



**HAL**  
open science

## Marine transmissible cancer navigates urbanised waters, threatening spillover

Maurine Hammel, Fanny Touchard, E a V Burioli, Laure Paradis, Frédérique Cerqueira, Elisa Chailler, Ianis Bernard, H. Cochet, Alexis Simon, Frédéric Thomas, et al.

### ► To cite this version:

Maurine Hammel, Fanny Touchard, E a V Burioli, Laure Paradis, Frédérique Cerqueira, et al.. Marine transmissible cancer navigates urbanised waters, threatening spillover. Proceedings of the Royal Society B: Biological Sciences, 2024, 10.1098/rspb.2023.2541 . hal-04116826v2

HAL Id: hal-04116826

<https://hal.science/hal-04116826v2>

Submitted on 10 Jan 2024

**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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 **Marine transmissible cancer navigates urbanised waters, threatening to spillover**

2

3 **Hammel M.<sup>1,2\*</sup>, Touchard F.<sup>1</sup>, Burioli E. A. V.<sup>1,2</sup>, Paradis L.<sup>1</sup>, Cerqueira F.<sup>1</sup>, Chailier E.<sup>1</sup>,**  
4 **Bernard I.<sup>3</sup>, Cochet H.<sup>4</sup>, Simon A.<sup>1</sup>, Thomas F.<sup>5</sup>, Destoumieux-Garzón D.<sup>2</sup>, Charrière G.**  
5 **M.<sup>2</sup>, Bierne N.<sup>1</sup>**

6 <sup>1</sup> ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

7 <sup>2</sup> IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan Via Domitia, Montpellier, France

8 <sup>3</sup> Eurêka Mer, Lézardrieux, France

9 <sup>4</sup> Cochet Environnement, 56550 Locoal, France

10 <sup>5</sup> CREEC/CANECEV (CREES), MIVEGEC, Unité Mixte de Recherches, IRD 224–CNRS  
11 5290–Université de Montpellier, Montpellier, France

12 \*Corresponding author address: [maurine.hammel@umontpellier.fr](mailto:maurine.hammel@umontpellier.fr)

13

14 **Abstract**

15 Inter-individual transmission of cancer cells represents a unique form of microparasites  
16 increasingly reported in marine bivalves. In this study, we sought to understand the ecology of  
17 the propagation of *Mytilus trossulus* Bivalve Transmissible Neoplasia 2 (MtrBTN2), a  
18 transmissible cancer affecting four *Mytilus* mussel species worldwide. We investigated the  
19 prevalence of MtrBTN2 in the mosaic hybrid zone of *M. edulis* and *M. galloprovincialis* along  
20 the French Atlantic coast, sampling contrasting natural and anthropogenic habitats. We observed  
21 a similar prevalence in both species, likely due to the spatial proximity of the two species in this  
22 region. Our results showed that ports had higher prevalence of MtrBTN2, with a possible hotspot  
23 observed at a shuttle landing dock. No cancer was found in natural beds except for two sites  
24 close to the hotspot, suggesting spillover. Ports may provide favourable conditions for the  
25 transmission of MtrBTN2, such as high mussel density, stressful conditions, sheltered and

26 confined shores, or buffered temperatures. Ships may also spread the disease through biofouling.  
27 Our results suggest ports may serve as epidemiological hubs, with maritime routes providing  
28 artificial gateways for MtrBTN2 propagation. This highlights the importance of preventing  
29 biofouling on docks and ship hulls to limit the spread of marine pathogens hosted by fouling  
30 species.

31

32 **Keywords:** Transmissible cancer, Bivalve Transmissible Neoplasia, epidemiology, biofouling,  
33 spillover, *Mytilus* mussels.

34

## 35 1. Introduction

36

37 Transmissible cancers are caused by malignant cell lineages that have acquired the ability  
38 to infect new hosts through the transmission of living cancer cells. Fourteen transmissible cancer  
39 lineages have been described to date: one in dogs (canine transmissible venereal tumour, CTVT  
40 [1, 2]), two in Tasmanian devils (devil facial tumour, DFT1 and DFT2 [3, 4]), and eleven in  
41 different marine bivalve species (bivalve transmissible neoplasia, BTNs [5–10]). While direct  
42 contact is necessary for CTVT and DFTs transmission, via coitus for the former and biting for  
43 the latter, the transmission of BTNs between shellfish is assumed to occur through the seawater  
44 column [11, 12]. *Mytilus* spp. mussels are affected by two BTN lineages known as MtrBTN1 and  
45 MtrBTN2 that originate from two different *M. trossulus* founder hosts. MtrBTN2 is distributed  
46 worldwide and has crossed the species barrier several times to infect four *Mytilus* species: *M.*  
47 *trossulus* in East Asia and Northern Europe [13–15], *M. chilensis* in South America [9], and *M.*  
48 *edulis* and *M. galloprovincialis* in Europe [9, 16]. Cancerous process in mussels was firstly  
49 described fifty years ago as disseminated neoplasia (DN, [17]) but the demonstration that some  
50 DNs are transmissible is recent [6, 7]. As non-transmissible neoplasia also occurs in mussels [9,  
51 16], a genetic diagnosis is required to ensure the identification of MtrBTNs. The oldest known  
52 genetically validated MtrBTN2 was sampled in 2009 [9]. Genetic data suggest that it could be at  
53 least one or two orders of magnitude more ancient, although dating is difficult due to an apparent  
54 acceleration of the mitochondrial clock [15, 16].

55

56 The global distribution of the MtrBTN2 lineage remains enigmatic [9, 14]. The other known  
57 transmissible cancer with such a worldwide distribution is CTVT in dogs. CTVT emerged 4,000  
58 to 8,000 years ago, probably in Eastern Asia, and has expanded rapidly worldwide over the past  
59 500 years, probably facilitated by the development of maritime transportation [18]. Transmission  
60 of MtrBTN2 within mussel populations occurs presumably through filter-feeding (discussed in  
61 [11, 19]). MtrBTN2 cells can survive for several days in seawater [11], which greatly increases  
62 their chances of infecting new hosts. However, it is unlikely that transport of cancer cells by  
63 marine currents alone explains such a global distribution. Indeed, the dilution of cancer cells in  
64 the marine environment may strongly limit the success of transmission over long distances.  
65 Shipping traffic has been proposed as the most likely explanation for the global distribution of  
66 MtrBTN2 -i.e., the transport of disease-carrying mussels on ship hulls [9, 14, 15]. To our  
67 knowledge, these claims have remained speculative and no formal evidence of the effect of  
68 shipping traffic and port habitat on the epidemiology of BTN has been provided to date.

69  
70 European mussels are affected by MtrBTN2, with a much higher prevalence in *M. edulis* (of the  
71 order of 1/100) than in *M. galloprovincialis* (of the order of 1/1000) [16]. These two host species  
72 are parapatric in Europe and can coexist in contact zones where hybridisation is taking place  
73 [20–22]. Associations between genetic backgrounds and environmental variables indicate that  
74 the environment partly influences the structure and maintenance of the hybrid zones: in the study  
75 area, *M. edulis* genotypes are more frequent in sheltered habitats under freshwater influence,  
76 while *M. galloprovincialis* genotypes are more frequent in habitats exposed to wave action and  
77 ocean salinities [23, 24]. In this context, the distribution and propagation of MtrBTN2 is likely to  
78 be influenced by host genetic background and population composition [16], as well as by other  
79 environmental factors such as population density, reduced hydrodynamics, and stressful  
80 environmental factors (e.g., heat, salinity or pollution stresses). MtrBTN2 global distribution  
81 indicates its resilience across a broad range of temperature with varying seasonal patterns [14],  
82 reflecting host species distribution range. Prior to evidence of horizontal transmission, pollutants  
83 have been suggested as a potential cause of DN; however, the association with disease  
84 prevalence remained inconclusive [25–27], and as we now know that some neoplasias are  
85 transmissible, it is no longer the initial carcinogenesis but the transmission that needs to be  
86 explained. Seasonal influence on DN prevalence was observed at a local scale (higher during

87 winter in *M. trossulus*, [28, 29], and *M. edulis*, [30]), whereas no significant influence of  
88 seasonality was noted at the European scale [31]. Unfortunately, in the absence of genetic  
89 diagnostics, we cannot know how many of these DNAs were indeed BTNs (i.e. transmissible) or  
90 not, and rigorous evaluation of MtrBTN2 prevalence seasonality is still lacking. In brief, little is  
91 known about the ecology of MtrBTN2 in relation to its host and the possible environmental  
92 factors influencing the epidemiology of this transmissible cancer.

93  
94 In the present study, we investigated the occurrence of MtrBTN2 in a survey area in Southern  
95 Brittany (France) which has contrasting natural and anthropogenic habitats and a mosaic  
96 distribution of *M. edulis* and *M. galloprovincialis*. We sampled 40 natural populations as well as  
97 9 mussel farms, 7 floating buoys, and 20 ports. To detect MtrBTN2-infected mussels, we used  
98 digital PCR and real-time PCR, amplifying one nuclear and two mitochondrial genes with  
99 published and newly designed primers targeting specific variants. We found a low prevalence of  
100 MtrBTN2 (23/1516), equally affecting the two coexisting *Mytilus* species. Interestingly, most  
101 affected sites (7/9) were ports, suggesting that ports could be favourable habitats for MtrBTN2  
102 transmission and/or that maritime transport could play a role as a vector for the spread of the  
103 disease. This highlights the possible role of ports as epidemiological hubs for MtrBTN2  
104 propagation.

105

## 106 **2. Material and methods**

107

### 108 **2.1. Sampling**

109 We collected 1516 *Mytilus* spp. mussels from 76 sites of the French Atlantic coast, from the Bay  
110 of Quiberon to Pornic, between January and March 2020. Approximately 20 mussels were  
111 sampled at each site (Table S1), allowing us to have a large number of closely located sites. We  
112 collected hemolymph from adductor muscle (with a 1 ml syringe, 26G needle) and a piece of  
113 mantle; both samples were fixed with 96% ethanol as described in [16]. DNA extraction was  
114 performed with the NucleoMag 96 Tissue kit (Macherey-Nagel) using a Kingfisher Flex  
115 extraction robot (serial number 711-920, ThermoFisher Scientific). We followed the kit protocol  
116 with a modified lysis time of 1 hour at 56°C and modified the volumes of some reagents: beads  
117 were diluted twice in water, 200 µl of MB3 and MB4, 300 µl of MB5 and 100 µl of MB6. DNA

118 concentration (ng/ $\mu$ l) was quantified using the Nanodrop 8000 spectrophotometer (ThermoFisher  
119 Scientific). DNA from hemolymph was used for MtrBTN2 detection and DNA from mantle was  
120 used for mussel genotyping (Figure S1).

121

## 122 **2.2. MtrBTN2 detection**

123 As the prevalence of MtrBTN2 is expected to be low [16, 32] and the number of samples is high  
124 ( $n = 1516$ ), we chose to pre-screen by pooling 12 samples (pooled screening step; Figure S1) and  
125 then demultiplexing positive pools to specifically target cancerous samples (simplex screening  
126 step; Figure S1). Prior to pooling, DNA concentrations were adjusted to 10 ng/ $\mu$ l to obtain an  
127 unbiased representation of each sample in the pool. This allowed us to considerably reduce the  
128 cost and time of detection, but we acknowledge that this protocol could result in some very early  
129 stages of the disease being missed. To partly circumvent this problem, MtrBTN2 detection in  
130 pools was performed by digital PCR (ddPCR) targeting one nuclear (Elongation Factor, EF) and  
131 one mitochondrial (Control region, described in [9]) marker (Table S2). ddPCR is a sensitive  
132 method, based on the Taqman method which requires two primers and one probe, and which  
133 directly estimates the copy number of the targeted sequence. In addition, the use of a  
134 mitochondrial marker should provide a more sensitive detection as the mitochondrial genome has  
135 more copies than the nuclear genome. ddPCR primers design and analysis was subcontracted to  
136 the company IAGE using the QIAcuity™ system (QIAGEN). Positive pools were then  
137 demultiplexed (simplex screening step; Figure S1) and MtrBTN2 detection was performed by  
138 real-time PCR targeting one nuclear (Elongation Factor, EF1 $\alpha$ -i3, described in [11]) and one  
139 mitochondrial (cytochrome c oxidase I, mtCOI-sub) marker (Table S2). To design the mtCOI-  
140 sub primers, we used sequences of *M. edulis*, *M. galloprovincialis*, *M. trossulus*, and MtrBTN2  
141 available from the National Center for Biotechnology Information (NCBI). Please note that  
142 mtCOI-sub primers were initially designed to amplify a polymorphism that discriminates  
143 between two MtrBTN2 sub-lineages “a” and “b” [16]. However, attempts to identify sub-  
144 lineages based on melting temperature ( $T_m$ ) did not yield conclusive results, most likely due to  
145 additional polymorphism. Amplification of EF1 $\alpha$ -i3 and mtCOI-sub markers was performed  
146 using a three-step cycling protocol (95°C for 10 s, 58°C for 20 s, 72°C for 25 s) for 40 cycles.  
147 We carried out real-time PCRs using the SensiFAST™ SYBR No-ROX Kit (Bioline) and the  
148 LightCycler 480 Real-Time PCR System (Roche Diagnostics). We also confirmed the positive

149 samples diagnosed by real-time PCR using ddPCR (simplex screening step; Figure S1). The  
150 specificity of both methods and markers was tested using various control samples of *M. edulis*,  
151 *M. galloprovincialis*, *M. trossulus*, and MtrBTN2 samples of early, moderate, or late stages  
152 (based on cytological observations, following [32], Tableau S3). To test the sensitivity, we used  
153 positive controls with MtrBTN2 DNA diluted 10-fold (x10) and 100-fold (x100) in *M. edulis*  
154 DNA.

