Influence of hydrodynamic connectivity on the genetic structure and gene flow of the common pandora Pagellus erythrinus

Anna Rita Rossi, Paolo Colangelo, L. Berline, Elisa Angiulli, Giandomenico Ardizzone, Chiheb Fassatoui, Luciana Sola

Influence of hydrodynamic connectivity on the genetic structure and gene flow of the common pandora *Pagellus erythrinus*

Anna Rita Rossi[1], Paolo Colangelo[1, 5,*], Léo Berline[2], Elisa Angiulli[1], Giandomenico Ardizzone[3], Chiheb Fassatoui[4], Luciana Sola[1]

[1] Department of Biology and Biotechnology "C. Darwin", Sapienza University of Rome, via Borelli 50, 00161 Rome, Italy

[2] Aix-Marseille Univ, Univ Toulon, CNRS, IRD, MIO UM 110, Mediterranean Institute of Oceanography, Marseille, France

[3] Department of Environmental Biology, Sapienza University of Rome, viale dell'Università 32, 00185 Rome, Italy

[4] Ecosystems and Aquatic Resources Research Unit (UR13AGRO1), National Agronomy Institute of Tunisia (INAT), University of Carthage, Charles Nicolle Avenue 43, 1082 Tunis, Tunisia

[5] National Research Council, CNR-IRET, Via Salaria km 29.300, 00015 Monterotondo (Rome), Italy.

*** Corresponding author:**

Paolo Colangelo

National Research Council, Via Salaria km 29.300, 00015 Monterotondo (Rome), Italy.

c/o

Department of Biology and Biotechnology "C. Darwin", Sapienza University of Rome, via Borelli 50, 00161 Rome, Italy

e-mail: paolo.colangelo@uniroma1.it

Telephone: +390649918122

Fax: +39 06 4457516

Running title: Genetic structure of common pandora

ABSTRACT

Many marine organisms have complex genetic patterns that cannot be easily resolved by data analysis on spatial distribution of variability usually applied in population genetic studies. We propose an analytical framework to evaluate the role of dispersal during early life stages that considers the actual hydrodynamic connectivity in the Mediterranean Sea, as a factor shaping the population structure of demersal fishes. To this purpose, and to test different scenarios of gene flow, genotypes of individuals of *Pagellus erythrinus* sampled at 12 sites in the central Mediterranean Sea were analyzed at ten microsatellite loci. The results show the lack of genetic structure in western Mediterranean basin and a pattern of gene flow that deviates from an isolation by distance model. The observed gene flow estimates appear to be scale-dependent. At local scale, it is likely the result of multifactorial components whereas at a larger scale it is mainly driven by the sea currents, directly influencing dispersal of larvae between sites not reachable by adult movements. Our results stress the importance of a quantitative analysis of potential early life stage dispersal in any study focusing on the population structure of fishes with a long larval stage.

KEYWORDS: demersal, dispersal, marine fishes, microsatellite, Sparidae, oceanic circulation

INTRODUCTION

Because of the absence of evident and pronounced physical barriers, the marine environment has long been considered an open environment with high potential species dispersal that prevents population subdivision (Hauser & Carvalho, 2008). However, many studies have provided evidence contrary to this hypothesis, i.e. the existence of pronounced population structures even on a limited spatial scale in vagile species (Knutsen et al., 2003; Ovenden at al., 2004; Palero et al., 2008; Ciannelli et al., 2010), indicating a substantial gap between expectation and observation in marine population genetics (Hedgecock & Pudovkin, 2011). Indeed, both biotic (including dispersal ability in different life stages, i.e. adults, juveniles, larvae and eggs) and abiotic (historical and oceanographic) factors components contribute to shape the genetic structure of marine species. Among the biotic factors that are related to dispersal ability, the existence of a pelagic larval phase is considered to play a key role in population connectivity and to be the primary dispersal phase in benthic, demersal and most coastal marine species (Cowen & Sponaugle, 2009). As an example planktonic larval duration (PLD), i.e. the amount of time that pelagic larvae spend in the open sea before settlement, was supposed to be a good predictor of the species' population connectivity and structure. However discordant results were obtained among taxa when this hypothesis was tested (see Bohonak, 1999; Jones et al., 2009; Weersing & Toonen, 2009; Kelly & Palumbi, 2010; Riginos et al., 2011; Modica et al., 2017; Nanninga & Manica, 2018).

These evidences, and the absence of correlation between PDL/dispersal/genetic structure in multispecies studies reveals the

77 complexity of the factors affecting population connectivity. Weersing &
78 Toonen (2009) suggested that “larval biology of individual species, not
79 PLD, is likely the primary determinant of realized dispersal”. This is
80 congruent with the idea that factors other than PLD drive population
81 connectivity (Félix-Hackradt et al., 2013) and define the distribution range
82 of fishes (Macpherson & Raventos, 2006; Robitzch et al., 2016; Faillettaz et
83 al., 2018). As an example, within the Mediterranean Sea the effect of
84 oceanographic discontinuities, like currents and fronts, on genetic
85 differentiation is not homogenous on different species, and depends on their
86 life history traits (Pascual et al., 2017). Even more, in coral reef species,
87 larval biology is crucial in connectivity; it determines both the high
88 percentages of local self-recruitment (Swearer et al., 1999; Almany et al.,
89 2007), or the high levels of genetic connectivity across large geographic
90 scales (Horne et al., 2013; Simpson et al., 2014).
91 As far as the abiotc factors are concerned, oceanographic processes (often
92 not properly considered in studies on fish population genetics) can greatly
93 influence, or even determine, the connectivity of marine populations (White
94 et al., 2010; Schunter et al., 2011; Simpson et al., 2014). Therefore,
95 quantitative analyses of hydrodynamic connectivity, i.e. the exchange of
96 particles due to coastal and oceanic circulations at different spatial and
97 temporal scales, may represent a useful framework to interpret the current
98 genetic exchange patterns of a given species (Mitarai et al., 2009; White et
99 al., 2010; Berline et al., 2014). This, in turn, can provide useful information
100 to optimize management of commercially important fishes (Berry et al.,
101 2012).

To date, a number of studies have effectively incorporated hydrodynamic connectivity in population genetic research on various marine organisms, using hydrodynamic models as proxy of passive dispersal. On this base different models of seascape connectivity have been developed with the main aims to (a) correlate historical and contemporary patterns of dispersal among populations (see Riginos & Liggins, 2013 for a review), (b) recognize common patterns related both to egg type, larval features and dispersal pathways (White et al., 2010; Crandall et al., 2012; Liggins et al., 2016; Faillettaz et al., 2018), and (c) identify the biological and physical drivers that shape genetic diversity (Simpson et al., 2014; Davies et al., 2015; Liggins et al., 2016). These studies suggested that, at ecological time scales, larval dispersal can show high variance and probably cover smaller spatial scales than reported (Crandall et al., 2012) and that, in co-distributed reef fishes, seascape features are better predictors of spatial genetic patterns than dispersal traits (Liggins et al., 2016). Despite the application of a variety of approaches (Jones et al., 2009), marine dispersal kernels remain still largely undescribed. Here we tested the role of hydrodynamic connectivity - (with different scenarios of gene flow) in determining the observed pattern of the population genetic structure of a demersal fish, the common pandora *Pagellus erythrinus* L. 1758.

The common pandora is a commercially important benthopelagic sparid fish that inhabits inshore waters of the continental shelf of the northeastern Atlantic Ocean, from Norway to Guinea-Bissau, and the Mediterranean and Black Sea (Bauchot & Hureau, 1986); it is present up to 200-300 m but most commonly at 20-100 m and occasionally at 100-200 m (Busalacchi et al., 2014). No catch in the 200-800 m stratum has been reported by the

international bottom trawl survey in the Mediterranean (MEDITS, 2001). The common pandora is a protogynous hermaphroditic fish that reproduces during late spring and summer (Valdés et al., 2004), when adults migrate from coastal waters to spawn in the open sea; trophic migration, in the opposite direction, occurs in summer-autumn (Županović & Jardas, 1989). Eggs and larvae are pelagic; the median larval stage lasts 44 days and larvae are distributed offshore (>1 mile) (Macpherson & Raventos, 2006). The species has a high commercial value and its catch has increased in the last 30 years, particularly along the North African coast of the Mediterranean Sea (FAO, 2018). It is also a candidate for marine fish aquaculture in the Mediterranean Sea, and a small-scale production already exists in Greece (Klaoudatos et al., 2004).

Previous mtDNA data (Angiulli et al., 2016) identified three sympatric mitochondrial lineages in specimens collected at 15 sites in the Mediterranean Sea and one in the Atlantic Ocean. The three lineages probably originated from isolation and divergence processes and then came into secondary contact during the Middle and Late Pleistocene glacial and interglacial cycling (Angiulli et al., 2016). Within the central Mediterranean Sea, their distribution does not show any evident geographic pattern, suggesting high connectivity even across the putative boundaries between the biogeographic sectors proposed in the investigated area (Villamor et al., 2014), and between the two Tunisian sites where previous gene enzyme electrophoresis analyses reported genetic differentiation (Fassatoui et al., 2009, 2011). The frequency distribution of mtDNA lineages across sampling sites suggested a more limited matrilinear gene flow only between the Mediterranean Sea and the Atlantic Ocean.

According to the idea that nuclear loci should usually provide greater resolution of genetic structure than mtDNA (Angers & Bernatchez, 1998), the main goal of this study was to analyze the present fine-scale population structure and connectivity of *P. erythrinus* within the Mediterranean Sea and to clarify the contribution of dispersal at different life history stages to delineate this structure. For this purpose, we analyzed the genetic variation at 10 microsatellite loci screened in specimens collected at 12 of the 16 sampling sites examined by mtDNA analysis (Angiulli et al., 2016). We focused on the Italian and Tunisian sites and estimated the hydrodynamic connectivity in the central Mediterranean Sea using a Lagrangian approach. We assumed that (a) adult common pandora occupy bottoms of the continental platform and are likely unable to move between sites separated by deep stretches of sea, implying that their possible migration routes are limited by the depth; (b) that eggs and larvae occupy superficial waters, as larvae of Sparidae in the Mediterranean Sea are mainly concentrated in the upper 10 m of the water column (Sabatés and Olivar, 1996). Thus, when connectivity exists between two sites separated by a deep stretch of sea (> 300 m in depth), we might expect that it is due to pelagic larval migration of early life history stages.

With this premise and on the base of the information derived from a Lagrangian model of hydrodynamic connectivity, we tested different gene flow scenarios in a Bayesian framework. In this way we aimed to determine whether the identified genetic structure (i) can be attributed mainly to the active dispersal capabilities of adults or (ii) is mainly shaped by early life history dispersal between pairs of sites due to the effects of hydrodynamic connectivity, or (iii) might be the result of both.

MATERIALS AND METHODS

Sampling and microsatellite genotyping

Common pandora adults (Standard Length > 13 cm) were collected from 12 sites along the Italian and Tunisian coasts in the Mediterranean Sea (Table 1, Fig 1). The samples were collected during MEDITS (Mediterranean International Trawl Survey, Bertrand et al., 2002) fishery research surveys or were obtained from commercial fishing landings.

