



HAL
open science

Distinguishing contemporary hybridization from past introgression with post-genomic ancestry-informative SNPs in strongly differentiated *Ciona* species

Sarah Bouchemousse, Cathy Liautard-Haag, Nicolas Bierne, Frédérique Viard

► **To cite this version:**

Sarah Bouchemousse, Cathy Liautard-Haag, Nicolas Bierne, Frédérique Viard. Distinguishing contemporary hybridization from past introgression with post-genomic ancestry-informative SNPs in strongly differentiated *Ciona* species. *Molecular Ecology*, 2016, 25 (21), pp.5527-5542 10.1111/mec.13854 . hal-01383805

HAL Id: hal-01383805

<https://hal.sorbonne-universite.fr/hal-01383805>

Submitted on 19 Oct 2016

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 **Distinguishing contemporary hybridization from past**
2 **introgression with post-genomic ancestry-informative SNPs in**
3 **strongly differentiated *Ciona* species**

4
5 **Sarah Bouchemousse^{1,2,*}, Cathy Liautard-Haag^{3,4}, Nicolas Bierne^{3,4} and Frédérique**
6 **Viard^{1,2,*}**

7 **¹Sorbonne Universités, UPMC Univ Paris 6, CNRS, UMR 7144, Equipe DIVCO, Station**
8 **Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France**

9 **²CNRS, UMR 7144, Adaptation et Diversité en Milieu Marin, Station Biologique de Roscoff,**
10 **Place Georges Teissier, 29680 Roscoff, France**

11 **³Université de Montpellier, Station Marine - OSU OREME, 2 Rue des Chantiers, 34200 Sète,**
12 **France**

13 **⁴CNRS-UM-IRD-EPHE, UMR 5554, Institut des Sciences de l'Evolution, Place Eugène**
14 **Bataillon, 34095 Montpellier, France**

15

16 *Correspondence : UMR7144, Equipe Diversité et Connectivité dans le paysage marin Côtier
17 (DIVCO), CNRS-UPMC, Station Biologique de Roscoff, Place Georges Teissier, 29680
18 Roscoff, France

19 Email: viard@sb-roscoff.fr

20

21 **Running title:** Introgression between *Ciona* species

22

23 **Keywords:** Secondary contacts, Biological invasions, Tunicates, Contemporary hybridization,
24 Past introgression, Population genomics

25 **Abstract**

26 Biological introductions bring into contact species that can still hybridize. The
27 evolutionary outcomes of such secondary contacts may be diverse (e.g. adaptive introgression
28 from or into the introduced species) but are not yet well examined in the wild. The recent
29 secondary contact between the non-native sea squirt *Ciona robusta* (formerly known as *C.*
30 *intestinalis* type A) and its native congener *C. intestinalis* (formerly known as *C. intestinalis*
31 type B), in the western English Channel, provides an excellent case study to examine. To
32 examine contemporary hybridization between the two species, we developed a panel of 310
33 ancestry-informative SNPs from a population transcriptomic study. Hybridization rates were
34 examined on 449 individuals sampled in 8 sites from the sympatric range and 5 sites from
35 allopatric ranges. The results clearly showed an almost complete absence of contemporary
36 hybridization between the two species in syntopic localities, with only one first generation
37 hybrid and no other genotype compatible with recent backcrosses. Despite the almost lack of
38 contemporary hybridization, shared polymorphisms were observed in sympatric and allopatric
39 populations of both species. Furthermore, one allopatric population from SE Pacific exhibited
40 a higher rate of shared polymorphisms compared to all other *C. robusta* populations. Altogether,
41 these results indicate that the observed level of shared polymorphism is more probably the
42 outcome of ancient gene flow spread afterwards at a worldwide scale. They also emphasise
43 efficient reproductive barriers preventing hybridization between introduced and native species,
44 which suggests hybridization should not impede too much the expansion and the establishment
45 of the non-native species in its introduction range.

46 **Introduction**

47 Speciation is a gradual spatio-temporal process during which geographical or ecological
48 isolation decrease gene flow between groups of individuals (Abbott *et al.* 2013). Species range
49 shifts can deeply modify the evolution of these emerging species by promoting the formation
50 of contact zones (Hewitt 2004; Maggs *et al.* 2008; Swenson & Howard 2005). In cases of
51 species that are not fully reproductively isolated, interspecific gene flow occurs across hybrid
52 zones (Barton 1979; Hewitt 2011). Hybridization and introgression processes between species
53 in contact zones are particularly interesting to provide insights about the relative role of intrinsic
54 and extrinsic barriers in the maintenance of species boundaries (Abbott *et al.* 2013; Harrison &
55 Larson 2014; Hewitt 1988; Orr & Smith 1998; Turelli *et al.* 2001).

56 In last few years, next generation sequencing techniques has revolutionized the study of
57 hybridization and speciation processes (for a review, see Seehausen *et al.* (2014)). For instance,
58 recent population genomic studies have provided evidence that adaptive introgression can occur
59 between divergent species and may be more common than previously expected (Abbott *et al.*
60 2013; Hedrick 2013). The evolutionary histories of the modern human (for a review, see
61 Racimo *et al.* (2015)), the malaria vector mosquito *Anopheles gambiae* (Fontaine *et al.* 2015),
62 *Heliconius* butterflies (Pardo-Diaz *et al.* 2012) and *Mytilus* mussels (Fraisse *et al.* 2016) are
63 particularly well-documented cases illustrating such processes.

64 Most of these studies are concerned with historical interspecific gene flow which
65 occurred over a long time during periods of range expansion (see Currat *et al.* (2008) for
66 theoretical supports and review of empirical evidences). And yet adaptive introgression may
67 occur on much shorter time scale, as exemplified by introduction of species by human activities
68 which modify species distribution at a global scale and at an unprecedented rate (e.g. in marine
69 ecosystems, see Molnar *et al.* (2008)). Biological introductions provide a window on the early
70 phase of secondary contacts between previously allopatric and non-reproductively isolated

71 species. A diverse set of consequences of hybridization between native and non-native taxa are
72 expected (Allendorf *et al.* 2001) for instance, the extinction of the native species (Rhymer &
73 Simberloff 1996) or the introgression of advantageous alleles from the native into the non-
74 native species facilitating local adaptation of the non-native species to its new colonized
75 environment (Ellstrand & Schierenbeck 2000; Schierenbeck & Ellstrand 2009) or also the
76 opposite situation, i.e. the rapid fixation of non-native alleles in the genome of native species,
77 for example between the non-native Barred Tiger salamanders and the native California one
78 (Fitzpatrick *et al.* 2010).

79 In this context, we consider two newly reclassified although strongly differentiated
80 species in the genus *Ciona*. These two species were considered as cryptic species of the *Ciona*
81 *intestinalis* species complex and formerly named *C. intestinalis* type A and *C. intestinalis* type
82 B (Nydam & Harrison 2007; Zhan *et al.* 2010). Following recent taxonomic revision, they are
83 now accepted as two distinct species (WoRMS database) and respectively named *C. robusta*
84 and *C. intestinalis* (Brunetti *et al.* 2015). They display a divergence estimated at ca. 4 Mya
85 (Roux *et al.* 2013). Currently, the two species, and particularly *C. robusta*, display a large
86 distribution over several distinct biogeographic regions because both have been introduced by
87 human-activities (see Supplementary Note in Bouchemousse *et al.* (2016a)). For instance, *C.*
88 *robusta*, assumed to be native to NW Pacific, has been reported as a non-native species in
89 almost all the oceans. This species lives in sympatry with *C. intestinalis*, native to the NE
90 Atlantic, in only one region, namely in the Western English Channel and South of Brittany. It
91 has been shown that *C. robusta* was introduced in this region probably in the early 2000s
92 (Bishop *et al.* 2015; Nydam & Harrison 2011).

93 Despite their high divergence (i.e. 14% of transcriptomic divergence (Roux *et al.* 2013)
94 and 12-14% of mitochondrial divergence (Bouchemousse *et al.* 2016a; Nydam & Harrison
95 2007; Zhan *et al.* 2010)), the two species are not reproductively isolated: first generation (F1)

96 hybrids are easily obtained under laboratory conditions, with however an asymmetry according
97 to the maternal lineage (Bouchemousse *et al.* 2016b; Suzuki *et al.* 2015): F1 hybrids produced
98 in laboratory experiments are obtained in one direction only corresponding to crosses involving
99 oocytes of *C. intestinalis* and sperm of *C. robusta* (ca. 80% of fertilization rate against < 6% in
100 the opposite direction (Bouchemousse *et al.* 2016b)). The question of the extent of hybridization
101 in nature is thus to be addressed. Recent molecular studies carried out in the only sympatric
102 range described so far (i.e. NE Atlantic) suggest contemporary hybridization happens at a small
103 rate: despite a close syntopy and reproductive synchrony, a few putative hybrids (i.e. individuals
104 showing shared alleles on putative species-diagnostic markers) were observed in the wild, with
105 a paucity of F1s (Bouchemousse *et al.* 2016b; Nydam & Harrison 2011). In addition, low levels
106 of introgression were detected and interpreted by the presence of some backcross genotypes in
107 the samples (i.e. between 4 and 6%; Bouchemousse *et al.* 2016b; Nydam & Harrison 2011; Sato
108 *et al.* 2014). If this interpretation is true, this contemporary hybridization could have a profound
109 effect on the expansion of the non-native species but also on the native species (i.e. adaptive
110 introgression processes, see above). However, these studies were based on few nuclear loci
111 assumed species-diagnostic. An alternative explanation is that the low level of admixture
112 measured nowadays in the contact zone is the outcome of historical introgression during past
113 secondary contacts. Roux *et al.* (2013) adjusted a secondary contact model to a population
114 transcriptomic dataset and inferred, under this model of a single contact, that introgression
115 lasted for 15,000 years (95% CI: 4,300 - 56,800) during which ca. 20% of loci presumably
116 crossed the species barriers in both direction. Most probably the two taxa repeatedly came into
117 contacts both in past- and present time. This situation is prone to the misinterpretation of
118 contemporary admixture when few loci are used. Among the samples studied by Bouchemousse
119 *et al.* (2016b), what appear to be a few individuals with hetero-specific alleles could indeed be
120 a consequence of a low genome-wide level of past introgression (e.g. only 1 to 4% of the

121 genome of present-day non-African humans derived from gene flow between Neanderthals and
122 modern humans (Green *et al.* 2010)). In order to ascertain the extent of the contemporary
123 hybridization between the two taxa, we used a population genomic approach based on 310
124 ancestry-informative SNPs derived from full transcriptomic sequences (Roux *et al.* 2013). By
125 studying such a large number of markers on an extensive sampling, we could also evaluate the
126 discriminating power of the few nuclear markers that have been used so far (e.g. Bouchemousse
127 *et al.* 2016b; Nydam & Harrison 2011; Sato *et al.* 2014). We studied a large number of
128 individuals from eight localities of the sympatric range (i.e. contemporary contact zone) and
129 two to three localities outside contact zones for each species: these allopatric populations were
130 used as a control for the absence of contemporary gene flow between the two species. The SNP
131 panel developed in this study should prove useful in an ascidian species with importance in
132 evolutionary biology, invasion biology, development biology and phylogeny (Procaccini *et al.*
133 2011; Satoh *et al.* 2014; Zhan *et al.* 2015).

134

135 **Materials and Methods**

136 *Sampling*

137 Sampling of *Ciona robusta* and *C. intestinalis* was done within their contemporary
138 sympatric range (i.e. Western English Channel and South of Brittany) in seven localities where
139 the two species are living in syntopy (i.e. living in the same habitat) and one locality where
140 surveys carried out over three years never reported the presence of *C. robusta* (i.e. only *C.*
141 *intestinalis* is present; no.9 in Table 1 (Bouchemousse *et al.* 2016b)). For comparison,
142 populations from localities outside of the contemporary contact zone (i.e. where a unique
143 species has been recorded so far) were sampled: for *C. robusta*, two localities of the SE Pacific
144 and Mediterranean Sea, and for *C. intestinalis*, two localities in the North Sea (one in shallow
145 water and one at 20-meters depth) and one in the NW Atlantic (Table 1). For each individual,

146 DNA extraction was performed with Nucleospin® 96 Tissue Kit according to the
147 manufacturer's protocol (Macherey-Nagel, Germany). A minimum of 24 individuals per
148 population was selected based on the DNA quality following extraction. Altogether a total of
149 449 individuals, 213 for *C. robusta* and 236 for *C. intestinalis* were further analyzed. A
150 preliminary assignment to either *C. robusta* or *C. intestinalis* was based both on morphological
151 features (Brunetti *et al.* 2015; Sato *et al.* 2012). In addition to specimens sampled in natural
152 populations, two F1-hybrids produced from experimental crosses (Bouchemousse *et al.* 2016b)
153 were included as control for F1-hybrid genotype.

154

155 *Loci selection and genotyping*

156 An Illumina BeadXpress® with Veracode™ technology (GoldenGate® Genotyping
157 Assay) was used to genotype 384 single nucleotide polymorphisms (SNPs) selected from a SNP
158 dataset detected in the full transcriptomes of 10 individuals of *C. robusta* and 10 individuals of
159 *C. intestinalis* (details in Roux *et al.* (2013)). The loci were first chosen to maximize their
160 genotyping success: Because we used transcriptome data, we identified exon borders by
161 aligning our data with *C. robusta* genome (vKH.71 from Ensembl; note that the genome name
162 is misleading as it is labelled “*C. intestinalis*” although it is from *C. robusta* following the recent
163 taxonomic revision (Brunetti *et al.* 2015)). Polymorphic sites closer than 20bp from exon border
164 were automatically excluded. Polymorphic positions were selected within exons to produce an
165 individual sequence for each given SNP compatible with the Assay Design Tool (ADT)
166 software available on Illumina webpage. Sites with a minor allele frequency lower than 0.1
167 were excluded. ADT software was used to choose primers for each SNP and estimate
168 probability of amplification of each marker before amplification. Only markers with a
169 probability of amplification greater than 40% were retained. We selected this low minimum
170 value because of the high divergence between *C. robusta* and *C. intestinalis* at the full

171 transcriptome level (i.e. 14% according to Roux *et al.* (2013)) and thus the poor number of
172 genomic regions likely to be conserved between the two species. The average probability
173 obtained for our final SNP panel of 384 markers was however reaching 74%. Based on the
174 results by Roux *et al.* (2013), loci could be sorted according to four categories of polymorphism
175 (Table S1): 1) SNPs differentially fixed between the two species (sf), 2) SNPs polymorphic in
176 *C. robusta* (sxA) but not in *C. intestinalis*, 3) SNPs polymorphic in *C. intestinalis* (sxB) but not
177 in *C. robusta* and 4) SNPs displaying polymorphism in the two species (ss). The full SNP panel
178 was intentionally not random, for instance including a substantial number of SNPs differentially
179 fixed between the two species and shared polymorphisms when compared to the genome-wide
180 expectation. Selecting loci showing a high genetic differentiation between reference
181 populations, such as allopatric populations located far away of a sympatric zone or from either
182 side of a hybrid zone is common practice for discriminating recent admixed individuals (Bierne
183 *et al.* 2011; Larson *et al.* 2014). However, these loci display restricted introgression and might
184 not randomly associate during reproduction (Harrison & Larson 2016). Thus, we expanded the
185 SNP panel with polymorphic loci showing less divergence and shared between the two species.
186 In addition, the 384 SNPs were selected to be spread over most of the chromosomes of the
187 published genome of *C. robusta* (Dehal *et al.* 2002) and 25 of them were localized in
188 introgression hotspots identified by Roux *et al.* (2013). At the end, we enriched the panel with
189 sf (101) and ss (47) SNPs and equalized the number of sxA (109) and sxB (127) SNPs as *C.*
190 *intestinalis* is more polymorphic than *C. robusta*. A subset of 70 SNPs that strictly reflect the
191 genome wide site frequency spectrum (i.e. randomly selected) was included in the SNP panel.
192 Genotyping was performed using Genome Studio software (Illumina Inc.). Out of the 384
193 SNPs, 324 SNPs amplified successfully and 310 SNPs were retained for further statistical
194 analyses. Despite the expected low amplification score (probability of amplification *in silico* of
195 74%), these 310 SNPs displayed a high rate of genotyping success: we obtained a minimum of

196 97% of the individuals without missing data, and an unambiguous genotype assignment. This
197 SNP panel included 58 SNPs randomly selected over the initial subset of 70 SNPs (the
198 remaining 12 SNPs have not been successfully genotyped).

199 In order to investigate the properties of putative species-diagnostic markers used in
200 previous studies (Bouchemousse *et al.* 2016b, Caputi *et al.* 2007, Nydam & Harrison 2010,
201 2011, Sato *et al.* 2014), we also genotyped all the individuals on three nuclear loci, namely
202 Hox5 (Caputi *et al.* 2007) vAChTP and CesaA (Nydam & Harrison 2010) by PCR and PCR-
203 RFLP (details in Caputi *et al.* (2007) and Nydam & Harrison (2010)) and a putative maternal
204 species-diagnostic mitochondrial locus (mtCOI; details in Nydam & Harrison (2007)). Details
205 regarding the physical mapping of the nuclear loci are provided in Figure S1 and Table S2.

206

207 *Intra-specific analysis*

208 In order to compare sympatric and allopatric populations, genetic studies were carried
209 out for both *C. robusta* and *C. intestinalis*. Only loci that were polymorphic in the targeted
210 species were used. For the few SNPs chosen in the same contig, only one SNP was selected,
211 the one showing the maximum value of the minor allele frequency. Totals of 111 and 150
212 polymorphic loci were retained for *C. robusta* and *C. intestinalis*, respectively.

213 *Genetic diversity.* At the intra-specific level, for each population, the number of
214 polymorphic loci and the expected heterozygosity (H_e) were estimated using GENETIX v.4.05
215 software (Belkhir *et al.* 2004). Fixation index (F_{IS}) was estimated and departures from Hardy-
216 Weinberg equilibrium were tested in each population using GENEPOP v4 with default
217 parameters for tests. P -values resulting of multiple tests were adjusted using the R package v.
218 3.1.3 (R Development Core Team 2014) QVALUE (Storey 2002).

219 *Genetic structure.* Genetic structure between populations was analyzed by estimating
220 the fixation index F_{ST} (Wright 1951) using GENEPOP. Exact G test for population

221 differentiation were carried out using 10,000 random permutations. To visualize the genetic
222 structure between populations, a Discriminant Analysis of Principal Components (DAPC;
223 Jombart *et al.* 2010)) was computed for each species separately using the R package
224 ADEGENET v.1.4 (Jombart & Ahmed 2011).

225

226 *Genome and population admixture analysis between the two species*

227 To identify putative F1- or recently introgressed individuals (product of several
228 generations of backcrosses within the sympatric range), a Bayesian clustering method
229 implemented in NEWHYBRID software v.1.1 was used (Anderson & Thompson 2002) using
230 the global dataset of 310 loci, the dataset of 58 random SNPs, as well as a dataset of the 105
231 most differentiated loci. Briefly, this method computes posterior probability for assigning a
232 given individual to different hybrid categories (i.e. F1, F2-hybrids and backcrosses with
233 parental *C. robusta* or *C. intestinalis* individuals) or parental species, using Markov chain
234 Monte Carlo algorithms (MCMC). Here, we considered allopatric populations as representative
235 of the parental species. We ran five independent analyses (each using a different random
236 starting value) with 500,000 MCMC after a period of 500,000 burn-in cycles using the Jeffreys-
237 like prior.

238 To examine inter-specific gene flow at the genome level, we used the R package
239 INTROGRESS (Gompert & Buerkle 2010). As the ancestral allelic state has to be defined
240 (based on the minor allele frequency in each species), we selected 105 loci that were the most
241 differentiated loci according to F_{ST} values ($F_{ST} > 0.9$) computed between the two species using
242 allopatric populations (Figure S2). Such a selection is common practice in hybridization studies
243 (e.g. between *Mytilus* species (Saarman & Pogson 2015) or *Gryllus* crickets (Larson *et al.* 2013,
244 2014)). For each individual, the maximum likelihood value of hybrid index was estimated. The
245 hybrid index (h) is defined as the proportion of *C. intestinalis* alleles over all loci ($h = 0$ for

246 individuals with *C. robusta* alleles only and $h = 1$ for individuals with *C. intestinalis* alleles
247 only). To visualize and compare the genomic architecture of interspecific admixture at the
248 individual level, the *mk.image* function implemented in INTROGRESS was used.

