



HAL
open science

Spatial gradients of introgressed ancestry reveal cryptic connectivity patterns in a high gene flow marine fish

Tony Robinet, Valérie Roussel, Karine Cheze, Pierre-Alexandre Gagnaire

► To cite this version:

Tony Robinet, Valérie Roussel, Karine Cheze, Pierre-Alexandre Gagnaire. Spatial gradients of introgressed ancestry reveal cryptic connectivity patterns in a high gene flow marine fish. *Molecular Ecology*, 2020, 10.1111/mec.15611 . hal-02925509

HAL Id: hal-02925509

<https://hal.science/hal-02925509>

Submitted on 29 Aug 2020

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.

[Introgression reveals cryptic connectivity in sea bass]

1 **Spatial gradients of introgressed ancestry reveal cryptic**
2 **connectivity patterns in a high gene flow marine fish**

3

4 Tony Robinet (1)*, Valérie Roussel (2), Karine Cheze (1) and Pierre-Alexandre
5 Gagnaire (3)

6

7

8 **ABSTRACT**

9

10 Assessing genetic connectivity among populations in high gene flow species is
11 sometimes insufficient to evaluate demographic connectivity. Genetic differentiation
12 quickly becomes zero as soon as a few dozen migrants are exchanged per generation.
13 This provides little information to determine whether migration can ensure
14 demographic coupling. The resulting difficulties in delineating conservation units for
15 the management of commercially exploited marine fish species are well illustrated in
16 the case of the European sea bass (*Dicentrarchus labrax*). Previous attempts to assess
17 connectivity patterns in the northeast Atlantic have been hampered by a lack of
18 spatial genetic structure. In contrast, mark-recapture data suggested low migration
19 rates between regional spawning areas. Here, we show how a spatial gradient of
20 introgressed Mediterranean ancestry across the northeast Atlantic reflects cryptic

1 1 Biologie des Organismes et Ecosystèmes Aquatiques (BOREA) ; Muséum National
2 d'Histoire Naturelle, CNRS, IRD, SU, UCN, UA ; Station marine de Concarneau,
3 Quai de la Croix, 29900 Concarneau, France

4 2 Université de Bretagne Occidentale, Institut GéoArchi EA7462, 6 avenue Le
5 Gorgeu, 29100 Brest, France

6 3 ISEM, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France.

7

8 * corresponding author : robinet@mnhn.fr

[Introgression reveals cryptic connectivity in sea bass]

21 patterns of genetic and demographic connectivity. Using a 1K SNP chip dataset in
22 827 individuals sampled from Portugal to the North Sea, we found null overall
23 genetic differentiation across the northeast Atlantic. We however detected a subtle
24 latitudinal admixture gradient originating at the edge of the contact zone with the
25 Mediterranean sea bass lineage. Two significant breaks in the ancestry gradient at the
26 tip of Galicia and northern Brittany indicated barriers to effective dispersal between
27 demographically distinct units. Moreover, a northward expansion signal in Irish and
28 North Seas was revealed by the surfing of rare Mediterranean alleles at the edge of
29 the species range. Our results show that introgressed ancestry gradients offer a
30 powerful alternative to assess genetic and demographic connectivity when the
31 neutral migration-drift balance is not informative. (250 words)

32

33 Keywords: admixture, expansion, genetic and demographic connectivity,
34 introgression, spatial structure, stock delineation

35 **1 | INTRODUCTION**

36

37 Assessment of genetic connectivity among populations is widely used in
38 conservation biology and species management to evaluate effective dispersal and the
39 consequences of spatially-dependent evolutionary processes on species persistence,
40 resilience and adaptability (Cayuela et al. 2018; Gagnaire 2020). Using polymorphism
41 data from individuals sampled across the landscape, the genetic approach can detect
42 local barriers to gene flow separating distinct populations that exchange measurable
43 amounts of migrants per generation. Demographic independence between such
44 populations can be assumed when the fraction of exchanged migrants is too small to
45 ensure population persistence in case of negative intrinsic growth rate (Pulliam, 1988;
46 Lowe & Allendorf, 2010). Measuring migratory exchanges using genetic data may,
47 however, be a difficult task.

48 The power of the population genetics approach depends on the relative intensity
49 of two opposed forces, migration (m) and genetic drift ($1/N_e$). When migration
50 overwhelms the effect of genetic drift (i.e. when the number of migrants exchanged
51 per generation ($N_e m$) exceeds a few dozen), genetic differentiation between
52 populations is close to zero. Null genetic differentiation can result from a variety of
53 equilibrium scenarios ranging from high connectivity among small populations to
54 nearly complete demographic independence among large populations.
55 Consequently, assessing demographic connectivity in species with both large
56 populations and high migration rates has been a long-standing challenge to
57 population genetic approaches (Waples, 1998). Such combinations of biological
58 parameters are frequently encountered in marine species with a dispersive larval
59 stage (Hedgecock et al., 2007). Therefore, a persistent gap between evolutionary and
60 ecological scales often prevents managers to use genetic connectivity information for

[Introgression reveals cryptic connectivity in sea bass]

61 fisheries management (Palumbi, 2003; Waples & Gaggiotti, 2006; Waples, Punt &
62 Cope, 2008).

63 Molecular markers influenced by selection can compensate the lack of signal at
64 neutral loci when the migration-drift equilibrium is not informative. The use of
65 adaptive differentiation signals has been proposed as a solution to improve the
66 delineation of management units (Funk et al., 2012). This approach is now greatly
67 facilitated by the availability of large polymorphism datasets to scan genomes for
68 outlier markers influenced by selection (Stapley et al., 2010; Savolainen et al., 2013;
69 Narum et al., 2013). Several studies in marine species that typically show weak to no
70 genetic differentiation at neutral markers have found stronger signals of spatial
71 structure at different scales using outlier loci (e.g. Bekkevold et al., 2016; Benestan et
72 al., 2016; Van Wyngaarden, 2017). Although these could help in delineating cryptic
73 evolutionary management units, the variation patterns displayed by outlier loci often
74 remain challenging to translate into quantitative assessments of demographic
75 connectivity. One of the main reasons is that observed genetic variation patterns can
76 be generally attributed to a variety of possible selective mechanisms (Bierne et al.
77 2013), potentially involving complex relationships between local adaptation and gene
78 flow (Tigano & Friesen, 2016). Contextualizing the eco-evolutionary history of the
79 studied species thus helps to discern the mechanisms underlying the patterns of
80 differentiation displayed by outlier loci, toward a better understanding of
81 connectivity (Liggins et al. 2020).

82 Genetic admixture between differentiated lineages is a particularly informative
83 evolutionary context to learn about connectivity. When two divergent taxa come into
84 secondary contact and exchange genes, the spatial diffusion of foreign genetic
85 material can be used to reveal the population genetic connectivity within each
86 introgressed lineage (Sedghifar et al., 2015; Bertl et al., 2018; Duranton et al., 2019). If

[Introgression reveals cryptic connectivity in sea bass]

87 the recipient lineage is not genetically structured, the spatial homogenization of
88 foreign allele frequencies is a quick process. However, local barriers to gene flow
89 may slow down the spread of foreign alleles and generate steps in the admixture and
90 introgression gradients within the recipient lineage (Gagnaire et al., 2015). Although
91 allele frequencies equilibrate quickly at neutral loci following secondary contact,
92 barrier loci involved in partial reproductive isolation have a reduced effective
93 migration rate between lineages and therefore retain their signal of differentiation for
94 longer (Sedghifar et al., 2016). Semi-permeable species boundaries characterized by
95 heterogeneous rates of introgression among loci thus provide favorable conditions to
96 reveal cryptic barriers to gene flow within introgressed lineages, even after several
97 thousands of generations of introgression (Gagnaire et al., 2015). Post-glacial lineages
98 that are nowadays in contact and exchange genes through natural hybridization
99 represent suitable cases for using spatial introgression gradients to study
100 connectivity.

101 The European sea bass (*Dicentrarchus labrax* L.) is a marine fish species that
102 ranges from northwestern Africa to southern Norway in the Atlantic, and
103 throughout the Mediterranean and Black Seas (Vandeputte et al. 2019). The species is
104 genetically subdivided into two distinct lineages, one Atlantic (ATL) and one
105 Mediterranean (MED), which naturally hybridize at the Atlantic-Mediterranean
106 transition zone (Naciri et al., 1999; Lemaire et al., 2005; Souche et al., 2015).
107 Population genomic studies based on RAD-sequencing (Tine et al., 2014) and whole-
108 genome resequencing (Duranton et al., 2018) have revealed several important aspects
109 of the of the European sea bass evolutionary history: (i) genetic divergence between
110 ATL and MED lineages is the result of about 270 kyrs of allopatric divergence, (ii)
111 secondary contact has started after the last glacial retreat about 11.5 kyrs ago, and
112 since then (iii) gene flow occurs at variable rates across the genome due to partial

[Introgression reveals cryptic connectivity in sea bass]

113 reproductive isolation between the two lineages. (*iv*) For a still unknown reason,
114 postglacial gene flow has been more pronounced from the Atlantic into the
115 Mediterranean than in the opposite direction. As a result, contemporary ATL
116 genomes are made of ~5% of MED genetic material, while western and eastern MED
117 genomes contain ~31% and 13% of ATL ancestry, respectively (Duranton et al., 2018).
118 Such levels of introgression are sufficiently high to provide information on genetic
119 connectivity within both sea bass lineages. In the Mediterranean, Duranton et al.
120 (2019) analyzed the neutral decay of introgressed haplotype length as a function of
121 distance from the contact zone to estimate a mean per-generation dispersal distance
122 of 5 to 50 km. This quantitative approach, based on a small number of whole-genome
123 sequences, did not have the spatial resolution required to provide a detailed map of
124 connectivity. Here, we used a different strategy based on moderate genome coverage
125 but extensive spatial sampling to identify local barriers to gene flow within the
126 Atlantic lineage. Our objective was to study both the neutral and non-neutral
127 diffusion of MED alleles across Atlantic sea bass populations to evaluate their fine-
128 scale spatial genetic structure.

129 Atlantic sea bass populations are of high economic importance to European
130 fisheries, but overfishing probably combined with other factors have led to a decline
131 in landings since 2009-2010, particularly north of the 48th parallel (ICES, 2019). Sea
132 bass stock assessment by scientific authorities (ICES: International Council for the
133 Exploration of the Sea) in the northeastern Atlantic currently relies on four
134 presumably distinct stocks: northern Atlantic, southern Ireland/western Scotland,
135 Biscay and Iberia. However, stock delineation remains poorly understood, making it
136 difficult to properly assess population connectivity and implement effective
137 management programs. Population genetic studies generally found non-significant
138 genetic structure across the northeastern Atlantic (Frisch et al., 2007; Coscia &

[Introgression reveals cryptic connectivity in sea bass]

139 Mariani, 2011; Souche et al., 2015), possibly due to insufficient spatial and genomic
140 coverage. In parallel, mark-recapture data indicated restricted individual movement
141 between the Bay of Biscay and the English Channel-Celtic Sea region (Fritsch et al.,
142 2007), a result partly explained by the fidelity to winter spawning areas recently
143 evidenced with tagging data (de Pontual et al., 2019). A small number of effective
144 migrants per generation may account for the absence of genetic differentiation at
145 neutral markers despite high fidelity to spawning grounds. Therefore, it is possible
146 that independent demographic units have remained undetected by previous
147 population genetic studies. Here, we investigate fine-scale connectivity patterns
148 among Atlantic sea bass populations using the spatial diffusion of Mediterranean
149 alleles from southern Portugal toward the northern part of the species range. We
150 document the existence of two significant steps in the Mediterranean ancestry
151 gradient along the Atlantic coast, which most likely delineate demographically
152 independent populations that slightly differ from current management units. We also
153 find molecular signatures of a recent northward expansion, consistent with recently
154 expanding sea bass fisheries in the northern part of the species range.