155  
156 Real-time PCR results were analysed using the LightCycler480 software. Two parameters were  
157 extracted: threshold cycle (CT), obtained by Absolute Quantification analysis using the Second  
158 Derivative Maximum method, and T<sub>m</sub>, obtained by T<sub>m</sub> calling analysis. Positive samples were  
159 defined as those with a CT of less than 35 cycles and a T<sub>m</sub> between 77.46 to 78.71 (based on T<sub>m</sub>  
160 values of the positive control samples; see result, Figure S2). ddPCR results were analysed with  
161 the QuantaSoft™ Analysis Pro software (Bio-Rad), which provides the copy number for each  
162 sample. Samples were considered positive when there were more than two positive droplets.

163

### 164 **2.3. Mussel genotyping**

165 We used 10 biallelic SNPs known to discriminate between *M. edulis* and *M. galloprovincialis*  
166 mussel species (Table S4). These markers were identified as being ancestry-informative (fixed  
167 for alternative alleles in *M. edulis* and *M. galloprovincialis*) in the [33] dataset and were  
168 subsequently confirmed as near diagnostic by analysis of larger datasets [22, 34]. Genotyping  
169 was performed using the Kompetitive Allele Specific PCR (KASP) method [35]. As we wanted  
170 to know the mean ancestry of each sample (G ancestry) and reduce the cost of genotyping, we  
171 developed a multiSNPs marker by multiplexing the 10 SNPs. Rapidly, 1µL of assay mix (KASP-  
172 TF V4.0 2X Master Mix, 1X target concentration; primers, 1µM target concentration;  
173 HyClone™ HyPure water) and 0.5 µL of DNA at 10 ng/µL were mixed in qPCR 384-well plates  
174 using Labcyte Echo525 (BECKMAN COULTER). KASP analysis was performed on the  
175 LightCycler 480 Instrument (Roche Diagnostics) with the following thermal cycling conditions:  
176 initialisation 15 min at 95°C, first amplification 20 s at 94°C and 1 min at 61°C to 55°C with  
177 steps of 0.6°C for 10 cycles, second amplification 20 s 94°C and 1 min at 55°C for 29 cycles,  
178 read 1 min at 37°C and 1 s at 37°C.

179 KASP results are a combination of two fluorescence values, one for allele X, and another for  
180 allele Y. Here, we oriented all SNPs so that allele X corresponds to *M. edulis* and allele Y to *M.*  
181 *galloprovincialis*. Following [36], we transformed the data to obtain a single measure of the  
182 relative fluorescence of the two alleles, using the following formula:  $y_{i,j} = Y_{i,j}/(X_{i,j} + Y_{i,j})$ ;  
183 where  $X_{i,j}$  and  $Y_{i,j}$  are the fluorescence values of allele X and allele Y, respectively for,  $i^{\text{th}}$  SNP of  
184 the  $j^{\text{th}}$  sample. To scale  $y_{i,j}$  value from 0 when *M. edulis* allele fluorescence dominates at the 10  
185 SNPs to 1 when *M. galloprovincialis* allele fluorescence dominates, we used the following  
186 formula:  $y'_{i,j} = (y_{i,j} - \min(y_i))/(\max(y_i) - \min(y_i))$ ; where  $y_{i,j}$  is the relative fluorescence  
187 value for the  $i^{\text{th}}$  SNP of the  $j^{\text{th}}$  sample. In the results section and figures,  $y'_{i,j}$  corresponds to “G  
188 ancestry”.

189

## 190 **2.4. Environmental data**

### 191 ***Source of environmental data***

192 The 76 sampling sites were selected to cover four different habitats (ports, natural beds, farms  
193 and floating buoys) presenting contrasted environments and genetic backgrounds (*M. edulis*, *M.*  
194 *galloprovincialis* and hybrids). The description of environmental data considered in this study is  
195 reported in Table S5 and all site information is available in Table S1. Note that natural beds are  
196 intertidal sites, while the other habitats are subtidal. During sampling, we collected two habitat  
197 discrete descriptive variables: population density (1: isolated, 2: patchy beds, 3: continuous beds)  
198 and wave action (1: exposed, 2: sheltered). We used the E.U. Copernicus Marine Service  
199 Information (doi: [10.48670/moi-00027](https://doi.org/10.48670/moi-00027)) to access temperature (°C), salinity (PSU 1e-3) and  
200 surface current velocity (m.s-1) data for each site (monthly recorded from 2021, resolution of  
201 0.028°x0.028°). The velocity variable was formed by combining its Northward (N) and Eastward  
202 (E) components, as follow:  $\sqrt{E^2 + N^2}$ . We also used EMODnet Human Activities (source:  
203 EMSA Route Density Map) to retrieve the maritime traffic density of fishing, passenger, and  
204 recreational boats in the study area (seasonal record from 2018, 2019 and 2020, resolution 1x1  
205 km).

206

### 207 ***Preparation of environmental data***

208 Among the 76 sampled sites, 13 fell outside the coverage scope of the Copernicus model (i.e.  
209 situated beyond the model’s pixel boundaries). Consequently, for these sites (see Table S1), we



210 were unable to extract temperature, salinity, or velocity data. Certain sites recorded a velocity of  
211 0 for all month (Table S1), raising concerns about the model resolution for this variable. To  
212 validate the accuracy of Copernicus data, we computed the mean values of neighbouring pixels  
213 (MNP, within a 0.05° buffer zone) and compared them to the pixel values at the respective sites  
214 of interest (SP, referred to as the 'central' pixel, Table S6). We note that, due to the proximity of  
215 our sampled sites to the coastline, the number of neighbouring pixels varied within a certain  
216 range (Table S6). Notably, there was no correlation between these variations and the absolute  
217 difference between SP and MNP values (Figure S3C). The comparison between SP and MNP  
218 values was carried out for all sites, encompassing temperature, salinity, and surface current  
219 velocity, and the results are presented in Figure S3. We observed a robust positive correlation for  
220 temperature and salinity variables, providing confidence in the reliability of the extracted data for  
221 our small-scale analysis. However, as expected, the inconclusive correlation for surface current  
222 velocity led us to exclude this environmental variable from the analyses.

223 In the statistical analysis (see section 2.5) we used the annual average of salinity and boat  
224 density (maritime traffic). For temperature, we used the annual variance as we identified strong  
225 seasonal variations in temperature that would be misinterpreted if only the annual average  
226 temperature were used (Figure S4).

227

## 228 **2.5. Statistical analysis: relationship between MtrBTN2 prevalence, host genotype and** 229 **environmental variables**

230

231 To ascertain whether cancer prevalence varied among the four habitats sampled (ports, natural  
232 beds, buoys and farms), we performed Fisher's exact test on the contingency table (*fisher.test()*  
233 function of the *stats* package *v4.1.3* [37]). Subsequently, we aimed to determine whether specific  
234 environmental variables could provide a more comprehensive explanation of MtrBTN2  
235 prevalence across these habitats. To do so, we pursued the following steps: an exploratory  
236 analysis with a Principal Component Analysis (PCA), an examination of data dependencies  
237 (Spearman correlation test and Moran's I test), and lastly, a regression model adapted to our  
238 dataset (spatial logistic regression model). All statistical analyses were performed in R *v4.1.3*  
239 [37].

240

241 ***Principal component analysis***

242 As an exploratory analysis, we firstly performed a PCA to investigate the relationship between  
243 the quantitative environmental variables. Missing values were estimated using *imputePCA()*  
244 (*missMDA* package v1.18 [38]), prior to performing the PCA using *PCA()* (*FactoMineR* package  
245 v2.6 [39]).

246

247 ***Spearman correlation test***

248 To test for correlation between environmental variables, we performed a non-parametric test of  
249 Spearman using *rcorr()* function from *Hmisc* package [40], that automatically excludes missing  
250 data.

251

252 ***Moran I test for spatial autocorrelation***

253 Given the proximity of the sampled sites, which constitutes a fine spatial matrix, we used the  
254 Moran's I statistic to test for spatial autocorrelation among the environmental variables. After  
255 converting longitude and latitude data into UTM coordinates (*spTransform()* function of *rgdal*  
256 package v1.6-7 [41]), this analysis was performed using the *moran.test()* function (*spdep*  
257 package version 1.2-8 [42]) with the k-nearest neighbour classification.

258

259 ***Spatial logistic regression model***

260 The association of MtrBTN2 prevalence and environmental variables was tested at the site level  
261 using a multivariate spatial logistic regression (*fitme()* function, *spaMM* package v4.3.0 [43]).  
262 Spatial autocorrelation was accounted for by including a Matern covariance function as a random  
263 effect (i.e. *Matern(1 | Longitude +Latitude)*). MtrBTN2 prevalence was represented as a binary  
264 matrix ( $N, N_{site}-N$ ) with  $N$  and  $N_{site}$  indicating site-specific MtrBTN2 case counts and total  
265 number of sampled mussels, respectively. The following explanatory variables were included  
266 into the full model: temperature variance, salinity mean, population abundance, wave action,  
267 average *M. galloprovincialis* ancestry of the mussel host population (G ancestry), maritime  
268 traffic (as described above, section 2.4). We also included a binary variable "port" into the  
269 model, where 1 denotes port habitat and 0 represents other habitats. Model suitability was  
270 evaluated using the residual diagnostics tool implemented in the "DHARMA" package [44]. To  
271 assess the significance of fixed effects (environmental variables), we performed Bartlett-

272 corrected likelihood ratio tests (LR tests) by comparing the full model with a model that  
273 excluded the variable of interest (*fixedLRT()* function of *spaMM* package *v4.3.0* [45]). To  
274 account for bias associated with rare events and collinearity of environmental variables (see  
275 results), LR tests were performed using 999 bootstraps.

276

### 277 **3. Results**

#### 278 **3.1. Validation of the screening method**

279 In order to detect MtrBTN2 in thousands of individuals, we developed dedicated ddPCR and  
280 qPCR screening tools. The design of real-time PCR and ddPCR primers and probes allowed us to  
281 amplify nuclear and mitochondrial markers specific to *M. trossulus* mussels and MtrBTN2  
282 (Figure S5, TableS7). However, for the mtCOI-sub locus, the melting temperature ( $T_m$ ) range of  
283 MtrBTN2 was different from *M. trossulus* controls and allowed us to specifically detect  
284 MtrBTN2 based on  $T_m$  values (Figure S2). With ddPCR, all MtrBTN2 control samples were  
285 detected, regardless of marker or dilution, except for one MtrBTN2 sample diluted at x100 which  
286 was negative for the EF locus (Figure S5). Copy numbers appeared higher for the mtCR marker,  
287 showing a higher sensitivity for this mitochondrial marker. The same pattern was observed for  
288 EF1 $\alpha$ -i3 and mtCOI-sub in real-time PCR, with a higher sensitivity of the mitochondrial marker  
289 based on the lower CT values observed (Figure S5). However, two MtrBTN2 samples diluted at  
290 x100 were negative for mtCOI-sub and positive for EF1 $\alpha$ -i3. Even though the mitochondrial  
291 marker appears to be more sensitive in both methods, we chose to screen for MtrBTN2 using  
292 both markers in order to have comparisons and additional validation.

293

#### 294 **3.2. MtrBTN2 screening reveals a low prevalence**

295 We amplified nuclear and mitochondrial markers using ddPCR (EF, mtCR) and real-time PCR  
296 (EF1 $\alpha$ -i3, mtCOI-sub) to detect the presence of MtrBTN2 in the 1516 mussels sampled from the  
297 76 sites of the study area (Figure S1). All results of the pooled and simplex screening are  
298 presented in TableS8.

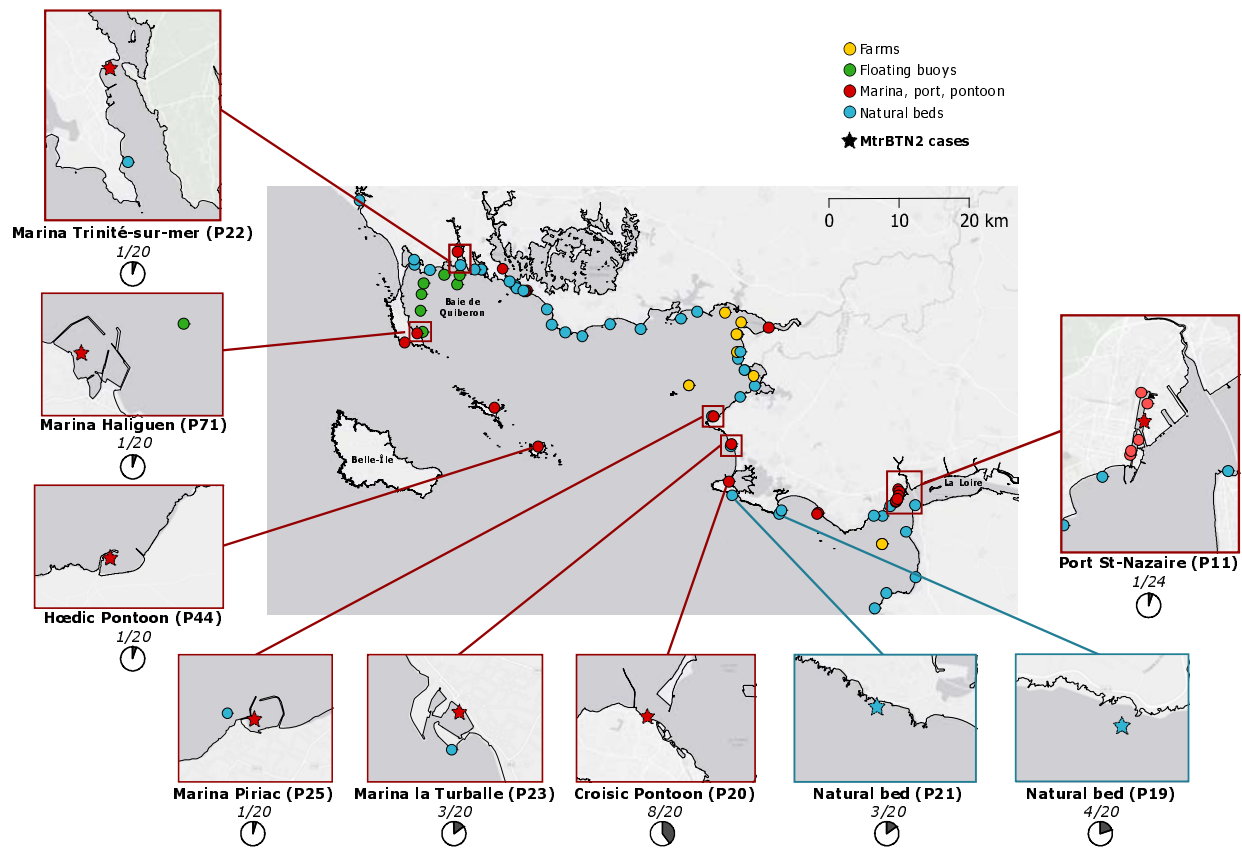
299 The pooled screening step using ddPCR revealed 25/127 positive pools: 9/127 with EF, 14/127  
300 with mtCR, and 2/127 were positive for both markers (Figure S5). As the 14 mtCR-positive  
301 pools showed low copy numbers (median = 7.59), we assume that EF did not amplify because of  
302 a detection threshold. Although the 9 EF-positive pools were not expected, they were included

303 for further investigation by simplex screening (see below). For pools positive for both markers,  
304 the EF copy numbers appear relatively low compared to mtCR copy numbers. This is due to the  
305 high mitochondrial fluorescence masking the nuclear fluorescence and underestimating EF copy  
306 numbers. To eliminate this bias in the simplex ddPCR step (Figure S1), we amplified the two  
307 markers separately. In order to maximise cancer detection in the simplex screening step, we  
308 chose to keep the pools positive for at least one gene, corresponding to 25 positive pools of 12  
309 samples ( $n = 300$ ).