249 NEWHYBRIDS and INTROGRESS analyses were also done using a small dataset
250 made of the three nuclear markers used in previous studies as putatively species-diagnostic (i.e.
251 Hox5, vAChTP and CesA, see “*Loci selection and genotyping*” section above).

252 The inter-specific admixture rate was investigated on the total dataset (i.e. 310 loci)
253 using a Bayesian clustering method implemented in STRUCTURE v.2.3 (Pritchard *et al.* 2000)
254 and a Principal Component Analysis (PCA) using the R package ADEGENET v.1.4 (Jombart
255 & Ahmed 2011). The method implemented in STRUCTURE method used MCMC to generate
256 posterior probabilities of assignment of each individual genotype to a number of clusters (K).
257 Ten replicates of 500,000 MCMC after a period of 500,000 burn-in cycles were ran for K values
258 ranging from 1 to 4; K = 2 is corresponding to the two species clusters. Results were
259 summarized across all replicate runs using CLUMPP v.1.1.2 (Jakobsson & Rosenberg 2007)
260 and visualized with DISTRUCT v1.1 (Rosenberg 2004). To check for the absence of biases due
261 to marker selection, we also ran STRUCTURE for K = 2 using additional dataset: 1) a subset
262 of 58 loci selected to represent a random sampling of the genome, 2) a subset of 105 loci
263 selected for the INTROGRESS analysis, 3) 245 SNPs corresponding to all the SNPs except
264 those differentially fixed between the two species, 4) only those SNPs that were polymorphic
265 in the two species (42 SNPs).

266 To better evaluate the evolutionary history between *C. robusta* and *C. intestinalis*
267 populations, we used a population graph approach implemented in the TREEMIX program
268 (Pickrell & Pritchard 2012), which infers patterns of splitting and migration between
269 populations. By using the matrix of allele frequency covariance between pair of populations,
270 this method generates maximum likelihood population trees under the hypothesis of an absence

271 of migration or the alternative hypothesis of migration event(s) (that are sequentially added).
272 By comparison with other methods commonly used to make demographic inferences or to
273 picture population relationships over time (e.g. Beaumont *et al.* 2002; Gutenkunst *et al.* 2009;
274 Hey & Nielsen 2001), TREEMIX has the advantage to be applicable to a large number of
275 populations by using a tree-construction based approach and, at once, testing for gene flow
276 between populations (Pickrell & Pritchard 2012). To avoid noises due to small sample sizes
277 and intra-specific migration (i.e. infra-specific admixture), we pooled populations according to
278 the region of sampling (i.e. no.4a and 4b for *C. intestinalis*; no.5 and 6 and no. 7 to 12 for each
279 species). Using the total dataset (i.e. 310 loci), we search for the best tree to fit the data testing
280 for a range of migration events from 0 to 8 (reaching an asymptotic value, Figure S3). Based
281 on these inferences, we used a block-jackknife procedure with blocks of 10 SNPs to determine
282 which migration events significantly improved the model fit. A complementary analysis based
283 on f_3 -statistic test, developed by Reich *et al.* (2009), was done to test the null hypothesis that
284 the evolutionary history of *Ciona* populations was consistent with the absence of migration
285 events between populations. The f_3 -statistics evaluates the deviation of the null hypothesis using
286 the same block-jackknife procedure for all combinations of three populations (one used as the
287 target and two tested as putative ancestral populations).

288

289 **Results**

290 *Diversity of the SNP panel*

291 Overall, 451 individuals (including the two F1-hybrids from experimental crosses) were
292 genotyped successfully at 310 SNPs defined from a transcriptome dataset of *Ciona robusta* and
293 *C. intestinalis* (Roux *et al.* 2013). Following this genotyping, the distribution of SNPs across
294 the categories, defined from a small sample of 10 specimens for each species, was modified, as
295 shown in Table S1. The most substantial change was a decrease of the sf and sxA categories

296 (i.e. a decrease of 31% and 22%, respectively) and a concomitant increase of the sxB and ss
297 categories (17% and 22%, respectively). We considered these new categories in the analyses
298 below.

299

300 *Population genetic structure and little heterozygosity variations in the two study* 301 *species*

302 The analyses aiming at comparing allopatric and sympatric populations of each species
303 were carried out separately for *C. robusta* and *C. intestinalis*, using the set of loci polymorphic
304 in each species, i.e. 111 and 150 SNPs respectively. Results are summarized in Table 2 and
305 Table 3.

306 *Diversity and genetic structure in populations of C. robusta*

307 Values of *He* were similar across populations of *C. robusta*, ranging from 0.234 (no.2) to 0.288
308 (no.1). No departure from Hardy-Weinberg equilibrium (HWE) was found in any of the study
309 populations. Exact test of differentiation revealed significant differences in allele frequencies
310 among all populations sampled and among populations of the sympatric range (Table 3). The
311 highest genetic differentiation was observed between the SE Pacific population (no.1) and all
312 of the other populations (pairwise comparisons provided in Table S3a) as well-illustrated by
313 the DAPC (Fig.1a) along the first discriminant axis. The second discriminant axis pointed out
314 the differentiation between populations of UK (i.e. no.5 and 6) and Mediterranean Sea (no.2)
315 which is confirmed by significant pairwise estimates of F_{ST} (Table S3a). Populations of Brittany
316 were relatively poorly differentiated between them (non-significant F_{ST} values in most of
317 pairwise comparisons, Table S3a). Altogether, SE Pacific and to a lesser extent UK and
318 Mediterranean Sea populations were the most different genetically.

319 *Diversity and genetic structure in populations of C. intestinalis*

320 Values of H_e were similar among the study populations, ranging from 0.240 (no.12) to 0.229
321 (no.5 and 6), except for the populations from the North Sea which exhibited lower values of H_e
322 (i.e. 0.194 and 0.172 for no.4a and 4b respectively). As for *C. robusta*, no departure from HWE
323 was observed in any study populations. Exact test of differentiation between *C. intestinalis*
324 populations indicated significant differences among all populations but was non-significant
325 between populations of the sympatric range (Table 3). The overall significant genetic structure
326 was mainly due to a strong and significant genetic differentiation of the populations sampled in
327 the two allopatric regions (no.3, 4a and 4b) and of one population sampled in the sympatric
328 range (no.6) with almost all other populations (pairwise comparisons are provided in Table
329 S3b). These patterns are pictured by the DAPC (Fig.1b).

330

331 *Low hybrid index disregarding the regional category and population status*

332 A total of 105 loci, showing a F_{ST} strictly superior to 0.9, were used with the R package
333 INTROGRESS to examine the patterns of shared polymorphism between the two species in the
334 contact zone and in allopatric populations. At the species level (i.e. across all individuals for
335 each species), values of h were very low, with an average value across individuals of 0.0029
336 for *C. robusta* and 0.0055 for *C. intestinalis*. Table 1 is providing the average values of the
337 hybrid index (h) for each population of *C. robusta* and *C. intestinalis*. A noticeable result was
338 the presence of one individual in this latter population with an h value of 0.5. When removing
339 this individual from the h estimation, the value in Camaret dropped to 0.006, a value close to
340 the average values for *C. intestinalis* populations. This individual was assigned with a
341 probability of 1 to a 'F1 hybrid' with NEWHYBRIDS (Table 1) with the 105 SNPs dataset; a
342 result confirmed with the full (310 SNPs) and the random (58 SNPs) dataset. The two F1-
343 hybrids obtained experimentally were also assigned with a probability of 1 to the 'F1-hybrid'
344 category which ascertains the robustness of NEWHYBRIDS for detecting individuals derived

345 from recent crosses. It is noteworthy that all the other study individuals were assigned to their
346 respective parental ‘species’ categories (Table 1). We also examined the relationship between
347 h -value and the heterozygosity rate across the 105 loci used with INTROGRESS, by using a
348 triangle plot displayed in Figure 2: all except one individual displayed extreme h -values (closed
349 to 0 or 1) and an extremely low proportion of heterozygote loci for *C. robusta* and *C. intestinalis*
350 alleles. The only exception is the individual sampled from Camaret (no.11) that was assigned
351 by NEWHYBRIDS as a F1-hybrid: this individual showed both a high h -value and a high
352 heterozygosity rate (i.e. 99%); these values were similar to the values observed for the two F1-
353 hybrids experimentally produced in the laboratory (Fig.2, i.e. 96% and 99%). STRUCTURE
354 analyses with $K = 2$ assigned equally the putative wild F1-individual and the two experimental
355 F1-hybrids to the two species clusters (Fig.3b); and the results of the PCA showed a clear
356 distribution of the overall genetic variance between the two study species with the natural and
357 experimental F1- hybrids at an intermediate position (Fig.3a). This finding is also observed on
358 STRUCTURE analyses done with the additional subset of loci (Figure S4; see *Material and*
359 *Methods* section) suggesting an absence of biases due to marker selection. Note that increasing
360 K values ($K = 3$ and $K = 4$, Fig.3b) in the STRUCTURE analysis confirmed intraspecific
361 variance observed with DAPC (Fig.1), notably the genetic differentiation of the SE Pacific
362 population (no.1) with all of the other populations of *C. robusta* and of the two sub-populations
363 of Fiskebackskil with all of the other populations of *C. intestinalis*.

364 Using the three nuclear markers used as species-diagnostic markers in previous studies
365 (i.e. Hox 5, vAChTP and CesA; see *Material & Methods*), an interesting pattern was observed:
366 the single F1 individual otherwise identified with the complete set of SNPs displayed a
367 heterozygote genotype at two loci (i.e. vAChTP and Hox5) but a homozygote genotype (two
368 *C. robusta* alleles) at CesA locus; a pattern clearly inconsistent with a F1-hybrid genotype if
369 loci are assumed fully diagnostic. In addition, NEWHYBRIDS assigned with a high probability

370 ($P = 0.95$) this F1-hybrid individual to the category of individuals backcrossed with *C. robusta*.
371 This result illustrates the inherent difficulty to accurately account for the sampling variance
372 when alleles are fixed in the samples or nearly so. Among the other individuals, 416 (92.8%)
373 were assigned to their respective parental 'species' categories with a probability above to 0.95,
374 while 32 individuals (7.2%) obtained ambiguous results with a posterior probability of being
375 parental genotypes ranging from 0.63 to 0.95. At the mtCOI locus (i.e. a putative maternal
376 species-diagnostic mitochondrial locus), there is a strict association between the preliminary
377 morphological assignment and the mitochondrial type (Fig.4b). As expected based on
378 experimental studies (i.e. asymmetry in reproductive success according to the maternal type, see
379 Introduction section), the F1 hybrid showed a *C. intestinalis*-mitochondrial type.

380

381 *Heterogeneous polymorphism rates at the genome level*

382 The subset of 105 SNPs used for examining the patterns of shared polymorphism
383 between the two species (i.e. showing F_{ST} values higher than 0.9) showed the following patterns
384 over the whole dataset: 65 were differentially fixed (i.e. sf loci), 39 with private polymorphisms
385 (i.e. sxA or sxB; 8 polymorphic in *C. robusta* and 31 in *C. intestinalis*) and only one locus
386 showing shared polymorphism (i.e. ss locus) between populations of the two species (snp18 on
387 the chromosome 1). The 40 loci showing shared and private polymorphisms were distributed
388 randomly along the genome of the two species (Fig.4a), as previously observed in Roux *et al.*
389 (2013). Among the 105 SNPs, some were found in two introgression hotspots defined by Roux
390 *et al.* (2013). Note that we did not have SNPs localized in the other two introgression hotspots.
391 Interestingly, we found shared or polymorphic SNP in these introgression hotspots: 1) one SNP
392 in the introgression hotspot of chromosome 1 showed shared polymorphism in the two species,
393 and 2) six loci showed private polymorphism in one or the other of the two species in the
394 introgression hotspot on the chromosome 2.

395 Polymorphism patterns were also informative regarding the status of the study
396 populations (i.e. allopatric and sympatric). Details of allele frequencies at each of the 40
397 polymorphic loci in the allopatric and sympatric ranges of the two species are provided in Table
398 S4. When comparing of populations for *C. robusta* and *C. intestinalis*, the rates of shared and
399 private polymorphism appeared to be remarkably stable across populations: for the two species,
400 individuals of each population carried a small number of heterozygous sites, but not always at
401 the same genome location (Fig.4a). A noteworthy exception was the allopatric population of *C.*
402 *robusta* from Chile (no.1) which shared for some loci more polymorphism with *C. intestinalis*
403 populations than with other populations of *C. robusta* (Fig.4a): heterozygous sites were more
404 important, for example, at the snp18 (chromosome 1), snp290 (chrom. 2) and snp237 (chrom.
405 10), the two first being in introgression hotspots defined by Roux *et al.* (2013). This finding
406 was already visible in the results of the PCA (Fig.3a) as the population of Chile was slightly
407 shifted towards the *C. intestinalis* points

408 The random distribution of few shared and private polymorphisms was also observed
409 when using the three putative species-diagnostic markers (Fig.4b): minor allele frequency
410 observed for Hox5, vAChTP and CesA was 0.2, 0.2 and 1.3% for *C. intestinalis*, and 6.4, 1.2
411 and 0.5% for *C. robusta*, respectively. None of them were localized in introgression hotspots.
412 As for the 40 polymorphic SNPs discussed above, the rate of polymorphism was quite stable
413 across populations with a small number of heterozygous sites.

414

415 *Admixture events between the two species revealed by a population tree approach*

416 The population tree inferred from TREEMIX without migration explained 88.5% of the
417 variance in the population covariance matrix. Note that in the population tree without migration
418 events, the population of Chile (no.1) showed a position shifted towards *C. intestinalis*
419 populations (Figure S5). The variance explained was increased when migration events were

420 added (Figure S3). The best fit to the data was obtained with two migration events, which
421 significantly improved the model ($P < 0.001$, Fig.5). This population tree, explaining 98.5% of
422 the variance (Figure S3), indicated significant gene flow in *C. robusta* population of Chile
423 (no.1) and in the *C. intestinalis* populations group of Brittany (no.7 to 12). These two migration
424 events were also supported by the f_3 statistics analysis (Table S5) with significant negative
425 values for almost all combinations of three populations involving as targets the *C. robusta*
426 population of Chile (no.1) and the *C. intestinalis* populations group of Brittany (no.7 to 12). f_3
427 statistics also showed significant negative values for combinations of three populations
428 involving as targets *C. robusta* populations groups of UK (no.5 and 6) and Brittany (no.7 to 12)
429 (Table S5). These negative f_3 statistics are consistent with the hypothesis that the tested
430 populations were the results of admixture with ancestors in the two tested population sources
431 (Reich *et al.* 2009).

432

433 **Discussion**

434 In this study we used 310 ancestry-informative SNPs to clarify relative contribution of
435 contemporary hybridization *versus* past introgression in the level of shared polymorphism
436 observed between *Ciona robusta* and *Ciona intestinalis*, and to analyze the introgression
437 patterns within allopatric and sympatric ranges of the two species. These two points are
438 discussed in turn below.

439

440 *Absence of contemporary interspecific gene flow in the sympatric range*

441 In previous studies that analysed interspecific gene flow in the sympatric range,
442 admixture have been observed between the two species although at low rates: 4.2% (including
443 one putative F1-hybrid) over 730 individuals sampled in Nydam & Harrison (2011), 6.3% over
444 288 individual sampled in one locality by Sato *et al.* (2014), 4.3% (including one putative F1

445 hybrid) over ca. 3,000 individuals by Bouchemousse *et al.* (2016b). For examining the extent
446 of hybridization between the two species, these authors used few nuclear markers (between 3
447 and 6 loci according to the study) which were supposed to be species-diagnostic. Consequently,
448 discriminating the footprint left by historical introgression *versus* contemporary hybridization
449 was particularly difficult as acknowledged in these studies (Bouchemousse *et al.* 2016b; Nydam
450 & Harrison 2011). Using a large number of informative loci with a large sample, we found that
451 with the exception of a single individual all the other occurrence of shared polymorphism are
452 the likely consequence of a low level of past introgression, i.e. ancient gene flow between the
453 two species, rather than contemporary hybridization. At a given locus some individuals can be
454 found heterozygotes at quasi diagnostic loci, but averaging at many such loci shows every
455 individual have the same hybrid index value. The only contemporary hybrid was a F1, as
456 supported by both NEWHYBRIDS (whatever the dataset used) and INTROGRESS analyses
457 (Table 1, Fig.2 and 4a). The mtDNA type of this individual is typical of *C. intestinalis* which
458 corresponds in many studies to crosses easily produced in laboratory experiments
459 (Bouchemousse *et al.* 2016b; Suzuki *et al.* 2005). The presence of one F1 hybrid only in our
460 study confirms the hypothesis by Sato *et al.* (2014) and Bouchemousse *et al.* (2016b) of the
461 existence of pre-zygotic isolation mechanisms preventing contemporary hybridization in the
462 wild. However, our new interpretation that recent backcrosses of a few generations are
463 completely lacking from the sympatric range, suggest that strong post-zygotic selection is also
464 occurring, which is the least one can expect for two highly divergent species (i.e. 14% of
465 divergence based on transcriptomic data (Roux *et al.* 2013)). Dobzhansky-Muller
466 incompatibilities expressed by recessive mutations in subsequent generations of hybridization
467 (e.g. Bierne *et al.* (2006); Fishman & Willis (2001) and see for a review Maheshwari & Barbash
468 (2011)) are likely the cause of this isolation. Altogether these results confirm that contemporary
469 gene flow is almost inexistent between the two species.

470

471 *The footprint of past introgression between the two species*

472 Hybrid index and interspecific heterozygosity values were similar whatever the region
473 (sympatric or allopatric) and locality status (syntopic vs. non-syntopic, Table 1, Fig. 2). This
474 finding validates that shared polymorphism were observed in all populations including in
475 localities of allopatric regions (Fig.4a). Footprints of gene flow were also observed and were
476 significant for some of them according to TREEMIX and f_3 -statistics analyses.

477 For a given locus, shared polymorphism or high derived allele frequency between two
478 species may result from incomplete lineage sorting of ancestral polymorphism, contemporary
479 or past secondary introgression, or homoplasy. In the case of the two *Ciona* species studied
480 here, the contemporary introgression hypothesis can be reasonably excluded as discussed
481 above. Concerning incomplete lineage sorting of ancestral polymorphism, it would have meant
482 that the polymorphism observed nowadays would have been maintained randomly across loci
483 after the allopatric divergence estimated to have occurred during the Pliocene (between 2.59
484 and 5.33 My (Roux *et al.* 2013)). Considering the long time elapsed since the divergence, the
485 probability of occurrence of the two ancestral alleles in both daughter species is likely to be
486 extremely low under a neutral model (Pamilo & Nei 1988). High effective population sizes
487 moderates the effect of genetic drift and so the probability of fixation of alleles over the time
488 (Maddison 1997; Pamilo & Nei 1988). *Ciona* species and their common ancestor were
489 characterized by high effective population sizes, estimated in Roux *et al.* (2013), as between
490 115,000 and 395,000 for *C. robusta*, 748,000-1,022,000 for *C. intestinalis* and 1,606,000-
491 2,220,000 for the common ancestor. However, the analysis of Roux *et al.* (2013) showed that
492 the strong excess of shared polymorphism between the two species cannot be obtained without
493 secondary introgression. The secondary contact has been estimated to have occurred 15,500
494 years ago (95% CI: 4,300-56,800), during which ca. 20% of loci crossed the species barrier in

495 both directions. Besides similarities in admixture levels across localities, the hypothesis of an
496 ancient admixture event is also well-supported 1) by significant admixture events between
497 populations of *C. robusta* and *C. intestinalis* according to TREEMIX and f_3 -statistics analyses
498 and 2) the presence of admixed loci in introgression hotspots (i.e. loci pointed by an asterisk in
499 Fig.4a).

500 Our finding is also interesting to consider in light of previous studies based on a few
501 markers used as species-diagnostic markers of the two study species (e.g. (Bouchemousse *et al.*
502 2016b; Nydam & Harrison 2011; Sato *et al.* 2014)). Analyzing a small number of such loci can
503 easily result in the erroneous interpretation that some individuals are more admixed than other
504 and cast doubts about the ability of these markers to reliably distinguish the two species. This
505 was shown in our study by comparing the results obtained with 310 SNPs vs. three markers
506 supposed to be species-diagnostic. In particular, the *CesA* locus showed a homozygote
507 genotype with two *C. robusta* alleles in the single F1 individual otherwise identified with the
508 complete set of SNPs. With the subset of three markers, this F1-hybrid individual was
509 consequently mistaken as a backcrossed individual and not assigned to the F1-hybrid category,
510 with NEWHYBRIDS.

