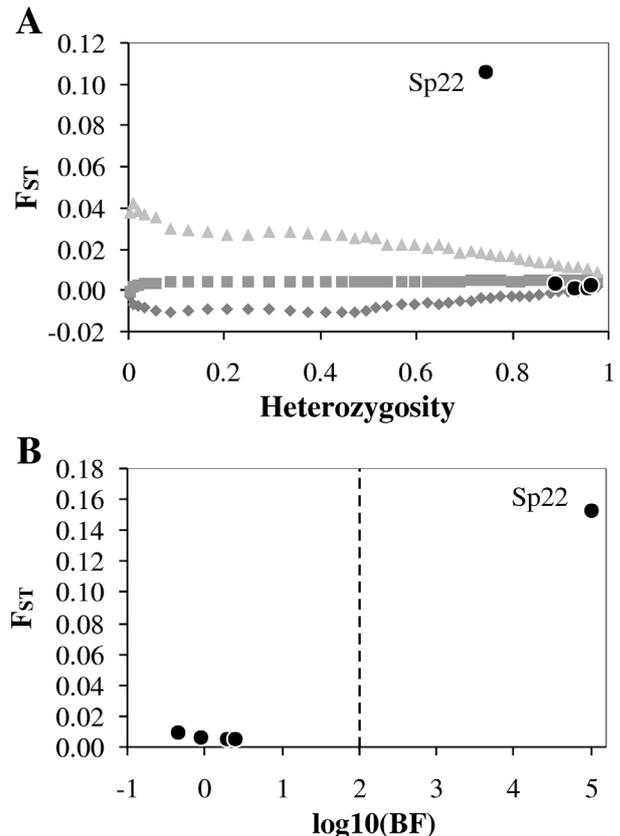
Fig. 3 – A, simulation results from the FDIST2 program. Upper and lower lines are the 95% confidence intervals and the middle line is the median. The estimated F_{ST} and heterozygosity values for each locus (black circles) have been superimposed on the graph. B, results of the BAYESCAN approach. F_{ST} values for each locus are plotted against their log-transformed Bayes factor. Broken line marks Log10(BF) of 2, which corresponds to posterior probability of 0.99.

Tests for correlation between allele frequencies of microsatellite locus Sp22 (4 main alleles, 19 in total) and allozymic locus SOD (2 alleles) in samples collected from the same geographical areas showed that the microsatellite allele Sp22[216] and allozymic allele SOD*100 (named as in Chlaida *et al.* 2009) exhibited the strongest correlation ($R=0.86$, $P=0.0020$) (Fig. 4).

The effect of environmental factors on genetic structure was estimated by the posterior probabilities of models that include a given factor, with software GESTE. The analysis using 16 samples and including the outlier locus Sp22 showed that the highest posterior probability was assigned to temperature (Pr=0.345). When locus Sp22 was removed from the analysis the highest probability was assigned to the model for longitude (Pr=0.274), while the model for temperature had a very low probability (Pr=0.027).

DISCUSSION

The analysis of 21 sardine samples for 5 microsatellite loci revealed a weak but statistically significant genetic structure over the species' distribution range
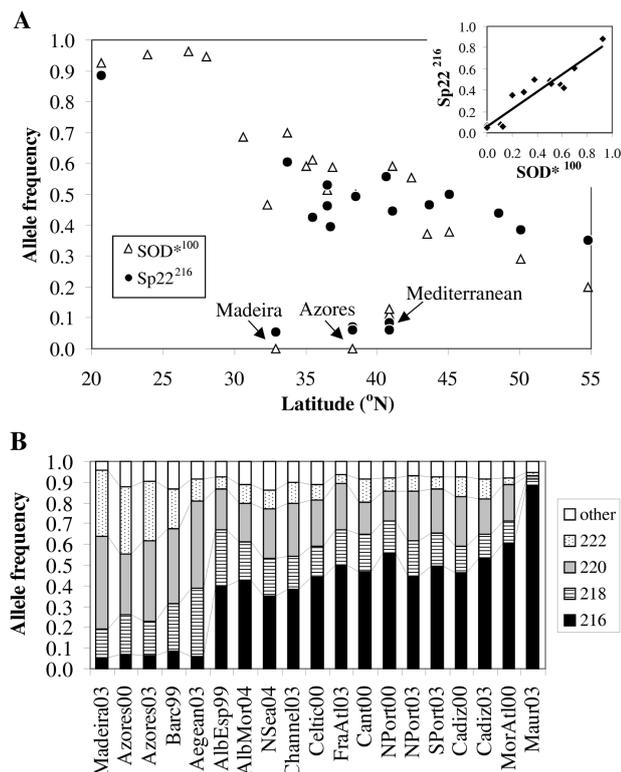
FIG. 4 – A, allelic frequencies for allele 216 of microsatellite locus Sp22 and for allele 100 of allozymic locus SOD* from different sardine samples plotted against the latitude. Data for SOD* are from Chlaida *et al.* (2006) and Laurent *et al.* (2007). The embedded graph displays the correlation between these 2 alleles. B, allelic frequencies of microsatellite locus Sp22 in the different samples.

