



Population genetic structure of *Mugil cephalus* in the Mediterranean and Black Seas: a single mitochondrial clade and many nuclear barriers

J. D. Durand^{1,*}, H. Blel², K. N. Shen^{3,4}, E. T. Koutrakis⁵, B. Guinand⁶

¹Institut de Recherche pour le Développement (IRD), UMR 5119 ECOSYM, Université Montpellier II, Place Eugène Bataillon, CC 93, 34095 Montpellier cedex 5, France

²Unité de Recherche de Génétique: Biodiversité et Valorisation des Bio ressources (UR: 09/30), Institut Supérieur de Biotechnologie de Monastir, Université de Monastir, 5000 Monastir, Tunisia

³Department of Environmental Biology and Fisheries Science, and ⁴Center of Excellence for Marine Bioenvironment and Biotechnology, National Taiwan Ocean University, Keelung 20224, Taiwan

⁵Fisheries Research Institute-NAGREF, 640 07 Nea Peramos, 64007 Kavala, Greece

⁶Institut des Sciences de l'Evolution de Montpellier, Evolution des Poissons, CNRS-UMR 5554, Université Montpellier II, Place Eugène Bataillon, CC 65, 34095 Montpellier cedex 5, France

ABSTRACT: The population structure and evolutionary history of *Mugil cephalus* were investigated across 18 sampling sites in the NE Atlantic Ocean, Mediterranean and Black Seas, using 2 classes of genetic markers: sequence polymorphism of an 857 bp fragment of mitochondrial (mtDNA) cytochrome *b*, and allele size variation at 7 nuclear loci. The level of nucleotide diversity recovered with the mtDNA marker was very low (~0.6% divergence), indicating the presence of a single clade over the entire area. Mismatch distribution, Bayesian skyline plots and associated statistics revealed a recent demographic crash followed by population expansion, but nuclear data indicated population constancy in the area covered in this study. While a single clade was detected, significant mtDNA genetic differentiation was, however, observed between the samples from the Black Sea and the samples from other (sub-) basins ($\Phi_{ST} = 0.17$; $p = 0.029$). The nuclear loci also revealed significant genetic differentiation and isolation-by-distance in *M. cephalus*. Patterns of genetic structure were, however, significantly more pronounced with nuclear than with mtDNA markers; the former indicated the presence of 3 (Bayesian clustering) to 6 (Monmonnier's method) populations. The highest levels of genetic heterogeneity at nuclear markers occurred at the well-known Almeria-Oran Front, but also in the Bosphorus Strait. Thus, both sets of markers revealed the importance of this strait as a barrier to gene flow, probably during the Pleistocene. The results also revealed genetic heterogeneity in the eastern Mediterranean basin, and suggested that the population expanded from this sub-basin towards the Atlantic Ocean and Black Sea.

KEY WORDS: Microsatellites · Mitochondrial DNA · Gene flow · Demography · Mediterranean Sea · Mugilidae · Bayesian clustering · Bayesian skyline plot

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

It is now well-established that marine organisms can be genetically structured across a broad range of geographic scales, despite being in an environment that is considered to be highly interconnected (Hauser

& Carvalho 2008). Estimated levels of connectivity among marine populations reflect events ranging from distant historical allopatric divergence or climate shifts, to current-day environmental features, oceanographic patterns, or behaviours that constrain genetic exchanges among populations (Selkoe et al.

*Email: jean-dominique.durand@ird.fr

2008). The nature and patterns of inheritance of genetic markers can provide information about how these processes affect marine connectivity: mitochondrial DNA (mtDNA) is traditionally considered to illustrate historical processes, while nuclear markers such as microsatellites are better at revealing contemporary processes and resolving population structure on a finer scale (e.g. Gonzalez & Zardoya 2007). Comparative studies of both types of markers are, therefore, useful for elucidating population structure of marine organisms, and possibly shedding light on the relative effects of genetic drift, mutation and migration, selection, differences in effective population size, or sex-biased dispersal (Buonaccorsi et al. 2001, Canino et al. 2010b). Collectively, this information is important for the efficient management of marine resources.

The flathead (or striped) grey mullet *Mugil cephalus* L., 1758 is considered to be a single species with a worldwide distribution comprising most of the coastal and estuarine environments between ca. 42° N and 42° S (Thomson 1997). This distribution includes major biogeographic zones and barriers to dispersal (Thomson 1997). The ecological success of the flathead mullet across such a broad geographic range has raised doubts about its true taxonomic status (Crosetti et al. 1994, Durand et al. 2012) and stimulated phylogeographical studies at an inter-oceanic scale (Crosetti et al. 1994, Rossi et al. 1998, Rocha-Olivares et al. 2000, Heras et al. 2009, Livi et al. 2011). While the inter-oceanic structure has been extensively investigated, further population genetic studies are now necessary at smaller oceanic scales. To date, these have found genetic homogeneity within the NW Atlantic and the Gulf of Mexico (Campton & Mahmoudi 1991, Rocha-Olivares et al. 2000). Studies performed in the Mediterranean Sea have also reported genetic homogeneity among *M. cephalus* samples (allozymes: Rossi et al. 1998; mtDNA: Livi et al. 2011; microsatellites: Blel et al. 2010). By contrast, genetic studies conducted in the NW Pacific demonstrated the existence of large genetic heterogeneity (review in Shen et al. 2011). Indeed, using mtDNA and microsatellite loci simultaneously, Shen et al. (2011) reported that the NW Pacific population of flathead mullet, which had been considered to harbour a single mitochondrial lineage (Crosetti et al. 1994, Heras et al. 2009), consisted of 3 mitochondrial lineages corresponding to distinct parapatric species. These results warrant a re-evaluation of the genetic structure of the flathead mullet in other geographic areas, based on data from multiple genetic markers. This is especially true for

the Mediterranean Sea and adjacent NE Atlantic (i.e. Moroccan, Spanish and Portuguese coastlines) because previous studies of genetic variation in *M. cephalus* only analysed limited sets of samples ($n \leq 5$; Rossi et al. 1998, Blel et al. 2010, Livi et al. 2011) that did not fully cover the marine phylogeographic regions recognized there (NE Atlantic, western and eastern Mediterranean basins, and the Black Sea; e.g. Patarnello et al. 2007). A wider picture of genetic differentiation for reliable monitoring of this important marine resource is lacking in this area, where mullet has been fished and cultured for centuries, providing an important income for artisanal fishermen in Mediterranean and other countries (Koutrakis et al. 2007, Whitfield et al. 2012), but which is now heavily impacted by climate change and human activities.

In the present study, the population genetic structure of *Mugil cephalus* was investigated in the Mediterranean Sea, the adjacent area of the NE Atlantic and the Black Sea, using 2 classes of genetic markers: sequence polymorphism of a mitochondrial gene (cytochrome *b*), plus allele size variation at 7 nuclear loci (6 microsatellite loci and 1 exon-primed intron-crossing [EPIC] locus). Based on extensive sampling coverage, we investigated (1) the null hypothesis of no genetic structure at each class of markers, (2) the past demographic history of flathead grey mullet in this area based on mtDNA and nuclear markers, and (3) estimates of number of effective migrants and patterns of asymmetric migration among phylogeographic areas based more specifically on nuclear markers.

MATERIALS AND METHODS

Sampling

Eighteen locations are considered in this study, from the Atlantic Ocean to the Black Sea, hence covering the 2 main biogeographic regions of the Mediterranean (western and eastern basins; Table 1, Fig. 1). *Mugil cephalus* specimens, comprising a sample of pectoral fins preserved in 95% ethanol, were collected in 2005 and 2006 from the landings of artisanal fisheries (cast nets, fish traps, purse seines and drift nets).

Molecular methods

Total genomic DNA was extracted from muscle samples or fin clips using standard phenol-chloroform

Table 1. *Mugil cephalus*. Sampling locations in the NE Atlantic and the Mediterranean and Black Seas, with the sample size (N) and some diversity indices estimated with microsatellite loci (*µsats*) and cytochrome *b* sequences (*cytb*). L.: lagoon; R.: river; *z*: mean number of alleles from the 7 loci used in this study (*µsats*) or allele number (*cytb*); H_{obs} : observed heterozygosity; *h*: haplotype diversity; π : nucleotide diversity; *D*: Tajima's D ; F_S : Fu's F_S . * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

No.	Code	Location	Country	Biogeographic area	Date	<i>µsats</i>		<i>cytb</i>				F_S	
						N	<i>n</i>	N	<i>n</i>	<i>h</i>	π		<i>D</i>
1	BSS	Sevastopol	Ukraine	Black Sea	Nov 2005	33	12.14	5	3	0.7	0.0012	0.243	-0.475
2	BSI	Istanbul	Turkey	Black Sea	May 2005	55	13	7	4	0.81	0.0017	-0.598	-0.78
3	ASH	Homa L.	Turkey	Aegean Sea	Aug/Nov 2006	24	9.86	11	5	0.62	0.0011	-1.791*	-2.309**
4	EML	Latakia	Syria	Eastern Med.	Oct 2006	33	14.14	11	6	0.8	0.0014	-1.22	-3.102**
5	EMH	Hifa	Israel	Eastern Med.	Jun 2005	25	13.29	15	1	0	0	0	-
6	EMM	Mesolonghi L.	Greece	Eastern Med.	Oct 2006	14	10.14	14	5	0.51	0.0007	-1.798*	-3.143***
7	EMV	Vontas R.	Greece	Eastern Med.	Oct 2006	38	13	10	6	0.84	0.0013	-1.388	-3.412**
8	EMZ	Zarzis	Tunisia	Western Med.	Dec 2005	31	13.57	3	2	0.67	0.0008	0	0.201
9	EMK	Khmis	Tunisia	Western Med.	Oct 2005	31	12.57	10	4	0.73	0.0013	-0.762	-0.615
10	WMG	Goulette	Tunisia	Western Med.	Dec 2005	30	14.71	10	5	0.76	0.0012	-0.943	-2.096*
11	WMO	Oristano	Italy	Western Med.	Mar 2006	46	15.57	11	3	0.35	0.0004	-1.430*	-1.246*
12	WMB	Berre L.	France	Western Med.	Jul 2006	45	15.57	15	6	0.72	0.0012	-1.549*	-4.190***
13	WMT	Thau L.	France	Western Med.	Mar/Jul 2006	25	12.71	6	5	0.93	0.0019	-1.337	-2.158*
14	WME	Ebro R.	Spain	Western Med.	Jul 2006	32	14.14	15	4	0.37	0.0006	-1.816*	-1.716*
15	WMA	Alicante	Spain	Western Med.	Oct 2006	30	14.14	15	5	0.63	0.0014	-1.188	-1.027
16	AOC	Cadiz	Spain	Atlantic	Nov 2006	39	14.43	15	4	0.47	0.0008	-1.518*	-1.287
17	AOM	Mejfa Zerga	Morocco	Atlantic	Feb 2006	47	15.57	9	4	0.69	0.0014	-0.69	-0.627
18	AOT	Tagé	Portugal	Atlantic	Jul 2006	0	-	2	1	0	0	0	-
					Total	578		184	24	0.7	0.0013	-0.960**	-21.562***

