



HAL
open science

An introgression breakthrough left by an anthropogenic contact between two ascidians

Alan Le Moan, Charlotte Roby, Christelle Fraisse, Claire Daguin-Thiébaud,
Nicolas Bierne, Frédérique Viard

► **To cite this version:**

Alan Le Moan, Charlotte Roby, Christelle Fraisse, Claire Daguin-Thiébaud, Nicolas Bierne, et al..
An introgression breakthrough left by an anthropogenic contact between two ascidians. *Molecular Ecology*, In press, 10.1111/mec.16189 . hal-03355754v1

HAL Id: hal-03355754

<https://hal.science/hal-03355754v1>

Submitted on 27 Sep 2021 (v1), last revised 3 Aug 2023 (v2)

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.

1 Accepted in Molecular Ecology (Sept. 2021)

2 Available on BioRxiv: doi: <https://doi.org/10.1101/2021.08.05.455260>

3

4

5 **An introgression breakthrough left by an anthropogenic contact between two**
6 **ascidians**

7 Alan Le Moan^{1,2,*}, Charlotte Roby¹, Christelle Fraisse³, Claire Daguin-Thiébaud¹, Nicolas
8 Bierne⁴, Frédérique Viard^{1,4,*}

9 ¹ Sorbonne Université, CNRS, UMR 7144, Station Biologique de Roscoff, Place Georges Teissier,
10 29680 Roscoff, France

11 ² Department of Marine Sciences, Tjärnö Marine Laboratory, University of Gothenburg,
12 Laboratorievägen 10, 452 96 Strömstad, Sweden

13 ³ CNRS, Univ. Lille, UMR 8198 – Evo-Eco-Paleo, F-59000 Lille, France

14 ⁴ ISEM, Univ. Montpellier, CNRS, EPHE, IRD, Montpellier, France

15 *corresponding authors: alan.le.moan@gmail.com, frederique.viard@umontpellier.fr

16

17 **Running title**

18 Introgression footprint of a sea squirt invasion

19 **Abstract**

20 Human-driven translocations of species have diverse evolutionary consequences such as
21 promoting hybridization between previously geographically isolated taxa. This is well-
22 illustrated by the solitary tunicate, *Ciona robusta*, native to the North East Pacific and
23 introduced in the North East Atlantic. It is now co-occurring with its congener *C. intestinalis* in
24 the English Channel, and *C. roulei* in the Mediterranean Sea. Despite their long allopatric
25 divergence, first and second generation crosses showed a high hybridization success between
26 the introduced and native taxa in the laboratory. However, previous genetic studies failed to
27 provide evidence of recent hybridization between *C. robusta* and *C. intestinalis* in the wild.
28 Using SNPs obtained from ddRAD-sequencing of 397 individuals from 26 populations, we
29 further explored the genome-wide population structure of the native *Ciona* taxa. We first
30 confirmed results documented in previous studies, notably i) a chaotic genetic structure at
31 regional scale, and ii) a high genetic similarity between *C. roulei* and *C. intestinalis*, which is
32 calling for further taxonomic investigation. More importantly, and unexpectedly, we also
33 observed a genomic hotspot of long introgressed *C. robusta* tracts into *C. intestinalis* genomes
34 at several locations of their contact zone. Both the genomic architecture of introgression,
35 restricted to a 1.5 Mb region of chromosome 5, and its absence in allopatric populations
36 suggest introgression is recent and occurred after the introduction of the non-indigenous
37 species. Overall, our study shows that anthropogenic hybridization can be effective in
38 promoting introgression breakthroughs between species at a late stage of the speciation
39 continuum.

40 **Keywords:** Biological introductions, non-indigenous species, anthropogenic hybridization,
41 introgression hotspots, population genomics, tunicates

42 Introduction

43 Anthropogenic hybridizations can arise from human-mediated translocations of
44 species outside their natural distribution range, which promote secondary contact between
45 previously geographically isolated taxa (McFarlane & Pemberton, 2019). Such circumstances
46 enable the initial phase of the hybridization to be studied in a natural context (Grabenstein &
47 Taylor, 2018, Hufbauer et al., 2012, Faust, Halvorsen, Andersen, Knutsen, & André, 2018;
48 Popovic, Matias, Bierne, & Riginos, 2020). Anthropogenic hybridizations thus provide unique
49 opportunities to examine gene flow between species experiencing incomplete reproductive
50 isolation, even at a late stage of the speciation process (Viard, Riginos, & Bierne, 2020).

51 Species introductions are common and occur at increasing rates in the marine
52 realm (Seebens et al., 2017). Ports and marinas, one component of the increasing marine
53 urbanization, are one point-of-entry of many non-native species (Firth et al., 2016), where
54 they can co-occur with native congeners (e.g., Bouchemousse, Lévêque, Dubois, & Viard,
55 2016b). Consequently, these habitats are prone to facilitating hybridization between
56 introduced and native species. They provide a suitable system to examine secondary gene
57 flow in case of anthropogenic hybridization. This is illustrated by a recent study from Simon et
58 al. (2020), which documented the presence of a singular lineage, named “dock mussels”,
59 originating from a recent admixture between *Mytilus edulis*, native to the North Atlantic, and
60 *Mytilus galloprovincialis*, native to the Mediterranean Sea. These admixed populations are
61 restricted to port habitats in European waters.

62 The secondary contact between the solitary tunicates *Ciona robusta* and *Ciona*
63 *intestinalis* in the English Channel is another case study, but with very different outcomes.
64 *Ciona robusta* (formerly known as *C. intestinalis* type A; Gissi et al., 2017), native to Asia, was

65 introduced in the early 2000s to the English Channel (Bouchemousse, Bishop, & Viard, 2016a),
66 the native range of *C. intestinalis* (formerly known as *C. intestinalis* type B; Gissi et al., 2017).
67 The two species are found in syntopy (Nydam & Harrison, 2010, Bouchemousse et al., 2016b),
68 display similar life-cycles (Bouchemousse, Lévêque, & Viard, 2017), and can be easily crossed
69 in the laboratory (Bouchemousse et al., 2016b; Malfant, Coudret, Le Merdy, & Viard, 2017)
70 despite their high molecular divergence (12% of net synonymous divergence, Roux,
71 Tsagkogeorga, Bierne, & Galtier, 2013, Roux et al., 2016). Successful hybridization nonetheless
72 occurs in one direction only, with *C. intestinalis* as the maternal lineage (Bouchemousse et al.,
73 2016b; Malfant, Darras, & Viard, 2018). Demographic inferences based on few individuals but
74 high number of markers derived from 852 coding sequences (total length of 270kb) suggested
75 the presence of several introgression hotspots between the two species (Roux et al., 2013).
76 However, by using 100 ancestry-informative SNPs, this introgression was later shown to be
77 the outcome of past introgression, far preceding the contemporary secondary contact
78 nowadays observed in the English Channel (Bouchemousse, Liautard-Haag, Bierne, & Viard,
79 2016c). In Europe, where both species occur in sympatry, Bouchemousse et al. (2016c) have
80 found limited evidence for hybridization (i.e., one F1 hybrid out of 449 individuals), and no
81 sign of contemporary introgression (i.e., no F2s or backcrosses). Thus, despite a high
82 hybridization potential, efficient reproductive barriers seem to prevent hybridization in the
83 wild between the native and non-native species. Although these results are based on low
84 genomic coverage, they suggest that introgression between *C. intestinalis* and *C. robusta* is far
85 less common than in *Mytilus* species. Indeed, admixture was effectively detected in dock
86 mussels using similar genomic coverage (Simon et al., 2020). Nevertheless, high genomic
87 coverage can reveal subtler introgression patterns as exemplified in model systems in

88 speciation (sticklebacks: Ravinet et al., 2018; rock periwinkle: Stankowski et al., 2020;
89 drosophila: Turissini & Matute, 2017), as well as in native-invasive systems (cotton bollworm:
90 Valencia-Montoya et al., 2020; honey bee: Calfee, Agra, Palacio, Ramírez, & Coop, 2020).

91 In this study, we further explored the genome-wide population structure of the
92 native tunicate *C. intestinalis* in the North Atlantic using a large number of SNPs provided by
93 a ddRAD-sequencing approach. Our study expands upon the work conducted by Hudson et al.
94 (2020), which described multiple glacial lineages of *C. instestinalis* within Europe. Here, we
95 aim to evaluate the consequences of anthropogenic hybridization with its congener *C. robusta*
96 that has been introduced in the range of one *C. intestinalis* glacial lineage, in the English
97 Channel. As a control, and for the sake of comparison, we also examined one population of
98 the native species *Ciona roulei* from the Mediterranean Sea. The species status of *C. roulei* has
99 been repeatedly questioned (Lambert, Lafargue, & Lambert, 1990; Nydam & Harrison, 2010;
100 Malfant et al., 2018), and it might better be described as an isolated population of *C.*
101 *intestinalis*. Interestingly, *C. roulei* can be found in sympatry with *C. robusta*, also introduced
102 in the Mediterranean Sea. Based on genome-wide SNPs, we 1. recover the population
103 structure described from previous studies for *C. intestinalis* both at fine and large geographical
104 scales 2. provide genome-wide support for a revision of the taxonomic status of *C. roulei*, and
105 3. provide the first evidence in favor of recent introgression events from *C. robusta* towards
106 *C. intestinalis* in their contact zone, but not in allopatric populations. However, introgression
107 is restricted to a 1.5 Mb region of chromosome 5. Overall, our study shows that anthropogenic
108 hybridization can be effective in promoting gene flow even between species at a late stage of
109 speciation, but in this case introgression can be restricted to localized breakthroughs in the
110 receiving genome.

111

112 **Materials and methods**

113 *Sample collection*

114 We studied 397 individuals of *Ciona* spp., previously sampled across the North
115 Atlantic. The sampling locations are shown in Figures 1B (fine-scale) and 2A (large-scale) with
116 details provided in Table S1 in the Supporting Information file. Most individuals (N=346) were
117 *C. intestinalis* sampled from 22 locations in 2012 by Bouchemousse et al. (2016a), except for
118 one site (Jer) that was sampled in 2014 by Hudson, Viard, Roby, & Rius (2016). This sampling
119 includes two locations (REK, Iceland, and NAH, US) where *C. intestinalis* is most likely
120 introduced although its status remains debated (i.e., cryptogenic; See Appendix 1 in
121 Bouchemousse et al., 2016a). The sampling scheme aims at covering the known geographic
122 range of this species, with a focus on the English Channel where the species coexists with its
123 introduced congener *C. robusta*. In addition, 19 specimens of *C. roulei*, native to the
124 Mediterranean Sea, were included, along with 32 individuals from the introduced species *C.*
125 *robusta*, of which 16 were sampled from the Mediterranean Sea (in sympatry with *C. roulei*)
126 and 16 from the English Channel (in sympatry with *C. intestinalis*).