310 We performed the simplex screening step using real-time PCR and then confirmed positive  
311 samples with ddPCR (Figure S1). In total, 44/300 samples were positive for at least one real-time  
312 PCR marker: 16 with EF1 $\alpha$ -i3 only, 23 with mtCOI-sub only, and 5 with both markers (Figure  
313 S5, Table S9). Pooled and simplex screening results are mostly consistent, as 20 of the 25  
314 positive pools had at least one individual detected by real-time PCR amplification (Table S10).  
315 The 5 pools without positive samples had low ddPCR copy numbers (less than 3.37 for EF and  
316 less than 6.65 for mtCR). These false positives were to be expected as we intentionally applied  
317 low threshold values in the pooled screening step. Two samples from mtCR-positive pools were  
318 found to be positive for both markers at the simplex level, which was expected as the  
319 mitochondrial marker is likely to be more sensitive than the nuclear marker. On the contrary, we  
320 did not expect to observe samples positive for EF1 $\alpha$ -i3/EF and negative for mtCOI-sub/mtCR.  
321 However, all but one of the individuals from the EF-positive pools were positive only for the  
322 EF1 $\alpha$ -i3 marker (Table S10). The absence of mitochondrial genes in these samples does not  
323 seem to be related to a detection threshold issue, especially as ddPCR and real-time PCR target  
324 two different genes (mtCR and mtCOI-sub). We suspect that some host alleles that were present  
325 at low frequencies in the ‘dock mussels’ populations (the hybrid lineage found in the port of  
326 Saint-Nazaire [34]) may sometimes be amplified with EF1 $\alpha$ -i3. This could be explained by  
327 shared polymorphism between host species rather than the presence of MtrBTN2. Therefore, the  
328 16 samples positive only for EF1 $\alpha$ -i3 were conservatively considered as false positives (note  
329 they were mostly sampled in the port of Saint-Nazaire, which does not affect our main result).  
330 For the remaining 28 samples, we performed ddPCR with both markers to confirm the real-time  
331 mitochondrial PCR diagnosis (Figure S1, TableS9). For 23/28 samples, mtCR ddPCR confirmed  
332 the mtCOI-sub real-time PCR diagnosis. For 10/23 samples positive only for mtCOI-sub, ddPCR  
333 revealed few EF copies. Finally, we considered all samples positive for mitochondrial markers

334 with both methods to be cancerous samples. We found 23/1516 positive samples from 9/76 sites  
335 in which the number of cases varied from 1 to 8 (Figure 1). The highest proportion of cancerous  
336 mussels was found at site P20 (Croisic Pontoon), and the next highest in the vicinity of this site  
337 (P19, P21, and P23; Figure 1).

338



339

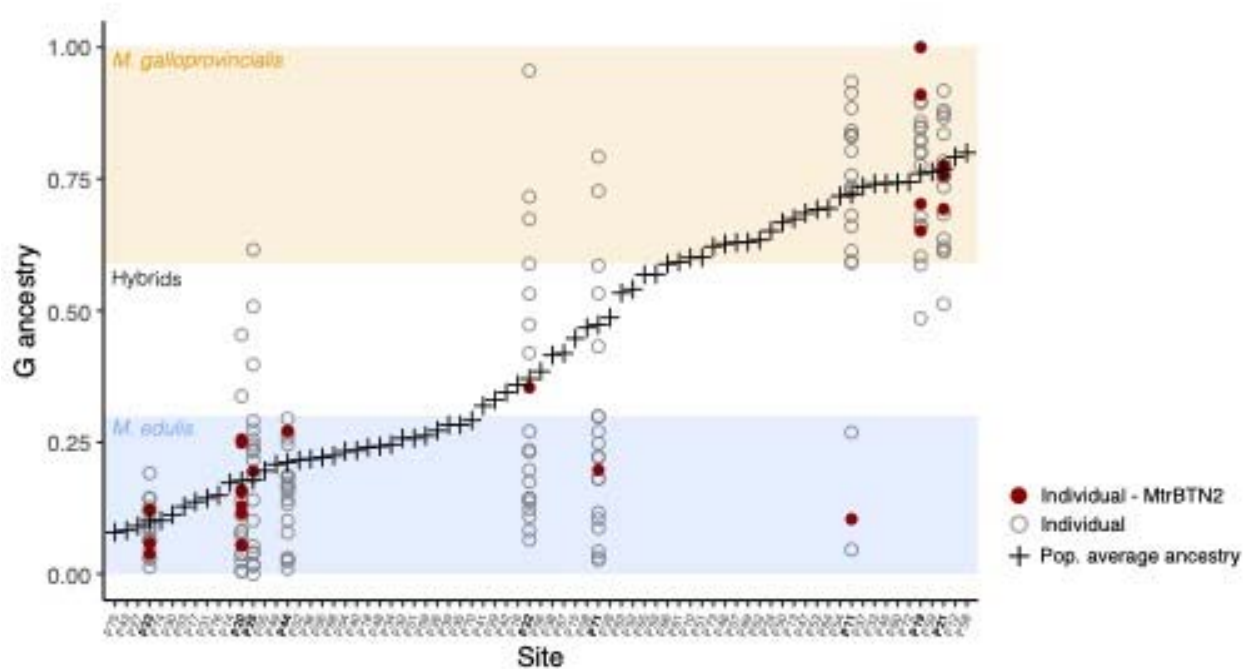
340 **Figure 1: Location of sampling sites.** Zoomed-in sites correspond to sites where mussels are  
341 affected by MtrBTN2. Pie graphs and numbers in italics correspond to the proportion of  
342 cancerous samples in each site. Coloured points represent the type of site, as indicated in the  
343 figure's legend. Map origin: Ersi gray (light).

344

345

346 **3.3. *M. edulis* and *M. galloprovincialis* are equally affected by MtrBTN2**

347 To determine whether host genetic background or population composition influences cancer  
348 prevalence and distribution, mussels were genotyped using 10 SNPs known to be diagnostic  
349 between *M. edulis* and *M. galloprovincialis* (the two species coexisting in the studied area).  
350 To reduce the cost of genotyping, we developed a multiplex assay, mixing 10 SNPs that directly  
351 give the average ancestry of the sample (G ancestry, Figure S1). This approach was validated by  
352 comparing the multiplex fluorescence and average simplex fluorescence of SNPs from a set of  
353 samples (Figure S6a). We genotyped each individual from the 9 sites affected by MtrBTN2 (180  
354 mussels) and estimated the average ancestry of each site by genotyping pools of individuals (76  
355 pools). The positive correlation between the average individual fluorescence of a site and the  
356 fluorescence of the corresponding pool validated our approach (Figure S6b). Mussels with G  
357 ancestry values below 0.3 were assigned to *M. edulis*, those with values above 0.59 to *M.*  
358 *galloprovincialis*, and those in between were assigned as hybrids (Figure 2, FigureS7). Of the  
359 mussels affected by MtrBTN2, 15/23 were assigned to *M. edulis*, 7/23 to *M. galloprovincialis*  
360 and 1/23 was a hybrid (Table S11). Considering the sites affected by MtrBTN2, MtrBTN2  
361 prevalence did not appear to be significantly different between species with 15/105 (14%) in *M.*  
362 *edulis*, 7/60 (11%) in *M. galloprovincialis* and 1/15 (6%) in hybrids (Fisher's exact test, p-value  
363 = 0.83). Most *M. edulis* mussels affected by MtrBTN2 were found in populations with a majority  
364 of *M. edulis* (P44, P23, P20), only two were found in populations with some hybrids and *M.*  
365 *galloprovincialis* mussels (P25, P71) and only one in a population with a majority of *M.*  
366 *galloprovincialis* (P11, Figure 2, Figure S7). Interestingly, all *M. galloprovincialis* mussels  
367 affected by MtrBTN2 were found in *M. galloprovincialis* populations (P19, P21, Figure 2,  
368 FigureS7), which are close to the site with the highest prevalence (Croisic pontoon, Figure 1).  
369 Finally, the hybrid individual affected by MtrBTN2 was found in a population where the two  
370 species coexist (P22, Figure 2, FigureS7).  
371



372

373 **Figure 2: Individual and population ancestry.** G ancestry of healthy (empty grey circle) and  
374 MtrBTN2 affected (dark red) hosts is shown for the 9 sites with cancer (bold labels). Population  
375 average ancestry for all sampled sites is represented by black crosses. Blue, white and orange  
376 rectangles correspond to the G ancestry range of *M. edulis*, hybrids and *M. galloprovincialis*  
377 respectively.

378

#### 379 **3.4. *MtrBTN2* infected mussels are mostly found in ports**

380 After the MtrBTN2 detection step, which revealed 23/1516 individuals infected with this  
381 transmissible cancer, we investigated associations with habitat types. Cancerous samples were  
382 found in 7 ports (16/23 infected mussels) and 2 natural beds (7/23) near the Croisic pontoon  
383 (P20, Figure 1). MtrBTN2 was not detected in the 180 individuals from the 9 mussel farms or in  
384 the 140 individuals from the 7 floating buoys. Mussels affected by MtrBTN2 were significantly  
385 more frequent in ports than in other habitats (Fisher's exact test, p-value = 0.0001, Table S12). In  
386 addition, the sampled ports are not significantly spatially autocorrelated (Moran I test, p-value =  
387 0.25, Table S13).

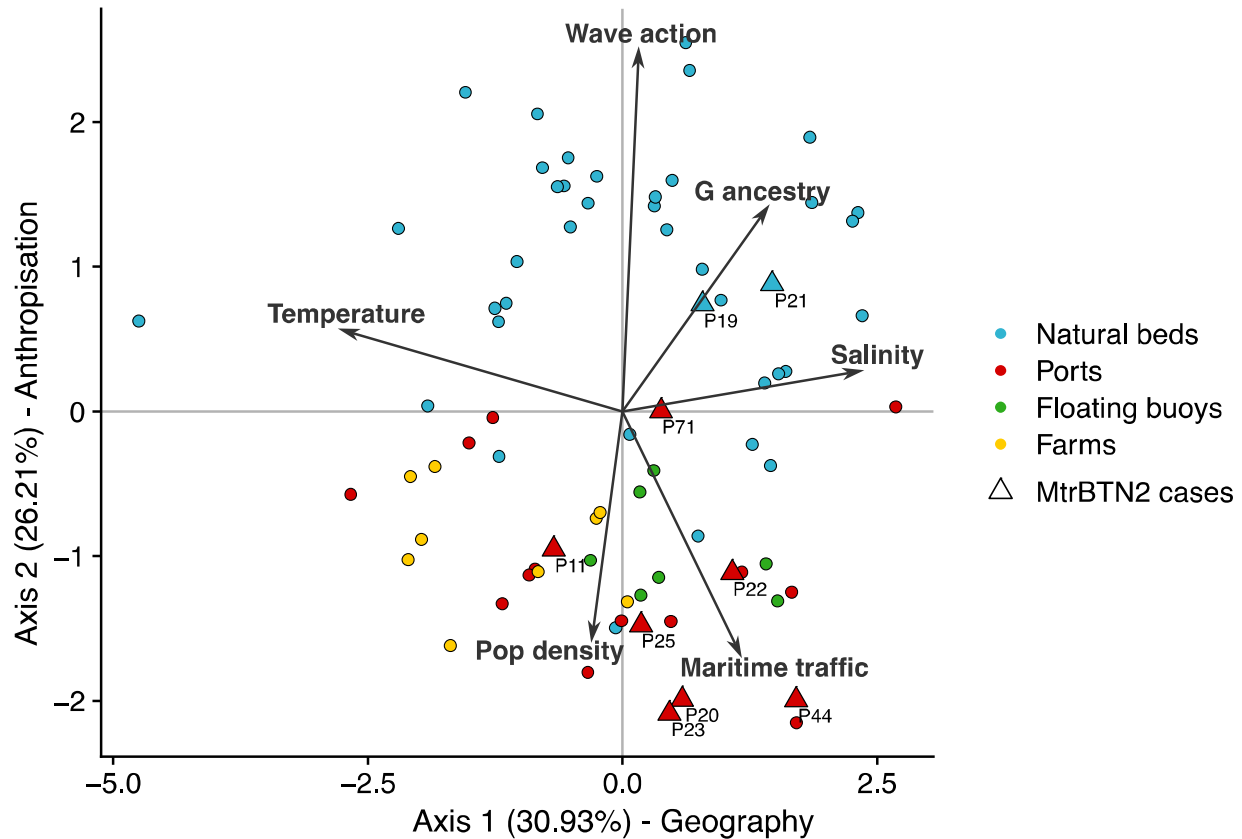
388

389 **3.5. Exploring the effect of some environmental variables, none other than the port habitat**  
390 **explains *MtrBTN2* prevalence after accounting for spatial autocorrelation**