protocols (Sambrook et al. 1989). The nuclear markers comprised 6 microsatellite loci (*MCS1EH*, *MCS2FH*, *MCS15AM*, *MCS15CM*, *MCS2DM* and *MCS16DM*; Miggiano et al. 2005) and the first intron of the prolactin 1 gene locus (*prl-1*; Blel et al. 2010). Primers and polymerase chain reaction (PCR) protocols are reported in Miggiano et al. (2005) and Blel et al. (2010), respectively. PCR products were used undiluted and mixed with an equal volume of formamide loading dye (95% formamide, 20 mM EDTA, 0.05% xylene cyanol and 0.05% bromophenol blue), and denatured at 94°C for 5 min. Then, 6 µl of this mixture was loaded into 6% denaturing polyacrylamide gel and run using 1× TBE buffer at 50 W for 2 h. The gel was then laser scanned, and the fluorescent bands were visualized in an FMBio II fluorescence imaging apparatus (Hitachi Instruments). Several individuals of known genotype were used as additional allele-size standards on each gel. The mitochondrial marker was an 857 base pair (bp) fragment of the cytochrome *b* (*cytb*) gene. MtDNA amplification was performed with the primers *Fishcytob-F* (Sevilla et al. 2007) and *MixCytob-0937-1-R* (5'-GGK CGG AAT GTY AGK CYT CG-3'). The PCR was carried out in a 50 µl reaction volume containing 5 µl of 10× reaction buffer (Promega), 1.5 µl of MgCl₂ (25 mM), 2 µl of dNTP (5 mM), 0.5 µl of each primer (10 µM), 1 U of GoTaq DNA polymerase (Promega) and 1 µl of template DNA of unknown concentration. PCR amplification conditions were as follows: preliminary denaturation at 92°C (5 min), strand denaturation at 92°C (1 min), primer annealing at 52°C (45 s) and primer extension at 72°C (1.5 min) repeated for 35 cycles and final extension at 72°C (5 min). All sequencing reactions were performed according to the manufacturer's protocol (Applied Biosystems). Sequences were deposited in GenBank (accession numbers JN390976 to JN391159).

Data analysis

Diversity indices (haplotype and nucleotide diversity) were estimated using ARLEQUIN v3.5 (Excoffier & Lischer 2010), and phylogenetic relationships among *Mugil cephalus* haplotypes were depicted through a phylogenetic

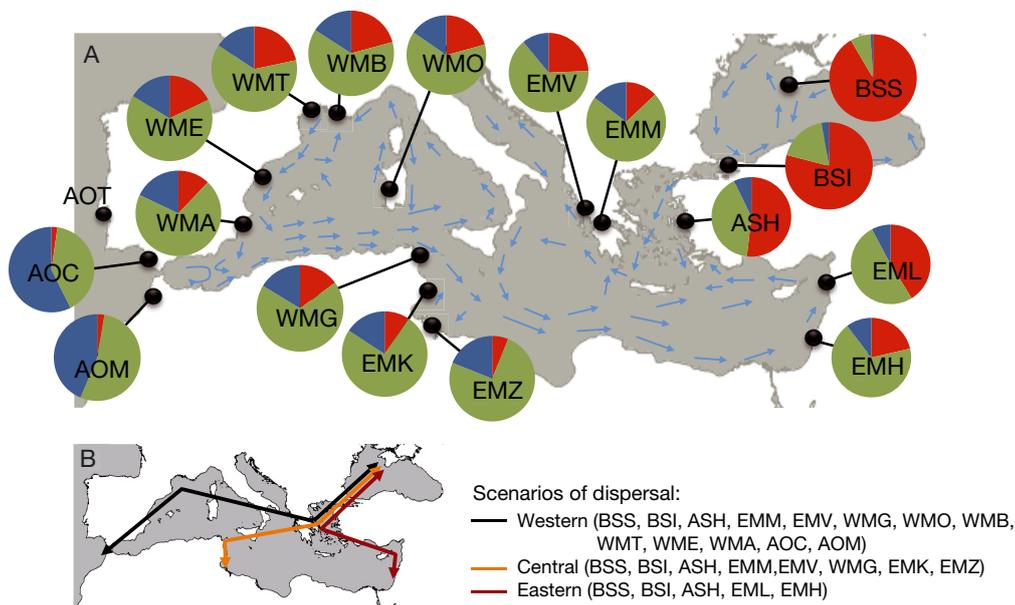


Fig. 1. Sampling locations of *Mugil cephalus* investigated in the NE Atlantic and the Mediterranean and Black Seas. Location details and codes are given in Table 1. (A) Admixture proportions of the 3 genetic clusters recovered by TESS software in *M. cephalus* samples from the NE Atlantic (blue) and the Mediterranean (green) and Black Seas (red). Light blue arrows represent contemporary currents. (B) Scenarios of dispersal through the Mediterranean area used for analysis of isolation-by-distance patterns at nuclear and mtDNA markers (see 'Results: Nuclear loci')

analysis based on a partitioned maximum-likelihood (ML) method implemented in MEGA 5 (Tamura et al. 2011). The HKY + G model of substitutions was used as it better fit our data set than any other models according to the model test implemented in ARLEQUIN v.3.5 (Bayesian information criterion = 6938.825). To provide an overview of the genetic diversity of *M. cephalus* and highlight phylogenetic relationships among mtDNA lineages observed in the Atlantic and in the East Pacific (Durand et al. 2012), we included in the phylogenetic analyses the *cytb* haplotypes described in Durand et al. (2012). The tree was rooted using 5 *cytb* sequences of *M. bananensis* as outgroups (GenBank accession numbers JQ060267, JQ060268) and *M. capurii* (JQ060269, JQ060270, JQ060271).

Observed heterozygosities (H_{obs}) were estimated from raw nuclear genotype data at each nuclear locus using GENETIX v.4.05 software (<http://kimura.univ-montp2.fr/genetix/>), while nucleotide and haplotype diversity indices at *cytb* were estimated using ARLEQUIN v.3.5. Deviations from Hardy-Weinberg expectations (HWE) within samples were investigated using Weir & Cockerham's (1984) estimate of f (quoted as \hat{f}) with GENETIX. A test for significant departure from HWE ($\hat{f} = 0$) was performed by randomly permuting alleles from the original matrix of genotypes using the appropriate procedure in GENETIX. Critical significance levels for multiple testing were corrected following the sequential Bonferroni procedure. Linkage disequilibrium between pairs of nuclear loci (i.e. non-random associations of particular genotypes) was tested with GENETIX.

The occurrence of bottlenecks among *Mugil cephalus* samples was investigated using BOTTLENECK on the nuclear genotypes (Piry et al. 1999). Tests (1000 replicates) were performed using the stepwise mutation model that best described the data. A Wilcoxon sign-ranked test was used to assess the significance of results. This test is recognized as having better statistical power when a low number of loci (<15) is considered (Luikart et al. 1998). For mtDNA, further insight into the historical demography of *M. cephalus* was achieved by mismatch distributions (i.e. the observed number of differences between pairs of haplotypes), a Bayesian skyline plot model (Drummond et al. 2005) and classical statistical tests of neutrality (Tajima 1989, Fu 1997) over observed mtDNA haplotypes. A mismatch distribution is usually multimodal for populations at demographic equilibrium (i.e. stationary), and unimodal for populations that experienced a recent demographic expansion. Harpending's (1994) raggedness test (RI for 'raggedness index') was used to determine if the observed mismatch distribution was drawn from an expanded or a stationary population (small versus large RIs, respectively). Past population demography was also estimated using a Bayesian skyline plot (BSP) model that generated a posterior distribution of effective population size through time using a Markov chain Monte Carlo (MCMC) sampling procedure. The parameter m that represents the number of grouped intervals was set to 5. The MCMC analysis was run for 20×10^6 generations (sampled every 1000 iterations), of which the first 10% was discarded as burn-in. The substitution

model used was HKY + G. The BSP model was computed using BEAST v.1.7.2 (http://beast.bio.ed.ac.uk/Main_Page). The median and corresponding credibility intervals of the BSP were depicted using TRACER v.1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>). BSPs are now a common tool in phylogeographic studies of marine fishes (reviewed in Grant et al. 2012), but over-interpretations concerning population growth, stasis, or population decrease and timing of demographic events can easily occur because of multiple factors. These include inadequate sampling (number of individuals and incomplete sampling of haplotypes), but molecular clock calibrations are also crucial to unravel uncertainties in demographic timelines (Ho et al. 2008, Ruzzante et al. 2008). We adopted a conservative approach, and, in this study, we have not, to date, searched events using BSP, as no reliable calibration of mutation rate was accessible. We only observed the trends in estimates of effective population size as a function of inferred mutational units (Grant et al. 2012). Tajima's D and Fu's F_S tests were used to test neutrality, with significant D values due to factors such as population expansion, bottlenecks and selection (Tajima 1989). Fu's F_S test compares the number of haplotypes expected in a random sample under the assumption of an infinite-site model without recombination (Fu 1997). This test is sensitive to the population demographic expansion that generally leads to large negative values of F_S . Analyses of mtDNA data were carried out using ARLEQUIN v.3.5. For nuclear loci, features of past demographic events were examined using the interlocus g -test (Reich et al. 1999), which is appropriate for the low number of loci considered in this study. The interlocus g -test assumes that, in an expanding population, the variance in the widths of the allele length distribution across loci is usually lower than in a population of constant size; an unusually low value of the g -statistic may be interpreted as an indication of expansion. The KGTTESTS Excel macro provided by Bilgin (2007) was used to compute the interlocus g -test. Significance was assessed by computer simulations that established cut-off values for a given number of samples and loci as described in Reich et al. (1999). As results of the interlocus g -test are sensitive to variation in mutation rate (Reich et al. 1999), data were analysed with and without the *pri-1* locus.

For both nuclear and mtDNA loci, pairwise levels of population differentiation were investigated using Weir & Cockerham's (1984) $\hat{\theta}$. Population structure was further examined using the analysis of molecular variance method (AMOVA; Excoffier et al. 1992) as implemented in ARLEQUIN v.3.5, with significance

determined by 10 000 random permutations of the data. The number of different alleles and of pairwise differences were used to estimate distances in molecular variance analyses using nuclear and mtDNA loci, respectively. In these analyses molecular variance is decomposed as follows: the variation observed within samples (Φ_{ST}), among samples with groups (Φ_{SC}) and among groups (Φ_{CT}). Isolation-by-distance (IBD) was tested across mtDNA and nuclear loci using the Mantel test as implemented in GENETIX, with significance assessed by permuting data (5000 iterations). The minimum distance over the continental plateau (*Mugil cephalus* is primarily a demersal species foraging over this plateau) and $\hat{\theta}/(1 - \hat{\theta})$ (nuclear data) or $\Phi_{ST}/(1 - \Phi_{ST})$ (mtDNA data) were considered among all pairs of populations in the genetic distance matrix (Rousset 1997).