Pectoral fin clips were removed and preserved in 96% ethanol at 4°C. Voucher specimens were deposited in the Ichthyology Collection of the Department of Biology and Biotechnology of Sapienza University of Rome (ICDBB 970-974). Genomic DNA was extracted according to the procedure of Aljanabi & Martinez (1997) and then used as a template in polymerase chain reactions (PCR) for 10 microsatellite loci expressly isolated for *Pagellus erythrinus* (MS2, MS3, MS4, MS6, Ramsak et al., 2003) or for other sparids (SL7, SL17, SL26, SL29 and SL33, Chopelet et al., 2009; PbOviD102, Piñera et al., 2006). Loci were amplified in two multiplex reactions in a total volume of 10 µl containing 1 µl of 10x buffer, 0.3 µl of $MgCl_2$ (50 mM), 0.2 µl of dNTP (2.5 mM), 0.1 µl of each primer (100 µM), 0.07 µl of BIOTAQ^{TM} DNA polymerase and 10-100 ng of DNA template. The PCRs were performed in a Thermocycler T1 (Biometra, Göttingen, Germany) with an initial denaturation of 2 min at 94°C, followed by 40 cycles of 94°C for 30 s, a locus-specific annealing temperature (for each primer as described by the above-mentioned authors) for 45 s and 72°C for

30 s and a final extension at 72°C for 5 min. PCR products were run on 6.5% denaturing polyacrylamide gel using a LI-COR DNA 4200 automated sequencer (LI-COR Biosciences, Cambridge, UK). Allele sizes were screened using GENE-IMAGE v4.05 software (Scanalytics, Fairfax, VA, USA). Approximately 10% of the samples (50 individuals) were re-scored at all 10 microsatellite loci to ensure repeatability.

Data analysis

Genetic variability and differentiation

Allele frequencies, expected and observed heterozygosity (H_{exp} and H_{obs}), average number of alleles (A), number of private alleles (N_P) and allelic richness (A_R) were estimated per locus and per sampling site using GENEPOP v3.4 (Raymond & Rousset, 1995) and FSTAT v2.9.3.2 (Goudet, 2001) software. Deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium were tested for each locus and pairs of loci respectively, and for sampling site. Statistical significance of both analyses was calculated with the Markov Chain Monte Carlo (MCMC) method using 10^4 permutations, $5x10^3$ dememorization steps and 500 batches. Sequential false discovery rate (FDR) correction for multiple tests was applied for the HWE tests and for the linkage disequilibrium significance because of the large number of tests involved (Benjamini & Hochberg, 1995).

The occurrence of putative null alleles, large allele drop-out and scoring errors was evaluated using MICROCHECKER (Van Oosterhout et al., 2004). Moreover, the frequency of null alleles was tested using the Expectation Maximization algorithm of Dempster et al., (1977), implemented in FreeNA software (Chapuis & Estoup, 2007).

The presence of loci under selection was evaluated by the coalescent approach (Beaumont & Nichols, 1996) and by both a global outlier test, implemented in LOSITAN (Antao et al., 2008), and the hierarchical method of Excoffier et al. (2009), using the software ARLEQUIN v3.5 (Excoffier & Lischer, 2010). The LOSITAN analyses were performed under the infinite allele model (IAM) and the stepwise mutation model (SMM) with $95x10^3$ permutations and 99% confidence interval, while the ARLEQUIN analyses were carried out under a hierarchical island model with $5x10^4$ coalescent simulations.

Genetic differentiation between samples was tested by F_{ST} (Weir & Cockerham, 1984) and R_{ST} (Slatkin, 1995) using ARLEQUIN v3.5 and RSTCALC v2.1 (Goodman, 1997) respectively, and the significance of pairwise values was calculated with the MCMC method using 10^4 permutations, $5x10^3$ dememorization steps and 500 batches. Sequential false discovery rate (FDR) correction for multiple tests was applied.

The simulation-based program POWSIM v4.1 (Ryman & Palm, 2006) was used to assess the statistical power of the 10 microsatellite loci to detect various levels of divergence among sampling sites. The simulations were performed using different combinations of N_e (effective population size) and t (time of divergence) leading to F_{ST} values in the range 0.005-0.010. The tested values of F_{ST} and N_e were selected according to the estimated values from data of this study (see Results section). The statistical power of the dataset was evaluated using Chi-squared and Fisher's exact tests. The estimate of power was reported as the proportion of significant outcomes ($p<0.05$) after 1000 replicates.

The correlation between least cost geographic distances and the genetic

distances, calculated as the $F_{ST}/(1- F_{ST})$, was analyzed with a Mantel test, using the software R statistical environment (R Development Core Team, 2011). Least cost distances among pairs of sites were calculated using the function shortest Path implemented in the "*gdistance"* package (van Etten, 2017). In order to obtain a least cost distance, we allowed connection only through areas covered by sea, preventing connection across a land surface.

Cluster analysis

Bayesian cluster analysis implemented in STRUCTURE 2.3 (Pritchard et al., 2000) was used to estimate the number of genetic clusters (*K*) within the dataset. Mean and variance of log likelihoods of the number of clusters for *K*=1 to *K*=12 (the number of sampling sites) were inferred from multilocus genotypes using the MCMC method. Twenty replicates for each set of values of *K* with $1.5x10^6$ iterations (burn-in period of $1x10^5$ iterations) were performed under the admixture ancestry model and the assumption of correlated allele frequencies among samples (Falush et al., 2003).

In addition to STRUCTURE, discriminant analysis of principal components (DAPC; Jombart et al., 2010) was used to detect the best number of genetic clusters using the function find.cluster. The correct number of clusters (K) was determined by testing K values from 1 to 12. Different clustering solutions were then compared using Bayesian Information Criterion (BIC). Subsequently DAPC was performed (the first 100 PC scores, which account for 80% of the total cumulative variance, were retained) to evaluate the percentage of posterior correct assignments. Both K-means and DAPC were implemented in the R package adegenet v1.3-1 (Jombart, 2008; Jombart & Ahmed , 2011).

Hydrodynamic model

A Lagrangian approach was used to quantify hydrodynamic connectivity. We define Oceanic distances (OD) as the mean connection time (MCT) taken by particles to connect two sites. The dispersal of a large ensemble of passive particles by the oceanic circulation was modeled using currents from a multiannual realistic ocean model simulation covering the whole Mediterranean basin. Particles were released in the surface layer (0-100 m) where eggs and larvae are generally found (Klimogianni et al., 2004 and references therein). For each pair of square sub-regions of 50 km per side, the time needed to go from one sub-region to the other (by the particles) was computed. The average time to connect site pairs found in a specific sub-region provides a measure of the oceanic distance. The resulting MCT matrix is asymmetric since the time required by a particle to go from site “a” to site “b” is not equal to the time to go from “b” to “a”. Details of the method and how to calculate MCT can be found in Berline et al. (2014). Here we only consider MCT below 200 days, a time window that includes the embryonic and larval phases as well as the early juvenile stage. The pairwise hydrodynamic connectivity was then used to propose different dispersal scenarios of gene flow (see below).

Gene flow and hydrodynamic connectivity

Gene flow within the Mediterranean Sea was estimated using a coalescent based approach implemented in MIGRATE-N v3.6 (Beerli & Felsenstein, 2001; Beerli, 2006), and alternative migration hypotheses (Beerli & Palczewski, 2010) were tested.

310 Complete panmixia (Fig. 2, Model 0: all the sampled specimens belong to a
311 unique large panmictic unit) was tested versus seven alternative migration
312 models (Fig. 2, Models 1-7). The evaluated scenarios allowed us to test
313 alternative population structures and migration dynamics (adults and early
314 life history dispersal) along different routes.
315 In the first migration model (Fig. 2: Model 1), sampled populations were
316 pooled according to their geographic sea origin: Ligurian Sea, Tyrrhenian
317 Sea, Sea of Sardinia, Sicilian Channel, Ionian Sea, Adriatic Sea. Migration
318 rates were allowed to be bi-directional between pairs of adjacent seas. All
319 the successive models considered the sampled geographic populations as
320 distinct genetic units, whereas migration rates among them were evaluated
321 according to different schemes. Indeed, in the second migration model (Fig.
322 2: Model 2) the gene flow was considered a bi-directional stepping stone
323 process related to active dispersal of *P. erythrinus*. Thus, the gene flow was
324 estimated only between sites whose distances and path could be covered by
325 adult migration routes along coastlines or across sea stretches less than 300
326 m deep.
327 This model allowed us to identify three groups of sites that can exchange
328 migrants (Fig. 2: Model 2). In the subsequent model (Fig. 2: Model 3) we
329 extended the previous model by considering additional early life history
330 dispersal routes mediated by sea currents. On the basis of oceanic distances
331 (see hydrodynamic connectivity) we identified two routes that allowed us to
332 connect the three groups of Model 2, thus allowing a potential continuous
333 exchange of genes across the sampled sites. In more detail we identified the
334 two unidirectional routes (EN->SN and AL->BZ) that allowed an early life
335 history dispersal between the three groups of units with the shortest period

(112 and 91 days respectively). These migration routes, identified on the basis of the oceanic distances, are congruent with the oceanic circulation known for the area (I.I.M., 1982; Millot & Taupier-Letage, 2005; Berline et al., 2014). In Models 4, 5 and 6 (Fig. 2) only a main effect of early life history gene flow, mediated by oceanic circulation, was considered. In these three models, different thresholds were used to calculate which localities are connected by sea currents respectively up to 100, 150 and 200 days. In these models, migration routes could be uni-directional because in some cases the adopted time thresholds allowed a particle to go from one locality to another but not vice versa. Finally, a further model (Fig. 2: Model 7) was tested considering all the sampled populations as independent genetic units; bi-directional migration was possible among all pairs of sampling sites.

Initially, a number of preliminary runs, using all sites and assuming a full migration matrix, was performed to identify the best parameters for the search. Identification of the optimal MCMC parameters was done until consecutive runs produced overlapping results and the estimated values for all the parameters showed a normal distribution (i.e. no double peaks and narrow intervals). The search parameters were then adopted for all the models tested. Initial migration (M) parameters were estimated using F_{ST}. Default MIGRATE-N priors were used for all the models. For each model two independent multiple MCMCs were run for 1000000 steps, sampling every 100 steps. For each MCMC chain the first 10000 trees were discarded and a static heating scheme was adopted (4 chains with temperatures: 1000000, 3, 1.20, 1). For each analysis the effective sample size (ESS) was evaluated and it was found to be much higher than 1000 for all the parameters in all the models tested, with the only exception for the most

complex model (n=144 parameters) showing seven parameters with an ESS slightly lower than 1000. The marginal likelihood of each model, as obtained by Bezier approximation, was used to calculate the model probability, estimated by dividing each marginal likelihood by the sum of the marginal likelihoods of all used models.

RESULTS

In total, 470 common pandora individuals from 12 sampling sites (Table 1) were screened for 10 microsatellite loci. The 50 DNA samples that were amplified and scored twice produced identical results in each trial.