127 *DNA extraction and library preparation*

128 For each individual, DNA was extracted using Nucleospin® 96 Tissue Kit according
129 to the manufacturer's protocol (Macherey-Nagel, Germany). Individual double-digest RAD-
130 seq libraries were constructed according to a protocol slightly modified from Brelsford,
131 Dufresnes, & Perrin (2016). A detailed step-by-step protocol is available at
132 [dx.doi.org/10.17504/protocols.io.bv4tn8wn](https://doi.org/10.17504/protocols.io.bv4tn8wn). Briefly, DNA was digested with PstI and MseI
133 after fluorometric quantification of DNA concentration with PicoGreen (Invitrogen, Carlsbad,

134 CA, USA) and normalization of the extracts. Each individual was labelled with a unique
135 barcode-index combination, with an inline barcode (incorporated in the PstI adaptor) and an
136 Illumina Truseq index (incorporated during the PCR carried out on the ligation products). Size
137 selection was carried out with 1.5% agarose cassettes in a pippin prep (Sage Science) to select
138 fragments between 280 and 600 base pairs. A total of three pooled libraries were sequenced,
139 each containing 184 individuals, with replicates (two individuals per library and two across the
140 three libraries). Each library was sequenced in two lanes of an Illumina HiSeq 2500 v4 high
141 throughput flow cell generating 125 base single-end reads at Eurofins Genomics (Ebersberg,
142 Germany).

143 *Bioinformatics pipeline*

144 The reads were demultiplexed based on their individual index-barcode with the
145 processRADtags programme of Stacks v2 (Rochette, Rivera- Colón, & Catchen, 2019). Reads
146 with ambiguous barcodes or low sequencing quality were removed. Overall, an average of
147 1.7M good quality reads per sample was kept for the downstream analyses. The reads were
148 trimmed to 80 base pairs, and mapped on the *C. robusta* genome (KH79 version, Dehal et al.,
149 2002, NCBI assembly GCF_000224145.1) using the default parameters implemented in BWA
150 software (Li & Durbin, 2009). Note that this genome is improperly referenced under the name
151 *C. intestinalis* in several databases because of a recent taxonomic revision (Gissi et al., 2017,
152 and references herein). Between 60 and 67% of the reads were mapped for *C. roulei* and *C.*
153 *intestinalis* individuals, and between 83 and 87% of the reads for *C. robusta* individuals. The
154 aligned RAD data were then processed using the reference mapping pipeline in Stacks v2 set
155 with default parameters (Rochette et al., 2019). Only the SNPs sequenced for at least 80% of
156 the individuals within a location, present in all the populations and with a maximum

157 heterozygosity of 80% were called using the *population* function. Additional filtering steps
158 were performed using vcfTools (Danecek et al., 2011) in order to remove SNPs with a minor
159 allele count of two or showing significant deviations from Hardy-Weinberg equilibrium (P-
160 value threshold of 0.05) in more than 60% of the population, with the function
161 *filter_hwe_by_pop.pl* implemented in dDocent pipeline (Puritz, Hollenbeck, & Gold, 2014). We
162 additionally removed all polymorphisms private to *C. robusta* populations, as such
163 polymorphisms are neither informative to describe *C. intestinalis* population structure nor to
164 evaluate the extent of introgression across species (which results in shared polymorphisms).
165 At the end of the filtration process, the dataset included 397 individuals (see Table S1 across
166 sampled locations) genotyped at 51,141 SNPs (17,280 with a Minor Allele Frequency (MAF)
167 above 5%) derived from 5,599 RAD-locus with an average depth across all samples of 59 reads
168 per locus, which was then exported into variant call format (VCF). The VCF was then
169 statistically phased using Beagle v5.2 (Browning & Browning, 2007) in order to extract
170 haplotypes and conduct phylogenetic analyses as explained below.

171 *Population structure analyses*

172 **Fine-scale population structure (dataset 1):** Because of our interest in identifying
173 introgression in the contact zone, 280 individuals of *C. intestinalis* from 18 sampling sites (70%
174 of the data) were obtained from the English Channel, Iroise Sea and the Bay of Biscay. We first
175 explored the fine-scale population structure of these populations (dataset 1; Table S1; Figure
176 1B). The dataset 1 was further filtered to keep SNPs with a MAF above 5%, and additionally
177 thinned by keeping one random SNP over bin of 1kb to take into account physical linkage,
178 using vcfTools (Danecek et al., 2011). In total, 13,603 linked SNPs and 3,510 unlinked SNPs
179 were used to study the fine-scale population structure. For the unlinked SNPs, we used the

180 function `find.clusters` in `adegenet` (Jombart, Devillard, & Balloux, 2010) to find the best
181 number of clusters (lower BIC value) describing the population structure on the 50 first
182 principal components of a PCA (Figure S1). These clusters were then used as discriminant
183 factors to compute a discriminant analysis of principal components (DAPC) with two
184 discriminant functions. We then used the `snmf` function of the R package LEA (Frichot &
185 François, 2015), using the number of cluster inferred from the `find.clusters` function in
186 `adegenet` to examine the admixture proportions within each location. We estimated pairwise
187 F_{ST} values with 95% confidence interval among sampling sites by bootstrapping (10,000
188 replications) using the R package `StAMPP` (Pembleton, Cogan, & Froster, 2013) and
189 significance was tested after accounting for multiple testing with Bonferroni correction. The
190 linked SNPs were used to compute another DAPC using the group inferred by the function
191 `find.clusters` on the unlinked SNPs, extract the eigenvalue of each individual SNP, and evaluate
192 the genomic distribution of the markers responsible for the population structure.

193 **Large-scale population structure (dataset 2):** For the large-scale population structure
194 analyses, and to achieve a more balanced sampling structure, we kept only one sampling
195 location per genetic cluster inferred with the fine-scale analyses (see above). In total, 129
196 individuals from nine sampling sites for *C. intestinalis*, 32 individuals from two sampling sites
197 for *C. robusta* and 19 individuals from one sampling site for *C. roulei* were included (dataset 2;
198 Table S1; Figure 2A). This second dataset was filtered to keep SNPs with a MAF above 5%, and
199 further filtered for physical linkage (one random SNP per kb). Overall, 17,138 linked and 3,828
200 unlinked SNPs were used depending on the analysis. Using the unlinked SNPs, we conducted
201 a PCA analysis from the R package `adegenet` (Jombart & Ahmed, 2011) and admixture analyses
202 (detailed on Figure S2) from the R package LEA (Frichot & François, 2015). We computed the

203 pairwise F_{ST} value with 95% confidence interval from bootstrapping (10,000 replications) using
204 the R package StAMPP (Pembleton et al., 2013), and significance was tested after accounting
205 for multiple testing using Bonferroni correction. A second PCA analysis was performed on the
206 linked SNPs and the eigenvalue of each SNPs were extracted and plotted against their physical
207 position on the reference genome in order to characterize the genomic distribution of the
208 markers responsible for the population structure. Finally, we used the statistically phased VCF
209 file to create a pseudo sequence of all SNPs, and transformed it into a fasta file containing two
210 haplotypes per individual using a custom R script available in the Zenodo archive (see the Data
211 Accessibility section). From this fasta file, we computed the pairwise genetic distance between
212 each haplotype, and represented a neighbor joining tree based on the GTR substitution model
213 using the R package phangorn (Schliep, 2011).

214 **Investigating introgression between *Ciona intestinalis* and *C. robusta* (dataset 3):** To assess
215 the variability of introgression along the genome, we first used the R package LEA using $k=2$
216 to compute the individual ancestry of *C. robusta* and *C. intestinalis* individuals, over the whole
217 genome, and for each chromosome independently. Pairwise F_{ST} values between the two
218 sampling locations (pooled) of *C. robusta* and each other locations (*C. intestinalis*) were then
219 computed for each SNP following the method of Weir and Cockerham (1986) using vcfTools
220 (Danecek et al., 2011). For each chromosome, we calculated the maximum F_{ST} value taken
221 over a sliding window of 100kb, which was then smoothed with the R package ggplot2
222 (Wickham, 2011). Additionally, we created an ancestry informative set of SNPs by extracting
223 the markers differentially fixed (i.e., F_{ST} value of one) between *C. robusta* and the Gul
224 population (6,849 SNPs in total). The later population comes from a region and an

225 environment (deep natural habitat) where *C. robusta* have never been reported, thus, is the
226 least likely population to have recently hybridized with *C. robusta*.

227 To detect potential introgression tracts of *C. robusta* within the genome of *C.*
228 *intestinalis*, we extracted the F_{ST} value of the 6,849 ancestry informative SNPs calculated
229 between *C. robusta* and each 22 *C. intestinalis* populations (Rek excluded due to low sampling
230 size). We then used the Hidden Markov Model (HMM) developed by Hofer, Foll & Excoffier
231 (2012) to infer the position of genomic islands. Briefly, the HMM characterizes and sorts
232 genomic regions according to their level of differentiation, and is generally used to detect
233 island of divergence (e.g. Soria-Carrasco et al., 2014, Shi et al., 2021). Here, the HMM was
234 applied to detect regions of introgression by contrasting regions with high-background
235 differentiation (with an F_{ST} value of one) from regions with intermediate differentiation (with
236 F_{ST} normally distributed around the 15% lower quantile) and low differentiation (with F_{ST}
237 normally distributed around the 5% lower quantile). The HMM was run using a modify version
238 of the R script by Marques et al. (2016). To avoid any biases by comparing introgressed
239 population from non-introgressed populations, the quantiles were extracted from the
240 distribution of all F_{ST} values calculated for the 22 pairwise comparisons. Regions of low
241 differentiation covering more than 4 consecutive SNPs, and regions of intermediate
242 differentiation covering more than 10 consecutive SNPs in a given pair of *C. robusta* and *C.*
243 *intestinalis* were considered as candidate for introgression. Finally, the genotypes of all the
244 individuals at each diagnostic SNP was visualized using a modified version of mk.image of the
245 R package introgress (Gompert & Buerkle, 2010) developed by Simon et al. (2021). This
246 analysis was performed independently along each chromosome.