The population genetic structure was further investigated using 2 different methods aimed at detecting spatial population structure and locating discontinuities in allele frequencies. The Bayesian clustering method implemented in the software TESS (Chen et al. 2007) was used. TESS is a spatially explicit Bayesian admixture model implementing a MCMC algorithm that estimates individual ancestry proportions by incorporating spatial trends and autocorrelation in the prior distribution. As an 'admixture model', TESS assumes that the data originate from the admixture of K putative parental populations. In these models, the parameters of interest are the ancestry coefficients (or admixture proportions), computed for each individual in the sample and stored in a matrix, \mathbf{Q} , with elements (q_{ik} ; i = individuals, k = sample [Table 1]) representing the proportion of the individual's genome that originates from the parental population (or cluster) K . In TESS, spatial trends are also incorporated in the \mathbf{Q} matrix together with admixture proportions and are allowed to vary over space at both local and global scales. In other words, the fraction of an individual's genome originating in the putative parental population A is proportional to its distance to A (Durand et al. 2009). Consideration of spatial trends and autocorrelation should limit the fact that non-spatial Bayesian algorithms may be confounded by discontinuous spatial sampling. TESS includes a Voronoï diagram describing spatially neighbouring and non-neighbouring populations (i.e. a 'tessellation'), and a prior distribution on cluster labels based on the Potts model (i.e. a statistical model in which the state [multilocus genotype] of each individual is influenced only by the states of its neighbours in the same sample and in close geographical samples). TESS has been shown to display results comparable to or better than the

results from other types of software based on admixture models (Chen et al. 2007, François & Durand 2010). Like other Bayesian clustering techniques, TESS not only aims to minimise the Wahlund effect when grouping individuals in K putative parental populations, but also incorporates spatial local dependencies when performing clustering. To determine the most probable value of K , the maximum number of clusters, K_{\max} , was sequentially increased until the final inferred number of clusters, K , was less than K_{\max} . First the non-admixture model was used with a burn-in period of 20 000 cycles, and estimation was performed using 30 000 additional cycles. The maximal number of clusters from $K_{\max} = 2$ to 6 (30 replicates for each value) was increased. Then, the conditional auto-regressive Gaussian model of admixture with linear trend surface (Durand et al. 2009) was used for $K_{\max} = 3$. One-hundred runs (burn-in 3×10^4 , run length 10^5 , $\rho = 0.6$, $\alpha = 1$) were performed, and the estimated admixture coefficients (\mathbf{Q} matrix) were averaged over the 20 runs with the smallest deviance information criterion (DIC). To account for label switching among runs, the CLUMPP v.1.1.1 software (Jakobsson & Rosenberg 2007) was used.

In order to compare and to confirm outcomes of TESS with results from an independent method, the BARRIER v.2.2 program was also used to identify the geographic areas associated with genetic discontinuities at nuclear loci (Manni et al. 2004). Monmonier's algorithm implemented in the program identifies boundaries associated with the highest genetic heterogeneity on a map where the samples are represented according to their geographical coordinates and are connected by Delaunay triangulation, with edges associated by genetic differentiation measures ($\hat{\theta}$). Analyses were conducted both for each of the 7 nuclear loci separately as well as for the whole nuclear data set. As there is no objective criterion to define the number of putative genetic boundaries, 4 to 6 implicit boundaries were tested to estimate reliability of the method. A bootstrapping procedure over the 7 genetic matrices was used to determine how many loci supported each barrier identified, and thus to verify the extent of data redundancy. To date, no published studies have applied TESS and BARRIER together, and their relative performance in locating discontinuities in allele frequencies have not been rigorously compared.

Finally, the MIGRATE-n v.3.0 program (<http://popgen.scs.fsu.edu/Migrate/Migrate-n.html>) was used to infer the mutation-scaled migration rate M ($M = m/\mu$, where m is the immigration rate per generation and μ is the mutation rate) among the most relevant groups

of the populations identified by the AMOVA, TESS and BARRIER using nuclear loci. The method assumes that the populations are at equilibrium with respect to drift and migration and that the population sizes and migration rates have remained constant over the coalescent period ($\sim 4N_e$ generations). A Brownian mutation model was used, and the mutation considered was constant for all loci. The coalescent ML based on MCMC with Hastings Metropolis importance sampling was used to infer the various parameters (Beerli & Felsenstein 1999, 2001). F_{ST} estimates ($\hat{\theta}$) among groups of populations were used as initial parameters for the estimation of Θ ($= 4N_e\mu$) and M in MIGRATE-n v.3.0. For each locus, the ML was run for 10 short and 5 long chains with 50 000 and 100 000 recorded genealogies, respectively, after discarding the first 10 000 genealogies (burn-in) for each chain. One of every 20 reconstructed genealogies was sampled for both short and long chains. An adaptive heating scheme with 4 concurrent chains was used; the analyses were run on a cluster computer using 1 master and 15 compute nodes. Analyses using MIGRATE-n were performed in 3 replicates. Preliminary tests were performed with or without the *prl-1* locus to ensure consistency of the results (i.e. assuming that this intronic locus may influence the results compared to the anonymous microsatellite loci), and results did not appear to be affected by the presence of this locus (not reported). Hence, we hereby report results based on the 6 nuclear loci. Unlike estimates of gene flow using F -statistics, MIGRATE-n allows asymmetrical migration rates among pairs of populations (e.g. $M_{A \rightarrow B}$ and $M_{B \rightarrow A}$) separated by barriers to gene flow as identified by TESS and BARRIER. Thus, the ratio of these estimates was used to report the estimated main direction of gene flow across recognised population breaks.

RESULTS

mtDNA

The genetic diversity and historical demography of *Mugil cephalus* in the NE Atlantic area were assessed by sequencing a 857 bp fragment of the *cytb* sequence in 184 individuals from 18 locations (31.83% of the 578 individuals studied for nuclear DNA; Table 1). Twenty-three sites were polymorphic, and 5 were parsimony informative in this fragment. Twenty-four mtDNA haplotypes of 29 recorded in this NE Atlantic lineage (82.76%) were identified in this study (Fig. 2), with the highest difference among

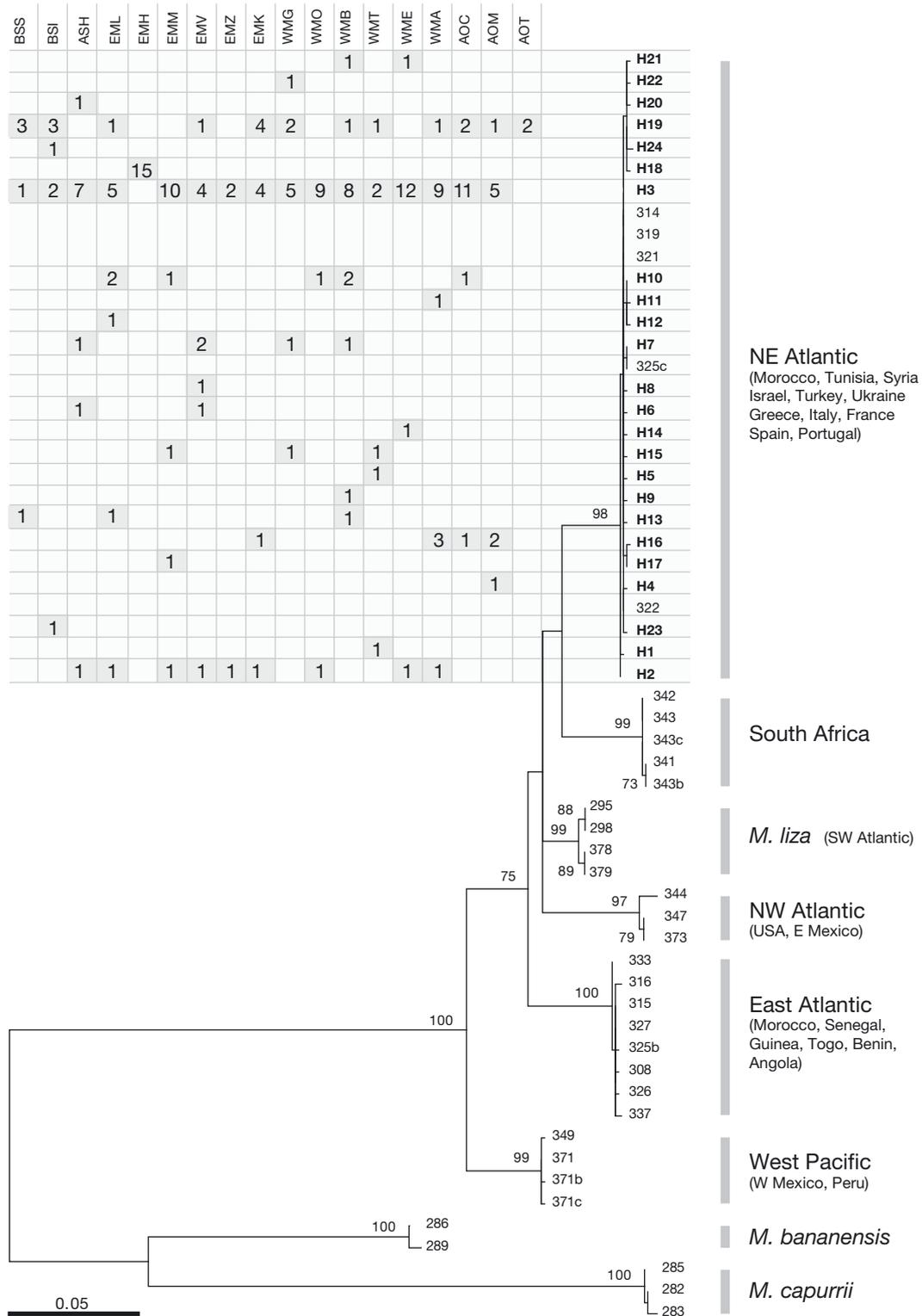


Fig. 2. *Mugil cephalus*. Phylogenetic tree depicting relationships among *M. cephalus* haplotypes sampled in the NE Atlantic and Mediterranean Sea, and some *M. cephalus* lineages (including *M. liza*) described in other geographic areas of the *M. cephalus* distribution range (Durand et al. 2012). Haplotypes found in this study are indicated in bold (H1 to H24) and respective number of observations at each sampling site are shown in tabular form. Relationships were inferred using partitioned maximum-likelihood (ML) analysis of 784 aligned nucleotides from the *cytb* gene. Branch lengths are proportional to the number of substitutions (scale bar: 0.05 substitutions per site) under the HKY + G model. Numbers on the branches are ML bootstrap values; those below 50% are not shown. A minimum spanning tree presenting the haplotype network is provided in Fig. A1 (Appendix 1)

expansion. On the overall population (or mtDNA clade) scale, F_S estimates were negative and significant, indicating departure from neutrality (Table 1).

