



The hidden side of a major marine biogeographic boundary: a wide mosaic hybrid zone at the Atlantic-Mediterranean divide reveals the complex interaction between natural and genetic barriers in mussels

Tahani El Ayari, Najoua Trigui El Menif, Bojan Hamer, Abigail E. Cahill,
Nicolas Bierne

► To cite this version:

Tahani El Ayari, Najoua Trigui El Menif, Bojan Hamer, Abigail E. Cahill, Nicolas Bierne. The hidden side of a major marine biogeographic boundary: a wide mosaic hybrid zone at the Atlantic-Mediterranean divide reveals the complex interaction between natural and genetic barriers in mussels. *Heredity*, 2019, 122 (6), pp.770-784. 10.1038/s41437-018-0174-y . hal-02403713

HAL Id: hal-02403713

<https://hal.science/hal-02403713>

Submitted on 10 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**1The hidden side of a major marine biogeographic boundary: a wide
2mosaic hybrid zone at the Atlantic-Mediterranean divide reveals the
3complex interaction between natural and genetic barriers in mussels**

4Tahani El Ayari^{1,2}, Najoua Trigui El Menif², Bojan Hamer³, Abigail E. Cahill⁴ and
5Nicolas Bierne¹

6¹ISEM, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

7² University of Carthage, Faculty of Sciences of Bizerta. Laboratory of Environment
8Bio-monitoring, 7021- Zarzouna, Bizerta, Tunisia

9³ Ruđer Bošković Institute, Center for Marine and Environmental Research, Grad
10Zagreb, Croatia

11⁴ Biology Department, Albion College, Michigan 49224, USA

12

13Name of the corresponding author: Nicolas Bierne

14Institut des Sciences de l'Évolution - Montpellier, Université de Montpellier, Place
15Eugène Bataillon

1634095 Montpellier Cedex 05 FRANCE

17Telephone: +33(0)4 67 14 36 88; nicolas.bierne@umontpellier.fr

18

19Running title: A mosaic tension zone at the Atlantic-Mediterranean divide

20

21

22Abstract

23The Almeria-Oran Front (AOF) is a recognised hotspot of genetic differentiation in
24the sea, with genetic discontinuities reported in more than 50 species. The AOF is
25a barrier to dispersal and an ecological boundary, and both facts can explain the
26position of these genetic breaks. However, the maintenance of genetic
27differentiation is likely reinforced by genetic barriers. A general drawback of
28previous studies is an insufficient density of sampling sites at the transition zone
29with a conspicuous lack of samples from the southern coastline. We analysed the
30fine-scale genetic structure in the mussel *Mytilus galloprovincialis* with a few
31ancestry-informative loci previously identified from genome scans. We discovered
32a 600 km wide mosaic hybrid zone eastward of the AOF along the Algerian
33coasts. This mosaic zone provides a new twist to our understanding of the
34Atlantic-Mediterranean transition because it demonstrates the two lineages can
35live in sympatry with ample opportunities to interbreed in a large area, but hardly
36do. This implies some form of reproductive isolation must exist to maintain the
37two genetic backgrounds locally cohesive. The mosaic zone ends with an abrupt
38genetic shift at a barrier to dispersal in the Gulf of Bejaia, Eastern Algeria.
39Simulations of endogenous or exogenous selection in models that account for the
40geography and hydrodynamic features of the region support the hypothesis that
41sister hybrid zones could have been differentially trapped at two alternative
42barriers to dispersal and/or environmental boundaries, at Almeria in the north
43and Bejaia in the south. A preponderantly unidirectional north-south gene flow
44next to the AOF can also maintain a patch of an intrinsically maintained genetic
45background in the south and the mosaic structure, even in the absence of local
46adaptation. Our results concur with the coupling hypothesis that suggests natural
47barriers mostly explain the position of genetic breaks while their maintenance
48must additionally require genetic barriers.

49

50**Keywords:** Barrier to gene flow, mosaic hybrid zone, reproductive isolation,
51marine connectivity, Almeria-Oran Front, *Mytilus galloprovincialis*.

52

53Introduction

54Natural populations are often subdivided into discrete patches within which
55apparent panmixia is observed or where genetic differentiation smoothly
56increases with geographic distance (Charlesworth *et al*, 2003). These roughly
57homogeneous genetic patches are bordered by abrupt genetic discontinuities
58maintained by “barriers to gene flow” (Arnold, 2006 ; Slatkin, 1987; Hewitt,
591989). Barriers to gene flow allow divergence between populations and possibly
60the evolution of reproductive isolation and speciation (Barton and Bengtsson,
611986; Palumbi, 1994; Charlesworth *et al*, 2003). They also guarantee the
62maintenance of the beta component (i.e. between-populations component) of
63genetic and species diversity. Barriers to gene flow are therefore crucial in
64structuring and maintaining biodiversity. The study of their origin, their location
65and the processes involved in their functioning, is therefore a central issue in
66evolutionary biology and ecology.

67When genetic discontinuities are spatially concordant among different species, a
68parsimonious explanation is that the same factors contribute to the barrier across
69species. For instance, species can share a common history (e.g. shared zones of
70secondary contact sometimes called suture zones; [Remington, 1968](#); [Hewitt, 1996](#);
711996; Avise, 2000; Hewitt, 2000; Swenson and Howard, 2005), a comparable
72mode of dispersal (e.g. planktonic dispersers should be similarly affected by
73marine connectivity; Pelc *et al*, 2009) or a similar ecological niche (e.g.
74adaptation to the same habitat such as brackish water in the Baltic Sea;
75Johannesson and Andre, 2006). In the marine realm, decades of research in
76population genetics have allowed researchers to identify and describe hotspots of
77genetic differentiation (Rocha and Bowen, 2008; Hellberg, 2009; Riginos and
78Liggins, 2013; see Table 1). Probably because many marine species have large
79population sizes and high dispersal potential via the larval phase, genetic
80homogeneity is almost always the rule within patches delineated by these
81barriers and the concentration of genetic breaks at some locations is obvious in
82the sea (Hellberg, 2009; Riginos *et al*, 2011; Gagnaire *et al*, 2015). Barriers to
83gene flow in the marine environment are found at locations of strong
84environmental gradients or oceanographic currents which prevent dispersal along
85shores (Hellberg *et al*, 2002; Teske *et al*, 2014). Many factors can be involved:
86physical (oceanic fronts, local and global oceanic currents; Patarnello *et al*, 2007),
87ecological (temperature, salinity, tide level; Johannesson *et al*, 2010; Stanley *et*

88al. 2018), historical (secondary contact; [Quesada et al., 1995](#); Avise, 2000), or
89genetic (reproductive isolation; Bierne et al, 2011). While barriers to gene flow
90are more likely to result from a combination of factors acting synergistically
91(Barton and Hewitt, 1985; Bierne et al, 2011; Cordero et al, 2014), a single factor
92is often emphasised to explain the genetic structure in any one case. In addition,
93the origin (primary divergence vs. secondary contact), localisation (ecotone vs.
94natural barrier vs. suture zone) and maintenance (dispersal barrier vs. local
95adaptation vs. reproductive isolation) of the genetic differentiation need to be
96treated as different questions. For instance, the position of a genetic cline at a
97barrier to dispersal does not necessarily mean that the genetic differentiation is
98due to low dispersal at this location; it is equally likely to be the legacy of
99divergence during past vicariance and/or the additional effect of a genetic barrier
100trapped by the natural barrier (Barton 1979). Demographic reconstruction can
101provide insight about the origin of the differentiation that the simple description
102of the contemporary genetic structure does not reveal. However, to better
103understand the maintenance of genetic breaks, the genetic analysis of distant
104samples far from the transition zone is hardly sufficient, even with high-
105throughput genomic data (Cruickshank and Hahn, 2014; Meirmans, 2015;
106Harrison and Larson, 2016). Such genome-wide analyses allow reconstructing the
107long term history of populations and identifying highly differentiated loci or
108genomic islands (Tine et al, 2014; Fraïsse et al, 2016), but tell little about how
109genetic lineages have the opportunity to interbreed at the boundary. A major
110shortcoming of many genome-wide studies is that they do not undertake detailed
111sampling at these points of geographic transition, though such sampling is critical
112for teasing apart competing explanations. Ultimately, one also needs to conduct
113experiments in the lab and in the field to obtain direct evidence of the isolation
114mechanisms at play (Harrison, 1993).

115The Almeria-Oran Front (AOF) is a well-known marine barrier to gene flow
116between populations of the Atlantic Ocean and the Mediterranean Sea (Quesada
117et al, 1995b; Borsa et al, 1997; Borrero-Pérez et al, 2011; Chevolot et al, 2006;
118Patarnello et al, 2007; Palero et al, 2008; Schunter et al. 2011; Tine et al, 2014).
119The front is generated by the convergence of Mediterranean and Atlantic waters.
120Atlantic waters enter into the Alboran Sea through the strait of Gibraltar and
121produce anticyclonic gyres westward of the AOF, the Western and Eastern
122Alboran Gyres, and a north-south unidirectional current at the AOF (see Figure 1).
123In addition, part of the flow of the Eastern Alboran Gyre is trapped at the AOF to

124contribute to the Algerian current that flows eastward of Oran along the Algerian
125coastline (Viúdez and Tintoré, 1995).