391 Given that MtrBTN2 was mostly found in ports, we aimed to gain a deeper understanding of  
392 whether environmental variables (including population ancestry) could explain this pattern.  
393 Univariate correlation between MtrBTN2 prevalence and all environmental variables across  
394 habitats are presented in Figure S8. The two best correlations are found with population density  
395 and maritime traffic, but we also have strong multicollinearity among variables that needs to be  
396 accounted for (see Table S14). The PCA reveals an association between environmental variables  
397 and habitat types rather than the presence of cancer itself (Figure 3). Indeed, anthropogenic  
398 habitats (ports, floating buoys and farms) with high population density and sheltered from wave  
399 action are well separated from natural beds on the second axis of the PCA. Interestingly,  
400 maritime traffic (i.e., annual mean of passenger, fishing, and recreational boat density) explains a  
401 direction of variation towards human-altered habitats, but also towards sites affected by  
402 MtrBTN2. MtrBTN2 seems to be found in sites with buffered temperature (low annual variance)  
403 and higher mean annual salinity, environmental variables that are associated with ports in our  
404 study. Water temperature indeed tends to be buffered in ports, being warmer in winter and cooler  
405 in summer (Figure S4). However, these variables are difficult to interpret as they also explain  
406 well the inertia of axis 1 of the PCA which differentiates the sites located near estuaries (La  
407 Loire or La Vilaine) from the others. To investigate more rigorously whether some  
408 environmental variables could explain prevalence of MtrBTN2, we used a spatial logistic  
409 regression model.

410





411

412 **Figure 3: Principal Component Analysis on environmental variables.** Axis 1 mainly  
413 captures the variance associated with annual temperature variance and annual salinity mean and  
414 Axis 2 mainly captures the variance associated with wave action, population density and  
415 maritime traffic.

416

417 As MtrBTN2 was mostly found in ports (see section 3.4), we included an environmental variable  
418 “port” (1 for port habitat, 0 for others) in the regression model analysis. To test for spatial  
419 structure, we conducted Moran’s I tests on all environmental variables. The results indicate  
420 significant spatial autocorrelation in most environmental variables (Table S13), which was  
421 expected as we selected this area due to its mosaic of contrasted environments. However,  
422 “population density” and “port” variables were found randomly distributed in space (Table S13).  
423 To overcome spatial autocorrelation bias we performed a logistic regression model that corrects  
424 for spatial effect (*spaMM* package, [43]) and assessed environmental variables significance via  
425 bootstrapped Likelihood Ratio (LR) tests. Bootstrapped LRTs are expected to handle the  
426 problem of variables multicollinearity. MtrBTN2 prevalence is significantly explained by the

427 port habitat variable (LR test:  $\chi^2 = 7.21$ ,  $df = 1$ ,  $p\text{-value} = 0.0072$ ; Table S15), and none of  
428 the other environmental variables significantly explain MtrBTN2 prevalence after accounting for  
429 spatial effects (Table S15). Despite the residuals following the distributional assumptions of the  
430 model (Figure S9) and the significance of the ‘Port’ variable (also supported by the Fisher’s  
431 exact test, see 3.4), we acknowledge that quantifying the magnitude of this trend remains  
432 challenging and requires further investigations.

433

#### 434 **4. Discussion**

435

436 Our results show that MtrBTN2 is present at low prevalence along the French Atlantic coast  
437 ( $23/1516 = 1.5\%$ ). Strikingly, it is mostly found in ports and is equally distributed in *M. edulis*  
438 and *M. galloprovincialis*. Our analysis shows that ports are likely epidemiological hubs for  
439 MtrBTN2 propagation. However, the precise environmental factors influencing MtrBTN2  
440 epidemiology have yet to be fully characterised. In this context, we discuss various  
441 environmental variables and assess their relevance to the port environment and marine parasite  
442 epidemiology.

443

444 Most MtrBTN2-infected mussels were found in ports and a prevalence hotspot was found around  
445 the Croisic pontoon. Given the low prevalence of MtrBTN2 previously described [16, 32] and  
446 confirmed in this study, we acknowledge that our sample sizes per site may be too small ( $n = 20$ )  
447 to assess MtrBTN2 prevalence with confidence at a site level. However, this was not the scale of  
448 resolution we were interested in. Instead, our overall sampling design, capturing replications of  
449 habitat features, was chosen to evaluate the effect of habitats and genetic composition of host  
450 populations. This allowed us to reveal similar MtrBTN2 prevalence in both species of the mosaic  
451 hybrid zone, and a higher prevalence in ports.

452

453 The highest prevalence of MtrBTN2 in ports could be explained by two non-exclusive  
454 hypotheses. Firstly, ports could offer a favourable environment as it is conducive to disease  
455 development and propagation. Indeed, ports are often polluted, confined and permanently  
456 immersed habitats (without tidal constraints) with high density of mussels. It is important to note  
457 here that our objective is not to explain the emergence of a transmissible cancer, i.e. the initial

458 carcinogenesis in the founder host, but its maintenance and spread. The emergence probably took  
459 place long before the industrial period and, despite studying a cancer, the mutagenic effect of  
460 pollutants does not have to be put forward to explain our results, but rather their effect on host  
461 resistance to infection by transmissible cancer cells. Secondly, there is increased connectivity  
462 between ports due to maritime traffic and ports are connected to other sites (other ports or wild  
463 sites) through biofouling from vessels.

464

465 Ports are often confined areas with longer particle residence times which could increase the  
466 contact time between parasites and hosts [46, 47], thereby increasing the probability of  
467 transmission and favouring the persistence of a passively dispersing parasite. However, the  
468 possible hotspot on Croisic Pontoon, which is affected by strong tidal currents, provides  
469 evidence against this hypothesis given that it is not a confined area.

470

471 Host population density is an important epidemiological parameter, as high density is known to  
472 favour parasite persistence in host populations [48, 49]. Our results suggest that ports are prone  
473 to harbouring high population density (Figure 3). However, we found no evidence of MtrBTN2  
474 cases in mussel farms or floating buoys despite high mussel population density in these sites as  
475 well. Although population density could favour MtrBTN2 persistence in an affected population,  
476 this parameter alone cannot explain the higher prevalence of MtrBTN2 in ports.

477

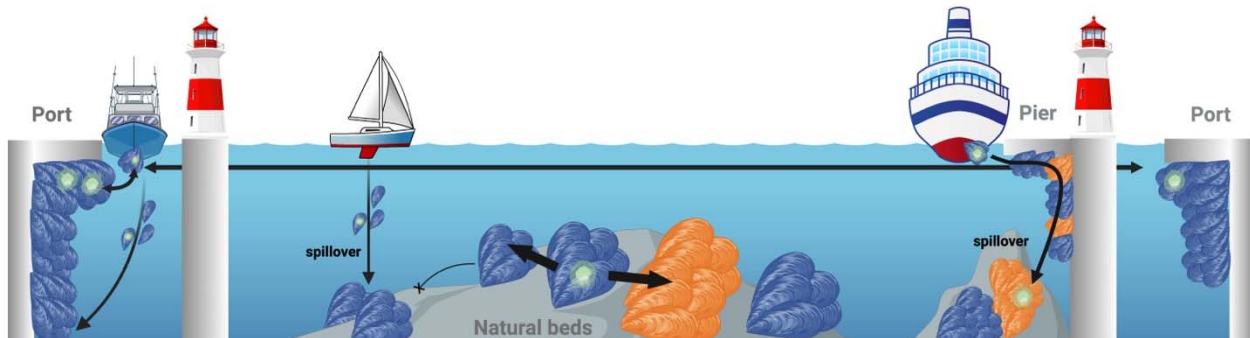
478 Port-based shipping activities are major contributors to water pollution, including issues such as  
479 oil spills, plastic contamination, wastewater discharge, and the use of antifouling paint [50, 51].  
480 Pollution is a major environmental stressor that could affect host physiology and increase  
481 susceptibility to parasites, particularly by altering its immune response (reviewed in [52]). As  
482 filter-feeders, species of bivalve molluscs are particularly concerned [53–55]. However, in our  
483 study, most of the affected individuals (15/23) were found in three open sea sites known to be  
484 healthy and protected ecosystems with abundant resources (Croisic Pontoon, Natural beds 19 and  
485 21 part of Natura 2000 area; Figure 1), suggesting that pollution alone cannot explain the higher  
486 prevalence of MtrBTN2 cancers in ports. Conversely, it can be hypothesised that a healthy, less  
487 stressful environment may increase the host permissiveness and allow MtrBTN2 (and possibly  
488 other BTN) to persist at a higher detectable prevalence. Host nutrition can affect disease

489 outcome by driving host immunity and parasite resource availability [56–58]. In invertebrates,  
490 increased host nutrition can lead to increased parasite load [59] and improved immunity against  
491 parasites [57, 58]. This balance could apply to mussel-MtrBTN2 interaction, and a healthy  
492 environment with abundant resources could increase host carrying capacity and/or slow down  
493 disease progression. A polluted environment could also affect the survival of MtrBTN2 cells in  
494 seawater and reduce their chances of infecting a new host. However, the impact of pollution on  
495 mussel hosts and cancer cells needs to be further investigated to clearly establish its influence on  
496 the transmission and persistence of such transmissible cancers.

497

498 Connectivity through maritime traffic appears to be the main common feature between affected  
499 ports, although this explanatory variable does not withstand spatial logistic regression analysis  
500 any better than others. Vessels could facilitate MtrBTN2 propagation through transport of  
501 affected mussels fixed on the hull. Maritime traffic has been identified as a vector of marine  
502 mollusk pathogens (for review see [60]), microorganisms [61] and non-indigenous species [62,  
503 63], through ballast water and/or hull biofouling. As such, ports are considered hotspots of  
504 marine non-indigenous species introductions [62, 64, 65] and of marine parasites [66]. *Mytilus*  
505 mussels are commonly reported as principal taxa composing hull biofouling (fixed adult mussels;  
506 [67–69]). If a port is affected by a local outbreak, it could become a source of contamination to  
507 other ports through biofouling of vessels, while natural beds will probably remain isolated by  
508 density troughs that could halt the propagation (Figure 4). It is also conceivable that prevalence  
509 hotspots exist at other unsampled locations along the coast, either due to environmental factors  
510 that have yet to be characterised, or simply due to stochastic local outbreaks. These local  
511 outbreaks can be the source of contamination to the anthropogenic connectivity network  
512 providing the disease reaches a port. While this does not disqualify the factors discussed above –  
513 e.g., pollution, population density and confinement- as drivers of contamination within ports, our  
514 study suggests that maritime routes between ports could be anthropogenic gateways to MtrBTN2  
515 propagation. Furthermore, despite a high density of maritime traffic around the floating buoys  
516 sampled, these sites remained free of MtrBTN2. Long mooring times in ports increase the  
517 probability of hull surface colonisation [69–71] as well as the risk of pathogen transmission [60].  
518 This suggests that, in addition to cruising, ships may have to dock for the spread of MtrBTN2 to  
519 be effective, either through transmission of cancer cells or detachment of infected mussels

520 (Figure 4). This could provide an explanation as to why maritime traffic is not a sufficient  
521 explanatory variable for MtrBTN2 prevalence. Further investigations would be necessary to  
522 ascertain the requisite propagule density (comprising both numerical and time-scale components)  
523 for effective transmission, as well as to evaluate possible disruption of byssus quality in affected  
524 hosts.  
525



526  
527 **Figure 4: Diagram illustrating the discussion on the propagation of MtrBTN2 across**  
528 ***Mytilus sp.* populations.** Arrows' widths correspond to the probability of MtrBTN2 propagation.  
529 Blue and orange mussels correspond to *M. edulis* and *M. galloprovincialis*, respectively.  
530 Diagram created with BioRender.

531  
532 A possible hotspot of MtrBTN2 prevalence was found in the area around the Croisic pontoon.  
533 Even though this possible hotspot will need to be confirmed, this site could represent a source for  
534 other ports and surrounding natural beds. Indeed, this pontoon is frequently used by maritime  
535 shuttles, as well as by recreational and fishing boats, and is connected by maritime traffic to  
536 many ports in the studied area. In particular, this pontoon is the departure site of boats travelling  
537 to Hoedic (site P44 in Figure 1), an isolated island for which the presence of cancer is most  
538 likely explained by anthropogenic transport given the distance from the mainland. MtrBTN2 is  
539 also found in foreshores near Croisic. This reveals a possible spillover from ports to adjacent  
540 wild populations (Figure 4). Maritime traffic has been identified as a vector for non-indigenous  
541 species from commercial ports to natural populations via marinas [72, 73]. In our study, when  
542 sites outside standard enclosed ports (e.g., enclosed marinas) were sampled, MtrBTN2 was never  
543 detected, suggesting that contagion often remains confined to the port area. Spillover may  
544 therefore depend on the degree of confinement of a port.