(F_{ST}=0.017). Locus by locus F_{ST} analysis showed that this structure was mainly due to locus Sp22 (F_{ST}=0.089), while the other 4 loci had F_{ST} values 50 to 100-fold smaller. The approaches of Beaumont and Nichols (1996) with FDIST2 software and of Foll and Gaggiotti (2006) with BAYESCAN software both showed that this locus is not neutral but probably under directional selection.

Genetic structure based on neutral variation

When locus Sp22 was removed from the analysis, the 4 remaining neutral loci showed that only Madeira and Azores were significantly different from the other areas. The lack of genetic heterogeneity over most of the sardine's distribution range was in accordance with the findings of Gonzalez and Zardoya (2007a), who had a similar sample coverage with the exception of Madeira and the Azores. The 8 microsatellite loci used in that study (Gonzalez and Zardoya, 2007b), had similar F_{ST} values (0.0003-0.0079) to the 4 neutral loci of the current study. The authors concluded that there was a single evolutionary unit with a weak genetic structuring due to isolation by distance (IBD). Concordant results have also been found with allozymes in a similar geographical area to that of the present study (Laurent *et al.* 2007). In that study, the presumably neutral loci

exhibited a weak IBD pattern and a lack of genetic subdivision, except for SOD*, which seemed to be under selection. In the present study, no IBD was detected using only the neutral loci. Another study by Atarhouch *et al.* (2006), who used mitochondrial DNA sequences, showed a general lack of phylogeographic structure in sardine.

Therefore, neutral genetic variation studied so far in sardine with allozymes, microsatellites and mitochondrial DNA mostly point to a lack of genetic population structure. This may be due to historical factors, i.e. a recent postglacial population expansion of sardine (Atarhouch *et al.* 2006), possibly from a single refugial population, as well as to the absence of strong barriers to gene flow. The weak but significant differentiation of the peripheral populations of Madeira and the Azores may be the result of genetic drift due to isolation, which is likely maintained by the deep oceanic waters that separate them from populations of the continental shelf.

Genetic structure based on non-neutral variation and correlation with environmental factors

The inclusion of the non-neutral locus Sp22 in the population genetic analysis revealed a strong genetic structure, which consisted of the following distinct population groups: 1) the population off the African coasts of Mauritania, 2) the area from the Atlantic coast of Morocco along the east Atlantic continental shelf to the North Sea, including the Alboran Sea, 3) the Mediterranean Sea, 4) the Azores and 5) the Madeira.

Allele frequencies of locus Sp22 varied greatly among these groups, especially for allele Sp22[216]. This allele had a frequency of 0.88 in Mauritania, 0.60-0.35 along the Atlantic continental shelf, from the Atlantic Moroccan coasts to the North Sea, including the Alboran Sea, and 0.05-0.08 in the Azores, Madeira and the rest of Mediterranean (Fig. 4). A similar pattern had been found in allozymic locus SOD* (Laurent *et al.* 2007) and these 2 loci may be linked, because there is a very strong correlation in frequency between allozymic allele SOD*[100] and microsatellite allele Sp22[216] in samples from the same locations.

This non-neutral variation exhibits abrupt changes in allele frequencies from the Alboran Sea to the western Mediterranean and from the eastern Atlantic coast to the Azores and Madeira. Chlaida *et al.* (2006) found an additional steep change in SOD* off the African Atlantic coast at 28-30°N, which could not be shown for locus Sp22 owing to lack of samples from the respective areas. They also found a latitudinal cline in allele frequency along the Atlantic continental shelf from Morocco to the North Sea, which was not so evident in locus Sp22.