511 These results highlight the risks of using putative species-diagnostic markers without
512 preliminary knowledge about the likelihood of past introgression between two study taxa. The
513 species complex of *Mytilus* species is another well-known case study: *Glu* and *mac-1* loci were
514 mistakenly considered as diagnostic makers for *M. galloprovincialis* and *M. edulis* at a global
515 scale (Borsa *et al.* 2007; Borsa *et al.* 2012), but were later shown to have been historically
516 introgressed during secondary contact(s) caused by glacial oscillations (Roux *et al.* 2014).

517

518 *Difference of introgression rate in Chile caused by adaptive introgression?*

519 Admixture profiles were remarkably stable across populations of allopatric and
520 sympatric ranges. This widespread interspecific admixture suggest that range expansion of the
521 two species, through both natural range shifts (with long-term environmental changes) and/or
522 human-mediated introductions, occurred after a primary episode of contact between the two
523 taxa, during which interspecific gene flow occurred. Genetic differentiations are however
524 reported between allopatric and sympatric populations for the two species (Table 3) suggesting
525 that intraspecific divergence history for each species influence more the genetic differentiation
526 between populations than different rates of introgression between species. For example, the two
527 sub-populations of North Sea (i.e. no.4a and 4b sampled both at Fiskebackskil at the surface
528 and at 20m depth, respectively) exhibited a strong genetic differentiation with the other *C.*
529 *intestinalis* populations and also between them (Fig.1b, Table S3) while they showed similar
530 hybrid index values (Table 1). This strong genetic differentiation could be explained by a
531 reduced gene flow between the two sub-populations in North Sea, a result which echoed to the
532 pattern described in the doctoral thesis of Elin Renborg
533 (<https://gupea.ub.gu.se/handle/2077/35128>). The poor connectivity is hypothesized to result of
534 density discontinuity of sea water which separates shallow and deep populations of *C.*
535 *intestinalis*. Such patterns of population differentiation have already been documented in other
536 coastal marine species showing extended distribution along depth gradient (Jennings *et al.*
537 2013; Pivotto *et al.* 2015).

538 A noteworthy exception of the stability of admixture profiles is the *C. robusta*
539 population from Chile which showed the highest number of loci with shared polymorphism
540 with *C. intestinalis* (Fig.4a) and the highest *h*-values over all *C. robusta* populations (Table 1).
541 Moreover, the position of the Chilean population on the first axis of the PCA (Fig.2a) first
542 suggests residual genotypic covariance best explained by a higher level of introgression by *C.*
543 *intestinalis* than other *C. robusta* populations. This is formally tested using TREEMIX and f_3

544 statistical analyses (Fig.5, Table S5) which highlighted significant migration events between *C.*
545 *intestinalis* ancestor and the Chilean population. Incomplete lineage sorting of ancestral
546 polymorphism is not expected to create such asymmetry of shared polymorphism between
547 populations, but point out evidence of local introgression in the Chilean population (Fraisse *et*
548 *al.* 2016; Martin *et al.* 2013; Pickrell & Pritchard 2012). This pattern of local introgression is
549 not uniformly distributed among loci (Figure S6), which is usually not accounted for in
550 demographic inferences such as TREEMIX and could explain why the source of admixture is
551 not a contemporary *C. intestinalis* population. This pattern could be a consequence of adaptive
552 introgression in the genomic region of these introgressed loci, a process documented in several
553 recent studies (Fontaine *et al.* 2015; Mendez *et al.* 2012; Pardo-Diaz *et al.* 2012). A similar
554 pattern was observed in the *Mytilus* mussel complex of species where local introgression proved
555 to be heterogeneous across loci (Fraisse *et al.* 2016). However, other processes can generate
556 heterogeneous introgression rates such as heterogeneous load of deleterious mutations in
557 migrant tracks (Christe *et al.* 2016; Harris & Nielsen 2016). None of the loci identified matched
558 with genes coding for a known phenotypic or physiological trait (Table S4).

559 It is important to note that the first report of *C. robusta* along the Chilean coasts (with
560 the name of *C. intestinalis* used until the recognition of *C. robusta* as a valid species) dates back
561 to the middle of the 20th century (Van Name 1945). We thus cannot exclude that local
562 introgression have occurred in the source population(s) of the populations introduced in Chile
563 rather than after the introduction (as an outcome of selection in the Chilean introduction range).
564 A recent phylogeographic study based on mtDNA data (Bouchemousse *et al.* 2016a) pointed
565 out a low genetic differentiation between populations of Chile and populations sampled in
566 Japan, the putative native range of *C. robusta*. Further analyses are needed to investigate if this
567 pattern could be due to adaptive introgression, using for instance modelling methods such as
568 those performed by Fraisse *et al.* (2014) in a *Mytilus* sp. hybrid zone to examine the likelihood

569 of adaptive introgression. A much larger number of population representatives of the global
570 distribution of *C. robusta*, particularly populations of the Asian range, is also needed to
571 investigate the processes that occurred in the SE Pacific as compared to the other regions where
572 *C. robusta* is nowadays distributed.

573

574 In conclusion, our study confirmed the almost complete absence of contemporary gene
575 flow in the human-mediated contact zone wherein *C. robusta* and *C. intestinalis* co-exist in
576 sympatry/syntopy. Efficient reproductive barriers seem to prevent hybridization in the wild
577 between the two species. These results are casting doubts that hybridization could impede the
578 spread of the non-native. Ecological processes (e.g. niche displacement, trophic competition)
579 might thus be more important to determine the fate of the two species in the sympatric range.
580 Even if efficient reproductive isolation mechanisms are acting, few crosses involving an
581 advantageous allele can be sufficient to favor its transmission in subsequent generations of the
582 non-native species (Hedrick 2013). Our density of markers was clearly not sufficient to detect
583 local signatures of adaptive introgression at genomic level. High-throughput genome analyses
584 will be needed to definitively exclude, or confirm, that invasion potential of *C. robusta* is
585 facilitated by adaptive introgression with *C. intestinalis* in the Northeast Atlantic. This approach
586 might also allow us to identify genomic regions completely devoid of introgression, which may
587 correspond to impassable reproductive barriers. Altogether, our study provides evidence that
588 what was inferred to be recently introgressed individuals are more likely the outcome of a low
589 level of residual historical introgression redistributed at global scale by natural range shifts and
590 human-mediated introductions. Local introgression patterns, mostly concentrated on a few
591 genome regions, were observed in the population sampled in the SE Pacific, a population far
592 from the current distribution range of *C. intestinalis*. This result paves the way for further work

593 to investigate adaptive introgression processes in other regions, in light of the range shift history
594 of *C. robusta*.

595

596 **Acknowledgment**

597 The authors are very grateful to C. Roux and N. Galtier for making available RNA
598 sequences from the Pophyl Project, K. Belkhir of the Montpellier Bioinformatics Biodiversity
599 computing platform for his precious help to the optimization of loci selection and the ADN^{id}
600 society (Montpellier) for the genotyping of SNPs. We are also grateful to all our colleagues
601 who contributed to the collection of samples: the divers of the Marine Operations department
602 (*Service Mer & Observation*) at the Roscoff Biological Station, J.D.D. Bishop, S. Krueger-
603 Hadfield, B. Lundve, J. Pechenik. The authors kindly acknowledge C. Roux and C. Fraisse for
604 help and advices on R packages and scripts and three anonymous reviewers for their helpful
605 comments and advices on the manuscript. This work benefitted from funding of the ANR
606 project HYSEA (no. ANR-12-BSV7-0011) and the Interreg IVa Marinexus project, and a
607 Languedoc-Roussillon Region “Chercheur(se)s d’avenir 2011” grant to NB.

608

609 **Data accessibility**

610 Full dataset of the 310 SNPs were deposited into the DRYAD database (DOI:
611 10.5061/dryad.1h9b1).

612

613 **Author contributions**

614 SB, CLH, NB and FV designed the study. CLH and NB with contribution of SB and FV
615 performed the choice of SNP panel. SB and FV performed the choice of populations for
616 genotyping. SB, CLH, NB and FV analyzed the data and wrote the article.

617 **References**

- 618 Abbott R, Albach D, Ansell S, *et al.* (2013) Hybridization and speciation. *Journal of*
619 *Evolutionary Biology* **26**, 229-246.
- 620 Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting
621 conservation guidelines. *Trends in Ecology & Evolution* **16**, 613-622.
- 622 Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids
623 using multilocus genetic data. *Genetics* **160**, 1217-1229.
- 624 Barton N (1979) Dynamics of hybrid zones. *Heredity* **43**, 341-359.
- 625 Beaumont MA, Zhang W, Balding DJ (2002) Approximate bayesian computation in population
626 genetics. *Genetics* **162**, 2025-2035.
- 627 Belkhir K, Borsa P, Goudet J, *et al.* (2004) Genetix 4.05, logiciel sous Windows™ pour la
628 génétique des populations. *Laboratoire Génome, Populations, Interactions, CNRS*
629 *UMR 5171, Université de Montpellier II, Montpellier (France).*
- 630 Bierne N, Bonhomme F, Boudry P, Szulkin M, David P (2006) Fitness landscapes support the
631 dominance theory of post-zygotic isolation in the mussels *Mytilus edulis* and *M.*
632 *galloprovincialis*. *Proceedings of the Royal Society B-Biological Sciences* **273**, 1253-
633 1260.
- 634 Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why
635 genome scans may fail to map local adaptation genes. *Molecular Ecology* **20**, 2044-
636 2072.
- 637 Bishop JDD, Wood CA, Yunnice ALE, Griffiths CA (2015) Unheralded arrivals: non-native
638 sessile invertebrates in marinas on the English coast. *Aquatic Invasions* **10**, 249-264.
- 639 Borsa P, Daguin C, Bierne N (2007) Genomic reticulation indicates mixed ancestry in
640 Southern-Hemisphere *Mytilus* spp. mussels. *Biological Journal of the Linnean Society*
641 **92**, 747-754.
- 642 Borsa P, Rolland V, Daguin-Thiebaut C (2012) Genetics and taxonomy of Chilean smooth-
643 shelled mussels, *Mytilus* spp. (Bivalvia: Mytilidae). *Comptes Rendus Biologies* **335**, 51-
644 61.
- 645 Bouchemousse S, Bishop JDD, Viard F (2016a) Contrasting global genetic patterns in two
646 biologically similar, widespread and invasive *Ciona* species (Tunicata, Ascidiacea).
647 *Scientific Reports* **6**, 24875.
- 648 Bouchemousse S, Lévêque L, Dubois G, Viard F (2016b) Co-occurrence and reproductive
649 synchrony do not ensure hybridization between an alien tunicate and its interfertile
650 native congener. *Evolutionary Ecology* **30**, 69-87.
- 651 Brunetti R, Gissi C, Pennati R, *et al.* (2015) Morphological evidence that the molecularly
652 determined *Ciona intestinalis* type A and type B are different species: *Ciona robusta*
653 and *Ciona intestinalis*. *Journal of Zoological Systematics and Evolutionary Research*
654 **53**, 186-193.
- 655 Caputi L, Andreakis N, Mastrototaro F, *et al.* (2007) Cryptic speciation in a model invertebrate
656 chordate. *Proceedings of the National Academy of Sciences of the United States of*
657 *America* **104**, 9364-9369.
- 658 Christe C, Stölting KN, Bresadola L, *et al.* (2016) Selection against recombinant hybrids
659 maintains reproductive isolation in hybridizing *Populus* species despite F1 fertility and
660 recurrent gene flow. *Molecular Ecology* **25**, 2482-2498.
- 661 Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: Massive
662 introgression by local genes. *Evolution* **62**, 1908-1920.
- 663 Dehal P, Satou Y, Campbell RK, *et al.* (2002) The draft genome of *Ciona intestinalis*: Insights
664 into chordate and vertebrate origins. *Science* **298**, 2157-2167.

665 Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of
666 invasiveness in plants? *Proceedings of the National Academy of Sciences of the United*
667 *States of America* **97**, 7043-7050.

668 Fishman L, Willis JH (2001) Evidence for Dobzhansky-Muller incompatibilities contributing to
669 the sterility of hybrids between *Mimulus guttatus* and *M. nasutus*. *Evolution* **55**, 1932-
670 1942.

671 Fitzpatrick BM, Johnson JR, Kump DK, *et al.* (2010) Rapid spread of invasive genes into a
672 threatened native species. *Proceedings of the National Academy of Sciences of the*
673 *United States of America* **107**, 3606-3610.

674 Fontaine MC, Pease JB, Steele A, *et al.* (2015) Extensive introgression in a malaria vector
675 species complex revealed by phylogenomics. *Science* **347** (6217), 1258524.

676 Fraisse C, Belkhir K, Welch J, Bierne N (2016) Local interspecies introgression is the main
677 cause of extreme levels of intraspecific differentiation in mussels. *Molecular Ecology*
678 **25**, 269-286.

679 Fraisse C, Roux C, Welch JJ, Bierne N (2014) Gene-flow in a mosaic hybrid zone: is local
680 introgression adaptive? *Genetics* **197**, 939-951.

681 Gompert Z, Buerkle CA (2010) INTROGRESS: a software package for mapping components
682 of isolation in hybrids. *Molecular Ecology Resources* **10**, 378-384.

683 Green RE, Krause J, Briggs AW, *et al.* (2010) A draft sequence of the Neandertal genome.
684 *Science* **328**, 710-722.

685 Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint
686 demographic history of multiple populations from multidimensional SNP frequency
687 data. *Plos Genetics* **5**(10), e1000695.

688 Harris K, Nielsen R (2016) The genetic cost of Neanderthal introgression. *Genetics* **203**, 881-
689 891.

690 Harrison RG, Larson EL (2014) Hybridization, introgression, and the nature of species
691 boundaries. *Journal of Heredity* **105**, 795-809.

692 Harrison RG, Larson EL (2016) Heterogeneous genome divergence, differential introgression
693 and the origin and structure of hybrid zones. *Molecular Ecology* **25**, 2454-2466.

694 Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new
695 mutation and standing variation as sources of adaptive variation. *Molecular Ecology* **22**,
696 4606-4618.

697 Hewitt GM (1988) Hybrid zones : Natural laboratories for evolutionary studies. *Trends in*
698 *Ecology & Evolution* **3**, 158-167.

699 Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary.
700 *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*
701 **359**, 183-195.

702 Hewitt GM (2011) Quaternary phylogeography: the roots of hybrid zones. *Genetica* **139**, 617-
703 638.

704 Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates
705 and divergence time, with applications to the divergence of *Drosophila pseudoobscura*
706 and *D. persimilis*. *Genetics* **167**, 747-760.

707 Iannelli F, Pesole G, Sordino P, Gissi C (2007) Mitogenomics reveals two cryptic species in
708 *Ciona intestinalis*. *Trends in Genetics* **23**, 419-422.

709 Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program
710 for dealing with label switching and multimodality in analysis of population structure.
711 *Bioinformatics* **23**, 1801-1806.

712 Jennings RM, Etter RJ, Ficarra L (2013) Population differentiation and species formation in the
713 deep sea: The potential role of environmental gradients and depth. *Plos One* **8**(10),
714 e77594.

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

717 Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new
718 method for the analysis of genetically structured populations. *Bmc Genetics* **11**: 94.

719 Larson EL, Andrés JA, Bogdanowicz SM, Harrison RG (2013) Differential introgression in a
720 mosaic hybrid zone reveals candidate barrier genes. *Evolution* **67**, 3653-3661.

721 Larson EL, White TA, Ross CL, Harrison RG (2014) Gene flow and the maintenance of species
722 boundaries. *Molecular Ecology* **23**, 1668-1678.

723 Maddison WP (1997) Gene trees in species trees. *Systematic Biology* **46**, 523-536.

724 Maggs CA, Castilho R, Foltz D, *et al.* (2008) Evaluating signatures of glacial refugia for North
725 Atlantic benthic marine taxa. *Ecology* **89**, 108-122.

726 Maheshwari S, Barbash DA (2011) The genetics of hybrid incompatibilities. In: *Annual Review*
727 *Genetics, Vol 45* (eds. Bassler BL, Lichten M, Schupbach G), pp. 331-355.

728 Martin SH, Dasmahapatra KK, Nadeau NJ, *et al.* (2013) Genome-wide evidence for speciation
729 with gene flow in *Heliconius* butterflies. *Genome Research* **23**, 1817-1828.

730 Mendez FL, Watkins JC, Hammer MF (2012) A haplotype at STAT2 introgressed from
731 Neanderthals and serves as a candidate of positive selection in Papua New Guinea.
732 *American Journal of Human Genetics* **91**, 265-274.

733 Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of
734 invasive species to marine biodiversity. *Frontiers in Ecology and the Environment* **6**,
735 485-492.

736 Nydam ML, Harrison RG (2007) Genealogical relationships within and among shallow-water
737 *Ciona* species (Asciacea). *Marine Biology* **151**, 1839-1847.

738 Nydam ML, Harrison RG (2010) Polymorphism and divergence within the ascidian genus
739 *Ciona*. *Molecular Phylogenetics and Evolution* **56**, 718-726.

740 Nydam ML, Harrison RG (2011) Introgression despite substantial divergence in a broadcast
741 spawning marine invertebrate. *Evolution* **65**, 429-442.

742 Orr MR, Smith TB (1998) Ecology and speciation. *Trends in Ecology & Evolution* **13**, 502-
743 506.

744 Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology*
745 *and Evolution* **5**, 568-583.

746 Pardo-Diaz C, Salazar C, Baxter SW, *et al.* (2012) Adaptive introgression across species
747 boundaries in *Heliconius* butterflies. *Plos Genetics* **8**(6), e1002752.

748 Pickrell JK, Pritchard JK (2012) Inference of population splits and mixtures from genome-wide
749 allele frequency data. *Plos Genetics* **8**(11), e1002967.

750 Pivotto ID, Nerini D, Masmoudi M, *et al.* (2015) Highly contrasted responses of Mediterranean
751 octocorals to climate change along a depth gradient. *Royal Society open science* **2**,
752 140493.

753 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus
754 genotype data. *Genetics* **155**, 945-959.

755 Procaccini G, Affinito O, Toscano F, Sordino P (2011) A new animal model for merging
756 ecology and evolution. In: *Evolutionary biology: Concepts, biodiversity,*
757 *macroevolution and genome evolution* (ed. Pontarotti P), pp. 91-106.

758 R Development Core Team (2014) R v.3.1.3.: A language and environment for statistical
759 computing. *R Fondation for Statistical Computing*, Vienne, Austria. ISBN: 3-900051-
760 07-0. <http://www.R-project.org>.

761 Racimo F, Sankararaman S, Nielsen R, Huerta-Sanchez E (2015) Evidence for archaic adaptive
762 introgression in humans. *Nature Reviews Genetics* **16**, 359-371.

763 Reich D, Thangaraj K, Patterson N, Price AL, Singh L (2009) Reconstructing Indian population
764 history. *Nature* **461**, 489-494.