247

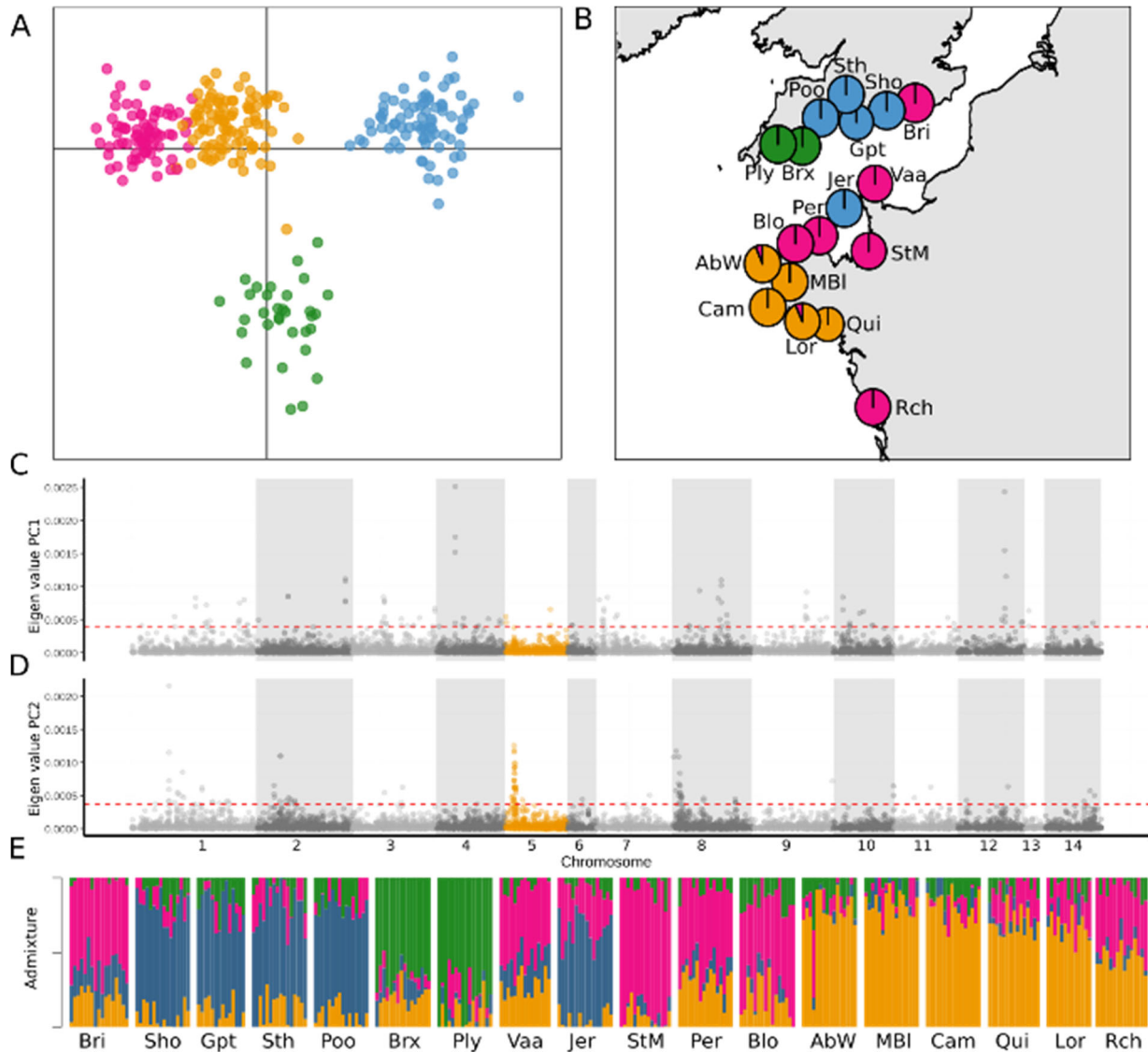
248 **Results**

249 *Fine-scale population structure of *Ciona intestinalis* in France and UK*

250 Four genetic clusters, which explained 25% of the variance in the DAPC, were
251 identified in the English Channel, Iroise Sea and Bay of Biscay (Figure 1A). The genetic
252 clustering of the sampling sites was consistent with their geographical proximity, except for
253 three of them grouping with geographically distant ones (Figure 1B): 1) the individuals from
254 Jer (Jersey), geographically close to the northern Brittany populations (deep pink cluster) but
255 clustered with eastern UK sites (blue cluster), 2) the individuals from Rch (La Rochelle) and Bri
256 (Brighton), which belong to the northern Brittany cluster, despite being geographically closer
257 to Western France (yellow cluster) and eastern UK (blue cluster), respectively. This mosaic
258 structure was also supported by the LEA analysis (Figure 1E), notably for Bri and Jer. However,
259 Rch showed evidence of admixture between nearby locations (Qui and Lor in Western
260 Brittany) and distant ones (Per and Blo in Northern Brittany). Additionally, the two
261 westernmost English populations (Ply and Brx) were genetically differentiated from all the
262 other sampling sites (green cluster in Figure 1B, admixture analysis, Figure 1E).

263 The SNPs contributing to the population structure observed along the first axis
264 were distributed genome-wide (Figure 1C). Conversely, the SNPs structuring the second axis
265 (and thus differentiating the westernmost English populations) were over-represented at the
266 start of two chromosomes (5 and 8, Figure 1D). Pairwise F_{ST} values were overall low ($0.001 <$
267 $F_{ST} < 0.028$) but significant across most sites (Table S2), except among four sites of northern
268 and western Brittany (AbW, MBI, Cam, and Lor), and among four sites of the eastern UK (Sth,

269 Sho, Gpt and Poo). The highest values of F_{ST} (0.028) were found among the most distant sites
 270 following the coastline, i.e. Rch versus several sampling sites in the UK (Ply, Sho, Poo).



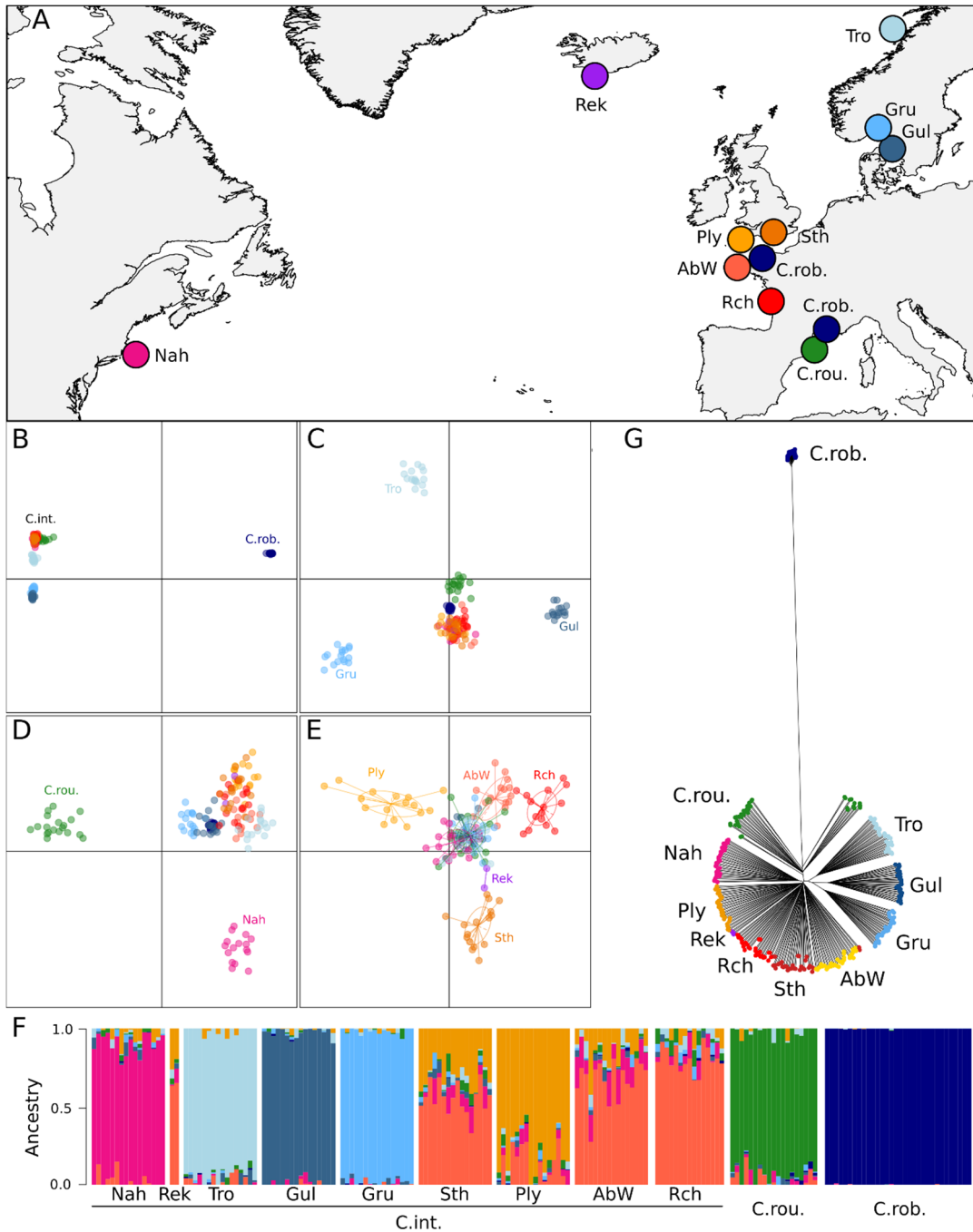
271
 272 **Figure 1:** Fine-scale structure analysis of the 18 populations of *C. intestinalis* sampled in Bay of Biscay,
 273 Iroise Sea and English Channel. Further details about the sampling sites and the code names are
 274 provided in Table S1. Figure 1A shows the clustering of 280 individuals genotyped at 3,510 unlinked
 275 SNPs with a DAPC set for $K=4$ clusters identified by the find.clusters function in adegenet (Jombart et
 276 al. 2010), each cluster is pictured by a different color. The same color scheme is used in the other plots.
 277 Figure 1B shows the proportion of individuals per sampling site assigned to each cluster with the DAPC.
 278 Figure 1C and 1D display the contribution (eigenvalue) of the 13,603 linked-SNP to the first and second
 279 axes of the DAPC, respectively; chromosome 5, which is carrying an introgression hotspot, is
 280 highlighted in orange; the dashed red lines show the 95% quantiles above which are located the top
 281 5% eigenvalues. Figure 1E shows the outcome of an admixture analysis made with the 3,510 unlinked
 282 SNPs using the *snmf* function of the LEA package for $K=4$ clusters; the ancestry proportion to each
 283 cluster is indicated for each individual, which are sorted according to their locations.

284
285 *Large-scale population structure across the Northern Atlantic:*

286 The non-indigenous individuals of *C. robusta* are highly divergent from the
287 individuals of the two species native to European waters, *C. intestinalis* and *C. roulei*, as
288 illustrated on the first axis of the PCA, which explained 51.80% of the inertia (Figure 2B).
289 Populations with less genetic divergence are then distinguished in the following axes: the
290 northern European samples of *C. intestinalis* (Gru, Gul and Tro, in blue) are distinguished from
291 the other locations, and from *C. roulei*, by the second axis (2.56%, figure 2B). The following
292 axes (3 to 8) distinguished individuals sampled from different locations but with much smaller
293 and decreasing inertia, from 1.65% to 0.52% (Figure 2C-E). *C. roulei* was distinct from *C.*
294 *intestinalis* only along the axis 5 (1.29%, Figure 2D). The two sampling sites Rek and Nah, in
295 which *C. intestinalis* has an undetermined status (putatively introduced), were very similar to
296 populations from the English Channel and Bay of Biscay, being distinguished on axes six and
297 eight (Figure 2D,E).