155

156

157 **2 | MATERIALS AND METHODS**

158

159 **2.1 | Sampling**

160 Fin clips from 846 individuals were sampled from southern Portugal to the Irish Sea
161 and North Sea in the period Oct. 2012 to May 2015 (Supplementary Table S1).
162 Samples were mostly collected in fish markets, only from fresh fish with a known
163 geographical origin indicated by its ICES statistical rectangle (precision 0.5° in
164 latitude and 1° in longitude) and a known fishing day. The quality of the spatial

[Introgression reveals cryptic connectivity in sea bass]

165 fishing information of each individual was checked with the name of the fishing
166 vessel by tracking its position history on <http://www.marinetraffic.com>.

167 Tissues were preserved in 85 % ethanol at -20 °C. Individual genomic DNA was
168 extracted following cellular lysis and proteinase K digestion using the Qiagen
169 DNeasy Blood & Tissue kit, and then conserved in TE. DNA extraction quality was
170 controlled using a NanoDrop spectrophotometer keeping samples with an
171 A260/A280 ratio between 1.8 and 2.0. Concentration of double-stranded DNA was
172 then measured with the Qbit dsDNA Broad Range kit with a standard benchmark (0-
173 100 ng/ μ l) on a qPCR BioRAD. Individual dsDNA were finally normalized at 25 ng/
174 μ l and randomly distributed across nine 96-well plates before genotyping.

175

176 2.2 | Genotyping and data quality control

177 Genotyping was performed by the *Labogena* platform (Jouy-en-Josas, France) using
178 an iSelect Custom Infinium Illumina array specifically developed in the European sea
179 bass (Faggion et al., 2019). The chip contains 1531 validated SNPs covering the whole
180 genome while being homogeneously distributed along the recombination map, and
181 presents no ascertainment bias between Atlantic and Mediterranean sea bass
182 lineages. Individual genotypes and hybridization intensities generated by Illumina's
183 BeadStudio software were reanalyzed using the R package *Argyle* (Morgan et al.,
184 2016) to perform SNP quality control. Markers were filtered based on their rate of no-
185 call genotypes (per locus missingness < 0.1), rate of heterozygosity (Het < 0.55), and
186 minor-allele frequency within the whole Atlantic dataset (MAF > 0.01), in order to
187 include even variants that are frequent or fixed within the Mediterranean, but rare in
188 the Atlantic. Individuals with an excess of missing genotypes were also excluded
189 (per sample missingness < 0.1).

[Introgression reveals cryptic connectivity in sea bass]

190 In order to remove a few miscalled markers which variation profile was
191 correlated with samples arrangement in the 96-well plates, we performed a
192 genotype-plate association analysis in *Plink* (Purcell et al., 2007, p-value exclusion
193 threshold=1e-4).

194 Reference Mediterranean genotypes of 10 MED individuals (four from eastern
195 MED and six from western MED) were merged to our final dataset to compare allele
196 frequencies between ATL and MED populations. These reference genotypes were
197 extracted from high-quality whole-genome resequencing data without any missing
198 genotype at the retained loci (Duranton et al., 2018).

199

200 **2.3 | Spatial genetic variation on different datasets**

201 The 827 quality-filtered specimens were stratified into three different datasets to
202 reach variable levels of precision in spatial genetics analyses: (i) a 'regional dataset' of
203 827 specimens distributed into 7 regions; (ii) a 'main dataset' in which the 827
204 specimens were assigned to 21 localities corresponding to single ICES rectangles or
205 groups of adjacent rectangles, with a minimum precision of 1.5° in latitude and 3° in
206 longitude, and a number of individuals per locality ranging from 11 to 111 (mean=40,
207 Figure 1, Supplementary Figure S1 and Table S1 for more details); (iii) a 'refined
208 dataset', more precise in space, consisting of 761 specimens distributed into 31
209 localities, with a minimum precision of 30' in latitude and 1° in longitude, and a
210 minimal number of 8 individuals per locality (from 8 to 43, mean = 24,
211 Supplementary Figure S2). The main dataset was used for all analyses, unless stated
212 otherwise.

213 The proportion of polymorphic markers among the 1012 retained SNPs was
214 calculated for each of the 21 localities. A regression model of the proportion of
215 polymorphic markers as a function of sample size per location was fitted using a

[Introgression reveals cryptic connectivity in sea bass]

216 nonlinear least squares analysis with two parameters in order to account for the
217 effect of sample size. We then tested for a linear relationship between the localities'
218 residuals to the fitted model and the latitude of localities. We also tested the existence
219 of a linear correlation between the mean observed heterozygosity per sample locality
220 and latitude. Sample groupings from the regional dataset were then used to test for
221 deviations from expectations under Hardy-Weinberg Equilibrium within each of the
222 seven regions defined above with the *Pegas* R package (Paradis, 2010).

223 A redundancy analysis (RDA) was performed using the R package *Vegan*
224 (Oksanen et al., 2019) in order to evaluate the extent to which SNP variation is
225 influenced by geographic and temporal factors. We first fitted the model Y
226 (individual genotype) \sim (Latitude + Longitude + Month + Year of capture) and
227 assessed the significance of each explanatory factor using 1000 permutations of
228 genotypic data before refitting the model.

229 Overall F_{ST} was calculated among all of the 21 Atlantic locations using all loci
230 and pairwise F_{ST} were then calculated for each pair of sampling locations including
231 the two reference MED populations, as well as for pairs of regions using the 'regional
232 dataset'. We used the *Hierfstat* R package (Goudet, 2005) to calculate F_{ST} values and
233 assess their significance using 5000 random permutations of genotypes among
234 localities and among regions. A sequential Bonferroni correction (Hommel, 1989) was
235 applied to empirical p-values with the function *p.adjust()* of the R-base package *Stats*.
236 In order to assess the presence of an isolation-by-distance pattern among ATL
237 locations, we tested the linear correlation between [$F_{ST} / (1 - F_{ST})$] and geographic
238 distance between pairs of localities. We also performed a Mantel test between
239 matrices of pairwise genetic and geographic distances with 10.000 permutations
240 using ADEGENET (Jombart 2008). Pairwise geographic distances were calculated as

[Introgression reveals cryptic connectivity in sea bass]

241 the shortest path by the continental plateau between two localities (i.e. bathymetry <
242 200m) using the R package *Marmap* (Pante & Simon-Bouhet, 2013).

243 We used the directionality index statistics Ψ (Peter & Slatkin, 2013) to test
244 whether genetic differentiation among pairs of localities was due to isolation-by-
245 distance at equilibrium or to a recent range expansion in the Atlantic distribution
246 area of *D. labrax*. Briefly, the directionality index uses comparisons of two-
247 dimensional site frequency spectra to capture deviations from symmetric migration,
248 such as those generated during range expansions. We used the R package
249 *rangeExpansion* (<https://github.com/BenjaminPeter/rangeexpansion>) to calculate Ψ
250 for all pairs of Atlantic locations, and used pairwise geographic distances between
251 locations to infer the most likely spatial origin of a possible range expansion.

252

253 **2.4 | MED introgression in ATL genomes**

254 In order to estimate the fraction of MED alleles present in ATL individuals, we used
255 the program *Admixture* (Alexander, Novembre & Lange, 2009), which infers
256 individual ancestry proportions from K ancestral populations (here K was set to 2).
257 Our unsupervised ancestry inference included 10 MED reference individuals in
258 addition to the 827 ATL samples, so that the two ancestral populations correspond to
259 the ATL and MED sea bass lineages. We set the termination criterion for the
260 optimization algorithm to 100 iterations to ensure convergence and used 1000
261 bootstrap replicates to estimate the standard error of individual ancestry
262 proportions.

263 The effects of sampling date and geographic distance to the southernmost
264 Atlantic location (Sines in southern Portugal, code "SINE") on the extent of
265 Mediterranean introgression were tested using a Global Linear Model (GLM,
266 family=Gaussian), with the percentage of MED ancestry as the dependent variable.

[Introgression reveals cryptic connectivity in sea bass]

267 The date of capture (using the 'as.date()' class in R) and the distance to the SINE
268 locality (calculated as the shortest path by the continental plateau) were used as
269 explanatory factors.

270 The effect of geographic distance to the southernmost ATL location (SINE) on the
271 extent of MED ancestry was then tested alone using a linear regression model. The
272 distance to the SINE location was calculated as the shortest path by the continental
273 plateau. In order to test for the existence of potential breaks in the admixture
274 gradient, we then performed a piecewise regression of MED ancestry as a function of
275 distance to SINE. Breaks separating different linear regression models were
276 introduced one by one and, and at each step, every possible break position was
277 examined by calculating the residual standard error of the piecewise regression
278 model. The significance of the reduction in the residual sum of squares of each
279 piecewise regression model compared to the simple regression model was tested
280 using an ANOVA, taking into account the number of additional parameters in the
281 piecewise regression model.

282

283 **2.5 | Genome scans**

284 We then evaluated the impact of variable rates of introgression between ATL and
285 MED sea bass lineages on allele frequency gradients within the Atlantic. We more
286 specifically tested the theoretical prediction that the genomic regions with reduced
287 rates of introgression between lineages tend to exhibit the strongest spatial allele
288 frequency patterns within the ATL range due to propitious ratios of between- to
289 within-lineage gene flow intensities (Gagnaire et al., 2015). We thus expected that the
290 loci showing the strongest genetic differentiation within the Atlantic would generally
291 also be F_{ST} outliers between ATL and MED lineages. Given the relatively low density
292 of markers in our data set (about 1 SNP per cM), the candidate loci detected by

[Introgression reveals cryptic connectivity in sea bass]

293 genomes scans for differentiation are most likely to be indirectly influenced by
294 selection through linkage with a nearby selected locus. A previous quantification
295 indicated that the genomic island regions under the influence of the strongest
296 reproductive isolation barriers between Atlantic and Mediterranean lineages occupy
297 about 4% of the sea bass genome (Duranton et al. 2018). Therefore, our SNP panel is
298 expected to contain at least 40 SNPs in strong linkage with reproductive isolation
299 barriers between the two sea bass lineages, plus variants showing intermediate
300 degrees of linkage with them. To identify those loci, we used two different genome
301 scan methods to detect F_{ST} outlier SNPs. Both methods were applied to the analysis
302 of genetic differentiation at two different scales: (i) within the ATL range among the
303 21 localities, and (ii) between the ATL and MED sea bass lineages.

304 The first method used was *Lositan* (Beaumont & Nichols, 1996; Antao et al., 2008),
305 which performs coalescent simulations under the symmetric island model to
306 generate a neutral distribution of F_{ST} conditioned on the mean expected
307 heterozygosity among populations. The mean F_{ST} calculated across all markers was
308 used as a target value to simulate 1 million SNPs, from which we empirically
309 determined the 99.5th percentile of the F_{ST} distribution to identify candidate outlier
310 loci showing an excess of genetic differentiation among locations.

311 We then used *BayeScan* (Foll & Gaggiotti, 2008), a Bayesian outlier detection
312 method that relies on the multinomial-Dirichlet model. The difference in allele
313 frequency between every sampling location and a common theoretical gene pool was
314 measured by a subpopulation-specific F_{ST} coefficient in order to account for
315 differences in effective sizes and migration rates among subpopulations. Selection
316 was introduced by decomposing for each locus the subpopulation-specific F_{ST}
317 coefficient into a subpopulation-specific component shared by all loci and a locus-
318 specific component shared by all subpopulations. Deviation from neutrality for a

[Introgression reveals cryptic connectivity in sea bass]

319 given locus was assumed when the locus-specific component was needed to explain
320 the observed diversity pattern. The posterior probability that a locus is influenced by
321 selection after accounting for multiple testing was determined by the Posterior Odds
322 (PO), which is the ratio of posterior probabilities between the selection and neutral
323 models. The minimum false discovery rate (FDR) at which a given locus reached
324 significance was determined using the q-value of each locus. A q-value threshold of
325 0.05 was used for outlier detection in most analyses, corresponding to an FDR
326 threshold of 5%. *BayeScan* chain parameters were set to 5.000 outputted iterations
327 with 5.000 steps for pilot runs and 50.000 burn-in steps.