- 765 Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review*
766 *of Ecology and Systematics* **27**, 83-109.
- 767 Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure.
768 *Molecular Ecology Notes* **4**, 137-138.
- 769 Roux C, Fraisse C, Castric V, *et al.* (2014) Can we continue to neglect genomic variation in
770 introgression rates when inferring the history of speciation? A case study in a *Mytilus*
771 hybrid zone. *Journal of Evolutionary Biology* **27**, 1662-1675.
- 772 Roux C, Tsagkogeorga G, Bierne N, Galtier N (2013) Crossing the species barrier: Genomic
773 hotspots of introgression between two highly divergent *Ciona intestinalis* species.
774 *Molecular Biology and Evolution* **30**, 1574-1587.
- 775 Sato A, Satoh N, Bishop JDD (2012) Field identification of 'types' A and B of the ascidian
776 *Ciona intestinalis* in a region of sympatry. *Marine Biology* **159**, 1611-1619.
- 777 Sato A, Shimeld SM, Bishop JDD (2014) Symmetrical reproductive compatibility of the two
778 species in the *Ciona intestinalis* (Acidiacea) species complex, a model for marine
779 genomics and developmental biology. *Zoological Science* **31**, 369-374.
- 780 Saarman NP & Pogson GH (2015) Introgression between invasive and native blue mussels
781 (genus *Mytilus*) in the central California hybrid zone. *Molecular Ecology* **24**, 4723-
782 4738.
- 783 Satoh N, Rokhsar D, Nishikawa T (2014) Chordate evolution and the three-phylum system.
784 *Proceedings of the Royal Society B-Biological Sciences* **281**, 20141729.
- 785 Schierenbeck KA, Ellstrand NC (2009) Hybridization and the evolution of invasiveness in
786 plants and other organisms. *Biological Invasions* **11**, 1093-1105.
- 787 Seehausen O, Butlin RK, Keller I, *et al.* (2014) Genomics and the origin of species. *Nature*
788 *Reviews Genetics* **15**, 176-192.
- 789 Storey J (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical*
790 *Society, Series B* **64**, 479-498.
- 791 Suzuki MM, Nishikawa T, Bird A (2005) Genomic approaches reveal unexpected genetic
792 divergence within *Ciona intestinalis*. *Journal of Molecular Evolution* **61**, 627-635.
- 793 Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and
794 phylogeographic breaks in North America. *American Naturalist* **166**, 581-591.
- 795 Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends in Ecology & Evolution*
796 **16**, 330-343.
- 797 Van Name WG (1945) The North and South American ascidians. *Bulletin of the American*
798 *Museum of Natural History* **84**, 1-476 pls. 471-431.
- 799 Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323-354.
- 800 Zhan A, Briski E, Bock DG, Ghabooli S, MacIsaac HJ (2015) Ascidiaceae as models for studying
801 invasion success. *Marine Biology* **162**, 2449-2470.
- 802 Zhan A, MacIsaac HJ, Cristescu ME (2010) Invasion genetics of the *Ciona intestinalis* species
803 complex: from regional endemism to global homogeneity. *Molecular Ecology* **19**, 4678-
804 4694.

805 Table 1. **Study localities, hybrid index (*h*) and number of hybrids *sensu lato*** (i.e. F1, F2 hybrids and backcrosses with parental species according to
 806 NEWHYBRID software) **in each population of *Ciona robusta* and *C. intestinalis*.**

N°	Locality	Region	Coordinates (Long., Lat.)	Regional status	Locality status	Sampling year	Nind	Hybrid index (mean ± SD)	Number of hybrids
<i>C. robusta</i>									
1-	Guanaqueros, Ch	South Eastern Pacific	-30.1945, -71.4300	Allopatric	Non-syntopic	2012	24	0.0056 ± 0.0039	0
2-	Etang de Thau, Fr	Mediterranean Sea	43.4014, 3.6582	Allopatric	Non-syntopic	2013	23	0.0010 ± 0.0020	0
5-	Falmouth, UK	English Channel	50.1543, -5.0579	Sympatric	Syntopic	2013	24	0.0024 ± 0.0031	0
6-	Plymouth, UK	English Channel	50.3583, -4.1228	Sympatric	Syntopic	2011	24	0.0024 ± 0.0028	0
7-	St Vaast, Fr	English Channel	49.5897, -1.2648	Sympatric	Syntopic	2012	23	0.0031 ± 0.0034	0
8-	Perros Guirec, Fr	English Channel	48.8112, -3.4295	Sympatric	Syntopic	2011	24	0.0038 ± 0.0037	0
10-	Moulin blanc, Fr	Bay of Brest	48.3906, -4.4318	Sympatric	Syntopic	2012	24	0.0034 ± 0.0041	0
11-	Camaret, Fr	Bay of Brest	48.2799, -4.5961	Sympatric	Syntopic	2011, 2012	24	0.0020 ± 0.0031	0
12-	Quiberon, Fr	Bay of Biscay	47.4858, -3.0999	Sympatric	Syntopic	2012, 2013	23	0.0025 ± 0.0024	0
Total							213	0.0029 ± 0.0034	
<i>C. intestinalis</i>									
3-	Nahant, USA	North Western Atlantic	42.4569, -70.9414	Allopatric	Non-syntopic	2013	24	0.0064 ± 0.0056	0
4a-	Fiskebackskil – surface, Sw	North Sea	58.2502, 11.4579	Allopatric	Non-syntopic	2010	12	0.0016 ± 0.0024	0
4b-	Fiskebackskil - 20m depth, Sw	North Sea	58.2502, 11.4579	Allopatric	Non-syntopic	2010	12	0.0012 ± 0.0022	0
5-	Falmouth, UK	English Channel	50.1543, -5.0579	Sympatric	Syntopic	2011	24	0.0042 ± 0.0043	0
6-	Plymouth, UK	English Channel	50.3583, -4.1228	Sympatric	Syntopic	2011	24	0.0060 ± 0.0058	0
7-	St Vaast, Fr	English Channel	49.5897, -1.2648	Sympatric	Syntopic	2012	23	0.0081 ± 0.0075	0
8-	Perros Guirec, Fr	English Channel	48.8112, -3.4295	Sympatric	Syntopic	2011, 2012	24	0.0070 ± 0.0069	0
9-	Aber Wrac'h, Fr	English Channel	48.5987, -4.5622	Sympatric	Non-syntopic	2011, 2012	23	0.0064 ± 0.0047	0
10-	Moulin blanc, Fr	Bay of Brest	48.3906, -4.4318	Sympatric	Syntopic	2011	24	0.0071 ± 0.0056	0
11-	Camaret, Fr	Bay of Brest	48.2799, -4.5961	Sympatric	Syntopic	2011, 2012	22	0.0286 ± 0.1065	1 (F1-hybrid)
	<i>without F1 -hybrid</i>						21	0.0059 ± 0.0052	0
12-	Quiberon, Fr	Bay of Biscay	47.4858, -3.0999	Sympatric	Syntopic	2011, 2012	24	0.0062 ± 0.0065	0
Total							236	0.0055 ± 0.0241	1 (F1-hybrid)
Total (without F1-hybrid)							235	0.0044 ± 0.0050	0

807 *Regional status* and *locality status* indicate if the two species have been reported to co-exist at a regional scale (allopatric vs. sympatric) or at the locality level (syntopic
 808 vs. non-syntopic). In this table, *h* is defined as the proportion of alleles from one species in the genetic background of the other species (i.e. proportion of *C. intestinalis*
 809 alleles over all loci in *C. robusta* individuals and proportion of *C. robusta* alleles over all loci in *C. intestinalis* individuals). *h* values were averaged across individuals
 810 for each sampled localities. Analyses were done with 105 SNPs selected for inter-specific gene flow analyses ($F_{ST} > 0.9$; see Material and Methods).

811 Table 2. **Genetic diversity indices and fixation index** of each study populations for *Ciona*
 812 *robusta* and *C. intestinalis*.

N°	Locality	Introduced vs. native status	P _{loc}	H _e	F _{IS}
<i>C. robusta</i>					
1-	Guanaqueros	Introduced	103	0.288	-0.013
2-	Etang de Thau	Introduced	77	0.234	-0.012
5-	Falmouth	Introduced	79	0.247	-0.055
6-	Plymouth	Introduced	80	0.240	-0.043
7-	St Vaast	Introduced	79	0.238	-0.076
8-	Perros Guirec	Introduced	81	0.246	-0.024
10-	Moulin blanc	Introduced	81	0.236	-0.043
11-	Camaret	Introduced	81	0.247	-0.022
12-	Quiberon	Introduced	83	0.254	-0.029
	Total (Sympatric pop.)		86	0.253	
	Total		111	0.265	
<i>C. intestinalis</i>					
3-	Nahant	Cryptogenic	117	0.233	-0.018
4a-	Fiskebackskil - surface	Native	97	0.194	0.048
4b-	Fiskebackskil - 20m depth		85	0.172	0.001
5-	Falmouth	Native	118	0.229	-0.012
6-	Plymouth	Native	125	0.229	0.002
7-	St Vaast	Native	123	0.233	0.032
8-	Perros Guirec	Native	125	0.234	-0.035
9-	Aber Wrac'h	Native	122	0.239	-0.005
10-	Moulin blanc	Native	125	0.239	0.014
11-	Camaret	Native	120	0.234	-0.011
12-	Quiberon	Native	120	0.240	-0.014
	Total (Sympatric pop.)		148	0.242	
	Total		150	0.244	

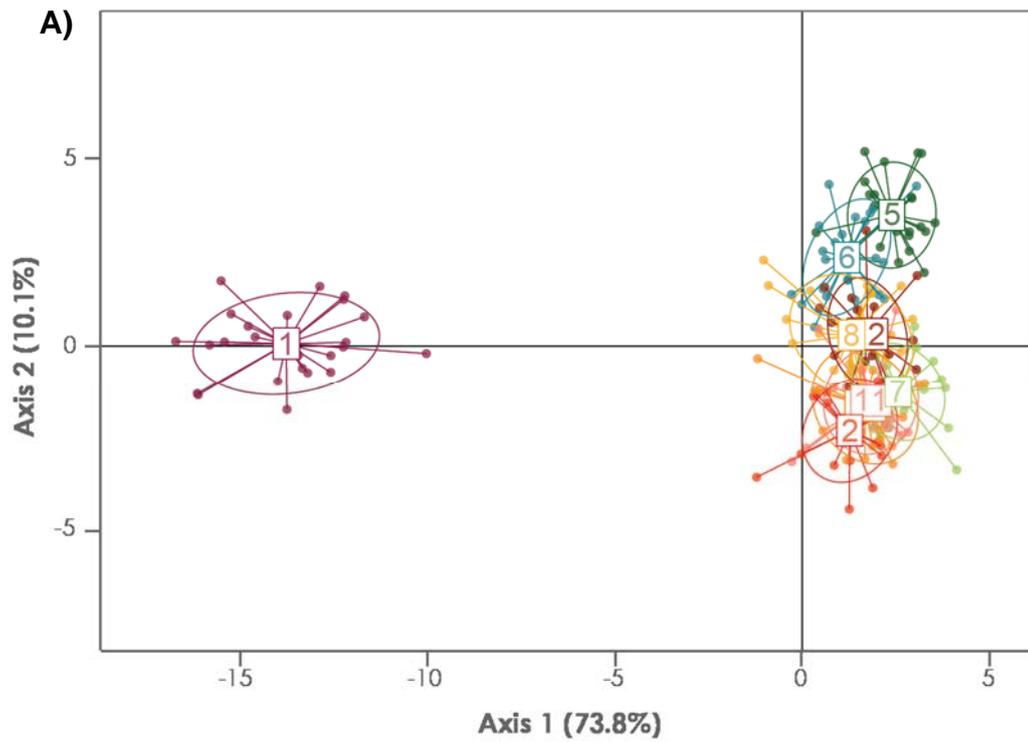
813

814 P_{loc}: number of polymorphic loci and H_e: expected heterozygosity over 111 and 150 polymorphic loci
 815 retained for intra-specific analyses in *C. robusta* and *C. intestinalis*, respectively (see *Materials and*
 816 *Methods*); F_{IS}: fixation index calculated (no deviation from Hardy-Weinberg equilibrium; exact test, *P*
 817 < 0.05).

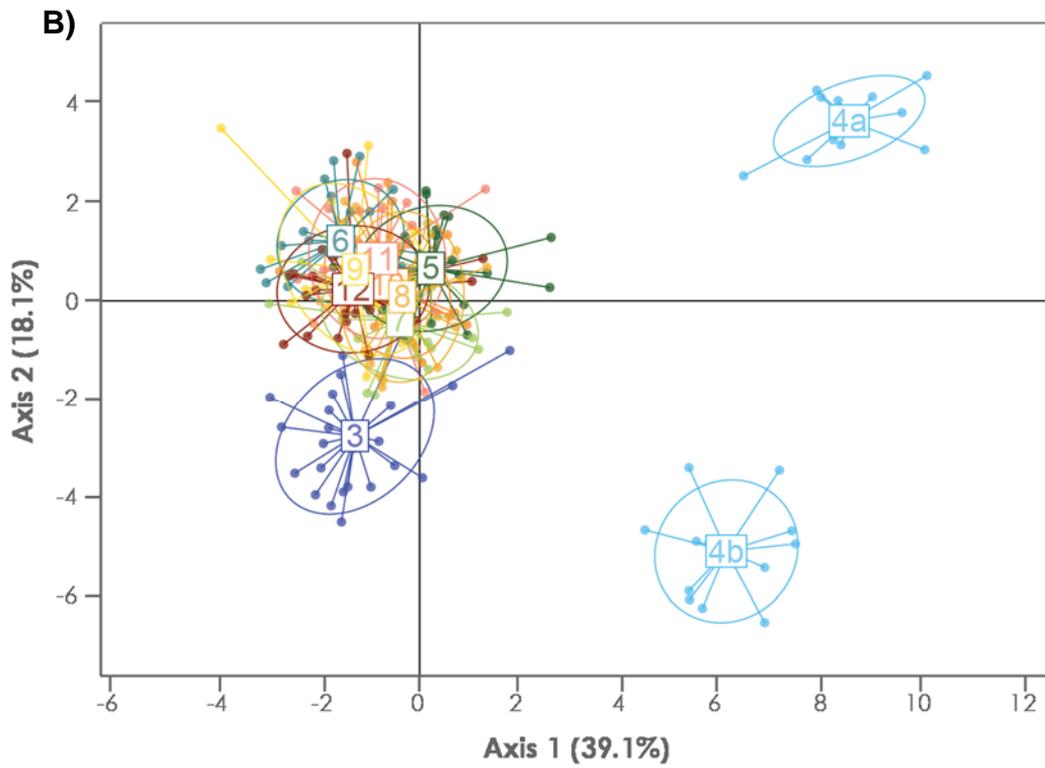
818 Table 3. Genetic structure among populations for *Ciona robusta* and *C. intestinalis*.

	F_{ST}	P -value
<i>C. robusta</i>		
All sampled populations (9 populations)	0.054	$P < 0.001$
All populations without Guanaqueros (all except no.1)	0.023	$P < 0.001$
Sympatric populations (all except no.1 and 2)	0.021	$P < 0.001$
<i>C. intestinalis</i>		
All sampled populations (11 populations)	0.045	$P < 0.001$
All populations without Fiskebackskil (all except no.4a and 4b)	0.021	$P < 0.001$
Sympatric populations (all except no.3, 4a and 4b)	0.014	$P = 0.020$

819



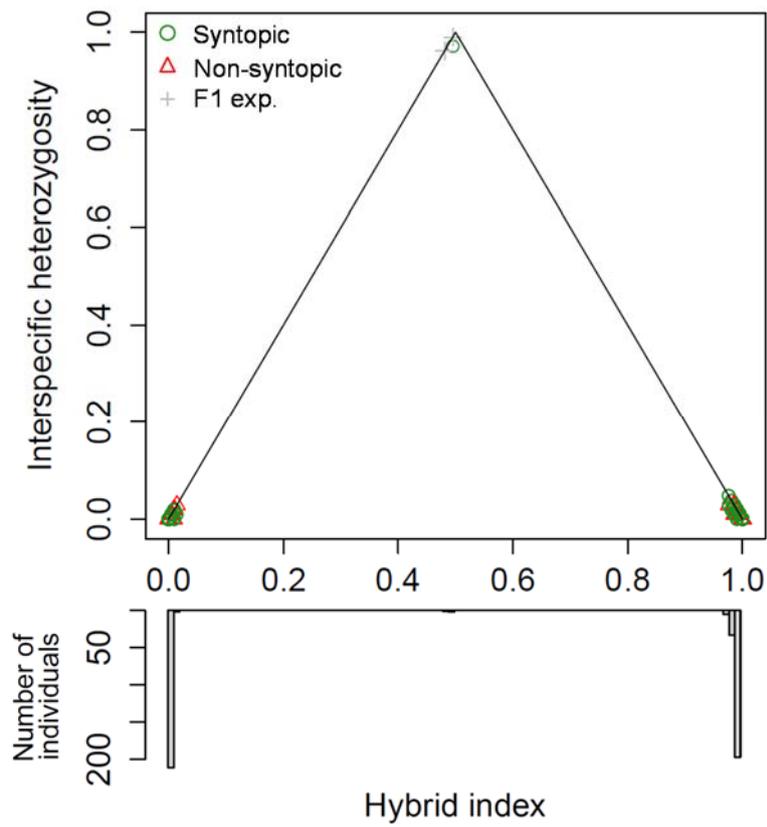
820



821

822 Figure 1. **Discriminant Analysis of Principal Components (DAPC) among populations of**
 823 ***C. robusta* (A) and *C. intestinalis* (B).** Only the two first axis showing the two higher
 824 discriminant eigenvalues are presented here.

825



826

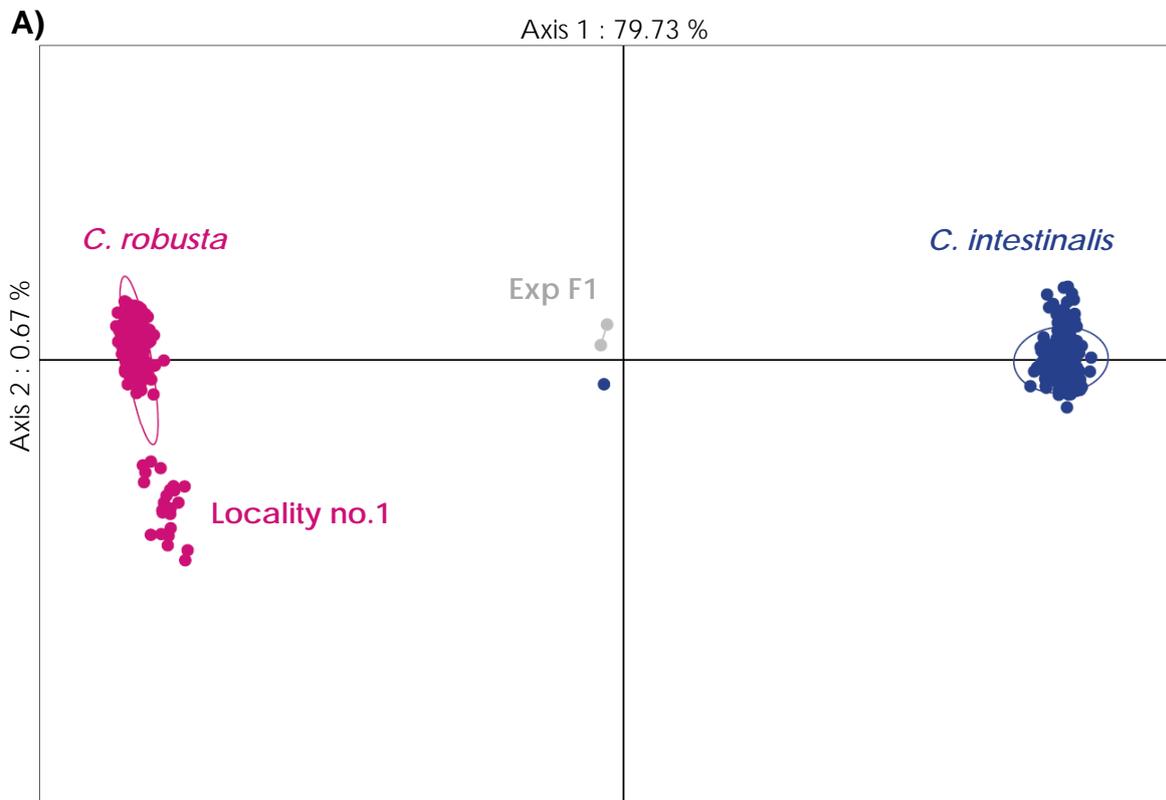
827 **Figure 2. Triangle plot showing the relationship between heterozygosity rate across loci**

828 **and hybrid index for each individual.** At the top of the triangle, one green circle is picturing

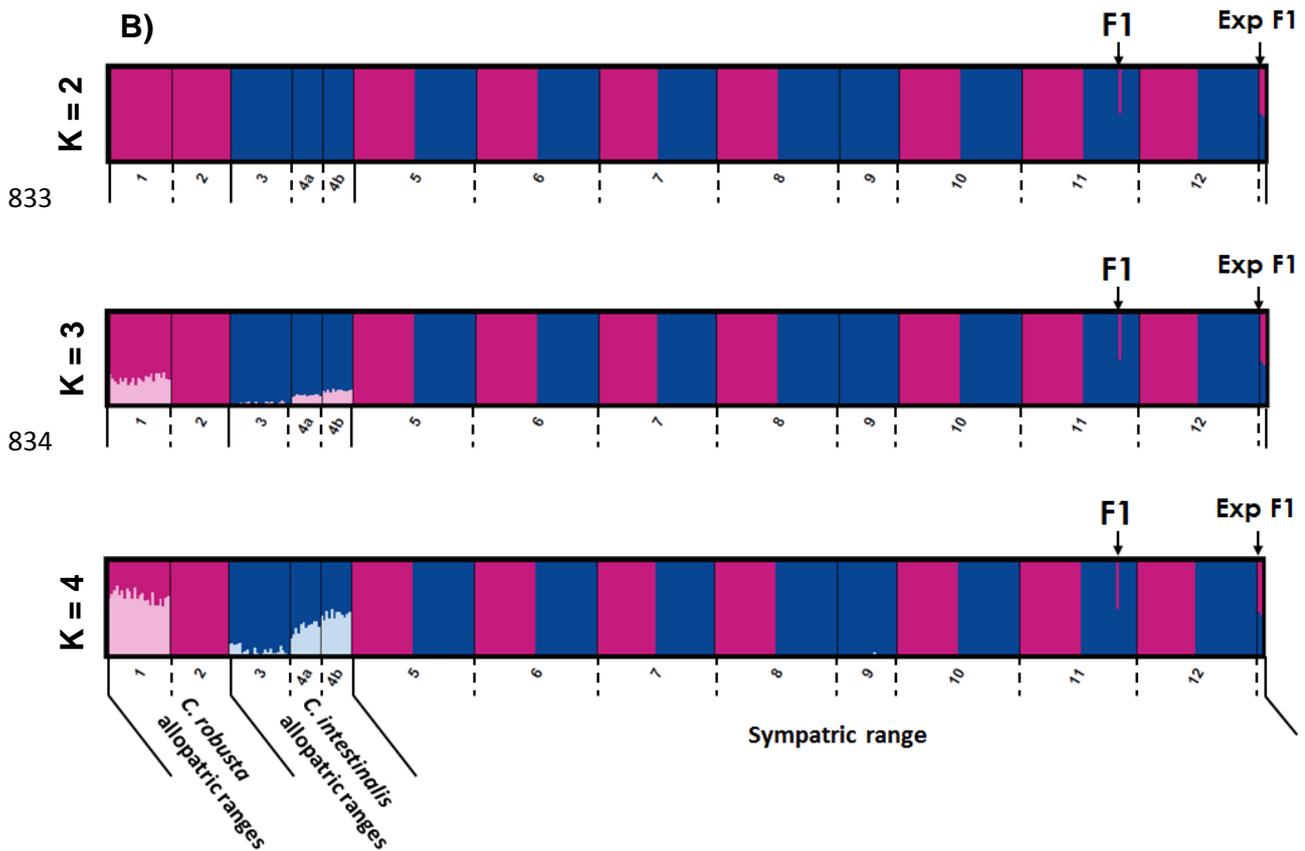
829 one individual from the locality no. 11 and the two gray crosses are F1-hybrids from

830 experimental crosses.