Genetic variability

Genetic variability parameters vary considerably among the 10 loci (Online resource 1): MS6 is the locus with the lowest number of alleles (14 distinct alleles, A= 6.4), while MS2 is the most variable (83 distinct alleles, A=35.6). Private alleles (65) constitute 14% of the total alleles (464), ranging from 7.8% at locus SL17 to 35.3% at locus SL33. The number of alleles per sampling site is medium/high and quite homogeneously distributed across samples (A ranges from 22 to 25.4, Table 1). Private allele distribution across sites ranges from 0.42% at sampling site LM to 3.75% at AN. Levels of mean H_{obs} are similar among all samples, ranging from 0.748 to 0.802, but lower than the H_{exp} (Table 1).

Significant departures from HWE were observed in 50 of 120 locus/site combinations after FDR correction (Online resource 2). MS2 and SL33 are the most affected loci (Online resource 2). All deviations are characterized by heterozygote deficit, and according to MICROCHECKER half of them

388 are associated with the presence of putative null alleles, whose frequencies
389 in each locus are not uniformly distributed across different sampled
390 populations (Online resource 2). According to this, the deviation from HWE
391 due to heterozygote deficit observed in our dataset cannot be simply
392 explained by a widespread presence of null alleles but rather as the result of
393 a combination of factors including other biological processes (e.g.
394 inbreeding and/or the Wahlund effect). Indeed, heterozygote deficit and
395 positive F_{IS} are also observed when data are calculated for each population
396 (data not shown). The different components that act on heterozygote deficit
397 cannot be easily discerned by a multilocus data analysis (Dharmarajan et al.,
398 2013). Carlsson (2008) demonstrated that the occurrence of null alleles
399 could only slightly affect the assignment test and F_{ST} estimation. Moreover,
400 a conservative approach of discarding loci deviating from HWE
401 expectations could eliminate informative markers and thus affect the
402 inferences on biological phenomena derived from the data (Dharmarajan et
403 al., 2013). According to these considerations and to the non-uniform
404 frequency distribution of null alleles across different populations, we
405 retained one of the two most affected loci, MS2, in the analysis; the other
406 one, SL33, was excluded from the genetic structure analysis as it is also
407 under selection (see below).
408 No linkage disequilibrium was detected for each pair of loci (585 pairwise
409 comparisons) by the FDR method.
410
411 **Population structure**
412 The POWSIM analysis showed that the combination of microsatellite loci
413 and sample sizes rendered a statistical power sufficient to detect a low level

of differentiation. In fact, about 95% (χ^2) and more than 90% (Fisher's) of the tests where the N_e/t combination led to F_{ST} = 0.0010 and 100% (both χ^2 and Fisher's) of the tests where N_e/t led to a F_{ST} value of ≥0.0025 were statistically significant (Online resource 3). The F_{ST} values were significant for all loci with the exception of PbOviD102. The global outlier test revealed directional selection at one locus, SL33, using both LOSITAN ($p < 0.01$) under IAM and SMM models and ARLEQUIN ($p<0.001$). These results suggested that SL33 is under positive selection, and for this reason this locus was eliminated from subsequent analyses.

Multilocus F_{ST} among all samples = 0.00395 ($p<0.001$). Multilocus pairwise estimates of F_{ST} and R_{ST} are reported in Table 2. Only two pairwise comparisons were found to be significant after FDR correction, AN-BA and GA-BA, for F_{ST} and R_{ST}, respectively.

The Mantel test applied to all sampling sites revealed the absence of correlation between genetic distances, expressed as $F_{ST}/(1- F_{ST})$, and geographic distances ($z=231132.8$, $p=0.175$).

Cluster analysis

Cluster analyses by means of STRUCTURE identified, among the 12 tested K, the greatest posterior probability value at K = 1, i.e. a single genetic cluster.

In contrast, according to BIC the best number of clusters inferred by the k-means analyses within the Mediterranean Sea sites is four, and DAPC provided >0.96% of posterior correct group assignments. The four genetic clusters can be easily distinguished by the first two discriminant axes, and

each includes a different number of individuals (n=112, 110, 93 and 121 respectively for cluster 1, 2, 3 and 4). All four clusters are not related to specific geographic areas and were found at all the sampling sites (Fig. 3, Online resource 4). Frequency of cluster across the twelve sites is not significantly different according to chi-square test (p=0.08)

Gene flow and hydrodynamic connectivity

Lagrangian model returns that only 59 possible connections (versus a total of 144 hypothetical ones) can occur within the threshold of 200 days.

The MCTs range between a maximum of 199 days and a minimum of 53 days, with an average of 134 days. Shortest MCT were found between Sardinian (SN-AL) and Tyrrhenian sites (GA-FM, FM-EN) as well as between Jonian and southern Tyrrhenian sites (JO-VM).

As expected, the largest MCTs were found between sites located in different sea basins. For every pair of sites, a strong asymmetry exists between the two directions of transport (Online resource 5).

Among the different migration models, according to the marginal likelihoods, the highest support (relative probability, P=1) was obtained for Model 3 (Online resource 6). All the estimated migration rates (M) show narrow 95% confidence intervals with density distributions following a normal distribution. Migration rates between pairs of sites are scattered across a wide range of values and are summarized in Fig. 4a and b. Estimated migration rates are almost symmetrical between all pairs of sites, with the exception of FM<->EN and LM<->ZA (Fig 4a). The estimated migration rates between pairs of sites (EN->SN and AL->BZ) that cannot be reached by adult *P. erythrinus* are comparable or even higher than those

observed between geographically adjacent pairs of sites that can be reached by adult dispersal. Very low migration rates were observed among a few sites and particularly characterized Adriatic (AN, BA) and Ionian (JO) sites.

DISCUSSION

Population structure

Our results highlight a lack of spatial population structure in the common pandora, similar to what observed in other Mediterranean demersal species (Schunter et al., 2011; Franchini et al., 2012). Partial inconsistencies among results obtained by the different analysis/programs employed could derive from the different assumptions on which the various methods rely, such as a priori grouping of individuals or their assumed pure/mixed ancestry. As an example, simulated data show that below F_{ST} = 0.1, STRUCTURE is unable to correctly identify the number of subpopulations, providing false certainty regarding the K (Latch et al., 2006).

STRUCTURE suggests potential panmixia in the common pandora whereas this scenario, that consider random genetic exchange among individuals and population with the same probability, is rejected by the MIGRATE-N model comparison. The k-means clustering and DAPC methods identify four different genetic clusters (Online resource 4), which however are spread across all sites and thus fail to identify a spatial structuring of genetic diversity. A previous mtDNA sequence analysis of the same samples within the Mediterranean Sea depicted a similar pattern characterized by the absence of a spatial genetic structure (Angiulli et al., 2016). Due to their different mutation rates and uni- or bi-parental inheritance, mitochondrial DNA and microsatellites usually provide different sensitivities at different

temporal scales. Mitochondrial DNA is more suitable to record signatures of historical processes (Avise, 2000), related to the last few million years. Microsatellites perform better in unraveling contemporary and recent processes, i.e. occurring in the last 10 thousand years (Hewitt, 2004). The advance of combining of the two kind of markers has been proved in marine fishes (Limborg et al., 2012), although sometimes they can yield different results, i.e. identify, different numbers of subpopulations (Durand et al., 2013; Lemos et al., 2005; Sala-Bozano et al., 2009; Silva et al., 2014). In the common pandora mtDNA and microsatellites provide similar pattern of genetic diversity with differentiated genetic clusters, occurring in sympatry. This pattern could be explained by group differentiation due to vicariance events, with the groups subsequently entering into secondary contact once the disappearance of putative barriers permitted complete connectivity in the Mediterranean basin.

Impact of hydrodynamic connectivity on gene flow pattern

Knowledge of the influence of hydrodynamic patterns on population connectivity of many Mediterranean species is still scarce, whereas it could be particularly important for the study of the genetic structure of demersal fishes. In these species, adults are usually associated with benthonic environments and, although capable of long-distance dispersion, are limited in crossing deep sea stretches, that act as natural barriers. As a consequence, long-range dispersal cannot be attributed to adult movements alone, but rather also to what occurs during the earliest ontogenetic stages, when eggs and larvae are dispersed by sea currents.

We employed an analytical method that allowed us to gain insight into the

gene flow and connectivity among common pandora populations. We tested different scenarios in which we assumed a different role of passive and active dispersal in determining the observed population structure in the western Mediterranean Sea. Our model selection approach suggests a scenario in which the gene flow over long distances is due to dispersal of eggs and larvae by oceanic currents, whereas at a small scale can effectively be affected by adult's movement. This scenario agrees with what we know about the biology of this species. Thus, it is likely that gene flow between neighboring sites (on the same coast or connected by a body of water less than 300 m deep) can occur by active movement of adults, even though a contribution by sea currents to the gene flow among neighboring sites cannot be completely ruled out. We tried to determine the potential contribution of adult and larval migration over short distances by assessing the influence of both oceanic and geographic distances on the estimated mean gene flow in the best-supported migration scenario of Model 3. Therefore, our results suggest that over long distances hydrodynamic connectivity and gene flow in the common pandora are affected by dispersal of eggs and larvae, and thus directly by oceanic circulations, gyres and barriers present within the Mediterranean Sea (Limborg et al., 2012). On the other hand, over short distances, it is likely that a mixed contribution by sea currents, adult movements and/or other biophysical phenomena determines the exchange among populations. As an example, recently simulated scenarios on other sparids showed that larval behavior could actually affect connectivity: swimming abilities of just few cm s^{-1}, at the end of the pelagic phase, can impact on the rate of recruitment and change connectivity patterns at regional scale (Faillettaz et al., 2018).

All these phenomena act at different scales determining a lack of spatial genetic structure. Further studies are necessary for a generalization of this connectivity pattern to other demersal fishes, but it is important to stress that a similar pattern of genetic structure was reported in the same geographic area for another sparid, the coastal *S. aurata* (Franchini et al., 2012), which has a comparable long-lasting larval stage (about 50 days).

Finally, it is important to frame the excess of homozygotes observed in the sampled populations within the gene flow scenario supported by our analyses. Under a scenario of weak population structure and extensive gene flow (as confirmed by F_{ST} and MIGRATE-N analyses), a high proportion of heterozygoses at most of the sampled loci is expected. On the other side, a recent study on the evolutionary consequences of alternative sex change pathways by Benvenuto et al. (2017) demonstrated a reduction of the effective population in protogynous fishes and also reported a reduction of heterozygotes in *P. erythrinus*. Further data indicate that biophysical phenomena can cause a long-term aggregation of larval fish siblings during dispersal, causing a deviation from Hardy-Weinberg proportions in samples containing a high proportion of related individuals (Ottmann et al., 2016). Thus, in *P. erythrinus* the extensive gene flow could be compatible with an excess of homozygotes, considering an effective low number of breeders or other phenomena that allow the passive dispersal of large numbers of siblings.