126Among the examples of genetic breaks at the AOF, one of the first described was
127found in the marine mussel *Mytilus galloprovincialis* (Quesada *et al*, 1995a;
1281995b, Sanjuan *et al*. 1994). Since the early reports, the origin of the genetic
129break has been considered a secondary contact between two lineages isolated in
130Mediterranean and Atlantic glacial refugia (Quesada *et al*, 1995a), as for other
131species such as sea bass (Lemaire *et al*, 2005; Tine *et al*, 2014; Duranton *et al*,
1322018). The genetic differentiation between the two mussel lineages proved to be
133highly heterogeneous among loci (Gosset and Bierne, 2013; Fraïsse *et al*, 2016).
134Heterogeneous differentiation is expected for a semi-permeable genetic barrier
135maintained by partial reproductive isolation (Barton and Hewitt, 1985; Harrison,
1361993) or because of heterogeneous divergence caused by linked selection
137(Cruickshank and Hahn 2014; Ravinet *et al*. 2017), or both (Duranton *et al*. 2018).
138To date only the northern coast near Almeria in Spain has been intensively
139sampled, and contains an abrupt shift in allele frequency on each side of a 290
140km wide no-mussel zone between Almeria and Alicante (Quesada *et al*, 1995a).
141Although some studies have identified weak genetic differentiation between
142Europe and North Africa east of the AOF (Ouagajjou and Preza 2015; Lourenço *et*
143al. 2015), the geographic structure of the transition zone along the southern
144coast in Algeria was not well characterised. This is an important sampling gap to
145fill, especially considering that Atlantic waters enter the Mediterranean Sea
146through the South by the Algerian current described above. The southern side of
147the transition is indeed a ubiquitous sampling lacuna in the study of the Atlantic-
148Mediterranean genetic divide.

149In the present study we aimed to advance our understanding of the marine
150transition zone at the AOF by analysing four ancestry-informative loci (one
151mitochondrial and three nuclear) at a fine-grained spatial scale. These loci were
152previously identified to be among the ten most differentiated loci in two genome
153scans, one with 388 AFLP loci (Gosset and Bierne, 2013), and another with 1269
154targeted contigs (~50,000 SNPs, Fraïsse *et al*, 2016). We would like to make it
155clear that our aim in the present study was not to infer the history of gene flow
156between populations nor the genomic heterogeneity of differentiation, which
157requires the analysis of a high number of loci and has been treated elsewhere
158(Fraïsse *et al*, 2016). Our objective was to investigate the spatial structure of the

159transition zone and how the two genetic backgrounds meet and interact when in
160contact, which requires analysing an extensive geographic sample with
161informative loci (Vines *et al*, 2016). We accomplished this in three steps. First, as
162a precaution we re-analysed the shift in allele frequency in South-Eastern Spain
163that was reported twenty years ago. We observed the same position of the break
164in Spain as earlier studies, confirming results from Diz and Presa (2008). Second,
165we sampled a new area along the Algerian coast between Oran and Tunisia, over
166a distance of 1330 km. We report a 600 km wide mosaic hybrid zone with
167populations of mixed ancestry. Third, we explored simple stepping-stone
168simulation models which show that the mosaic structure observed can be
169explained by the local geographic and hydrodynamic characteristics of the
170coastlines. We used a Y-shaped bi-unidimensional structure of the coastlines to
171reflect the shape of the coastlines that allows north-south gene exchange west of
172the AOF but only along coastlines to the east of the AOF because the distance
173between Algerian and Spanish coasts become too large for larval dispersal. We
174also included two alternative barriers to larval dispersal and/or environmental
175boundaries (one at the AOF and one in the Gulf of Bejaia), and a preponderantly
176unidirectional north-south gene flow at the AOF.

177

178

179 MATERIALS AND METHODS

180 *Sampling and molecular markers*

181 *Mytilus galloprovincialis* (Lamarck, 1819) samples were collected from 13
182 locations on the southern coast of the Alboran and Mediterranean Seas from
183 Nador (Morocco) to Bizerte (Tunisia) and in 7 locations on the Northern coast from
184 Manilva (Spain) to Peniscola (Table S1, Figure 1). Mitochondrial DNA sequences
185 proved that all the samples were *Mytilus galloprovincialis* (see below) as
186 expected in the study area. We also analysed reference samples already reported
187 in previous studies (Gosset and Bierne, 2013) in Faro (Portugal) and Sète
188 (France). DNA was extracted from gills using the QIAGEN DNeasy Blood & Tissue
189 Kit following the instructions of the manufacturer. DNA concentration was
190 measured for each sample using a NanoDrop8000 Spectrophotometer (Thermo
191 Scientific) and standardised to a DNA concentration of 100 ng μ L⁻¹.

192 A 350-bp fragment of the F-mtDNA cytochrome oxidase subunit III was PCR
193 amplified and sequenced with primers *FOR1* (5'-TATGTACCAGGTCCAAGTCCGTG-
194 3') and *REV1* (5'-TGCTCTTCTTGAATATAA GCGTA-3') (Zouros *et al*, 1994).
195 Sequence reactions were precipitated using a standard EDTA/ethanol protocol,
196 suspended in 15 μ l Hi-Di formamide and sequenced on an ABI 3130XL automated
197 sequencer.

198 Based on previous genome scans, three nuclear markers were chosen based on
199 both their strong differentiation between the Atlantic and Mediterranean
200 backgrounds of *M. galloprovincialis* and for their ease of analysis. All three
201 produced length-polymorphism of PCR products. *EFbis* is a widely used marker in
202 mussels that was amplified with the following two primers: *EFbis-F* 5'-
203 ACAAGATGGACAATACCGAACCACC-3' and *EFbis-R* 5'-CTCAAT
204 CATGTTGTCTCCATGCC-3' (Bierne *et al*, 2002). *EF2* has been described in Gosset
205 and Bierne (2013) and was amplified with the following two primers: *EF2-F* 5'-
206 GGAAATCCCATGGGTGATTAGCGG-3' and *EF2-R* 5'-
207 GTCAAATAAATACTGAAACACAGTGACTTC-3'. A contig containing the *Precollagen-D*
208 gene was identified from a genome scan of 1269 contigs as the most
209 differentiated contig between Atlantic and Mediterranean *M. galloprovincialis*
210 populations (Fraïsse *et al*, 2016). Using publicly available sequences (Qin *et al*,
211 1997) we designed primers to amplify the fifth intron of the gene with the 3' end
212 of each primer positioned at the exon-intron junction, which prevented

213 amplification of paralogous sequences of the gene family. The primers used to
 214 amplify *Precol-D* locus were *Precol-D-F* 5'-GACAAGGACCAGCAGGTACCATATT-3'
 215 and *Precol-D-R* 5'-GAGTGGTCCGGCTGGTCCTAAGAA-3'. Indel polymorphisms in
 216 the intron produced nine length alleles in our samples with strong allele
 217 frequency differences between Atlantic and Mediterranean samples. Standard
 218 PCR protocols were used with annealing temperatures set at 54°C for *EFbis* and
 219 *EF2*, and 49°C for *Precol-D*. 2 µl of PCR products were mixed with Hi-Di
 220 formamide/ROX 500 size standard (12.8 µl formamide and 0.2 µl ROX), and the
 221 mixtures were then loaded on an ABI 3130XL capillary automated sequencer.
 222 GeneMapper® v4.5 software (Applied Biosystems) was used to read the resulting
 223 chromatograms.

224 **Data analysis**

225 *COIII* sequences were aligned using the ClustalW multiple alignment algorithm in
 226 BIOEDIT V7.1.3.0 and revised manually. Aligned sequences were then used to
 227 construct a Neighbour-Joining tree with MEGA software v6 (Tamura *et al.*, 2013).
 228 Allelic and genotyping frequencies were computed using Genetix software
 229 v4.0.5.2 (Belkhir *et al.*, 2002). Number of haplotypes (*h*), nucleotide diversity (π),
 230 average number of nucleotide substitutions per site between populations (*D_{xy}*)
 231 and net number of nucleotide substitutions per site between populations (*D_a*)
 232 were calculated using DnaSP V5.10.1 (Librado and Rozas, 2009). Individuals were
 233 assigned to the Mediterranean and Atlantic lineages with a discriminant analysis
 234 of principal components (DAPC) implemented in the adegenet package in R
 235 (Jombart and Ahmed, 2011). The individual ancestries were examined by a
 236 Bayesian clustering method implemented in STRUCTURE 2.3.4 (Falush *et al.*,
 237 2003). Analysis with STRUCTURE was conducted for various numbers of clusters
 238 *K* under the admixture model, assuming that the allele frequencies are correlated
 239 across populations, with a burn-in period of 50,000 steps and a run length of
 240 100,000 iterations. We used NewHybrids software (Anderson and Thompson,
 241 2002) to determine the hybrid status of individuals in the hybrid zones. We used
 242 a model with six categories of genotypes corresponding to two parental lineages
 243 and four categories of hybrids: F1 and F2 hybrids and the two types of
 244 backcrosses. Note that there are many more categories of hybrid genotypes in
 245 natural hybrid zones. However, this analysis allows investigating the frequency of
 246 early generation hybrids (expected to be rare in natural bimodal hybrid zones).
 247 Individual posterior probabilities of belonging to each category were obtained by

248running NewHybrids with uniform priors, a burn-in period of 50 000 steps
 249followed by a run length of 100 000 iterations. Because parental lineages tend to
 250be more introgressed locally within a mosaic hybrid zone than outside of the zone
 251in allopatric populations (Bierne *et al*, 2003) we did not use allopatric populations
 252as references in the analysis. In this context including allopatric samples would
 253have inflated the posterior probabilities of locally introgressed parental genotype
 254to be backcrosses, which is not desired. To depict allele frequency clines, all four
 255loci were transformed into bi-allelic loci by pooling alleles according to their
 256frequencies in reference samples (sample 1 Faro Portugal for the Atlantic lineage
 257and sample N9 Sète France for the Mediterranean lineage, two localities often
 258studied in previous genetic studies of *M. galloprovincialis*). Finally, the average
 259departure from Hardy-Weinberg equilibrium was measured by F_{is} and the average
 260pairwise linkage disequilibrium across all loci, D , was estimated using the
 261variance in the hybrid index as described in Barton and Gale (1993). Departure
 262from Hardy-Weinberg and linkage equilibrium was tested by permutations with
 263the Genetix software combining p-values using Fisher's method.