Nuclear loci

A total of 578 *Mugil cephalus* from 17 locations were screened for nuclear loci (except for the Tagus population [AOT]; Table 1). Allelic richness and

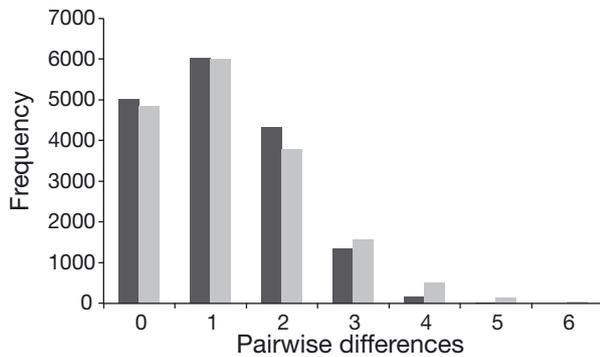


Fig. 3. *Mugil cephalus*. Mismatch distributions among mtDNA *cytb* sequences of *M. cephalus* in the NE Atlantic and Mediterranean Sea: the observed distribution (dark grey) versus the simulated distribution (light grey) under a population expansion model. Using the raggedness test, the population expansion assumption could not be rejected (raggedness index = 0.05, $p = 0.5$)

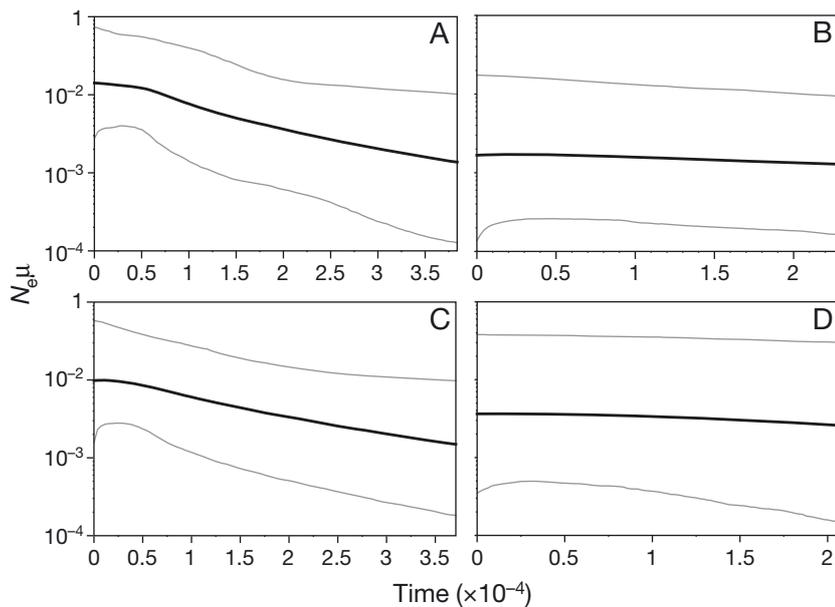


Fig. 4. *Mugil cephalus*. Bayesian skyline plots estimated from an alignment of *M. cephalus* sequences collected in (A) the NE Atlantic and Mediterranean Sea, (B) the Atlantic, (C) the Mediterranean Sea and (D) the Black Sea. The black line indicates the median estimate, whereas the grey lines indicate the 95% credibility intervals

observed levels of gene diversity were quite variable among these loci, ranging from 4 alleles at *prl-1* (H_{obs} : range 0.385 to 0.710) to 43 alleles at *MCS15AM* (H_{obs} : range 0.871 to 1.0). Eleven out of 119 sample-locus combinations deviated significantly from HWE after correction for multiple tests, and 5 out of 7 nuclear loci had such deviations (Table A1 in Appendix 1). This distribution makes the probability of a null allele specifically affecting a given locus very low. No significant linkage disequilibrium was detected within the populations ($p > 0.05$; details not reported). Also, no sign of a bottleneck based on Wilcoxon sign-rank tests was detected at the nuclear loci, irrespective of the model of mutation used for analysis (not shown).

A low but significant level of differentiation across samples was found at nuclear loci ($\hat{\theta} = 0.016$, $p < 0.001$). After correction for multiple tests, pairwise estimates of population differentiation revealed a genetic distinctness of the Black Sea samples (BSS, BSI), similar to mtDNA, but also of the 2 remaining Atlantic samples (AOC, AOM) (Table 2). Results associated with the classical estimation of population differentiation indicated 3 main genetic units in flathead mullet: the Atlantic, the Mediterranean, and the Black Sea. These 3 units were apparent when using BARRIER (Fig. 5), but more so when using TESS with $K_{max} = 3$ (this value was found to minimize DIC; details for alternative models not reported). The results can be interpreted as a differential introgression among parental genomes, for example, low but existing contributions of Atlantic and Black Sea parental genomes to the Mediterranean pool (Fig. 1; see below for additional results).

Besides these well-supported units, the eastern Mediterranean samples (EMM, EMV, ASH, EML and EMH) were different from each other when the corrections for multiple tests were not considered (Table 2). This indicates different population structures, being panmictic in the western but differentiated in the eastern Mediterranean. Supplementary subdivisions among eastern Mediterranean samples were confirmed by BARRIER, which identified genetic breaks between Greece and Libya and in the far eastern Mediterranean Sea, isolating the EML population (although only 4 loci showed differentiation from the neighbouring EMH sample; Fig. 5). TESS was inconclusive in identifying those areas as

observed levels of gene diversity were quite variable among these loci, ranging from 4 alleles at *prl-1* (H_{obs} : range 0.385 to 0.710) to 43 alleles at *MCS15AM* (H_{obs} : range 0.871 to 1.0). Eleven out of 119 sample-locus combinations deviated significantly from HWE after correction for multiple tests, and 5 out of 7 nuclear loci had such deviations (Table A1 in Appendix 1). This distribution makes the probability of a null allele specifically affecting a given locus very low. No significant linkage disequilibrium was detected within the populations ($p > 0.05$; details not reported). Also, no sign of a bottleneck based on Wilcoxon sign-rank tests was detected at the nuclear loci, irrespective of the model of mutation used for analysis (not shown).

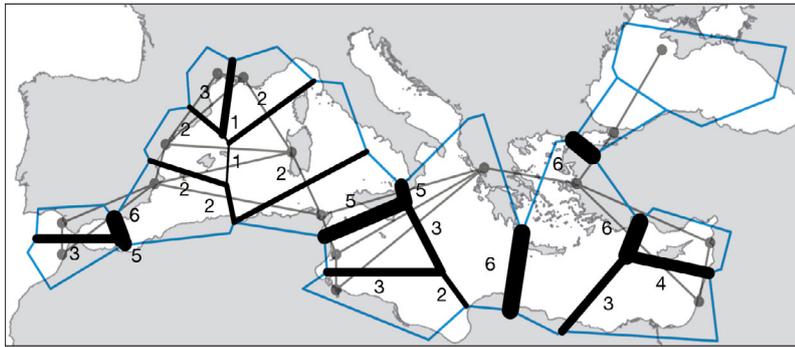


Fig. 5. *Mugil cephalus*. Genetic differences in *M. cephalus* in the Mediterranean Sea and adjacent seas, recovered with 6 microsatellite loci and 1 exon-primed intron-crossing locus. Sampled locations (grey dots) with corresponding Voronoï tessellation (connecting grey lines) and Delaunay triangulation (blue and black lines) were calculated using BARRIER 2.2. Reliability of gene flow barriers was estimated using a bootstrap procedure on 7 θ matrices (1 per nuclear locus). Line thickness of gene flow barriers is proportional to the bootstrap value, and numbers indicate the number of loci that show significant genetic heterogeneity (i.e. number of loci supporting a genetic break)

additional barriers to gene flow (i.e. supporting $K_{\max} = 6$, no improvement of DIC; details not reported).

A significant proportion of the explained genetic variance among groups was recovered using different models of population structure (i.e. different groups of samples based on basins, currents and different numbers of barriers) in AMOVA (Table 3). Among those models, only the partition of samples that included the 5 main geographic barriers estimated with BARRIER resulted in a non-significant nuclear genetic differentiation among groups (Φ_{SC} ; apparent decrease of the Wahlund effect; Table 3). Concurrently, the partition according to the 3 clusters delineated by TESS did not allow minimizing Φ_{SC} (Table 3).

Table 3. *Mugil cephalus*. AMOVA results considering different groups of samples based on basins, currents and different numbers of barriers; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Φ_{ST} : fixation index within samples; Φ_{SC} : fixation index among samples; Φ_{CT} : fixation index among groups. Sampling codes and other abbreviations are given in Table 1

Groups considered in the AMOVA		Fixation index	μsat		cytb	
			Φ	Variation (%)	Φ	Variation (%)
Black Sea	[BSS/BSI]	Φ_{ST}	0.0126***	98.74	0.2633***	73.66
East Mediterranean Sea	[ASH/EML/EMH/EMM/EMV/EMK/EMZ]	Φ_{SC}	0.0017**	0.16	0.2759***	28.07
West Mediterranean Sea	[WMO/WMB/WMT/WME/WMA/WMG]	Φ_{CT}	0.011***	1.10	-0.0174	-1.74
Atlantic Ocean	[AOC/AOM/AOT ^a]					
Black Sea	[BSS/BSI]	Φ_{ST}	0.0111***	98.89	0.2735***	72.65
Aegean Sea	[ASH]	Φ_{SC}	0.0026**	0.25	0.2245***	21.03
East Mediterranean Sea	[EMM/EMV]	Φ_{CT}	0.0086**	0.86	0.0632	6.32
Ionian Sea	[EML/EMH]					
SW Mediterranean Sea	[WMG/EMK/EMZ]					
NW Mediterranean Sea	[WMO/WMB/WMT/WME/WMA]					
Atlantic Ocean	[AOC/AOM/AOT ^a]					
Black Sea	[BSS/BSI]	Φ_{ST}	0.014***	98.59	0.2858***	71.77
Aegean & Ionian Sea	[ASH/EML/EMH]	Φ_{SC}	0.002***	0.16	0.2336***	21.77
Mediterranean Sea	[WMO/WMB/WMT/WME/WMA/WMG/EMM/EMV/EMK/EMZ]	Φ_{CT}	0.013***	1.25	0.0681	6.81
Atlantic Ocean	[AOC/AOM/AOT ^a]					
Black Sea	[BSS/BSI]	Φ_{ST}	0.012***	98.80	0.2778***	72.22
Aegean Sea	[ASH]	Φ_{SC}	0.001	0.05	0.2175***	20.07
East Mediterranean Sea	[EML/EMH]	Φ_{CT}	0.012***	1.15	0.077	7.7
Middle Mediterranean Sea	[EMK/EMZ/EMV/EMM]					
West Mediterranean Sea	[WMO/WMB/WMT/WME/WMA/WMG]					
Atlantic Ocean	[AOC/AOM/AOT ^a]					
Black Sea	[BSS/BSI]	Φ_{ST}	0.012***	98.81	0.3021***	69.79
Aegean Sea	[ASH]	Φ_{SC}	0.0001	0.01	0.0467*	3.42
NE Mediterranean Sea	[EML]	Φ_{CT}	0.012***	1.18	0.268**	26.8
SE Mediterranean Sea	[EMH]					
Middle Mediterranean Sea	[EMK/EMZ/EMV/EMM]					
West Mediterranean Sea	[WMO/WMB/WMT/WME/WMA/WMG]					
Atlantic Ocean	[AOC/AOM/AOT ^a]					

^asample only includes *cytb* test

Despite the problem associated with the EMH sample (see previous subsection), this grouping also minimized the percentage of variation associated with Φ_{SC} for mtDNA, compared to other models considered (Table 3). This indicates that results from BARRIER based on nuclear loci were applicable to mtDNA, despite some sampling bias at this genetic marker.