545

546 *Mytilus galloprovincialis* was thought to be more resistant to MtrBTN2 because of lower  
547 prevalence in this species [16], and population patches of *M. galloprovincialis* could also have  
548 acted as barriers to disease propagation between patches of *M. edulis*. However, our results  
549 suggest that *M. galloprovincialis* population patches close to *M. edulis* populations may also  
550 show enhanced prevalence. This does not contradict the idea that *M. galloprovincialis* population  
551 patches could impose a slightly higher resistance to MtrBTN2 propagation. Under the hypothesis  
552 that natural propagation is hindered by density troughs between population patches, or by  
553 population patches of a more resistant species, ports and ship traffic would create anthropogenic  
554 gateways that favour connections between patches of host populations (Figure 4). This  
555 hypothesis could also suggest that BTN2 spread naturally and progressively from one host to  
556 another by close contact rather than drifting with the currents, making enhanced R0 by  
557 anthropogenic dispersal a threat to host populations.

558

559 Our results represent a significant advance in our understanding of MtrBTN2 epidemiology, both  
560 by proposing ports as epidemiological hubs and by suggesting that the natural propagation of this  
561 transmissible cancer could be naturally halted by density troughs in the distribution of host  
562 mussel populations. Detection of MtrBTN2 presence on boat biofouling would provide more  
563 direct evidence of its role as a vector. It would also be interesting to investigate whether  
564 prevalence hotspots are more commonly located in areas with particular environmental  
565 conditions, or whether they are stochastic local outbreaks. Sampling vessel biofouling in  
566 international ports linking known MtrBTN2 sites (East and North Europe, North West and South  
567 Pacific) and nearby marinas, along with genetic comparison of cancer cells, would provide a  
568 better understanding of the global and local transmission routes of this transmissible cancer since  
569 its emergence. While MtrBTN2 is not associated with extensive mortality events [30], other  
570 pathogens could be translocated and affect wild and farmed mussel populations. Various  
571 measures are employed to mitigate biofouling, including antifouling painting, dry docking, in-  
572 water cleaning with capture, and ballast water filtration (as discussed in [60]). While biofouling  
573 management for international vessel exchanges is established [74–76], regulations for biofouling  
574 control in local recreational boats remain to be improved [67, 77]. This is a significant concern,  
575 as the involvement of recreational boats in the spread of invasive species and possible associated

576 pathogens is increasingly documented [78–80]. Prioritising the survey of vessels that stay in port  
577 for long periods [81] or identifying nodes of busy maritime route [72] could contribute  
578 considerably to local and international biofouling management.

579

## 580 **5. Conclusion**

581

582 Our study reveals that the prevalence of MtrBTN2 along the South Brittany French coast is low,  
583 but that the most affected sites are ports (marinas, commercial ports, and docks). This result  
584 suggests that ports are epidemiological hubs and that maritime routes are anthropogenic  
585 gateways for the propagation of MtrBTN2. Investigations into the presence of other BTNs in  
586 boat biofouling would provide more direct evidence of its role as vector, although mussels are  
587 the most common species of bivalve mollusc in biofouling. The spread of marine  
588 (micro)parasites and pathogens through vessel biofouling is a legitimate concern, and results  
589 such as ours -i.e., on a smaller scale than translocations identified in the literature on marine  
590 bioinvasion or by [9]- highlight the critical need for policy regulation to limit the effects of  
591 biofouling, both on ship hulls and port docks, in addition to ballast water control.

592

## 593 **Data accessibility**

594 All data used in this study are available in the supplementary material.

## 595 **Author's contribution**

596 M.H., Conceptualization, Data curation, Formal analysis, Investigation, Methodology,  
597 Supervision, Visualization, Writing – original draft, Writing – review and editing. F.To.,  
598 Investigation, Data curation, Writing – review and editing. E.A.V.B., Formal analysis,  
599 Investigation, Writing – review and editing. L.P., Data curation, Investigation, Writing – review  
600 and editing. F.C., Investigation, Writing – review and editing. E.G., Investigation, Writing –  
601 review and editing. I.B., Resources, Writing – review and editing. H.C., Resources, Writing –  
602 review and editing. A.S., Formal analysis, Writing – review and editing. F.Th.,  
603 Conceptualization, Funding acquisition, Writing – review and editing. D.D-G.,  
604 Conceptualization, Funding acquisition, Writing – review and editing. G.M.C.,  
605 Conceptualization, Project administration, Funding acquisition, Writing – review and editing.

606 N.B., Conceptualization, Funding acquisition, Formal analysis, Investigation, Methodology,  
607 Project administration, Supervision, Writing – original draft, Writing – review and editing.

608

### 609 **Conflict of interest declaration**

610 We declare we have no competing interests.

611

### 612 **Fundings**

613 This work was supported by Montpellier Université d'Excellence (BLUECANCER project) and  
614 Agence Nationale de la Recherche (TRANSCAN project, ANR-18-CE35-0009 and DockEvol  
615 project, ANR-23-CE02-0020-01). This study falls within the framework of the “Laboratoires  
616 d'Excellence (LABEX)” CEMEB (10-LABX-0004) and Tulip (ANR-10-LABX-41). FT was  
617 supported by the EVOSEXCAN project (ANR-23-CE13-0007), the MAVA Foundation and the  
618 HOFFMANN Family.

619

### 620 **Acknowledgement**

621 We are grateful to the IAGE company for carrying out the ddPCR assay of this study. We thank  
622 the GENSEQ and QPCR HAUT-DEBIT platforms for access to equipments and their expertise.  
623 Credit is owed to Olivier Gimenez, François Rousset and an anonymous reviewer for their  
624 valuable advice on statistical analysis. Finally, we thank Cécile Perrin for her valuable reading of  
625 the manuscript.

626

### 627 **References**

628

- 629 [1] Murgia C, Pritchard JK, Kim SY, et al. Clonal Origin and Evolution of a Transmissible  
630 Cancer. *Cell* 2006; 126: 477–487.
- 631 [2] Rebbeck CA, Thomas R, Breen M, et al. Origins and evolution of a transmissible cancer.  
632 *Evolution (N Y)* 2009; 63: 2340–2349.
- 633 [3] Pearse AM, Swift K. Allograft theory: Transmission of devil facial-tumour disease. *Nature*  
634 2006; 439: 549.
- 635 [4] Pye RJ, Pemberton D, Tovar C, et al. A second transmissible cancer in Tasmanian devils.  
636 *Proceedings of the National Academy of Sciences* 2016; 113: 374–379.



- 637 [5] Garcia-Souto D, Bruzos AL, Diaz S, et al. Mitochondrial genome sequencing of marine  
638 leukaemias reveals cancer contagion between clam species in the Seas of Southern  
639 Europe. *Elife*; 11. Epub ahead of print 1 January 2022. DOI: 10.7554/ELIFE.66946.
- 640 [6] Metzger MJ, Reinisch C, Sherry J, et al. Horizontal transmission of clonal cancer cells  
641 causes leukemia in soft-shell clams. *Cell* 2015; 161: 255–263.
- 642 [7] Metzger MJ, Villalba A, Carballal MJ, et al. Widespread transmission of independent  
643 cancer lineages within multiple bivalve species. *Nature* 2016; 534: 705–709.
- 644 [8] Michnowska A, Hart SFM, Smolarz K, et al. Horizontal transmission of disseminated  
645 neoplasia in the widespread clam *Macoma balthica* from the Southern Baltic Sea. *Mol*  
646 *Ecol* 2022; 31: 3128–3136.
- 647 [9] Yonemitsu MA, Giersch RM, Polo-Prieto M, et al. A single clonal lineage of transmissible  
648 cancer identified in two marine mussel species in South America and Europe. *Elife*; 8.  
649 Epub ahead of print 1 November 2019. DOI: 10.7554/eLife.47788.
- 650 [10] Yonemitsu MA, Sevigny JK, Vandepas LE, et al. Multiple lineages of transmissible  
651 neoplasia in the basket cockle (*Clinocardium nuttallii*) with repeated horizontal transfer of  
652 mitochondrial DNA. *bioRxiv* 2023; 2023.10.11.561945.
- 653 [11] Burioli EAV, Hammel M, Bierne N, et al. Traits of a mussel transmissible cancer  
654 are reminiscent of a parasitic life style. *Scientific Reports 2021 11:1* 2021; 11: 1–11.
- 655 [12] Giersch RM, Hart SFM, Reddy SG, et al. Survival and Detection of Bivalve Transmissible  
656 Neoplasia from the Soft-Shell Clam *Mya arenaria* (MarBTN) in Seawater. Epub ahead of  
657 print 2022. DOI: 10.3390/pathogens11030283.
- 658 [13] Skazina M, Odintsova N, Maiorova M, et al. Two lineages of bivalve transmissible  
659 neoplasia affect the blue mussel *Mytilus trossulus* Gould in the subarctic Sea of Okhotsk.  
660 *Curr Zool* 2022; 00: 1–12.
- 661 [14] Skazina M, Ponomartsev N, Maiorova M, et al. Genetic features of bivalve transmissible  
662 neoplasia in blue mussels from the Kola Bay (Barents Sea) suggest a recent trans-Arctic  
663 migration of the cancer lineages. *Mol Ecol*. Epub ahead of print 5 October 2023. DOI:  
664 10.1111/MEC.17157.
- 665 [15] Skazina M, Odintsova N, Maiorova M, et al. First description of a widespread *Mytilus*  
666 *trossulus*-derived bivalve transmissible cancer lineage in *M. trossulus* itself. *Sci Rep*; 11.  
667 Epub ahead of print 2021. DOI: 10.1038/s41598-021-85098-5.

- 668 [16] Hammel M, Simon A, Arbiol C, et al. Prevalence and polymorphism of a mussel  
669 transmissible cancer in Europe. *Mol Ecol* 2022; 31: 736–751.
- 670 [17] Farley CA. Sarcomatoid proliferative disease in a wild population of blue mussels  
671 (*Mytilus edulis*). *J Natl Cancer Inst* 1969; 43: 509–516.
- 672 [18] Baez-Ortega A, Gori K, Strakova A, et al. Somatic evolution and global expansion of an  
673 ancient transmissible cancer lineage. *Science (1979)*; 365. Epub ahead of print 2 August  
674 2019. DOI: 10.1126/science.aau9923.
- 675 [19] Caza F, E B, FJ V, et al. Hemocytes released in seawater act as Trojan horses for spreading  
676 of bacterial infections in mussels. *Sci Rep* 2020; 10: 19696–19696.
- 677 [20] Skibinski DOF, Beardmore JA, Cross TF. Aspects of the population genetics of *Mytilus*  
678 (*Mytilidae*; Mollusca) in the British Isles. *Biological Journal of the Linnean Society* 1983;  
679 19: 137–183.
- 680 [21] Coustau C, Renaud F, Delay B. Genetic characterization of the hybridization between  
681 *Mytilus edulis* and *M. galloprovincialis* on the Atlantic coast of France. *Mar Biol* 1991;  
682 111: 87–93.
- 683 [22] Simon A, Fraïsse C, El Ayari T, et al. How do species barriers decay? Concordance and  
684 local introgression in mosaic hybrid zones of mussels. *J Evol Biol* 2021; 34: 208–223.
- 685 [23] Bierne N, Bonhomme F, David P. Habitat preference and the marine-speciation paradox.  
686 *Proc R Soc Lond B Biol Sci* 2003; 270: 1399–1406.
- 687 [24] Bierne N, David P, Langlade A, et al. Can habitat specialisation maintain a mosaic hybrid  
688 zone in marine bivalves? *Mar Ecol Prog Ser* 2002; 245: 157–170.
- 689 [25] Lowe DM, Moore MN. Cytology and Quantitative Cytochemistry of a Proliferative  
690 Atypical Hemocytic Condition in *Mytilus edulis* (Bivalvia, Mollusca). *J Natl Cancer Inst*  
691 1978; 60: 1455–1459.
- 692 [26] Landsberg JH. Neoplasia and biotoxins in bivalves: Is there a connection? *J Shellfish Res*  
693 1996; 15: 203–230.
- 694 [27] Green M, Alderman DJ. Neoplasia in *Mytilus edulis* L. from United Kingdom waters.  
695 *Aquaculture* 1983; 30: 1–10.
- 696 [28] Mix MC. Haemic neoplasms of bay mussels, *Mytilus edulis* L., from Oregon: occurrence,  
697 prevalence, seasonality and histopathological progression. *J Fish Dis* 1983; 6: 239–248.

- 698 [29] Rasmussen LPD. Occurrence, prevalence and seasonality of neoplasia in the marine  
699 mussel *Mytilus edulis* from three sites in Denmark. *Marine Biology: International Journal*  
700 *on Life in Oceans and Coastal Waters* 1986; 92: 59–64.
- 701 [30] Charles M. Etude des pathogènes, des conditions physiologiques et pathologiques  
702 impliqués dans les mortalités anormales de moules (*Mytilus* sp.), [https://tel.archives-](https://tel.archives-ouvertes.fr/tel-03053331)  
703 [ouvertes.fr/tel-03053331](https://tel.archives-ouvertes.fr/tel-03053331) (2019, accessed 13 October 2021).
- 704 [31] Bramwell G, Schultz AG, Sherman CDH, et al. A review of the potential effects of climate  
705 change on disseminated neoplasia with an emphasis on efficient detection in marine  
706 bivalve populations. *Science of The Total Environment* 2021; 775: 145134.
- 707 [32] Burioli EAV, Trancart S, Simon A, et al. Implementation of various approaches to study  
708 the prevalence, incidence and progression of disseminated neoplasia in mussel stocks. *J*  
709 *Invertebr Pathol* 2019; 168: 107271.
- 710 [33] Fraïsse C, Belkhir K, Welch JJ, et al. Local interspecies introgression is the main cause of  
711 extreme levels of intraspecific differentiation in mussels. *Mol Ecol* 2016; 25: 269–286.
- 712 [34] Simon A, Arbiol C, Nielsen EE, et al. Replicated anthropogenic hybridisations reveal  
713 parallel patterns of admixture in marine mussels. *Evol Appl* 2020; 13: 575–599.
- 714 [35] Semagn K, Babu R, Hearne S, et al. Single nucleotide polymorphism genotyping using  
715 Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application  
716 in crop improvement. *Molecular Breeding* 2014; 33: 1–14.
- 717 [36] Cuenca J, Aleza P, Navarro L, et al. Assignment of SNP allelic configuration in polyploids  
718 using competitive allele-specific PCR: application to citrus triploid progeny. *Ann Bot*  
719 2013; 111: 731–742.
- 720 [37] R Core Team. R: A language and environment for statistical computing, [http://www.r-](http://www.r-project.org/index.html)  
721 [project.org/index.html](http://www.r-project.org/index.html) (2020).
- 722 [38] Josse J, Husson F. missMDA: A Package for Handling Missing Values in Multivariate  
723 Data Analysis. *J Stat Softw* 2016; 70: 1–31.
- 724 [39] Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. *J Stat*  
725 *Softw* 2008; 25: 1–18.
- 726 [40] Harrell JF. *\_Hmisc: Harrell Miscellaneous\_*. *R package version 51-0* .
- 727 [41] Bivand R, Keitt T, Rowlingson B. rgdal: Bindings for the ‘Geospatial’ Data Abstraction  
728 Library.