The Bayesian analysis of environmental factors with the genetic data showed that temperature had the highest posterior probability in explaining genetic structure when all loci were used. This probability

diminished when locus Sp22 was excluded from the analysis, which could indicate that this environmental factor is the selective force responsible for the outlier behaviour of the locus (see Gaggiotti *et al.* 2009).

Nevertheless, it is rather unlikely that a selection factor such as temperature would be more similar between, for example, the Alboran Sea and the North Sea, which have similar allele frequencies for Sp22, than between the Alboran Sea and the northwestern Mediterranean. Although selection seems to drive the observed pattern, the abrupt changes in allele frequencies of this locus seem to be related more to barriers to gene flow than to steep changes in selection pressure. We do know that such hydrodynamic barrier exists between the Alboran Sea and the Mediterranean (Almeria-Oran front)(Patarnello *et al.* 2007), and the deep ocean separating the Azores and Madeira from the nearby continental shelf probably also acts as a barrier. The abrupt change in allozymic frequency of SOD* off the African Atlantic coast at ~30°N has also been attributed to a hydrodynamic barrier (Chlaida *et al.* 2009), which may also hold for locus Sp22.

However, these barriers do not show any differentiation in neutral variation of sardine, with the exception of Madeira and the Azores. This may be attributed to the combined effects of recent separation of the populations, weak genetic drift associated with the large population size, and a reduced gene flow that is nevertheless sufficient to homogenize neutral genetic variation. It is known that the number of migrants required to homogenize neutral genetic variation is far less than that required to overcome selection. Hence, gene flow may be high enough to prevent differentiation of neutral markers but not so high as to prevent local adaptation (Conover *et al.* 2006).

Directional selection is expected to decrease within-population diversity and increase between-population differentiation in comparison with neutral expectations. However, it will only affect locus-specific patterns in neutral polymorphisms as compared to the genome-wide effects of population history and demographic events (Simonsen *et al.* 1995). In sardine, all 4 neutral microsatellite loci exhibited similar levels of diversity among the different populations studied. On the contrary, locus Sp22 had reduced diversity in Mauritania (H=0.21), where allele Sp22[216] had the highest frequency (0.88), whereas diversity was much higher (H=0.57-0.74) in all other areas where this allele was in lower frequency (Table 1). This indicates that in the Mauritanian population a genomic region linked to locus Sp22 may be under strong directional selection.

On the basis of these assumptions, a hypothesis that can be formulated, but remains to be tested, is that in the southern part of the sardine's distribution range (off the Mauritanian coast) there is directional selection related to temperature for a locus linked to Sp22. In the other parts of the distribution, selection might be reduced or absent so there is higher diversity, while gene flow creates a cline along the NE Atlantic coast.

Strong barriers to gene flow towards the Mediterranean (Almeria-Oran front) and towards the Azores and Madeira, possibly coupled with negative selection against the allele linked to Sp22[216], keep its frequency low in these regions.

Inferring stock structure based on non-neutral variation and non-genetic evidence

The non-neutral variation detected in sardine exhibits a pattern of abrupt change in allele frequencies in certain areas, which seems to be caused mainly by barriers to gene flow than by steep changes in selection, as discussed earlier. The unveiling of these barriers, most of which were not evident in the neutral variation, can be used to define the following genetic stocks in sardine: 1) an African one (south of 30°N, as has been demonstrated by Chlaida *et al.* 2009); 2) a NE Atlantic one (north of 30°N to the North Sea and the Alboran Sea); 3) a Mediterranean one (east of the Almeria-Oran front); and 4) one with the Azores and Madeira, which may actually consists of 2 different stocks.

Areas of discontinuity outlined by non-neutral genetic variation are generally consistent with areas of phenotypic change in sardine. The genetic boundary between the Atlantic and the Mediterranean separates sardines with shallow but significant morphometric differences and substantial differences in growth performance, with individuals from Mediterranean showing a lower head-to-body ratio (Silva 2003, Silva *et al.* 2008), lower length-at-age and lower maximum length (Silva *et al.* 2008). Other morphological and life-history traits were reported to vary between the 2 regions, such as the number of gill rakers (Andreu 1969), relative fecundity and spawning frequency (Ganias *et al.* 2003; Ganias *et al.* 2004) and length at first maturity (Silva *et al.* 2006): all of these showed lower values in the Mediterranean Sea, possibly reflecting a combination of poorer growth and lower global productivity in the area.