298 The SNPs contributing to the major divergence between *C. robusta* and the two
299 native species *C. intestinalis* and *C. roulei* were distributed genome-wide (i.e., Figure S3).
300 However, at the start of chromosome 5, a reduction of the divergence between *C. robusta*
301 and *C. intestinalis* was observed, as shown by a slight decline of eigenvalues in the PC1. The
302 population structure depicted by the PCA was corroborated by the admixture analysis (Figure
303 2F, Figure S2) showing high support for all the clusters described from the axes one to six of
304 the PCA (Figure 2B-E). All the pairwise F_{ST} were significantly different from 0 except for two
305 comparisons (between the two *C. robusta* sites, and between Rek and Sth, Table S3). The F_{ST}
306 values ranged from 0.02 to 0.169 among *C. intestinalis* sampling sites, which is similar to the
307 range (from 0.069 to 0.166) observed between *C. roulei* and any of the *C. intestinalis*

308 populations. The two sites Nah and Rek showed the lowest F_{ST} values with the populations
309 from the English Channel (e.g., F_{ST} Nah vs AbW = 0.033 and F_{ST} Rek vs Sth non-significant from
310 0), and were less differentiated than pairwise comparisons between northern and southern
311 North Atlantic sites. Very large values (i.e., 0.761 to 0.813) were observed between *C. robusta*
312 and *C. intestinalis*/*C. roulei* (Table S3). The deep divergence of *C. robusta* from the other
313 populations, and that each population clustered in separate groups, is also confirmed by the
314 phylogenetic tree (Figure 2G). *C. roulei* individuals formed a distinct group closely related to
315 *C. intestinalis*.



316

317 **Figure 2:** Large-scale population structure analysis of *Ciona intestinalis* (*C. int.*; native to Europe)
 318 populations, *C. roulei* (*C.rou.*; native to Europe) and *C. robusta* (*C. rob.*, introduced to Europe). The
 319 locations examined are shown in Figure 2A. The location name, and further details, associated to each
 320 code are provided in Table S1. The colour code used to picture each sampling location in A) is used in
 321 the other plots. PCA plots, based on 180 individuals genotyped at 3,828 unlinked SNPs, are shown in
 322 figures 2B to 2E, which are displaying 8 different axes (B: 1 vs. 2; C-E: 3 vs. 4, 5 vs. 6, 7 vs. 8), associated
 323 to 51.80, 2.56, 1.65, 1.45, 1.29, 0.75, 0.55 and 0.52% of the total inertia, respectively. Figure 2F displays
 324 the outcome of the admixture analysis made with the same SNPs using the *snmf* function of the LEA
 325 package for a *k*=8 clusters (detailed on cluster selection available in Figure S2); the ancestries of each
 326 individual are sorted by species and, for *C. intestinalis*, by sampling locations (left part). A neighbor-
 327 joining tree of 360 phased haplotypes built with the 17,138 linked SNPs is pictured in figure 2G.

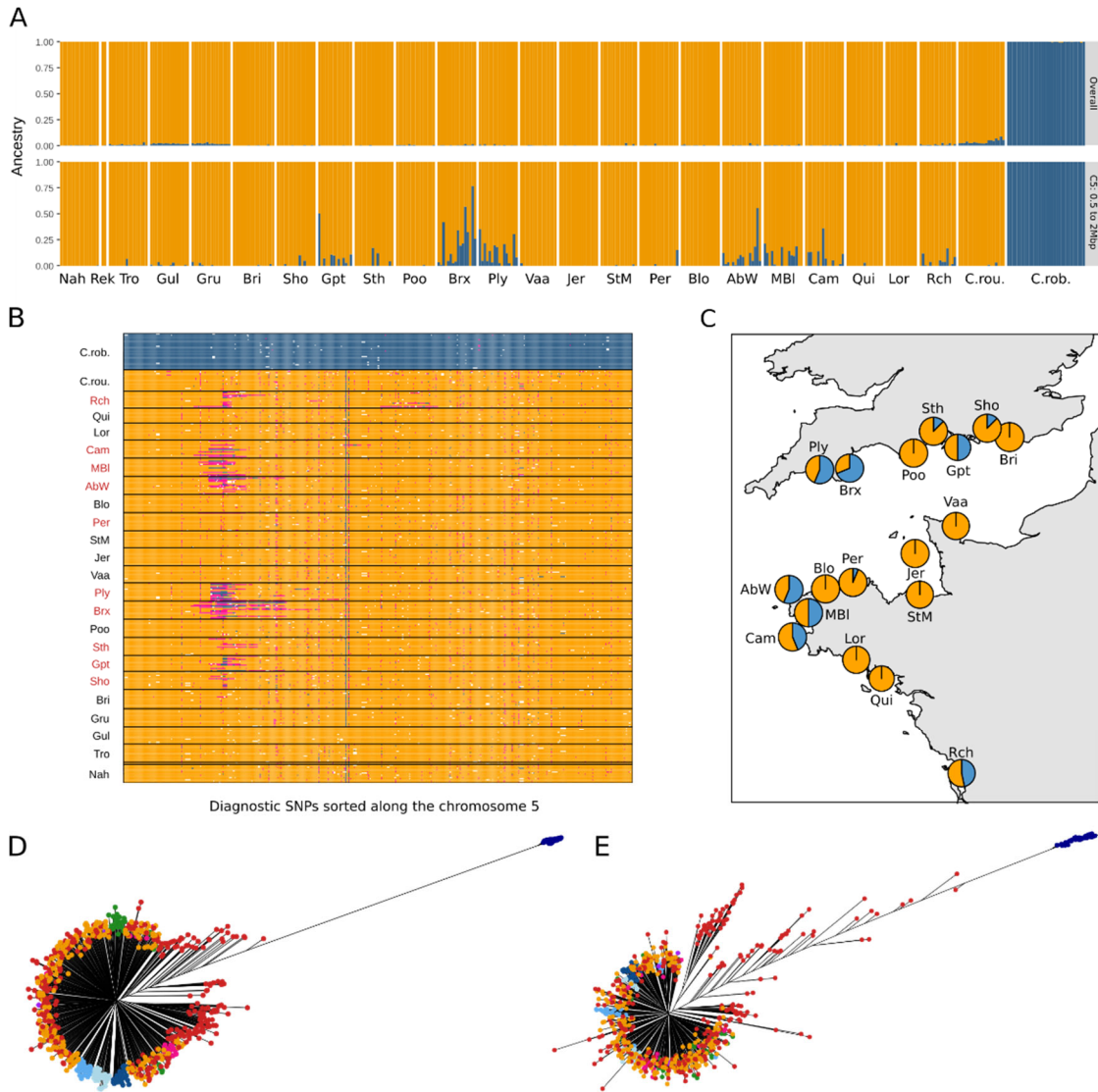
328 *Introgression between C. robusta and C. intestinalis*

329 *Ciona robusta* and *C. intestinalis* had consistent high divergence across the
330 genome (Figure S4), with ~39.64% of the SNPs, with a MAF above 5%, being diagnostic using
331 Gul population as reference ($F_{ST}=1$). No sign of genome-wide admixture was detected
332 between the two species (Figure 3A, top panel), but the *C. roulei* individuals appeared admixed
333 with a *C. robusta* ancestry ranging from 1.96 to 8.77% (Figure 3A, top-panel).

334 The same results were obtained when each chromosome was analyzed
335 independently (Figure S5), with one noticeable exception found on chromosome 5. On this
336 chromosome, 82 *C. intestinalis* individuals showed a signal of admixture with *C. robusta* (up
337 to 8.62%). Chromosome 5 was also the only chromosome where regions of introgression with
338 a low differentiation between *C. intestinalis* and *C. robusta* at diagnostic SNPs were detected
339 by the HMM, all of which being located between 0.61 and 1.58 Mb (tracts sizes ranging from
340 64.83kb to 0.49Mb, Table S4). These regions were found in nine populations, and shared an
341 80kb fragment located from 0.81 to 0.88 Mb of chromosome 5. Large regions of intermediate
342 differentiation, with tracts sizes ranging from 40.04kb to 1.25Mb, were also found only on
343 chromosome 5, with 90% of them located around the region of low differentiation (from 0.40
344 to 2.22Mb). In this portion of chromosome 5, *C. robusta* ancestry among *C. intestinalis*
345 individuals reached up to 76.77% (Figure 3A - bottom panel).

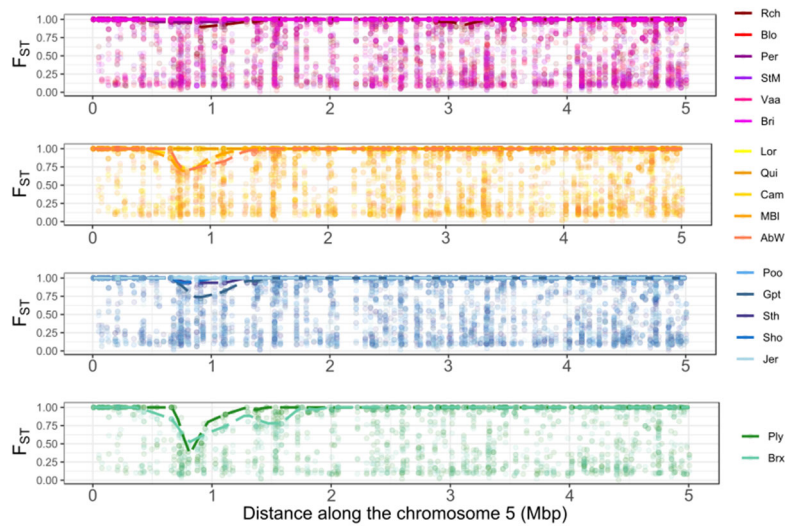
346 The presence of admixed individuals were detected at sites located in the contact
347 zone in the English Channel, Iroise Sea and Bay of Biscay (Figure 3B,C, Table S4). In agreement
348 with the HMM analysis, chromosome-wide F_{ST} values calculated between those populations
349 and *C. robusta* showed a striking decline in F_{ST} on chromosome 5 (Figure 4). The most extreme
350 drop was localized around 0.87Mb, i.e. within the 80kb region of low differentiation shared

351 among the most introgressed populations. The largest decline in F_{ST} was found in the south
352 western part of the English Channel, in the two populations assigned to the green cluster in
353 the fine-scale analyses (Brx and Ply, Figure 1). Here, F_{ST} decreased below 0.5 (Figure 4 – bottom
354 panel), and 29 out of 32 individuals carried at least one *C. robusta* tract (Figure 3A-B-C). The
355 decline in F_{ST} was not observed in every locations of the contact zone (Figure 4). For instance,
356 F_{ST} values of one was found throughout the entire chromosome 5 in five out of six comparisons
357 involving sites from the pink cluster identified in the fine-scale analysis (Figure 1), except in
358 Rch (Figure 4 – top panel). In this latter population, a second F_{ST} decline was visible on the
359 same chromosome around 3.1 Mb, which is also a region of intermediate differentiation
360 identified by the HMM analysis (Table S4). The admixture signal, and the long tracts of *C.*
361 *robusta* ancestry were absent from all the sites outside the contact zone, or from other
362 chromosomes (Figure 3A-B, and Table S4), except in *C. roulei* where few small tracts of
363 introgression (<77kb) were detected on chromosome 7 and 10.