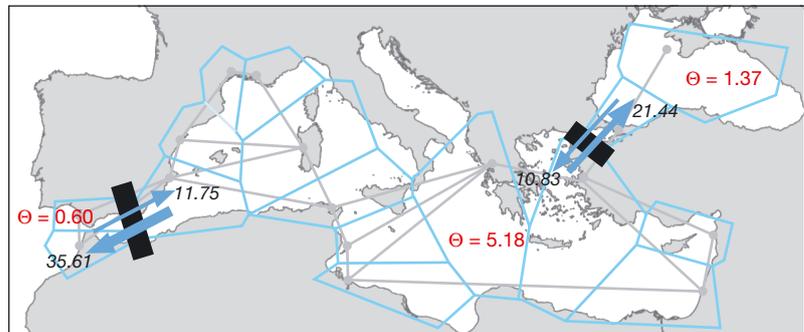
The use of TESS was nevertheless informative of the relative proportion of each cluster ($K_{max} = 3$) based on an average q_{ik} of individuals belonging to each sample. The relative proportions of each putative parental genome varied spatially (Fig. 1A). This allowed the testing of IBD using 3 alternative scenarios of dispersal: western, central and eastern (Fig. 1B). When considering the whole set of samples, there was no significant correlation between the genetic and geographic distances at nuclear loci ($Z = 35.21$, $p = 0.184$). Excluding peripheral populations from IBD analyses according to groupings in AMOVA (e.g. the Black Sea or Atlantic samples) did not reveal a significant IBD relationship either (not reported). A significant IBD pattern was found when the so-called western axis of dispersal ($Z = 21.55$, $p = 0.015$) was considered, but no significant IBD pattern was detected in the other 2 scenarios (central: $Z = 5.96$, $p = 0.096$; eastern: $Z = 1.47$, $p = 0.212$), possibly because the number of populations was lower and/or overall distance shorter than (i.e. the IBD slope/gradient was not sufficient to counteract sampling noise in estimates of $\hat{\theta}$) (Fig. 1B). When reconsidering IBD patterns at the mtDNA locus according to these 3 scenarios, the results were similar to those for microsatellites. Indeed, there was a significant IBD pattern for the so-called western dispersal scenario ($Z = 78.17$, $p = 0.048$ [10 000 permutations instead of 5000 to ensure significance at $\alpha = 0.05$]), whereas no IBD pattern was found for the other scenarios (central: $Z = 13.36$, $p = 0.264$; eastern: $Z = 29.70$, $p = 0.218$). This western IBD scenario was therefore indicated by both nuclear and mtDNA markers and provides further support for a differential introgression among the 3 main population clusters.

Interlocus g -tests unravelling past demography from nuclear data were computed according to the full data

set, then to the main results provided by TESS and BARRIER (i.e. for data sets composed of 3 and 6 population clusters, respectively). Regardless of the data set considered, results were globally negative to detect population expansion, when considering locus *prl-1* or not. Values of the g -statistic ranged from 0.86 to 1.62, when considering all 6 nuclear loci, and from 0.98 to 1.87, when excluding *prl-1*. None of these estimated g -statistics were found to report significant population expansion when compared with the fifth-percentile rejection values (≤ 0.10 and ≤ 0.08 , respectively) appropriate for the number of loci and sample sizes given in Reich et al. (1999; their Table 1).

Results from MIGRATE-n were also tested according to the main results provided by TESS and BARRIER (Fig. 6). While analyses were not performed on the same number of clusters, results were similar. Higher Θ values occurred in the Mediterranean in each case, and more specifically in the eastern Mediterranean basin when considering the larger number of genetic boundaries detected with BARRIER (Fig. 6). Consequently, lower Θ values occurred in peripheral

A) $K = 3$ (TESS)



B) $K = 6$ (BARRIER)

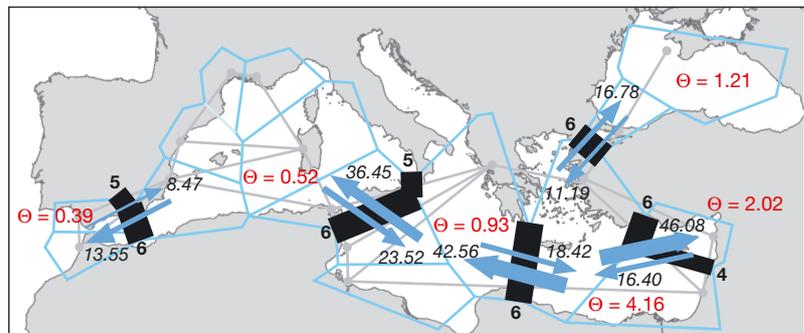


Fig. 6. *Mugil cephalus*. Estimates of historical effective population size (numbers in red) as estimated by Θ using MIGRATE-n among the most relevant groups of populations identified by both (A) TESS and (B) BARRIER v.2.2, using nuclear loci. Arrows depict the main direction of mutation-scaled migration rate among population groups. Italic numbers represent asymmetrical estimated number of migrants among adjacent groups of populations. Thickness of arrows roughly illustrates the strength of asymmetric migration rate from one population group to nearby groups. K : number of nuclear clusters recovered

Atlantic and Black Sea populations. In each case, the ratios of estimated numbers of migrants across the boundaries indicated a net flux of grey mullet from the Mediterranean population through other basins (Fig. 6), and especially from the eastern Mediterranean to other basins when considering $K = 6$ (Fig. 6B).

DISCUSSION

This study revealed patterns of population differentiation at mtDNA and nuclear markers in *Mugil cephalus* over an area spanning from the NE Atlantic to the Mediterranean and Black Seas. Previous genetic studies focussed on reciprocal monophyly and used limited individual samples, or considered smaller sampling areas and did not aim to detect genetic differentiation (Rossi et al. 1998, Blel et al. 2010, Livi et al. 2011). As the genetic structure of marine organisms (plants, invertebrates and vertebrates) has been extensively studied in the NE Atlantic and the Mediterranean and Black Seas (Patarnello et al. 2007), patterns of genetic differentiation for the flathead mullet can be compared with a plethora of other data. Only a few species have, however, been analysed for both mtDNA and nuclear markers in this area (e.g. 8 of 75 studies reviewed in Patarnello et al. 2007).

Confirmation of a single clade and mito-nuclear signals of population expansion

In agreement with Crosetti et al. (1994) and Livi et al. (2011), the mtDNA data in our study confirms the existence of a single clade over the Mediterranean and Black Seas, and extends it further to the NE Atlantic (Portugal, Morocco), although the latter requires further validation. The occurrence of a single clade is not surprising because the enclosed geography of the Mediterranean Sea and associated basins, and their complex paleo-climatic history (glacial–interglacial cycling) during the Pleistocene era, makes the whole area particularly prone to major demographic impacts (Patarnello et al. 2007). Similar to other fishes and invertebrates, which also exhibit unimodal mismatch distributions at mtDNA markers (review in Patarnello et al. 2007), the flathead mullet has probably undergone recent population expansion in this area, as also demonstrated by the star-like phylogeny of haplotypes, BSPs and the significant negative values of D and F_S . Hence, missing haplotypes probably do not affect the outcomes

of BSPs, the demographic results of which are qualitatively congruent to more classical approaches. The BSPs do, however, suggest rather continuous growth and did not indicate a sudden increase in effective population size (Grant et al. 2012). The BSPs have to be considered with caution because the star-like phylogenies seen here seem especially prone to biased inferences of quantitative aspects of past demographic processes (see Grant et al. 2012), while standard statistics such as Tajima's D appeared to be more buffered against biased inference when, for example, $\hat{\theta}$ is low (St Onge et al. 2012).

The average level of nucleotide diversity was low in this study ($\pi = 0.0013$), which is also consistent with some demographic crashes, in a manner comparable to the intra-lineage diversity revealed with the same mitochondrial gene in the northernmost NW Pacific *Mugil cephalus* lineage (Shen et al. 2011). A correlation between the distribution range of the different lineages and their nucleotide diversity has been established in the NW Pacific. A northernmost lineage, distributed across the Sea of Japan, was less genetically diverse ($\pi = 0.0003$, NWP1 clade in Shen et al. 2011) and demonstrated signs of demographic crashes during the Pleistocene era (non-equilibrium conditions), while temperate NW Pacific lineages had larger diversity (i.e. $\pi > 0.0030$; Shen et al. 2011) and exhibited patterns of population differentiation that reflected equilibrium conditions (Jamandre et al. 2009, Shen et al. 2011). Hence, results obtained independently in other lineages support the thesis that flathead mullet populations located at or close to the species' northernmost range limit (NW Pacific or Mediterranean and Black Seas) underwent rapid population expansion. While timing of population expansion was not evaluated in this study because of limitations to molecular clock calibrations (e.g. Ho et al. 2008, Grant et al. 2012), expansion revealed by the mtDNA marker in Mediterranean flathead mullet very probably arose during the Pleistocene, as signatures of older demographic events have shown to be erased by more recent events in marine fishes (Grant & Bowen 1998, Grant et al. 2012).

In the present study, patterns of genetic diversity at nuclear markers did not, however, suggest any bottleneck, and the interlocus g -tests did not provide nuclear support for population expansion either. When testing for bottlenecks as well as for population expansion at the nuclear level, negative results may be due to an insufficient number of loci tested (usually $n = 15$ or 20). Although it performs better than other methods when the number of loci is low, the Wilcoxon sign-rank test also has more power

when the number of loci increases (Luikart et al. 1998). Regarding the interlocus g -test, Reich et al. (1999) also reported that the use of more loci increased the power to reject size constancy. To our knowledge, the interlocus g -test has rarely been used in marine species, and, in each case of use it was with a relatively low number of microsatellite loci (e.g. Babbucci et al. 2010, Canino et al. 2010a,b, Gaffney et al. 2010). It is therefore difficult to reconcile patterns affecting the demographic history of mtDNA and nuclear data in this study, maybe because of an insufficient number of nuclear markers. Such results are, however, not entirely irrelevant because biparentally inherited nuclear genes are expected to have effective population sizes (N_e) that are 4 times those of uniparentally inherited genes such as mtDNA markers. This makes nuclear genes less susceptible to erosion of neutral genetic variation during population bottlenecks and less prone to detect population expansion resulting from founder events or population recovery events after range contraction. At some values of N_e , random drift may reduce mtDNA gene diversities without affecting nuclear gene diversities (e.g. Canino et al. 2010b). This may have occurred in *Mugil cephalus*, possibly explaining why mtDNA did not appear to be at population equilibrium as inferred by summary statistics such as, e.g. Tajima's D and mismatch distributions, while nuclear markers may have reached this equilibrium as illustrated by detection of significant IBD, at least under a scenario involving larger distances among populations (the so-called 'western scenario of dispersal'). Babbucci et al. (2010) reported a similar pattern of population expansion for mtDNA and population stasis for nuclear marker in the spiny lobster *Palinurus elephas* in the NE Atlantic.