264**Simulations**

265We used a model of evolution in a metapopulation of 2×60 demes arranged in
 266two parallel linear stepping-stones (Bierne *et al.*, 2011; 2013). The purpose of the
 267simulations was to account for the geographic and hydrodynamic characteristics
 268of the region and to compare intrinsic reproductive isolation with local adaptation
 269at two different spatial scales (fine- or coarse-grained environmental
 270heterogeneity). At generation zero, two partially reproductively isolated
 271backgrounds meet and start to exchange genes. The auto-recruitment rate was 1
 272- $2m$, and migration to adjacent demes was m . A weak barrier to dispersal was
 273set between demes 20 and 21 in the northern chain to simulate the AOF, and
 274another barrier to dispersal in the southern chain between demes 40 and 41 to
 275simulate the Gulf of Bejaia Barrier (GBB). The migration rate at barriers to
 276dispersal was set to a value m_{bar} , with $m_{bar} < m$. Bidirectional migration was
 277possible between the northern and the southern chain for the first 20 demes (left
 278of the first barrier to dispersal, i.e. in the Alboran Sea) with the same rate of
 279migration m . For a few more demes (from 1 to 10) bidirectional or unidirectional
 280north-south gene flow was possible between the north and the south and then
 281stopped for the remaining demes. The aim of the parameterization was to
 282account for the fact that the Spanish and Algerian coasts east of the AOF become

283 too distant to be connected directly by a single generation of larval dispersal,
 284 which is thought to be roughly around 50 km gen⁻¹ in mussels (McQuaid and Phillips
 285 2000, Gilg and Hilbish 2003). Selection acted at a local adaptation locus
 286 (exogenous selection) or against recombinant genotypes (intrinsic selection) at
 287 two freely recombining reproductive isolation loci (Bierne *et al.*, 2011). Allelic
 288 fitnesses at the exogenous locus were $W(C) = 1$ and $W(c) = 1 - t$ in habitat 1 and
 289 $W(C) = 1 - t$ and $W(c) = 1$ in habitat 2. The two-locus fitnesses at the
 290 incompatible alleles were $W(AB) = W(ab) = 1$ and $W(Ab) = W(aB) = 1 - s$. We did
 291 not consider the interaction between the two types of selection here (see Bierne
 292 *et al.* 2011) but consider either exogenous or endogenous selection alone. A
 293 neutral marker was positioned at a recombination map distance of 1 cM to a
 294 selected locus.

295 In order to simulate the trapping of a hybrid zone by a barrier to dispersal or an
 296 environmental boundary, we need clines to overlap with barriers at their initial
 297 position (Goldberg and Lande, 2007) or to move toward the barrier (Barton and
 298 Turelli, 2011). We could have chosen the initial conditions and the parameter
 299 values such that the northern cline overlapped with the AOF (between demes 20
 300 and 21) and the southern cline overlapped with the GBB (between demes 40 and
 301 41 of the southern chain). However, we have chosen to exemplify the arguably
 302 more realistic situation of initially moving hybrid zones. Local adaptation clines
 303 move to environmental boundaries driven by local selection. However, tension
 304 zones maintained by selection against recombinant genotypes stay at their initial
 305 contact position in a purely deterministic genetic model with equal fitnesses of
 306 parental genotypes and isotropic migration. There are three main reasons why
 307 tension zones could move. First, they are expected to move down gradients of
 308 population density and to be trapped by density troughs (Abbott *et al.*, 2013;
 309 Barton, 1979; Hewitt, 1975). This demographic effect is not accounted for in our
 310 purely genetic model, although the tension zone can be trapped by barriers to
 311 dispersal providing it moves towards the barriers for other reasons. Second,
 312 tension zones can move because one parental genotype is fitter. These moving
 313 tension zones can nonetheless easily be halted by density troughs or barriers to
 314 dispersal owing to their bi-stable nature (Barton and Turelli, 2011). Random drift
 315 is expected to ultimately free such tension zones (Barton 1979, Pialek and Barton
 316 1997) but the trapping can last for very long time. Finally, tension zones can
 317 move haphazardly because of stochastic processes - genetic drift within
 318 populations and stochastic dispersal between populations. Here, we used the

latter two processes to simulate tension zone movement. We used a slight fitness advantage a ($W(AB)=1+a$) of the Mediterranean parental genotype starting from a contact within the Mediterranean Sea (at deme 50), or a slight fitness advantage of the Atlantic parental genotype ($W(ab)=1+a$) starting from a contact on the Atlantic side (at deme 10). Alternatively we introduced random drift by using multinomial sampling of genotypes within each deme at each generation ($N = 500$ per deme). This second way of simulating tension zone movement results in various outcomes from the same starting position, because the movement is not directional. We used this alternative only to show that results are qualitatively similar regardless of the “engine” of the tension zone movement and we did not intend to complete an extensive analysis of the effect of drift on the outcomes, as this is far from the purpose of the present work.

331

332**RESULTS**

333**Analysis of genetic diversity**

334The analysis of 568 *COIII* sequences revealed 150 haplotypes and a high
335nucleotide diversity ($\pi = 0.023$). Average and net nucleotide divergence
336measures between Atlantic and Mediterranean populations were also high ($D_{xy} =$
3370.029, $D_a = 0.011$). The NJ-tree is presented in the supplementary Figure S1. Two
338clades of haplotypes could be defined according to their phylogenetic
339relationships and their frequency in Atlantic and Mediterranean samples (Figure
340S1). The clade nearly fixed in Atlantic samples is known to be due to introgression
341of alleles from the sister species *M. edulis* (Quesada *et al*, 1998). We used
342phylogenetic relationships to define the Mediterranean and the Atlantic
343haplogroups. Nine size-alleles were detected at the new locus *Precol-D* and allele
344frequencies are presented in Table S2. *EF2* was found to be bi-allelic as in Gosset
345and Bierne (2013). The use of capillary electrophoresis allowed us to distinguish
346*EFbis* size-alleles with 1bp differences that were not detected previously with
347acrylamide gels. We observed 22 alleles and the correspondence with previous
348allele names is given in Table S2. Clustering analysis was performed with multi-
349allelic data. In order to represent allele frequency clines we transformed multi-
350allelic loci into bi-allelic loci by pooling alleles according to their frequency in
351reference samples, sample 1 (Faro, Portugal) for the Atlantic lineage and sample
352N9 (Sète, France) for the Mediterranean lineage (see Table S2).

353**Spatial genetic structure**

354The frequencies of the Mediterranean allele at the mitochondrial *COIII* locus and
355the three nuclear markers (*Precol-D*, *EFbis* and *EF2*) in each sample are presented
356in Figure 2A and 2B. Clustering methods did not provide meaningful results with a
357number of genetic clusters parameterized above $K=2$ (no improvement of the
358likelihood in the STRUCTURE analysis, low contribution of secondary axes in
359multivariate analyses, and lack of spatial structure of additional clusters) and we
360provide the results with $K=2$. As reported 20 years ago (Quesada *et al*, 1995b,
361Sanjuan *et al*. 1994) and 10 years ago (Diz and Presa 2008), the northern
362transect along the Spanish coast contained an abrupt change in allele
363frequencies in the zone of transition between Almeria and Cartagena with all four
364loci. The three samples collected in Atlantic waters (1-Faro, N2-Manilva and N3-
365Almeria) have high frequencies of the Atlantic alleles (Figure 2A), uniformly high

366Atlantic cluster membership probabilities in the DAPC analysis (Figure 1, sup file
 367text S1), and a high proportion of Atlantic ancestry in the STRUCTURE analysis
 368(Figure 2E). Conversely, samples from Mediterranean waters (N6-Tabarca, N7-
 369Nules, N8-Peñíscola del Pinatar, and N9-Sète) have high frequencies of
 370Mediterranean alleles (Figure 2B), uniformly high Mediterranean cluster
 371membership probabilities in the DAPC analysis (Figure 1), and a high proportion
 372of Mediterranean ancestry in the STRUCTURE analysis (Figure 2F). Two
 373populations (N4-Cartagena and N5-San Pedro) were obtained from the region
 374reported to be devoid of mussels 20 years ago. These samples were
 375predominantly of Mediterranean ancestry, but some individuals were assigned to
 376the Atlantic genetic cluster with high confidence and a few mussels with a
 377genome of mixed ancestry were also present (Figure 1 and 2E). Departure from
 378Hardy-Weinberg and linkage equilibrium was maximal in these intermediate
 379samples (Figure 2C) as expected from hybrid zone theory (Barton and Gale, 1993)
 380and the departure was significantly different from zero in sample N4-Cartagena.
 381These samples are crucial for our understanding of the functioning of the AOF
 382barrier to gene flow as they show Atlantic mussels are found in observable
 383numbers east of the AOF within the Mediterranean Sea.

