



HAL
open science

Water eDNA metabarcoding is effective in detecting non-native species in marinas, but detection errors still hinder its use for passive monitoring

Marjorie Couton, Laurent Lévêque, Claire Daguin-Thiébaud, Thierry Comtet, Frédérique Viard

► To cite this version:

Marjorie Couton, Laurent Lévêque, Claire Daguin-Thiébaud, Thierry Comtet, Frédérique Viard. Water eDNA metabarcoding is effective in detecting non-native species in marinas, but detection errors still hinder its use for passive monitoring. *Biofouling*, In press, 10.1080/08927014.2022.2075739 . hal-03669095

HAL Id: hal-03669095

<https://hal.science/hal-03669095>

Submitted on 16 May 2022

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

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

1 Accepted version of the paper published in Biofouling with the following reference:

2
3 Marjorie Couton, Laurent Lévêque, Claire Daguin-Thiébaud, Thierry Comtet
4 & Frédérique Viard (2022): Water eDNA metabarcoding is effective in detecting non-native
5 species in marinas, but detection errors still hinder its use for passive monitoring, Biofouling,
6 DOI:10.1080/08927014.2022.2075739

7
8 <https://doi.org/10.1080/08927014.2022.2075739>

9

10 **Water eDNA metabarcoding is effective in detecting non-native species in**
11 **marinas, but detection errors still hinder its use for passive monitoring**

12 Marjorie Couton^a, Laurent Lévêque^b, Claire Daguin-Thiébaud^a, Thierry Comtet^a, Frédérique
13 Viard^{a,c*}

14 ^a Sorbonne université, CNRS, UMR 7144, Station Biologique de Roscoff, Place Georges
15 Teissier, 29680 Roscoff, France

16 ^b Sorbonne université, CNRS, FR 2424, Station Biologique de Roscoff, Place Georges
17 Teissier, 29680 Roscoff, France

18 ^c ISEM, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

19 * Corresponding author: frederique.viard@umontpellier.fr

20 ORCID ID:

21 Marjorie Couton : 0000-0001-9880-8646

22 Frédérique Viard: 0000-0001-5603-9527

23 Claire Daguin-Thiébaud 0000-0003-4665-6411

24 **Abstract**

25 Marinas are high-priority targets for marine non-indigenous species (NIS), where
26 they compose a large portion of the biofouling communities. The practicality of water
27 samples collection makes environmental DNA (eDNA) metabarcoding an interesting tool for
28 routine NIS surveys. Here the effectiveness of water-eDNA-metabarcoding to identify
29 biofouling NIS, in ten marinas from western France, was examined. Morphological
30 identification of specimens collected in quadrats brought out 18 sessile benthic NIS beneath
31 floating pontoons. Water-eDNA-metabarcoding detected two thirds of them, failing to detect
32 important NIS. However, sampling and bioinformatics filtering steps can be optimized to
33 identify more species. In addition, this method allowed the detection of additional NIS from
34 neighboring micro-habitats. Caution should however be taken when reporting putative novel
35 NIS, because of errors in species assignment. This work highlights that water-eDNA-
36 metabarcoding is effective for active (targeted) NIS surveys, and could be significantly
37 improved for its further use in marine NIS passive surveys.

38

39 **Keywords:** biological invasions, environmental DNA, metabarcoding, morphology, ports,
40 quadrats

41

42 **Introduction**

43 As the number of marine introductions is growing worldwide (Seebens et al. 2017),
44 much effort is directed towards non-indigenous species (NIS) management. The implemented
45 strategies can vary greatly from, limiting the spread of new species by applying regulation on
46 vectors of introduction, to eradication protocols to remove invasive species from a particular
47 area (Robertson et al. 2020). Nevertheless, vectors of introductions are not efficiently
48 regulated (Carney et al. 2017), and measures of eradication are difficult to apply when non-
49 indigenous species (NIS) have become too abundant or widespread (Ojaveer et al. 2015).
50 Monitoring NIS to detect new arrivals or to evaluate the spread of already established species,
51 is thus crucial to support efficient management strategies (Robertson et al. 2020). It is
52 especially true in the primary introduction sites, such as ports or marinas. In such habitats,
53 fouling communities from artificial hard substrates are often targeted since they include many
54 NIS (Dafforn et al. 2009; Leclerc et al. 2020).

55 Diverse traditional approaches exist to monitor biofouling communities, from Rapid
56 Assessment Surveys or full species inventories (Lehtiniemi et al. 2015) to passive sampling
57 strategies using settlement plates (Dafforn et al. 2009; Outinen et al. 2021). Their inherent
58 issues, however, might prevent their broad use by practitioners and hinder their efficiency.
59 Most of the problems with traditional approaches arise from the morphological identification
60 of organisms. They require taxonomic expertise, and the time constraint of field sessions can
61 lead to erroneous identification for complex taxa (eg the ascidians *Botrylloides* spp., Viard et
62 al. 2019). Furthermore, these approaches do not allow the detection of introductions by
63 (pseudo-)cryptic species, common among NIS, especially in biofouling assemblages (Viard et

64 al. 2016; Turon et al. 2020). Finally, they are not always sufficient to discriminate species at
65 early-settlement stages. There is thus a particular need for a new method solving these issues.

66 Molecular approaches, especially DNA metabarcoding, have been shown to overcome
67 these limitations (Zaiko et al. 2015; Rey et al. 2020). Metabarcoding has been successfully
68 applied to various types of samples in the context of NIS surveys (eg plankton: Couton et al.
69 2019, sediment: Shaw et al. 2019, water: Borrell et al. 2017, homogenized organisms:
70 Pearman et al. 2021), which provide different results, reflecting the composition of different
71 communities (Koziol et al. 2018; Fernandez et al. 2021). Rey et al. (2020) suggested that, for
72 NIS surveys, all types of samples (or a combination of several of them) should be
73 simultaneously processed (see also Holman et al. 2019). However, this might be hardly
74 feasible for regular monitoring, as it would require expensive sampling tools and expertise
75 that is not always available, and would be very time-consuming. In that context, water
76 environmental DNA (eDNA) would be the sample type of choice as seawater can be easily
77 and rapidly sampled in ports, even by non-experts with a minimal training, and the sampling
78 protocol can be highly standardized (Morissette et al. 2021). Moreover, being non-destructive,
79 water sampling can be conducted as often as required, allowing fine temporal surveys, which
80 is recommended as many NIS are short-lived and show seasonal patterns of abundance
81 (Couton et al. 2019). Finally, water might allow to get a broader picture of NIS present in a
82 given marina, encompassing at the same time species living in the water column or on the
83 bottom (Stat et al. 2017).

84 Very few studies investigated the reliability of water-eDNA-metabarcoding to study
85 benthic communities, even less when considering hard-bottom, and they provided contrasting
86 results. For example, while Leduc et al. (2019) and Antich et al. (2021) reported many false

87 negatives (only up to 15% benthic species/MOTUs retrieved in water-eDNA samples), West
88 et al. (2021) recovered 70% of coral taxa. Many authors thus preferentially used bulk samples
89 from scraping of either hard bottoms or settlement plates (Duarte et al. 2021). This approach,
90 however, is not free of bias, and requires processing a large amount of biomass, which is
91 time-consuming and destructive for the community (van der Loos and Nijland 2021). As an
92 example, von Ammon et al. (2018) did not recover 7 NIS out of 48 after metabarcoding of
93 bulk communities from settlement plates. On the other hand, the fact that several authors
94 identified DNA of benthic species in water samples (eg Borrell et al. 2017; Cordier et al.
95 2019), particularly with species-specific approaches (eg Ardura et al. 2015; Simpson et al.
96 2017), is encouraging. The practicality of water-eDNA sampling drives the need to determine
97 its reliability to monitor biofouling communities in marinas, to understand why
98 metabarcoding is not always able to detect benthic species and to evaluate how these
99 limitations could be overcome (Trebitz et al. 2017; Blackman et al. 2019; Darling et al. 2020;
100 Makiola et al. 2020).

101 Some of the limitations inherent to the use of water-eDNA metabarcoding are already
102 well known, and previous works made suggestions to circumvent them. One of these
103 limitations is the degradation of DNA in water over time which can be partially addressed by
104 using smaller marker regions (Taberlet et al. 2012). Others, such as DNA transport and
105 dilution or differential shedding rates between species can also be decreased by sampling as
106 close as possible to the target community (Antich et al. 2021). However, even when all these
107 factors are taken into account to maximize the chances of species detection, the likelihood of
108 false detections or non-detections is still not negligible. It is therefore crucial to evaluate the
109 remaining biases and investigate possible explanations. In addition, the whole process, from

110 sampling to taxonomic assignment, has rarely been considered at once to determine which
111 steps are the most critical to be optimized.