328

329 **2.6 | Analysis of introgression gradients around barrier loci**

330 Simulations of secondary contact between two genetically subdivided lineages
331 exhibiting partial reproductive isolation show that neutral markers linked to barrier
332 loci can display substantial allele frequency steps within the introgressed lineage
333 (Gagnaire et al., 2015). These steps appear at the place where local barriers to gene
334 flow slow down the diffusion of introgressed alleles. They are dynamically
335 maintained by differences in flow intensity, and are amplified when gene flow
336 between lineages is slightly higher than the rate of homogenization within the
337 introgressed lineage. Depending on the age of the secondary contact, the steps in
338 allele frequencies magnified by foreign introgression may appear at variable
339 recombination distances from the barrier loci.

340 To explore that effect, we first focused on the subset of SNPs that were the most
341 strongly associated to barrier loci in our dataset, using only outlier loci that were
342 detected in the *BayeScan* MED-ATL genome-scan with a stringent FDR threshold of
343 1%. Since genetic differentiation is strong between ATL and MED within genomic
344 islands, the minor frequency allele found at each of these loci in the ATL most often

[Introgression reveals cryptic connectivity in sea bass]

345 corresponds to the major frequency allele in the MED. Therefore, gradients of
346 introgression of Mediterranean alleles at those loci are expected to generate minor
347 allele frequency (MAF) clines within the Atlantic distribution range. In order to test
348 for a cryptic population structure within the Atlantic, we thus compared the mean
349 MAF of MED-ATL outlier loci to that of neutral SNPs as a function of the distance to
350 the southernmost ATL location (SINE).

351 Finally, we evaluated how different degrees of linkage to barrier loci affect MAF
352 gradients within the Atlantic. We assigned SNPs to different categories defined by
353 their level of allele frequency difference (Δp) between ATL and MED. Our rationale
354 was that stronger Δp values between ATL and MED are associated to stronger
355 linkage to barrier loci, which was supported from earlier works (Duranton et al.,
356 2018).

357

358

359 **3 | RESULTS**

360

361 **3.1 | SNP genotyping and quality filtering**

362 After excluding non-variable markers corresponding to polymorphisms private to
363 the Mediterranean sea bass lineage as well as marker genotyping failures, 1361 SNPs
364 were successfully genotyped in 846 specimens over the 1531 markers present on the
365 array, representing an overall call-rate of 90.7%. SNPs with a minor allele frequency
366 (MAF) lower than 0.01 within the Atlantic were subsequently excluded, as were
367 individuals with more than 10% of non-scored genotypes. A few additional SNPs
368 showing variation profiles significantly correlated with the samples' mapping within
369 the 96-wells plates used for genotyping were also discarded to avoid experimental

[Introgression reveals cryptic connectivity in sea bass]

370 effects. After these steps of quality-filtering, 1012 bi-allelic SNP markers that were
371 genotyped in 827 individuals were retained in the final main dataset (Supplementary
372 Table S1), resulting in an overall genotyping call rate of 96.87%.

373

374 **3.2 | Genetic diversity gradient and spatial structure**

375 The average fraction of polymorphic markers per sampling location calculated across
376 the 1012 SNPs was negatively correlated to latitude after standardizing for the
377 number of specimens genotyped by locality (adjusted $R^2 = 0.77$; $p < 0.001$; Figure 2a),
378 providing evidence of a latitudinal decrease in genetic diversity across the ATL sea
379 bass range. Similarly, the mean observed heterozygosity per sample location was
380 negatively correlated to latitude (adj. $r^2 = 0.66$; $p < 0.001$; Supplementary Figure S3),
381 decreasing by nearly 10 % from Southern Portugal to northeastern UK
382 (Supplementary Table S2).

383 The proportion of the total genotypic variance explained by the redundancy
384 analysis constrained by (latitude + longitude + month + year of capture) was low
385 (0.64 %) but highly significant (p -value < 0.001). However, constrained ordination
386 only revealed significant marginal effects on SNP variation for the two spatial
387 variables (latitude: p -value < 0.001 ; longitude: p -value = 0.037). The RDA1 axis
388 received a large contribution of latitude, as illustrated by the gradient in individual
389 coordinates distributions among the 7 regions from Portugal to the North Sea (Figure
390 2b).

391 The overall F_{ST} calculated over all loci among the 21 Atlantic locations was not
392 significant ($F_{ST} = 0.0002$, $p > 0.05$). Pairwise F_{ST} values calculated between pairs of
393 ATL locations or regions were consistently low and most often non-significant
394 (Supplementary Tables S3-S5), and only few pairwise F_{ST} values reached significance
395 among the 7 ATL regions considered (Supplementary Table S6). By contrast, genetic

[Introgression reveals cryptic connectivity in sea bass]

396 differentiation was much higher between ATL and MED sea bass lineages (average
397 $F_{ST} = 0.1474$) using the same marker dataset, consistently with previous studies (Tine
398 et al. 2014; Duranton et al. 2018). A weak but significant positive correlation was
399 found between geographic distance separating pairs of ATL localities and genetic
400 distance estimated by $F_{ST} / (1 - F_{ST})$ (adj. $r^2 = 0.078$; $p < 0.001$; Figure 2c). This
401 isolation-by-distance (IBD) pattern was remarkably stronger when the analysis was
402 restricted to the northern part of the Atlantic range, that is, excluding Portugal,
403 Biscay and southwestern Channel locations (adj. $r^2 = 0.236$; $p\text{-value} < 0.001$;
404 Supplementary Figure S4). This was also confirmed by the Mantel test
405 (Supplementary Table S7).

406 The analysis of directionality index Ψ based on the method by Peter and Slatkin
407 (2013) detected a strongly significant deviation from an isolation-by-distance at
408 equilibrium model. Instead, a spatial expansion scenario was supported with an
409 inferred origin located in the northwestern part of the Celtic Sea ($p\text{-value} < 10^{-10}$,
410 Figure 2d). The gradient in the relative fit as expansion origin showed a steep
411 decrease in likelihood location towards the southern Celtic Sea and northern Biscay,
412 and a smoother gradient towards the northeastern part of the range (Irish Sea,
413 English Channel and North Sea). Together with the decreasing latitudinal gradient in
414 heterozygosity, these results support a recent northeastern range expansion from a
415 region located in the northwestern Celtic Sea.

416

417 **3.3 | Spatial gradient of MED ancestry in the ATL**

418 Our inferences of individual MED ancestry proportions showed a clear northward-
419 decreasing gradient (Figure 3). The southernmost location in Portugal (SINE), which
420 is the closest to the ATL-MED transition zone, showed a significantly higher average

[Introgression reveals cryptic connectivity in sea bass]

421 level of MED ancestry (11.2 %) compared to all other locations and a stronger
422 variance in MED ancestry among individuals.

423 The GLM analysis showed that the explanatory variable "distance to SINE" had a
424 very significant negative effect ($p < 2e-16$) on the percentage of MED ancestry,
425 whereas no significant effect could be detected for the date of sampling ($p=0.135$).

426 The spatial gradient in MED ancestry was even more obvious when individual
427 ancestry proportions were averaged by sampling location (Figure 4a). The mean
428 fraction of MED ancestry per location was negatively correlated with the distance to
429 SINE in the simple linear regression model (adj. $r^2 = 0.6026$, $p\text{-value} < 0.001$), and the
430 correlation was strengthened when the SINE location (which is highly admixed) was
431 removed from the analysis (adj. $r^2 = 0.7405$; $p\text{-value} < 0.001$).

432 The piecewise regression approach detected two significant break points in the
433 spatial admixture gradient. Taking into account the presence of two additional
434 parameters, the model with three different linear regression displayed a significantly
435 better fit to the data than the simple linear regression model (ANOVA for model
436 comparison: $p\text{-value} = 3.094e-9$; best model with 3 regression lines: adj. $r^2 = 0.93$, $p\text{-value} = 1.3e-9$). The strongest break occurred in the southwestern English Channel
438 near the GONB locality in the Gulf of Saint-Malo, more than 2000 km northward to
439 SINE by the plateau (Figure 4b). The second break was detected at the tip of Galicia
440 between CORU and ASTU sampling locations (Figure 4c). Consistent with the
441 dilution of MED ancestry toward the north, the average fraction of MED alleles
442 decreased from 5.7 % [0.037, 0.076] in the Portugal region, to 2.6 % [0.013, 0.041] in
443 the Bay of Biscay and southwestern English Channel, and 1.2 % [0.004, 0.022] in the
444 north of Brittany from Celtic Sea to Irish and North Seas (Figure 3). Each of these two
445 steps was characterized by a more than two-fold reduction in the mean MED
446 ancestry from south to north (Figure 4d). Although we found negative regression

[Introgression reveals cryptic connectivity in sea bass]

447 slopes indicating a decreasing mean MED ancestry with increasing distance to SINE
448 in both the Portugal and Biscay-southwestern Channel regions, the slope was
449 significantly positive to the north of the Gulf of Saint-Malo breakpoint. This striking
450 result, indicating an inversed latitudinal gradient in MED ancestry in the northern
451 part of the range, remained visible separately in a western Channel-Irish Sea transect
452 and an eastern Channel-North Sea transect using the more spatially 'refined dataset'
453 containing 31 localities (Supplementary Figure S5a).

454

455 **3.4 | Within-ATL outliers are enriched for between-lineages outliers**

456 Genome scans for highly differentiated SNPs were performed both at the within-ATL
457 scale to detect outlier loci within the northeastern Atlantic and between ATL and
458 MED samples to detect outlier loci showing an excess of differentiation between the
459 two European sea bass lineages.

460 *Lositan* detected 32 outliers exceeding the 99.5th percentile of the neutral
461 distribution of F_{ST} at the within-ATL scale (Supplementary Figure S6a) and 183
462 outliers between ATL and MED lineages (Supplementary Figure S6b). Within-ATL
463 outliers were highly significantly enriched for between-lineages outliers (21 out of 32
464 within-ATL outliers belong to the 183 ATL-MED outliers, hypergeometric test p-
465 value = 1.9e-9).

466 Comparatively, the *BayeScan* analysis was more stringent in detecting outlier loci
467 showing q-values < 0.05, with only 7 candidate outliers being found at the within-
468 ATL scale (Supplementary Figure S6c) and 74 outliers detected between ATL and
469 MED lineages (Supplementary Figure S6d). All of the outliers identified with
470 BAYESCAN were also detected by LOSITAN. Here again, the outliers detected at the
471 within-ATL scale were highly significantly enriched for between-lineages outliers (5

[Introgression reveals cryptic connectivity in sea bass]

472 out of 7 within-ATL outliers belong to the 74 ATL-MED outliers, hypergeometric test
473 p-value = $3.4e-5$.

474

475 **3.5 | Minor allele frequency steps at genomic islands associated loci**

476 We used a stringent subset of 52 outlier loci that were detected in the BAYESCAN
477 MED-ATL genome-scan with an FDR < 0.01 to characterize the gradient in mean
478 MAF as a function of the distance to SINE. Major steps were observed directly
479 northward to CORU and GONB in these mean MAF profiles, which were not
480 detected using non-outlier loci (Figure 5). In order to refine the spatial positions of
481 these steps, the same analysis was made based on the spatially refined dataset (31
482 localities). Mean MAF were represented separately along two transects: one from
483 Portugal toward the Irish Sea and one toward the North Sea (Supplementary Figure
484 S5b). Both profiles revealed low MAF values in the western Channel area (after
485 PBRE-CRNW and PBRE-GONB), with the lowest values being found in IRIS,
486 CRNW2 and 3, ISWI2 and BSEI1. Northward to these regions, the mean MAF profiles
487 steadily increased towards WALE-SGEO (Irish Sea) and DOVE-ANGL-NEUK (North
488 Sea), on both sides of Britain.