384The southern transect along the Algerian coast revealed a complex structure on a
 385much wider geographic scale, consisting of a mosaic pattern of alternation
 386between mostly Atlantic, mostly Mediterranean, and intermediate samples
 387(Figure 1 and 2). Individuals assigned to the Atlantic cluster predominated in the
 388sample from S2-Nador (Morocco) situated in the Alboran Sea (Figure 1 and 2).
 389The Mediterranean cluster predominated in three samples from eastern Algeria
 390(S11-Zaima Mansouriah, S12-Collo and S13-Skikda) and Tunisia (S14-Bizerta)
 391(Figure 1 and 2). From S3-Oran to S10-Bejaia, a mosaic hybrid zone was observed
 392with a tendency for the Mediterranean ancestry to decrease eastward along the
 393zone. The zone ends with an abrupt genetic shift in the Gulf of Bejaia (Gulf of
 394Bejaia Barrier, GBB), west of which are preponderantly Atlantic samples and east
 395of which is populated by Mediterranean mussels. Within the mosaic zone,
 396samples were found to be a mixture of individuals belonging to both clusters and
 397with mixed ancestry (Figure 1 and 2F). Again strong departures from Hardy-
 398Weinberg and linkage equilibrium (Figure 2D) attest to a strong deficit of hybrids
 399relative to random mating. The departure was strongly significant in S3-Oran and
 400S5-Sidi Lakdher and significant in S8-Tipaza and S9-Algiers (Figure 2D). We did
 401not obtain support for the existence of early generation hybrids in Algerian

populations as none of the 183 individuals sampled in the hybrid zone obtained a cumulative posterior probability above 0.8 to be F1, F2 or backcross in NewHybrids (Supplementary Table S3). This is not surprising as early generation hybrids must be extremely rare in nature for the genetic differentiation to be so efficiently maintained between two backgrounds. Evidence for hybridization nonetheless comes from allele frequencies and the comparison of ancestry values which show Atlantic mussels from the hybrid zone have a higher fraction of Mediterranean ancestry and Mediterranean mussels a higher fraction of Atlantic ancestry than mussels from peripheral allopatric populations (Figure 2F). This local introgression of parental backgrounds within the mosaic hybrid zone also explains the lack of power to characterise hybrid genotypes despite the fact that our markers are strongly differentiated between reference samples well outside of the hybrid zone.

Modeling a Y-shape bi-unidirectional stepping stone that accounts for the seascape features of the AOF transition zone

Our simple model in a Y-shape bi-unidimensional stepping stone allowed us to obtain useful indications about how the principal features of the seascape around the AOF area could produce the mosaic structure observed.

Endogenous selection- Let us begin with endogenous selection. It is important to remember that in the tension zone model, selection against hybrids maintains the clinal structure but, in contrast to exogenous selection, the position of the cline is not stabilized by selection. Theory predicts that tension zones should be trapped by and coincide with natural barriers to dispersal, i.e. zones of low population density such as mountains, rivers or oceanic fronts (Barton, 1979). First, we describe the results observed with deterministic simulations at a selected locus in the case of a tension zone moving as a consequence of a fitness advantage of one parental genotype. The objective is not to explore the conditions in which the trapping process occurs, which are well known (Barton and Turelli, 2011; Barton, 1979), but to illustrate how tension zone trapping can occur in an unusual spatial and connectivity context such as the one in the AOF area. Second, we will provide an example with a neutral locus introgressing through the barrier and incorporating genetic drift. Again, we do not treat the effect of linkage and random drift that are known (Barton, 1986; Polechová and

436Barton 2011); our aim is simply to exemplify that the model can retrieve a
437pattern which resembles the true data.

438Figure 3 presents the results of the deterministic model: movement of
439endogenous clines along the northern coast, characterized by a barrier to
440dispersal between demes 20 and 21 (called AOF in Figure 3), and the southern
441coast, characterized by a barrier to dispersal between demes 40 and 41 (called
442GBB in Figure 3). We obtained qualitatively similar results when we added a
443barrier to dispersal between demes 20 and 21 of the southern chain (i.e. Oran)
444and provide the results with no barrier to exemplify the trapping of the cline even
445when the barrier is restricted to the northern coast (i.e. Almeria). The most
446interesting result is a contact within the Mediterranean Sea (between demes 50
447and 51) and a westward propagation of the cline owing to a slight fitness
448advantage of the Mediterranean parental genotype. The clines propagate along
449the southern and northern coasts and a barrier to dispersal is first met by the
450southern cline between demes 40 and 41 (GBB), where it is trapped while the
451northern cline continues to propagate westward (Figure 3A). The northern cline is
452subsequently trapped by the AOF barrier between demes 20 and 21 when it
453comes to overlap with it (blue cline). When migration is symmetrical between the
454north and the south, we end with a pair of shifted clines, one trapped by the AOF
455in the North and one trapped by the GBB in the south (see inset in Figure 3A). In
456this case the Algerian coast is inhabited by the Atlantic lineage. However, when
457migration is asymmetrical from the north to the south, the Mediterranean
458genotype can establish and be maintained in the south, generating a zone of
459coexistence around the AOF in the south (Figure 3A bottom panel). This
460simulation shows that a mosaic hybrid zone can be produced in the South with
461purely intrinsic reproductive isolation when the major hydrodynamic
462characteristics of the region are taken into account. We also considered the
463situation of a contact on the Atlantic side and an eastward propagation of clines
464owing to a slight fitness advantage of the Atlantic parental genotype. The clines
465propagate and meet the AOF barrier between demes 20 and 21 where they are
466trapped (Figure 3B). The southern cline remains trapped at the AOF, even when
467the barrier is simulated only in the northern chain (migration between demes 20
468and 21 is fixed at m_{bar} in the northern chain, at Almeria, and at m in the
469southern chain, at Oran). The southern cline does not have the opportunity to
470meet the barrier to dispersal at the GBB. It is well established that tension zones
471are expected to be trapped by the first minor barrier encountered (Barton 1979,

472Barton and Turelli 2011). Despite this prediction, it is usually not well
473acknowledged that hotspots for hybrid zones (i.e. suture zones) are more likely to
474be localised at the barrier to dispersal that is closest to an area of frequent
475contact rather than at the strongest barrier to dispersal across the overall
476distribution area.

477The deterministic model shows that sister tension zones can be differentially
478trapped at two alternative barriers to larval dispersal, at Almeria in the North and
479Bejaia in the South and that a preponderantly asymmetrical north-south
480migration rate at the AOF can maintain a “pocket” of the Mediterranean
481background east of Oran in the South. In order to model the true data more
482closely, we considered a neutral locus linked to a selected locus, which allowed
483us to account for introgression, and introduced genetic drift. A simulation output
484is presented in Figure 4A, with the two coastlines superposed in the same figure
485and compared to true data in Figure 4B. In this simulation the two parental
486genotypes have the same fitness, which shows that the pocket of Mediterranean
487background in the South is not due to a fitness advantage but to the balance
488between migration from the North and selection against hybrids. A comparison
489with Figure 3 shows that introgression of Mediterranean alleles onto the Atlantic
490background and of Atlantic alleles onto the Mediterranean background is stronger
491in the mosaic zone between the two barriers (AOF and GBB) than in peripheral
492populations. This is expected in mosaic hybrid zones as already explained in
493another mussel hybrid zone (Bierne *et al*, 2003). Small patches of parental
494populations enclosed within a mosaic hybrid zone and flanked by patches of
495populations of the alternative genetic background are expected to introgress
496faster than large and broadly distributed peripheral populations.