112 Here, the capacity of water-eDNA-metabarcoding to detect NIS from biofouling
113 communities in marinas was investigated. The goals were to 1) evaluate the extent of species
114 detection errors, and 2) explore the possible causes for these errors, and propose possible
115 solutions. To this end, water sampling was carried out along floating pontoons in ten marinas
116 and two seasons. This sampling strategy allowed to examine detectability at a regional scale
117 and over time, the latter being particularly important as many benthic sessile NIS display
118 short life cycles. Metabarcoding was performed using three markers (18S, COI, and 16S) to
119 increase the detection capacity. Simultaneously, a list of sessile benthic NIS present in the
120 study marinas was established by identifying morphologically and *in situ* species sampled
121 within quadrats beneath the same floating pontoons. This list was used to evaluate the number
122 of false negatives (species observed at the time of sampling but not recovered with eDNA),
123 and explore various causes for metabarcoding detection failures, such as sampling effort or
124 bioinformatics processing.

125

126 **Materials and methods**

127 *Sampling*

128 Biofouling communities immersed beneath floating pontoons (ie floating structures
129 used as docks for mooring leisure boats) were sampled in ten marinas around Brittany
130 (France; Fig. 1). These locations are known to harbour different fouling communities since
131 they are spread over two ecoregions *sensu* Spalding et al. (2007), namely the Celtic seas and

132 the North European Atlantic Shelf, characterized by different hydrological and physical
133 parameters and species assemblages, and encompassing a biogeographic transition zone
134 between the Western English Channel and the NE Atlantic (Gallon et al. 2014). Although
135 native species should particularly differ between these two regions, differences were also
136 observed for NIS, as shown by previous Rapid Assessment Surveys (Bishop et al. 2015; FV
137 and LL, unpublished data). Since many biofouling species are short lived and seasonal,
138 sampling was performed in fall (October) 2017 and spring (May-June) 2018. Two pontoons
139 were selected in each marina, to include potential differences in environmental conditions
140 such as water flow, salinity or type of boats docked (long distance travellers vs. local boats).

141 Seawater was collected every 5 m along a 50-m transect per pontoon. At each
142 sampling point, 200 mL of seawater were collected in a plastic jar using a sterile 100-mL
143 syringe fixed to a sampling rod and immersed at 1 m below the surface. A total volume of 2 L
144 was thus collected per pontoon (after pooling all sampling points). To investigate the effect of
145 sampled volume, in spring 2018, an additional 2-L replicate was collected, resulting in a total
146 volume of 4 L per pontoon (the two 2-L replicates were processed separately). The seawater
147 content of each jar was immediately filtered on site on a Millipore® Sterivex™ filter unit
148 (0.22 µm) using a Masterflex® L/S® economy drive peristaltic pump. After filtration, 2 mL
149 of preservation buffer (sucrose 0.75 M, Tris 0.05 M pH 8, EDTA 0.04 M) were added before
150 storing the filter at -20°C in a portable freezer, until transfer to the laboratory. All
151 consumables not sold as DNA-free were previously immersed in 12.5 % commercial bleach
152 (ca. 0.65 % hypochlorite) for at least 30 min, rinsed with ultrapure water and placed under UV
153 light for at least 15 min. The sampling rod was immersed in 12.5 % commercial bleach for at
154 least 30 min before sampling in a given marina, and rinsed with tap water between each
155 pontoon. Six controls (ie one in three different marinas for each season) were performed by

156 sampling, for every control, 10 replicates of 200 mL of ultrapure water every 5 m along a 50-
157 m transect on a pontoon, with the same equipment, and after applying the aforementioned
158 decontamination protocol, thus mimicking the seawater sampling protocol. The 10 replicates
159 from a pontoon were then pooled, as were the real samples, to reach a total volume of 2 L per
160 control.

161 A list of sessile benthic NIS present, at the time of water collection, in the biofouling
162 community was established. Four taxonomic classes were targeted, namely Gastropoda,
163 Bivalvia, Ascidiacea and Gymnolaemata known to include a substantial number of NIS in the
164 study region (Bishop et al. 2015) (see below for an explanation). Biofoulers were collected
165 beneath pontoons at the same points where water samples were collected, to maximize the
166 probability to sample the same taxa, as shown by Antich et al. (2021). At each of the ten
167 sampling points, all organisms within a 0.25-m² quadrat were scraped by divers and placed in
168 a closing bag. Importantly, the scrapping was performed after the water sampling to avoid any
169 contamination. Specimens (> ca. 5 mm) were morphologically identified on site, to the lowest
170 possible taxonomic level (family, genus or species). The identified organisms were not
171 systematically collected, nor barcoded but, as part of a previous work (Couton et al. 2019),
172 regular barcoding was carried out to get local references on the four taxonomic groups studied
173 here. In order to refine field identification for erected bryozoans from the class
174 Gymnolaemata, specimens were collected at each site for further laboratory identification.

175 *DNA extraction*

176 Caution was taken to avoid any contamination of DNA extracts, as recommended by
177 Goldberg et al. (2016). All equipment and bench surfaces were DNA decontaminated before
178 use and consumables not sold as DNA-free were processed as detailed above. DNA was

179 extracted from the Sterivex units using the NucleoSnap® Finisher Midi extraction kit
180 (Macherey-Nagel) following the manufacturer's protocol with a few modifications
181 (Supplementary Protocol 1).

182 *Library preparation and sequencing*

183 All steps were performed following strict rules for preventing contamination, such as
184 UV irradiation of all tips, plates and tubes, and use of filter tips. Libraries were prepared using
185 a dual-barcoded, dual-indexed two-step PCR procedure. Each DNA sample was amplified 15
186 times with each of the three primer pairs targeting a different marker: i) a 313-bp COI portion
187 amplified using primers designed by Leray et al. (2013), ii) a 350-bp to 450-bp portion of the
188 18S rDNA V1-V2 region amplified using primers designed by Fonseca et al. (2010), and iii) a
189 115-bp portion of the 16S rDNA amplified using primers designed by Kelly et al. (2016)
190 (fragment sizes excluding primers). Combining these markers would increase the detection
191 ability, as they differ in their taxonomic coverage and resolution, and conservation of primer
192 binding sites (Couton et al. 2019; Makiola et al. 2020). Detailed PCR protocols are available
193 in Supplementary Protocol 2. Five tagged PCR replicates were performed for each sample and
194 marker (Table S1). After this first PCR, the five tagged replicates from the same DNA
195 extraction sample were pooled according to their intensity on the agarose gel visualisation. A
196 second PCR was performed to complete Illumina® adapters and to tag samples with a unique
197 dual-index combination (Table S2). Sequencing was performed for each marker separately on
198 an Illumina® MiSeq sequencer. The COI and 18S markers were sequenced using a 600 cycles
199 v3 protocol and the 16S marker was sequenced using a 300 cycles v2 protocol, both with two
200 index reads.

201 *Reads processing and filtering*

202 After demultiplexing performed by the sequencing machine based on index
203 combination (defining the samples), reads were assigned to each of the five PCR replicates
204 according to their tag combination. Primers and tags were then removed using CUTADAPT v-
205 2.8. A set of amplicon sequence variants (ASVs) was produced using DADA2 v-1.13.1
206 (Callahan et al. 2016); see Table S3 for detailed parameter values). When demultiplexing,
207 reads can be falsely assigned to a sample when an index combination is not correctly
208 recognized, which is referred to as ‘index-jump’ (Taberlet et al. 2018). To assess this
209 phenomenon, 12 index combinations, not used in the described PCR experiments, and chosen
210 among the 96 available, were added to the MiSeq sequencing sample sheet in order to get the
211 corresponding fastq files. For each ASV, the number of reads assigned to one of these internal
212 control index combinations was divided by the total number of reads in the complete dataset
213 for this same ASV. The maximal proportion occurring in an index control was recorded for
214 each marker (Table S3). Any ASV that did not account for more than this maximal proportion
215 in a sample was discarded from this sample. Furthermore, only ASVs found in at least two out
216 of the five PCR replicates per sample were retained.

217 *Taxonomic assignment*

218 A first taxonomic assignment was performed to assess the proportion of reads assigned
219 to metazoans. For that purpose, each dataset was aligned against references retrieved from the
220 GenBank nt database using the ecotag command from OBITOOLS v-1.2.11 (Boyer et al. 2016)
221 with no minimum identity threshold. To perform taxonomic assignment at lower taxonomic
222 levels, assignment was later performed by aligning all ASVs against a restricted reference
223 database using the BLAST® command line tool (Altschul et al. 1990). This database was

224 composed of sequences retrieved from GenBank and sequences produced locally (Couton et
225 al. 2019), targeting ten metazoan phyla commonly found in marinas: Annelida, Arthropoda
226 (only Crustacea), Bryozoa, Chordata (only Actinopterygii, Chondrichthyes, and Tunicata),
227 Cnidaria, Echinodermata, Mollusca, Nemertea, Platyhelminthes, and Porifera. Only
228 alignments covering 99% of the subject sequence and displaying a minimum identity
229 threshold (18S: 99%; COI: 92%; 16S: 97%) based on barcoding gaps computed from
230 reference sequences (Figs. S1-S3), were considered. If one ASV matched with several
231 references, it was assigned to the one with the highest identity percentage. If two alignments
232 with different references had the same identity, the ASV was assigned to the lowest common
233 taxonomic rank. All assignments to a rank higher than the family were classified as
234 ‘unassigned’.