It has also been suggested that lower mtDNA diversity may result from various selective effects at such markers (reviewed in Galtier et al. 2009). Some selective events acting on mtDNA may lead to star-like phylogenies and mimic demographic population expansion (Babbucci et al. 2010, Canino et al. 2010b). Demographic and selective scenarios are not mutually exclusive, and are, therefore, difficult to disentangle when using a single mtDNA marker (Haney et al. 2010) and when regarding a single lineage as done here. Haney et al. (2010) demonstrated that distinct mtDNA markers provided different pictures of demographic history in the Caribbean reef fish *Halochoeres bivittatus* and that inference of population growth based on the 2 mtDNA markers they used (*COX1*, *ATPase6-8*) could be an artefact of selection on these mitochondrial proteins compared to the picture pro-

vided by a non-coding (neutral) mtDNA region (control region, *CR*) and a nuclear-encoded marker. Population expansion of *Mugil cephalus* in the NE Atlantic and the Mediterranean should therefore be confirmed using other mtDNA markers such as *CR* and/or sequences of nuclear genes. However, it should be noted that in 2 of the 3 known lineages of the flathead mullet inhabiting the NW Pacific, the *COX1* and *cytb* markers exhibited reduced nucleotide diversities compared to *CR* diversity (not assessed for the third lineage; data compiled in Shen et al. 2011). However, as for *cytb*, *CR* also showed a star-like phylogeny and unimodal mismatch distribution for the northern lineage (Jamandre et al. 2009), indicating in this case that varying levels of nucleotide diversity did not translate into conflicting population histories as reported in Haney et al. (2010). Furthermore, numerous studies of mtDNA diversity in marine fish species within the Mediterranean area were based on *CR* polymorphisms, and most of them also reported a single mtDNA lineage and rapid population expansion in this area, as found for *M. cephalus* with *cytb* (e.g. Atarhouch et al. 2006, Charrier et al. 2006). It is unlikely that marine fish species with very distinct life-histories such as sardines (Atarhouch et al. 2006) and anglerfishes (Charrier et al. 2006) were affected by selection acting on mtDNA, producing similar star-like phylogenies. In light of these observations, conformity to neutral expectation and interpreting results as a signature of population expansion appears more parsimonious than a selection-based hypothesis.

Overall, demographic inferences suggest that the reduction in population size followed by the population expansion that occurred in *Mugil cephalus* and in some marine invertebrates such as spiny lobster in this area was possibly not strong enough to affect nuclear variability (see also Canino et al. 2010b for a reported case in the NW Pacific of Atka mackerel *Pleurogrammus monopterygius*). Repeated observations of possible, distinct demography should motivate further comparative studies of mito-nuclear genetic variation to decipher the role of neutral and selective processes.

MtDNA population differentiation

Although a single mtDNA clade/lineage was recovered, significant population differentiation was found at the *cytb* marker between Mediterranean and Black Sea samples ($\Phi_{ST} = 0.17$; $p = 0.029$). The presence in the Black Sea of a distinct population has already been reported in other marine species, including sea

grasses, fishes and invertebrates using mtDNA (Patarnello et al. 2007). To our knowledge, among the marine fish species found in the area, only the anchovy *Engraulis encrasicolus*, the sprat *Sprattus sprattus* and the pipefish *Syngnathus typhle* have been screened for genetic differentiation between the Mediterranean and Black Seas (Magoulas et al. 2006, Debes et al. 2008, Wilson & Eigenmann Veraguth 2010, respectively). The particular status of this basin was confirmed in the present study, but contrary to anchovy and pipefish (Magoulas et al. 2006, Wilson & Eigenmann Veraguth 2010) and other studies on invertebrates (e.g. Peijnenburg et al. 2006), reduction of mtDNA diversity in Black Sea samples was not observed. In the Black Sea samples of *Mugil cephalus* 3 to 4 haplotypes were found, whereas the eastern Mediterranean contained 5 to 6 haplotypes (except at EMH). This could only be due to low samples sizes for mtDNA in this study, with less sampling effort in the Black Sea compared to the Mediterranean Sea explaining lower observed diversity.

It should also be noted that Haplotype H18 was restricted to the Israeli (EMH) sample, while others were widely distributed. Together with Black Sea samples, this haplotype significantly affected mtDNA genetic differentiation, while nuclear markers did not show this sample to be genetically distinct. This haplotype clearly belongs to the single mtDNA clade depicted in our study, and was not related to other known foreign clades (Fig. 2). The presence of this apparently private haplotype in a dispersive species such as *Mugil cephalus* needs further examination, but it probably did not arise from repeated sequencing errors in all individuals from this sampling location.

Differentiation at nuclear markers and possible gender-biased dispersal

The inclusion of nuclear markers shed complementary but distinct light on the population structure revealed by the mtDNA in the flathead mullet. Indeed, while levels of genetic differentiation were ca. 10 to 15 times lower than those observed with mtDNA, nuclear markers demonstrated finer genetic structuring and more genetic breaks than the uniparentally inherited marker. They identified the NE Atlantic, the Mediterranean and the Black Seas as genetically differentiated units. Genetic breaks located at the Almeria-Oran front, in the Siculo-Tunisian Strait and south of the Greek Peninsula are widely reported in the literature (Patarnello et al. 2007).

Hence, the geographical patterns of nuclear genetic differentiation in *Mugil cephalus* match those already reported in numerous marine species. Nevertheless, possibly due to insufficient sampling in some areas (e.g. eastern Mediterranean Sea; see below), these genetic breaks are typically not detected simultaneously in a single species (but see Rolland et al. 2007). Furthermore, when differentiation is recovered with nuclear markers in these areas, it is also detected with mtDNA (e.g. Calderón et al. 2008; more references in Patarnello et al. 2007). That is clearly not the case in *M. cephalus*, as the only pattern of genetic differentiation shared by both types of markers was the distinctness of the Black Sea. As previously mentioned for mtDNA, studies of genetic differentiation between the Mediterranean and Black Seas using nuclear genetic markers are also rare, being restricted so far to pipefish (Wilson & Eigenmann Veraguth 2010), flat oyster *Ostrea edulis* (Launey et al. 2002) and seagrass *Zostera marina* (Olsen et al. 2004). Studies reported nuclear genetic differentiation between the Mediterranean and Black Seas, together with lower allelic richness and/or gene diversity in Black Sea samples. As for mtDNA, this decrease in diversity was not really observed in *M. cephalus*, even though the Istanbul sample (BSI) had the lowest gene diversity.

Partial discrepancy between levels of diversity and/or inferred patterns and levels of genetic differentiation at each class of loci is not unusual and has been reported on several occasions for organisms from the Atlantic and Mediterranean (e.g. flat oyster: Diaz-Almela et al. 2004; sea bass: Lemaire et al. 2005; spiny lobster: Babbucci et al. 2010), or other marine species distributed worldwide (e.g. Buonaccorsi et al. 2001, Canino et al. 2010b). Because of the 4-fold lower N_e , estimates of population differentiation are expected to be higher for mitochondrial markers (e.g. Lemaire et al. 2005). At migration–drift equilibrium, an expected 4:1 ratio in mtDNA versus nuclear levels of genetic differentiation should theoretically be observed. Deviations from this ratio of the levels of population differentiation among markers may result from the differential effects of genetic drift, mutation and migration on a marker class, or may result from selection (i.e. general violations of the migration–drift equilibrium assumptions), or sex-biased dispersal (e.g. Buonaccorsi et al. 2001). Indeed, sexes often differ in their degree of dispersal and, hence, in their contribution to spatial genetic structure both within and among populations (Handley & Perrin 2007). Asymmetric migration rates among sexes, but also spatio-temporal variation in sex ratio will then facili-

tate differential genetic structuring between nuclear and mitochondrial markers (e.g. Consuegra & García de Leániz 2007). Some authors are proponents of purely neutral processes. Differences in the magnitude of estimated population subdivision from nuclear and mtDNA markers could be accounted for entirely by differences in N_e and variance in estimates of population differentiation among loci (e.g. Buonaccorsi et al. 2001, Wilson & Eigenmann Veraguth 2010). In contrast, other authors have indicated that such differences cannot be accounted for without selection (Bensch et al. 2006, Canino et al. 2010b) and/or without sex-biased dispersal or sex-biased philopatry (FitzSimmons et al. 1997). As discussed for other marine species (e.g. *Pleuronectes platessa*: Hoarau et al. 2004), assumptions of the mutation-drift equilibrium in flathead mullet are certainly violated, as indicated by probable recent population expansion (above), selection is poorly supported (above) and large differences in levels of population differentiation resulting in different genetic structures at marker loci cannot be easily interpreted. Data suggest, however, that male philopatry (female-biased gene flow) could be a component of increased genetic differentiation at nuclear loci. Unfortunately, there is poor empirical support for this interpretation as, while movements and migrations of flathead mullet have been extensively investigated, sex or sex ratio of migrant individuals was not considered (review in Whitfield et al. 2012).

Reliability of nuclear genetic breaks

Results obtained by TESS and BARRIER did not match exactly, as TESS identified $K = 3$ versus $K = 6$ population clusters with BARRIER, but the 3 main clusters were identified by both methods. Indeed, spatial information is not incorporated in the same manner in each method. Basically, TESS considers both local and global trends of allelic variation that would allow finer modelling of admixture proportions (François & Durand 2010). BARRIER only considers local trends among neighbouring populations without considering distant ones; hence, it considers less available information than TESS. Detailed comparisons of the outcomes provided by each method using simulations are necessary. In the present case, however, the 3 additional clusters found by BARRIER are biologically relevant, as the Siculo-Tunisian Strait and the area south of the Peloponnesus are recognized barriers to gene flow in numerous marine fish species (e.g. Bahri-Sfar et al. 2000, Yebra et al. 2011).

Differentiation of populations within the eastern Mediterranean Sea, such as the isolation of the EML sample revealed by BARRIER, has rarely been reported. This area is poorly studied in the literature because of a bias in sampling, with western and central areas of the Mediterranean overrepresented in the published literature. When considered, sampling in the Aegean Sea and the eastern Mediterranean basin is often restricted to 1 or 2 samples that strongly limit our knowledge of the genetic structure of taxa within this basin. By considering 5 distinct samples of *Mugil cephalus* in this basin, our study is a notable exception. To our knowledge, differentiation at nuclear loci within the eastern Mediterranean basin has only been reported for sea bass (Bahri-Sfar et al. 2000, Castilho & Ciftci 2005). These studies support genetic differentiation of fish populations within the eastern Mediterranean basin, such as differentiation between Tunisian and Aegean populations (Bahri-Sfar et al. 2000) and differentiation among Aegean and Turkish populations (Castilho & Ciftci 2005), as reported there for *M. cephalus*. Results indicate the need for further investigation of the genetic structure of taxa within the eastern Mediterranean basin, in order to better manage important marine resources. For example, *M. cephalus* is highly prized in adjoining countries such as Tunisia, Egypt and Israel (Whitfield et al. 2012).