- 729 [42] Bivand R. R Packages for Analyzing Spatial Data: A Comparative Case Study with Areal  
730 Data. *Geogr Anal* 2022; 54: 488–518.
- 731 [43] Rousset F, Ferdy J-B. Testing environmental and genetic effects in the presence of spatial  
732 autocorrelation. *Ecography* 2014; 37: 781–790.
- 733 [44] Hartig F. *\_DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed)  
734 Regression Models\_*, <https://CRAN.R-project.org/package=DHARMA> (2022, accessed 19  
735 December 2023).
- 736 [45] Thomas F, Fisher D, Fort P, et al. Applying ecological and evolutionary theory to cancer: a  
737 long and winding road. *Evol Appl* 2013; 6: 1–10.
- 738 [46] Morrisey D, Gadd J, Page M, et al. *In-water cleaning of vessels: biosecurity and chemical  
739 contamination risks. New Zealand Ministry for Primary Industries Technical Paper No,*  
740 <https://www.researchgate.net/publication/276920418> (November 2013, accessed 3  
741 November 2023).
- 742 [47] Gadd JB, Lewis JA. In-water cleaning of vessels: biosecurity and chemical contamination  
743 risks, <https://www.researchgate.net/publication/276920418> (2011, accessed 23 October  
744 2023).
- 745 [48] Bommarito C, Thieltges DW, Pansch C, et al. Effects of first intermediate host density,  
746 host size and salinity on trematode infections in mussels of the south-western Baltic Sea.  
747 *Parasitology* 2021; 148: 486–494.
- 748 [49] Hopkins SR, Fleming-Davies AE, Belden LK, et al. Systematic review of modelling  
749 assumptions and empirical evidence: Does parasite transmission increase nonlinearly with  
750 host density? *Methods Ecol Evol* 2020; 11: 476–486.
- 751 [50] Široka M, Piličić S, Milošević T, et al. A novel approach for assessing the ports’  
752 environmental impacts in real time – The IoT based port environmental index. *Ecol Indic*  
753 2021; 120: 106949.
- 754 [51] Roberts T, Williams I, Preston J, et al. Ports in a Storm: Port-City Environmental  
755 Challenges and Solutions. *Sustainability* 2023, Vol 15, Page 9722 2023; 15: 9722.
- 756 [52] Destoumieux-Garzón D, Mavingui P, Boetsch G, et al. The one health concept: 10 years  
757 old and a long road ahead. *Front Vet Sci* 2018; 5: 14.

- 758 [53] Canesi L, Ciacci C, Bergami E, et al. Evidence for immunomodulation and apoptotic  
759 processes induced by cationic polystyrene nanoparticles in the hemocytes of the marine  
760 bivalve *Mytilus*. *Mar Environ Res* 2015; 111: 34–40.
- 761 [54] Girón-Pérez MI. Relationships between innate immunity in bivalve molluscs and  
762 environmental pollution. *ISJ* 2010; 7: 149–156.
- 763 [55] Renault T. Immunotoxicological effects of environmental contaminants on marine  
764 bivalves. *Fish Shellfish Immunol* 2015; 46: 88–93.
- 765 [56] Chandra RK. Nutrition, immunity and infection: From basic knowledge of dietary  
766 manipulation of immune responses to practical application of ameliorating suffering and  
767 improving survival. *Proc Natl Acad Sci U S A* 1996; 93: 14304–14307.
- 768 [57] Tseng M, Myers JH. The Relationship between Parasite Fitness and Host Condition in an  
769 Insect - Virus System. *PLoS One* 2014; 9: e106401.
- 770 [58] Pike VL, Lythgoe KA, King KC. On the diverse and opposing effects of nutrition on  
771 pathogen virulence. *Proceedings of the Royal Society B*; 286. Epub ahead of print 10 July  
772 2019. DOI: 10.1098/RSPB.2019.1220.
- 773 [59] Cressler CE, Nelson WA, Day T, et al. Disentangling the interaction among host resources,  
774 the immune system and pathogens. *Ecol Lett* 2014; 17: 284–293.
- 775 [60] Georgiades E, Scianni C, Davidson I, et al. The Role of Vessel Biofouling in the  
776 Translocation of Marine Pathogens: Management Considerations and Challenges .  
777 *Frontiers in Marine Sciences*; 8. Epub ahead of print 2021. DOI:  
778 10.3389/fmars.2021.660125.
- 779 [61] Ruiz GM, Rawlings TK, Dobbs FC, et al. Global spread of microorganisms by ships.  
780 *Nature* 2000 408:6808 2000; 408: 49–50.
- 781 [62] Massé C, Viard F, Humbert S, et al. An Overview of Marine Non-Indigenous Species  
782 Found in Three Contrasting Biogeographic Metropolitan French Regions: Insights on  
783 Distribution, Origins and Pathways of Introduction. *Diversity (Basel)* 2023; 15: 161.
- 784 [63] Molnar JL, Gamboa RL, Revenga C, et al. Assessing the global threat of invasive species  
785 to marine biodiversity. *Front Ecol Environ* 2008; 6: 485–492.
- 786 [64] Bishop JDD, Wood CA, Yunnice ALE, et al. Unheralded arrivals: non-native sessile  
787 invertebrates in marinas on the English coast. *Aquat Invasions* 2015; 10: 249–264.

- 788 [65] Couton M, Lévêque L, Daguin-Thiébaud C, et al. Water eDNA metabarcoding is effective  
789 in detecting non-native species in marinas, but detection errors still hinder its use for  
790 passive monitoring. *Biofouling* 2022; 38: 367–383.
- 791 [66] Pagenkopp Lohan KM, Darling JA, Ruiz GM. International shipping as a potent vector for  
792 spreading marine parasites. *Divers Distrib* 2022; 28: 1922–1933.
- 793 [67] Clarke Murray C, Pakhomov EA, Therriault TW. Recreational boating: a large  
794 unregulated vector transporting marine invasive species. *Divers Distrib* 2011; 17: 1161–  
795 1172.
- 796 [68] Salta M, Chambers L, Wharton J, et al. Marine fouling organisms and their use in  
797 antifouling bioassays.
- 798 [69] Davidson IC, Brown CW, Sytsma MD, et al. The role of containerships as transfer  
799 mechanisms of marine biofouling species. *Biofouling* 2009; 25: 645–655.
- 800 [70] Bouda A, Bachari NI, Nacef L, et al. Risk Analysis of Invasive Species Introduction in the  
801 Port of Arzew, by Calculation of Biofouling Surface on Ships’ Hulls. *Environmental*  
802 *Modeling and Assessment* 2018; 23: 185–192.
- 803 [71] Sylvester F, Kalaci O, Leung B, et al. Hull fouling as an invasion vector: can simple  
804 models explain a complex problem? *Journal of Applied Ecology* 2011; 48: 415–423.
- 805 [72] Floerl O, Inglis GJ, Dey K, et al. The importance of transport hubs in stepping-stone  
806 invasions. *Journal of Applied Ecology* 2009; 46: 37–45.
- 807 [73] Guzinski J, Ballenghien M, Daguin-Thiébaud C, et al. Population genomics of the  
808 introduced and cultivated Pacific kelp *Undaria pinnatifida*: Marinas—not farms—drive  
809 regional connectivity and establishment in natural rocky reefs. *Evol Appl* 2018; 11: 1582–  
810 1597.
- 811 [74] International Maritime Organization [IMO]. *International Convention for the Control and*  
812 *Management of Ships’ Ballast Water and Sediments*. Lodon, 2004.
- 813 [75] International Maritime Organization [IMO]. *Guidelines for the Control and Management*  
814 *of Ships’ Biofouling to Minimize the Transfer of Invasive Aquatic Species*. London, 2011.
- 815 [76] McClay T, Zabin C, Davidson I, et al. *Vessel biofouling prevention and management*  
816 *options report*. 2015.
- 817 [77] GEF-UNDP-IMO GloFouling Partnerships Project 2022. Biofouling Management for  
818 Recreational Boating: Recommendations to Prevent the Introduction and Spread of

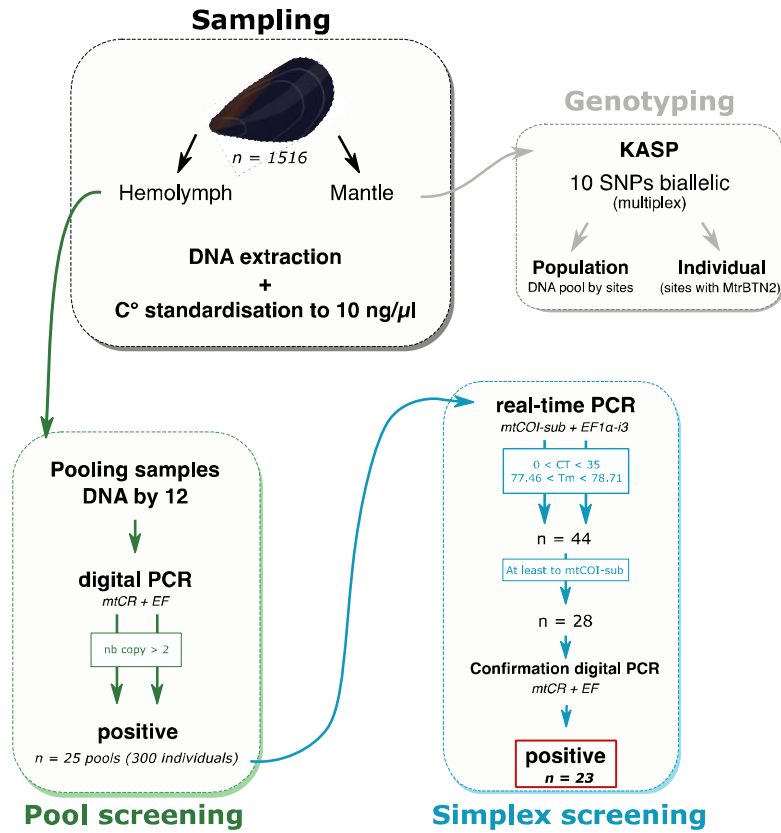
- 819 Invasive Aquatic Species. *GEF-UNDP-IMO GloFouling Partnerships Project, 2022.*  
820 *Biofouling Management for Recreational Boating: Recommendations to Prevent the*  
821 *Introduction and Spread of Invasive Aquatic Species.*
- 822 [78] Briski E, Wiley CJ, Bailey SA. Role of domestic shipping in the introduction or secondary  
823 spread of nonindigenous species: biological invasions within the Laurentian Great Lakes.  
824 *Journal of Applied Ecology* 2012; 49: 1124–1130.
- 825 [79] Ulman A, Ferrario J, Forcada A, et al. Alien species spreading via biofouling on  
826 recreational vessels in the Mediterranean Sea. *Journal of Applied Ecology* 2019; 56:  
827 2620–2629.
- 828 [80] Ashton G V., Zabin CJ, Davidson IC, et al. Recreational boats routinely transfer organisms  
829 and promote marine bioinvasions. *Biol Invasions* 2022; 24: 1083–1096.
- 830 [81] Lim CS, Leong YL, Tan KS. Managing the risk of non-indigenous marine species transfer  
831 in Singapore using a study of vessel movement. *Mar Pollut Bull* 2017; 115: 332–344.