235 To evaluate the ability of water-eDNA-metabarcoding to retrieve species observed in
236 the field, the metabarcoding dataset was restricted to the four classes targeted in the quadrats
237 survey (ie Ascidiacea, Bivalvia, Gastropoda, and Gymnolaemata). These classes were chosen
238 because they are known to include a substantial number of NIS in marinas of the study region
239 (Bishop et al. 2015), and because the involved team members have a good expertise for their
240 morphological identification. This expertise is further reinforced by a number of previous
241 studies combining morphological and molecular approaches (Bishop et al. 2013; Bishop et al.
242 2015; Viard et al. 2019; Turon et al. 2020). In addition, for these four classes, local reference
243 data for native species and novel NIS were previously obtained (Couton et al. 2019). This
244 selection thus allowed minimizing issues due to the incompleteness of reference databases,
245 and better assess the effect of the taxonomic assignment step in the metabarcoding process (ie,
246 use of a curated database).

247

248 **Results**

249 *NIS observed in the field*

250 From the 400 quadrat samples collected, 93 taxa were identified, with differences
251 across sites and seasons (Table S4). Most (n=62) were identified to species, 20 to genus, 8 to
252 family, and 3 to higher levels. Forty-seven sessile or poorly mobile taxa belonged to the four
253 targeted classes (ie Ascidiacea, Bivalvia, Gastropoda, and Gymnolaemata). In these, 15 NIS
254 (incl. two pairs pooled in the field: *Botrylloides diegensis* and *B. violaceus*, and *Didemnum*
255 *vexillum* and *D. pseudovexillum*) and three cryptogenic species were identified (Table S4), all
256 previously reported in the study area. Among the 47 taxa observed, only 43 were considered
257 for assessing the ability of metabarcoding to detect them. *Ruditapes* sp. was removed from the
258 positive control list because it lives buried in soft sediments and is thus likely to be
259 inconsistently recovered in quadrat surveys. *Membraniporella nitida*, *Scruparia ambigua*, and
260 *Amathia gracilis* were excluded because they usually did not fulfil the chosen size criteria of 5
261 mm, and thus the observed species list might also be biased across locations.

262 *Taxonomic assignment of sequence reads*

263 The MiSeq runs yielded 15,754,233 reads, 13,108,322 reads, and 13,396,162 reads for
264 18S, COI, and 16S, respectively, resulting in 3972, 4258, and 347 unique ASVs. After
265 filtering, all samples (all tagged PCR replicates combined for each sample) contained a mean
266 of 253,600 reads, 188,853 reads, and 199,910 reads for 18S, COI, and 16S, respectively.

267 Only a small proportion of reads were assigned to Metazoa for 18S (23%) and COI
268 (24%), contrasting with the results for 16S (98%; Fig. 2). The most abundant metazoan phyla
269 identified across markers were Arthropoda, Chordata, Cnidaria, Annelida, and Mollusca,
270 totalizing 88%, 35% and 76% of the reads for 18S, COI and 16S, respectively, with sharp
271 differences between markers. For instance, bryozoans accounted for 14% of 16S reads but for
272 less than 1% of 18S and COI reads. Differences were also observed when assigning ASVs to
273 the ten targeted metazoan phyla. For instance, almost three times more taxa were identified
274 with 18S, and almost four times more species were identified with COI, as compared to 16S
275 (Table 1). Overall (60 2-L samples), 557 taxa including 409 species were identified by at least
276 one marker, over the ten selected phyla (Table S5-S8). Among the 154 taxa from the four
277 classes compared with morphological data, 114 were identified at the species-level across all
278 markers (Table S5).

279 ***Recovery of species known to be present by water-eDNA-metabarcoding***

280 Half of the 43 taxa (n=21) observed in quadrats were unambiguously recovered by
281 water-eDNA-metabarcoding (Fig. 3, column Tot). A higher number were identified with 18S
282 (14) as compared to COI (10) and 16S (5), with 18S being more efficient for Ascidiacea, and
283 the two latter being more suited for Bivalvia and Gymnolaemata (Fig. 3). Four species could
284 not be recovered because reference sequences were lacking for any of the three markers. Five
285 taxa were not identified because of identical reference sequences (ie lack of taxonomic
286 resolution with the markers used). For example, *Ascidiella aspersa* and *A. scabra* were
287 distinguished in the quadrats sampling, based on morphology, when metabarcoding only
288 identified them at the genus level, because of identical 18S references for the two species, and
289 amplification failure for this genus with COI and 16S (Figs. S4-S5).

290 Focusing on the 18 NIS and cryptogenic taxa known to be present, a large proportion
291 (66.7 %) were detected in eDNA with at least one marker (Table 2). Out of the six taxa not
292 recovered, four (*Botrylloides violaceus*, *Crassostrea gigas*, *Bugulina stolonifera*, and *B.*
293 *simplex*) had reference sequences identical to sequences of closely related species, resulting in
294 the assignment to the genus. For instance, *Bugulina stolonifera* and *B. simplex* share the same
295 sequence for the only marker able to retrieve them (16S). The other two NIS (*Didemnum*
296 *vexillum*, and *Tricellaria inopinata*) remained undetected despite the availability of references
297 for at least one marker, their large abundance in some locations, and their presence in almost
298 all localities at both seasons, suggesting PCR amplification biases with the markers used.

299 The recovery rate of NIS by water-eDNA-metabarcoding was much lower when
300 examining in details the distribution of taxa, across localities and seasons, with only 23.9 %
301 overlap (119 out of 498 cases; Fig. 3). Focusing on NIS or cryptogenic species, out of the 239
302 occurrences in quadrats, only 73 (30.5 %) were successfully recovered with water-eDNA-
303 metabarcoding (Fig. 3).

304 ***Additional detections with water-eDNA-metabarcoding***

305 Within the four targeted classes, 114 species were detected with water-eDNA-
306 metabarcoding. All species identified were classified as ‘expected’ vs. ‘unexpected’, based on
307 previous knowledge on their presence/absence in the study area, as reported in the Roscoff
308 biological species inventory (<http://abims.sb-roscoff.fr/inventaires/>) and in the literature, or as
309 reported before by the team members involved in this study. Eighteen species were
310 ‘unexpected’ species, and thus might be novel NIS for the study area. Sixteen of them were
311 assigned with more than 99% identity (Table S9). All were, however, identified with only one

312 marker, except *Mytilus trossulus*, identified with COI and 16S. None of them were seen in the
313 quadrats.

314 Regarding the distribution of taxa across localities, water-eDNA-metabarcoding
315 identified taxa in localities where they were not reported based on quadrats sampling in eight
316 cases (eg *Ascidia mentula* in SM and BLO during spring; Fig. 3). Of these eight cases, only
317 three are associated to a NIS or a cryptogenic species (*Crepidula fornicata* in MB,
318 *Watersipora subatra* in ET, and *Bugulina fulva* in ET). Moreover, water-eDNA-
319 metabarcoding efficiently identified three *Molgula* species (*Molgula socialis*, *M. bleizi*, and
320 *M. complanata*) when only the genus could be confidently recovered in the field. Finally, this
321 method allowed the identification of two additional NIS, *Mya arenaria* and *Ruditapes*
322 *philippinarum*, known to be present in the study region but living buried in soft sediments.

323 ***Effect of the sampling strategy on taxa detection by water-eDNA-metabarcoding***

324 Less than half of the taxa from the ten metazoan phyla (42.5 %) were shared between
325 the two pontoons sampled in a given locality with water-eDNA-metabarcoding (Fig. 4). The
326 proportion of shared taxa ranged between 22.6 % (CAM, fall) and 50.8 % (PG, fall).
327 Similarly, the total number of taxa was slightly higher when merging the results obtained with
328 the two 2-L replicates (482 taxa) than when using only one of them (436 and 428, for the two
329 series analysed separately), over all marinas, in the spring sampling. This result is consistent
330 at every pontoon, with a significantly higher number of taxa when using 4 L of water than
331 when using one of the two 2-L samples (Wilcoxon tests for paired data: $V = 0$, $P < 0.001$ for
332 the two series). When focusing on the four targeted classes, the number of taxa unique to one
333 2-L replicate ranged from 0 to 8 for 16S and 18S and 0 to 7 for COI (Fig. 5). The proportion
334 of reads assigned to these unique taxa was very low for 16S (ranged from 0 to 24 %; mean

335 value: 1.6 %), but was substantially higher for 18S (ranged from 0 to 66.1 %; mean value: 6.8
336 %), and even higher for COI (ranged from 0 to 100 %; mean value: 29.5 %). This proportion
337 was largely influenced by the total number of reads associated to all species from the four
338 target groups within each replicate for the different markers. These numbers ranged from
339 5,439 to 381,749 reads with 16S (mean: 116,582.1), 109 to 37,921 reads with 18S (mean:
340 8,370.6), and 1 to 6,191 reads with COI (mean: 1,764.7). For example, in the replicates of the
341 pontoon 2 from AW, only two species from the four groups were identified with COI, one
342 different in each replicate associated with 13 and 51 reads respectively (Fig. S6), leading to
343 the proportion of 100% (Fig. 5).