Conclusion

A quantitative method that evaluates the effect of hydrodynamic

connectivity on gene flow can be a powerful tool for an understanding of the population genetics and dynamics of demersal fishes. Our seascape genetic analyses in the common pandora allowed us to elucidate population dynamics and dispersal patterns, confirming the usefulness of genetic approaches in testing predictions about transport mechanisms in the sea (Selkoe et al., 2016). Our results suggest that gene flow across the central Mediterranean Sea in the common pandora, and likely in other demersal fishes, is extremely complex and driven by a combination of factors acting at different scales, including movements of both adults and eggs and larval migration. The complex pattern detected in the central Mediterranean Sea suggests that in fish population genetic studies, hydrodynamic connectivity estimates can help not only to understand the distribution of genetic diversity but also to study dispersal patterns of demersal fishes. Moreover, clarifying the interactions between hydrodynamic connectivity and population genetics could provide useful information for fisheries management and for the better design of marine protected areas (Andrello et al., 2015)

ACKNOWLEDGEMENTS

We thank the MEDITS coordinators and all of our colleagues for their invaluable help with sample collection, specifically Maria Teresa Spedicato and Giuseppe Lembo, COISPA (Bari); Angelo Cau, Dept. of Life and Environmental Sciences, Cagliari University; Corrado Piccinetti, Dept. of Biological, Geological and Environmental Sciences, Bologna University; Angelo Tursi, Dept. of Biology, Bari University; Fabio Fiorentino, Coastal Marine Environment Institute, Mazara del Vallo and Mohamed Salah

596 Romdhane, INAT/Université de Carthage, Tunisia. Financial support for
597 this work was provided by Sapienza University of Rome, Sapienza Projects
598 2014, grant C26A143AL3

599

600**REFERENCES**

601

602 Aljanabi, S. M. & I. Martinez,1997. Universal and rapid salt extraction of
603 high quality genomic DNA for PCR-based techniques. Nucleic Acids
604 Research 25:4692–4693.

605 Almany, G. R., M. L. Berumen, S. R. Thorrold, S. Planes & G. P. Jones,
606 2007. Local replenishment of coral reef fish populations in a marine
607 reserve. Science 316: 742-744.

608 Andrello, M., M. N. Jacobi, S. Manel, W. Thuiller & D. Mouillot, 2015.
609 Extending networks of protected areas to optimize connectivity and
610 population growth rate. Ecography 38:273–282.

611 Angers, B. & L. Bernatchez, 1998. Combined use of SMM and non-SMM
612 methods to infer fine structure and evolutionary history of closely related
613 brook charr (*Salvelinus fontinalis*, Salmonidae) populations from
614 microsatellites. Molecular Biology and Evolution 15:143–159

615 Angiulli, E., L. Sola, G. Ardizzone, C. Fassatoui & A. R. Rossi, 2016.
616 Phylogeography of the common pandora *Pagellus erythrinus* in the
617 central Mediterranean Sea: sympatric mitochondrial lineages and genetic
618 homogeneity. Marine Biology Research 12:4-15.

619 Antao, T., A. Lopes, R. J. Lopes, A. Beja-Pereira & G. Luikart,2008.
620 LOSITAN: a workbench to detect molecular adaptation based on a F*st*-
621 outlier method. BMC Bioinformatics 9:323.

622 Avise, J. C., 2000. Phylogeography. The history and formation of species.
623 Harvard University Press: Cambridge, MA.

624 Bauchot, M. L. & J. C. Hureau,1986. Sparidae. In: Whitehead PJP, Bauchot
625 ML, Hureau JC, Nielsen J, Tortonese E (eds), Fishes of the north-eastern

626 Atlantic and the Mediterranean, UNESCO, Paris. Vol 2: 883-907.
627 Beaumont, M. A. & R. A. Nichols,1996. Evaluating loci for use in the
628 genetic analysis of population structure. Proceedings of the Royal
629 Society of London B: Biological Sciences 263:1619-1626.
630 Beerli, P., 2006. Comparison of Bayesian and maximum likelihood
631 inference of population genetic parameters. Bioinformatics 22:341–345.
632 Beerli, P. & J. Felsenstein, 2001. Maximum likelihood estimation of a
633 migration matrix and effective population sizes in n subpopulations by
634 using a coalescent approach. Proceeding of the National Academy of
635 Science USA 98:4563–4568.
636 Beerli, P. & M. Palczewski, 2010. Unified framework to evaluate panmixia
637 and migration direction among multiple sampling locations. Genetics
638 185:313–326
639 Benjamini, Y. & Y. Hochberg,1995. Controlling the false discovery rate: a
640 practical and powerful approach to multiple testing. Journal of the Royal
641 Statistical Society: Statistical methodology 57: 289–300.
642 Benvenuto, C., I. Coscia, J. Chopelet, M. Sala-Bozano & S. Mariani, 2017.
643 Ecological and evolutionary consequences of alternative sex change
644 pathways in fish. Scientific Reports| 7: *9084*
645 Berline, L., A. M. Rammou, A. Doglioli, A. Molcard& A. Petrenko, 2014. A
646 Connectivity-Based Eco-Regionalization Method of the Mediterranean
647 Sea. PLoSONE 9:e111978.
648 Berry, O., P. England, D. Fairclough, G. Jackson& J. Greenwood, 2012.
649 Microsatellite DNA analysis and hydrodynamic modelling reveal the
650 extent of larval transport and gene flow between management zones in an

651 exploited marine fish (*Glaucosoma hebraicum*). Fisheries Oceanography
652 21: 243-254.
653 Bertrand, J., L. Gil de Sola, C. Papaconstantinou, C. Relini & A.
654 Souplet,2002. The general specifications of the MEDITS surveys.
655 Scientia Marina 66:9–17.
656 Bohonak, A. J., 1999. Dispersal, gene flow, and population structure. The
657 Quarterly Review of Biology 74: 21-45.
658 Busalacchi, B., T. Bottari,D. Giordano, A. Profeta & P. Rinelli, 2014.
659 Distribution and biological features of the common pandora, *Pagellus*
660 *erythrinus* (Linnaeus, 1758), in the southern Tyrrhenian Sea (Central
661 Mediterranean). Marine Research 68:491–501
662 Carlsson, J., 2008. Effects of microsatellite null alleles on assignment
663 testing. Journal of Heredity 99: 616–623.
664 Chapuis, M. P. & A. Estoup,2007. Microsatellite null alleles and estimation
665 of population differentiation. Molecular Biology and Evolution 24:621-
666 631.
667 Chopelet, J., S. Helyar, B. Mann, & S. Mariani, 2009. Novel polymorphic
668 microsatellite loci for the ptotogynous hermaphrodite slinger sea bream
669 (*Chrysoblephus puniceus*, Sparidae). Molecular Ecology Research
670 9:1223–1226.
671 Lorenzo Ciannelli, L., H. Knutsen, E. M. Olsen, S. H. Espeland, L. Asplin,
672 A. Jelmert, J. A. Knutsen and N. C. Stenseth, 2010. Small-scale genetic
673 structure in a marine population in relation to water circulation and egg
674 characteristics. Ecology 91: 2918–2930.
675 Cowen, R. & S. Sponaugle, 2009. Larval dispersal and marine population
676 connectivity. Annual Review of Marine Science 1:443–466.

677 Crandall, E. D., E. A. Treml & P. H. Barber, 2012. Coalescent and
678 biophysical models of stepping-stone gene flow in neritid snails.
679 Molecular Ecology 21: 5579–5598
680 Davies, S. W., E. A. Treml, C. D. Kenkel & M. V. Matz, 2015. Exploring the
681 role of Micronesian islands in the maintenance of coral genetic diversity
682 in the Pacific Ocean. Molecular Ecology 24: 70–82.
683 Dempster, A. P., N. M. Laird, & D. B. Rubin,1977. Maximum likelihood
684 from incomplete data via the EM algorithm. Journal of the Royal
685 Statistical Society: Statistical methodology 39:1-38.
686 Dharmarajan, G., W. S. Beatty & E. R. Olin Jr, 2013. Heterozygote
687 deficiencies caused by a Wahlund effect: Dispelling unfounded
688 expectations. The Journal of Wildlife Management 77: 226–234.
689 Durand, J. D., H. Blel, K. N. Shen, E. T. Koutrakis &B. Guinand, 2013.
690 Population genetic structure of *Mugil cephalus* in the Mediterranean and
691 Black Seas: a single mitochondrial clade and many nuclear barriers.
692 Marine Ecology Progress Series 474:243–261.
693 Excoffier, L. & H. E. L Lischer,2010. Arlequin suite ver 3.5: a new series of
694 programs to perform population genetics analyses under Linux and
695 Windows. Molecular Ecology Resources 10:564-567.
696 Excoffier, L., T. Hofer & M. Foll,2009. Detecting loci under selection in a
697 hierarchically structured population. Heredity 103:285–298.
698 Faillettaz, R., C. B. Paris & J. O. Irisson, 2018. Larval fish swimming
699 behavior alters dispersal patterns from marine protected areas in the
700 north-western Mediterranean Sea. Frontiers in Marine Science 5:97.
701 Falush, D., M. Stephens & J. K. Pritchard, 2003. Inference of population
702 structure using multilocus genotype data: linked loci and correlated allele

703 frequencies. Genetics 164:1567–1587.
704 FAO, 2018. FishStat plus. Capture production 1950-2012, Aquaculture
705 production 1950-2016. Available from
706 http://www.fao.org/fishery/statistics/software/fishstatj/en Accessed
707 March 2018.
708 Fassatoui, C., E. Mdelgi & M. S. Romdhane, 2009. A preliminary
709 investigation of allozyme genetic variation and population structure in
710 common pandora (*Pagellus erythrinus*, Sparidae) from Tunisian and
711 Libyan coasts. Ichthyological Research 56:301-307.
712 Fassatoui, C., E. Mdelgi& M. S. Romdhane, 2011. Short-term temporal
713 genetic investigation of two populations of *Pagellus erythrinus*
714 (Linnaeus, 1758, Sparidae) from Western and Eastern Mediterranean
715 basins off the Tunisian coast. Marine Biology Research 7:147–158.
716 Faurby, S. P. H. Barber,2012. Theoretical limits to the correlation between
717 pelagic larval duration and population genetic structure. Molecular
718 Ecology 21: 3419–3432.
719 Félix-Hackradt, F. C., C. W. Hackradt, Á. Pérez-Ruzafa & J. A. García-
720 Charton, 2013. Discordant patterns of genetic connectivity between two
721 sympatric species, *Mullus barbatus* (Linnaeus, 1758) and *Mullus*
722 *surmuletus* (Linnaeus, 1758), in south-western Mediterranean Sea.
723 Marine Environmental Research 92:23-34.
724 Franchini, P., L. Sola, D. Crosetti, V. Milana& A. R. Rossi, 2012. Low
725 levels of population genetic structure in the gilthead seabream, *Sparus*
726 *aurata*, along the Italian coasts. ICES Journal of Marine Science 69:41-
727 50.
728 Galarza, J. A., J. Carreras-Carbonell, E. Macpherson, M. Pascual, S.