Off northwest Africa, 2 sardine populations, the Moroccan and Saharan populations, have been considered on the basis of differences in growth, maturation length, otolith and body morphology, parasitic infection and seasonal migrations (FAO, 2008). A boundary was assumed at the 28°N parallel, close to the genetic break recently described by Chlaida *et al.* (2009). However, formal statistical testing of these differences has not been carried out. A sharp morphometric distinction was found between sardine from north Morocco and Mauritania (Anonymous 2006), but given the lack of samples from the area between the 2 regions, it cannot be determined whether this difference is the consequence of sampling distant populations from a gradient.

Despite the general concordance of phenotypic and genetic changes in sardine, the existence of 2 morphological subspecies based on meristic studies was not supported by the current genetic analysis nor by Gonzalez and Zardoya (2007a).

For fisheries management, a single stock of sardine is considered within the European Atlantic, the Atlanto-Iberian stock distributed between the Bay of Biscay and the Gulf of Cadiz (ICES 2009). Three stocks are considered off the African coast: a smaller stock off northern Morocco (between 32°N and 35°45'N), a large central stock between 26°N and 32°N and an intermediate southern stock from the southern extent of the species range to 26°N (FAO 2008). Within the Mediterranean Sea, 8 stocks are considered, the largest ones located off the northern Spanish waters and the Gulf of Lions (FAO 2009). The genetic stocks proposed in the present study, as well as the available data on sardine phenotypic and life-history traits, show no major conflicts with the above fisheries stock definition; the main genetic discontinuities are already taken into account by separate management of European Atlantic, Mediterranean and Moroccan sardine stocks. Nevertheless, the combination of genetic and phenotypic data on sardine raises 3 topics for further debate and research regarding fisheries stock structure: (1) the northern limit of the Atlanto-Iberian stock is challenged both by the lack of genetic and phenotypic discontinuities and by the existence of an adjacent large sardine population in the Bay of Biscay (ICES 2006); (2) the boundary between European and African sardine populations seems to lay well within the African waters (around 30°N), suggesting that the extent of mixing between the Atlanto-Iberian and northern Morocco stocks warrants further attention; and (3) a higher number of stocks than suggested by genetic and phenotypic data is considered for fisheries management. Although this may partly reflect administrative decision, it is also a consequence of gaps in biological knowledge regarding connectivity patterns at small and large spatial scales.

The results presented in this paper corroborate earlier evidence of weak neutral genetic structure in sardine and point to the existence of selection, which needs to be further investigated. The low number of microsatellite loci analyzed reduces the statistical power of the methods used to detect selection, so a larger number of markers should be tested in the future to confirm these results. Moreover, the next generation sequencing technologies, which allow for extensive sequencing of the genomes with low cost, could greatly assist in establishing the assumed linkage between loci SOD and Sp22 and in finding the aetiology behind the observed pattern.

ACKNOWLEDGEMENTS

REFERENCES

Allendorf F.W., Phelps S.R. 1981. Use of allelic frequencies to describe population structure. *Can. J. Fish Aquat. Sci.* 38: 1507-1514.

Andreu B. 1969. Las branquispinhas en la caracterizacion de las poblaciones de *Sardina pilchardus* (Walb.). *Inv. Pesq.* 33: 1-607.

Anonymous. 2006. Sardine dynamics and stock structure in the North-eastern Atlantic. Final Report. DGXIV Fisheries, European Commission, Brussels. Q5RS/2002/000818, available from yorgos@pimar.pt.

Atarhouch T., Rüber L., Gonzalez E.G., Albert E.M., Rami M., Dakkak A., Zardoya R. 2006. Signature of an early genetic bottleneck in a population of Moroccan sardines (*Sardina pilchardus*). *Mol. Phyl. Evol.* 39: 373-383.

Balding D.J., Nichols R.A. 1995. A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. *Genetica* 96: 3-12.

Batargias C., Dermitzakis E., Magoulas A., Zouros E. 1999. Characterization of 6 polymorphic microsatellite markers in gilthead seabream, *Sparus aurata*. *Mol. Ecol.* 8: 897-899.

Beaumont M.A., Nichols R.A. 1996. Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B. Biol. Sci.* 263: 1619-1626.

Belkhir K., Borsa P., Chikhi L., Raufaste N., Bonhomme F. 1996-2004. *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Univ. Montpellier II, Montpellier (France).