344

345 **Discussion**

346 The growing ‘hardening’ of marine coastal areas offers a wide range of new artificial
347 substrates for fouling organisms, many being non-indigenous species (NIS). Regular
348 monitoring of ports and marinas is thus critical for implementing efficient management
349 strategies (Trebitz et al. 2017). In that sense, species detection through water-eDNA-
350 metabarcoding is an appealing approach promising fast and easy sampling, and reduced costs.
351 However, its reliability and the extent of its biases when focusing on benthic NIS are still
352 unclear (Holman et al. 2019; Antich et al. 2021). In particular, one need to assess the risks of
353 both false negatives (which species are not detected and why; eg Hansen et al. 2018) and false
354 positives (detecting species which are not truly present; Darling et al. 2021). Here, the
355 effectiveness of water-eDNA-metabarcoding to identify the presence of biofouling animals
356 (particularly NIS) over the study area was shown, as it correctly recovered half of the 43 taxa,
357 and two third of the 18 NIS or cryptogenic species observed in the field. In addition, as

358 expected, water-eDNA-metabarcoding retrieved species from adjacent and non-targeted
359 micro-habitats, refined the assignment at the species level for 2 taxa identified at higher
360 taxonomic level in the field, and detected one previously unreported species, putatively a
361 novel NIS, in the study area. However, others species remained undetected in eDNA, with
362 important discrepancies locally. These limitations, and solutions to circumvent them, are
363 discussed below.

364 ***Improving the quality of reference data and targeting several markers is needed to increase***
365 ***the detection effectiveness***

366 Half of the species observed in quadrats, and one-third of NIS, were not recovered
367 with water-eDNA-metabarcoding, including abundant ones. For instance, the solitary tunicate
368 *Asciidiella aspersa*, native in the study area, and invasive in other coastal areas in the world,
369 was found in every quadrat sample, reaching up to 384 individuals m⁻² in MB (Table S4), but
370 remained unassigned with metabarcoding. The detection failure was due to both a lack of
371 resolution with 18S for this genus (same sequence as *A. scabra*), and amplification failures for
372 COI and 16S (Figs. S4-S5). Six other abundant NIS (tunicates: *Botrylloides violaceus*,
373 *Didemnum vexillum*; bryozoans: *Tricellaria inopinata*, *Bugulina simplex*, and *Bugulina*
374 *stolonifera*; bivalves: *Crassostrea gigas*) also remained undetected in the molecular dataset.
375 At a first glance, these results contrast with those by Antich et al. (2021) in which abundant
376 benthic MOTUs were more easily detected in water-eDNA than rare ones. However, in their
377 study, the benthic data were obtained from bulk metabarcoding, and abundant benthic species
378 might have also been missed because of amplification biases or low marker resolution.

379 The substantial proportion of false negatives observed highlights flaws that may still
380 impede the routine use of water-eDNA-metabarcoding. Those include species-specific DNA

381 shedding rate and decay (Andruszkiewicz et al. 2021; Holman et al. 2021), differential primer
382 binding efficiency between taxa (Piñol et al. 2018), and issues related to taxonomic
383 assignment. As several studies targeting benthic species showed the occurrence of their DNA
384 in water using species-specific methods (eg Simpson et al. 2017) their shedding rate must be
385 sufficient for detection. This is why more effort should be dedicated to design new primers
386 targeting markers with high resolution, where the sensitivity of each primer should be
387 reported to identify easily species that cannot be detected with them (Xia et al. 2021). For
388 instance, although abundant in quadrats from every location, the invasive colonial ascidian
389 *Didemnum vexillum*, was not recovered in water-eDNA, whereas Simpson et al. (2017)
390 detected it in water samples from Australia using species-specific primers. Similarly, the
391 invasive tunicate *B. violaceus* remained undetected because of identical sequences with
392 congeneric species, although it can be distinguished from the introduced *B. diegensis* and
393 native *B. leachii* using genus-specific primers (Couton et al. 2021). Issues related to
394 taxonomic resolution/accuracy were here (partly) addressed by combining several markers, as
395 suggested by Zhang et al. (2018), and by producing reference sequences for at least one
396 marker for each NIS observed within quadrats.

397 The availability of reference sequences is another well-known problem, which
398 depends on the marker, COI being the most populated in public databases (Porter and
399 Hajibabaei 2018), or on the taxonomic group. For instance, a barcode search in BOLD
400 revealed that only 14% and 5% of the 2,925 and 5,651 accepted species, for the classes
401 Ascidiacea and Gymnolaemata, respectively, were associated to a COI sequence (WoRMS
402 and BOLD search in August 2020), a proportion that could even be lower when considering
403 quality (Paz and Rinkevich 2021). This proportion seems higher for NIS: Briski et al. (2016)
404 reported proportions of 59% and 72% of 45 and 36 NIS, for the same two classes and marker,

405 respectively. These higher values might be explained by the long-lasting and still increasing
406 use of DNA barcoding in the study of marine NIS (Duarte et al. 2020). Such an imbalance in
407 the proportion of sequences available for native and non-indigenous species might, in turn,
408 increase the risk of false detections of new introductions, as will be discussed below.

409 ***Improving sampling strategy, data production and data processing to overcome limitations***

410 Increasing the sampling effort and optimizing data processing can substantially
411 decrease the detection failures of water-eDNA-metabarcoding. Sampling water at two
412 pontoons in each marina allowed increasing (sometimes doubling) the number of metazoan
413 taxa recovered. For example, in the three marinas where it was detected with water-eDNA-
414 metabarcoding in spring (BL, CAM, and SM), *Asterocarpa humilis* was identified from one
415 pontoon only. Furthermore, doubling the volume of water filtered improved the detection of
416 rare species (ie low number of reads), which might be particularly useful for NIS early
417 detection. Although the sampling effort (volume and number of replicates) is within the range
418 of that recommended and/or mostly used in other studies (eg Grey et al. 2018; Holman et al.
419 2019; Duarte et al. 2021), it was not sufficient to recover all the species observed in the field.
420 The sampling standards (eg water volume, sampling points) should thus be increased
421 substantially when targeting specific benthic species, such as particular NIS, because their
422 DNA might be scarce in the water, especially those lacking pelagic stages (Leduc et al. 2019).
423 This would still keep reduced labor- and time-costs in the field as compared to traditional
424 methods (sampling and analyzing 400 quadrats in the field required eight people over two
425 weeks per season). The results also underlined the importance of seasonal sampling, a much
426 less addressed issue. Because many biofoulers, such as tunicates or erected bryozoans, are
427 short-lived, their abundance can vary greatly depending on the season. Their DNA shedding
428 and decay rates can also differ throughout the year (eg reproduction period; Salter 2018;

429 Wacker et al. 2019). The tunicate *Ciona robusta*, for example, would not have been identified
430 from spring samples only, which is congruent with what is known about its settlement period
431 (mostly in late summer/autumn) in the study area (Bouchemousse et al. 2017). In this regard,
432 sampling effort should be maximized, with an emphasis on spatial and temporal replicates.