364

365 **Figure 3:** Evidence for introgression along chromosome 5 from the introduced species *C. robusta* into
 366 the native species *C. intestinalis*. Figure 3A shows the admixture plots computed for K=2 i) on 397
 367 individuals genotyped at 17280 linked-SNPs from the overall dataset (top), and ii) on a subset of 354
 368 SNPs located on chromosome 5, from 0.5 to 2.0 Mb (bottom). Figure 3B displays the introgress plot
 369 showing the genotypes of the 397 individuals at 545 SNPs chosen to be diagnostic between *C. robusta*
 370 and *C. intestinalis* along chromosome 5. Individuals (y-axis) are ordered from top to bottom per species
 371 (*C. robusta*: *C. rob.*, *C. roulei*: *C. rou.*, *C. intestinalis*); *C. intestinalis* individuals are sorted by location.
 372 Populations codes where introgression was detected by the HMM are colored in red. Dark blue boxes
 373 indicate homozygote genotype on *C. robusta* alleles; yellow, homozygote genotype on *C. intestinalis*
 374 alleles; pink, heterozygotes; and white boxes, missing values. Figure 3C shows the proportion of
 375 individuals per site displaying a *C. robusta* tract (blue) in chromosome 5. Figure 3D and 3E display a
 376 neighbor-joining trees build on 794 phased haplotypes showing the similarities between some *C.*
 377 *intestinalis* haplotypes from admixed locations (red dots) and the haplotypes obtained for *C. robusta*
 378 (dark blue dots on the right divergent branch), when using data for chromosome 5 (D) and a zoom
 379 from 0.5 to 2.0 Mb of the same chromosome (E); such similarities are not observed for the haplotypes
 380 obtained in locations from the contact zone with no introgression (yellow dots), or from other sites
 381 located outside the contact zone and from *C. roulei* (colored according to the color code used in Figure
 382 2).



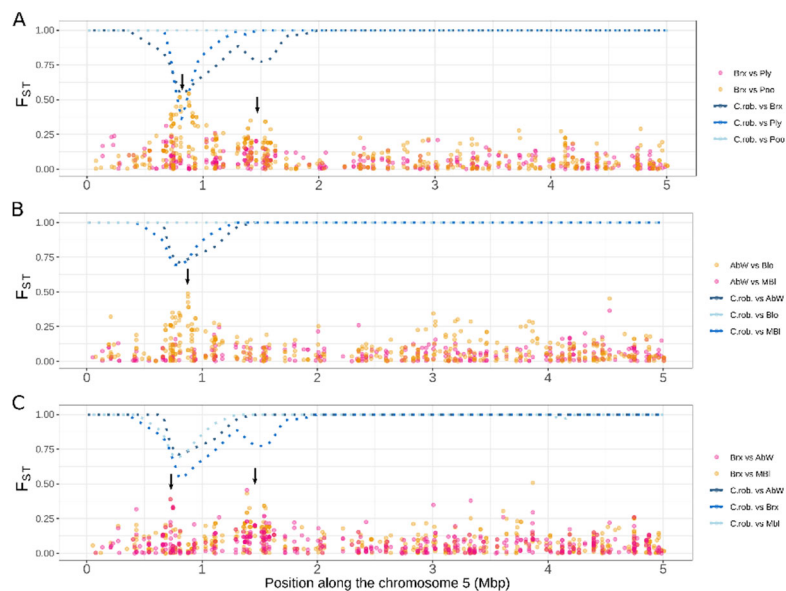
383

384 **Figure 4:** Pairwise F_{ST} values along chromosome 5 between *C. robusta* and populations of *C. intestinalis*
 385 sampled in the contact zone between the two species, in the English Channel, Iroise Sea and Bay of
 386 Biscay. Each graph corresponds to a comparison made with populations from each of the four genetic
 387 clusters identified in the fine-scale analysis (clusters are pictured with the same color code as in Figure
 388 1). The population that belongs to each cluster is listed on the right. Each dot represents the F_{ST} value
 389 for the 1235 SNPs with a MAF of 5% on chromosome 5, and the dashed lines show the maximum F_{ST}
 390 value computed over bins of 100kb.

391

392 The neighbor-joining tree built with SNPs from chromosome 5 showed that some
 393 phased haplotypes from admixed sites (red haplotypes in Figure 3D) were genetically closer
 394 to *C. robusta* than the haplotypes from non-admixed sites (yellow haplotypes in Figure 3D).
 395 This pattern is exacerbated when zooming on the region spanning from 0.5 to 2.0 Mb (Figure
 396 3E). Interestingly, 23 haplotypes (including 10 from 5 homozygous individuals) completely
 397 overlapped with *C. robusta* haplotypes when focusing on the portion between 0.7 and 1.2 Mb
 398 of chromosome 5 (Figure S6B). In this small region, 19 SNPs were species-diagnostic between
 399 Gul and *C. robusta*, and three SNPs were polymorphic in both species (red arrows in Figure
 400 S6A). Two of these three SNPs might reflect incomplete lineage sorting or parallel mutation.
 401 One was indeed polymorphic in several *C. robusta* and one *C. intestinalis* individual from non-
 402 admixed sites (Nah), and the other one was polymorphic in *C. robusta* and in 10 *C. intestinalis*

403 sites, two of which are localized outside the contact zone. Conversely, the last of these three
 404 SNPs (position 1020271), was polymorphic in *C. robusta* (MAF = 9%), with the minor allele only
 405 found in three phased haplotypes of the 23 *C. intestinalis* haplotypes identical to *C. robusta*
 406 haplotypes. The three haplotypes carrying the *C. robusta* minor allele at this specific SNP were
 407 all collected from the east of UK (one from Ply and two from Gpt). These three haplotypes,
 408 different from other *C. intestinalis* haplotypes, all clustered with other *C. robusta* in the
 409 phylogenetic tree (black arrow in Figure S6B). Thus, it seems that different *C. robusta*
 410 haplotypes have introgressed admixed *C. intestinalis* populations. In addition, the
 411 introgression tracts were variable in size (Figure 3B). This size variation is associated with twin
 412 peaks of differentiation among differentially introgressed populations of *C. intestinalis*, as
 413 illustrated by pairwise F_{ST} values computed between Brx and MBI or AbW (Figure 5C, below
 414 the arrows), and also visible when comparing the sites of Ply and Brx (Figure 5A).



415

416 **Figure 5:** F_{ST} values (orange and pink colors) between pairs of *C. intestinalis* populations and maximum
 417 value of F_{ST} (dashed lines in blue colors) between these *C. intestinalis* populations and *C. robusta* over
 418 bins of 100kb, computed using 1235 SNPs with MAF of 5% located along chromosome 5. F_{ST} is
 419 calculated between pairs of sites either A) geographically close in the UK or B) in Brittany (France). In
 420 C), F_{ST} is calculated between introgressed sites in France vs. UK. Pink color is used to show F_{ST} between
 421 two introgressed sites and orange color between an introgressed and a non introgressed site.

422 **Discussion**

423 Using a genome-wide approach based on ddRAD-sequencing, we investigated the
424 fine-scale and large-scale population structure of *Ciona intestinalis* in native (NE Atlantic) and
425 possibly introduced range (US and Iceland). Our results provide supports to previous studies
426 about the influence of human-mediated transports on population structure at both regional
427 and global scales. Comparisons with its congeners *C. robusta* and *C. roulei* offered new insights
428 about the past and recent history of these three species. Our study finally provided the first
429 evidence of contemporary introgression from *C. robusta* into *C. intestinalis*, in the introduction
430 range of the former. This introgression was not homogeneously distributed in the genome but
431 rather forms a breakthrough located in an introgression hotspot on chromosome 5.

432

433 **Chaotic genetic structure and cosmopolitanism: a footprint of human-mediated dispersal**

434 The overall genetic structure of *C. intestinalis* was in line with the results from previous
435 studies (Bouchemousse et al., 2016a,c; Hudson et al., 2016; Johannesson et al., 2018; Hudson,
436 Johannesson, McQuaid, & Rius, 2020, Johannesson, Le Moan, Perini, & André, 2020). In
437 particular, we confirmed that the populations of *C. intestinalis* are highly differentiated over
438 large geographical scales, which is likely due to the presence of different glacial lineages in
439 Europe (Hudson et al., 2020). At smaller geographical scales, *C. intestinalis* is much less
440 genetically structured, and more importantly shows discrepancies between genetic clustering
441 and geographic distance, leading to a mosaic structure. Such a mosaic structure had been
442 previously reported in the study area, with microsatellite markers, and attributed to human-
443 mediated connectivity among harbors (Hudson et al., 2016). *Ciona intestinalis* has indeed
444 limited natural dispersal capacity, being characterized by very short-lived larvae (<24 hours

445 under laboratory conditions). Similar mosaic structures have been documented in introduced
446 species inhabiting ports, and characterized by low natural dispersal ability, for example the
447 seaweed *Undaria pinnatifida* (Guzinski, Ballenghien, Daguin-Thiébaud, Lévêque, & Viard,
448 2018).

449 The individuals sampled from locations with unclear native vs. non-native status (in
450 the North Western Atlantic and in Iceland) were genetically more similar to the populations
451 sampled in France and England (average F_{ST} of 0.032) than to other European populations (F_{ST}
452 up to 0.176). They were also more similar among one another than several comparisons within
453 Europe, suggesting that they derived recently from somewhere near the English Channel. A
454 similar situation has been described by Hudson et al. (2020) in a Canadian site located further
455 north than our study population (Nah). However, the Canadian individuals appeared admixed
456 between Swedish and English Channel lineages, while Nah here appeared as a mostly pure
457 cluster genetically close to the English Channel lineage. Population structure along the
458 western Atlantic Coast is common, even in invasive species such as the green crab (Pringle,
459 Blakeslee, Byers, & Roman, 2011; Jeffery, et al., 2017). The Canadian site is located in Nova
460 Scotia, a major suture zone for marine species living along the Western Atlantic coast, while
461 our sampling site is located further south where populations display usually less admixture
462 (Standley et al., 2018). The different admixture pattern observed here and in Hudson et al.
463 (2020) suggest population structure of *C. intestinalis* in North America that might reflect
464 multiple introduction events from the native range, which is often reported in marine
465 introduced species (Viard, David, & Darling, 2016). Additional sampling sites along the
466 Western Atlantic coast are needed to further explore the population structure of *C. intestinalis*
467 and reconstruct its introduction history. Our study concurs with previous results by

468 Bouchemousse et al. (2016a) and Hudson et al. (2020), and suggests that *C. intestinalis* can be
469 described as a neo-cosmopolitan species according to the terminology by Darling & Carlton
470 (2018), with a trans-Atlantic distribution due to human introductions, rather than to an
471 eucosmopolitan species with relict populations that took advantage of new available habitats
472 (ports and marinas).