Carvalho G., Hauser L. 1998. Advances in the molecular analysis of fish population structure. *Ital. J. Zool.* 65(Suppl): 21-33

Chlaida M., Kifani S., Lenfant P., Ouragh L. 2006. First approach for the identification of sardine populations *Sardina pilchardus* (Walbaum 1792) in the Moroccan Atlantic by allozymes. *Mar. Biol.* 149: 169-175.

Chlaida M., Laurent V., Kifani S., Benazzou T., Jaziri H., Planes S. 2009. Evidence of a genetic cline for *Sardina pilchardus* along the Northwest African coast. *ICES J. Mar. Sci.* 66: 264-271.

Conover D.O., Clarke L.M., Munch S.B., Wagner G.N. 2006. Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *J. Fish Biol.* 69 (Suppl C): 21-47.

Coombs S.H., Smyth T.J., Conway D.V.P., Halliday N.C., Bernal M., Stratoudakis Y., Alvarez P. 2006. Spawning season and temperature relationships for sardine (*Sardina pilchardus*) in the eastern North Atlantic. *J. Mar. Biol. Assoc. UK* 86: 1245-1252.

Dupanloup I., Schneider S., Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11: 2571-81.

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

Excoffier L., Laval G., Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinformatics Online* 1: 47-50.

Excoffier L., Hofer T., Foll M. 2009. Detecting loci under selec-

tion in a hierarchically structured population. *Heredity* 103: 285-298.

FAO 2008. Report of the FAO Working Group on the Assessment of Small Pelagic Fish off Northwest Africa. Saly, Senegal, 6-15 May 2008. *FAO Fish Rep.* No. 882.

FAO 2009. Report of the SCSA working group on stock assessment of small pelagic species. GFCM: SAC12/2010/Inf.14.

Foll M., Gaggiotti O.E. 2006. Identifying the environmental factors that determine the genetic structure of populations. *Genetics* 174: 875-891.

Foll M., Gaggiotti O.E. 2008. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180: 977-993.

Ganias K., Somarakis S., Machias A., Theodorou A. 2003. Evaluation of spawning frequency in a Mediterranean sardine population. *Mar. Biol.* 142: 1169-1179.

Ganias K., Somarakis S., Machias A., Theodorou A. 2004. Pattern of oocyte development and batch fecundity in the Mediterranean sardine. *Fish. Res.* 67: 13-23.

Gaggiotti O.E., Bekkevold D., Jorgensen H.B.H., Foll M., Carvalho G.R., Andre C., Ruzzante D.E. 2009. Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. *Evolution* 63: 2939-2951.

Gauldie R.W. 1988. Tagging and genetically isolated stocks of fish: a test of one stock hypothesis and the development of another. *J. Appl. Ichth.* 4: 168-173.

Goudet J. 1995. Fstat version 1.2: a computer program to calculate Fstatistics. *J. Heredity* 86: 485-486.

GFCM 2006. Report of the eighth session of the sub-committee on stock assessment (SCSA). GFCM:SAC9/2006/Inf.8. ftp://ftp.fao.org _/DOCUMENT/gfcm/sac9/inf8e.pdf

Gonzalez E.G., Zardoya R. 2007a. Relative role of life-history traits and historical factors in shaping genetic population structure of sardines (*Sardina pilchardus*). *BMC Evol. Biol.* 7: 197- 209.

Gonzalez E.G., Zardoya R. 2007b. Isolation and characterization of polymorphic microsatellites for the sardine, *Sardina pilchardus* (Clupleidae). *Mol. Ecol. Notes* 7: 519-521.

Guinand B., Lemaire C., Bohnomme F. 2004. How to detect polymorphisms undergoing selection in marine fishes? A review of methods and case studies, including flatfishes. *J. Sea Res.* 51: 167-182.

Hemmer-Hansen J., Nielsen E.E., Frydenberg J., Loeschcke V. 2007. Adaptive divergence in a high gene flow environment: *Hsc70* variation in the European flounder (*Platichthys flesus* L.). *Heredity* 99: 592-600.

ICES 2006. Report of the Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine, and Anchovy. ICES Document CM 2006/ACFM: 36.

ICES 2009. Report of the Working Group on Anchovy and Sardine. ICES Document CM 2009/ACOM:13.