433 Data filtering steps are crucial in any metabarcoding analysis, notably to avoid false
434 positives, but it may decrease detection sensitivity. Here, all taxa identified were retained after
435 index-jump and replicate filtering in the whole dataset, with three exceptions (*Musculus* sp.,
436 *Bugulina stolonifera*, and *B. simplex*; Figs. S7-S8) but the detection capacity was much lower
437 at the unit level (ie ‘species-site-date’ combination). Before any corrections, water-eDNA-
438 metabarcoding recovered a given taxon in 41% of all units. This number dropped to 22% after
439 corrections, with 106 units of water-eDNA-metabarcoding detection disappearing.
440 Considering NIS or cryptogenic species, this proportion decreased from 56% to 30% (Figs.
441 S7-S8). PCR replicate filtering, used to limit false positives due to PCR errors (Alberdi et al.
442 2018; Tsuji et al. 2020), resulted in 32% loss of detection (Figs. S4-S5), while most losses
443 (68%) occurred after index-jump filtering. Choosing filtering parameters is thus critical and
444 depends on the study aim. For example, von Ammon et al. (2018) kept all OTUs, even
445 singletons, to increase their chance of detecting rare species, advocating that the cost of
446 missing even one rare NIS would be higher than that of a false positive. This option might,
447 however, lead to a dramatic increase in the number of false detections due to index-jump.
448 Indeed, in the dataset, all index-jump controls exhibited reads belonging to several ASVs.
449 Index-jump is an important issue very delicate to account for in metabarcoding studies
450 (Taberlet et al. 2012), and more work needs to be done to find the right compromise between
451 the proportion of false positives and false negatives in the context of benthic NIS detection.
452 Another way to circumvent filtering issues would be to increase the proportion of reads

453 associated to the target species. Most filtering procedures indeed rely on the fact that errors
454 are in low abundance. In most studies, ours included, using universal primers (Coward et al.
455 2015; Grey et al. 2018; Collins et al. 2019), the vast majority of either 18S or COI reads were
456 not assigned to metazoans (in this study, 77% and 76%, respectively). As a consequence, the
457 DNA from many biofoulers was likely not amplified in the first place. For example, each
458 species detected within the four targeted classes represented less than 1% of read abundances,
459 except for *Ciona intestinalis* with 18S (2.9%; Table S9). More effort should thus be directed
460 toward the development of new primer sets for more restricted groups of organisms that could
461 be multiplexed to increase the taxonomic range (Zhang et al. 2018; Corse et al. 2019; West et
462 al. 2020).

463 ***Erroneous species assignment challenges the opportunity to detect novel introduced species***

464 One expected outcome of water-eDNA-metabarcoding is the detection of species new
465 to a given area, either introduced or expanding their current distribution range, and not
466 actively looked for (ie passive surveillance). Over the four targeted classes, 18 species yet
467 unreported in the study area were recovered (Table S9). Most were, however, likely the result
468 of erroneous assignments. For example, the ascidian *Clavelina meridionalis*, known from the
469 southern hemisphere and yet unreported in Europe, was identified by 18S only, with 99.16 %
470 identity. No reference sequence, however, was available for the closely-related native
471 *C. lepadiformis*, observed in the field and identified with COI (100% identity) that reliably
472 distinguishes these species (Pérez-Portela and Turon 2008). In total, eight species were
473 identified with 18S only, including six showing identity below 100%. Considering the low
474 taxonomic resolution of 18S, and that most of them have native congeners in the study area,
475 these eight species were not considered as novel introduced species, in the absence of other
476 evidence. Three other species were assigned with 100% identity with COI, and might thus be

477 true identifications, yet still not necessarily novel introduced species. Indeed, the tunicate
478 *Polycarpa tenera* might have been previously misreported in the study area as the
479 Mediterranean species *P. gracilis* (Monniot 1974; Vazquez et al. 1995). Similarly, the
480 gastropod *Cerithiopsis petanii*, described from the Croatian coast (Prkić and Mariottini 2009),
481 belongs to the *C. tubercularis* species complex (Modica et al. 2013), reported in the study
482 area. It could have thus been overlooked by previous surveys. The only likely new record
483 would then be *Haminoea orteai*, reported in the Mediterranean Sea and the southern part of
484 the Northeast Atlantic (Garabedian et al. 2017).

485 Knowing the evolutionary history of the targeted species is also important to avoid
486 false positives. In this study's dataset, ASVs from both mitochondrial markers were assigned
487 to *Mytilus trossulus* although very similar to *M. edulis* references (only one gap). The
488 reference sequences for *M. trossulus* corresponded to individuals from the Baltic Sea, a well-
489 known hybrid zone between *M. trossulus* and *M. edulis*, where the mitochondrial genome of
490 *M. trossulus* has been fully introgressed by the mitochondrial genome of *M. edulis* (Kijewski
491 et al. 2006). Aligning the COI ASVs with sequences of *M. trossulus* from Scotland, which
492 have retained their ancestral mitochondrial genome (Zbawicka et al. 2010), led to 83.4%
493 identity only, confirming they were not *M. trossulus* but *M. edulis*.

494 These case studies demonstrate that passive detection of novel NIS by using water-
495 eDNA-metabarcoding could lead to false positives, as pointed by Darling et al. (2020).
496 Therefore, results need to be confirmed using alternative methods, either traditional and/or
497 more sensitive and targeted molecular approaches such as individual barcoding, qPCR or
498 digital PCR (Borrell et al. 2018; Harper et al. 2018; Wood et al. 2019; Ardura et al. 2021).

499 ***Benefitting from metabarcoding of water samples to monitor a wider range of habitats***

500 While water-eDNA-metabarcoding did not recover all species observed, it detected
501 many more taxa by encompassing a wider range of environments. Marinas are composed of
502 diverse micro-habitats, such as pillars, seawalls, floating pontoons, and soft-sediment
503 bottoms, displaying different faunal assemblages and NIS contributions (Leclerc et al. 2020),
504 which all contribute to the local pool of eDNA. Traditional standardized surveillance
505 protocols (eg Hewitt and Martin 2001; HELCOM and OSPAR 2013) usually target water,
506 sediment, and various hard substrates to encompass the whole diversity of marinas, but with
507 important time and labor costs. This study showed that water-eDNA-metabarcoding can
508 provide data over a wide range of habitats. Considering the 96 expected species identified
509 with water-eDNA-metabarcoding over the four target classes (Table S9), only 35 were
510 biofoulers, four might eventually live fixed (occasionally producing a byssus; *Aequipecten*
511 *opercularis*, *Limaria hians*, *Mimachlamys varia*, and *Venerupis corrugata*), 28 were soft-
512 bottom species, and 30 were motile gastropods. Focusing on NIS, the bivalves *Mya arenaria*
513 and *Ruditapes philippinarum*, known from the study area, were recovered by water-eDNA-
514 metabarcoding, while neither targeted nor expected in quadrats because of their burrowing
515 lifestyle.

516 A wide variety of management strategies exists, depending on the targeted species but
517 also on the phase of the introduction process this species has reached (Robertson et al. 2020).
518 These include measures on transport vectors to limit new arrivals, containment measures to
519 avoid the spread of already established NIS, or suppression or eradication measures to limit
520 the impact of introduced NIS. Whatever the management strategy implemented, however,
521 supporting activities (*sensu* Robertson et al. 2020) are required. In particular there is a dire
522 need to identify NIS and monitor them over time and space, either to report new arrivals or to

523 investigate their spread and putative impact of reported ones. According to the results, water-
524 eDNA-metabarcoding could be a perfect tool for monitoring species in this context. The wide
525 variety of targeted habitats and the ease of sampling allow both for the investigation of NIS
526 spread and putative impact on the communities, as well as allowing for regular sampling.

527 As a conclusion, the fact that two thirds of NIS present at the time of water sampling
528 were effectively recovered by water-eDNA-metabarcoding supports that this technique is a
529 valuable fast and cost-efficient monitoring method, although not yet ready to be used as a
530 unique tool for marine NIS passive surveillance programs. This results also show that species
531 detection errors (false positives or false negatives) can be limited thanks to the use of multiple
532 markers, the design of suitable primer sets (as performed in Westfall et al. 2020), an increased
533 sampling effort, and the building of reliable curated reference databases, including databases
534 specific to NIS (Darling et al. 2017).

535

536 **Acknowledgments**

537 We thank the Diving and Marine core service from the Roscoff Biological Station,
538 particularly Yann Fontana, Wilfried Thomas and Mathieu Camusat for the quadrat sampling,
539 and Jérôme Coudret and Stéphane Loisel for their help for the field work. We thank Gwenn
540 Tanguy from the Biogenouest Genomer Platform for advice and access to the sequencing
541 facilities, and the Biogenouest ABIMS Platform for access to the calculation and storage
542 resources. This project was supported by the TOTAL Foundation (project AquaNIS2.0). MC
543 acknowledges a PhD grant by Région Bretagne (ENIGME ARED project) and Sorbonne
544 Université (ED 227 ‘Sciences de la Nature et de l’Homme’).