729 Roques, G. F. Turner & C. Rico,2009. The influence of oceanographic
730 fronts and early-life-history traits on connectivity among littoral fish
731 species. Proceedings of the National Academy of Science USA
732 106:1473–1478. 6
733 Goodman, S. J., 1997. RSTCALC: A collection of computer programs for
734 calculating unbiased estimates of genetic differentiation and determining
735 their significance for microsatellite data. Molecular Ecology 6:881-885.
736 Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and
737 fixation indices (version 2.9.3). Institute of Ecology, University of
738 Lausanne, Lausanne, Switzerland
739 Hauser, L, & J. R. Carvalho, 2008. Paradigm shifts in marine fisheries
740 genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries
741 9:333–362.
742 Hedgecock, D. & A. Pudovkin, 2011. Sweepstakes reproductive success in
743 highly fecund marine fish and shellfish: a review and commentary.
744 Bulletin of Marine Science. 87: 971–1002.
745 Hewitt, G. M., 2004. The structure of biodiversity – insights from molecular
746 phylogeography. Frontiers in Zoology 1:4.
747 Horne, J. B., L. van Herwerden, S. Abellana & J. L. McIlwain, 2013.
748 Observations of migrant exchange and mixing in a coral reef fish
749 metapopulation link scales of marine population connectivity. Journal of
750 Heredity 104: 532–546.
751 I.I.M. Istituto Idrografico della Marina 1982 Atlante delle correnti
752 superficiali dei mari italiani 3068.
753 Jombart, T., 2008. Adegenet: a R package for the multivariate analysis of
754 genetic markers. Bioinformatics 24:1403−1405.

755 Jombart, T. & I. Ahmed, 2011. Adegenet 1.3-1: new tools for the analysis of
756 genome-wide SNP data. Bioinformatics 27:3070-1.
757 Jombart, T., S. Devillard & F. Balloux, 2010. Discriminant analysis of
758 principal components: a new method for the analysis of genetically
759 structured populations. BMC Genetics 11:94.
760 Jones, G. P., G. R. Almany, G. R. Russ, P. F. Sale, R. S. Steneck, M. J. H. van
761 Oppen & B. L. Williset, 2009. Larval retention and connectivity among
762 populations of corals and reef fishes: history, advances and challenges.
763 Coral Reefs 28: 307–325.
764 Kelly, R. P. & S. R. Palumbi, 2010. Genetic structure among 50 species of
765 the northeastern Pacific rocky intertidal community. PLoS One. 2010 Jan
766 7;5:e8594.
767 Klaoudatos, S. D., G. Iakovopoulos & D. S. Klaoudatos, 2004. *Pagellus*
768 *erythrinus* (common pandora): a promising candidate species for
769 enlarging the diversity of aquaculture production. Aquaculture
770 International 12:299–320.
771 Klimogianni, A., G. Koumoundouros, P. Kaspiris & M. Kentouri, 2004.
772 Effect of temperature on the egg and yolk-sac larval development of
773 common pandora, *Pagellus erythrinus*. Marine Biology 145: 1015–1022.
774 Knutsen, H., P. E. Jorde, C. Andre& N. C. Stenseth, 2003. Fine-scaled
775 geographic population structuring in a highly mobile marine species: the
776 Atlantic cod. Molecular Ecology 12:385–39.
777 Latch, E. K., G. Dharmarajan, J. C. Glaubitz& O. E. Rhodes, 2006. Relative
778 performance of Bayesian clustering software for inferring population
779 substructure and individual assignment at low levels of population
780 differentiation. Conservation Genetics 7:295–302.

781 Lemos, A., A. I. Freitas, A. T. Fernandes, R. Goncalves, J. Jesus & C.
782 Andrade, 2006. Microsatellite variability in natural populations of the
783 blackspot seabream *Pagellus bogaraveo* (Brünnich, 1768): a database to
784 access parentage assignment in aquaculture. Aquaculture Research
785 37:1028–1033.
786 Liggins, L., E. A. Trem, H.P. Possingham & C. Riginos, 2016. Seascape
787 features, rather than dispersal traits, predict spatial genetic patterns in co-
788 distributed reef fishes. Journal of Biogeography 43: 256–267.
789 Limborg, M. T., R. Hanel, P. V. Debes, A. K. Ring, C. André, C. S.
790 Tsigenopoulos & D. Bekkevold, 2012. Imprints from genetic drift and
791 mutation imply relative divergence times across marine transition zones
792 in a pan-European small pelagic fish (*Sprattus sprattus*). Heredity
793 109:96-107.
794 Macpherson, E. & N. Raventos, 2006. Relationship between pelagic larval
795 duration and geographic distribution of Mediterranean littoral fishes.
796 Marine Ecology Progress Series 327:257–265.
797 MEDITS Biological Report 2001 available at
798 http://www.sibm.it/SITO%20MEDITS/principaleprogramme.htm).
799 Millot, C. & I. Taupier-Letage, 2005. Circulation in the Mediterranean Sea.
800 In: Saliot A (ed.) The Mediterranean Sea Vol 5k of Handbook of
801 Environmental Chemistry. Springer: Berlin. p 29–66.
802 Mitarai, S., D. A. Siegel, J. R. Watson, C. Dong, C. & J. C. McWilliams,
803 2009. Quantifying connectivity in the coastal ocean with application to
804 the Southern California Bight. Journal of Geophysical Research 114:
805 C10026.
806 Modica, M.V., V. Russini, G. Fassio & M. Oliverio, 2017. Do larval types

affect genetic connectivity at sea? Testing hypothesis in two sibling marine gastropods with contrasting larval development. Marine Environmental Research 127 92e101

Nanninga, G. B. & A. Manica, 2018. Larval swimming capacities affect genetic differentiation and range size in demersal marine fishes. Marine Ecology Progress Series 589: 1–12.

Ovenden, J. R., J. Salini, S. O'Connor & R. Street, 2004. Pronounced genetic population structure in a potentially vagile fish species (*Pristipomoides multidens*, Teleostei;Perciformes; Lutjanidae) from the East Indies triangle. Molecular Ecology 13**:** 1991–1999.

Palero, F., P. Abello, E. Macpherson, M. Gristina & M. Pascual, 2008. Phylogeography of the European spiny lobster (*Palinurus elephas*): influence of current oceanographical features and historical processes. Molecular Phylogenetics and Evolution 48:708–717.

Pascual, M., B. Rives, C. Schunter & E. Macpherson,2017. Impact of life history traits on gene flow: A multispecies systematic review across oceanographic barriers in the Mediterranean Sea. PLoSONE 12:e0176419.

Piñera, J. A., D. Bernardo, G. Blanco, E. Vazquez &J. A. Sanchez, 2006. Isolation and characterization of polymorphic microsatellite markers in *Pagellus bogaraveo*, and cross-species amplification in *Sparus aurata* and *Dicentrarchus labrax*. Molecular Ecology Notes 6:33–35.

Pinet, P. R., 2009. Invitation to Oceanography, 5th edn. Jones and Bartlett Publishers, Burlington, MA, USA.

Piry, S., G. Luikart & J. M. Cornuet,1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size

using allele frequency data. Journal of Heredity 90:502–503.

Pritchard, J. K., M. Stephens, P. Donnelly,2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Ramsak, A., F. Garoia, I. Guarniero, P. Mannini& F. Tinti,2003. Novel polymorphic microsatellite markers for the common pandora (*Pagellus erythrinus*). Molecular Ecology Notes 3:553–555.

Raymond, M. & F. Rousset,1995. An exact test for population differentiation. Evolution 49:1280-1283.

R Development Core Team, 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0, URL http://www.R-project.org

Riginos, C., K. E. Douglas, Y. Jin, D. F. Shanahan, & E. A. Treml, 2011. Effects of geography and life history traits on genetic differentiation in benthic marine fishes. Ecography 34: 566-575.

Riginos, C. & L. Liggins, 2013. Seascape genetics: populations, individuals, and genes marooned and adrift. Geography Compass 7: 197–216.

Robitzch, V. S. N., D. Lozano-Cortés, N. M. Kandler, E. Salas & M. L. Berumen, 2016. Productivity and sea surface temperature are correlated with the pelagic larval duration of damselfishes in the Red Sea. Marine Pollution Bulletin 105: 566–574.

Ryman, N. & S. Palm, 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6:600–602.

Sabatés, A. & M. P. Olivar, 1996. Variation of larval fish distributions associated with variability in the location of a shelf-slope front. Marine Ecology Progress Series 135: 11–20.

859 Sala-Bozano, M., V. Ketmaier & S. Mariani, 2009. Contrasting signals from
860 multiple markers illuminate population connectivity in a marine fish.
861 Molecular Ecology 18:4811-4826.
862 Schunter, C., J. Carreras-Carbonell, E. Macpherson, J. Tintore, E. Vidal-
863 Vijande, & A. Guidetti &M. Pascual, 2011. Matching genetics with
864 oceanography: directional gene flow in a Mediterranean fish species.
865 Molecular Ecology 20:5167-5181.
866 Selkoe, K. A., K. T. Scribner& H. M. Galindo, 2016. Waterscape Genetics –
867 Applications of landscape genetics to rivers, lakes, and seas. In:
868 Balkenhol N, Cushman SA, Storfer AT, Waits LP (eds) Landscape
869 Genetics - concepts, methods, applications, Wiley & Sons, Chichester,
870 UK. p 220-246.
871 Silva, G., J. B. Horne& R. Castilho,2014. Anchovies go north and west
872 without losing diversity: post-glacial range expansions in a small pelagic
873 fish. Journal of Biogeography 41:1171–1182.
874 Simpson, S.D., H. B. Harrison, M. R. Claereboudt &S. Planes, 2014. Long-
875 Distance dispersal via ocean currents connects Omani clownfish
876 populations throughout entire species range. PLoSONE 9:e107610.
877 Slatkin, M., 1995. A measure of population subdivision based on
878 microsatellite allele frequencies. Genetics 139: 457-462
879 Swearer, S. E., J. E. Caselle, D. W. Lea & R. R. Warner, 1999. Larval
880 retention and recruitment in an island population of a coral-reef fish.
881 Nature 402: 799-802.
882 Weersing, K. & R. J. Toonen, 2009. Population genetics, larval dispersal,
883 and connectivity in marine systems. Marine Ecology Progress Series 393:
884 1–12.

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

White, C., K. A. Selkoe, J. Watson, D. A. Siegel, D. C. Zacher & R. J. Toonen, 2010. Ocean currents help explain population genetic structure. Proceedings of the Royal Society B 277: 1685-1694.

Valdés, P., A. García-Alcázar, I. Abdel, M. Arizcun, C. Suárez & E. Abellán, 2004. Seasonal changes on gonadosomatic index and maturation stages in common pandora *Pagellus erythrinus* (L.). Aquaculture International 12: 333–343.

Van Etten J., 2017. R Package gdistance: Distances and Routes on Geographical Grids. Journal of Statistical Software 76: 1-21.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills & P. Shipley, 2004. Micro- checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538.

Villamor, A., F. Costantini & M. Abbiati, 2014. Genetic Structuring across marine biogeographic boundaries in rocky shore invertebrates. PLoSONE, **9**, e101135.

Županović, Š. & I. Jardas, 1989. Fauna i flora Jadrana. Logos Split. 526 pp.

FIGURE LEGENDS

Fig. 1 Fig. 1 Map of the sampling sites of *P. erythrinus*, abbreviated as in Table1. Grey outline indicates the 300m isobath within which *P. erythrinus* adults can actively move and above which they cannot be found.

Fig. 2 Different models of population connectivity. Model 0: sampled populations belong to a unique panmictic unit; Model 1-7: populations are pooled according to different criteria or are considered as independent units, and different option of migration are allowed (see text for details).