832

833

834 Supplementary material

835



836

837 **Figure S1: Diagram of screening and genotyping experimental steps.**

838

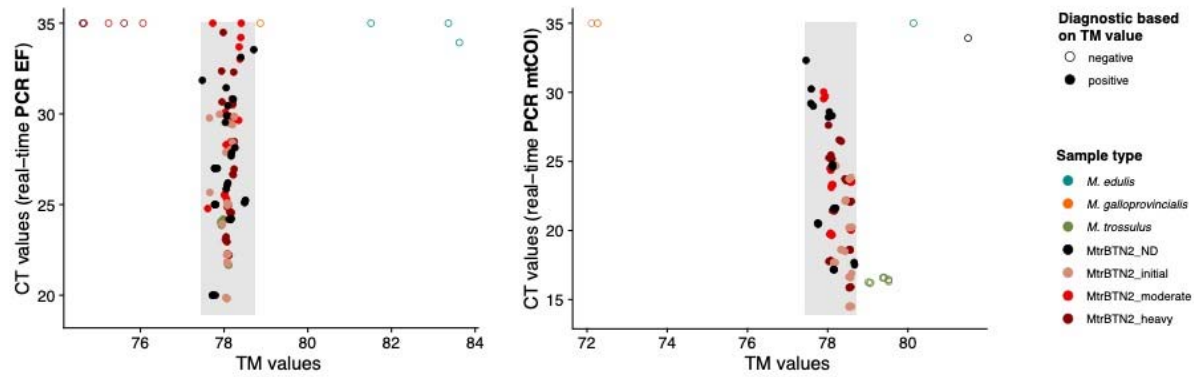
839

840

841

842





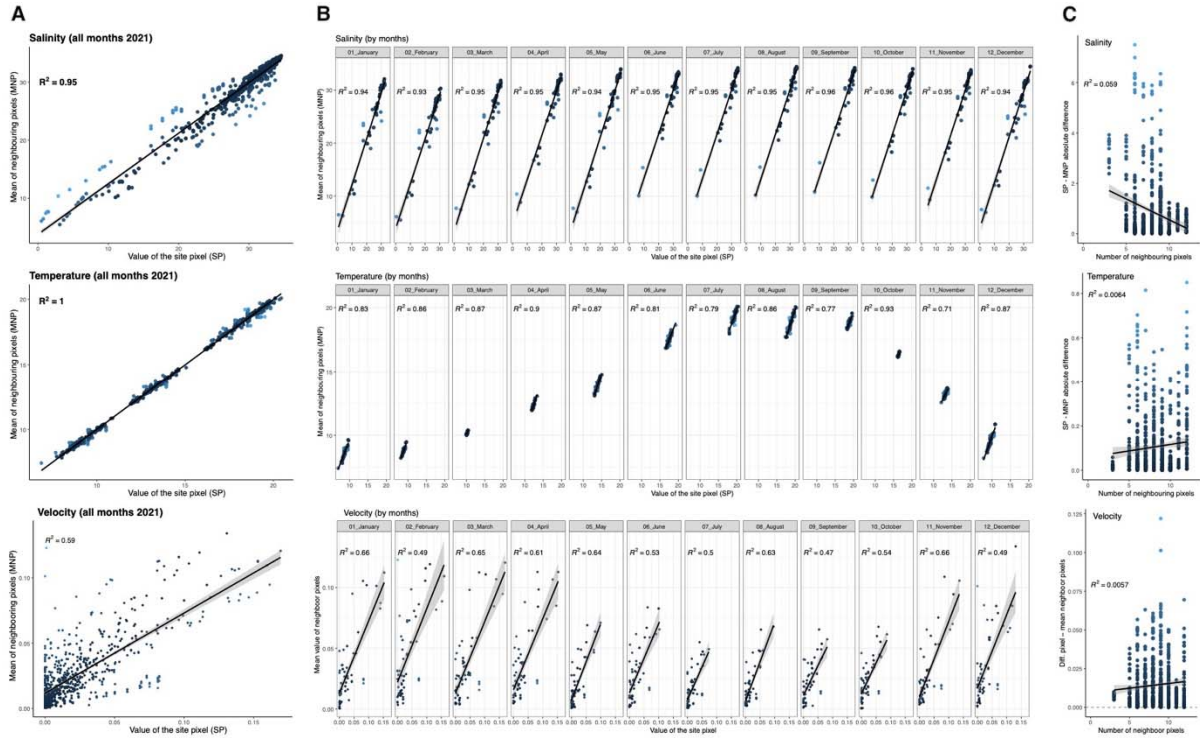
843

844 **Figure S2: TM threshold determination for EF1 $\alpha$ -i3 (left) and mtCOI-sub (right) real-time**  
845 **PCR markers.** Grey rectangles correspond to TM threshold based on MtrBTN2 positive control  
846 samples (77.46<TM<78.71). Sample type colour and diagnostic based on TM values are  
847 indicated in the legend. *M. edulis* and *M. galloprovincialis* correspond to negative controls, *M.*  
848 *trossulus* are *M. trossulus* controls, and MtrBTN2 samples are positive controls (ND: cancer  
849 stage not defined, initial: early stage, moderate: moderate stage, heavy: late stage, see TableS3)

850

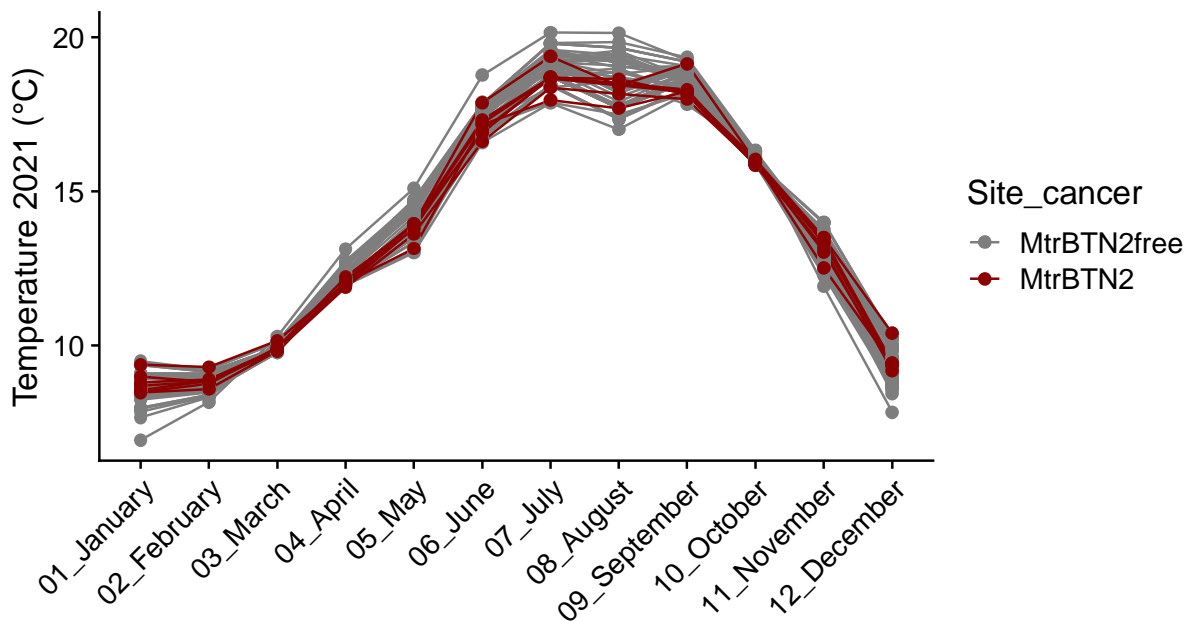
851

852



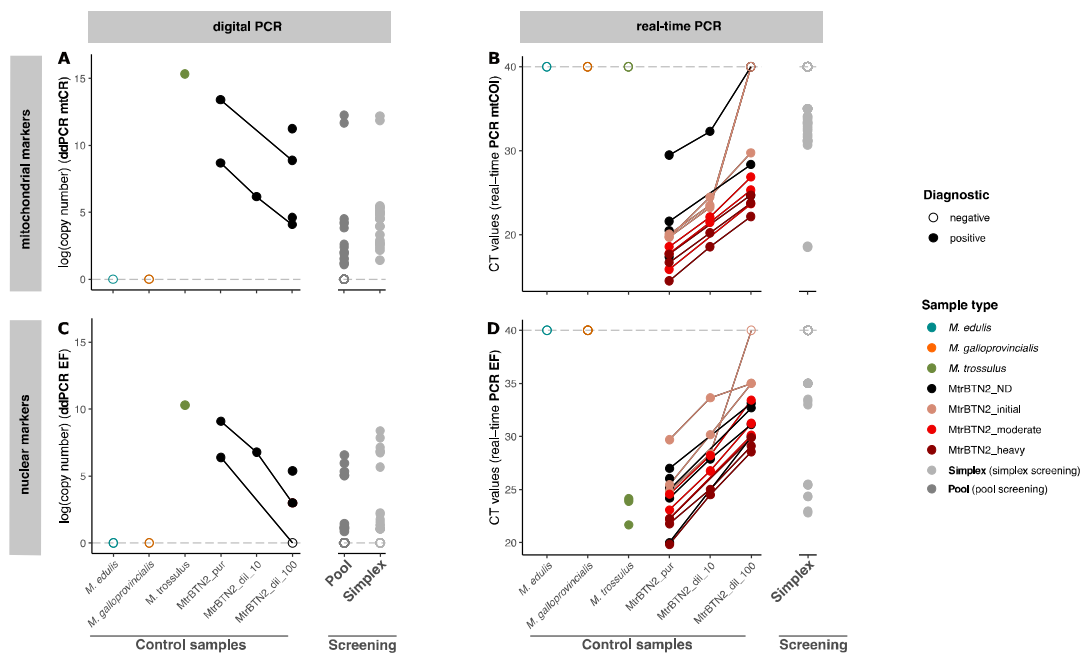
853

854 **Figure S3: Evaluation of Copernicus model resolution for salinity, temperature and**  
 855 **velocity data.** Correlation between the value of the site pixel (SP) and the mean value of  
 856 neighbouring pixels (MNP) for each variable. **(A)** Correlation including all months of year 2021  
 857 and **(B)** each month separately. No correlation observed between the absolute difference of SP  
 858 and MNP values and the number of neighbouring pixels **(C)**.



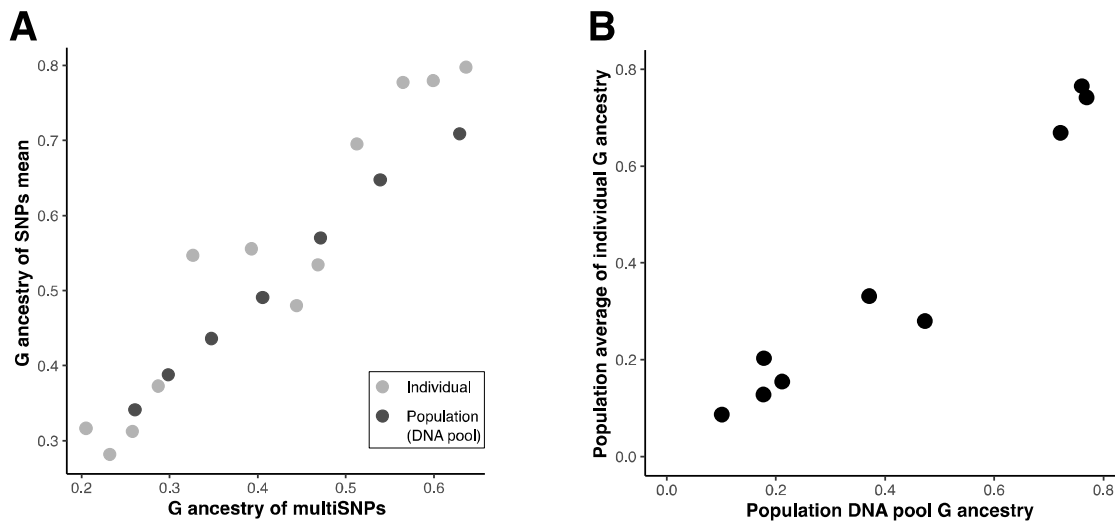
859  
 860 **Figure S4: Sampled sites temperature, recorded monthly in 2021.** Points correspond to a  
 861 temperature measurement, solid lines refer to MtrBTN2-free sites (grey) and MtrBTN2 affected  
 862 sites (darkred).

863  
 864  
 865



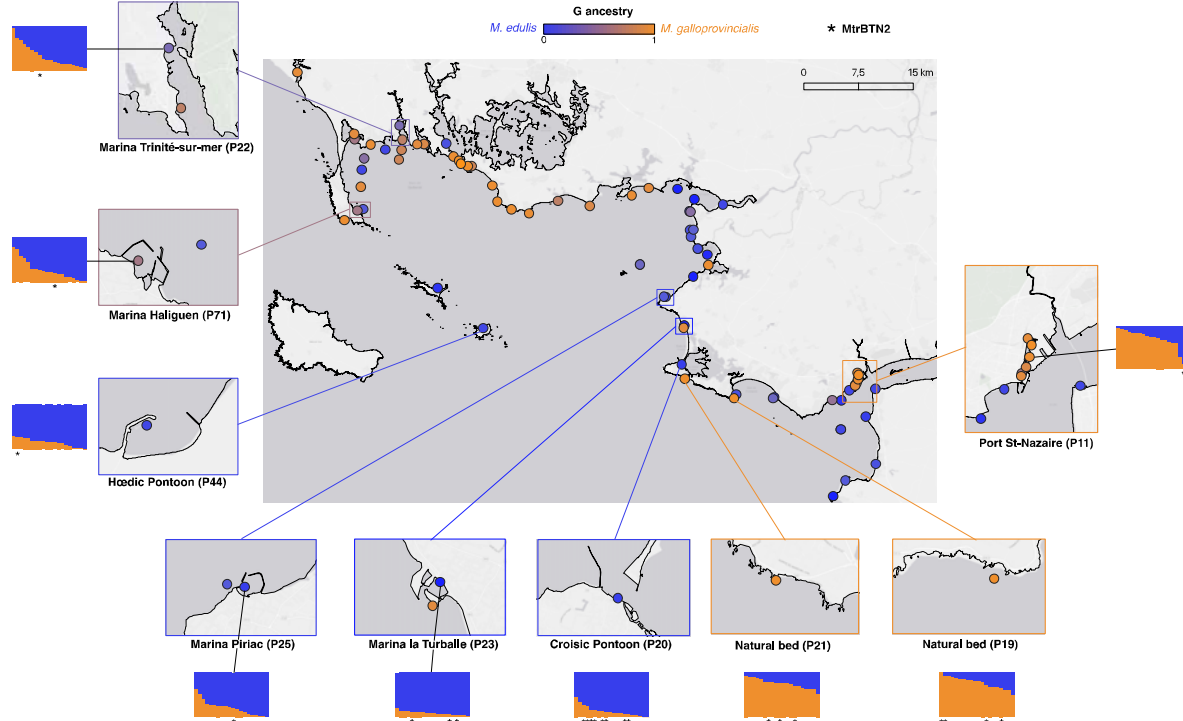
866

867 **Figure S5: Real-time PCR and digital PCR results.** (A) and (C) represent ddPCR results of EF  
868 and mtCR, and (B) and (D) represent real-time PCR results of EF1 $\alpha$ -i3 and mtCOI-sub,  
869 respectively. Dotted lines correspond to the negative thresholds, 0 for no DNA copy detected by  
870 ddPCR and 40 for no amplification in real-time PCR. Solid lines in each graph refer to the same  
871 sample diluted to different concentrations (pure, x10, x100). *M. edulis* and *M. galloprovincialis*  
872 correspond to negative controls, *M. trossulus* are *M. trossulus* controls, and MtrBTN2 samples  
873 are positive controls (ND: cancer stage not defined, Initial: early stage, moderate: moderate  
874 stage, heavy: late stage, see TableS2). Screening results using both method and markers are  
875 shown for Pool and Simplex. Sample type colours and diagnostic status are indicated in the  
876 figure legend.  
877



878  
879 **Figure S6: Genotyping method validation.** (A) Positive correlation between multiSNPs G  
880 ancestry and the mean of the 10 single SNPs from a subset of Individual and Pool samples. (B)  
881 Positive correlation between population DNA pool G ancestry and the mean individual G  
882 ancestry of individuals from the same population.