545 **Disclosure statement**

546 The authors report that there are no competing interests to declare.

547 **Data availability statement**

548 Raw sequences were deposited on NCBI SRA: BioProject PRJNA744352. The

549 metabarcoding and analyses scripts used in this study are available at

550 https://github.com/joarwrie/benthic_NIS_detection

551 **References**

- 552 Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K. 2018. Scrutinizing key steps for reliable
553 metabarcoding of environmental samples. *Methods Ecol Evol.* 9(1):134-147.
- 554 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search
555 tool. *J Mol Biol.* 215(3):403-410.
- 556 Andruszkiewicz EA, Zhang WG, Lavery AC, Govindarajan AF. 2021. Environmental DNA
557 shedding and decay rates from diverse animal forms and thermal regimes. *Environ DNA.*
558 3(2):492-514.
- 559 Antich A, Palacín C, Cebrian E, Golo R, Wangensteen OS, Turon X. 2021. Marine
560 biomonitoring with eDNA: Can metabarcoding of water samples cut it as a tool for surveying
561 benthic communities? *Mol Ecol.* 30:3175-3188.
- 562 Ardura A, Rick J, Martinez JL, Zaiko A, Garcia-Vazquez E. 2021. Stress resistance for
563 unraveling potential biopollutants. Insights from ballast water community analysis through
564 DNA. *Mar Pollut Bull.* 163:111935.
- 565 Ardura A, Zaiko A, Martinez JL, Samulioviene A, Semenova A, Garcia-Vazquez E. 2015.
566 eDNA and specific primers for early detection of invasive species – A case study on the
567 bivalve *Rangia cuneata*, currently spreading in Europe. *Mar Environ Res.* 112:48-55.
- 568 Bishop JDD, Roby C, Yunnice ALE, Wood CA, Lévêque L, Turon X, Viard F. 2013. The
569 Southern Hemisphere ascidian *Asterocarpa humilis* is unrecognised but widely established in
570 NW France and Great Britain. *Biol Invasions.* 15(2):253-260.
- 571 Bishop JDD, Wood CA, Lévêque L, Yunnice ALE, Viard F. 2015. Repeated rapid assessment
572 surveys reveal contrasting trends in occupancy of marinas by non-indigenous species on
573 opposite sides of the western English Channel. *Mar Pollut Bull.* 95(2):699-706.
- 574 Blackman R, Mächler E, Altermatt F, Arnold A, Beja P, Boets P, Egeter B, Elbrecht V, Filipe
575 A, Jones J et al. 2019. Advancing the use of molecular methods for routine freshwater
576 macroinvertebrate biomonitoring – the need for calibration experiments. *Metabarcoding and*
577 *Metagenomics.* 3:e34735.
- 578 Borrell YJ, Miralles L, Do Huu H, Mohammed-Geba K, Garcia-Vazquez E. 2017. DNA in a
579 bottle—Rapid metabarcoding survey for early alerts of invasive species in ports. *PLOS ONE.*
580 12(9):e0183347.
- 581 Borrell YJ, Miralles L, Martínez-Marqués A, Semeraro A, Arias A, Carleos CE, García-
582 Vázquez E. 2018. Metabarcoding and post-sampling strategies to discover non-indigenous
583 species: A case study in the estuaries of the central south Bay of Biscay. *Journal for Nature*
584 *Conservation.* 42:67-74.
- 585 Bouchemousse S, Lévêque L, Viard F. 2017. Do settlement dynamics influence competitive
586 interactions between an alien tunicate and its native congener? *Ecol Evol.* 7(1):200-213.

- 587 Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E. 2016. OBITOOLS: a UNIX-
588 inspired software package for DNA metabarcoding. *Mol Ecol Resour.* 16(1):176-182.
- 589 Briski E, Ghabooli S, Bailey SA, MacIsaac HJ. 2016. Are genetic databases sufficiently
590 populated to detect non-indigenous species? *Biol Invasions.* 18(7):1911-1922.
- 591 Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2:
592 High-resolution sample inference from Illumina amplicon data. *Nat Methods.* 13:581-583.
- 593 Carney KJ, Minton MS, Holzer KK, Miller AW, McCann LD, Ruiz GM. 2017. Evaluating
594 the combined effects of ballast water management and trade dynamics on transfers of marine
595 organisms by ships. *PLOS ONE.* 12(3):e0172468.
- 596 Collins RA, Bakker J, Wangensteen OS, Soto AZ, Corrigan L, Sims DW, Genner MJ,
597 Mariani S. 2019. Non-specific amplification compromises environmental DNA
598 metabarcoding with COI. *Methods Ecol Evol.* 10(11):1985-2001.
- 599 Cordier T, Frontalini F, Cermakova K, Apothéoz-Perret-Gentil L, Treglia M, Scantamburlo
600 E, Bonamin V, Pawlowski J. 2019. Multi-marker eDNA metabarcoding survey to assess the
601 environmental impact of three offshore gas platforms in the North Adriatic Sea (Italy). *Mar*
602 *Environ Res.* 146:24-34.
- 603 Corse E, Tougard C, Archambaud-Suard G, Agnès J-F, Messu Mandeng FD, Bilong Bilong
604 CF, Duneau D, Zinger L, Chappaz R, Xu CCY et al. 2019. One-locus-several-primers: A
605 strategy to improve the taxonomic and haplotypic coverage in diet metabarcoding studies.
606 *Ecol Evol.* 9(8):4603-4620.
- 607 Couton M, Baud A, Daguin-Thiébaud C, Corre E, Comtet T, Viard F. 2021. High-throughput
608 sequencing on preservative ethanol is effective at jointly examining infraspecific and
609 taxonomic diversity, although bioinformatics pipelines do not perform equally. *Ecol Evol.*
610 11:5533-5546.
- 611 Couton M, Comtet T, Le Cam S, Corre E, Viard F. 2019. Metabarcoding on planktonic larval
612 stages: an efficient approach for detecting and investigating life cycle dynamics of benthic
613 aliens. *Manage Biol Inv.* 10(4):657-689.
- 614 Cowart DA, Pinheiro M, Mouchel O, Maguer M, Grall J, Miné J, Arnaud-Haond S. 2015.
615 Metabarcoding is powerful yet still blind: a comparative analysis of morphological and
616 molecular surveys of seagrass communities. *PLOS ONE.* 10(2):e0117562.
- 617 Dafforn KA, Johnston EL, Glasby TM. 2009. Shallow moving structures promote marine
618 invader dominance. *Biofouling.* 25(3):277-287.
- 619 Darling JA, Galil BS, Carvalho GR, Rius M, Viard F, Piraino S. 2017. Recommendations for
620 developing and applying genetic tools to assess and manage biological invasions in marine
621 ecosystems. *Mar Policy.* 85:54-64.
- 622 Darling JA, Jerde CL, Sepulveda AJ. 2021. What do you mean by false positive? *Environ*
623 *DNA.* 3:879-883.

- 624 Darling JA, Pochon X, Abbott CL, Inglis GJ, Zaiko A. 2020. The risks of using molecular
625 biodiversity data for incidental detection of species of concern. *Divers Distrib.* 26(9):1116-
626 1121.
- 627 Duarte S, Vieira P, Costa F. 2020. Assessment of species gaps in DNA barcode libraries of
628 non-indigenous species (NIS) occurring in European coastal regions. *Metabarcoding and*
629 *Metagenomics.* 4:e55162.
- 630 Duarte S, Vieira PE, Lavrador AS, Costa FO. 2021. Status and prospects of marine NIS
631 detection and monitoring through (e)DNA metabarcoding. *Sci Total Environ.* 751:141729.
- 632 Fernandez S, Miller DL, Holman LE, Gittenberger A, Ardura A, Rius M, Mirimin L. 2021.
633 Environmental DNA sampling protocols for the surveillance of marine non-indigenous
634 species in Irish coastal waters. *Mar Pollut Bull.* 172:112893.
- 635 Fonseca V, Carvalho G, Sung W, Johnson HF, Power D, Neill S, Packer M, Blaxter ML,
636 Lamshead PJ, Thomas W et al. 2010. Second-generation environmental sequencing unmasks
637 marine metazoan biodiversity. *Nat Commun.* 1:98.
- 638 Gallon RK, Robuchon M, Leroy B, Le Gall L, Valero M, Feunteun E. 2014. Twenty years of
639 observed and predicted changes in subtidal red seaweed assemblages along a biogeographical
640 transition zone: inferring potential causes from environmental data. *J Biogeogr.* 41(12):2293-
641 2306.
- 642 Garabedian K, Malaquias M, Crocetta F, Zenetos A, Kavadas S, Valdés Á. 2017. *Haminoea*
643 *ortei* Talavera, Murillo and Templado, 1987 (Mollusca: Gastropoda: Heterobranchia:
644 Cephalaspidea), a widespread species in the Mediterranean and northeastern Atlantic. *Cah*
645 *Biol Mar.* 58:107-113.
- 646 Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Spear SF,
647 McKee A, Oyler-McCance SJ, Cornman RS et al. 2016. Critical considerations for the
648 application of environmental DNA methods to detect aquatic species. *Methods Ecol Evol.*
649 7(11):1299-1307.
- 650 Grey EK, Bernatchez L, Cassey P, Deiner K, Deveney M, Howland KL, Lacoursière-Roussel
651 A, Leong SCY, Li Y, Olds B et al. 2018. Effects of sampling effort on biodiversity patterns
652 estimated from environmental DNA metabarcoding surveys. *Sci Rep.* 8(1):8843.
- 653 Hansen BK, Bekkevold D, Clausen LW, Nielsen EE. 2018. The sceptical optimist: challenges
654 and perspectives for the application of environmental DNA in marine fisheries. *Fish Fish.*
655 19(5):751-768.
- 656 Harper LR, Lawson Handley L, Hahn C, Boonham N, Rees HC, Gough KC, Lewis E, Adams
657 IP, Brotherton P, Phillips S et al. 2018. Needle in a haystack? A comparison of eDNA
658 metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*).
659 *Ecol Evol.* 8:6330-6341.
- 660 HELCOM, OSPAR. 2013. Joint Harmonised Procedure for the Contracting Parties of OSPAR
661 and HELCOM on the Granting of Exemptions Under International Convention for the Control
662 and Management of Ships' Ballast Water and Sediments. Regulation A-4.