489 Finally, the evolution of mean MAF profiles was evaluated as a function of the
490 degree of marker linkage to MED-ATL islands of divergence assessed with difference
491 in allele frequencies between ATL and MED (Δp). Progressively flattening MAF
492 profiles were found for groups of loci showing from high to low Δp values
493 (Supplementary Figure S7).

494

495

496 **4 | DISCUSSION**

497

498 Genetic connectivity assessment in high gene flow species generally provides little
499 information on the degree of demographic connectivity among populations (Waples,
500 1998; Lowe & Allendorf, 2010). This shortcoming typically arises when genetic
501 differentiation at neutral markers is null or close to zero, so that no real signal can be
502 used to determine whether migration rates are sufficiently high to ensure
503 demographic coupling among populations. Here, we found non-significant overall
504 genetic differentiation across the vast majority of the European sea bass northeastern
505 Atlantic range. The sea bass thus represents one of the many borderline cases where
506 neutral markers may remain silent to conservation and management issues.

507 As for many other species, however, the spatial distribution of genetic diversity
508 in sea bass has retained signals of historical and contemporary admixture. Two sea
509 bass lineages, one Atlantic and one Mediterranean, have diverged in allopatry during
510 ca. 300,000 years and have subsequently experienced a postglacial secondary contact
511 which started less than 15,000 years BP (Tine et al., 2014; Duranton et al., 2018). Since
512 that time, gene flow has been strongly asymmetrical from the Atlantic into the
513 Mediterranean lineage. However, gene flow has been also occurring in the opposite
514 direction, accounting for less than 5% of introgressed Mediterranean alleles within
515 Atlantic sea bass genomes (Duranton et al., 2018). Here, we show that the northward
516 diffusion of Mediterranean alleles within the Atlantic sea bass populations can be
517 used to reveal two dispersal barriers separating demographically independent units,
518 and an ongoing expansion in the northern part of the species range. These findings
519 could provide useful new elements to improve the delineation of European sea bass
520 stocks in fisheries management programs.

521

522 **4.1 | Cryptic dispersal barriers revealed by introgression clines**

[Introgression reveals cryptic connectivity in sea bass]

523 Introgression between differentiated lineages can be used as an alternate way to
524 assess within-species genetic and demographic connectivity (Gagnaire et al., 2011;
525 Gagnaire et al., 2015; Duranton et al., 2019). Several marine species display large
526 spatial scale allele frequency patterns that can be attributed to ongoing gene flow
527 between divergent units (e.g. Milano et al., 2014; Dahle et al., 2018; Lehnert et al.,
528 2018; Wennevik et al., 2019). These admixture gradients originate at the edge of
529 contact zones and sometimes attenuate over large distances beyond the contact
530 zones. They are dynamically maintained by continuing gene flow between lineages,
531 combined with a delayed rate of homogenization within the introgressed lineages
532 due to spatially limited dispersal. Selection against introgressed fragments may also
533 contribute to maintaining these admixture gradients (Sedghifar et al., 2016). Spatial
534 discontinuities in admixture gradients indicate local reductions in the spread of
535 introgressed alleles (Gagnaire et al., 2015). Therefore, they can be used to identify
536 cryptic dispersal barriers that are not strong enough to produce genetic
537 differentiation at migration-drift equilibrium.

538 Here, the analysis of individual ancestry proportions within the Atlantic revealed
539 the existence of a wide admixture gradient extending from southern Portugal to the
540 British Isles. Our closest sampling location from the contact zone with the
541 Mediterranean sea bass lineage (Sines, South Portugal) displayed the highest mean
542 and variance in individual Mediterranean ancestry (mean=0.11, s.d.=0.09). This was
543 mainly driven by the presence of admixed genotypes with significantly higher
544 proportions of Mediterranean ancestry above the background population level (e.g. 6
545 individuals between 0.15 and 0.40 in Sines), which most likely represent late-
546 generation backcrosses. Therefore, our results support that ongoing gene flow
547 between the two sea bass lineages contributes to maintain the admixture gradient.

[Introgression reveals cryptic connectivity in sea bass]

548 Northward to Sines (SINE), the level of Mediterranean ancestry mostly remained
549 below 5%, displaying a general decreasing trend with increasing latitude. A detailed
550 examination of this admixture gradient revealed the existence of two significant
551 breakpoints located at the tip of Galicia (CORU) and near the Gulf of Saint-Malo
552 (GONB). Two-fold reductions in mean Mediterranean ancestry were observed
553 between sampling sites directly located in the southern vs. northern sides of these
554 two breaks. Such a magnitude of change cannot be easily explained by insufficient
555 sampling coverage generating spurious breaks in an otherwise continuous gradient.
556 Instead, it indicates the presence of cryptic barriers that significantly reduce gene
557 flow over relatively short distances, on a scale of a hundred kilometers or less (i.e. the
558 distance between two consecutive sampling sites).

559 The observed discontinuities in admixture proportions are predicted under
560 spatial admixture models, and can be maintained as long as differences in allele
561 frequencies exist between hybridizing lineages (Gagnaire et al., 2015). For marine
562 species with large population sizes and a planktonic larval stage, it is often assumed
563 that a rapid erosion of allele frequency differences occurs upon secondary contact.
564 However, the existence of genetic barriers to gene flow between Atlantic and
565 Mediterranean sea bass lineages have potentially delayed the dissipation of allele-
566 frequency gradients for a substantial amount of time (Tine et al., 2014; Duranton et
567 al., 2018). Our results support this interpretation since most outlier loci found within
568 the Atlantic were likely explained by the introgression of Mediterranean alleles
569 (Figure 5). Moreover, loci showing delayed rates of introgression between lineages
570 also displayed the largest steps in allele frequencies within the Atlantic, as predicted
571 by theory (Gagnaire et al., 2015; Supplementary Figure S7).

572 Our results suggest that migration rates between populations on either side of
573 the identified breaks are too low to ensure demographic coupling, otherwise the

[Introgression reveals cryptic connectivity in sea bass]

574 breaks would not have been observed. However, quantifying the rate of effective
575 migration across these breaks would require additional data on the intensity of
576 selection acting at the barrier loci and their recombination rate with the neutral
577 markers we used here. This remains challenging, and therefore calls for the use of
578 complementary information on individual movement, as inferred for instance from
579 archival tags (de Pontual et al., 2019) or scale microchemistry (Cambiè et al., 2016).
580 Such a coupling between life-history migration and genetic data has the potential to
581 refine our understanding of sea bass demographic connectivity in the northeastern
582 Atlantic.

583

584 **4.2 | Surfing of introgressed alleles owing to northward range expansion**

585 Northward to the break identified near the Gulf of Saint-Malo, the mean
586 Mediterranean ancestry surprisingly increased towards the northernmost localities of
587 WALE-SGEO (Irish Sea) and DOVE-ANGL-NEUK (North Sea) on both sides of
588 Britain (Supplementary Figure S5). This result was explained by an increased
589 variance in Mediterranean ancestry among individuals (Figure 3). The mean
590 Mediterranean ancestry in the English Channel is close to 1%, and therefore the
591 increased variance in ancestry in the Irish Sea and North Sea results in increased
592 average ancestry due to border effect (i.e. individual ancestries can deviate more
593 strongly when they lie above versus below the 1% average ancestry).

594 A possible mechanism underlying the increased variance in Mediterranean
595 ancestry is a recent range expansion in the northern part of the species distribution
596 area. This spatial expansion hypothesis was supported both by the decreasing
597 latitudinal gradient in heterozygosity and the directionality index analysis. A
598 number of direct observations also support the recent increase in sea bass abundance
599 around Britain (Wright et al. 2020). This is exemplified by the rapid increase of

[Introgression reveals cryptic connectivity in sea bass]

600 Eastern Channel, North and Irish Seas populations attributed to increased seawater
601 winter temperatures (Henderson, 2007; Pawson et al., 2007), and the new or
602 expanding sea bass fisheries in UK and Irish seas (Cheung et al., 2012).

603 In a spatially expanding population, the intensity of genetic drift is amplified at
604 the wave front, which leads to increased variance in allele frequencies compared to
605 the core population (Edmonds et al., 2004; Klopstein et al., 2006). Putting this in the
606 sea bass context may help explain how range expansion affects Mediterranean
607 ancestry in the northern part of the species range. Alleles of Mediterranean origin
608 segregate at low frequencies at the place where spatial expansion occurs. Therefore,
609 most of these low-frequency Mediterranean alleles are likely to be lost by random
610 genetic drift or to remain at low frequency during range expansion. Some alleles of
611 Mediterranean origin, however, can successfully surf on the wave of advance and
612 reach significantly higher frequencies. These two opposite effects of genetic drift (i.e.
613 small decrease in frequency for most Mediterranean alleles but successful surfing for
614 part of them) is expected to increase the variance in ancestry among individuals
615 living in the expansion zone. Thus, our results are consistent with the surfing of a
616 fraction of the rare Mediterranean alleles, as a consequence of ongoing northward
617 expansion of sea bass populations on both sides of Britain. A more thorough
618 examination of this hypothesis would require denser genome-wide polymorphism
619 data to test the surfing of local ancestry blocks of Mediterranean origin.

620 A possible alternative explanation for the observed ancestry patterns is selection
621 against weakly deleterious introgressed fragments of Mediterranean origin. Since the
622 rate of introgression is particularly reduced in low-recombining regions of the sea
623 bass genome, selection against introgression is thought to rely on many loci of small
624 individual effects (Duranton et al., 2018). In the large populations of Portugal, Biscay
625 and western Channel, selection against introgressed tracts of Mediterranean origin

[Introgression reveals cryptic connectivity in sea bass]

626 could be sufficiently efficient to progressively decrease Mediterranean ancestry as the
627 distance to the contact zone increases. By contrast, lower effective population sizes in
628 the northern margins of the species range could reduce the efficacy of selection
629 against introgression, resulting in slightly increased mean Mediterranean ancestry in
630 these regions. The two proposed mechanisms (i.e. surfing of introgressed alleles and
631 reduced efficacy of selection against introgression) are not mutually exclusive and
632 may collectively explain the spatial ancestry patterns documented here.

633

634

635 **4.3 | Applications to conservation and management**

636

637 Our results mirror previous findings that support the existence of dispersal barriers
638 in the two identified regions of Galicia and Brittany for other marine taxa.

639 The northwestern coast of the Iberian Peninsula is characterized by complex
640 seasonal oceanographic circulation patterns that may act as potential barriers to
641 dispersal (Varela et al., 2005). Mixed evidence from genetic studies point out that, on
642 the one hand, most marine species lack genetic structure across northwestern Iberia
643 (reviewed in Gomes-Gesteira et al., 2011). On the other hand, some exceptions have
644 been found such as in *Littorina saxatilis* (Piñeira et al., 2008) or in *Fucus ceranoides*
645 (Neiva et al., 2012), which both lack planktonic dispersive stages. In these species,
646 however, genetic breaks may also be attributed to past secondary contacts.

647 Brittany represents a less controversial biogeographic transition zone between
648 marine biotas. In some species like *Fucus vesiculosus*, the genetic break occurs near the
649 Ushant thermal front that lies off the coast of western Brittany (e.g. Almedia et al.,
650 2017), and separates cold inshore surface waters from warm stratified offshore

[Introgression reveals cryptic connectivity in sea bass]

651 waters (Chevallier et al., 2014). Another important transition zone, visible in species
652 assemblages of benthic macrofauna (Gaudin et al., 2018), and contact zones between
653 divergence lineages (Bierne et al., 2003; Jolly et al., 2005), corresponds to the Gulf of
654 Saint-Malo that separates western from eastern English Channel.