497**Exogenous selection-** We now consider exogenous selection. We explored two
498kinds of seascape, a fine-grained mosaic structure in the hybrid zone, or coarse-
499grained environmental variation. In the first model we have a mosaic of habitat 1
500(Atlantic-like) that alternates randomly with habitat 2 (Mediterranean-like) in the
501central area of the southern stepping stone (corresponding to the area between
502the AOF and the GBB in Algeria). This model corresponds to a fine-grained mosaic
503of habitats. As expected, notwithstanding that genotypes adapted to habitat 1
504and genotypes adapted to habitat 2 flow from peripheral populations, this kind of
505selection generates a mosaic structure at the selected locus that correlates with
506habitat variation (Figure 4A). Despite the strong selection coefficient ($s=0.2$) and

507with the recombination rate we used (1 cM) introgression proceeds quickly at the
508neutral locus in the mosaic zone (Figure 4A presents the results at generation
5091000) and the neutral marker loses the association with both the selected locus
510and the habitat. We can of course increase selection and decrease recombination
511to maintain the association for longer periods, but the parameters we used are
512already strong selection and tight linkage. Our objective here was to briefly verify
513the well-established result that micro-geographic adaptation generates a very
514weak barrier to neutral gene flow (Flaxman et al. 2012; Thiebert-Plante and
515Hendry 2015). Secondly, we simulated coarse-grained environmental variation
516(Figure 3B). As for endogenous selection, we observed two clines at the two
517environmental boundaries we simulated, the AOF in the North and the GBB in the
518south, when north-south migration was symmetrical. In order to obtain the
519maintenance of genotypes adapted to habitat 2 (i.e. Mediterranean-like) in the
520south, asymmetrical north-south migration at the AOF was required. The result
521obtained is very similar to that obtained with endogenous selection. Introgression
522proceeds much more slowly than in the fine-grained habitat model, although
523faster than with endogenous selection (with the same recombination rate of 1
524cM). Another difference between the exogenous and endogenous models is that
525under exogenous selection, the habitat 2 genetic background (Mediterranean-
526like) cannot be maintained without migration from the north, because it is
527counter-selected in the south. In the endogenous model however, we can
528imagine that a sufficiently big patch of Mediterranean background can persist in
529the south, at least transiently, as soon as it has been established by an episode
530of north-south migration.

531

532

533DISCUSSION

534The Almeria-Oran Front (AOF) is arguably one of the most famous hotspots of
535genetic differentiation in the sea (Patarnello *et al.*, 2007), and *Mytilus*
536*galloprovincialis* one of the iconic species exhibiting a pronounced genetic break
537at the AOF (Quesada *et al.*, 1995a,b, Sanjuan *et al.* 1994). We already had indirect
538evidence that reproductive isolation might have evolved between Mediterranean
539and Atlantic refugia (i) in order to maintain genetic differentiation upon contact
540(Quesada *et al.*, 1995a), (ii) in order to maintain a sharp genetic break (Quesada
541*et al.*, 1995b), and (iii) in order to explain the semi-permeability of the barrier to
542gene flow (Fraisse *et al.* 2016). However, the story could not be definitively
543settled without more direct evidence that interbreeding opportunities exist at the
544transition zone, even if such opportunities are rare. Maybe due to this lack of
545evidence of reproductive isolation, the idea that the Atlantic-Mediterranean divide
546could simply be explained either by the AOF itself acting as a barrier to larval
547dispersal or by the ecological differences between Atlantic and Mediterranean
548waters remains strongly anchored in the marine literature. In addition, knowing
549Atlantic waters enter the Mediterranean Sea through the South by the Algerian
550current, the description of the southern side of the transition zone was a
551sampling lacuna that needed to be filled. Sampling that coast revealed far more
552than one could have expected. The unexplored side of the biogeographic
553boundary was hiding a 600 km wide mosaic hybrid zone. We will argue here that
554the existence of this large mosaic hybrid zone provides evidence-based support
555for the existence of partial reproductive isolation, which prevents gene flow when
556the two lineages live in sympatry/syntopy with the opportunity to interbreed and
557exchange genes in the absence of such reproductive isolation. We believe that
558the existence of partial reproductive isolation between heterogeneously-
559differentiated genetic backgrounds should supplement the disruption of
560population connectivity and selection against maladaptive migrants to provide a
561fuller explanation of genetic subdivision in marine (Bierne *et al.*, 2011, Gagnaire
562*et al.* 2015, Riginos *et al.* 2016) as well as terrestrial species (Barton and Hewitt
5631985, Hewitt 1989).

564The previously-reported absence of mussels between Almeria and Alicante
565(Quesada *et al.*, 1995b) on the Spanish coast suggested that there was no strong
566opportunity for sympatry and interbreeding between the Atlantic and the
567Mediterranean lineages, in accordance with the AOF as an efficient barrier to

568dispersal. However, two populations were sampled south of Alicante in Cartagena
569and San Pedro (N4 and N5 in Figure 1) with a low but non-negligible proportion of
570these individuals assigned to the Atlantic genetic cluster. This is a sufficient
571proportion of migrant mussels to refute the barrier to dispersal as strong enough
572to explain the maintenance of the genetic structure. Along with the results from
573the southern coast presented here, there remains no ambiguity that the Atlantic
574lineage not only enters the Mediterranean Sea, but co-exists with the
575Mediterranean lineage, sometimes in balanced proportions, across a widespread
576area where mussels are numerous. Described as a strong barrier to larval
577dispersal, the AOF is in fact rather a mix of Atlantic waters and Alboran surface
578waters called 'Modified Atlantic water' (MAW, Font *et al*, 1998) in the Eastern
579Alboran anticyclone and along the Algerian coast thanks to the Algerian current.
580There are several species without a genetic break at the AOF -e.g. the European
581flat oyster *Ostrea edulis* (Launey *et al*, 2002), the Mediterranean rainbow wrasse
582*Coris julis* (Aurelle *et al*, 2003), or the saddled seabream *Oblada melanura*
583(Galarza *et al*, 2009). These species share similar biological and ecological
584characteristics with species that do exhibit a break and this observation was
585already a strong argument against the hypothesis of a sufficiently reduced
586dispersal rate at the AOF to explain the genetic differentiation observed in many
587species.

588We now turn to the more complex issue of the nature of the genetic barrier
589(extrinsic, intrinsic, or both). Given the contrasting environments inhabited by the
590two lineages in the Mediterranean Sea and the Atlantic Ocean, there is little
591doubt that differential adaptation must exist between the two lineages (including
592phenotypic plasticity and epigenetic responses). However, the large zone of co-
593existence in Algeria modifies our view of how the genetic barrier operates. We no
594longer have to explain adaptation in the two seas but have to explain that the
595two lineages are maintained genetically cohesive in sympatry along 600 km of
596Algerian coasts. Without being able to refute exogenous selection, we argue here
597that intrinsic pre- or post-zygotic isolation is equally likely, if not more likely. We
598have shown with simple simulations that the mosaic genetic structure observed
599in Algeria can indeed be obtained with intrinsic reproductive isolation alone
600(Figure 3 and 4). It only requires asymmetric north-south dispersal at the AOF,
601which seems a reasonable assumption. In addition, simulations of exogenous
602selection to a coarse-grained environment also require the assumption of
603asymmetric north-south dispersal at the AOF (Figure 5B), or a patch of

maladapted Mediterranean mussels cannot be maintained in the south. Adaptation to a patchy fine-grained environment might better explain the mosaic structure but does not generate a strong barrier to gene flow if neutral markers are not tightly linked to adaptive loci. We probably have not sampled the genome sufficiently densely during our previous genome scans to argue for strong physical linkage between our markers and local adaption genes (Hoban *et al.* 2016), and we therefore need a process able to generate a barrier to gene flow on a large proportion of the genome (Bierne *et al.* 2011). Selection against migrants that are maladapted to the local environment is often proposed as a process able to generate a genome-wide barrier to gene flow (Marshall *et al.*, 2010; Nosil *et al.*, 2005). However, it is not well appreciated that local selection needs to be extremely strong to generate a sufficiently effective barrier to gene flow (Barton and Bengtsson, 1986; Feder and Nosil, 2010; Slatkin, 1973), producing a strong segregation of each ecotype in its own favoured habitat. As soon as lineages coexist in balanced proportions somewhere, as observed here in Algeria, some sort of selection against hybrids must exist. It can still be exogenous selection (Kruuk *et al.* 1999), but hybrids must perform poorly in both habitats. Hybrids are the bridges used by neutral genes to flow across the barrier and they either need to be produced in low proportion or to be all unfit for a barrier to gene flow to have broad genomic effects (Barton and Bengtsson 1986). Here, in the Algerian hybrid zone, this is indeed the case as we observed a strong deficit of intermediate genotypes and strong departures from Hardy-Weinberg and linkage equilibrium among physically unlinked loci (Figure 2). Finally, the coarse-grained environment we simulated is not very likely. Although Atlantic waters enter the Algerian current, Algerian waters are much more similar to other Mediterranean waters than to Atlantic waters. Fine-grained habitat heterogeneity was also far from evident between our sampling sites in Algeria as we sampled very similar microsites: high-shore rocks protected from wave action by artificial port breakwaters and with no evident influence of freshwater from nearby estuaries. Taken altogether, we believe these arguments suggest intrinsic reproductive isolation, maybe pre-zygotic given the apparent paucity of F1 hybrids, is a parsimonious evidence-based hypothesis that should be considered.

Overall our results provide additional support for the coupling hypothesis, which suggests that genetic breaks are often secondary contact semi-permeable tension zones between heterogeneously differentiated genomes trapped by physical barriers to dispersal or environmental boundaries (Bierne *et al.*, 2011).