- 663 Hewitt CL, Martin RB. 2001. Revised protocols for baseline port surveys for introduced
664 marine species: survey design sampling protocols and specimen handling. Hobart, Australia:
665 Centre for research on introduced marine pests. Technical report no. 22.
- 666 Holman LE, Chng Y, Rius M. 2021. How does eDNA decay affect metabarcoding
667 experiments? *Environ DNA*. 4(1):108-116.
- 668 Holman LE, de Bruyn M, Creer S, Carvalho G, Robidart J, Rius M. 2019. Detection of
669 introduced and resident marine species using environmental DNA metabarcoding of sediment
670 and water. *Sci Rep*. 9(1):11559.
- 671 Kelly RP, O'Donnell JL, Lowell NC, Shelton AO, Samhoury JF, Hennessey SM, Feist BE,
672 Williams GD. 2016. Genetic signatures of ecological diversity along an urbanization gradient.
673 *PeerJ*. 4:e2444.
- 674 Kijewski TK, Zbawicka M, Väinölä R, Wenne R. 2006. Introgression and mitochondrial
675 DNA heteroplasmy in the Baltic populations of mussels *Mytilus trossulus* and *M. edulis*. *Mar*
676 *Biol*. 149(6):1371-1385.
- 677 Koziol A, Stat M, Simpson T, Jarman S, DiBattista JD, Harvey ES, Marnane M, McDonald J,
678 Bunce M. 2018. Environmental DNA metabarcoding studies are critically affected by
679 substrate selection. *Mol Ecol Resour*. 19(2):366-376.
- 680 Leclerc J-C, Viard F, González Sepúlveda E, Díaz C, Neira Hinojosa J, Pérez Araneda K,
681 Silva F, Brante A. 2020. Habitat type drives the distribution of non-indigenous species in
682 fouling communities regardless of associated maritime traffic. *Divers Distrib*. 26(1):62-75.
- 683 Leduc N, Lacoursière-Roussel A, Howland KL, Archambault P, Sevellec M, Normandeau E,
684 Dispas A, Winkler G, McKindsey CW, Simard N et al. 2019. Comparing eDNA
685 metabarcoding and species collection for documenting Arctic metazoan biodiversity. *Environ*
686 *DNA*. 1(4):342-358.
- 687 Lehtiniemi M, Ojaveer H, David M, Galil B, Gollasch S, McKenzie C, Minchin D,
688 Occhipinti-Ambrogi A, Olenin S, Pederson J. 2015. Dose of truth—Monitoring marine non-
689 indigenous species to serve legislative requirements. *Mar Policy*. 54:26-35.
- 690 Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ.
691 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region
692 for metabarcoding metazoan diversity: application for characterizing coral reef fish gut
693 contents. *Front Zool*. 10(1):34.
- 694 Makiola A, Compson ZG, Baird DJ, Barnes MA, Boerlijst SP, Bouchez A, Brennan G, Bush
695 A, Canard E, Cordier T et al. 2020. Key questions for next-generation biomonitoring. *Front*
696 *Env Sci*. 7:197.
- 697 Modica MV, Mariottini P, Prkić J, Oliverio M. 2013. DNA-barcoding of sympatric species of
698 ectoparasitic gastropods of the genus *Cerithiopsis* (Mollusca: Gastropoda: Cerithiopsidae)
699 from Croatia. *J Mar Biol Assoc UK*. 93(4):1059-1065.

- 700 Monniot C. 1974. Ascidies littorales et bathyales récoltées au cours de la campagne Biaçores:
701 Phlébobranches et Stolidobranches. Bulletin du Museum National d'Histoire Naturelle 3ème
702 série. 251(173):1327-1352.
- 703 Morisette J, Burgiel S, Brantley K, Daniel WM, Darling J, Davis J, Franklin T, Gaddis K,
704 Hunter M, Lance R. 2021. Strategic considerations for invasive species managers in the
705 utilization of environmental DNA (eDNA): steps for incorporating this powerful surveillance
706 tool. *Manage Biol Inv.* 12(3):747-775.
- 707 Ojaveer H, Galil BS, Campbell ML, Carlton JT, Canning-Clode J, Cook EJ, Davidson AD,
708 Hewitt CL, Jelmert A, Marchini A et al. 2015. Classification of non-indigenous species based
709 on their impacts: considerations for application in marine management. *PLoS Biol.*
710 13(4):e1002130.
- 711 Outinen O, Puntilla-Dodd R, Barda I, Brzana R, Hegele-Drywa J, Kalnina M, Kostanda M,
712 Lindqvist A, Normant-Saremba M, Ścibik M et al. 2021. The role of marinas in the
713 establishment and spread of non-indigenous species in Baltic Sea fouling communities.
714 *Biofouling.* 37(9-10):984-997.
- 715 Paz G, Rinkevich B. 2021. Gap analysis of DNA barcoding in ERMS reference libraries for
716 ascidians and cnidarians. *Environ Sci Eur.* 33(1):4.
- 717 Pearman JK, von Ammon U, Laroche O, Zaiko A, Wood SA, Zubia M, Planes S, Pochon X.
718 2021. Metabarcoding as a tool to enhance marine surveillance of nonindigenous species in
719 tropical harbors: A case study in Tahiti. *Environ DNA.* 3(1):173-189.
- 720 Pérez-Portela R, Turon X. 2008. Phylogenetic relationships of the Clavelinidae and
721 Pycnoclavellidae (Ascidacea) inferred from mtDNA data. *Invertebr Biol.* 127(1):108-120.
- 722 Piñol J, Senar MA, Symondson WOC. 2018. The choice of universal primers and the
723 characteristics of the species mixture determine when DNA metabarcoding can be
724 quantitative. *Mol Ecol.* 28:407-419.
- 725 Porter TM, Hajibabaei M. 2018. Over 2.5 million COI sequences in GenBank and growing.
726 *PLOS ONE.* 13(9):e0200177.
- 727 Prkić J, Mariottini P. 2009. Description of two new *Cerithiopsis* from the Croatian coast, with
728 comments on the *Cerithiopsis tubercularis* complex (Gastropoda, Cerithiopsidae).
729 *Aldrovandia.* 5:3-27.
- 730 Rey A, Basurko OC, Rodriguez-Ezpeleta N. 2020. Considerations for metabarcoding-based
731 port biological baseline surveys aimed at marine nonindigenous species monitoring and risk
732 assessments. *Ecol Evol.* 10(5):2452-2465.
- 733 Robertson PA, Mill A, Novoa A, Jeschke JM, Essl F, Gallardo B, Geist J, Jarić I, Lambin X,
734 Musseau C et al. 2020. A proposed unified framework to describe the management of
735 biological invasions. *Biol Invasions.* 22(9):2633-2645.