655 Most interestingly, a genetic connectivity study in the marine trematode parasite
656 *Bucephalus minimus*, whose final and most vagile host is *Dicentrarchus labrax*, revealed
657 significant population structure along the northeastern Atlantic coastline (Feis et al.,
658 2015). Population samples of parasites taken from both sides of the two dispersal
659 barriers evidenced here exhibited genetic differentiation, which in turn suggested
660 some level of isolation among their host sea bass populations. By contrast,
661 population genetic studies conducted in sea bass generally failed to detect any
662 significant genetic structure between populations of the Bay of Biscay and English
663 Channel-Celtic Sea (Frisch et al., 2007; Coscia & Mariani, 2011; Souche et al., 2015).
664 On the other hand, mark-recapture data indicated restricted movement between
665 these two populations, thus supporting their demographic independence (Fritsch et
666 al., 2007). An important component in the maintenance of independent demographic
667 units is the fidelity to winter spawning areas, recently evidenced with tagging data
668 (de Pontual et al., 2019). Limited effective dispersal of larvae across barriers in Galicia
669 and Brittany is also implicitly supported by our findings, and could be explained by
670 the location of spawning areas in relation to the main oceanic current features.

671 The results presented here are thus broadly consistent with previous findings.
672 They confirm that (i) Atlantic sea bass populations exhibit strong genetic connectivity
673 at neutral markers, which implies that (ii) gene flow is high enough to ensure the
674 spread of adaptive variants among populations. (iii) However, the number of
675 migrants per generation among populations separated by barriers does not seem
676 sufficiently large to homogenize the latitudinal gradient in Mediterranean ancestry

[Introgression reveals cryptic connectivity in sea bass]

677 across the whole genome. This supports the existence of already suspected
678 demographically independent units, maintained by behavioral processes and
679 spatially limited larval dispersal. Finally, (iv) northern Atlantic sea bass populations
680 seem to exhibit genetic signatures of ongoing northward expansion on both sides of
681 Britain. The demographic independence of these expanding populations with respect
682 to neighboring populations from Celtic Sea and English Channel could not be
683 specifically addressed here and will need further examination. However, the finding
684 of weak but significant genetic differentiation between Irish-Celtic Seas and North
685 Sea populations, indicates demographic independence over long distances.

686 These conclusions potentially have direct implications for the management of
687 European sea bass populations in the northeastern Atlantic. The most recent report of
688 the ICES advisory committee for sea bass populations (ICES, 2019) states that
689 spawning-stock biomass has been declining since 2009-2010 and is now just above or
690 below the maximum sustainability yield, depending on the region considered. The
691 need for precise stock assessment is therefore more crucial than ever for a proper
692 management of fisheries. To this end, the spatial delineation of at least three main
693 stocks would need to be revised to account for these new findings. Namely, the
694 northern management unit comprising populations from the North Sea, Irish Sea,
695 English Channel, Bristol Channel, and Celtic Sea includes areas that most likely
696 belong to the northern part of the Bay of Biscay population. Moreover, the Bay of
697 Biscay south, which is currently grouped with the Atlantic Iberian Waters, most
698 likely belongs to the Bay of Biscay population.

699 To conclude, the approach presented in this study supports that introgression
700 gradients consisting of less than 5% of introgressed ancestry are sufficient to detect
701 cryptic contemporary barriers to gene flow, even using moderate genome coverage.
702 Estimation of migration rates across these barriers and identification of ideal

[Introgression reveals cryptic connectivity in sea bass]

703 conditions for quantitative assessments will require further development. However,
704 we believe that these will be encouraged by the frequent discovery of admixture in
705 many species, thus expanding the possibilities of studying genetic connectivity in
706 other taxa.

707

708 **DATA ACCESSIBILITY STATEMENT**

709 - Raw data, genotypes per site, variables associated with sampling sites have
710 been deposited here : <https://doi.org/10.5281/zenodo.3899247>

711 - All scripts used have been deposited here :
712 <https://github.com/tonyrobinet/introgression>

713

714

715 **ACKNOWLEDGEMENTS**

716 This study was funded by France Filière Pêche (FFP, convention PH/2013/11)
717 and the Muséum National d'Histoire Naturelle (MNHN, SJ 833-13). Authors thank
718 Morgane Ramonet (FFP) and Thierry Guigues (Pêcheurs de Bretagne) for supporting
719 this work from the project (2012) to the final report (2018). Thanks to producer's
720 organizations for their collaboration (Coopérative Maritime Étaploise, Cobrenord,
721 Pêcheurs de Bretagne, Pêcheurs d'Aquitaine, Plymouth Trawler Agents), and to the
722 34 fish markets that kindly opened their doors, nights and days. Special thanks to:
723 Mickaël Drogou and the Ifremer LBH lab in Brest for sharing tissues from Biscay;
724 Ifremer sea campaign in the Bay of Biscay (N/O Thalassa, EVOHE 2012-14); the
725 fishing vessel Les Calanques (South Ireland); Jasmin Hilien and Filip Volckaert
726 (Leuven Univ.) for sharing some tissues from UK; David Righton, Stephen Shaw
727 (CEFAS Lowestoft), John Ashworth (Newlyn) for collecting fin-clips in UK waters (in

[Introgression reveals cryptic connectivity in sea bass]

728 fish markets and on N/O Endeavour); Alison M. Pessell (Plymouth Trawler Agents);
729 Eva Maria Velasco Gil (Instituto Espanol de Oceanografia, Gijon) and Eneko Aierbe
730 Sarasola (Gijon) for collecting fin-clips in Asturian fish markets; Alexandre Alonso-
731 Fernandez, Fran Saborido-Rey and David Villegas-Rios (Instituto de Investigaciones
732 Marinas, CSIC, Vigo-Pontevedra) for their kind collaboration in Galicia; Joana Robalo
733 and Carla Sousa Santos (ISPA MARE, Lisboa) for collecting fin-clips in Portuguese
734 fish markets. Thanks to Daniel Sellos (MNHN Concarneau) for his kind help in the
735 lab. The genotyping chipset was designed by both François Allal (Ifremer Palavas)
736 and PAG. The last special-special-thanks go to the numerous little hands from the
737 marine station MNHN of Concarneau, for their much appreciated effort (Claudie
738 Guenguant, Thomas Barreau, Bernard Bourlès, Marie Cadou, Guillaume Dirberg,
739 Lou Frotté, Samuel Iglésias, Pierre-Yves Lebon, Sébastien Mayot, Julien Ringelstein,
740 Héloïse You!). The manuscrit was improved by three anonymous reviewers, that we
741 would like to acknowledge for their time, positive assessment and helpful comments.

742

743

744 **AUTHORS CONTRIBUTIONS**

745 TR and VR wrote the proposal, raised funds; TR collected specimens with many
746 others; PAG conducted and supervised the analysis, with the help of TR and VR; TR
747 and KC managed collections and prepared DNA for genotyping; TR and PAG
748 analyzed and interpreted the results; TR and PAG wrote the paper, with the help of
749 VR.

750

751

752 **REFERENCES**

- 753 Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of
754 ancestry in unrelated individuals. *Genome research*, 19(9), 1655-1664.
- 755 Almeida, S. C., Nicastro, K. R., Zardi, G. I., Pearson, G. A., Valero, M., & Serrão, E. A.
756 (2017). Reproductive strategies and population genetic structure of *Fucus* spp.
757 across a northeast Atlantic biogeographic transition. *Aquatic Living Resources*, 30,
758 16.
- 759 Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: a
760 workbench to detect molecular adaptation based on a FST-outlier method. *BMC*
761 *bioinformatics*, 9(1), 323.
- 762 Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of
763 population structure. *Proceedings of the Royal Society of London, Series B: Biological*
764 *Sciences*, 263, 1619– 1626.
- 765 Bekkevold, D., Gross, R., Arula, T., Helyar, S. J., & Ojaveer, H. (2016). Outlier loci
766 detect intraspecific biodiversity amongst spring and autumn spawning herring
767 across local scales. *PloS one*, 11(4).
- 768 Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., ...
769 & Bernatchez, L. (2016). Seascape genomics provides evidence for thermal
770 adaptation and current-mediated population structure in American lobster
771 (*Homarus americanus*). *Molecular Ecology*, 25(20), 5073-5092.
- 772 Bertl, J., Ringbauer, H., & Blum, M. G. (2018). Can secondary contact following range
773 expansion be distinguished from barriers to gene flow?. *PeerJ*, 6, e5325.
- 774 Bierne, N., Borsa, P., Daguin, C., Jollivet, D., Viard, F., Bonhomme, F., & David, P.
775 (2003). Introgression patterns in the mosaic hybrid zone between *Mytilus edulis*
776 and *M. galloprovincialis*. *Molecular Ecology*, 12(2), 447-461.

- 777 Bierne, N., Roze, D., & Welch, J. J. (2013). Pervasive selection or is it...? why are FST
778 outliers sometimes so frequent? *Molecular ecology*, 22(8), 2061-2064.
- 779 Cambiè, G., Kaiser, M. J., Marriott, A. L., Fox, J., Lambert, G., Hiddink, J. G., ... &
780 McCarthy, I. D. (2016). Stable isotope signatures reveal small-scale spatial
781 separation in populations of European sea bass. *Marine Ecology Progress Series*, 546,
782 213-223.
- 783 Cayuela, H., Rougemont, Q., Prunier, J. G., Moore, J. S., Clobert, J., Besnard, A., &
784 Bernatchez, L. (2018). Demographic and genetic approaches to study dispersal in
785 wild animal populations: A methodological review. *Molecular Ecology*, 27(20),
786 3976-4010.
- 787 Cheung, W. W., Pinnegar, J., Merino, G., Jones, M. C., & Barange, M. (2012). Review
788 of climate change impacts on marine fisheries in the UK and Ireland. *Aquatic
789 Conservation: Marine and Freshwater Ecosystems*, 22(3), 368-388.
- 790 Chevallier, C., Herbette, S., Marié, L., Le Borgne, P., Marsouin, A., Péré, S., ... &
791 Reason, C. (2014). Observations of the Ushant front displacements with
792 MSG/SEVIRI derived sea surface temperature data. *Remote sensing of environment*,
793 146, 3-10.
- 794 Coscia, I., & Mariani, S. (2011). Phylogeography and population structure of
795 European sea bass in the north-east Atlantic. *Biological Journal of the Linnean Society*,
796 104(2), 364-377.
- 797 Dahle, G., Quintela, M., Johansen, T., Westgaard, J. I., Besnier, F., Aglen, A., ... &
798 Glover, K. A. (2018). Analysis of coastal cod (*Gadus morhua* L.) sampled on
799 spawning sites reveals a genetic gradient throughout Norway's coastline. *BMC
800 genetics*, 19(1), 42.