Physical and environmental factors mostly explain the position of the genetic breaks but the maintenance of genome-wide genetic differentiation is best explained by reproductive isolation. Our model implies the differential coupling of sister tension zones at two different geographic positions, at the AOF in the North and at the GBB in the south. The abrupt genetic shift at the GBB is a new observation in an understudied area. However, it fits well with the position of a barrier to dispersal identified with oceanographic modelling in three recent analyses (Andrello *et al*, 2015; Berline *et al*, 2014; Rossi *et al*, 2014). Note in addition that these three hydrodynamic studies identified much more efficient barriers to larval dispersal than the AOF within the Mediterranean Sea that are not hotspots of genetic differentiation. Differential coupling of sister tension zones can occur in coastal or river species living in linear landscapes with tree-like reticulation that are not well-described by standard 1D or 2D stepping stone models (Fourcade *et al*. 2013). For instance, some hybrid zones secondarily trapped at the entrance of the Baltic Sea are likely to have produced sister hybrid zones which have moved further north along the Norwegian coast to be trapped somewhere in northern Scandinavia (e.g. in *Mytilus trossulus*, *Macoma balthica*, *Gadus morhua*, *Platichthys flesus* and *Gammarus zaddachi*; reviewed in Bierne *et al*, 2011). Recently the Baltic - North Sea contact zone between the strongly divergent mussel species *Mytilus trossulus* and *M. edulis* has been investigated in the southern coast area and was identified to be positioned well inside the Baltic Sea in Germany (Stuckas *et al*. 2017) while it is localised in the Oresund in the North (Väinölä and Hvilsom 1991).

To conclude, the mosaic hybrid zone between Oran and Bejaia observed in *M. galloprovincialis* mussels contributes to definitively refute the hypothesis that the AOF itself generates a sufficiently strong barrier to dispersal to maintain genetic differentiation in this species. It also opens the debate about the nature of the genetic barrier, which is not necessarily related to differential adaptation to Atlantic and Mediterranean waters but could imply intrinsic pre- or post-zygotic reproductive isolation. The next step will be to conduct lab experiments (hybrid crosses) and field studies (settlement and reproduction) to put the alternatives to the test. Our results also call for new genetic studies along Algerian coasts in other marine species. A few samples of sea bass from eastern Algeria proved to be admixed between the Atlantic and Mediterranean lineages (Duraton *et al*. 2018), although we do not yet know the spatial variation of admixture proportions in this system. Finally our study reveals how complex the interplay

676between connectivity (hydrography), environmental variation (seascape) and
677reproductive isolation (including both extrinsic and intrinsic mechanisms) can be.
678Without minimising the importance of hydrography and local selection, we
679believe the excessive emphasis on oceanographic features in the interpretation
680of the genetic structure in the sea sometimes results in an incomplete
681understanding of the underlying processes explaining its origin and maintenance.

682

683**Acknowledgments**

684It is a pleasure to thank Marie-Thérèse Augé and Cathy Haag-Liautard for their
685technical help in the lab, and Marina Albentosa as well as Carlos Saavedra for
686Spanish samples. We also thank Pierre-Alexandre Gagnaire, Christelle Fraïsse,
687François Bonhomme, Nathalie Charbonnel, Sophie Arnaud-Haond and Lilia Bahri-
688Sfar for discussions, as well as Cynthia Riginos, Heiko Stuckas and 6 anonymous
689referees for their insightful comments on earlier drafts. Molecular data were
690produced through the ISEM platform Génomique des Populations Marines
691(PGPM7) at the Station Marine de Sète (OSU OREME, Observatoire de Recherche
692Méditerranéen de l'Environnement) and the platform Génomique
693Environnementale of the LabEx CeMEB (Laboratoire d'Excellence Centre
694Méditerranéen de l'Environnement et de la Biodiversité). This work was funded by
695a Languedoc-Roussillon "Chercheur(se)s d'Avenir" grant (Connect7 project). This
696is article 2018-286 of Institut des Sciences de l'Evolution de Montpellier.

697

698**Conflict of Interest**

699The authors declare no conflict of interest.

700

701**Data Archiving**

702Sequence data have deposited in GenBank: accession numbers MK259098 -
703MK259664

704Genotype data have been deposited in Dryad: doi:10.5061/dryad.m3t6p5g

705

706References

- 707Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N *et al* (2013).
708Hybridization and speciation. *Journal of Evolutionary Biology* **26**: 229-246.
- 709Anderson E, Thompson E (2002). A model-based method for identifying species
710hybrids using multilocus genetic data. *Genetics* **160**: 1217-1229.
- 711Andrello M, Jacobi MN, Manel S, Thuiller W, Mouillot D (2015). Extending networks
712of protected areas to optimize connectivity and population growth rate.
713*Ecography* **38**: 273-282.
- 714Arnold (2006). Evolution through genetic exchange. *Oxford, UK: Oxford*
715*University Press*.
- 716Aurelle D, Guillemaud T, Afonso P, Morato T, Wirtz P, Santos RS *et al* (2003).
717Genetic study of *Coris julis* (Osteichthyes, Perciformes, Labridae) evolutionary
718history and dispersal abilities. *Comptes Rendus Biologies* **326**: 771-785.
- 719Avice JC (2000). *Phylogeography: The history and formation of species*. Harvard
720University Press: Cambridge, MA.
- 721Barton N, Turelli M (2011). Spatial waves of advance with bistable dynamics:
722cytoplasmic and genetic analogues of Allee effects. *The American Naturalist* **178**:
723E48-E75.
- 724Barton NH (1979). The dynamics of hybrid zone. *Heredity* **43**: 341-359.
- 725Barton NH (1986). The effects of linkage and density-dependant regulation on
726gene flow. *Heredity* **57**: 415-426.
- 727Barton NH, Bengtsson BO (1986). The barrier to genetic exchange between
728hybridising populations. *Heredity* **56**: 357-376.
- 729Barton NH, Gale KS (1993). Genetic analysis of hybrid zones. In: HARRISON RG
730(ed) *Hybrid zones and the evolutionary process*. Oxford University Press: New
731York, Oxford, pp 13-45.
- 732Barton NH, Hewitt GM (1985). Analysis of hybrid zones. *Annual Review of Ecology*
733*and Systematics* **16**: 113-148.
- 734Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. (2002). GENETIX 4.05,
735logiciel sous Windows TM pour la génétique des populations. Université de
736Montpellier 2: Montpellier, France.
- 737Berline L, Rammou A-M, Doglioli A, Molcard A, Petrenko A (2014). A Connectivity-
738Based Eco-Regionalization Method of the Mediterranean Sea. *PLOS ONE* **9**:
739e111978.
- 740Bialozyt R, Ziegenhagen B, Petit RJ (2006). Contrasting effects of long distance
741seed dispersal on genetic diversity during range expansion. *Journal of*
742*Evolutionary Biology* **19**: 12-20.
- 743Bierne N, Bonhomme F, Boudry P, Szulkin M, David P (2006). Fitness landscapes
744support the dominance theory of post-zygotic isolation in the mussels *Mytilus*
745*edulis* and *M. galloprovincialis*. *Proceedings of the Royal Society B-Biological*
746*Sciences* **273**: 1253-1260.

747Bierne N, Borsa P, Daguin C, Jollivet D, Viard F, Bonhomme F *et al* (2003).
748Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and M-
749galloprovincialis. *Molecular Ecology* **12**: 447-461.

750Bierne N, David P, Boudry P, Bonhomme F (2002). Assortative fertilization and
751selection at larval stage in the mussels *Mytilus edulis* and M-galloprovincialis.
752*Evolution* **56**: 292-298.

753Bierne N, Gagnaire P-A, David P (2013). The geography of introgression in a
754patchy environment and the thorn in the side of ecological speciation. *Current*
755*Zoology* **59**: 72-86.

756Bierne N, Welch J, Loire E, Bonhomme F, David P (2011). The coupling hypothesis:
757why genome scans may fail to map local adaptation genes. *Molecular Ecology*
758**20**: 2044-2072.

759Borrero-Pérez GH, González-Wangüemert M, Marcos C, Pérez-Ruzafa A (2011).
760Phylogeography of the Atlanto-Mediterranean sea cucumber *Holothuria*
761(*Holothuria*) *mammata*: the combined effects of historical processes and current
762oceanographical pattern. *Molecular Ecology* **20**: 1964-1975.

763Borsa P, Naciri M, Bahri L, Chikhi L, Garcia de Leon FJ, Kotoulas G *et al* (1997).
764Zoogéographie infra-spécifique de la mer Méditerranée. *Vie Milieu* **47**: 295-305.

765Charlesworth B, Charlesworth D, Barton NH (2003). The effects of genetic and
766geographic structure on neutral variation. *Annual Review of Ecology and*
767*Systematics* **34**: 99-125.

768Chevolot M, Hoarau G, Rijnsdorp AD, Stam WT, Olsen JL (2006). Phylogeography
769and population structure of thornback rays (*Raja clavata* L., Rajidae). *Molecular*
770*Ecology* **15**: 3693-3705.

771Cordero D, Peña JB, Saavedra C (2014). Phylogeographic analysis of introns and
772mitochondrial DNA in the clam *Ruditapes decussatus* uncovers the effects of
773Pleistocene glaciations and endogenous barriers to gene flow. *Molecular*
774*Phylogenetics and Evolution* **71**: 274-287.

775Cruickshank TE, Hahn MW (2014). Reanalysis suggests that genomic islands of
776speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*
777**23**(13): 3133-3157.

778Diz AP, Presa P (2008). Regional patterns of microsatellite variation in *Mytilus*
779galloprovincialis from the Iberian Peninsula. *Marine Biology* **154**: 277-286.

780Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using
781multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*
782**164**: 1567-1587.

783Feder JL, Nosil P (2010). The efficacy of divergence hitchhiking in generating
784genomic islands during ecological speciation. *Evolution* **64**: 1729-1747.

785Flaxman, S. M., Feder, J. L., & Nosil, P. (2012). Spatially explicit models of
786divergence and genome hitchhiking. *Journal of evolutionary biology*, **25**: 2633-
7872650.

788Fourcade, Y., Chaput-Bardy, A., Secondi, J., Fleurant, C., & Lemaire, C. (2013). Is
789local selection so widespread in river organisms? Fractal geometry of river

790networks leads to high bias in outlier detection. *Molecular Ecology*, **22**: 2065-
7912073.

792Font J, Millot C, Salas J, Julià A, Chic O (1998). The drift of Modified Atlantic Water
793from the Alboran Sea to the eastern Mediterranean. *Scientia Marina* 62: 211-216.

794Fraïsse C, Belkhir K, Welch JJ, Bierne N (2016). Local interspecies introgression is
795the main cause of extreme levels of intraspecific differentiation in mussels.
796*Molecular Ecology* **25**: 269-286.

797Gagnaire PA, Broquet T, Aurelle D, Viard F, Souissi A, Bonhomme F *et al* (2015).
798Using neutral, selected and hitchhiker loci to assess connectivity of marine
799populations in the genomic era. *Evolutionary Applications* **8**: 769-786..

800Galarza JA, Carreras-Carbonell J, Macpherson E, Pascual M, Roques S, Turner GF *et*
801*al* (2009). The influence of oceanographic fronts and early-life-history traits on
802connectivity among littoral fish species. *Proceedings of the National Academy of*
803*Sciences* **106**: 1473-1478.

804Gilg, M. R., & Hilbish, T. J. (2003). The geography of marine larval dispersal:
805Coupling genetics with fine-scale physical oceanography. *Ecology*, **84**: 2989-
8062998.

807Goldberg EE, Lande R (2007). Species' borders and dispersal barriers. *The*
808*American Naturalist* **170**: 297-304.

809Gosset C, Bierne N (2013). Differential introgression from a sister species explains
810high *F_{st}* outlier loci within a mussel species. *Journal of Evolutionary Biology* **26**:
81114-26.

812Harrison RG (1993). *Hybrid zones and the evolutionary process*. Oxford University
813Press: Oxford.

814Harrison RG, Larson EL (2016). Heterogeneous genome divergence, differential
815introgression, and the origin and structure of hybrid zones. *Molecular Ecology*: in
816press.

817Hellberg ME (2009). Gene flow and isolation among populations of marine
818animals. *Annual Review of Ecology, Evolution, and Systematics* **40**: 291-310.

819Hellberg ME, Burton RS, Neigel JE, Palumbi SR (2002). Genetic assesment of
820connectivity among marine populations. *Bulletin of Marine Science* **70**: 273-290.

821Hewitt GM (1975). A sex-chromosome hybrid zone in the grasshopper *Podisma*
822*pedestris* (Orthoptera: Acrididae). *Heredity* **35**: 375-387.

823Hewitt GM (1996). Some genetic consequences of ice ages, and their role in
824divergence and speciation. . *Biological Journal of the Linnean Society* **58**: 247-
825276.

826Hewitt GM (2000). The genetic legacy of the quaternary ice ages. *Nature* **405**:
827907-913.

828Ibrahim KM, Nichols RA, Hewitt GM (1996). patial patterns of genetic variation
829generated by different forms of dispersal during range expansion. *Heredity* **77**:
830282-291.

831Johannesson K, Andre C (2006). Life on the margin: genetic isolation and diversity
832loss in a peripheral marine ecosystem, the Baltic Sea. *Molecular Ecology* **15**:
8332013-2029.

834Johannesson K, Panova M, Kemppainen P, Andre C, Rolan-Alvarez E, Butlin RK
835(2010). Repeated evolution of reproductive isolation in a marine snail: unveiling
836mechanisms of speciation. *Philosophical Transactions of the Royal Society B*:
837*Biological Sciences* **365**: 1735-1747.

838Jombart T, Ahmed I (2011). adegenet 1.3-1: new tools for the analysis of genome-
839wide SNP data. *Bioinformatics* **27**: 3070-3071.

840Launey S, Ledu C, Boudry P, Bonhomme F, Naciri-Graven Y (2002). Geographic
841structure in the European flat oyster, *Ostrea edulis* L, as revealed by
842microsatellite polymorphism. *Journal of Heredity* **93**: 331-351.

843Le Corre V, Machon N, Petit RJ, Kremer A (1997). Colonization with long-distance
844seed dispersal and genetic structure of maternally inherited genes in forest trees:
845a simulation study. *Genetical Research* **69**: 117-125.

846Lemaire C, Versini JJ, Bonhomme F (2005). Maintenance of genetic differentiation
847across a transition zone in the sea: discordance between nuclear and cytoplasmic
848markers. *Journal of Evolutionary Biology* **18**: 70-80.

849Librado P, Rozas J (2009). DnaSP v5: a software for comprehensive analysis of
850DNA polymorphism data. *Bioinformatics* **25**: 1451-1452.

851Lourenço, C. R., Nicastro, K. R., Serrão, E. A., Castilho, R., & Zardi, G. I. (2015).
852Behind the mask: cryptic genetic diversity of *Mytilus galloprovincialis* along
853southern European and northern African shores. *Journal of Molluscan Studies*,
85481(3), 380-387.

855M'Gonigle LK, FitzJohn RG (2010). Assortative mating and spatial structure in
856hybrid zones. *Evolution* **64**: 444-455.

857Manel S, Holderegger R (2013). Ten years of landscape genetics. *Trends in*
858*Ecology and Evolution* **28**: 614-621.

859Marshall DJ, Monro K, Bode M, Keough MJ, Swearer S (2010). Phenotype-
860environment mismatches reduce connectivity in the sea. *Ecology Letters* **13**: 128-
861140.

862McQuaid, C. D., & Phillips, T. E. (2000). Limited wind-driven dispersal of intertidal
863mussel larvae: in situ evidence from the plankton and the spread of the invasive
864species *Mytilus galloprovincialis* in South Africa. *Marine Ecology Progress Series*,
865201: 211-220.

866Meirmans PG (2015). Seven common mistakes in population genetics and how to
867avoid them. *Molecular Ecology* **24**: 3223-3231.

868Nichols RA, Hewitt GM (1994). The genetic consequences of long distance
869dispersal during colonization. *Heredity* **72**: 312-317.

870Nosil P, Vines TH, Funk DJ (2005). Reproductive isolation caused by natural
871selection against immigrants from divergent habitats. *Evolution* **59**: 705-719.

872Ouagajjou, Y., & Presa, P. (2015). The connectivity of *Mytilus galloprovincialis* in
873northern Morocco: a gene flow crossroads between continents. *Estuarine, Coastal*
874*and Shelf Science*, 152, 1-10.

875Palero F, Abelló P, Macpherson E, Gristina M, Pascual M (2008). Phylogeography
876of the European spiny lobster (*Palinurus elephas*): influence of current
877oceanographical features and historical processes. *Molecular Phylogenetics and*
878*Evolution* **48**: 708-717.

879Palumbi SR (1994). Genetic divergence, reproductive isolation, and marine
880speciation. *Annual Review of Ecology and Systematics* **25**: 547-572.

881Patarnello T, Volckaert FA, Castilho R (2007). Pillars of Hercules: is the Atlantic-
882Mediterranean transition a phylogeographical break? *Molecular Ecology* 16: 4426-
8834444.

884Pelc RA, Warner RR, Gaines SD (2009). Geographical patterns of genetic structure
885in marine species with contrasting life histories. *Journal of Biogeography* **36**:
8861881-1890.

887Piálek J, Barton NH (1997). The spread of an advantageous allele across a barrier:
888the effects of random drift and selection against heterozygotes. *Genetics* **145**:
889493-504.

890Polechová J, Barton N (2011). Genetic drift widens the expected cline but narrows
891the expected cline width. *Genetics* **189**: 227-235.

892Qin XX, Coyne KJ, Waite JH (1997). Tough Tendons Mussel Byssus Has Collagen
893with Silk-Like Domains. *The Journal of Biological Chemistry* **51**: 32623-32627.

894Quesada H, Beynon CM, Skibinski DOF (1995a). A mitochondrial DNA
895discontinuity in the mussel *Mytilus galloprovincialis* : pleistocene vicariance
896biogeography and secondary intergradation. *Molecular Biology and Evolution* **12**:
897521-524.

898Quesada H, Warren M, Skibinski DOF (1998). Nonneutral evolution and differential
899mutation rate of gender-associated mitochondrial DNA lineages in the marine
900mussel *Mytilus*. *Genetics* **149**: 1511-1526.

901Quesada H, Zapata C, Alvarez G (1995b). A multilocus allozyme discontinuity in
902the mussel *Mytilus galloprovincialis*: the interaction of ecological and life-history
903factors. *Marine Ecology Progress Series* **116**: 99-115.

904Rand DM, Harrison RG (1989). Ecological genetics of a mosaic hybrid zone:
905mitochondrial , nuclear, and reproductive differentiation of crickets by soil type.
906*Evolution* **43**: 432-449.

907Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., ... &
908Westram, A. M. (2017). Interpreting the genomic landscape of speciation: a road
909map for finding barriers to gene flow. *Journal of evolutionary biology* **30**: 1450-
9101477.

911Remington CL (1968). Suture-zones of hybrid interaction between recently joined
912biotas. In: Dobzhansky T, Hecht MK and Steere WC (eds) *Evolutionary Biology*.
913Plenum Press: New York. Vol. 2, pp 321-428.

914Riginos C, Douglas KE, Jin Y, Shanahan DF, Trembl EA (2011). Effects of geography
915and life history traits on genetic differentiation in benthic marine fishes.
916*Ecography* **34**: 566-575.

917Riginos C, Liggins L (2013). Seascape genetics: populations, individuals, and
918genes marooned and adrift. *Geography Compass* **7**: 197-216.

919Rocha L, Bowen B (2008). Speciation in coral-reef fishes. *Journal of Fish Biology*
920**72**: 1101-1121.

921Rossi V, Ser-Giacomi E, López C, Hernández-García E (2014). Hydrodynamic
922provinces and oceanic connectivity from a transport network help designing
923marine reserves. *Geophysical Research Letters* **41**: 2883-2891.

924Sanjuan, A., Zapata, C., & Alvarez, G. (1994). *Mytilus galloprovincialis* and *M.*
925*edulis* on the coasts of the Iberian Peninsula. *Marine Ecology Progress Series*,
926**131**: 131-146.

927Schunter, C., Carreras-Carbonell, J., Macpherson, E., Tintoré, J., Vidal-Vijande, E.,
928Pascual, A., ... & Pascual, M. (2011). Matching genetics with oceanography:
929directional gene flow in a Mediterranean fish species. *Molecular Ecology*, **20**:
930**5167-5181**.

931Stanley, R. R., DiBacco, C., Lowen, B., Beiko, R. G., Jeffery, N. W., Van
932Wyngaarden, M., ... & Johnson, C. (2018). A climate-associated multispecies
933cryptic cline in the northwest Atlantic. *Science advances*, *4*(3), eaaq0929.

934Slatkin (1987). Gene flow and the geographic structure of natural populations.
935*Science* **236**: 787-792.

936Slatkin M (1973). Gene flow and selection in a cline. *Genetics* **75**: 733-756.

937Stuckas, H., Knöbel, L., Schade, H., Breusing, C., Hinrichsen, H. H., Bartel, M., ... &
938Melzner, F. (2017). Combining hydrodynamic modelling with genetics: can passive
939larval drift shape the genetic structure of Baltic *Mytilus* populations?. *Molecular*
940*Ecology*, **26**: 2765-2782.

941Swenson NG, Howard DJ (2005). Clustering of Contact
942Zones, Hybrid Zones, and Phylogeographic Breaks in North America. *The*
943*American Naturalist* **166**: 581-591.

944Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S (2013) MEGA6: Molecular
945Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**:
946**2725-2729**.

947Teske PR, Papadopoulos I, Barker NP, McQuaid CD, Beheregaray LB (2014).
948Mitonuclear discordance in genetic structure across the Atlantic/Indian Ocean
949biogeographical transition zone. *Journal of Biogeography* **41**: 293-401.

950Thibert-Plante X, Hendry A 2010. When can ecological speciation be detected
951with neutral loci? *Molecular Ecology* **19**: 2301-2314.

952Tine M, Kuhl H, Gagnaire P-A, Louro B, Desmarais E, Martins RST *et al* (2014). The
953European seabass genome and its variation provide insights into adaptation to
954euryhalinity and speciation. *Nature Communications* **5**.

955Väinölä, R., & Hvilsum, M. M. (1991). Genetic divergence and a hybrid zone
956between Baltic and North Sea *Mytilus* populations (Mytilidae: Mollusca). *Biological*
957*Journal of the Linnean Society* **43**: 127-148.

957Vines TH, Dalziel AC, Albert AY, Veen T, Schulte PM, Schluter D (2016). Cline
958coupling and uncoupling in a stickleback hybrid zone. *Evolution* **70**: 1023-1038.

959Viúdez Á, Tintoré J (1995). Time and space variability in the eastern Alboran Sea
960from March to May 1990. *Journal of Geophysical Research: Oceans* **100**: 8571-
9618586.

962Zouros E, Oberhauser Ball A, Saavedra C, Freeman KR (1994). A unusual type of
963mtDNA inheritance in the blue mussel *Mytilus*. *Proceedings of the National*
964*Academy of Science of the United States of America* **91**: 7463-7467.

965

966**Titles and legends to figures**

967Table 1. Examples of marine hotspots of genetic differentiation.

968

969Figure 1. Sampling localities of *Mytilus galloprovincialis* in the north-eastern
970Atlantic and the Mediterranean Sea. Proportion of individuals assigned to the
971Atlantic cluster (in red), the Mediterranean cluster (in green) and unassigned (in
972black), based on a DAPC analysis with the four loci (*COIII*, *Precol-D*, *EFbis* and
973*EF2*). Sample names and their GPS positions are given in the supplementary
974Table S1.

975

976Figure 2. *Mytilus galloprovincialis* Mediterranean allele frequencies at the four
977semi-diagnostic loci analysed (*COIII*, *Precol-D*, *EFbis* and *EF2*) for (A) the 9
978samples of the northern coastline and (B) the 14 samples of the southern
979coastline. AOF: Almeria Oran Front, GBB: Gulf of Bejaïa Barrier. (C) Average
980multilocus linkage disequilibrium (D , bold line, left axis) and Hardy-Weinberg
981disequilibrium (F_{is} , dashed line, right axis) in samples of the northern coastline
982and (D) samples of the southern coastline. Bar plot of the estimated ancestry
983proportions (Q -values) estimated by STRUCTURE with the three nuclear markers
984(*Precol-D*, *EFbis* and *EF2*) for (E) the samples of the northern coastline and (F) the
985southern coastline.

986

987Figure 3. (A) Simulation output obtained with a bi-locus deterministic model with
988selection against recombinant genotypes ($s=0.1$) and a slight advantage of the
989Mediterranean background ($a=0.01$) in a 60 demes Y-shape stepping stone model
990($m=0.3$), including two weak barriers to gene flow ($m_{bar}=0.1$) between demes
99120 and 21 (AOF) in the northern chain and between demes 40 and 41 (GBB) in

the southern chain of demes, a contact between demes 50 and 51, and unidirectional migration right to the AOF. Above: northern chain, below: southern chain. The insert shows an output when migration was bidirectional between the two chains. Clines are represented every 500 generations with a rainbow colour code from orange to dark blue. Blue clines superimposed green clines in the southern chain because they remain at the same position (trapping). (B) Simulation with a contact between demes 10 and 11 and an advantage of the Atlantic background ($a=0.01$). Clines are represented every 500 generations with a rainbow colour code from orange to dark blue. Dark blue clines superimposed light blue and green clines because they remain at the same position (trapping).

1002

Figure 4. (A) A simulation output obtained at a neutral locus at a genetic map length of 1cM to one intrinsic incompatibility after 1000 generations, with the Y-shape bi-unidimensional stepping stone model, including two weak barriers to gene flow between demes 25 and 26 (AOF) in the northern chain and between demes 40 and 41 (GBB) in the southern chain of demes and unidirectional migration right to the AOF. (B) Proportion of Mediterranean ancestry under an admixture model (STRUCTURE Q-values). Geographic positions are rescaled so that Almeria superposes to Oran.

1011

Figure 5. (A) Simulation output obtained with a deterministic model with exogenous selection in a fine-grained environment between AOF and GBB in the southern coasts (in which habitat type was assigned randomly). Other parameters the same as previous figure. Thick line: selected locus, thin line: neutral locus after 1000 generations ($r=1cM$) (B) Simulation output obtained with

1017a deterministic model with exogenous selection in a coarse-grained environment.

1018Thick line: selected locus, thin line: neutral locus after 1000 generations ($r=1\text{cM}$).

1019

Barrier to gene flow	Potential processes involved Vicariance (V) Temperature (T) Salinity (S) Hydrography (H)	References
Øresund and Danish Belts	S	Johannesson and Andre (2006)
Almeria Oran Front	V, H, S, T	Paternello et al (2007), Galarza et al (2009)
Balearic Front	H	Galarza et al (2009)
Siculo - Tunisian Strait	V, H, S, T	Bahri Sfar et al (2000), Mejri et al (2009)
Cape Cod (Massachusetts)	V, H	McGovern & Hellberg (2003), Kelly et al. (2006), (Haney et al. 2007)
Cape Canaveral (Florida)	V, H, T	Reeb and Avise (1990), Avise et al (1992), Pelc et al. (2009)
Nova Scotia 44-45°N	T	Stanley et al. (2018)
Cape Mendocino (California)	C, H	Blanchette et al. 2008, Sivasundar & Palumbi (2010)
Point Conception (California)	V, H	Burton (1998), Cassone & Boulding (2006), Dawson (2001), Pelc et al. (2009).
Cape Agulhas (South Africa)	H	Evans et al. 2004, Teske et al (2014)
30°S Chilean biogeographic break	H	Brante et al (2012)

1020

1021

1022