Jensen J.L., Bohonak A.J., Kelley S.T. 2005. Isolation by distance, web service. *BMC Genetics* 6: 13.

Laurent V., Caneco B., Magoulas A., Planes S. 2007. Isolation by distance and selection effects on genetic structure of sardines *Sardina pilchardus,* Walbaum. *J. Fish Biol.* 71: 1-17.

Mäkinen H.S., Cano J.M., Merilä J. 2008. Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. *Mol. Ecol.* 17: 3565-3582.

Maynard Smith J., Haigh J. 1974. The hitch–hiking effect of a favourable gene. *Genet. Res.* 23: 23-35.

Miller S.A., Dykes D.D., Polesky H.F. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16: 1215.

Moen T., Hayes B., Nilsen F., Delghandi M., Fjalestad K., Fevolden S.E., Berg P., Lien S. 2008. Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. *BMC Genetics* 9: 18.

Nielsen E.E., Hemmer-Hansen J., Larsen P.F., Bekkevold D. 2009. Population genomics of marine fishes: identifying adaptive variation in space and time. *Mol. Ecol.* 18: 3128-3150.

Nosil P., Funk D.J., Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* 18: 375-402.

Ovenden J.R. 1990. Mitochondrial DNA and marine stock assessment: a review. *Aust. J. Mar. Freshwater Res.* 41: 835-853.

Patarnello T., Volckaert A.M.J., Castilho R. 2007. Pillars of Hercules: is the Atlantic–Mediterranean transition a phylogeographical break? *Mol. Ecol.* 16: 4426-4444.

Parrish R.H., Serra R., Grant W.S. 1989. The monotypic Sardines, *Sardina* and *Sardinops*: their taxonomy, distribution, stock structure, and zoogeography. *Can. J. Fish. Aquat. Sci.* 46: 2019-2036.

Pogson G.H. 2001. Nucleotide polymorphism and natural selection at the pantophysin (*Pan I*) locus in the Atlantic cod, *Gadus morhua* (L.). *Genetics* 157: 317-330.

Raymond M., Rousset F. 1995. GENEPOP (version1.2): population genetics software for exact tests and ecumenism. *J. Heredity* 86: 248-249.

Rice W.R. 1989. Analyzing tables of statistical tests. *Evolution* 43: 223-225.

Schulte P.M., Glemet H.C., Fiebig A.A., Powers D.A. 2000. Adaptive variation in lactate dehydrogenase-B gene expression: role of a stress-responsive regulatory element. *Proc. Natl. Acad. Sci. USA* 97: 6597-6602.

Simonsen K.L., Churchill G.A., Aquadro C.F. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* 141: 413-429.

Silva A. 2003. Morphometric variation among sardine (*Sardina pilchardus*) populations from the northeastern Atlantic and the western Mediterranean. *ICES J. Mar. Sc*i. 60: 1352-1360.

Silva A., Santos M.B., Caneco B., Pestana G., Porteiro C., Carrera P., Stratoudakis Y. 2006. Temporal and geographic variability of sardine maturity at length in the north-eastern Atlantic and the western Mediterranean. *ICES J. Mar. Sci.* 63: 663-676.

Silva A., Carrera P., Massé J., Uriarte A.D., Santos M.B., Oliveira P.B., Soares E., Porteiro C., Stratoudakis Y. 2008. Geographic variability of sardine growth across the northeastern Atlantic and the Mediterranean Sea. *Fish. Res.* 90: 56-69.

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. *J. Fish Biol.* 35: 417-437.

Tinti F., Di Nunno C., Guarniero I., Talenti M., Tommasini S., Fabbri E., Piccinetti C. 2002. Mitochondrial DNA sequence variation suggests the lack of genetic heterogeneity in the Adriatic and Ionian stocks of *Sardina pilchardus*. *Mar. Biotech.* 4: 163-172.

Tsigenopoulos C.S., Hellemans B., Chistiakov D.A., Libertini A., Kotoulas G., Volckaert F. 2003. Eleven new microsatellites of the sea bass (*Dicentrarchus labrax* L.). *Mol. Ecol. Notes* 3: 352-354.

van Oosterhout C., Hutchinson W.F., Wills D.P.M., Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4: 535-538.

Waples R.S. 1998. Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *J. Heredity* 89: 438-450.

Ward R.D. 2000. Genetics in fisheries management. *Hydrobiologia* 420: 191-201.

Weir B.S., Cockerham C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.