831

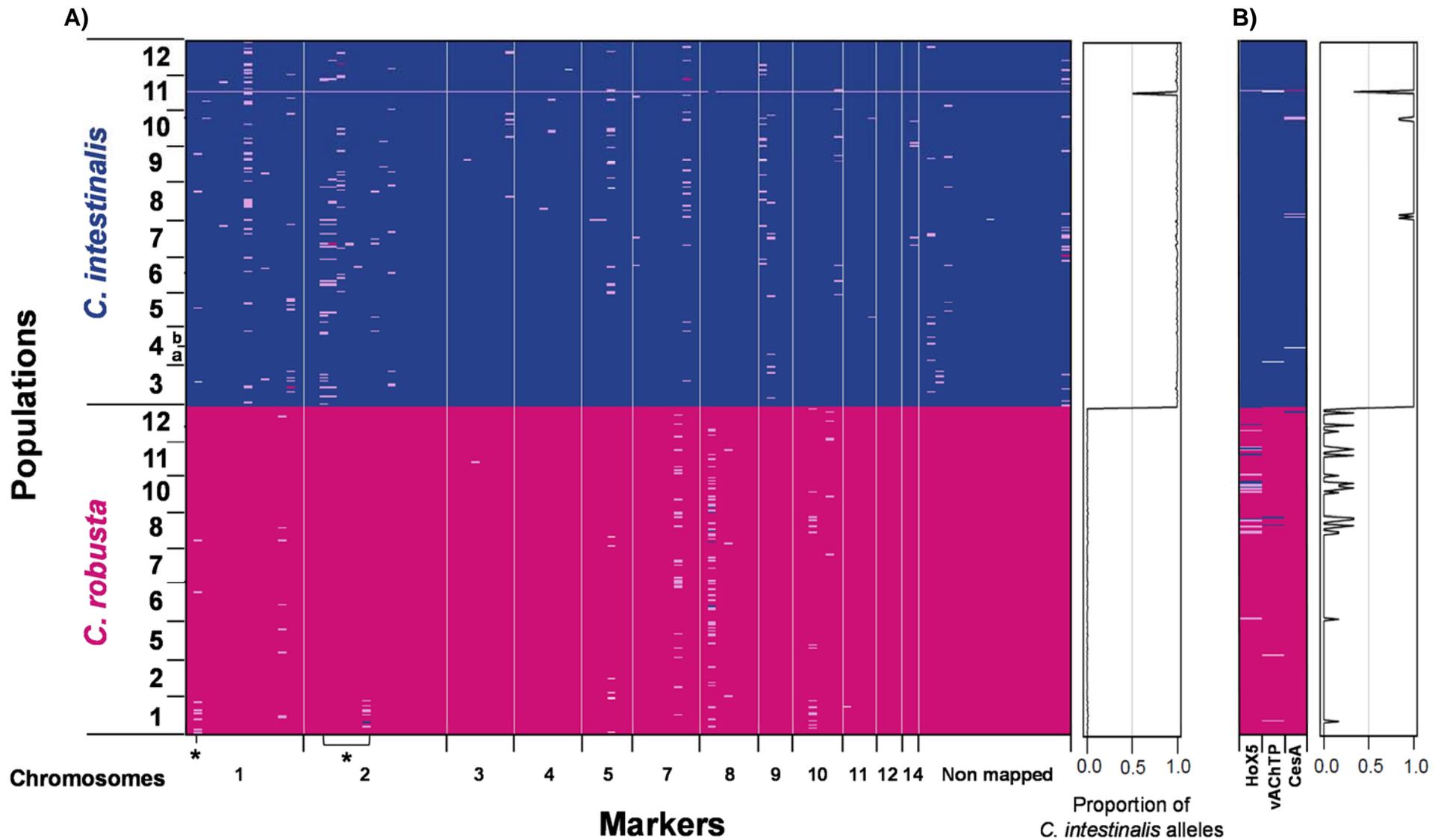


832



835

836 Figure 3. **A) Principal Components Analysis (PCA)** and **B) Individual Bayesian**
 837 **assignment proportion for clusters from K =2 to K = 4** using the total dataset (i.e. 310 SNPs,
 838 449 individuals from natural population and the two F1 hybrids from experimental crosses).



839

840

841

842

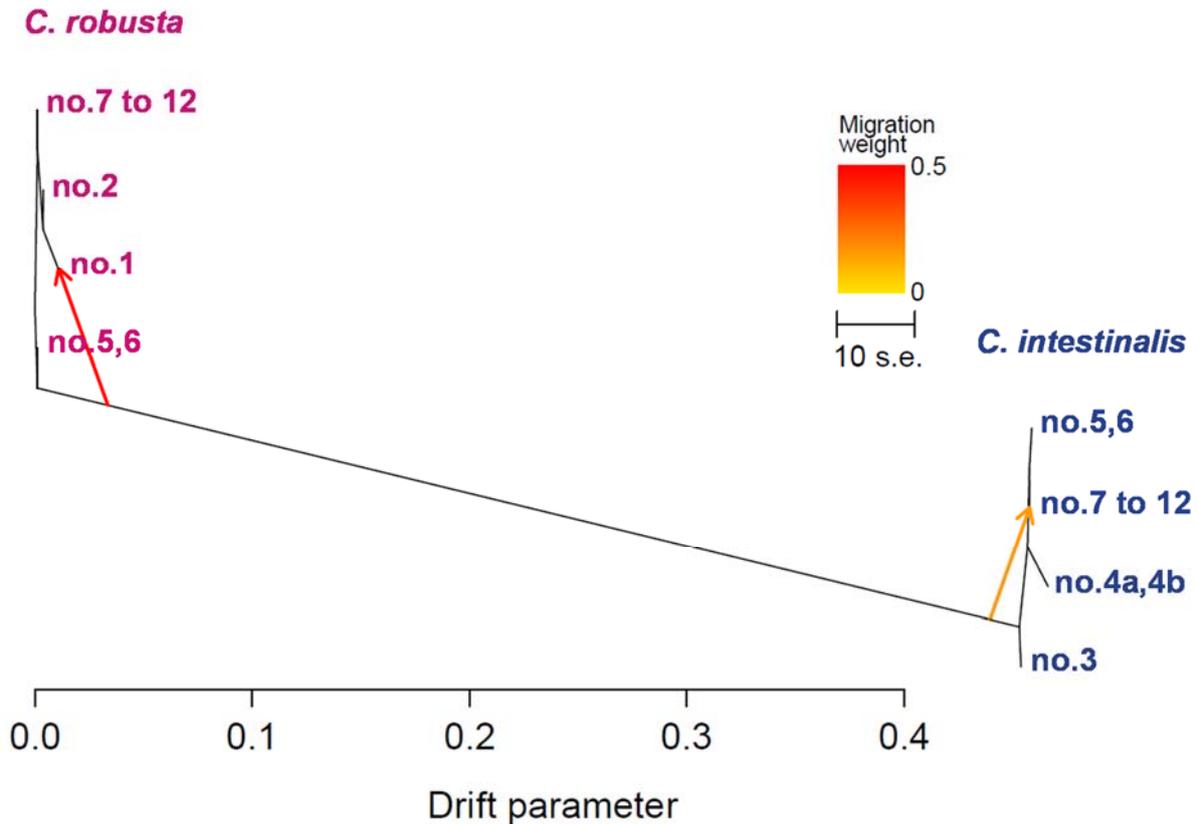
843

844

845

846

Figure 4. **A) Genomic architecture using 105 highly differentiated loci ($F_{st} > 0.9$) selected for inter-specific analyses.** Markers (x-axis) are ordered following physical position on chromosomes. Individuals (y-axis) are ordered per population. Dark pink cases indicate homozygote genotype on *C. robusta* alleles; dark blue, homozygote genotype on *C. intestinalis* alleles; light purple, heterozygotes for *C. robusta* and *C. intestinalis* alleles; and white cases, missing values. Asterisks indicate loci located in introgression hotspots defined by Roux *et al.* (2013). **B) Pattern of admixture for 3 nuclear loci (Hox5, vAChTP, CesA) and one mitochondrial locus (mtCOI) analyzed by PCR and PCR-RFLP,** already used in previous studies (Nydam & Harrison 2011; Sato *et al.* 2014; Bouchemousse *et al.* 2016).



847

848 Figure 5. **Population tree inferred by TREEMIX indicating two migration events between**
 849 ***Ciona robusta* and/or *C. intestinalis* populations using the total dataset** (i.e. 310 SNPs).
 850 Terminal nodes are labelled by locality number (Table 1). Note that we pooled populations
 851 according to regions of sampling (i.e. no.4a and 4b for *C. intestinalis*; no.5 and 6 and no. 7 to
 852 12 for each species) to avoid noises by intra-specific admixture events. Admixture arrows are
 853 colored according to the migration weight. The two admixture events significantly improved
 854 the model as compared to a situation without migration ($P < 0.001$).
 855

Supplementary Material

Distinguishing contemporary hybridization from past introgression with post-genomic ancestry-informative SNPs in strongly differentiated *Ciona* species

Sarah Bouchemousse^{1,2}, Cathy Haag-Liautard^{3,4}, Nicolas Bierne^{3,4} and Frédérique Viard^{1,2,*}

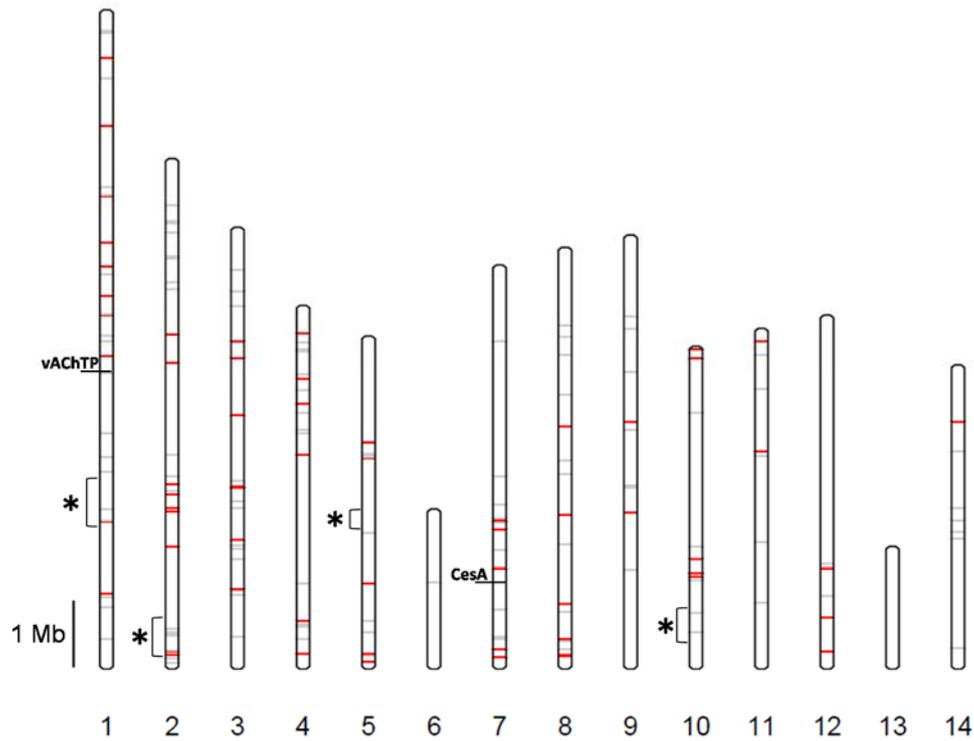


Figure S1. **Physical mapping of the 310 SNPs genotyped** (note that 62 of the 310 SNPs are not represented because located on non-mapped scaffolds). Red bars point the 105 loci selected for interspecific admixture analysis. Asterisks indicate regions of introgression hotspots as defined by Roux *et al.* (2013). Two genes (vaChTP and CesA) used in previous studies looking for hybrids (e.g. Bouchemousse *et al.* (2016); Nydam & Harrison (2011); Sato *et al.* (2014)) are also indicated (the remaining Hox5 locus, used in Caputi *et al.* (2007) and Bouchemousse *et al.* (2016), is located on a non-mapped scaffold).

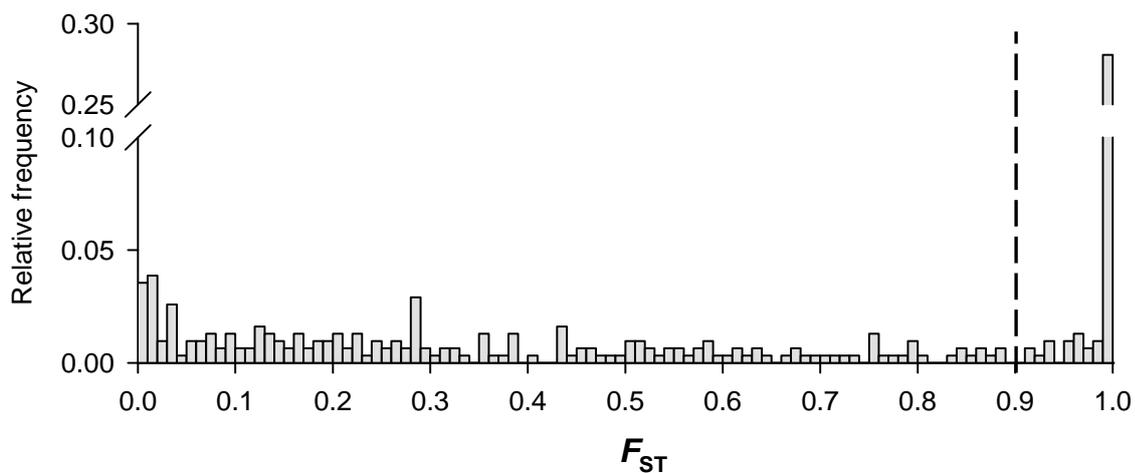


Figure S2. **Distribution of interspecific differentiation given by F_{ST} values** using populations in the allopatric ranges of the two species (no. 1 and 2 for *C. robusta* and no. 3, 4a and 4b for *C. intestinalis*) as a reference for absence of contemporary gene flow. Dashed line separates the subset of 105 loci with F_{ST} values strictly superior to 0.9.

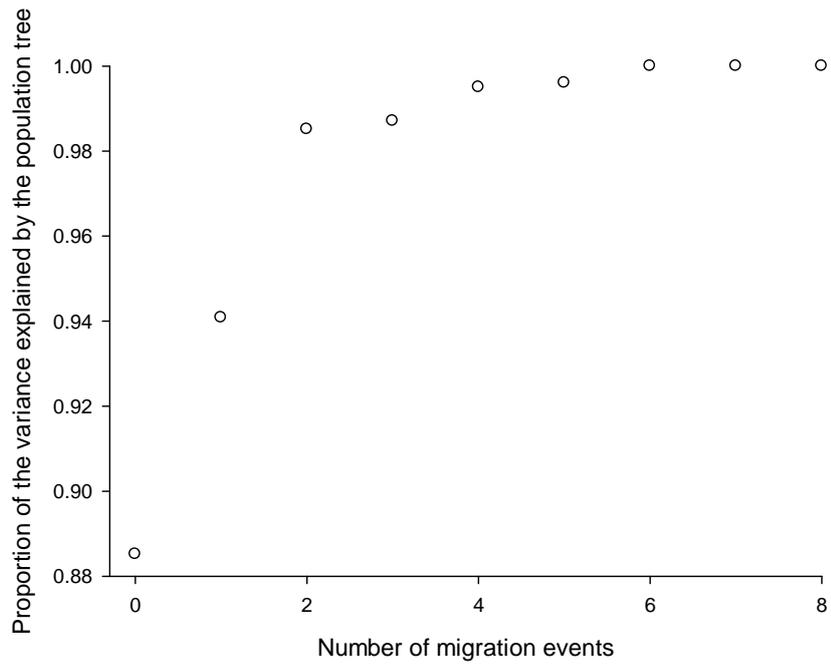
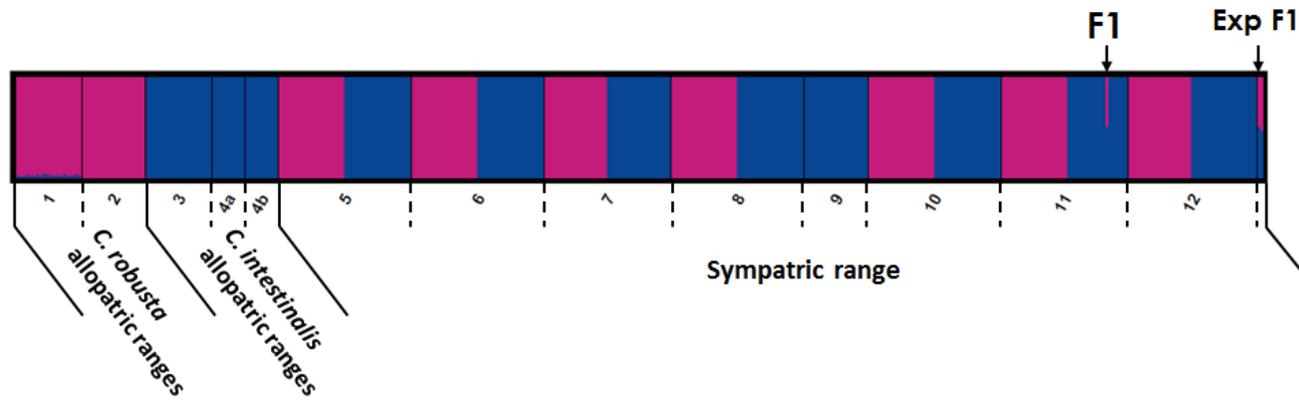
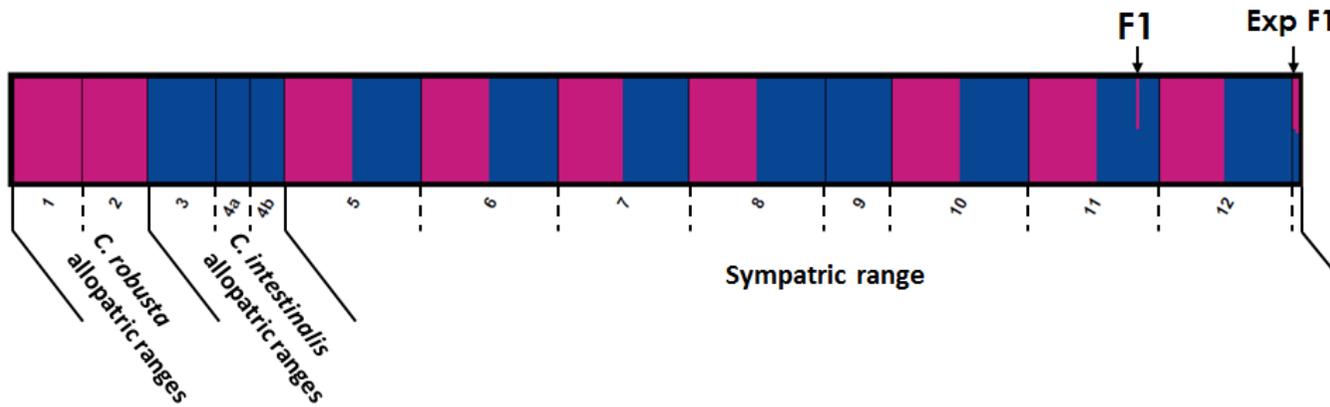


Figure S3. **Summary of TREEMIX analysis using the total dataset** (i.e. 310 SNPs). The scatterplot present the proportion of the variance explained by the population tree with different numbers of migration events (from 0 to 8).