Fig. 3 Frequency distribution of the four clusters identified by discriminant analysis of principal components across sampling sites.

Fig. 4 Migration rates between localities according to Model 3. (a) The median estimate is reported (vertical dash) with 50% (bold continuous line) and 95% (dotted line) confidence intervals; (b) Thickness of the arrows is proportional to the mean estimate.

ELECTRONIC SUPPLEMENTARY MATERIAL

Online resource 1 Summary statistics for 10 microsatellite loci over the 12 common pandora sampling sites analyzed.

Online resource 2 Locus/site combination tests for null alleles (NA; in parentheses their estimated frequencies) and for stuttering errors (ST). In grey, deviations from Hardy-Weinberg expectations.

Online resource 3 POWSIM simulations

Online resource 4 Discriminant Analysis of Principal Component (DAPC).

Online resource 5 Calculated mean connection time (MCT) between pairs of sites

Online resource 6 Comparison of the eight models. For each model, the marginal likelihood estimated using the Bezier approximation is reported.

Table 1 Details for sampling and summary statistics over ten loci. N: number of individuals, A: average number of alleles, N_P: number of private alleles, H_{exp}: mean expected heterozygosity, H_{obs}: mean observed heterozygosity, Fis: inbreeding coefficient (all the values are significant for $p<0.0001$).

Sea	Locality name	Geographic coordinates	Locality code[1]	Collection year	N	A	N_P	H_{exp}	H_{obs}	Fis
Adriatic Sea	Ancona	44°12'N 13°15'E	AN^M	2010	38	24	9	0.885	0.7746	0.138
	Bari	41°25'N 16°10'E	BA^M	2010	42	23.8	8	0.8614	0.7483	0.143
Ionian Sea	Acitrezza	38°20'N 16°90'E	JO^M	2010	39	25.4	7	0.8706	0.7576	0.143
Ligurian Sea	Elba	42°59'N 10°22'E	EN^M	2010	40	23.8	3	0.8953	0.7763	0.146
East Tyrrhenian Sea	Fiumicino	41°41'N 12°40'E	FM^C	2010	40	24.6	4	0.8721	0.7838	0.114
	Gaeta	41°11'N 13°26'E	GA^M	2010	32	22	6	0.867	0.7572	0.143
	Vibo Marina	38°45'N 16°80'E	VM^C	2010	40	22.9	6	0.8644	0.7382	0.158
Sardinian Sea	Castel Sardo	40°84'N 08°34'E	SN^M	2010	40	24.4	6	0.8825	0.7517	0.161
	Alghero	40°32'N 08°11'E	AL^M	2010	39	23.7	6	0.8793	0.7622	0.146
Sicilian Channel	Bizerte (Tunisia)	37°16'N 09°55'E	BZ^C	2009	40	24.2	3	0.8654	0.7678	0125
	Lampedusa	35°11'N 12°70'E	LM^C	2010	40	23.6	1	0.8745	0.7691	0.133
	Zarzis (Tunisia)	33°24'N 11°13'E	ZA^C	2009	40	24.1	6	0.8824	0.8025	0.103

[1] Superscript character indicate sampling source: M, MEDITS survey; C, commercial fishing

Table 2 F_{ST} values (Weir & Cockerham 1984, below the diagonal) and R_{ST} (Slatkin, 1995; above the diagonal) calculated for each pair of sampling locations. Significant values before and after FDR correction are indicated underlined and bold respectively. For sample codes see Table 1

	AN	BA	JO	EN	FM	GA	VM	SN	AL	BZ	LM	ZA
AN	-	0.0194	0.0016	0.0025	0.0014	-0.0031	0.0067	0.002	0.0038	-0.0083	0.0196	0.0097
BA	**0.0092**	-	0.0066	0.0169	0.0091	**0.0377**	0.0115	0.0049	0.0112	0.007	0.0079	0.0097
JO	0.0073	0.0055	-	0.0054	-0.0001	0.0119	0.0001	-0.0016	-0.0077	-0.0051	0.0067	0.0055
EN	0.0065	0.0064	0.0061	-	-0.0078	0.0052	0.0017	-902	0.0135	0.0024	0.0138	0.0125
FM	0.0055	0.0041	0.0025	0.0024	-	0.0014	-0.0056	-0.008	0.0043	-0.0006	0.0108	0.0131
GA	0.0024	0.0032	0.0023	0.0041	0.0005	-	0.0072	0.0063	0.015	0.003	0.0294	0.0332
VM	0.0075	0.0013	0.0046	0.0058	0.0008	0.0018	-	-0.0007	0.0012	0.0051	0.013	0.0048
SN	0.0063	0.004	0.0051	0.0003	0.0033	0.0042	0.0051	-	0.0055	-195	0.0103	0.0098
AL	0.0039	0.0011	-0.0009	0.0044	0.0002	-0.0016	0.0003	0.0059	-	-0.0025	0.0135	0.0004
BZ	0.004	0.0051	0.003	0.0072	0.0029	0.0017	0.0037	0.0063	0.0005	-	0.0027	0.0036
LM	0.0065	0.0027	0.004	0.0049	0.0017	0.0016	0.0039	0.0041	0.0022	0.0039	-	0.0082
ZA	0.0078	0.0042	0.0022	0.0067	0.0037	0.0057	0.006	0.0058	0.0002	0.0065	0.0058	-