[Introgression reveals cryptic connectivity in sea bass]

- 801 Duranton, M., Allal, F., Fraïsse, C., Bierne, N., Bonhomme, F., & Gagnaire, P. A.
802 (2018). The origin and remolding of genomic islands of differentiation in the
803 European sea bass. *Nature communications*, 9(1), 2518.
- 804 Duranton, M., Bonhomme, F., & Gagnaire, P. A. (2019). The spatial scale of dispersal
805 revealed by admixture tracts. *Evolutionary Applications*. 12(9) 1743-1756.
- 806 Edmonds, C. A., Lillie, A. S., & Cavalli-Sforza, L. L. (2004). Mutations arising in the
807 wave front of an expanding population. *Proceedings of the National Academy of*
808 *Sciences*, 101(4), 975-979.
- 809 Faggion, S., Vandeputte, M., Chatain, B., Gagnaire, P. A., & Allal, F. (2019).
810 Population-specific variations of the genetic architecture of sex determination in
811 wild European sea bass *Dicentrarchus labrax* L. *Heredity*, 122(5), 612.
- 812 Feis, M. E., Thieltges, D. W., Olsen, J. L., de Montaudouin, X., Jensen, K. T., Bazairi,
813 H., ... & Luttikhuisen, P. C. (2015). The most vagile host as the main determinant
814 of population connectivity in marine macroparasites. *Marine Ecology Progress*
815 *Series*, 520, 85-99.
- 816 Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci
817 appropriate for both dominant and codominant markers: a Bayesian perspective.
818 *Genetics* 180:977-93.
- 819 Fritsch, M., Morizur, Y., Lambert, E., Bonhomme, F., & Guinand, B. (2007).
820 Assessment of sea bass (*Dicentrarchus labrax*, L.) stock delimitation in the Bay of
821 Biscay and the English Channel based on mark-recapture and genetic data.
822 *Fisheries Research*, 83(2-3), 123-132.
- 823 Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing
824 genomics for delineating conservation units. *Trends in ecology & evolution*, 27(9),
825 489-496.

[Introgression reveals cryptic connectivity in sea bass]

- 826 Gagnaire, P. A., Minegishi, Y., Zenboudji, S., Valade, P., Aoyama, J., & Berrebi, P.
827 (2011). Within-population structure highlighted by differential introgression
828 across semipermeable barriers to gene flow in *Anguilla marmorata*. *Evolution:
829 International Journal of Organic Evolution*, 65(12), 3413-3427.
- 830 Gagnaire, P.A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F.,
831 Arnaud-Haond, S., & Bierne, N. (2015). Using neutral, selected, and hitchhiker loci
832 to assess connectivity of marine populations in the genomic era. *Evolutionary
833 Applications*, 8 (8), 769–786.
- 834 Gagnaire, P. A. (2020). Comparative genomics approach to evolutionary process
835 connectivity. *Evolutionary Applications*, 13(6), 1320-1334.
- 836 Gaudin, F., Desroy, N., Dubois, S. F., Broudin, C., Cabioch, L., Fournier, J., ... &
837 Thiébaud, É. (2018). Marine sublittoral benthos fails to track temperature in
838 response to climate change in a biogeographical transition zone. *ICES Journal of
839 Marine Science*, 75(6), 1894-1907.
- 840 Gomez-Gesteira, M., Beiras, R., Presa, P., & Vilas, F. (2011). Coastal processes in
841 northwestern Iberia, Spain. *Continental Shelf Research*, 31(5), 367-375.
- 842 Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-
843 statistics. *Molecular Ecology Notes*, 5(1), 184-186.
- 844 Hedgecock, D., Barber, P. H., & Edmands, S. (2007). Genetic approaches to
845 measuring connectivity. *Oceanography*, 20(3), 70-79.
- 846 Henderson, P. A. (2007). Discrete and continuous change in the fish community of
847 the Bristol Channel in response to climate change. *Journal of the Marine Biological
848 Association of the United Kingdom*, 87(2), 589-598.
- 849 Hommel, G. (1989). A comparison of two modified Bonferroni procedures.
850 *Biometrika*, 76(3), 624-625.

- 851 ICES. 2019. Sea bass (*Dicentrarchus labrax*) in Divisions 4.b–c, 7.a, and 7.d–h (central
852 and southern North Sea, Irish Sea, English Channel, Bristol Channel, and Celtic
853 Sea). In Report of the ICES Advisory Committee, 2019, bss.27.4bc7ad-h,
854 <https://doi.org/10.17895/ices.advice.4779>
- 855 Jolly, M. T., Jollivet, D., Gentil, F., Thiébaud, É., & Viard, F. (2005). Sharp genetic
856 break between Atlantic and English Channel populations of the polychaete
857 *Pectinaria koreni*, along the North coast of France. *Heredity*, 94(1), 23.
- 858 Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic
859 markers. *Bioinformatics*, 24(11), 1403-1405.
- 860 Klopstein, S., Currat, M., & Excoffier, L. (2006). The fate of mutations surfing on the
861 wave of a range expansion. *Molecular biology and evolution*, 23(3), 482-490.
- 862 Lehnert, S. J., DiBacco, C., Jeffery, N. W., Blakeslee, A. M., Isaksson, J., Roman, J., ... &
863 Hamilton, L. C. (2018). Temporal dynamics of genetic clines of invasive European
864 green crab (*Carcinus maenas*) in eastern North America. *Evolutionary Applications*,
865 11(9), 1656-1670.
- 866 Liggins, L., Treml, E. A., & Riginos, C. (2020). Seascape genomics: contextualizing
867 adaptive and neutral genomic variation in the ocean environment. *Population
868 Genomics: Marine Organisms*, 171-218.
- 869 Lowe, W.H., & Allendorf, F.W. (2010). What can genetics tell us about population
870 connectivity? *Molecular Ecology* 19, 3038-3051.
- 871 Lemaire, C., Versini, J. J., & Bonhomme, F. (2005). Maintenance of genetic
872 differentiation across a transition zone in the sea: discordance between nuclear
873 and cytoplasmic markers. *Journal of Evolutionary Biology*, 18(1), 70-80.

[Introgression reveals cryptic connectivity in sea bass]

- 874 Milano, I., Babbucci, M., Cariani, A., Atanassova, M., ... & Bargelloni, L. (2014).
875 Outlier SNP markers reveal fine-scale genetic structuring across European hake
876 populations (*Merluccius merluccius*). *Molecular Ecology* 23(1), 118-135.
- 877 Morgan, A. P. (2016). argyle: an R package for analysis of Illumina genotyping
878 arrays. *G3: Genes, Genomes, Genetics*, 6(2), 281-286.
- 879 Naciri, M., Lemaire, C., Borsa, P., & Bonhomme, F. (1999). Genetic study of the
880 Atlantic/Mediterranean transition in sea bass (*Dicentrarchus labrax*). *Journal of*
881 *Heredity*, 90(6), 591-596.
- 882 Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013).
883 Genotyping-by-sequencing in ecological and conservation genomics. *Molecular*
884 *ecology*, 22(11), 2841-2847.
- 885 Neiva, J., Pearson, G. A., Valero, M., & Serrão, E. A. (2012). Fine-scale genetic breaks
886 driven by historical range dynamics and ongoing density-barrier effects in the
887 estuarine seaweed *Fucus ceranoides* L. *BMC Evolutionary Biology*, 12(1), 78.
- 888 Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'hara, R. B., Simpson, G. L., ... &
889 Wagner, H. (2019). Vegan: community ecology package. R package version 2.5-6.
890 <https://cran.r-project.org/web/packages/vegan/> (accessed 2020.03.22)
- 891 Palumbi, S. R. (2003). Population genetics, demographic connectivity, and the design
892 of marine reserves. *Ecological applications*, 13(sp1), 146-158.
- 893 Pante, E., & Simon-Bouhet, B. (2013). marmap: a package for importing, plotting and
894 analyzing bathymetric and topographic data in R. *PLoS One*, 8(9), e73051.
- 895 Paradis, E. (2010). pegas: an R package for population genetics with an integrated-
896 modular approach. *Bioinformatics*, 26(3), 419-420.
- 897 Pawson, M. G., Kupschus, S., & Pickett, G. D. (2007). The status of sea bass
898 (*Dicentrarchus labrax*) stocks around England and Wales, derived using a separable

- 899 catch-at-age model, and implications for fisheries management. *ICES Journal of*
900 *Marine Science*, 64(2), 346-356.
- 901 Peter, B. M., & Slatkin, M. (2013). Detecting range expansions from genetic data.
902 *Evolution*, 67(11), 3274-3289.
- 903 Piñeira, J., Quesada, H., Rolán-Alvarez, E., & Caballero, A. (2008). Genetic
904 discontinuity associated with an environmentally induced barrier to gene
905 exchange in the marine snail *Littorina saxatilis*. *Marine Ecology Progress Series*, 357,
906 175-184.
- 907 Pontual (de), H., Lalire, M., Fablet, R., Laspougeas, C., Garren, F., Martin, S., Drogou,
908 M., & Woilez, M. (2019). New insights into behavioural ecology of European
909 seabass off the West Coast of France: implications at local and population
910 scales. *Ices Journal of Marine Science*, 76(2), 501-515.
- 911 Pulliam, H. R. (1988). Sources, sinks, and population regulation. *The American*
912 *Naturalist*, 132(5), 652-661.
- 913 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... &
914 Sham, P. C. (2007). PLINK: a tool set for whole-genome association and
915 population-based linkage analyses. *The American journal of human genetics*, 81(3),
916 559-575.
- 917 Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local
918 adaptation. *Nature Reviews Genetics*, 14(11), 807-820.
- 919 Sedghifar, A., Brandvain, Y., Ralph, P., & Coop, G. (2015). The spatial mixing of
920 genomes in secondary contact zones. *Genetics*, 201(1), 243-261.
- 921 Sedghifar, A., Brandvain, Y., & Ralph, P. (2016). Beyond clines: lineages and
922 haplotype blocks in hybrid zones. *Molecular ecology*, 25(11), 2559-2576.

[Introgression reveals cryptic connectivity in sea bass]

- 923 Souche, E. L., Hellemans, B., Babbucci, M., MacAoidh, E., Guinand, B., Bargelloni,
924 L., ... & Volckaert, F. A. (2015). Range-wide population structure of European sea
925 bass *Dicentrarchus labrax*. *Biological Journal of the Linnean Society*, 116(1), 86-105.
- 926 Stapley, J., Reger, J., Feulner, P. G., Smadja, C., Galindo, J., Ekblom, R., ... & Slate, J.
927 (2010). Adaptation genomics: the next generation. *Trends in ecology & evolution*,
928 25(12), 705-712.
- 929 Tigano, A., & Friesen, V.L. (2016). Genomics of local adaptation with gene flow.
930 *Molecular Ecology*, 25, 2144–64.
- 931 Tine, M., Kuhl, H., Gagnaire, P. A., Louro, B., Desmarais, E., Martins, R. S., ... &
932 Dieterich, R. (2014). European sea bass genome and its variation provide insights
933 into adaptation to euryhalinity and speciation. *Nature communications*, 5, 5770.
- 934 Vandeputte, M., Gagnaire, P. A., & Allal, F. (2019). The European sea bass: a key
935 marine fish model in the wild and in aquaculture. *Animal genetics*, 50(3), 195-206.
- 936 Van Wyngaarden, M., Snelgrove, P. V., DiBacco, C., Hamilton, L. C., Rodríguez-
937 Ezpeleta, N., Jeffery, N. W., ... & Bradbury, I. R. (2017). Identifying patterns of
938 dispersal, connectivity and selection in the sea scallop, *Placopecten magellanicus*,
939 using RAD seq-derived SNP s. *Evolutionary Applications*, 10(1), 102-117.
- 940 Varela, R. A., Rosón, G., Herrera, J. L., Torres-López, S., & Fernández-Romero, A.
941 (2005). A general view of the hydrographic and dynamical patterns of the Rías
942 Baixas adjacent sea area. *Journal of Marine Systems*, 54(1-4), 97-113.
- 943 Waples, R. S. (1998). Separating the wheat from the chaff: Patterns of genetic
944 differentiation in high gene flow species. *Journal of Heredity*, 89, 438–450.
- 945 Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation
946 of some genetic methods for identifying the number of gene pools and their
947 degree of connectivity. *Molecular Ecology*, 15(6), 1419–1439.

[Introgression reveals cryptic connectivity in sea bass]

948 Waples, R.S., Punt, A.E., & Cope J.M. (2008). Integrating genetic data into
949 management of marine resources: how can we do it better? *Fish and Fisheries*
950 2008(9), 423-449.