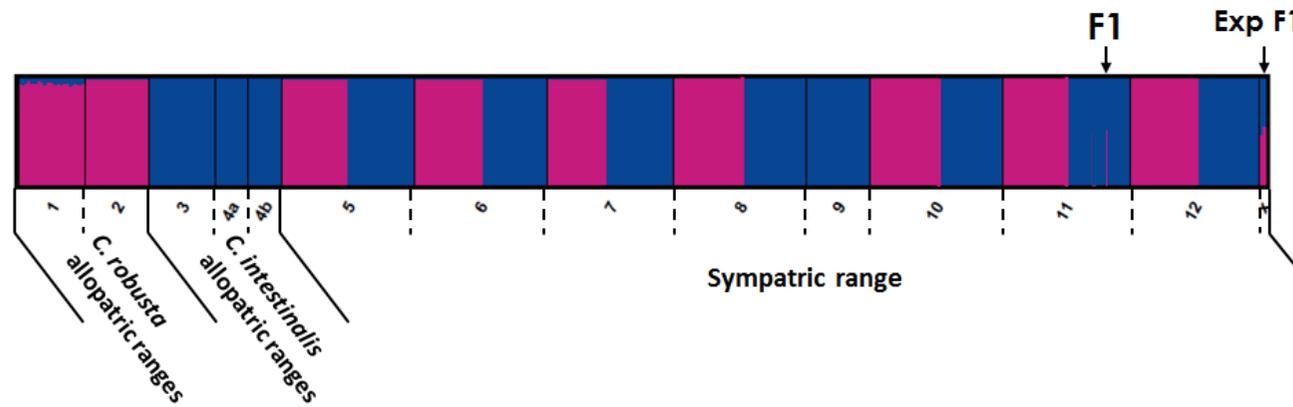
1) 58 SNPs randomly selected



2) 105 highly differentiated SNPs



3) 245 SNPs : all SNPs except those differentially fixed between species



4) 42 SNPs polymorphic in the two species

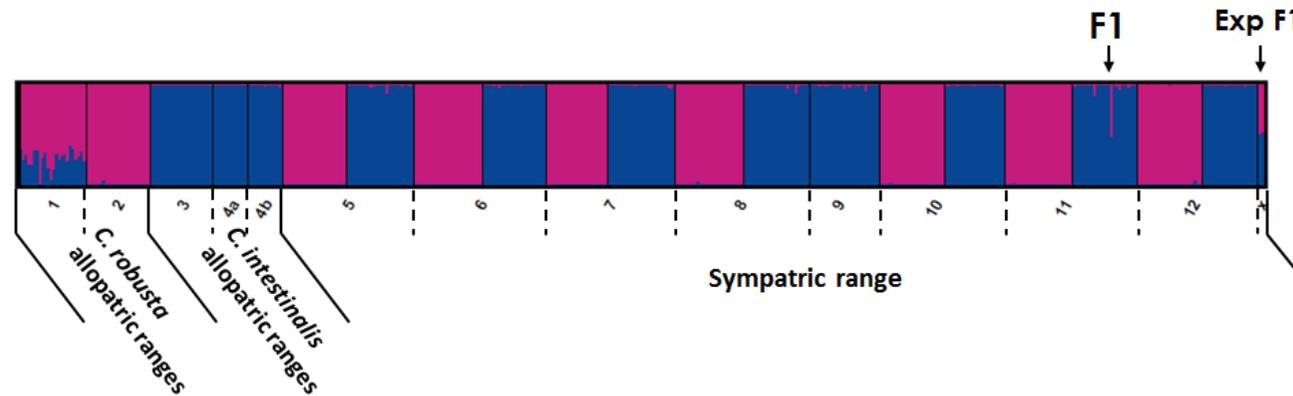


Figure S4. **Results of the STRUCTURE analysis on additional dataset.** STRUCTURE analyses were carried out for $K = 2$ using a subset of 1) 58 SNPs randomly selected, 2) 105 SNPs used in the INTROGRESS analysis (i.e. highly differentiated loci with $F_{ST} > 0.9$, see Fig.S2), 3) 245 SNPs: all except the SNPs that were differentially fixed between the two species, and 4) 42 SNPs polymorphic in the two species. Using different subset of markers do not influence the admixture patterns (see also Fig. 3b in the main text for the same analysis carried out on the full dataset). The F1-hybrid found in the Camaret (no.11) population and the F1- Hybrids experimentally produced (Exp F1) are well recovered by the analysis, including with the 42 SNP panels which is composed only of loci polymorphic in the two species, a property that is prone to false identification.

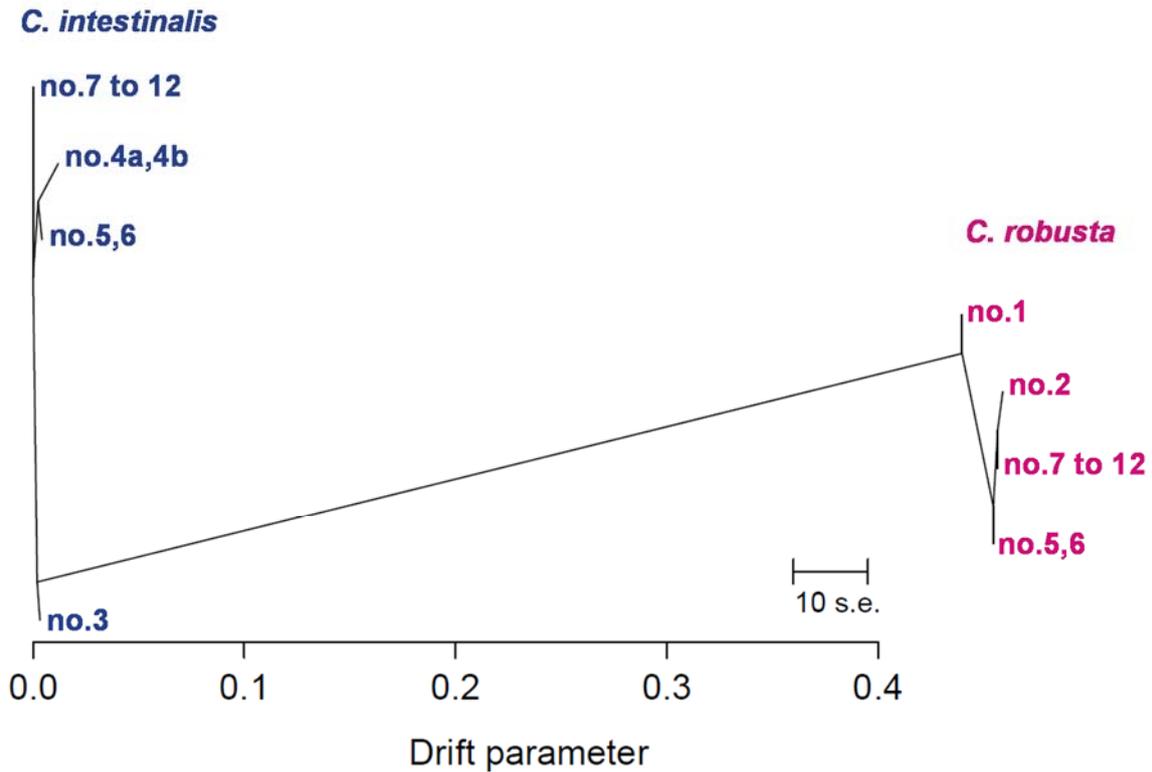


Figure S5. Population tree inferred by TREEMIX without migration events between *Ciona robusta* and *C. intestinalis* populations using the total dataset (i.e. 310 SNPs). Terminal nodes are labelled by locality number (Table 1 in the main text). Note that the single F1-hybrid sampled in natural populations (no.11) was removed of the analysis.

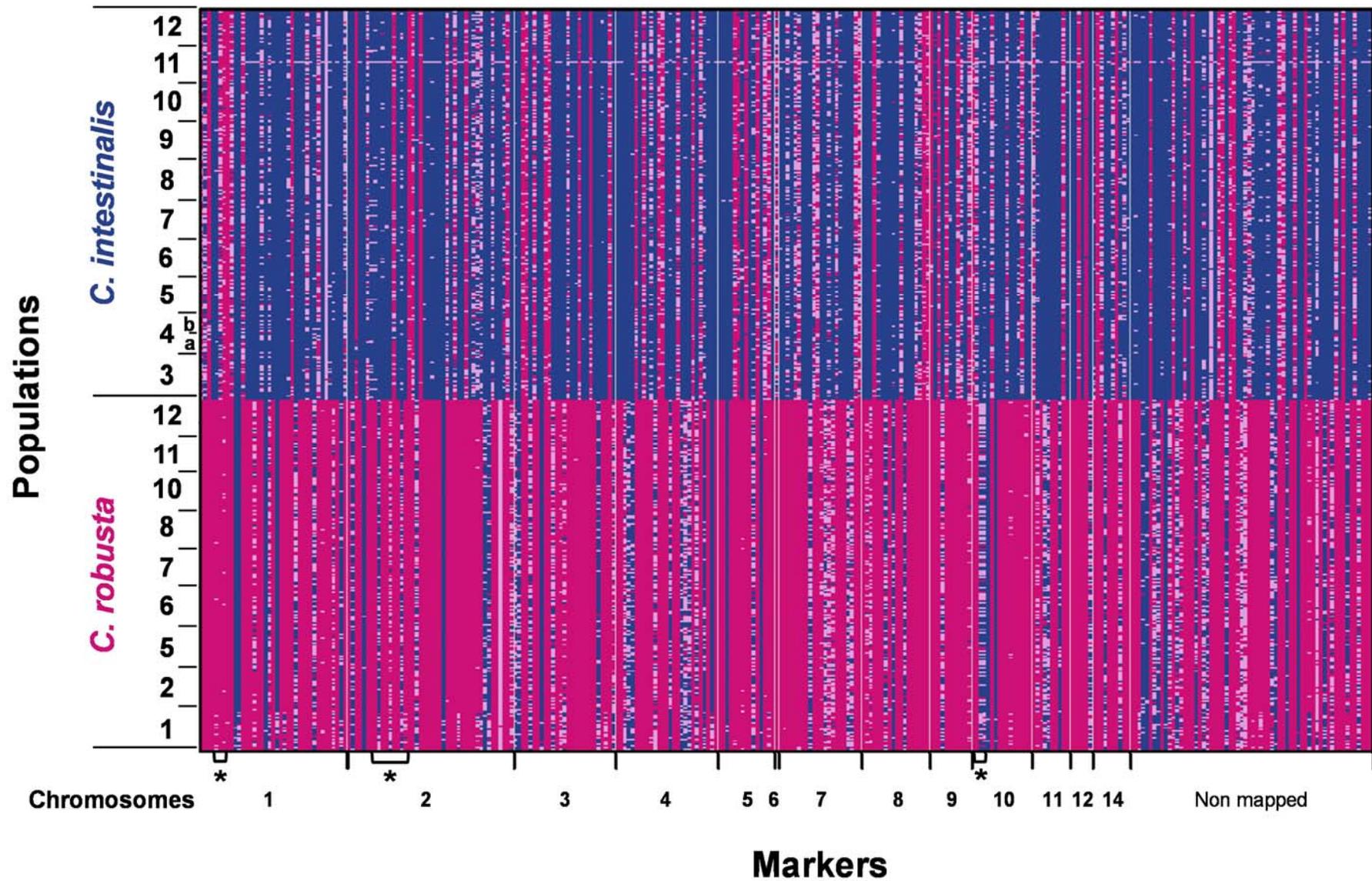


Figure S6. **Genomic architecture using the 310 loci.** Markers (x-axis) are ordered following physical position on chromosomes. Individuals (y-axis) are ordered per population. Dark pink cases indicate homozygote genotype on *C. robusta* alleles; dark blue, homozygote genotype on *C. intestinalis* alleles; light purple, heterozygotes for *C. robusta* and *C. intestinalis* alleles; and white cases, missing values. Asterisks indicate loci located in introgression hotspots defined by Roux *et al.* (2013).

Table S1. **Number of SNPs per categories defined according to their polymorphism patterns:** 1) sf: loci differentially fixed between the two species; 2) sxA: loci polymorphic in *Ciona robusta* only; 3) sxB: loci polymorphic in *C. intestinalis* only; 4) ss: polymorphism shared by the two species. The number of SNPs in each category is given according to the sorting made using the dataset from which the loci were designed (Roux *et al.* 2013) and according to the results of the present study.

		Distribution based on the results of the present study					Not genotyped or monomorphic	
		sf	sxA	sxB	ss	Total		Total
Distribution based on data from Roux <i>et al.</i> (2013)	sf	64	3	26	1	94	7	101
	sxA	0	71	1	17	89	20	109
	sxB	0	0	99	3	102	25	127
	ss	1	1	2	21	25	22	47
Total		65	75	128	42	310	74	384

Table S2. Characteristics of the 310 SNPs genotyped and the three markers genotyped by PCR-RFLP, their allele frequencies and F_{ST} values computed among allopatric localities (as references of no contemporary inter-specific gene flow) for *C. robusta* and *C. intestinalis*

cat.: category of SNP polymorphism in the study dataset (see caption of Table S1); Chr: chromosome number (after no.14, the number corresponds to a non-mapped scaffold); Position (exact nucleotide position along the full genome in base pairs); HS introg (indicate SNPs that are localized in an introgression hotspots defined by Roux *et al.* (2013)).

Loc. 1 and Loc. 2: SE Pacific and Mediterranean localities (presence of *C. robusta* only). Loc. 3, Loc. 4a and 4b: NW Atlantic and North Sea localities (presence of *C. intestinalis* only). The 105 SNPs kept ($F_{ST} > 0.9$) for interspecific admixture analyses (see Material and Methods in the main text) are indicated in bold.

Index	SNPs characteristics				Allele frequency in allopatric populations								F_{ST}
	cat.	Chr.	Position	HS introg	<i>C. robusta</i>			<i>C. intestinalis</i>					
					Loc. 1	Loc. 2	All	Loc. 3	Loc. 4a	Loc. 4b	All		
snp266	sxB	1	414791	no	1.000	1.000	1.000	0.688	0.792	0.917	0.771	0.219	
snp215	sxB	1	908672	no	1.000	1.000	1.000	0.271	0.458	0.417	0.354	0.640	
snp56	sxB	1	1063227	no	1.000	1.000	1.000	1.000	1.000	0.542	0.885	0.099	
snp22	sf	1	1109920	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp18	ss	1	2205837	yes	0.875	1.000	0.938	0.000	0.000	0.000	0.000	0.935	
snp367	sxB	1	2231134	yes	1.000	1.000	1.000	0.583	0.167	0.375	0.427	0.564	
snp108	ss	1	2402960	yes	0.979	0.978	0.979	0.958	0.917	0.917	0.938	0.011	
snp350	ss	1	2986331	no	0.979	1.000	0.990	1.000	1.000	1.000	1.000	0.000	
snp109	sxB	1	3195706	no	1.000	1.000	1.000	0.708	0.750	0.958	0.781	0.206	
snp159	sxA	1	3576763	no	0.938	1.000	0.969	1.000	1.000	1.000	1.000	0.022	
snp318	sxA	1	4483756	no	0.667	1.000	0.833	1.000	1.000	1.000	1.000	0.162	
snp260	sxB	1	4484320	no	1.000	1.000	1.000	0.750	0.708	0.292	0.625	0.365	
vAChTP	ss	1	4501921	no	0.979	1.000	0.989	0.000	0.000	0.000	0.000	0.989	
snp343	sxB	1	4762944	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp102	sf	1	4763382	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp329	sxA	1	4987764	no	0.792	0.913	0.852	0.000	0.000	0.000	0.000	0.851	
snp240	sxB	1	5073937	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp172	sxB	1	5074210	no	1.000	1.000	1.000	0.250	0.083	0.375	0.240	0.756	
snp363	sxA	1	5074378	no	0.979	1.000	0.990	1.000	1.000	1.000	1.000	0.000	
snp255	ss	1	5074414	no	0.479	0.783	0.631	0.208	0.083	0.375	0.219	0.284	
snp124	sf	1	5395001	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp115	sxA	1	5415221	no	0.688	1.000	0.844	1.000	1.000	1.000	1.000	0.151	
snp334	sf	1	5694948	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp143	sxA	1	5694998	no	0.542	1.000	0.771	0.000	0.000	0.000	0.000	0.764	
snp47	sxB	1	5695260	no	1.000	1.000	1.000	0.063	0.000	0.000	0.031	0.968	
snp338	sxB	1	6027477	no	1.000	1.000	1.000	0.938	0.958	1.000	0.958	0.031	
snp105	sxA	1	6028067	no	0.750	0.587	0.668	1.000	1.000	1.000	1.000	0.326	
snp25	sf	1	6136134	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp206	sxB	1	6506772	no	1.000	1.000	1.000	0.021	0.000	0.000	0.010	0.989	
snp113	sxB	1	7215261	no	1.000	1.000	1.000	0.417	0.083	0.125	0.260	0.734	
snp103	sf	1	7216496	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp339	ss	1	7242782	no	0.646	0.717	0.682	1.000	0.833	0.875	0.927	0.166	
snp328	sxB	1	7357460	no	1.000	1.000	1.000	0.563	0.917	0.167	0.552	0.439	
snp11	sxA	1	8307018	no	0.938	1.000	0.969	0.000	0.000	0.000	0.000	0.968	
snp230	sxB	1	9032746	no	1.000	1.000	1.000	0.521	0.500	0.500	0.510	0.487	
snp239	sxB	1	9032786	no	1.000	1.000	1.000	0.125	0.000	0.000	0.063	0.936	
snp281	ss	1	9032828	no	0.583	0.717	0.650	0.000	0.000	0.000	0.000	0.648	
snp188	sf	1	9350711	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
snp139	sxA	1	9718320	no	0.729	1.000	0.865	1.000	1.000	1.000	1.000	0.127	
snp68	sxB	1	9752321	no	1.000	1.000	1.000	0.396	0.708	0.458	0.490	0.502	