473

474 **A continued and needed appraisal of species status within the *Ciona* genus**

475 The many *Ciona* species belong to a complex genus, as shown by the recent
476 discovery of a new species in the Mediterranean Sea (i.e., *Ciona intermedia*, Mastrototaro et
477 al., 2020), and the recent in-depth taxonomic revision of the previously accepted *C. intestinalis*
478 species. Morphological observations (Brunetti et al., 2015), molecular data (Nydam &
479 Harrison, 2010; Zhan, Macisaac, & Cristescu, 2010; Bouchemousse et al., 2016a,c) and
480 experimental crosses (Lambert et al., 1990; Malfant et al., 2017) show that two deeply
481 diverged lineages, named type A and type B, co-exist under the accepted name *C. intestinalis*.
482 This taxonomic revision restored a previously synonymized species, *C. robusta*. In line with
483 these previous data, we observed differentially fixed SNPs between *C. robusta* and *C.*
484 *intestinalis* across most of their genome (average $F_{ST} > 0.780$ for all pairwise comparison). This
485 conclusion also holds for *C. robusta* and *C. roulei*. Our results show that the species status of
486 *C. roulei* needs revision. We showed a weak genome-wide differentiation when comparing *C.*
487 *roulei* with *C. intestinalis* populations sampled in the English Channel ($F_{ST} < 0.075$).
488 Interestingly, the differentiation is even weaker than those observed between the southern
489 and northern European populations of *C. intestinalis* studied here (F_{ST} up to 0.169). Lambert
490 et al. (1990) and Malfant et al. (2018) performed experimental crosses showing that ‘*C.*

491 *intestinalis* and *C. roulei* can be hybridized with no signs of outbreeding depression. In
492 addition, mitochondrial sequencing data shows that the two taxa display similar haplotypes
493 (Malfant et al., 2018). Our results agree with these previous studies suggesting that *C. roulei*
494 is a divergent lineage of *C. intestinalis*, likely trapped in cold waters of the northern
495 Mediterranean Sea after post-glacial warming, like other cold-adapted marine species
496 (*Platichthys flesus*: Borsa, Blanquer, & Berrebi, 1997; *Sprattus sprattus*: Debes, Zachos, Hanel,
497 2008; *Sagitta setosa*: Peijnenburg, Fauvelot, Breeuwer, & Menken, 2006).

498

499 **Evidence for contemporary introgression of the native species by its introduced congener**

500 Our analyses provided several novel results. We observed one peak of intra-specific
501 differentiation on chromosome 5 that corresponded to a decline of inter-specific
502 differentiation between sympatric populations of *C. robusta* and *C. intestinalis*. This decline
503 was only found in a subset of the populations located in the contact zone between the two
504 species, pointing toward introgression from *C. robusta* into *C. intestinalis* populations.
505 Introgression was confirmed by the presence of long *C. robusta* ancestry tracts in some *C.*
506 *intestinalis* individuals sampled across the English Channel and Iroise Sea, and in the Rch site
507 in the Bay of Biscay. These long introgression tracts were primarily found on chromosome 5
508 between positions 0.38 to 2.32 Mb (Table S4), which we refer as an introgression hotspot.
509 Other long tracts were found on chromosome 5 outside the main introgression hotspot in
510 three individuals from Rch and one from Cam. These other tracts close to the introgression
511 hotspot could be either due to independent introgression events or, more likely to tracts that
512 hitchhiked with the introgression at the hotspot. This introgression restricted to a single
513 hotspot explains why it was missed in previous studies using fewer ancestry-informative SNPs

514 (Bouchemousse et al., 2016b,c). Marker density is thus a key to obtain evidence for very
515 localized introgression (Ravinet et al., 2018, Turissini & Matute, 2017, Stankowski et al., 2020).

516 *C. roulei* did not show long tracts of *C. robusta* ancestry in chromosome 5, but showed
517 two small tracks (<78kb) on chromosome 7 and 10, and signs of admixture with *C. robusta*
518 were spread across all chromosomes. This admixture pattern and the absence of long
519 haplotypes from *C. robusta* into *C. roulei* suggests that introgression between *C. roulei* and *C.*
520 *robusta* is ancient. Previous studies examining gene flow between *C. intestinalis* and *C. robusta*
521 have interpreted the patterns of allele sharing as a consequence of historical rather than
522 recent introgression (Bouchemousse et al., 2016c). The *C. roulei* samples examined here could
523 thus be another example of historical introgression. Alternatively, such introgression could
524 involve another species closely related to *C. robusta*, which was present in the Mediterranean
525 Sea. Out of the 14 species currently accepted in the genus *Ciona* (Word Register of Marine
526 Species; <http://marinespecies.org/>), three other species not included in this study have been
527 reported in the Mediterranean Sea (*C. intermedia*, *Ciona* sp. C and *Ciona* sp. D). Based on
528 mitochondrial phylogeny, *Ciona* sp. C appears genetically close to *C. robusta* (Mastrototaro et
529 al., 2020). Thus, carrying out a genome-wide analysis on the *Ciona* taxa found in the
530 Mediterranean Sea is needed to confirm that the signs of admixture into *C. roulei* are truly
531 due to introgression with *C. robusta*.

532

533 **Spatial and temporal dynamics of the introgression tracts**

534 The recent invasion of *C. robusta* into the English Channel (Nydam & Harrison,
535 2010), the absence of large introgression tracts outside of the contact zone including in pop-
536 ulations outside European Seas (Rek and Nah), the large size of the introgression tracts (>0.5

537 Mb) and the genetic similarity between introgressed tracts in *C. intestinalis* and *C. robusta*
538 haplotypes (Figure 3), all point toward a recent introgression event (i.e., post-dating the intro-
539 duction of *C. robusta* in Europe in the late 20th or early 21st centuries). Dating the age of an
540 introgression event is not easy when using reduced representation DNA sequencing, such as
541 ddRAD-sequencing. Shchur et al. (2020) provided a theoretical framework for interpreting the
542 timing of introgression, based on the distribution of genomic admixture tract lengths, and in-
543 cluding positive selection effects. However, the underlying assumption of their method may
544 not hold in our case. For example, the level of genome-wide introgression should be suffi-
545 ciently high to estimate the baseline neutral introgression rate, while in our case it is around
546 0.1% on average (Fraise et al., unpublished results), with few tracts outside of the chromo-
547 some 5 hotspot. In addition, these methods assume a single pulse of admixture has occurred
548 while multiple admixture events have likely happened in *C. intestinalis*, as discussed below.

549 The localized pattern of long (0.84-1.24 Mb, Table S4), and thus likely young, tracts
550 of introgression in a single region of chromosome 5, is difficult to explain without invoking
551 some sort of selection. Although a localized desert of genes associated with low
552 recombination rates can potentially produce this pattern, such regions usually exist at several
553 places of a genome. In addition, 21 genes are located in the 80kb tract shared by the most
554 introgressed populations, from 0.81 to 0.88 Mb of chromosome 5 (listed in Table S5), which
555 does not support the gene desert hypothesis. Moreover, the high variation of *C. robusta* tracts
556 length in *C. intestinalis* suggests recombination is operating efficiently on the introgression
557 hotspot. In addition, the introgression is not fixed in any of the study populations. Simple
558 positive selection (adaptive introgression *sensu stricto*) also seems unlikely to explain

559 synchronous incomplete sweeps at many distant locations belonging to different genetic
560 clusters. Alternative selective scenarios therefore need to invoke other kind of selection such
561 as balancing selection, frequency dependence or heterosis. For instance, as the two species
562 studied here are highly divergent (12%, which translates into roughly 2,000 non-synonymous
563 substitutions per Mb), the introgressed tracts may carry many mutations affecting fitness.
564 Theory predicts an intrinsic benefit of heterozygosity and a cost of admixture (Schneemann,
565 De Sanctis, Roze, Bierne, & Welch, 2020). Introgressed tracts might provide a fitness
566 advantage when heterozygous at the start of the introgression, through heterosis effect, but
567 a fitness reduction when frequent and homozygote in a *C. intestinalis* background.
568 Furthermore, the heterosis effect could vanish when long tracts are broken down by
569 recombination (Harris & Nielsen, 2016). Thus, the dynamic of the introgression might change
570 over time, with some haplotypes being first positively selected, and then counter selected as
571 soon as recombination breaks the introgressed haplotypes into smaller pieces (Leitwein,
572 Duranton, Rougemont, Gagnaire, & Bernatchez, 2020). Examining changes in introgression
573 frequencies and distribution over time is now required to investigate if selection is truly acting
574 and continues to drive introgressed tracts to high frequency. In addition, genome sequencing
575 will better delineate the heart of the introgression hotspot and identify candidate genes on
576 the basis of their function.

577 The length of introgression tracts can be informative about where the contact
578 happened, as it is inversely correlated to the distance from where the introgression first
579 occurred (Leitwein, Duranton, Rougemont, Gagnaire, & Bernatchez et al., 2020). This effect
580 was documented between the Atlantic and the Mediterranean lineages of European sea bass,
581 *Dicentrarchus labrax*, where the size of the Atlantic introgression tracts into the

582 Mediterranean population is proportionally reduced with the distance from the contact zone
583 between the two lineages, at the Almeria-Oran front (Duranton, Bonhomme, & Gagnaire,
584 2019). Building on these observations, the length of *C. robusta* haplotypes in the genome of
585 *C. intestinalis* individuals was generally larger in western UK (1.24 Mb in Brx, Table S4) than in
586 other populations (e.g., ~0.5 Mb in Gpt/Sth/MBI), suggesting that the introgression occurred
587 there first. Interestingly, introgression was not found restricted to a particular area and shows
588 a chaotic spatial structure similar to what was observed in Hudson et al. (2016), and genome
589 wide in our study. Intra-specific gene flow promoted by human-mediated transport, as
590 previously suggested (Hudson et al., 2016), could explain the chaotic dispersal of introgression
591 tracts. In this context, introgression tracts could be a relevant tool to identify recent human-
592 mediated migration routes (Gagnaire et al., 2015). For instance, while the population Jer
593 seems genetically related to distant populations located in central-eastern UK at the genome-
594 wide level (Figure 2B), this population shares a lack of introgression tracts with StM (Figure
595 3C), a population in close vicinity, and strongly connected through leisure boating.

596 Another non mutually exclusive hypothesis might explain both the mosaic structure of
597 introgression and the variable size of introgression tracts: independent introgression events
598 occurring several times in different populations. For instance, the honey bee (*Apis mellifera*
599 *scutellata*) shows repeated introgressions from the African lineage into the European lineage
600 at one genomic location of chromosome 1 in two hybrid zones that are localized more than
601 5,000 kilometers apart (Calfee et al., 2020). Repeated adaptive introgression has also been
602 documented at a gene involved in insecticide resistance in the mosquito *Anopheles* spp. (Weill
603 et al., 2000; Norris et al., 2015) and in cotton bollworm *Helicoverpa* spp. (Valencia-Montoya
604 et al., 2020). Although *Ciona* species are much more divergent than any of the examples

605 presented above, the contact zones between *C. intestinalis* and *C. robusta* are not restricted
606 to a few locations, as they co-occurred in syntopy in several isolated harbors along coastlines
607 of Great-Britain and France, which increases the possibility of introgression. We showed that
608 one SNP distinguished two haplotypes of *C. robusta* found within *C. intestinalis* individuals,
609 suggesting that introgression from *C. robusta* into *C. intestinalis* could have happened at least
610 twice. This SNP was only carried by long introgressed tracts, and could be a footprint of more
611 recent introgression contributing to the twin peaks of differentiation observed in the
612 surrounding of the main introgression hotspot between populations that carry introgression
613 tracts of variable size (Figure 5).