883  
884  
885

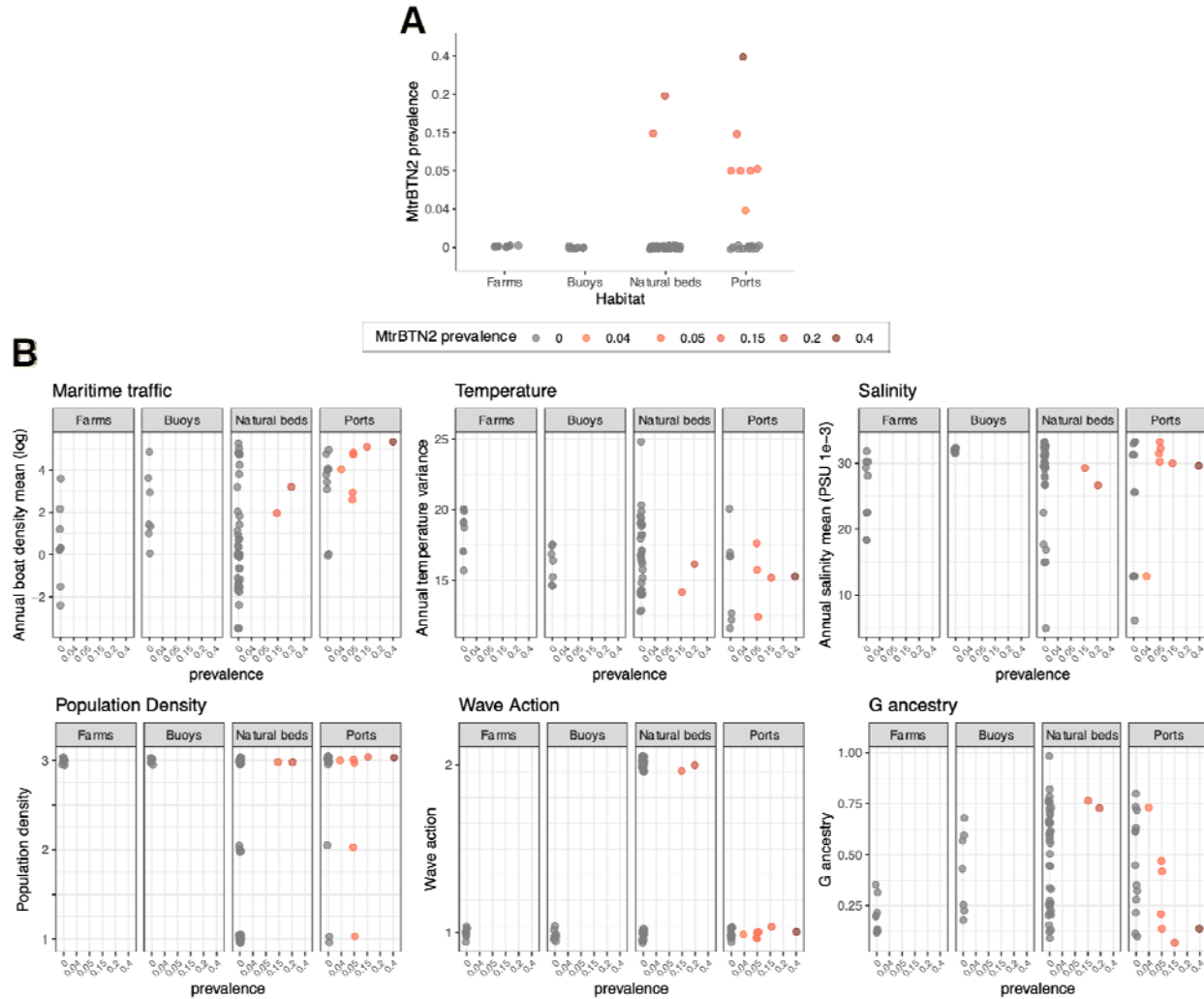


886

887 **Figure S7: Population and individual ancestry composition across the sampled area.**

888 Zoomed-in sites correspond to sites with mussels affected by MtrBTN2. Barplots represent the  
889 estimated ancestry of individuals based on the 10SNPs multiplex tool. MtrBTN2 individuals are  
890 indicated by stars under each barplot. Coloured points represent the population average ancestry,  
891 as indicated in the figure's legend. Map origin: Ersi gray (light).

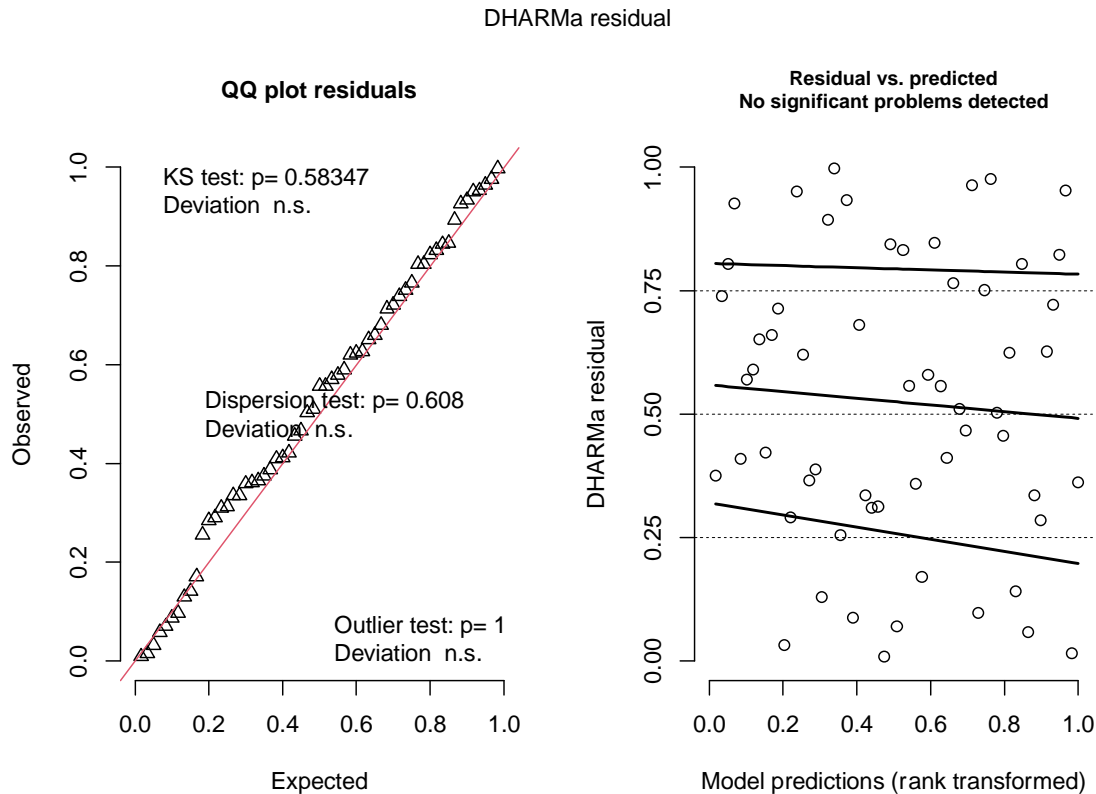
892



893

894 **Figure S8: Univariate correlations between the prevalence of MtrBTN2 and environmental**  
895 **variables.** (A) MtrBTN2 prevalence among habitats. (B) Correlations for the six environmental  
896 variables studied. Coloured points correspond to the prevalence of MtrBTN2 as indicated in the  
897 legend.

898



899

900 **Figure S9: Residual diagnostics of the logistic regression model accounting for spatial**  
901 **autocorrelation.** Model results are presented in Table S15. Plot generated using  
902 *simulateResiduals()* function of DHARMA package in R.

903

904

905 **Supplementary Tables (→ [See excel TableS1-15](#))**

906

907 **TableS1 : Description of sampled sites.** Sources of environmental variables are provided in  
908 TableS5. Site: site name; N: number of sampled mussels; MtrBTN2 prevalence: number of  
909 MtrBTN2 affected mussels; Temperature: annual variance; Salinity: annual mean; Maritime  
910 traffic: annual mean in log; temp: temperature; sal: salinity; vel\_E and vel\_N: Eastward and  
911 Northward current velocity, respectively; [fish-pass-other][19-21]\_[1-4] fish: fishing boat  
912 density, pass: passenger boat density, other: recreational boat density, 19 to 21 correspond to the  
913 year of record, and 1 to 4 correspond to spring, summer, autumn, winter, respectively.

914 Temperature and Salinity missing data (blank) correspond to sites outside of Copernicus model  
915 pixels (n = 13).

916

917 **Table S2: Primer and probe sequences used in real-time PCR and digital PCR.**

918

919 **Table S3: Control samples used to validate real-time PCR and ddPCR screening tools.** All  
920 samples are from our laboratory sample collection. *M. edulis* and *M. galloprovincialis*  
921 correspond to negative controls, *M. trossulus* are *M. trossulus* controls, and MtrBTN2 samples  
922 are positive controls. MtrBTN2 stages were defined by cytology of the hemolymph, according to  
923 the following classification: early (<15% cancer cells), moderate (15-75% cancer cells), late  
924 (>75% cancer cells), ND when cancer stage was not defined.

925

926 **Table S4: Bi-allelic nuclear SNPs used for the multiSNPs genotyping.**

927

928 **TableS5: Source and description of environmental variables collected for each habitat**  
929 **(ports, natural beds, farms, buoys).**

930

931 **TableS6: Data for the evaluation of Copernicus model resolution.** Sources of environmental  
932 variables are provided in TableS5. Site: site name; N\_pixels: number of neighbouring pixels  
933 within a buffer zone of 0.05°; SP: value of the site pixel; MNP: mean value of the neighbouring  
934 pixels; SP\_MNP: absolute difference between SP and MNP values; .s, .t and .v refer to salinity,  
935 temperature and velocity values, respectively. Velocity correspond to the combination of squared  
936 Northward and Eastward velocity values. Missing data (blank) correspond to sites outside of  
937 Copernicus model pixels (n = 13).

938

939 **Table S7: Control sample results validate real-time PCR and ddPCR screening tools.**  
940 EF\_CopyNb: copy number for EF ddPCR marker; mtCR\_CopyNb: copy number for mtCR  
941 ddPCR marker; CT: threshold cycle, TM: melting temperature for EF1 $\alpha$ -i3 and mtCOI-sub real-  
942 time PCR markers. NP: not performed.

943



944 **TableS8: Pooled and simplex screening results.** Pool\_CopyNb\_EF: copy number for EF  
945 ddPCR marker (pooled screening); Pool\_CopyNb\_mtCR: copy number for mtCR ddPCR marker  
946 (pooled screening); CT: threshold cycle, TM: melting temperature for EF1 $\alpha$ -i3 and mtCOI-sub  
947 real-time PCR markers; Simplex\_CopyNb\_EF: copy number for EF ddPCR marker (simplex  
948 screening); Simplex\_CopyNb\_mtCR: copy number for mtCR ddPCR marker (simplex  
949 screening). NP: not performed.

950

951 **TableS9: Simplex screening results for positive samples.** CT: threshold cycle, TM: melting  
952 temperature for EF1 $\alpha$ -i3 and mtCOI-sub real-time PCR markers; PCR\_result: marker name  
953 positive in real-time PCR; EF\_CopyNb: copy number for EF ddPCR marker; mtCR\_CopyNb:  
954 copy number for mtCR ddPCR marker; ddPCR\_result: marker name positive in ddPCR;  
955 MtrBTN2: final diagnostic status. NP: not performed.

956

957 **Table S10: Correspondence between pool and simplex screening.** Pool\_CopyNb\_EF : copy  
958 number for EF ddPCR marker (pooled screening); Pool\_CopyNb\_mtCR : copy number for  
959 mtCR ddPCR marker (pooled screening); Pool\_ddPCR\_result: marker name positive in ddPCR  
960 (pool); Nb\_ind\_positive: number of individuals in the pool found positive for one or two marker  
961 in simplex screening; Simplex\_PCR\_result: marker name of the positive individual in real-time  
962 PCR; Simplex\_ddPCR\_result: marker name of the positive individual in ddPCR, negative if the  
963 sample did not amplify. NP: not performed.

964

965 **Table S11: MultiSNPs genotyping results.** G ancestry: fluorescence value; Ancestry: ancestry  
966 defined based on G ancestry values.

967

968 **Table S12: Contingency table used for the Fisher exact test.**

969

970 **Table S13: Moran I test on environmental variables.** Results were obtained using moran.test()  
971 function in R.

972

973 **Table S14: Correlation coefficient and Spearman correlation test results for each**  
974 **environmental data pairs.**

975

976 **Table S15: Spatial logistic regression model testing the effect of environmental variables on**

977 **MtrBTN2 prevalence.** Model results were obtained using `fitme()` from `spaMM` package in R.

978 Matern covariance function was included to account for spatial autocorrelation (i.e. `Matern(1 |`

979 `Longitude + Latitude)`. Effects of environmental variables were tested using Likelihood ratio

980 (LR) tests using `fixedLRT()` function from the same package. Temperature: annual variance;

981 Salinity: annual mean; Maritime traffic: annual mean.

982

983

984

985

986

987