951 Wennevik, V., Quintela, M., Skaala, Ø., Verspoor, E., Prusov, S., & Glover, K. A.
952 (2019). Population genetic analysis reveals a geographically limited transition zone
953 between two genetically distinct Atlantic salmon lineages in Norway. *Ecology and*
954 *Evolution* , 9(12), 6901-6921.

955 Wright, P. J., Pinnegar, J. K., & Fox, C. (2020). Impacts of climate change on fish,
956 relevant to the coastal and marine environment around the UK. *MCCIP Science*
957 *Review 2020*, 354-381.

958

959

960 **FIGURES CAPTIONS**

961

962 Figure 1. Sampling map of the 827 Atlantic sea bass individuals analyzed in this
963 study. In the 'main dataset', the 827 specimens were assigned to 21 localities
964 corresponding to single ICES rectangles or groups of adjacent rectangles
965 represented with the same color. In the 'regional dataset', specimens were
966 grouped into 7 regions represented by colored polygons. Total number of samples
967 is indicated below each region name. Details about sample locations and the
968 spatially 'refined dataset' are provided in Fig. S1 and Table S1.

969

970 Figure 2. Spatial genetic variation. a: Negative correlation between standardized
971 polymorphism corrected for sampling size and latitude of the 21 sampling
972 localities. The blue line shows the linear regression (slope $p < 0.001$, $r^2 = 0.77$) and its

[Introgression reveals cryptic connectivity in sea bass]

973 0.95 confidence interval (grey shade). b: Distribution of individual coordinates on
974 RDA1 axis as a function of latitude for each of the 7 regions: Portugal (Po), South
975 Biscay (SB), North Biscay (NB), Celtic Sea (CS), Channel (Ch), Irish Sea (IS) and
976 North Sea (NS). The RDA was constrained using only spatial coordinates. For each
977 region, the boxplot horizontal line represents the mean, whiskers the standard
978 deviation and vertical lines the 0.95 confidence interval. Dots indicate outlier
979 individuals, and whisker width cover the latitudinal range of each region. c:
980 Positive correlation between genetic differentiation represented by $F_{st}/(1-F_{st})$ and
981 distance by the plateau (in km) between pairs of localities. The blue line shows the
982 linear regression (slope $p < 0.001$, $r^2 = 0.078$) and its 0.95 confidence interval (grey
983 shade). d: Results of the directionality index analysis showing the gradient in the
984 relative fit as expansion origin (color scale) and most likely origin of expansion in
985 southern Ireland (blue cross). For each of the 21 sampling localities, the inner circle
986 indicates the mean heterozygosity (grey scale) and the color of the outer circle
987 indicates the group identified with the admixture gradient analysis (blue:
988 Portugal, pink: Biscay, green: Celtic Sea - Channel - North Sea, see Fig. 3).

989

990 Figure 3. Individual ancestry proportions. Individual proportions of Mediterranean
991 ancestry inferred with ADMIXTURE (point estimation and standard error bars)
992 are displayed for the 10 reference Mediterranean samples (MED) and the 827
993 Atlantic individuals. Three groups of sampling locations that differ in their
994 average ancestry proportions (shown with the means of lower and upper values)
995 are delimited by vertical dashed bars: Portugal, Biscay, and Celtic Sea - Channel -
996 North Sea. Individuals are ranked by their specimen code number, as referred in
997 the dataset, so their ranking within the groups is meaningless.

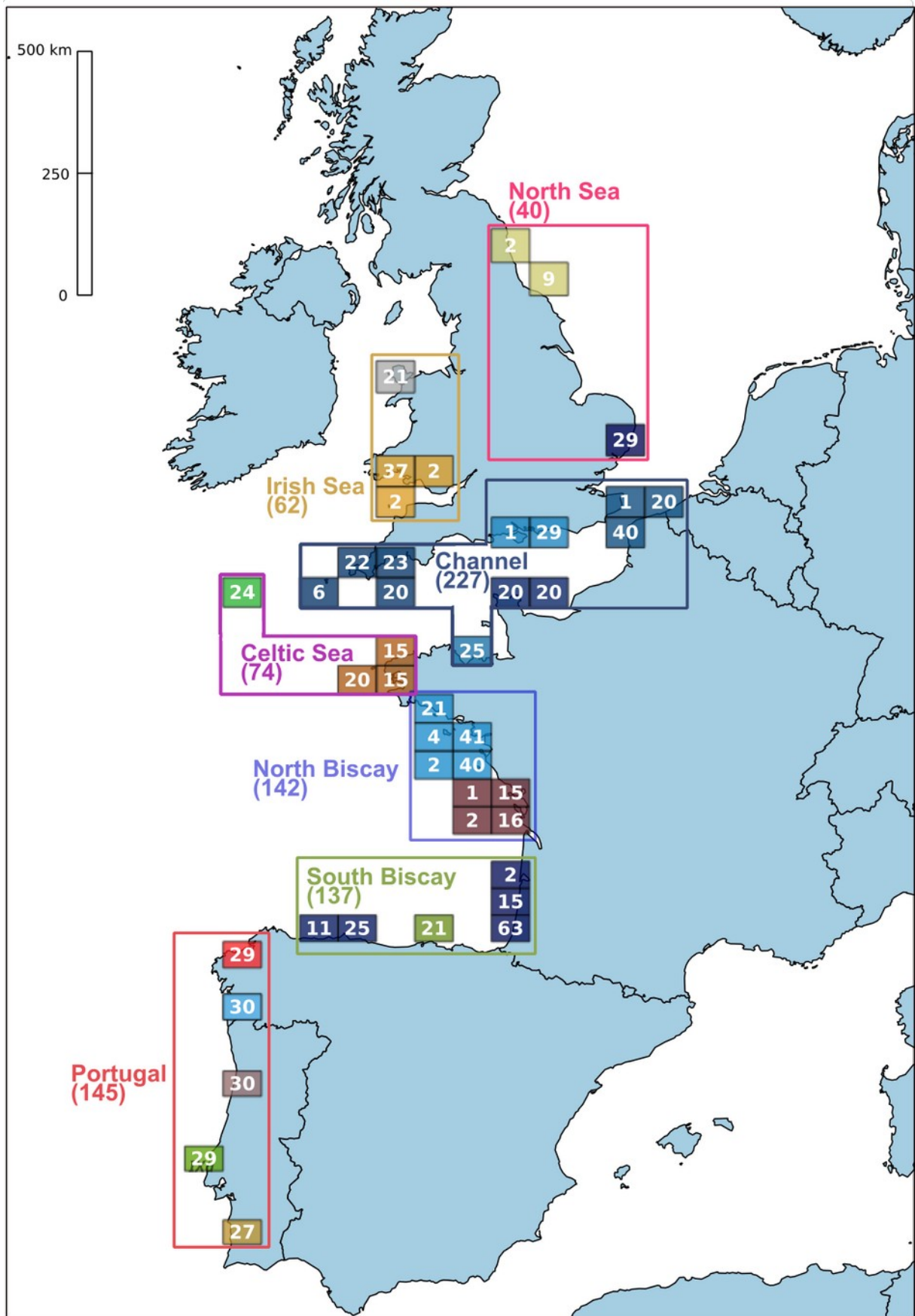
998

999 Figure 4. Determination of the positions of barriers to dispersal. a: Mean proportion
1000 of MED ancestry (shown with the means of lower and upper values) inferred with
1001 ADMIXTURE for each of the 21 localities, as a function of the distance to SINE by
1002 the plateau (km). b: Plot of the residual standard error obtained using different
1003 breakpoint values (x axis) between two linear models of mean MED ancestry as a
1004 function of distance. The position with the minimal residual standard error is
1005 directly northward to the GONB locality (red line). c: Detection of a second
1006 breakpoint directly northward to the CORU locality (red line). d: Representation
1007 of the best piecewise regression model with the two break positions determined in
1008 b and c. Colored dots: mean MED ancestry for each group of locations: Portugal
1009 (blue), Biscay (pink) and Celtic Sea - Channel - North Sea (green).

1010

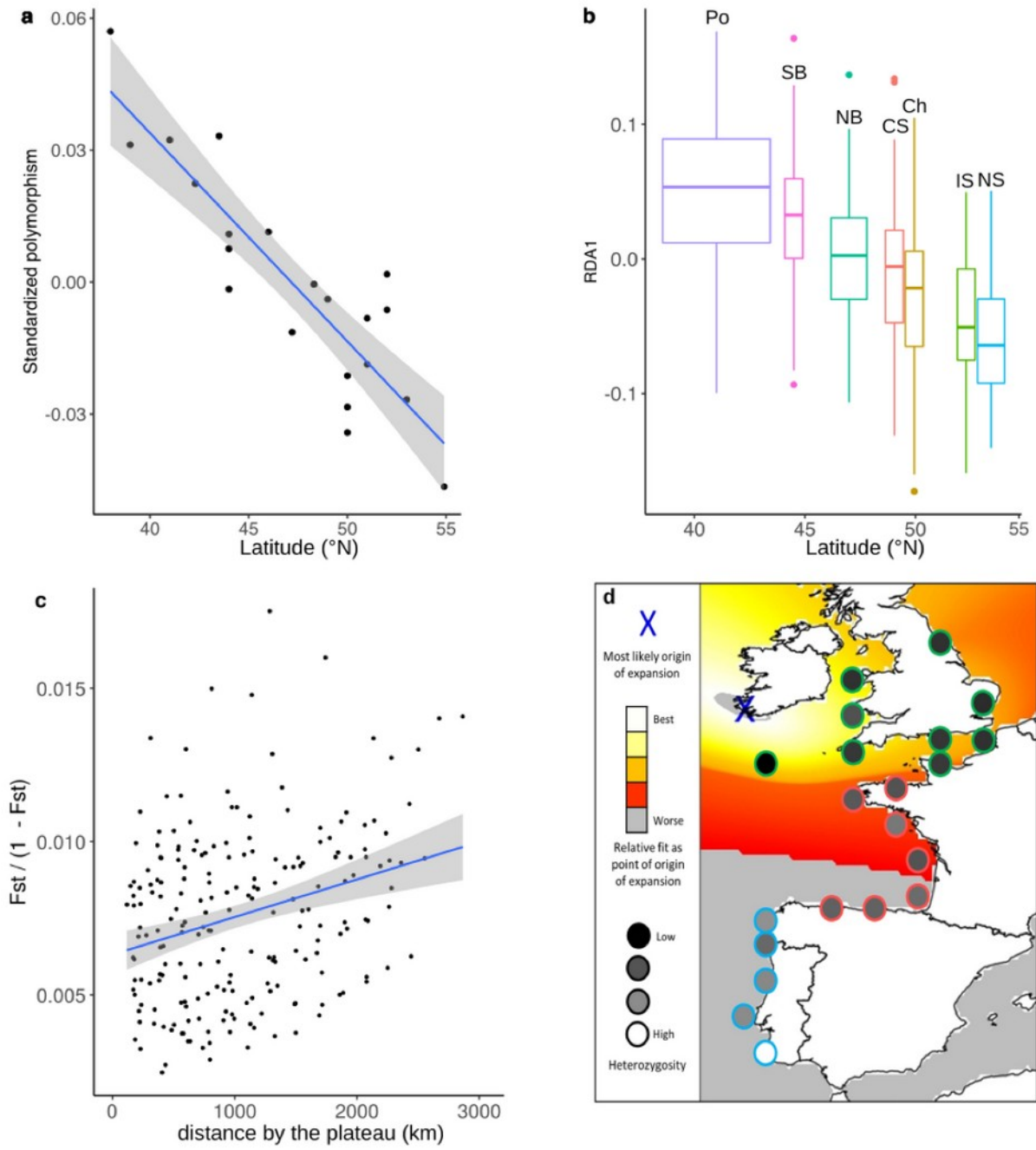
1011 Figure 5. Loci linked to genetic barriers to gene flow between ATL and MED lineages
1012 display stronger genetic differentiation within the Atlantic. a: Scaled distributions
1013 of per-locus F_{ST} values between ATL and MED lineages for non-outlier (N = 960,
1014 red density curve) and outlier SNPs (N=52, blue density curve). Candidate outliers
1015 for between-lineage divergence were identified with BayeScan using a q-value
1016 threshold of 0.01. b: Spatial profiles of mean minor allele frequency (MAF) for non-
1017 outlier (N = 960, red dots) and outlier SNPs (N=52, blue triangles), as a function of
1018 the distance to SINE by the plateau (km).