Figure 1

Click here to access/download;Figure;Fig1-Mappa-campionamento_con_batimetria.eps

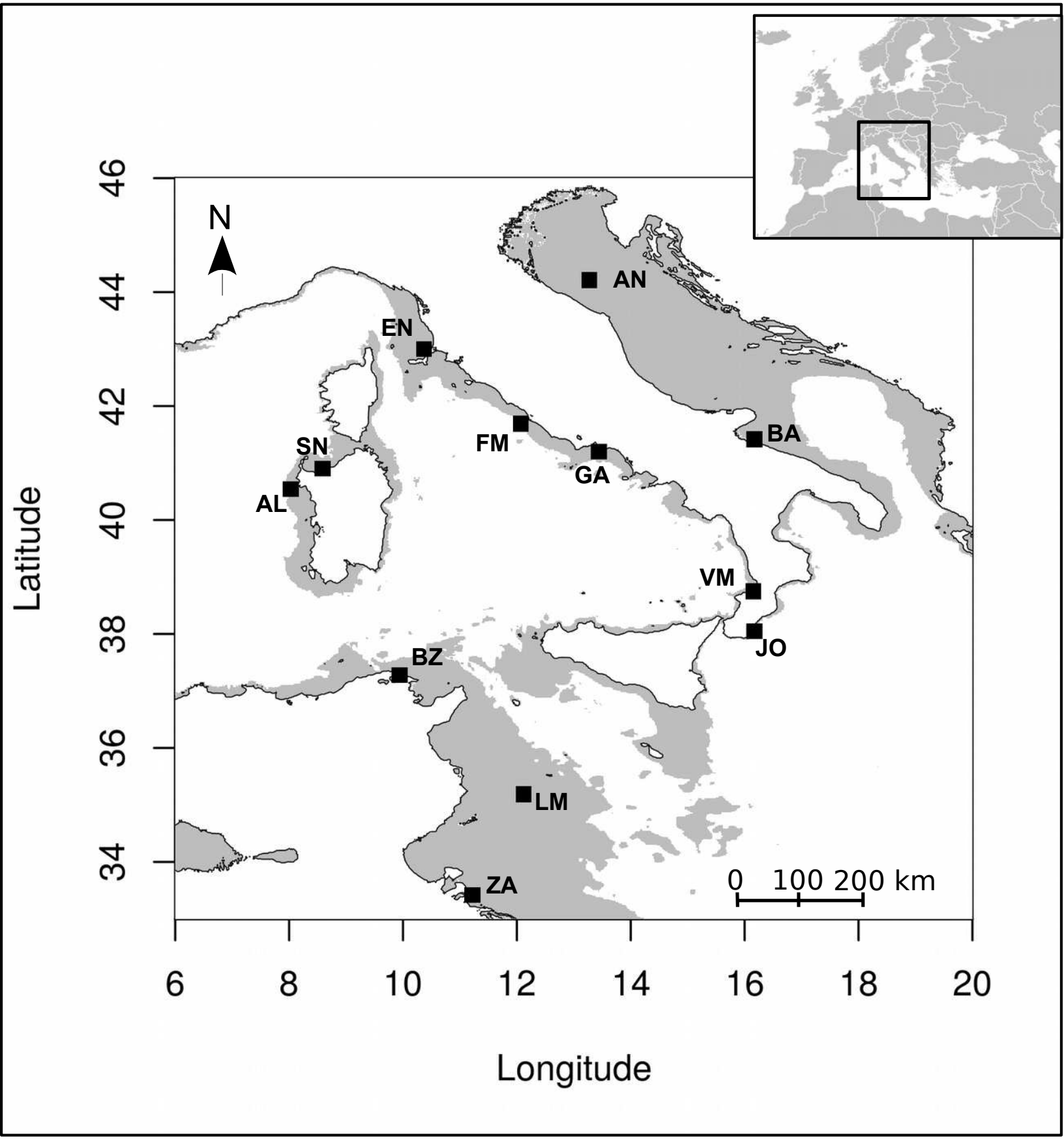


Figure 2

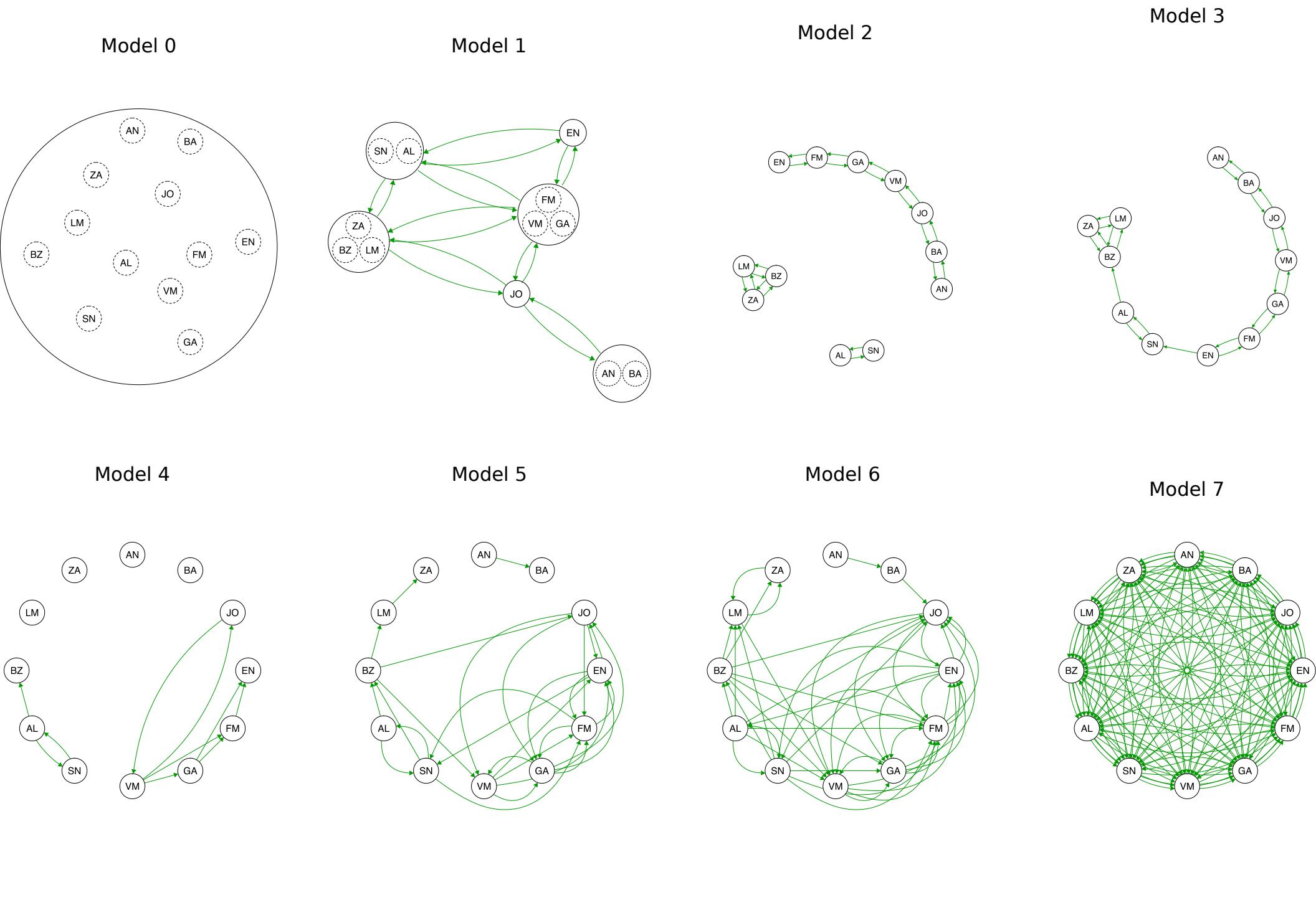
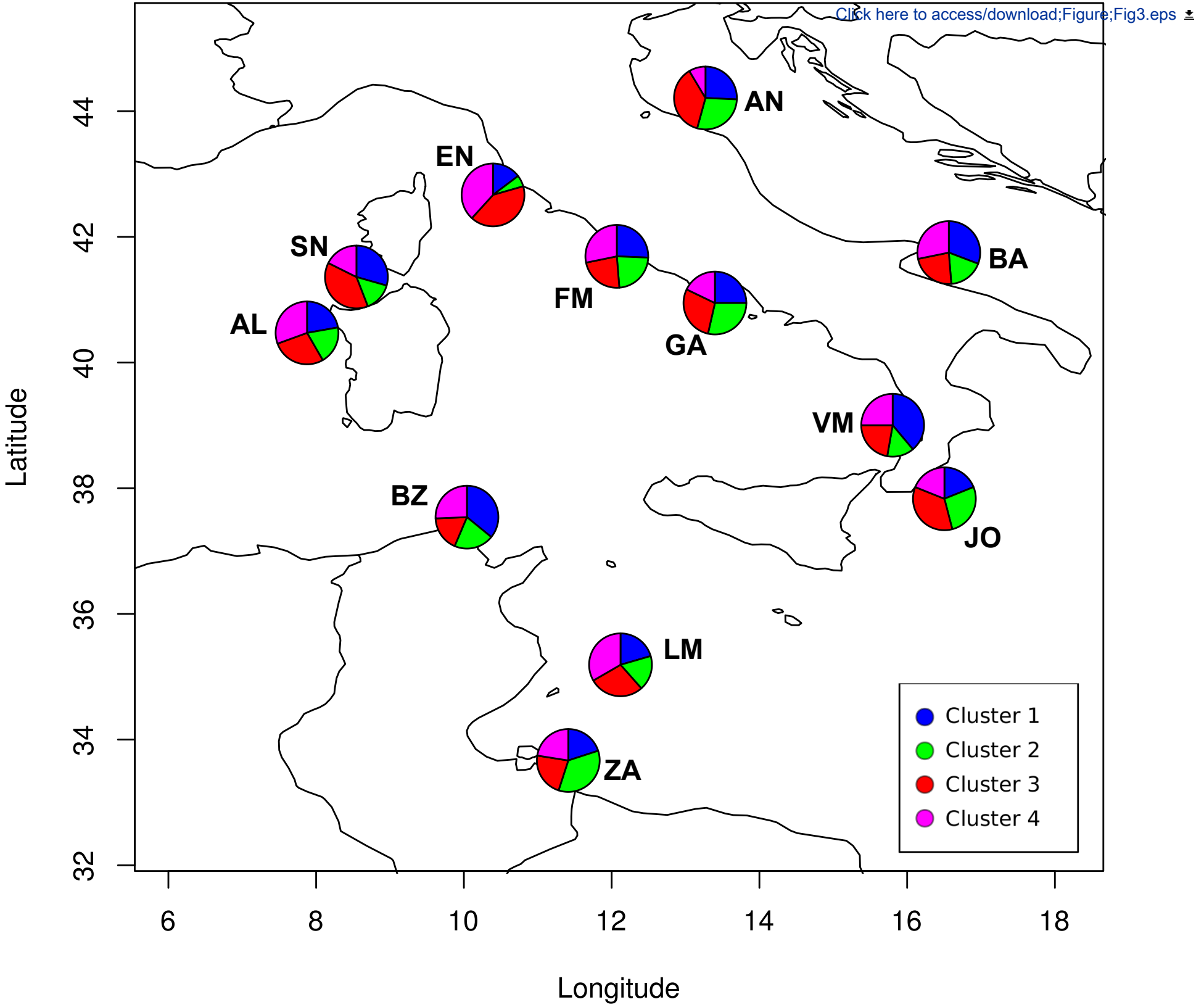
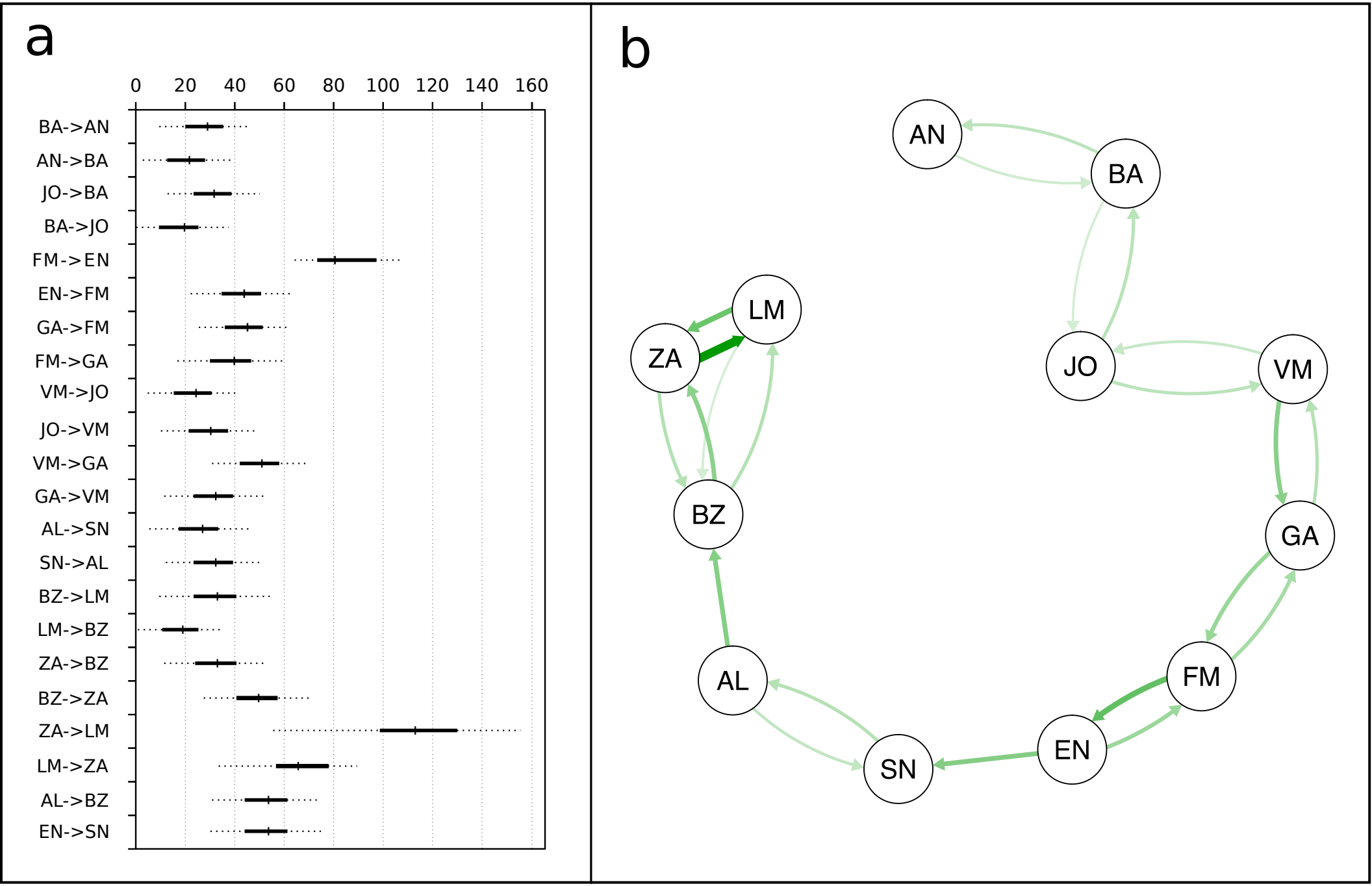


Figure 3





Electronic Supplementary Material

Influence of hydrodynamic connectivity on the genetic structure and gene flow of the common pandora *Pagellus erythrinus*

Anna Rita Rossi, Paolo Colangelo, Léo Berline, Elisa Angiulli, Giandomenico Ardizzone, Chiheb Fassatoui, Luciana Sola

Online resource 1. Summary statistics for ten microsatellite loci over the twelve common pandora sampling sites analysed

Locus	N	K	Np	A	A_R	H_{WE}	Fis	Null All
MS2	467	83	11	35.6	33.16	High sign	0.272	0.131
MS3	465	43	13	15.7	14.25	High sign	0.183	0.075
MS4	465	42	6	25.4	23.46	High sign	0.092	0.043
MS6	468	14	3	6.4	6.11	0.0000	0.089	0.047
SL7	464	55	5	34.9	31.58	High sign	0.130	0.063
SL17	469	77	6	33	28.75	High sign	0.099	0.050
SL26	465	34	7	17.8	17.58	High sign	0.081	0.035
SL29	464	61	4	37.1	33.69	High sign	0.071	0.033
SL33	466	17	6	8.3	7.89	High sign	0.354	0.160
PbOviD102	466	38	4	22.6	22.60	0.0280	0.032	0.016

N = number of individuals, K = allele number, NP = number of private alleles, A = average number of allelels, AR = Allelic richness, HWE = significance deviation from the Hardy-Weinberg equilibrium, Fis = inbreeding coefficient; Null All = null alleles frequencies

Electronic Supplementary Material

Influence of hydrodynamic connectivity on the genetic structure and gene flow of the common pandora *Pagellus erythrinus*