A possible history of dispersal: from the Mediterranean to the margins

Interestingly, the estimates of Θ based on clusters detected by TESS and BARRIER indicated larger effective population size for *Mugil cephalus* in the eastern Mediterranean basin compared to in more peripheral populations (basins). This indicates an asymmetric migration, with a net flux of migrants from the Mediterranean basin to adjacent basins with $K = 3$ as inferred with TESS. When K was set to 6, as suggested by BARRIER, the net flux of migrants was directed from the eastern to the adjacent western cluster (except between the 2 easternmost Mediterranean divisions). Values of Θ decreased from eastern (Aegean Sea) to western (Atlantic) population clusters.

However, these results obtained with nuclear markers have to be interpreted with caution because populations must be at demographic equilibrium to correctly estimate these values, and this equilibrium is equivocal between mtDNA and nuclear markers. The spatial variation of the coefficient of membership (i.e. mean q_{ik} values) of the Black Sea parental popu-

lation with TESS seems, at least intuitively, to indicate exchange of individuals from this basin to the Mediterranean basin and that the net flux of migrants estimated with MIGRATE-n could be reversed. However, the very recent origin of the Black Sea (~9000 yr BP; Major et al. 2006) does not support this interpretation, and genetic markers that are at low frequency in the Mediterranean Sea but that provide a far higher contribution to the Black Sea gene pool have already been reported. In the anchovy, Magoulas et al. (2006) reported very low frequency of Black Sea mtDNA haplotypes within the Mediterranean, whereas they had very high frequency within the Black Sea. We would interpret our flathead mullet results as an increase in frequency of some low-frequency Mediterranean alleles in the Black Sea, rather than the reverse. Furthermore, the eastern Mediterranean Sea is considered to have been a refuge during glaciations, the resident organisms of which dispersed when Atlantic and Mediterranean waters came into contact (Patarnello et al. 2007). Hence, as inferred with MIGRATE-n, a larger estimated population size in this area compared to more peripheral populations is expected under this scenario, together with migration predominantly occurring from this refuge to other areas. In marine fishes, the east to west migration was also suggested by, e.g., Aurelle et al. (2003) and Lemaire et al. (2005) when studying the Atlantic–Mediterranean divide at the Almeria-Oran front (i.e. the westernmost genetic break detected in this study). In *Syngnathus typhle*, Wilson & Eigenmann Veraguth (2010) also suggested migration from the Mediterranean to peripheral seas, but, as the eastern Mediterranean basin was poorly sampled, comparison with the flathead mullet cannot be made.

Overall, patterns of dispersal inferred by the nuclear markers in this study indicated that the predominant flux of migrants was against the direction of the main currents connecting those basins, as indicated in Fig. 1. This illustrates that the historical process—albeit certainly recent—shaped the genetic variation of *Mugil cephalus* in this area.

CONCLUSIONS

For the first time, nuclear genetic variation at the scale of a single basin was demonstrated in the cosmopolitan species *Mugil cephalus*. To date, such variation has only been detected at an inter-oceanic scale using allozymes (Crosetti et al. 1994), or in the NW Pacific, where several cryptic species are pres-

ent (e.g. Liu et al. 2009, Shen et al. 2011). As is the case for other marine species, the Atlantic, Mediterranean and Black Seas were the main units that structured genetic variation, although further genetic divisions were recorded within the Mediterranean Sea. All the population clusters found in this study with nuclear markers have been reported in the literature, but this is the first time they have been detected together in a single species. Patterns of population structure across markers are coherent despite less genetic structure being detected using mtDNA and past differences in demography being inferred by each class of markers. The data suggest the existence of a refugium in the eastern Mediterranean for *M. cephalus*, which then dispersed to the Black Sea and to the proximal Atlantic. Further work using more nuclear and mtDNA loci and considering relationships within the NE Atlantic and among the close *M. cephalus* clades delineated by Durand et al. (2012) are, however, necessary to gain information on the historical demography of this species.

Acknowledgements. We thank all the people who participated in the sampling: O. Akyol, A. P. Apostolidis, C. Perdicaris, J. A. Balbuena, P. Berrebi, H. Cabral, L. Euzet, G. Leptra, M. Hassan, P. Merella, B. Morales Nin, H. Nouiri, H. Rosenfeld, V. Sarabeev, Z. Smolenicka and E. Ünlü. We acknowledge M. Pagès for her help with Bayesian skyline plots, O. François and R. Vitalis for further counselling and pointing out the appropriate data analysis for the present data set. K. Belkhir is acknowledged for implementing programs on a bioinformatic cluster. Thanks to D. McKenzie for English language editing. This study was funded by IRD-UR 070, UMR5119, and the MUGIL program (INCO-DEV-SSA-1) and MULTRACE program (ORCHID 2011) of the National Science Council of Taiwan (NSC100-2911-I-291-501-MY2) and the French Ministry of Foreign Affairs.

LITERATURE CITED

- Atarhouch T, Rüber L, Gonzalez EG, Albert EM, Rami M, Dakkak A, Zardoya R (2006) Signature of an early genetic bottleneck in a population of Moroccan sardines (*Sardina pilchardus*). *Mol Phylogenet Evol* 39:373–383
- Aurelle D, Guillemaud T, Afonso P, Morato T, Wirtz P, Santos RS, Cancela ML (2003) Genetic study of *Coris julis* (Osteichthyes, Perciformes, Labridae) evolutionary history and dispersal abilities. *C R Biol* 326:771–785
- Babbucci M, Buccoli S, Cau A, Cannas R and others (2010) Population structure, demographic history, and selective processes: contrasting evidences from mitochondrial and nuclear markers in the European spiny lobster *Palinurus elephas* (Fabricius, 1787). *Mol Phylogenet Evol* 56: 1040–1050
- Bahri-Sfar L, Lemaire C, BenHassine OK, Bonhomme F (2000) Fragmentation of sea bass populations in the western and eastern Mediterranean as revealed by

- microsatellite polymorphism. *Proc R Soc Lond B* 267: 929–935
- Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* 152:763–773
- Beerli P, Felsenstein J (2001) Maximum likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Proc Natl Acad Sci USA* 98:4563–4568
- Bensch S, Irwin DE, Irwin JH, Kvist L, Åkesson S (2006) Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Mol Ecol* 15:161–171
- Bilgin R (2007) Kgttests: a simple excel macro program to detect signatures of population expansion using microsatellites. *Mol Ecol Notes* 7:416–417
- Blel H, Panfili J, Guinand B, Berrebi P, Said K, Durand JD (2010) Selection footprint at the first intron of the *Prl* gene in natural populations of the flathead mullet (*Mugil cephalus* L., 1758). *J Exp Mar Biol Ecol* 387:60–67
- Buonaccorsi VP, McDowell JR, Graves JE (2001) Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). *Mol Ecol* 10:1179–1196
- Calderón I, Giribet G, Turon X (2008) Two markers and one history: phylogeography of the edible common sea urchin *Paracentrotus lividus* in the Lusitanian region. *Mar Biol* 154:137–151
- Campton DE, Mahmoudi B (1991) Allozyme variation and population structure of striped mullet (*Mugil cephalus*) in Florida. *Copeia* 1991:485–492
- Canino MF, Spies IB, Cunningham KM, Hauser L, Grant WS (2010a) Multiple ice-age refugia in Pacific cod, *Gadus macrocephalus*. *Mol Ecol* 19:4339–4351
- Canino MF, Spies IB, Lowe SA, Grant WS (2010b) Highly discordant nuclear and mitochondrial DNA diversities in Atka Mackerel. *Mar Coast Fish Dynam Manag Ecosys Sci* 2:375–387
- Castilho R, Ciftci Y (2005) Genetic differentiation between close eastern Mediterranean *Dicentrarchus labrax* (L.) populations. *J Fish Biol* 67:1746–1752
- Charrier G, Chenel T, Durand JD, Girard M, Quiniou L, Laroche J (2006) Discrepancies in phylogeographical patterns of two European anglerfishes (*Lophius budegassa* and *Lophius piscatorius*). *Mol Phylogenet Evol* 38: 742–754
- Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol Ecol Notes* 7:747–756
- Consuegra S, García de Leániz C (2007) Fluctuating sex ratios, but no sex-biased dispersal, in a promiscuous fish. *Evol Ecol* 21:229–245
- Crosetti D, Nelson WS, Avise J (1994) Pronounced genetic structure of mitochondrial DNA among populations of the circumglobally distributed grey mullet (*Mugil cephalus* Linnaeus). *J Fish Biol* 44:47–58
- Debes PV, Zachos FE, Hanel R (2008) Mitochondrial phylogeography of the European sprat (*Sprattus sprattus* L., Clupeidae) reveals isolated climatically vulnerable populations in the Mediterranean Sea and range expansion in the Northeast Atlantic. *Mol Ecol* 17:3873–3888
- Diaz-Almela E, Boudry P, Launey S, Bonhomme F, Lapègue S (2004) Reduced female gene flow in the European flat oyster *Ostrea edulis*. *J Hered* 95:510–516
- Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* 22: 1185–1192
- Durand E, Jay F, Gaggiotti OE, François O (2009) Spatial inference of admixture proportions and secondary contact zones. *Mol Biol Evol* 26:1963–1973
- Durand JD, Shen KN, Chen WJ, Jamandre BW and others (2012) Systematics of the grey mullet (Teleostei: Mugiliformes: Mugilidae): molecular phylogenetic evidence challenges two centuries of morphology-based taxonomic studies. *Mol Phylogenet Evol* 64:73–92
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10: 564–567
- 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
- FitzSimmons NN, Moritz C, Limpus CJ, Pope L, Prince R (1997) Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male biased gene flow. *Genetics* 147:1843–1854
- François O, Durand E (2010) Spatially explicit Bayesian clustering models in population genetics. *Mol Ecol Resour* 10:773–784
- Fu YX (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics* 147:915–925
- Gaffney PM, Pascal CM, Barnhart J, Grant WS, Seeb JE (2010) Genetic homogeneity of weathervane scallops (*Patinopecten caurinus*) in the northeastern Pacific. *Can J Fish Aquat Sci* 67:1827–1839
- Galtier N, Nabholz B, Glémin S, Hurst GDD (2009) MtDNA as a marker of molecular diversity: a reappraisal. *Mol Ecol* 18:4541–4550
- Gonzalez EG, Zardoya R (2007) Relative role of life-history traits and historical factors in shaping genetic population structure of sardines (*Sardina pilchardus*). *BMC Evol Biol* 7:197, doi:10.1186/1471-2148-7-197
- Grant WAS, Bowen WS (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *J Hered* 89:415–426
- Grant WS, Liu M, Tao TX, Yanagimoto T (2012) Limits of Bayesian skyline plot analysis of mtDNA sequences to infer historical demographies in Pacific herring (and other species). *Mol Phylogenet Evol* 65:203–212
- Haney RA, Silliman BR, Rand DM (2010) Effects of selection and mutation on mitochondrial variation and inferences of historical population expansion in a Caribbean reef fish. *Mol Phylogenet Evol* 57:821–828
- Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum Biol* 66:591–600
- Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish Fish* 9:333–362
- Heras S, Roldan MI, Gonzalez Castro M (2009) Molecular phylogeny of Mugilidae fishes revised. *Rev Fish Biol Fish* 19:217–231
- Ho SYW, Saarma U, Barnett R, Haile J, Shapiro B (2008) The effect of inappropriate calibration: three case studies in molecular ecology. *PLoS One* 3:e1615