snp1	sf	2	63538	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp146	sxA	2	63961	no	0.917	0.870	0.893	1.000	1.000	1.000	1.000	0.099
snp70	sxB	2	64337	no	1.000	1.000	1.000	0.896	0.167	1.000	0.740	0.245
snp218	sf	2	118366	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp138	sxA	2	118516	no	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000
snp46	sxB	2	118564	no	1.000	1.000	1.000	0.667	0.167	0.000	0.375	0.617
snp55	sxB	2	173448	yes	1.000	1.000	1.000	0.125	0.083	0.000	0.083	0.915
snp154	sxB	2	179880	yes	1.000	1.000	1.000	0.042	0.000	0.000	0.021	0.979
snp170	sxA	2	227624	yes	0.625	0.957	0.791	1.000	1.000	1.000	1.000	0.205
snp66	sxB	2	228866	yes	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp39	sxB	2	251567	yes	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp186	sxA	2	251945	yes	0.479	0.696	0.587	0.000	0.000	0.000	0.000	0.583
snp49	sxB	2	251993	yes	1.000	1.000	1.000	0.896	0.958	0.292	0.760	0.226
snp35	sxB	2	476678	yes	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp63	ss	2	477077	yes	0.583	0.826	0.705	0.917	1.000	0.875	0.927	0.146
snp290	sxA	2	483036	yes	0.833	1.000	0.917	0.000	0.000	0.000	0.000	0.915
snp344	sxB	2	483351	yes	1.000	1.000	1.000	1.000	0.833	1.000	0.958	0.031
snp176	sxB	2	533683	no	1.000	1.000	1.000	0.563	0.958	1.000	0.771	0.219
snp73	ss	2	534013	no	0.750	0.457	0.603	1.000	1.000	1.000	1.000	0.390
snp198	sxB	2	578179	no	1.000	1.000	1.000	0.979	1.000	1.000	0.990	0.000
snp116	sxB	2	1834123	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp31	sxB	2	2372600	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp376	sxB	2	2417520	no	1.000	1.000	1.000	0.063	0.000	0.000	0.031	0.968
snp32	sf	2	2417784	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp308	sf	2	2627523	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp182	sxA	2	2684517	no	0.917	1.000	0.958	1.000	1.000	1.000	1.000	0.033
snp101	ss	2	2684820	no	1.000	0.978	0.989	0.625	0.208	0.083	0.385	0.589
snp110	sf	2	2786390	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp162	sxB	2	2900734	no	1.000	1.000	1.000	0.500	0.875	0.833	0.677	0.311
snp13	sxA	2	3247109	no	0.604	1.000	0.802	0.000	0.000	0.000	0.000	0.797
snp306	sf	2	4658149	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp340	sxB	2	5101067	no	1.000	1.000	1.000	0.833	0.917	1.000	0.896	0.095
snp383	sf	2	5101731	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp145	sxB	2	5795960	no	1.000	1.000	1.000	0.313	0.292	0.042	0.240	0.756
snp246	ss	2	5895850	no	0.771	1.000	0.885	0.229	0.167	0.083	0.177	0.663
snp84	ss	2	5896615	no	0.583	1.000	0.792	0.146	0.292	0.083	0.167	0.551
snp247	sxA	2	6262685	no	0.979	0.674	0.827	1.000	1.000	1.000	1.000	0.163
snp54	sxA	2	6294305	no	0.583	1.000	0.792	1.000	1.000	1.000	1.000	0.205
snp297	sxB	2	6294833	no	1.000	1.000	1.000	0.479	0.000	0.333	0.323	0.671
snp287	sf	2	6660331	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp165	sxA	2	6660358	no	0.521	0.500	0.510	0.000	0.000	0.000	0.000	0.513
snp370	sxB	2	6801805	no	1.000	1.000	1.000	0.087	0.250	1.000	0.356	0.631
snp78	sxB	2	6847076	no	1.000	1.000	1.000	0.646	0.625	0.500	0.604	0.388
snp276	sxA	2	7079634	no	0.792	0.804	0.798	0.000	0.000	0.000	0.000	0.798
snp249	ss	3	451355	no	0.896	0.609	0.752	0.917	1.000	1.000	0.958	0.146
snp341	sxA	3	1100321	no	0.938	0.913	0.925	1.000	1.000	1.000	1.000	0.066
snp369	sxB	3	1101017	no	1.000	1.000	1.000	0.917	0.833	1.000	0.917	0.074
snp27	sxB	3	1161884	no	1.000	1.000	1.000	0.896	0.875	1.000	0.917	0.071
snp288	sxA	3	1162010	no	0.167	0.413	0.290	0.000	0.000	0.000	0.000	0.281
snp256	sxB	3	1190641	no	1.000	1.000	1.000	0.604	0.750	0.875	0.708	0.281
snp187	sf	3	1190698	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp293	sxA	3	1644402	no	0.875	1.000	0.938	1.000	1.000	1.000	1.000	0.052
snp379	sxB	3	1798687	no	1.000	1.000	1.000	0.938	0.958	1.000	0.958	0.031
snp244	sxB	3	1848532	no	1.000	1.000	1.000	0.750	0.958	1.000	0.865	0.123

snp377	sxA	3	1848910	no	0.354	0.565	0.460	0.000	0.000	0.000	0.000	0.454
snp294	sf	3	1930758	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp155	sxA	3	2429568	no	0.750	0.935	0.842	0.000	0.000	0.000	0.000	0.841
snp194	sxA	3	2531188	no	0.500	0.674	0.587	0.000	0.000	0.000	0.000	0.582
snp220	sxB	3	2531736	no	1.000	1.000	1.000	0.417	0.375	0.708	0.479	0.511
snp120	sf	3	2728434	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp214	sxA	3	2743934	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp17	sxB	3	2830088	no	1.000	1.000	1.000	0.833	1.000	1.000	0.917	0.074
snp197	sf	3	3843873	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp45	sf	3	4740049	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp134	sxB	3	4982458	no	1.000	1.000	1.000	0.875	0.708	0.708	0.792	0.197
snp233	sf	3	5003239	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp41	sxA	3	5534447	no	0.708	0.913	0.811	1.000	1.000	1.000	1.000	0.185
snp59	sxB	3	5760414	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp371	sxA	3	5761572	no	0.688	1.000	0.844	0.000	0.000	0.000	0.000	0.840
snp82	sxB	3	5761929	no	1.000	1.000	1.000	0.979	0.958	1.000	0.979	0.010
snp205	sxA	3	6085809	no	0.188	0.543	0.365	0.000	0.000	0.000	0.000	0.357
snp136	sf	4	191930	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp72	sf	4	420193	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp37	sxA	4	420424	no	0.438	0.826	0.632	0.000	0.000	0.000	0.000	0.626
snp2	sxA	4	608398	no	0.792	0.739	0.765	1.000	1.000	1.000	1.000	0.227
snp259	ss	4	637562	no	0.875	0.870	0.872	1.000	1.000	1.000	1.000	0.119
snp50	sxB	4	638933	no	1.000	1.000	1.000	0.958	0.583	0.500	0.750	0.239
snp303	sf	4	707889	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp223	sxB	4	1270534	no	1.000	1.000	1.000	0.854	0.667	0.875	0.813	0.177
snp248	sxB	4	3237733	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp150	ss	4	3575332	no	0.958	1.000	0.979	0.208	0.125	0.000	0.135	0.835
snp258	sxA	4	3617354	no	0.479	0.413	0.446	1.000	1.000	1.000	1.000	0.552
snp225	sxB	4	3617924	no	1.000	1.000	1.000	0.875	0.917	0.958	0.906	0.082
snp77	sxB	4	3889603	no	1.000	1.000	1.000	0.958	0.958	0.875	0.938	0.053
snp349	sxB	4	4028088	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp362	ss	4	4235976	no	0.979	0.674	0.827	0.979	1.000	1.000	0.990	0.137
snp345	sf	4	4412519	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp270	sxB	4	4475096	no	1.000	1.000	1.000	0.708	0.750	1.000	0.792	0.197
snp309	sxA	4	4475480	no	0.875	0.848	0.861	1.000	1.000	1.000	1.000	0.130
snp75	ss	4	4831084	no	0.250	0.326	0.288	0.000	0.000	0.000	0.000	0.284
snp211	sxA	4	4845933	no	0.771	0.826	0.798	1.000	1.000	1.000	1.000	0.194
snp38	sxB	4	4846290	no	1.000	1.000	1.000	0.833	1.000	0.955	0.905	0.087
snp333	sxA	4	4852526	no	0.417	0.522	0.469	0.000	0.000	0.000	0.000	0.464
snp24	sxB	4	4867595	no	1.000	1.000	1.000	0.833	0.375	0.792	0.708	0.280
snp286	ss	4	4971806	no	0.813	1.000	0.906	0.854	0.917	0.750	0.844	0.005
snp12	sf	4	5111307	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp36	sxA	4	5111470	no	0.417	1.000	0.708	1.000	1.000	1.000	1.000	0.290
snp268	sf	4	5119511	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp161	sf	5	76363	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp44	sxB	5	153337	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp278	sxA	5	153688	no	0.958	1.000	0.979	1.000	1.000	1.000	1.000	0.011
snp118	sxB	5	197127	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp252	ss	5	529117	no	0.938	1.000	0.969	0.771	0.958	0.909	0.852	0.071
snp204	sxB	5	699006	no	1.000	1.000	1.000	0.917	0.542	1.000	0.844	0.137
snp178	sxB	5	1279954	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp16	sxB	5	2038176	no	1.000	1.000	1.000	1.000	1.000	0.917	0.979	0.000
snp274	sf	5	3191768	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp48	ss	5	3208641	no	0.583	0.565	0.574	0.896	0.958	0.958	0.927	0.278

snp51	sxB	5	3219354	no	1.000	1.000	1.000	1.000	0.917	0.833	0.938	0.053
snp359	sxA	5	3261936	no	0.979	1.000	0.990	1.000	1.000	1.000	1.000	0.000
snp33	sxB	5	3262071	no	1.000	1.000	1.000	0.583	0.417	0.542	0.531	0.459
snp283	ss	5	3414778	no	0.833	1.000	0.917	0.917	0.750	0.750	0.833	0.019
snp52	sf	5	3425248	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp241	sxB	6	1282558	no	1.000	1.000	1.000	0.500	0.958	0.458	0.604	0.384
snp236	sxB	7	136337	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp330	sf	7	259834	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp272	sxB	7	418779	no	1.000	1.000	1.000	0.250	0.292	0.167	0.240	0.756
snp121	sf	7	463451	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp85	sxB	7	463496	no	1.000	1.000	1.000	0.583	0.375	0.125	0.417	0.575
snp200	ss	7	881023	no	0.958	1.000	0.979	0.646	0.458	0.875	0.656	0.287
CesA	ss	7	1306948	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp179	sf	7	1493954	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp10	sf	7	1770784	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp381	ss	7	1770922	no	0.604	0.761	0.683	1.000	1.000	1.000	1.000	0.313
snp212	sxB	7	1771534	no	1.000	1.000	1.000	0.396	0.750	0.250	0.448	0.544
snp87	sxB	7	2098237	no	1.000	1.000	1.000	0.771	0.292	0.917	0.688	0.301
snp235	sxA	7	2098582	no	0.979	0.978	0.979	0.000	0.000	0.000	0.000	0.979
snp354	sxA	7	2203208	no	0.583	0.609	0.596	1.000	1.000	1.000	1.000	0.401
snp149	ss	7	2203361	no	0.792	0.761	0.776	0.167	0.042	0.333	0.177	0.525
snp238	sxA	7	2209315	no	0.833	0.587	0.710	0.000	0.000	0.000	0.000	0.712
snp196	sxB	7	2209624	no	1.000	1.000	1.000	0.333	0.500	0.500	0.417	0.575
snp184	sxB	7	2222540	no	1.000	1.000	1.000	0.021	0.000	0.000	0.010	0.989
snp29	sf	7	2222625	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp137	sxA	7	2268247	no	0.729	0.391	0.560	1.000	1.000	1.000	1.000	0.433
snp94	sxA	7	2465122	no	0.417	0.478	0.447	0.000	0.000	0.000	0.000	0.443
snp131	sxB	7	2915901	no	1.000	1.000	1.000	0.625	0.750	0.542	0.635	0.356
snp114	sxB	7	4992424	no	1.000	1.000	1.000	0.729	0.833	0.917	0.802	0.186
snp62	sf	8	160616	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp209	sxA	8	174450	no	0.938	0.935	0.936	0.000	0.000	0.000	0.000	0.936
snp384	sxA	8	256454	no	0.813	0.739	0.776	0.000	0.000	0.000	0.000	0.776
snp217	sxB	8	393893	no	1.000	1.000	1.000	0.958	1.000	1.000	0.979	0.010
snp314	sxB	8	413938	no	1.000	1.000	1.000	0.125	0.250	0.000	0.125	0.872
snp86	sf	8	413980	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp253	sxA	8	840507	no	0.375	0.761	0.568	1.000	1.000	1.000	1.000	0.431
snp90	sxA	8	962330	no	1.000	0.978	0.989	0.000	0.000	0.000	0.000	0.989
snp307	ss	8	1859346	no	0.979	0.913	0.946	0.917	0.958	0.917	0.927	0.000
snp332	sf	8	2313257	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp300	sxB	8	2948303	no	1.000	1.000	1.000	0.563	0.417	0.333	0.469	0.524
snp67	sxA	8	3154619	no	0.750	0.826	0.788	1.000	1.000	1.000	1.000	0.206
snp365	sf	8	3667430	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp92	sf	8	4161357	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp372	sxB	8	4161426	no	1.000	1.000	1.000	0.604	0.455	0.458	0.530	0.463
snp135	sxB	8	4786382	no	1.000	1.000	1.000	0.667	0.000	0.208	0.385	0.606
snp15	sxB	8	5073716	no	1.000	1.000	1.000	0.955	1.000	1.000	0.977	0.003
snp380	sxB	8	5232250	no	1.000	1.000	1.000	0.750	1.000	0.958	0.865	0.126
snp168	sxB	9	1484808	no	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000
snp133	sxB	9	2358923	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp125	sxB	9	2732667	no	1.000	1.000	1.000	0.833	0.917	0.833	0.854	0.132
snp296	sxA	9	2771243	no	0.813	0.587	0.700	1.000	1.000	1.000	1.000	0.292
snp265	sxB	9	3628263	no	1.000	1.000	1.000	0.833	0.500	0.750	0.729	0.260
snp117	sxB	9	3742199	no	1.000	1.000	1.000	0.042	0.000	0.083	0.042	0.958
snp5	sf	9	4514869	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000

snp89	sxB	9	4515172	no	1.000	1.000	1.000	0.729	0.417	0.792	0.667	0.325
snp317	sxB	9	5185795	no	1.000	1.000	1.000	0.646	0.250	0.250	0.448	0.543
snp219	sf	9	5379609	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp144	ss	9	5379938	no	0.938	0.891	0.914	0.688	1.000	0.875	0.813	0.031
snp152	sxB	10	526821	yes	1.000	1.000	1.000	0.875	0.875	0.917	0.885	0.105
snp8	sxB	10	533223	yes	1.000	1.000	1.000	0.354	0.292	0.167	0.292	0.703
snp373	sxA	10	828723	yes	0.646	0.848	0.747	1.000	1.000	1.000	1.000	0.248
snp185	ss	10	828753	yes	0.500	0.848	0.674	0.813	1.000	0.958	0.896	0.129
snp164	sxA	10	1326368	no	0.958	1.000	0.979	1.000	1.000	1.000	1.000	0.011
snp104	sxB	10	1326494	no	1.000	1.000	1.000	0.792	0.708	1.000	0.823	0.166
snp132	ss	10	1370640	no	0.958	1.000	0.979	1.000	1.000	1.000	1.000	0.011
snp229	sf	10	1370821	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp325	sf	10	1431323	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp156	sxB	10	1431944	no	1.000	1.000	1.000	0.271	0.625	0.417	0.396	0.597
snp237	sxA	10	1431979	no	0.854	1.000	0.927	0.000	0.000	0.000	0.000	0.926
snp190	sf	10	1644037	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp374	sxB	10	1829131	no	1.000	1.000	1.000	0.042	0.000	0.375	0.115	0.883
snp83	sxB	10	3881287	no	1.000	1.000	1.000	0.375	0.917	0.875	0.635	0.354
snp122	sxA	10	4729725	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp221	sxB	10	4866550	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp81	sxB	11	977658	no	1.000	1.000	1.000	0.771	0.708	0.708	0.740	0.251
snp166	ss	11	1904080	no	0.958	0.630	0.794	0.000	0.000	0.000	0.000	0.797
snp226	sf	11	1907473	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp203	sxA	11	1907554	no	0.625	0.652	0.639	1.000	1.000	1.000	1.000	0.357
snp6	sxA	11	3225028	no	0.771	0.696	0.733	1.000	1.000	1.000	1.000	0.260
snp151	sf	11	3291525	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp352	sf	11	4253899	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp175	sxA	11	4253954	no	1.000	0.978	0.989	1.000	1.000	1.000	1.000	0.000
snp262	sxB	11	4790343	no	1.000	1.000	1.000	0.896	0.792	0.875	0.865	0.121
snp315	sxB	11	4987955	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp174	sf	12	223704	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp210	sf	12	759131	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp348	sxB	12	1086526	no	1.000	1.000	1.000	0.938	0.917	1.000	0.948	0.042
snp222	sf	12	1503856	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp364	sxB	12	1572677	no	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000
snp299	sxA	12	1573004	no	0.500	0.739	0.620	1.000	1.000	1.000	1.000	0.378
snp353	sxB	14	283115	no	1.000	1.000	1.000	0.208	0.042	0.000	0.115	0.883
snp99	sxB	14	1944967	no	1.000	1.000	1.000	1.000	0.958	1.000	0.990	0.000
snp160	sxB	14	2052390	no	1.000	1.000	1.000	0.625	0.750	0.958	0.740	0.249
snp312	sxA	14	2235245	no	0.708	0.761	0.735	1.000	1.000	1.000	1.000	0.262
snp251	sf	14	2235545	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp326	sxB	14	2236220	no	1.000	1.000	1.000	0.250	0.125	0.583	0.302	0.693
snp357	sxB	14	2424063	no	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000
snp95	ss	14	2424137	no	0.208	0.565	0.387	0.021	0.042	0.000	0.021	0.332
snp254	sxB	14	3284624	no	1.000	1.000	1.000	0.396	0.042	0.000	0.208	0.787
snp167	sxB	14	3748763	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp19	sf	23	43134	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp20	sxB	23	48748	no	1.000	1.000	1.000	0.021	0.083	0.042	0.042	0.958
snp181	ss	37	84812	no	0.625	1.000	0.813	0.000	0.000	0.000	0.000	0.808
snp273	ss	37	152661	no	0.938	0.913	0.925	1.000	1.000	1.000	1.000	0.066
snp158	sxA	42	67444	no	1.000	0.957	0.978	1.000	1.000	1.000	1.000	0.011
snp271	sxB	49	140348	no	1.000	1.000	1.000	1.000	1.000	0.917	0.979	0.010
snp193	sxA	49	273227	no	0.813	0.609	0.711	1.000	1.000	1.000	1.000	0.283
snp106	sxA	52	936	no	0.979	0.935	0.957	1.000	1.000	1.000	1.000	0.033