614 The possibility of repeated introgression events does not rule out the hypothesis of
615 intra-specific diffusion of the introgression facilitated by human activities. Both mechanisms
616 might have jointly contributed to the current distribution of the introgression tracts. Overall,
617 our study shows that anthropogenic hybridization can be effective in promoting gene flow
618 between species even at late stage of speciation, contributing to the population structure
619 described among contemporary populations. Future work should focus on complete genome
620 sequencing, and temporal sampling of introgressed populations.

621 **Acknowledgements**

622 The authors are grateful to Sarah Bouchemousse and Sabrina Le Cam for providing samples
623 from Iceland and Norway, to Alan Brelsford and Pierre-Alexandre Gagnaire for providing
624 advice on ddRAD-sequencing and/or scripts for RAD-Seq analyses, and to James Reeve for help
625 in editing the last version of the manuscript. The authors are also most grateful to the
626 Biogenouest Genomer core facility for their technical support, and from the Biogenouest
627 ABIMS Platform for softwares and bioinformatics tools set-up, and access to computing
628 resources. This work benefitted from funding of the French National Research Agency (ANR)
629 with regards the ANR Project HYSEA (no. ANR-12-BSV7-0011).

630 **Data Accessibility**

631 Trimmed fastq files are available under the NCBI bioproject “PRJNA759193”. The filtered vcf
632 files used for the fine-scale, large-scale, and introgression analyses, respectively, as well as the
633 R scripts for population genetics and HMM analyses, and the custom scripts for phylogenetic
634 analyses are available as zenodo archives (doi: 10.5281/zenodo.5346932).

635 **Authors’ contributions**

636 F.V. and N.B. designed the research. F.V. obtained the funding grants. C.D. and C.R. performed
637 the lab-work. C.R. performed the raw-data filtration. A.LM. performed the bio-informatics
638 analyses with advices from N.B. and C.F.. A.LM., N.B., C.F., and F.V. interpreted the data. A.LM.
639 and F.V. wrote the first-draft of the paper, with substantial inputs from N.B. and C.F.. All
640 authors revised and edited the final manuscript.

641 **References**

- 642 Borsa, P., Blanquer, A., & Berrebi, P. (1997). Genetic structure of the flounders *Platichthys flesus* and *P. stellatus*
643 at different geographic scales. *Marine Biology*, *129*, 233–246. doi: 10.1007/s002270050164
- 644 Bouchemousse, S., Bishop, J.D.D., & Viard, F. (2016a). Contrasting global genetic patterns in two biologically
645 similar, widespread and invasive *Ciona* species (Tunicata, Ascidiacea). *Scientific reports*, *6*, 24875. doi:
646 10.1038/srep24875
- 647 Bouchemousse, S., Lévêque, L., Dubois, G., & Viard, F. (2016b). Co-occurrence and reproductive synchrony do
648 not ensure hybridization between an alien tunicate and its interfertile native congener. *Ecology and*
649 *Evolution*, *30*, 69–87. doi: 10.1007/s10682-015-9788-1
- 650 Bouchemousse, S., Lévêque, L. & Viard, F. (2017). Do settlement dynamics influence competitive interactions
651 between an alien tunicate and its native congener? *Ecology and Evolution*, *7*, 200–213. doi:
652 10.1002/ece3.2655
- 653 Bouchemousse, S., Liautard-Haag, C., Bierne, N., & Viard, F. (2016c). Distinguishing contemporary hybridization
654 from past introgression with postgenomic ancestry-informative SNPs in strongly differentiated *Ciona*
655 species. *Molecular Ecology*, *25*, 5527–5542. doi: 10.1111/mec.13854
- 656 Brelsford, A., Dufresnes, C., & Perrin, N. (2016). High-density sex-specific linkage maps of a European tree frog
657 (*Hyla arborea*) identify the sex chromosome without information on offspring sex. *Heredity*, *116*, 177–
658 181. doi: 10.1038/hdy.2015.83
- 659 Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for
660 whole-genome association studies by use of localized haplotype clustering. *American Journal of Human*
661 *Genetics*, *81*, 1084–1097. doi: 10.1086/521987
- 662 Brunetti, R., Gissi, C., Pennati, R., Caicci, F., Gasparini, F., & Manni, L. (2015). Morphological evidence that the
663 molecularly determined *Ciona intestinalis* type A and type B are different species: *Ciona robusta* and
664 *Ciona intestinalis*. *Journal of Zoological Systematics and Evolutionary Research*, *53*, 186–193. doi:
665 10.1111/jzs.12101
- 666 Calfee, E., Agra, M.N., Palacio, M.A., Ramírez, S.R., & Coop, G. (2020). Selection and hybridization shaped the
667 rapid spread of African honey bee ancestry in the Americas. *PLoS Genetics*, *16*, e1009038. doi:
668 10.1371/journal.pgen.1009038
- 669 Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., ... 1000 Genomes Project Analysis
670 Group (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. doi:
671 10.1093/bioinformatics/btr330
- 672 Darling, J. A., & Carlton, J. T. (2018). A framework for understanding marine cosmopolitanism in the
673 Anthropocene. *Frontiers in Marine Science*, *5*, 293. doi: 10.3389/fmars.2018.00293
- 674 Debes P.V., Zachos F.E., & Hanel, R. (2008). Mitochondrial phylogeography of the European sprat (*Sprattus*
675 *sprattus* L., Clupeidae) reveals isolated climatically vulnerable populations in the Mediterranean Sea and
676 range expansion in the northeast Atlantic. *Molecular Ecology*, *17*, 3873–88. doi: 10.1111/j.1365-
677 294X.2008.03872.x.
- 678 Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., ... Rokhsar, D.S. (2002). The draft
679 genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science*, *298*, 2157–67. doi:
680 10.1126/science.1080049
- 681 Duranton, M., Bonhomme, F. & Gagnaire, P.-A. (2019). The spatial scale of dispersal revealed by admixture tracts.
682 *Evolutionary applications*, *12*, 1743–1756. doi: 10.1111/eva.12829
- 683 Faust, E., Halvorsen, K. T., Andersen, P., Knutsen, H., & André, C. (2018). Cleaner fish escape salmon farms and
684 hybridize with local wrasse populations. *Royal Society Open Science*, *5*, 171752. doi :
685 10.1098/rsos.171752
- 686 Firth, L. B., Knights, A. M., Bridger, D., Evans, A., Mieskowska, N., Moore, P. J., . . . Hawkins, S. J. (2016). Ocean
687 sprawl: challenges and opportunities for biodiversity management in a changing world. *Oceanography*
688 *and Marine Biology: An Annual Review*, *54*, 193–269. doi:10.1201/9781315368597
- 689 Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in*
690 *Ecology and Evolution*, *6*, 925–929. doi: 10.1111/2041-210X.12382
- 691 Gagnaire, P. A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., ... Bierne, N. (2015). Using neutral,
692 selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era.
693 *Evolutionary Applications*, *8*, 769–786. doi: 10.1111/eva.12288

- 694 Gissi, C., Hastings, K.E.M., Gasparini, F., Stach, T., Pennati, R., & Manni, L. (2017). An unprecedented taxonomic
695 revision of a model organism: the paradigmatic case of *Ciona robusta* and *Ciona intestinalis*. *Zoologica*
696 *Scripta*, *46*, 521–522. doi: 10.1111/zsc.12233
- 697 Gompert, Z., & Alex Buerkle, C. (2010). introgress: a software package for mapping components of isolation in
698 hybrids. *Molecular Ecology Resources*, *10*, 378–384. doi: 10.1111/j.1755-0998.2009.02733.x
- 699 Grabenstein, K.C., & Taylor, S.A. (2018). Breaking barriers: causes, consequences, and experimental utility of
700 human-mediated hybridization. *Trends in Ecology & Evolution*, *33*, 198–212. doi:
701 10.1016/j.tree.2017.12.008
- 702 Guzinski, J., Ballenghien, M., Daguin-Thiébaud, C., Lévêque, L., & Viard, F. (2018). Population genomics of the
703 introduced and cultivated Pacific kelp *Undaria pinnatifida*: Marinas-not farms-drive regional
704 connectivity and establishment in natural rocky reefs. *Evolutionary Applications*, *11*, 1582–1597. doi:
705 10.1111/eva.12647
- 706 Harris, K., & Nielsen, R. (2016). The genetic cost of Neanderthal introgression. *Genetics*, *203*(2), 881–891. doi:
707 10.1534/genetics.116.186890
- 708 Hofer, T., Foll, M., & Excoffier, L. (2012). Evolutionary forces shaping genomic islands of population differentiation
709 in humans. *BMC genomics*, *13*, 107. <https://doi.org/10.1186/1471-2164-13-107>
- 710 Hudson, J., Johannesson, K., McQuaid, C.D., & Rius, M. (2020). Secondary contacts and genetic admixture shape
711 colonization by an amphiatlantic epibenthic invertebrate. *Evolutionary Applications*, *13*, 600–612. doi:
712 10.1111/eva.12893
- 713 Hudson, J., Viard, F., Roby, C., & Rius, M. (2016). Anthropogenic transport of species across native ranges:
714 unpredictable genetic and evolutionary consequences. *Biology Letters*, *12*, 20160620. doi:
715 10.1098/rsbl.2016.0620
- 716 Hufbauer, R. A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C. E., Rey, O., & Estoup, A. (2012).
717 Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered
718 habitats within the native range can promote invasions. *Evolutionary Applications*, *5*, 89–101. doi:
719 10.1111/j.1752-4571.2011.00211.x
- 720 Jeffery, N.W., DiBacco, C., Wringe, B.F., Stanley, R.R.E., Hamilton, L.C., Ravindran, P.N. & Bradbury, I.R. (2017).
721 Genomic evidence of hybridization between two independent invasions of European green crab
722 (*Carcinus maenas*) in the Northwest Atlantic. *Heredity*, *119*, 154-165. doi: 10.1038/hdy.2017.22
- 723 Johannesson, K., Le Moan, A., Perini, S., & André, C. (2020). A Darwinian laboratory of multiple contact zones.
724 *Trends in Ecology & Evolution*, *35*, 1021–1036. doi: 10.1016/j.tree.2020.07.015
- 725 Johannesson, K., Ring, A. K., Johannesson, K. B., Renborg, E., Jonsson, P. R., & Havenhand, J. N. (2018).
726 Oceanographic barriers to gene flow promote genetic subdivision of the tunicate *Ciona intestinalis* in a
727 North Sea archipelago. *Marine Biology*, *165*, 126. doi: 10.1007/s00227-018-3388-x
- 728 Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data.
729 *Bioinformatics*, *27*, 3070–3071. doi: 10.1093/bioinformatics/btr521
- 730 Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for
731 the analysis of genetically structured populations. *BMC Genetics*, *11*, 94. doi: 10.1186/1471-2156-11-94
- 732 Lambert, C., Lafargue, F., & Lambert, G. (1990). Preliminary note on the genetic isolation of *Ciona* species
733 (Ascidiacea, Urochordata). *Vie et Milieu*, *40*, 293–295. doi: hal-03036352
- 734 Leitwein, M., Duranton, M., Rougemont, Q., Gagnaire, P. A., & Bernatchez, L. (2020). Using haplotype information
735 for conservation genomics. *Trends in Ecology & Evolution*, *35*(3), 245–258. doi:
736 10.1016/j.tree.2019.10.012
- 737 Malfant, M., Coudret, J., Le Merdy, R. & Viard, F. (2017). Effects of temperature and salinity on juveniles of two
738 ascidians, one native and one invasive, and their hybrids. *Journal of Experimental Marine Biology and*
739 *Ecology*, *497*, 180-187. doi: 10.1016/j.jembe.2017.09.019
- 740 Malfant, M., Darras, S., & Viard, F. (2018). Coupling molecular data and experimental crosses sheds light about
741 species delineation: a case study with the genus *Ciona*. *Scientific Reports*, *8*, 1480. doi: 10.1038/s41598-
742 018-19811-2
- 743 Mastrototaro, F., Montesanto, F., Salonna, M., Viard, F., Chimienti, G., Trainito, E., & Gissi, C. (2020). An
744 integrative taxonomic framework for the study of the genus *Ciona* (Ascidiacea) and description of a new
745 species, *Ciona intermedia*. *Zoological Journal of the Linnean Society*, *190*, 1193–1216. doi:
746 10.1093/zoolinnean/zlaa042