Anna Rita Rossi, Paolo Colangelo, Léo Berline, Elisa Angiulli, Giandomenico Ardizzone, Chiheb Fassatoui, Luciana Sola

Online resource 2. Locus/site combination tests for null alleles (NA; in parentheses their estimated frequencies) and for stuttering errors (ST). In grey deviations from Hardy–Weinberg expectations

Locus	Locality											
	AN	BA	JO	EN	FM	GA	VM	SN	AL	BZ	LM	ZA
MS2	NA (0.18605)	NA (0.08541)	NA (0.09873)	NA (0.11984)	NA (0.09713)	NA (0.12473)	NA (0.15418)	NA (0.04502)	NA (0.12457)	NA (0.13913)	NA (0.16967)	NA (0.13864)
MS3	(0.01970)	(0.03759)	(0.05090)	(0.06050)	NA (0.12828)	NA/ST (0.13478)	(0.07617)	NA/ST (0.09982)	(0.07332)	(0.01328)	(0.09435)	(0.02852)
MS4	(0.06117)	(0.05013)	(0.01877)	(0.02947)	(0.05440)	(0.01664)	NA (0.07836)	(0.01451)	NA (0.06293)	(0.00000)	(0.00630)	(0.04910)
MS6	(0.00000)	(0.08305)	(0.00570)	(0.14292)	(0.00402)	(0.00000)	(0.13088)	(0.00465)	(0.09737)	(0.00000)	(0.00000)	(0.01637)
SL7	(0.03341)	NA (0.08705)	NA (0.07270)	(0.04892)	(0.06008)	NA (0.08435)	(0.03329)	NA (0.11093)	(0.03009)	NA (0.06539)	(0.04558)	(0.01979)
SL17	(0.02042)	(0.08967)	(0.10918)	(0.00063)	(0.00476)	(0.00334)	(0.06868)	(0.03620)	(0.06292)	(0.07935)	(0.06352)	(0.00000)
SL26	(0.04173)	(0.03611)	(0.03955)	NA (0.08419)	(0.00000)	(0.04894)	(0.00000)	(0.05848)	NA (0.07646)	(0.04591)	(0.00000)	(0.00000)
SL29	(0.02698)	(0.00000)	NA (0.06144)	(0.01355)	(0.00000)	(0.02539)	(0.00000)	(0.01637)	(0.04653)	NA (0.04492)	(0.03507)	(0.05382)
SL33	NA/ST (0.20935)	NA/ST (0.12897)	NA/ST (0.16804)	NA/ST (0.08890)	NA/ST (0.10385	NA/ST (0.11962)	NA/ST (0.16804)	NA/ST (0.22001)	NA/ST (0.06403)	NA/ST (0.15800)	NA/ST (0.14429)	NA/ST (0.19752)
PbOviD102	(0.03193)	(0.02964)	(0.00018)	(0.00610)	(0.00000)	(0.00018)	(0.00646)	(0.05340)	(0.00000)	(0.00464)	(0.00000)	(0.00002)

Electronic Supplementary Material

Influence of hydrodynamic connectivity on the genetic structure and gene flow of the common pandora *Pagellus erythrinus*

Anna Rita Rossi, Paolo Colangelo, Léo Berline, Elisa Angiulli, Giandomenico Ardizzone, Chiheb Fassatoui, Luciana Sola

Online resource 3: POWSIM simulations to assess the statistical power of microsatellite loci to differentiate populations at varying levels of expected FST. Results of X2 and Fisher's exact tests for the proportion of simulations out of 1000 significant with a critical value of 0.05.

Expected F_{ST}	N_e	t	X^2 test (%)	Fisher's test (%)
0.0005	1.000/ 5.000/ 10.000	1/ 5/ 10	49.7/ 52.7/ 50.5	47.5/ 50.3/ 49.3
0.0010	1.000/ 5.000/ 10.000	2/ 10/ 20	95.5/ 94.9/ 95.7	90.4/ 91.8/ 90.9
0.0025	1.000/ 5.000/ 10.000	5/ 25/ 50	100/ 100/ 100	100/ 100/ 100
0.0050	1.000/ 5.000/ 10.000	10/ 50/ 100	100/ 100/ 100	100/ 100/ 100
0.0100	1.000/ 5.000/ 10.000	20/ 100/ 200	100/ 100/ 100	100/ 100/ 100

Electronic Supplementary Material

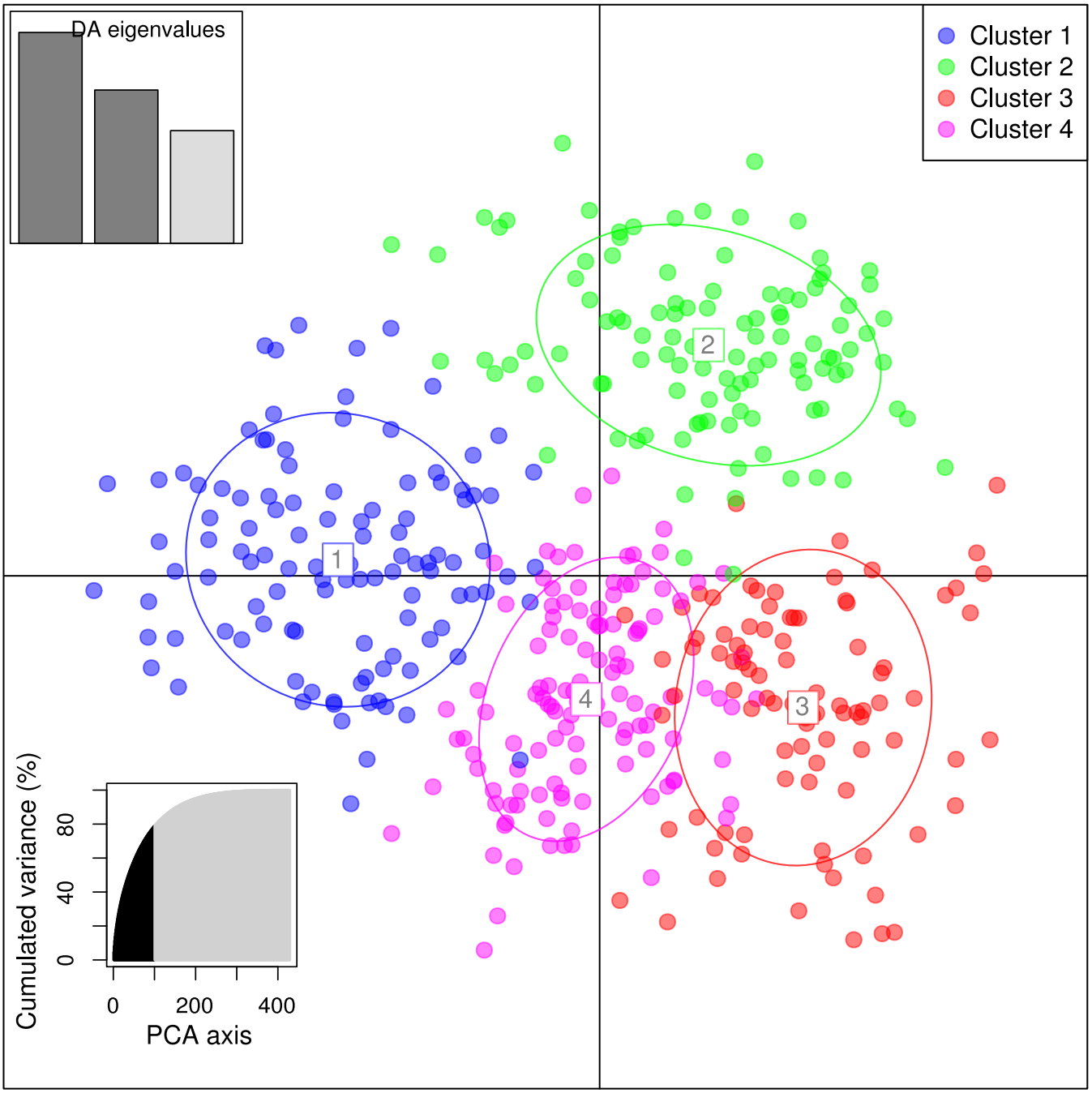
Influence of hydrodynamic connectivity on the genetic structure and gene flow of the common pandora Pagellus erythrinus

Anna Rita Rossi, Paolo Colangelo, Léo Berline, Elisa Angiulli, Giandomenico Ardizzone,

Chiheb Fassatoui, Luciana Sola

Online resource 4

Discriminant Analysis of Principal Component (DAPC). The obtained graph represents the individuals as dots and the groups (identify by k-means) as different colors. Eigenvalues of the discriminant axes (DA) and the comulative variances of the PC axes used are displayed in insets.



Electronic Supplementary Material

Influence of hydrodynamic connectivity on the genetic structure and gene flow of the common pandora *Pagellus erythrinus*

Anna Rita Rossi, Paolo Colangelo, Léo Berline, Elisa Angiulli, Giandomenico Ardizzone, Chiheb Fassatoui, Luciana Sola

Online resource 5:
Mean connection time (MCT) among pair of sites. Only MCT below 200 days were considered. NaN indicates MCTs above 200 days

	To											
From:	**AN**	**BA**	**JO**	**EN**	**FM**	**GA**	**VM**	**SN**	**AL**	**BZ**	**LM**	**ZA**
AN	-	115.343	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
BA	NaN	-	199.327	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
JO	NaN	NaN	-	148.421	120.661	105.241	59.278	NaN	NaN	NaN	NaN	NaN
EN	NaN	NaN	180.97	-	113.199	127.119	179.191	112.058	191.018	NaN	NaN	NaN
FM	NaN	NaN	156.431	74.65	-	102.64	170.539	145.889	NaN	NaN	NaN	NaN
GA	NaN	NaN	140.907	94.269	55.141	-	123.078	NaN	NaN	NaN	NaN	NaN
VM	NaN	NaN	63.732	128.283	92.117	82.791	-	NaN	NaN	NaN	NaN	NaN
SN	NaN	NaN	196.626	153.813	129.73	194.615		-	53.295	130.72	194.761	NaN
AL	NaN	NaN	182.998	NaN	186.644		168.05	77.288	-	91.82	154.64	NaN
BZ	NaN	NaN	134.09	NaN	167.053	161.51	108.496	249.425	NaN	-	103.632	184.872
LM	NaN	NaN	NaN	NaN	NaN	NaN	160.144	NaN	NaN	NaN	-	119.735
ZA	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	162.997	-

Electronic Supplementary Material

Influence of hydrodynamic connectivity on the genetic structure and gene flow of the common pandora *Pagellus erythrinus*

Anna Rita Rossi, Paolo Colangelo, Léo Berline, Elisa Angiulli, Giandomenico Ardizzone, Chiheb Fassatoui, Luciana Sola

Online resource 6: Model comparison of the eight models. For each model the marginal likelihood estimated using the Bezier approximation is reported. The probability of each model was calculated following Beerli and Palczewski (2010)

Model number	Bezier approximation	P model
Model 0	-7006182.17	0
Model 1	-546114.82	0
Model 2	-3788213.34	0
Model 3	-349000.46	1
Model 4	-3516744.21	0
Model 5	-1262798.92	0
Model 6	-442189.13	0
Model 7	-1218611.85	0