- Hoarau G, Piquet AMT, van der Veer HW, Rijnsdorp AD, Stam WT, Olsen JL (2004) Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: a comparison of resolving power between microsatellites and mitochondrial DNA data. *J Sea Res* 51:183–190
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806
- Jamandre BW, Durand JD, Tzeng WN (2009) Phylogeography of the flathead mullet *Mugil cephalus* in the north-west Pacific as inferred from the mtDNA control region. *J Fish Biol* 75:393–407
- Koutrakis ET, Conides A, Parpoura AC, Van Ham EH, Katselis G, Koutsikopoulos C (2007) Lagoon fisheries' resources in Hellas, Chapter 6. In: Papaconstantinou C, Zenetos A, Vassilopoulou V, Tserpes G (eds) State of Hellenic fisheries. Hellenic Centre for Marine Research, Athens, p 223–234
- Launey S, Ledu C, Boudry P, Bonhomme F, Naciri-Graven Y (2002) Geographic structure in the European flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. *J Hered* 93:331–338
- Lawson Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex-biased dispersal. *Mol Ecol* 16:1559–1578
- Lemaire C, Versini JJ, Bonhomme F (2005) Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. *J Evol Biol* 18:70–80
- Liu JY, Lun ZR, Zhang JB, Yang TB (2009) Population genetic structure of striped mullet, *Mugil cephalus*, along the coast of China, inferred by AFLP fingerprinting. *Biochem Syst Ecol* 37:266–274
- Livi S, Sola L, Crosetti D (2011) Phylogeographic relationships among worldwide populations of the cosmopolitan marine species, the striped gray mullet (*Mugil cephalus*), investigated by partial cytochrome *b* gene sequences. *Biochem Syst Ecol* 39:121–131
- Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Mol Ecol* 7:963–974
- Magoulas A, Castilho R, Caetano S, Marcato S, Patarnello T (2006) Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (*Engraulis encrasicolus*). *Mol Phylogenet Evol* 39:734–746
- Major CO, Goldstein SL, Ryan WBF, Lericolais G, Piotrowski AM, Hajdas I (2006) The co-evolution of Black Sea level and composition through the last deglaciation and its paleoclimatic significance. *Quat Sci Rev* 25:2031–2047
- Manni F, Guérard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Hum Biol* 76:173–190
- Miggiano E, Lyons RE, Li Y, Dierens LM, Crosetti D, Sola L (2005) Isolation and characterization of microsatellite loci in the striped mullet, *Mugil cephalus*. *Mol Ecol Notes* 5:323–326
- Olsen JL, Stam WT, Coyer JA, Reusch TBH and others (2004) North Atlantic phylogeography and large-scale population differentiation of the seagrass *Zostera marina* L. *Mol Ecol* 13:1923–1941
- Patarnello T, Volckaert FAMJ, Castilho R (2007) Pillars of Hercules: Is the Atlantic–Mediterranean transition a phylogeographical break? *Mol Ecol* 16:4426–4444
- Peijnenburg KT, Fauvelot C, Breeuwer JA, Menken SB (2006) Spatial and temporal genetic structure of the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. *Mol Ecol* 15:3319–3338
- Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered* 90:502–503
- Reich DE, Feldman MW, Goldstein DB (1999) Statistical properties of two tests that use multilocus data sets to detect population expansions. *Mol Biol Evol* 16:453–466
- Rocha-Olivares A, Garber NM, Stuck KC (2000) High genetic diversity, large interoceanic divergence and historical demography of the striped mullet. *J Fish Biol* 57:1134–1149
- Rolland JL, Bonhomme F, Lagardère F, Hassan M, Guinand B (2007) Population structure of the common sole (*Solea solea*) in the northeastern Atlantic and the Mediterranean Sea: revisiting the divide with EPIC markers. *Mar Biol* 151:327–341
- Rossi AR, Capula M, Crosetti D, Sola L, Campton DE (1998) Allozyme variation in global populations of striped mullet, *Mugil cephalus* (Pisces: Mugilidae). *Mar Biol* 131:203–212
- Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* 145:1219–1228
- Ruzzante DE, Walde SJ, Gosse JC, Cussac VE, Habit E, Zemlak TS, Adams ED (2008) Climate control on ancestral population dynamics: insight from Patagonian fish phylogeography. *Mol Ecol* 17:2234–2244
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Selkoe KA, Henzler CM, Gaines SD (2008) Seascape genetics and the spatial ecology of marine populations. *Fish Fish* 9:363–377
- Sevilla RG, Diez A, Noren M, Mouchel O and others (2007) Primers and polymerase chain reaction conditions for DNA barcoding teleost fish based on the mitochondrial cytochrome *b* and nuclear rhodopsin genes. *Mol Ecol Notes* 7:730–734
- Shen KN, Jamandre BW, Hsu CC, Tzeng WN, Durand JD (2011) Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet *Mugil cephalus*. *BMC Evol Biol* 11:83
- St Onge KR, Palmé AE, Wright SI, Lascoux M (2012) Impact of sampling schemes on demographic inference: an empirical study in two species with different mating systems and demographic histories. *G3* 2:803–814
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
- Thomson JM (1997) The Mugilidae of the world. *Mem Queensl Mus* 41:457–562

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* 38: 1358–1370

Whitfield AK, Panfili J, Durand JD (2012) A global review of the cosmopolitan flathead mullet *Mugil cephalus* Linnaeus, 1758 (Teleostei: Mugilidae), with emphasis on the biology, genetics, ecology and fisheries aspects of this apparent species complex. *Rev Fish Biol Fish* 22:

641–681

Wilson AB, Eigenmann Veraguth I (2010) The impact of Pleistocene glaciation across the range of a widespread European coastal species. *Mol Ecol* 19:4535–4553

Yebra L, Bonnet D, Harris RP, Lindeque PK, Peijnenburg KTCA (2011) Barriers in the pelagic: population structuring of *Calanus finmarchicus* and *C. euxinus* in European waters. *Mar Ecol Prog Ser* 428:135–149

APPENDIX 1

Table A1. *Mugil cephalus*. \hat{f} estimate of Weir & Cockerham's (1984) equivalent of Wright's fixation index; average values across all 7 nuclear loci. **Bold** values significant after Bonferroni correction. **p* < 0.025, ****p* < 0.001. Sampling codes are given in Table 1

	<i>Pri-1</i>	<i>MCS2FH</i>	<i>MCS15CM</i>	<i>MCS15AM</i>	<i>MCS1EH</i>	<i>MCS2DM</i>	<i>MCS16DM</i>
BSS	0.002	-0.002	-0.109	-0.074***	-0.049	0.000	0.060
BSI	0.277	-0.059	0.011	0.014	0.063	-0.017	0.034
ASH	-0.199	-0.003	0.044	-0.020	-0.079	-0.101***	-0.119
EML	-0.127	-0.063	-0.062	0.072	0.028	-0.045	0.039
EMH	0.077	-0.079	0.037	-0.059***	0.001	0.078	0.144
EMM	0.358	-0.080***	0.087	0.057	-0.036	0.175	-0.257***
EMV	0.127	-0.054	0.087	0.019	-0.061	-0.032	0.003
EMZ	-0.083	-0.045	-0.043	0.008	0.027	-0.015	-0.016
EMK	-0.065	-0.011	-0.040	-0.073***	0.001	-0.041	-0.157*
WMG	-0.055	0.000	-0.029	-0.029	-0.016	0.035	-0.153
WMO	0.200	-0.039	-0.028	-0.029	-0.106***	0.010	-0.074
WMB	-0.093	-0.056	-0.017	0.006	0.020	-0.004	-0.140
WMT	-0.068	-0.104***	0.100	-0.015	-0.034	-0.055***	-0.228***
WME	-0.184	0.022	0.016	-0.067***	-0.056	0.037	-0.173
WMA	0.180	0.080	0.010	0.052	0.043	-0.026	-0.170
AOC	0.036	-0.042	-0.016	0.053	-0.031	0.004	-0.040
AOM	-0.075	0.025	0.072	-0.026	-0.003	0.006	-0.084

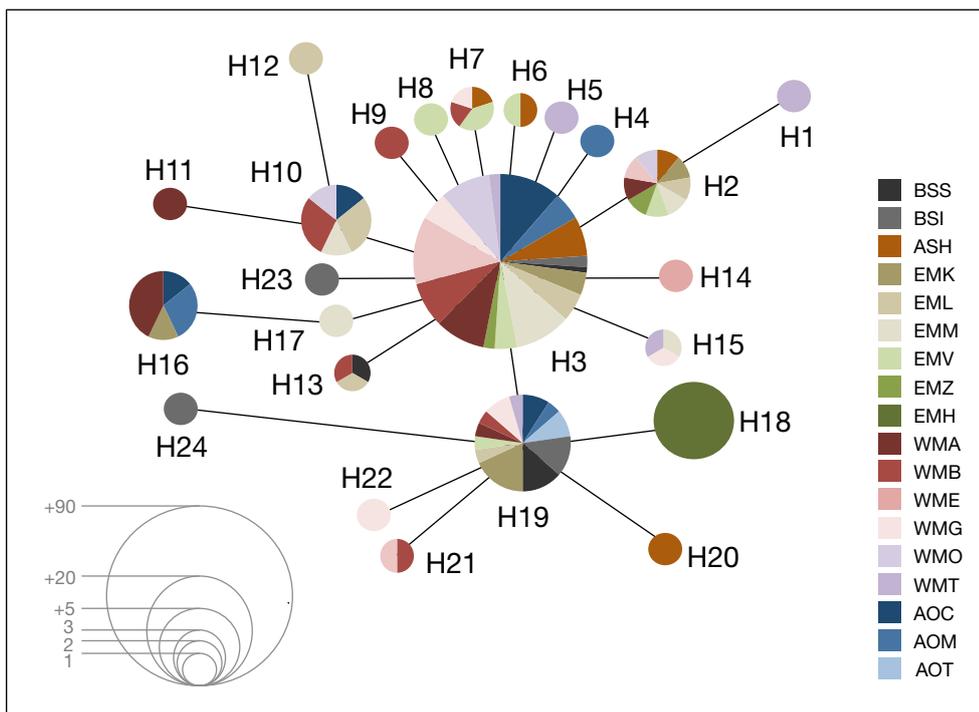


Fig. A1. *Mugil cephalus*. Network of cytochrome *b* haplotypes collected in the NE Atlantic and the Mediterranean and Black Seas, and considered in this study. The lengths of the connecting lines relate to the number of mutations between haplotypes. Each circle represents a haplotype, with the diameter of the circle proportional to the number of sequences of that haplotype. Populations are labelled as in Table 1