747 Marques, D. A., Lucek, K., Haesler, M. P., Feller, A. F., Meier, J. I., Wagner, C. E., Excoffier, L., & Seehausen, O.
748 (2017). Genomic landscape of early ecological speciation initiated by selection on nuptial colour.
749 *Molecular ecology*, 26(1), 7–24. <https://doi.org/10.1111/mec.13774>

750 Turissini, D. A., & Matute, D. R. (2017). Fine scale mapping of genomic introgressions within the *Drosophila*
751 *yakuba* clade. *PLoS Genetics*, 13, e1006971. doi: 10.1371/journal.pgen.1006971

752 McFarlane, S. E., & Pemberton, J. M. (2019). Detecting the true extent of introgression during anthropogenic
753 hybridization. *Trends in Ecology & Evolution*, 34, 315–326. doi: 10.1016/j.tree.2018.12.013

754 Norris, L. C., Main, B. J., Lee, Y., Collier, T. C., Fofana, A., Cornel, A. J., & Lanzaro, G. C. (2015). Adaptive
755 introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated
756 bed nets. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 815–
757 820. doi: 10.1073/pnas.1418892112

758 Nydam, M. L., & Harrison, R. G. (2010). Polymorphism and divergence within the ascidian genus *Ciona*. *Molecular*
759 *Phylogenetics and Evolution*, 56, 718–726. doi: 10.1016/j.ympev.2010.03.042

760 Peijnenburg K.T., Fauvelot C., Breeuwer J.A., & Menken, S.B. (2006). Spatial and temporal genetic structure of
761 the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear
762 DNA markers. *Molecular Ecology*, 15, 3319–38. doi: 10.1111/j.1365-294X.2006.03002.x.

763 Pembleton, L. W., Cogan, N. O., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic
764 differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, 13, 946–
765 952. doi : 10.1111/1755-0998.12129

766 Popovic, I., Matias, A.M.A., Bierne, N., & Riginos, C. (2020). Twin introductions by independent invader mussel
767 lineages are both associated with recent admixture with a native congener in Australia. *Evolutionary*
768 *Applications*, 13, 515–532. doi: 10.1111/eva.12857

769 Pringle, J. M., Blakeslee, A. M., Byers, J. E., & Roman, J. (2011). Asymmetric dispersal allows an upstream region
770 to control population structure throughout a species' range. *Proceedings of the National Academy of*
771 *Sciences of the United States of America*, 108, 15288–15293. doi: 10.1073/pnas.1100473108

772 Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2014). dDocent: a RADseq, variant-calling pipeline designed for
773 population genomics of non-model organisms. *PeerJ*, 2, e431. <https://doi.org/10.7717/peerj.431>

774 Ravinet, M., Yoshida, K., Shigenobu, S., Toyoda, A., Fujiyama, A., & Kitano, J. (2018). The genomic landscape at a
775 late stage of stickleback speciation: High genomic divergence interspersed by small localized regions of
776 introgression. *PLoS Genetics*, 14, e1007358. doi : journal.pgen.1007358

777 Rochette, N.C., Rivera-Colón, A.G., & Catchen, J.M. (2019). Stacks 2: Analytical methods for paired-end
778 sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28, 4737–4754. doi:
779 10.1111/mec.15253

780 Roux, C., Fraïsse, C., Romiguier, J., Ancaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of
781 speciation along a continuum of genomic divergence. *PLoS Biology*, 14, e2000234. doi:
782 10.1371/journal.pbio.2000234

783 Roux, C., Tsagkogeorga, G., Bierne, N., & Galtier, N. (2013). Crossing the species barrier: genomic hotspots of
784 introgression between two highly divergent *Ciona intestinalis* species. *Molecular Biology and Evolution*,
785 30, 1574–1587. doi: 10.1093/molbev/mst066

786 Shchur, V., Svedberg, J., Medina, P., Corbett-Detig, R., & Nielsen, R. (2020). On the Distribution of Tract Lengths
787 During Adaptive Introgression. *G3: Genes|Genomes|Genetics*, 10(10), 3663–3673.
788 doi:10.1534/g3.120.401616

789 Schliep, K.P. (2011). phangorn: phylogenetic analysis in R. *Bioinformatics*, 27, 592–593. doi:
790 10.1093/bioinformatics/btq706

791 Schneemann, H., De Sanctis, B., Roze, D., Bierne, N. & Welch, J.J. (2020). The geometry and genetics of
792 hybridization. *Evolution*, 74, 2575–2590. doi: 10.1111/evo.14116

793 Seebens, H., Blackburn, T.M., Dyer, E.E., Genovesi, P., Hulme, P.E., Jeschke, J.M., ... Essl, F. (2017). No saturation
794 in the accumulation of alien species worldwide. *Nature Communications*, 8, 14435. doi:
795 10.1038/ncomms14435

796 Shi, Y., Bouska, K. L., McKinney, G. J., Dokai, W., Bartels, A., McPhee, M. V., & Larson, W. A. (2021). Gene flow
797 influences the genomic architecture of local adaptation in six riverine fish species. *BioRxiv*. doi:
798 10.1101/2021.05.18.444736

799 Simon, A., Arbiol, C., Nielsen, E.E., Couteau, J., Sussarellu, R., Burgeot, T., ... Bierne, N. (2020). Replicated
800 anthropogenic hybridizations reveal parallel patterns of admixture in marine mussels. *Evolutionary*
801 *Applications*, 13, 575–599. doi: 10.1111/eva.12879

802 Simon, A., Fraïsse, C., El Ayari, T., Liautard-Haag, C., Strelkov, P., Welch, J. J., & Bierne, N. (2021). How do species
803 barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. *Journal of*
804 *Evolutionary Biology*, *34*, 208–223. doi: 10.1111/jeb.13709

805 Soria-Carrasco, V., Gompert, Z., Comeault, A. A., Farkas, T. E., Parchman, T. L., Johnston, J. S., Buerkle, C. A., Feder,
806 J. L., Bast, J., Schwander, T., Egan, S. P., Crespi, B. J., & Nosil, P. (2014). Stick insect genomes reveal
807 natural selection's role in parallel speciation. *Science*, *344*(6185), 738–742. doi:
808 10.1126/science.1252136

809 Stankowski, S., Westram, A. M., Zagrodzka, Z. B., Eyres, I., Broquet, T., Johannesson, K., & Butlin, R. K. (2020). The
810 evolution of strong reproductive isolation between sympatric intertidal snails. *Philosophical*
811 *Transactions of the Royal Society of London. Series B, Biological sciences*, *375*, 20190545. doi:
812 10.1098/rstb.2019.0545.

813 Stanley, R., DiBacco, C., Lowen, B., Beiko, R. G., Jeffery, N. W., Van Wyngaarden, ... Bradbury, I. R. (2018). A
814 climate-associated multispecies cryptic cline in the northwest Atlantic. *Science Advances*, *4*, eaaq0929.
815 doi: 10.1126/sciadv.aaq0929

816 Turissini, D. A., & Matute, D. R. (2017). Fine scale mapping of genomic introgressions within the *Drosophila*
817 *yakuba* clade. *PLoS Genetics*, *13*, e1006971. doi: 10.1371/journal.pgen.1006971

818 Valencia-Montoya, W. A., Elfekih, S., North, H. L., Meier, J. I., Warren, I. A., Tay, W. T., Gordon, K., Specht, A.,
819 Paula-Moraes, S. V., Rane, R., Walsh, T. K., & Jiggins, C. D. (2020). Adaptive introgression across
820 semipermeable species boundaries between local *Helicoverpa zea* and invasive *Helicoverpa armigera*
821 moths. *Molecular biology and evolution*, *37*, 2568–2583. doi: 10.1093/molbev/msaa108

822 Viard, F., David, P. & Darling, J. (2016). Marine invasions enter the genomic era: Three lessons from the past, and
823 the way forward. *Current Zoology*, *62*, 629-642. doi: 10.1093/cz/zow053

824 Viard, F., Riginos, C., & Bierne, N. (2020). Anthropogenic hybridization at sea: Three evolutionary questions
825 relevant to invasive species management. *Philosophical Transactions of the Royal Society of London.*
826 *Series B, Biological sciences*, *375*, 20190547. doi: 10.1098/rstb.2019.0547

827 Weill, M., Chandre, F., Brengues, C., Manguin, S., Akogbeto, M., Pasteur, N., Guillet, P., & Raymond, M. (2000).
828 The kdr mutation occurs in the Mopti form of *Anopheles gambiae* s.s. through introgression. *Insect*
829 *Molecular Biology*, *9*, 451–455. doi: 10.1046/j.1365-2583.2000.00206.x

830 Weir, B.S., & Cockerham, C.C. (1984). Estimating F-Statistics for the analysis of population structure. *Evolution*,
831 *38*, 1358–1370. doi: 10.2307/2408641

832 Wickham, H. (2011). ggplot2. *WIREs Computational Statistics*, *3*, 180–185. doi: 10.1002/wics.147

833 Zhan, A., Macisaac, H.J., & Cristescu, M.E. (2010). Invasion genetics of the *Ciona intestinalis* species complex:
834 from regional endemism to global homogeneity. *Molecular Ecology*, *19*, 4678–4694. doi:
835 10.1111/j.1365-294X.2010.04837.x