1019



1020 Figure 1.

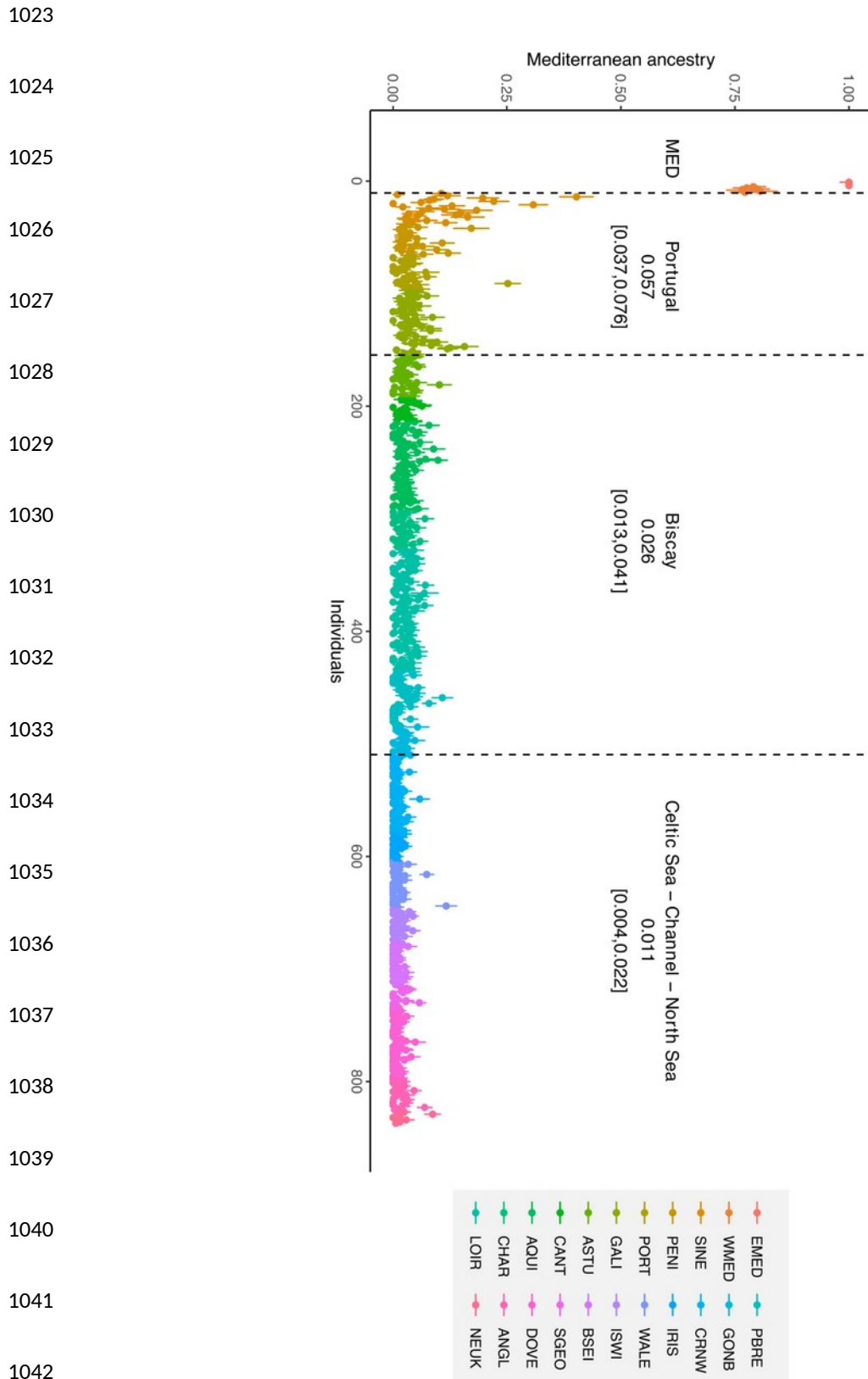
[Introgression reveals cryptic connectivity in sea bass]



1021 Figure 2.

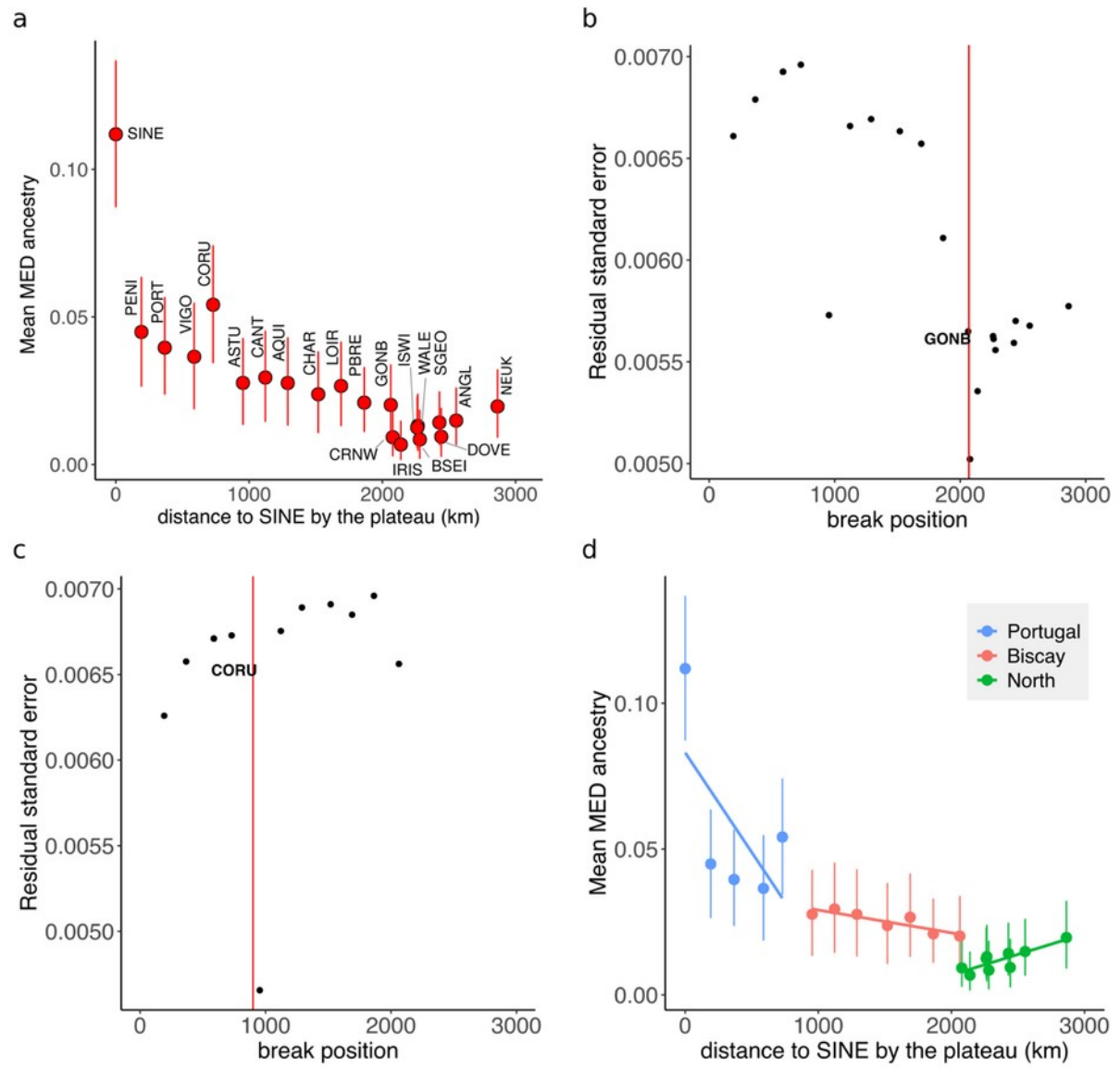
1022

[Introgression reveals cryptic connectivity in sea bass]



1044 Figure 3.

[Introgression reveals cryptic connectivity in sea bass]

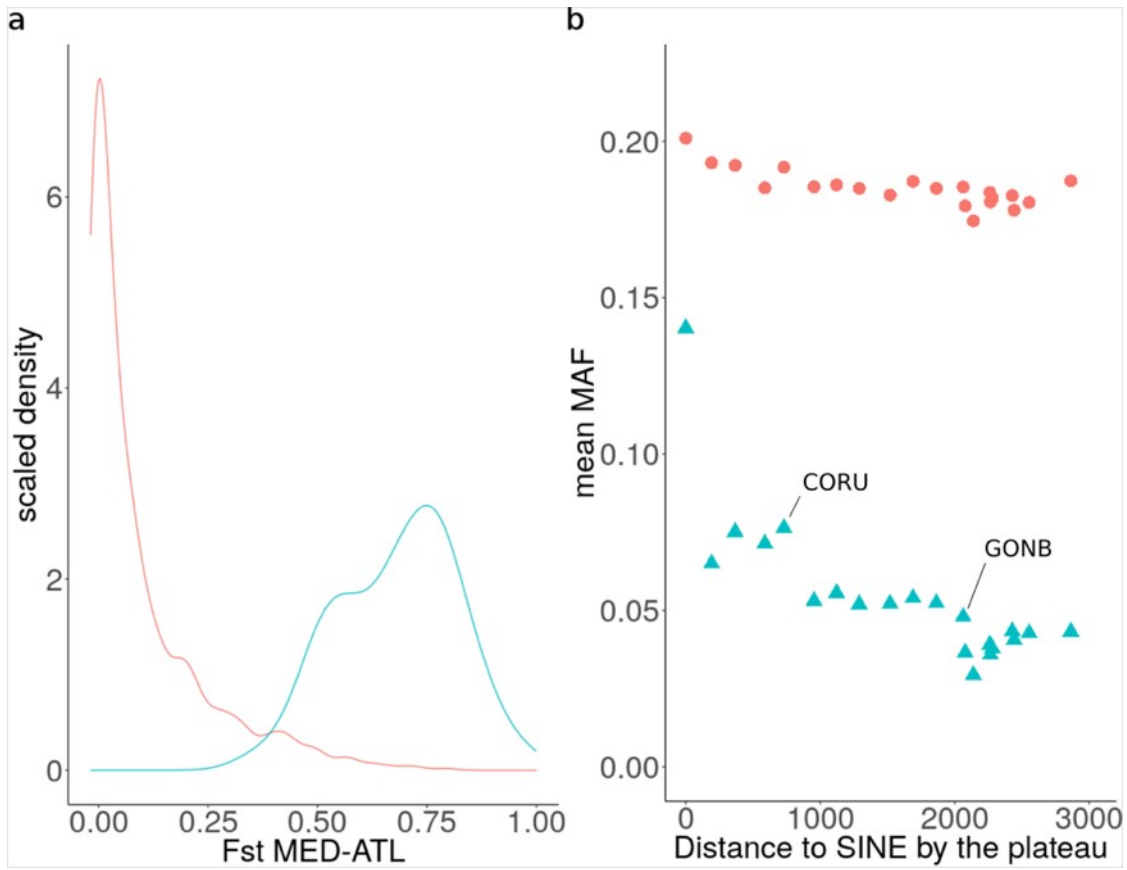


1046 Figure 4.

1047

[Introgression reveals cryptic connectivity in sea bass]

1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058



1059 Figure 5.