snp263	sxB	62	63214	no	1.000	1.000	1.000	0.063	0.000	0.000	0.031	0.968
snp142	ss	62	63338	no	0.958	1.000	0.979	1.000	1.000	1.000	1.000	0.011
snp177	ss	75	691527	no	0.625	0.761	0.693	0.000	0.000	0.000	0.000	0.690
snp282	ss	94	369979	no	0.917	1.000	0.958	0.958	0.875	0.708	0.875	0.031
snp30	sxA	98	54303	no	0.875	0.565	0.720	1.000	1.000	1.000	1.000	0.269
snp107	sxA	98	128272	no	0.688	0.826	0.757	0.000	0.000	0.000	0.000	0.754
snp169	sxB	98	128407	no	1.000	1.000	1.000	0.646	0.958	1.000	0.813	0.174
snp207	sxB	98	128494	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp351	sxB	98	166571	no	1.000	1.000	1.000	1.000	0.958	1.000	0.990	0.000
snp183	sxA	98	207095	no	0.917	1.000	0.958	1.000	1.000	1.000	1.000	0.033
snp360	sf	98	910443	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp257	ss	98	915948	no	0.708	0.587	0.648	0.708	1.000	1.000	0.854	0.098
snp323	sxA	103	62605	no	0.271	0.696	0.483	1.000	1.000	1.000	1.000	0.517
snp337	sxB	103	62716	no	1.000	1.000	1.000	0.458	0.542	0.500	0.490	0.507
snp3	sf	103	362542	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp366	sxB	103	362833	no	1.000	1.000	1.000	0.833	0.708	0.833	0.802	0.186
snp69	sxB	103	366420	no	1.000	1.000	1.000	0.875	0.875	0.917	0.885	0.105
snp119	sxA	103	366738	no	0.896	0.783	0.839	1.000	1.000	1.000	1.000	0.154
snp304	sxB	105	3452	no	1.000	1.000	1.000	0.771	0.958	1.000	0.875	0.114
snp331	sxB	105	3524	no	1.000	1.000	1.000	1.000	0.917	1.000	0.979	0.010
snp79	ss	120	170654	no	0.563	0.891	0.727	0.000	0.000	0.000	0.000	0.722
snp234	sxA	137	179389	no	0.542	0.457	0.499	1.000	1.000	1.000	1.000	0.498
snp23	ss	141	14705	no	0.688	0.652	0.670	0.146	0.667	0.292	0.313	0.220
snp285	sxB	166	224331	no	1.000	1.000	1.000	0.583	0.417	0.500	0.521	0.471
snp291	ss	166	224607	no	0.917	1.000	0.958	0.063	0.167	0.042	0.083	0.865
snp346	sf	166	224649	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp74	ss	166	224850	no	0.771	1.000	0.885	0.000	0.042	0.000	0.010	0.870
snp311	sf	170	35649	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp127	sxB	170	35904	no	1.000	1.000	1.000	0.250	0.375	0.417	0.323	0.671
snp191	sxA	170	36054	no	0.208	0.717	0.463	1.000	1.000	1.000	1.000	0.539
snp356	sxA	184	34695	no	0.958	0.891	0.925	1.000	1.000	1.000	1.000	0.066
snp141	sxB	192	4717	no	1.000	1.000	1.000	0.833	0.333	0.792	0.698	0.293
snp261	sxB	262	248150	no	1.000	1.000	1.000	0.792	0.833	0.458	0.719	0.271
snp9	sxA	290	7970	no	1.000	0.935	0.967	1.000	1.000	1.000	1.000	0.022
snp382	sxB	290	8037	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp65	sxB	290	8101	no	1.000	1.000	1.000	0.646	0.917	0.833	0.760	0.228
snp192	sxA	662	1336	no	0.917	1.000	0.958	1.000	1.000	1.000	1.000	0.028
snp231	sf	662	1378	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp80	sxB	1063	48786	no	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000
snp43	ss	1063	49359	no	0.104	0.978	0.541	0.188	0.750	0.542	0.417	0.010
snp313	sf	1063	49515	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp245	ss	1069	14160	no	0.417	0.543	0.480	0.021	0.375	0.000	0.104	0.287
snp336	sf	1111	36694	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp321	sxA	1114	20660	no	0.854	0.848	0.851	1.000	1.000	1.000	1.000	0.141
snp202	sf	1114	21133	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp324	sxA	1126	10524	no	0.521	0.457	0.489	1.000	1.000	1.000	1.000	0.509
snp280	sxB	1126	10961	no	1.000	1.000	1.000	0.729	0.458	0.500	0.604	0.387
snp327	sf	1126	36378	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp147	sxB	1144	26090	no	1.000	1.000	1.000	0.979	1.000	1.000	0.990	0.000
snp322	sxA	1144	26125	no	0.250	0.522	0.386	1.000	1.000	1.000	1.000	0.615
snp275	sf	1152	2594	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp88	sxB	1157	11870	no	1.000	1.000	1.000	0.417	0.458	0.917	0.552	0.437
snp195	ss	1157	12038	no	0.396	0.478	0.437	0.000	0.000	0.000	0.000	0.432
snp111	sf	1158	18976	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000

snp26	sf	1160	30058	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
Hox5	sxA	1177	-	no	1.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000
snp34	sxB	1272	9864	no	1.000	1.000	1.000	0.083	0.000	0.000	0.042	0.958

Table S3. Estimates of pairwise population genetic differentiation for *Ciona robusta* (a) and *C. intestinalis* (b).

Fixation indices F_{ST} were computed based 111 and 150 polymorphic loci for *C. robusta* and *C. intestinalis*, respectively. Bold numbers indicate statistical significance (P -value < 0.05).

a)C. robusta	1-Guanaqueros	2-Etang de Thau	5-Falmouth	6-Plymouth	7-St Vaast	8-Perros Guirec	10-Moulin Blanc	11-Camaret
2-Etang de Thau	0.140							
5-Falmouth	0.148	0.056						
6-Plymouth	0.131	0.039	0.017					
7-St Vaast	0.142	0.021	0.041	0.026				
8-Perros Guirec	0.137	0.029	0.024	0.025	0.018			
10-Moulin Blanc	0.152	0.028	0.046	0.035	0.018	0.018		
11-Camaret	0.137	0.022	0.027	0.027	0.016	0.015	0.005	
12-Quiberon	0.132	0.013	0.014	0.013	0.011	0.014	0.024	0.012

b)C. intestinalis	3-Nahant	4a-Fiske.- surface	4b-Fiske.- 20m	5-Falmouth	6-Plymouth	7-St Vaast	8-Perros Guirec	9-Aber Wrac'h	10-Moulin Blanc	11-Camaret
4a-Fiske.-surface	0.147									
4b-Fiske.-20m	0.153	0.172								
5-Falmouth	0.052	0.097	0.144							
6-Plymouth	0.061	0.128	0.170	0.022						
7-St Vaast	0.047	0.104	0.130	0.018	0.042					
8-Perros Guirec	0.042	0.106	0.130	0.015	0.031	0.002				
9-Aber Wrac'h	0.034	0.116	0.146	0.015	0.026	0.014	0.007			
10-Moulin Blanc	0.033	0.111	0.140	0.021	0.022	0.019	0.010	0.004		
11-Camaret	0.040	0.108	0.145	0.008	0.026	0.010	0.007	-0.004	0.003	
12-Quiberon	0.037	0.123	0.133	0.022	0.027	0.021	0.008	0.004	-0.003	-0.000

Table S4. **Characteristics of the 40 SNPs sharing polymorphism between the two species over the 105 SNPs selected for interspecific gene flow analyses ($F_{ST} > 0.9$).**

Allele frequencies in populations sampled in the allopatric (localities 1 and 2 for *C. robusta* and localities 3 and 4a/4b for *C. intestinalis*) and sympatric ranges (Loc. 5 to 12 for both species) of the two species are given. See caption of Table S2 for details of abbreviation. In bold are loci showing atypical allele frequencies (i.e. higher frequency of *C. intestinalis* allele) in the single populations of SE Pacific (Loc.1) compared to others populations of *C. robusta*.

Index	SNPs characteristics				Allele frequency								Ensembl Gene ID	Protein Description
	cat.	Chr	Position	HS introg	<i>C. robusta</i>			<i>C. intestinalis</i>						
					Loc. 1	Loc. 2	Loc. 5 to 12	Loc. 3	Loc. 4a	Loc. 4b	Loc. 5 to 12			
snp18	ss	1	2205837	yes	0.875	1.000	0.994	0.000	0.000	0.000	0.008	ENSCING00000004302	Uncharacterized protein	
snp343	sxB	1	4762944	no	1.000	1.000	1.000	0.000	0.000	0.000	0.005	ENSCING00000005260	Uncharacterized protein	
snp240	sxB	1	5073937	no	1.000	1.000	1.000	0.000	0.000	0.000	0.005	ENSCING00000019274	Proteasome Z submit (zeta-201)	
snp47	sxB	1	5695260	no	1.000	1.000	1.000	0.063	0.000	0.000	0.101	ENSCING00000012769	Zinc finger protein	
snp206	sxB	1	6506772	no	1.000	1.000	1.000	0.021	0.000	0.000	0.005	Unknown		
snp11	sxA	1	8307018	no	0.938	1.000	0.982	0.000	0.000	0.000	0.000	Unknown		
snp239	sxB	1	9032786	no	1.000	1.000	1.000	0.125	0.000	0.000	0.021	ENSCING00000000038	Uncharacterized protein	
snp55	sxB	2	173448	yes	1.000	1.000	1.000	0.125	0.083	0.000	0.040	Unknown		
snp154	sxB	2	179880	yes	1.000	1.000	1.000	0.042	0.000	0.000	0.045	ENSCING00000004185	Uncharacterized protein	
snp66	sxB	2	228866	yes	1.000	1.000	1.000	0.000	0.000	0.000	0.048	Unknown		
snp39	sxB	2	251567	yes	1.000	1.000	1.000	0.000	0.000	0.000	0.005	ENSCING00000004218	Uncharacterized protein	
snp35	sxB	2	476678	yes	1.000	1.000	1.000	0.000	0.000	0.000	0.003	Unknown		
snp290	sxA	2	483036	yes	0.833	1.000	1.000	0.000	0.000	0.000	0.000	Unknown		
snp116	sxB	2	1834123	no	1.000	1.000	1.000	0.000	0.000	0.000	0.013	Unknown		
snp31	sxB	2	2372600	no	1.000	1.000	1.000	0.000	0.000	0.000	0.005	Unknown		
snp376	sxB	2	2417520	no	1.000	1.000	1.000	0.063	0.000	0.000	0.016	ENSCING00000005003	Uncharacterized protein	
snp120	sxB	3	2728434	no	1.000	1.000	1.000	0.000	0.000	0.000	0.003	ENSCING00000023156	Uncharacterized protein	
snp214	sxA	3	2743934	no	1.000	1.000	0.997	0.000	0.000	0.000	0.000	ENSCING00000009577	Uncharacterized protein	
snp59	sxB	3	5760414	no	1.000	1.000	1.000	0.000	0.000	0.000	0.019	ENSCING00000019791	Uncharacterized protein	
snp248	sxB	4	3237733	no	1.000	1.000	1.000	0.000	0.000	0.000	0.003	Unknown		
snp349	sxB	4	4028088	no	1.000	1.000	1.000	0.000	0.000	0.000	0.008	ENSCING00000008884	Uncharacterized protein	
snp44	sxB	5	153337	no	1.000	1.000	1.000	0.000	0.000	0.000	0.003	ENSCING00000022288	Uncharacterized protein	
snp118	sxB	5	197127	no	1.000	1.000	1.000	0.000	0.000	0.000	0.003	ENSCING00000007496	Uncharacterized protein	
snp178	sxB	5	1279954	no	1.000	1.000	1.000	0.000	0.000	0.000	0.038	Unknown		
snp236	sxB	7	136337	no	1.000	1.000	1.000	0.000	0.000	0.000	0.008	ENSCING00000024960	Uncharacterized protein	

snp235	sxA	7	2098582	no	0.979	0.978	0.928	0.000	0.000	0.000	0.000	Unknown	
snp184	sxB	7	2222540	no	1.000	1.000	1.000	0.021	0.000	0.000	0.048	ENSCING00000007874	Uncharacterized protein
snp209	sxA	8	174450	no	0.938	0.935	0.855	0.000	0.000	0.000	0.000	ENSCING00000008350	Uncharacterized protein
snp90	sxA	8	962330	no	1.000	0.978	0.994	0.000	0.000	0.000	0.000	ENSCING00000020390	Ar20 protein
snp133	sxB	9	2358923	no	1.000	1.000	1.000	0.000	0.000	0.000	0.080	ENSCING00000014133	Eukaryotic translation initiation factor 6
snp117	sxB	9	3742199	no	1.000	1.000	1.000	0.042	0.000	0.083	0.016	ENSCING00000013017	60S ribosomal protein L18
snp237	sxA	10	1431979	no	0.854	1.000	0.976	0.000	0.000	0.000	0.000	ENSCING00000010029	Histone H2A
snp122	sxA	10	4729725	no	1.000	1.000	0.982	0.000	0.000	0.000	0.000	Unknown	
snp221	sxB	10	4866550	no	1.000	1.000	1.000	0.000	0.000	0.000	0.021	ENSCING00000023081	Uncharacterized protein
snp315	sxB	11	4987955	no	1.000	1.000	1.000	0.000	0.000	0.000	0.005	ENSCING00000000723	Uncharacterized protein
snp167	sxB	14	3748763	no	1.000	1.000	1.000	0.000	0.000	0.000	0.013	ENSCING00000004075	Transcription factor protein (hmg 1/2)
snp20	sxB	23	48748	no	1.000	1.000	1.000	0.021	0.083	0.042	0.016	Unknown	
snp263	sxB	62	63214	no	1.000	1.000	1.000	0.063	0.000	0.000	0.000	ENSCING00000001276	Uncharacterized protein
snp207	sxB	98	128494	no	1.000	1.000	1.000	0.000	0.000	0.000	0.016	ENSCING00000016399	Uncharacterized protein
snp34	sxB	1272	9864	no	1.000	1.000	1.000	0.083	0.000	0.000	0.043	ENSCING00000008321	Zinc finger protein

Table S5. Results of the f_3 statistics testing the null hypothesis that the evolutionary history of *Ciona* populations is consistent with absence of migration events between populations of the two species.

The f_3 -statistics was done for all combinations of three populations: the ‘target population’ is tested as resulting from admixture from sources 1 and 2 (ancestors). Numbers in the table corresponds to locality (Table 1 in the main text). Negative values of f_3 are interpreted as evidence of admixture (for more details see Reich *et al.* (2009)). Deviation of the f_3 mean to 0 was tested using a t -test of Student. In bold are f_3 statistics that are significantly less than 0 ($P < 0.001$).

Target population	Source 1 (<i>C. robusta</i>)	Source 2 (<i>C. intestinalis</i>)	f_3 statistics (mean)	Standard deviation
<i>C. robusta</i>				
no.1	no.7 to 12	no.7 to 12	-0.0011	0.003
no.1	no.7 to 12	no.5 and 6	-0.0009	0.003
no.1	no.7 to 12	no.4a and 4b	-0.0002	0.004
no.1	no.7 to 12	no.3	-0.0011	0.004
no.1	no.5 and 6	no.7 to 12	0.0001	0.004
no.1	no.5 and 6	no.5 and 6	0.0004	0.004
no.1	no.5 and 6	no.4a and 4b	0.0011	0.004
no.1	no.5 and 6	no.3	0.0000	0.004
no.1	no.2	no.7 to 12	-0.0026	0.004
no.1	no.2	no.5 and 6	-0.0024	0.004
no.1	no.2	no.4a and 4b	-0.0017	0.004
no.1	no.2	no.3	-0.0029	0.004
no.2	no.7 to 12	no.7 to 12	0.0026	0.001
no.2	no.7 to 12	no.5 and 6	0.0025	0.001
no.2	no.7 to 12	no.4a and 4b	0.0025	0.001
no.2	no.7 to 12	no.3	0.0027	0.001
no.2	no.5 and 6	no.7 to 12	0.0049	0.002
no.2	no.5 and 6	no.5 and 6	0.0050	0.003
no.2	no.5 and 6	no.4a and 4b	0.0050	0.002
no.2	no.5 and 6	no.3	0.0050	0.002
no.2	no.1	no.7 to 12	0.0188	0.005
no.2	no.1	no.5 and 6	0.0186	0.005
no.2	no.1	no.4a and 4b	0.0178	0.004
no.2	no.1	no.3	0.0190	0.005
no.5 and 6	no.7 to 12	no.7 to 12	-0.0004	0.001
no.5 and 6	no.7 to 12	no.5 and 6	-0.0005	0.001
no.5 and 6	no.7 to 12	no.4a and 4b	-0.0005	0.001
no.5 and 6	no.7 to 12	no.3	-0.0003	0.001
no.5 and 6	no.1	no.7 to 12	0.0157	0.005
no.5 and 6	no.1	no.5 and 6	0.0154	0.005
no.5 and 6	no.1	no.4a and 4b	0.0146	0.005
no.5 and 6	no.1	no.3	0.0158	0.005
no.5 and 6	no.2	no.7 to 12	-0.0010	0.002
no.5 and 6	no.2	no.5 and 6	-0.0010	0.002
no.5 and 6	no.2	no.4a and 4b	-0.0010	0.002

no.5 and 6	no.2	no.3	-0.0011	0.002
no.7 to 12	no.5 and 6	no.7 to 12	0.0019	0.001
no.7 to 12	no.5 and 6	no.5 and 6	0.0021	0.001
no.7 to 12	no.5 and 6	no.4a and 4b	0.0021	0.001
no.7 to 12	no.5 and 6	no.3	0.0019	0.001
no.7 to 12	no.1	no.7 to 12	0.0168	0.004
no.7 to 12	no.1	no.5 and 6	0.0166	0.004
no.7 to 12	no.1	no.4a and 4b	0.0160	0.004
no.7 to 12	no.1	no.3	0.0169	0.004
no.7 to 12	no.2	no.7 to 12	-0.0010	0.001
no.7 to 12	no.2	no.5 and 6	-0.0010	0.001
no.7 to 12	no.2	no.4a and 4b	-0.0009	0.001
no.7 to 12	no.2	no.3	-0.0012	0.001

C. intestinalis

no.3	no.7 to 12	no.7 to 12	0.0028	0.002
no.3	no.7 to 12	no.5 and 6	0.0003	0.003
no.3	no.7 to 12	no.4a and 4b	0.0008	0.004
no.3	no.5 and 6	no.7 to 12	0.0028	0.002
no.3	no.5 and 6	no.5 and 6	0.0005	0.003
no.3	no.5 and 6	no.4a and 4b	0.0010	0.004
no.3	no.1	no.7 to 12	0.0027	0.002
no.3	no.1	no.5 and 6	0.0001	0.003
no.3	no.1	no.4a and 4b	-0.0001	0.004
no.3	no.2	no.7 to 12	0.0030	0.002
no.3	no.2	no.5 and 6	0.0005	0.003
no.3	no.2	no.4a and 4b	0.0011	0.004
no.4a and 4b	no.7 to 12	no.7 to 12	0.0122	0.004
no.4a and 4b	no.7 to 12	no.5 and 6	0.0092	0.003
no.4a and 4b	no.7 to 12	no.3	0.0134	0.004
no.4a and 4b	no.5 and 6	no.7 to 12	0.0120	0.004
no.4a and 4b	no.5 and 6	no.5 and 6	0.0091	0.003
no.4a and 4b	no.5 and 6	no.3	0.0132	0.004
no.4a and 4b	no.1	no.7 to 12	0.0131	0.003
no.4a and 4b	no.1	no.5 and 6	0.0099	0.003
no.4a and 4b	no.1	no.3	0.0144	0.004
no.4a and 4b	no.2	no.7 to 12	0.0121	0.004
no.4a and 4b	no.2	no.5 and 6	0.0092	0.003
no.4a and 4b	no.2	no.3	0.0132	0.004
no.5 and 6	no.7 to 12	no.7 to 12	0.0043	0.002
no.5 and 6	no.7 to 12	no.4a and 4b	0.0021	0.003
no.5 and 6	no.7 to 12	no.3	0.0059	0.003
no.5 and 6	no.5 and 6	no.7 to 12	0.0041	0.002
no.5 and 6	no.5 and 6	no.4a and 4b	0.0022	0.003
no.5 and 6	no.5 and 6	no.3	0.0057	0.003
no.5 and 6	no.1	no.7 to 12	0.0044	0.002
no.5 and 6	no.1	no.4a and 4b	0.0015	0.003
no.5 and 6	no.1	no.3	0.0061	0.003
no.5 and 6	no.2	no.7 to 12	0.0042	0.002

no.5 and 6	no.2	no.4a and 4b	0.0022	0.003
no.5 and 6	no.2	no.3	0.0057	0.003
no.7 to 12	no.7 to 12	no.5 and 6	-0.0025	0.001
no.7 to 12	no.7 to 12	no.4a and 4b	-0.0016	0.003
no.7 to 12	no.7 to 12	no.3	0.0016	0.002
no.7 to 12	no.5 and 6	no.5 and 6	-0.0024	0.001
no.7 to 12	no.5 and 6	no.4a and 4b	-0.0014	0.003
no.7 to 12	no.5 and 6	no.3	0.0016	0.002
no.7 to 12	no.1	no.5 and 6	-0.0027	0.001
no.7 to 12	no.1	no.4a and 4b	-0.0024	0.003
no.7 to 12	no.1	no.3	0.0017	0.002
no.7 to 12	no.2	no.5 and 6	-0.0025	0.001
no.7 to 12	no.2	no.4a and 4b	-0.0015	0.003
no.7 to 12	no.2	no.3	0.0015	0.002