- 736 Salter I. 2018. Seasonal variability in the persistence of dissolved environmental DNA
737 (eDNA) in a marine system: The role of microbial nutrient limitation. PLOS ONE.
738 13(2):e0192409.
- 739 Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, Pagad S, Pyšek P,
740 Winter M, Arianoutsou M et al. 2017. No saturation in the accumulation of alien species
741 worldwide. Nat Commun. 8:14435.
- 742 Shaw JLA, Weyrich LS, Hallegraeff G, Cooper A. 2019. Retrospective eDNA assessment of
743 potentially harmful algae in historical ship ballast tank and marine port sediments. Mol Ecol.
744 28(10):2476-2485.
- 745 Simpson TJS, Dias PJ, Snow M, Muñoz J, Berry T. 2017. Real-time PCR detection of
746 *Didemnum perlucidum* (Monniot, 1983) and *Didemnum vexillum* (Kott, 2002) in an applied
747 routine marine biosecurity context. Mol Ecol Resour. 17(3):443-453.
- 748 Spalding MD, Fox HE, Allen GR, Davidson N, Ferdaña ZA, Finlayson M, Halpern BS, Jorge
749 MA, Lombana A, Lourie SA et al. 2007. Marine ecoregions of the world: A
750 bioregionalization of coastal and shelf areas. Bioscience. 57(7):573-583.
- 751 Stat M, Huggett MJ, Bernasconi R, DiBattista JD, Berry TE, Newman SJ, Harvey ES, Bunce
752 M. 2017. Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a
753 tropical marine environment. Sci Rep. 7(1):12240.
- 754 Taberlet P, Bonin A, Zinger L, Coissac E. 2018. Environmental DNA: For Biodiversity
755 Research and Monitoring. Oxford: Oxford University Press.
- 756 Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH. 2012. Environmental DNA. Mol Ecol.
757 21(8):1789-1793.
- 758 Trebitz AS, Hoffman JC, Darling JA, Pilgrim EM, Kelly JR, Brown EA, Chadderton WL,
759 Egan SP, Grey EK, Hashsham SA et al. 2017. Early detection monitoring for aquatic non-
760 indigenous species: Optimizing surveillance, incorporating advanced technologies,
761 and identifying research needs. J Environ Manage. 202:299-310.
- 762 Tsuji S, Miya M, Ushio M, Sato H, Minamoto T, Yamanaka H. 2020. Evaluating intraspecific
763 genetic diversity using environmental DNA and denoising approach: A case study using tank
764 water. Environ DNA. 2(1):42-52.
- 765 Turon X, Casso M, Pascual M, Viard F. 2020. Looks can be deceiving: *Didemnum*
766 *pseudovexillum* sp. nov. (Asciidiacea) in European harbours. Mar Biodivers. 50(4):48.
- 767 van der Loos LM, Nijland R. 2021. Biases in bulk: DNA metabarcoding of marine
768 communities and the methodology involved. Mol Ecol. 30(13):3270-3288.
- 769 Vazquez E, Ramos-Espla AA, Turon X. 1995. The genus *Polycarpa* (Asciidiacea, Styelidae)
770 on the Atlantic and Mediterranean coasts of the Iberian Peninsula. J Zool. 237(4):593-614.
- 771 Viard F, David P, Darling JA. 2016. Marine invasions enter the genomic era: three lessons
772 from the past, and the way forward. Curr Zool. 62(6):629-642.

- 773 Viard F, Roby C, Turon X, Bouchemousse S, Bishop JDD. 2019. Cryptic diversity and
774 database errors challenge non-indigenous species surveys: an illustration with *Botrylloides*
775 spp. in the English Channel and Mediterranean Sea. *Front Mar Sci.* 6:615.
- 776 von Ammon U, Wood SA, Laroche O, Zaiko A, Tait L, Lavery S, Inglis GJ, Pochon X. 2018.
777 Combining morpho-taxonomy and metabarcoding enhances the detection of non-indigenous
778 marine pests in biofouling communities. *Sci Rep.* 8(1):16290.
- 779 Wacker S, Fossøy F, Larsen BM, Brandsegg H, Sivertsgård R, Karlsson S. 2019. Downstream
780 transport and seasonal variation in freshwater pearl mussel (*Margaritifera margaritifera*)
781 eDNA concentration. *Environ DNA.* 1(1):64-73.
- 782 West KM, Adam AAS, White N, Robbins WD, Barrow D, Lane A, T. Richards Z. 2021. The
783 applicability of eDNA metabarcoding approaches for sessile benthic surveying in the
784 Kimberley region, north-western Australia. *Environ DNA.* 4(1):34-49.
- 785 West KM, Stat M, Harvey ES, Skepper CL, DiBattista JD, Richards ZT, Travers MJ,
786 Newman SJ, Bunce M. 2020. eDNA metabarcoding survey reveals fine-scale coral reef
787 community variation across a remote, tropical island ecosystem. *Mol Ecol.* 29(6):1069-1086.
- 788 Westfall KM, Therriault TW, Abbott CL. 2020. A new approach to molecular biosurveillance
789 of invasive species using DNA metabarcoding. *Global Change Biol.* 26(2):1012-1022.
- 790 Wood SA, Pochon X, Laroche O, von Ammon U, Adamson J, Zaiko A. 2019. A comparison
791 of droplet digital polymerase chain reaction (PCR), quantitative PCR and metabarcoding for
792 species-specific detection in environmental DNA. *Mol Ecol Resour.* 19:1407-1419.
- 793 Xia Z, Zhan A, Johansson ML, DeRoy E, Haffner GD, MacIsaac HJ. 2021. Screening marker
794 sensitivity: Optimizing eDNA-based rare species detection. *Divers Distrib.* 27(10):1981-1988.
- 795 Zaiko A, Samuiloviene A, Ardura A, Garcia-Vazquez E. 2015. Metabarcoding approach for
796 nonindigenous species surveillance in marine coastal waters. *Mar Pollut Bull.* 100(1):53-59.
- 797 Zbawicka M, Burzyński A, Skibinski D, Wenne R. 2010. Scottish *Mytilus trossulus* mussels
798 retain ancestral mitochondrial DNA: Complete sequences of male and female mtDNA
799 genomes. *Gene.* 456(1):45-53.
- 800 Zhang GK, Chain FJJ, Abbott CL, Cristescu ME. 2018. Metabarcoding using multiplexed
801 markers increases species detection in complex zooplankton communities. *Evol Appl.*
802 11(10):1901-1914.
803

804 **Tables**

805 **Table 1** Number and proportion (in parentheses) of ASVs and reads assigned to one of the
 806 reference sequences from the restricted database targeting ten metazoan phyla, for each
 807 marker (18S, COI, and 16S). The number indicated for ‘Taxa’ includes assignments to
 808 species-, genus- and family-level. The last three columns indicate the number of species,
 809 genera and families detected.

	<i>ASVs</i>	<i>Reads</i>	<i>Taxa</i>	<i>Species</i>	<i>Genus</i>	<i>Family</i>
<i>18S</i>	389 (9.8%)	1,881,029 (15.8%)	291	203	233	159
<i>CO</i>	943 (22.1%)	706,023 (7.0%)	286	235	221	177
<i>I</i>						
<i>16S</i>	173 (49.9%)	7,311,390 (54.6%)	104	65	78	70

810

811

812 **Table 2** List of non-indigenous (NIS) and cryptogenic (Crypto) species within the four target classes,
813 known from the study area, and either observed in quadrats or identified in water-eDNA samples.
814 Stars indicate that, even if the species was not identified in the dataset, sequences were assigned either
815 to the genus because of a lack of resolution, or to another species of the same genus. When applicable,
816 the absence of reference sequence is indicated.

<i>Class</i>	<i>Species</i>	<i>Status</i>	<i>Observed in quadrats</i>	<i>Detected in eDNA samples</i>		
				<i>18S</i>	<i>COI</i>	<i>16S</i>
<i>Ascidacea</i>	<i>Asterocarpa humilis</i>	NIS	yes	yes	no	no
<i>Ascidacea</i>	<i>Botrylloides diegensis</i>	NIS	yes	yes	yes	no
<i>Ascidacea</i>	<i>Botrylloides violaceus</i>	NIS	yes	no*	no*	no
<i>Ascidacea</i>	<i>Botryllus schlosseri</i>	Crypto	yes	yes	no	no
<i>Ascidacea</i>	<i>Ciona robusta</i>	NIS	yes	yes	yes	no
<i>Ascidacea</i>	<i>Corella eumyota</i>	NIS	yes	yes	no	no
<i>Ascidacea</i>	<i>Didemnum vexillum</i>	NIS	yes	no	no	no
<i>Ascidacea</i>	<i>Diplosoma listerianum</i>	Crypto	yes	no	yes	no
<i>Ascidacea</i>	<i>Perophora japonica</i>	NIS	yes	yes	no	no
<i>Ascidacea</i>	<i>Styela clava</i>	NIS	yes	yes	no	no
<i>Bivalvia</i>	<i>Crassostrea gigas</i>	NIS	yes	no*	no	no
<i>Bivalvia</i>	<i>Mya arenaria</i>	NIS	no	yes	no	no*
<i>Bivalvia</i>	<i>Ruditapes philippinarum</i>	NIS	no	yes	yes	yes
<i>Gastropoda</i>	<i>Crepidula fornicata</i>	NIS	yes	no	yes	no
<i>Gymnolaemata</i>	<i>Bugula neritina</i>	NIS	yes	yes	yes	yes
<i>Gymnolaemata</i>	<i>Bugulina fulva</i>	Crypto	yes	no	yes	yes
<i>Gymnolaemata</i>	<i>Bugulina simplex</i>	NIS	yes	no	no*	no*
<i>Gymnolaemata</i>	<i>Bugulina stolonifera</i>	NIS	yes	no	no*	no*
<i>Gymnolaemata</i>	<i>Tricellaria inopinata</i>	NIS	yes	no reference	no	no reference
<i>Gymnolaemata</i>	<i>Watersipora subatra</i>	NIS	yes	yes	yes	no reference *

817

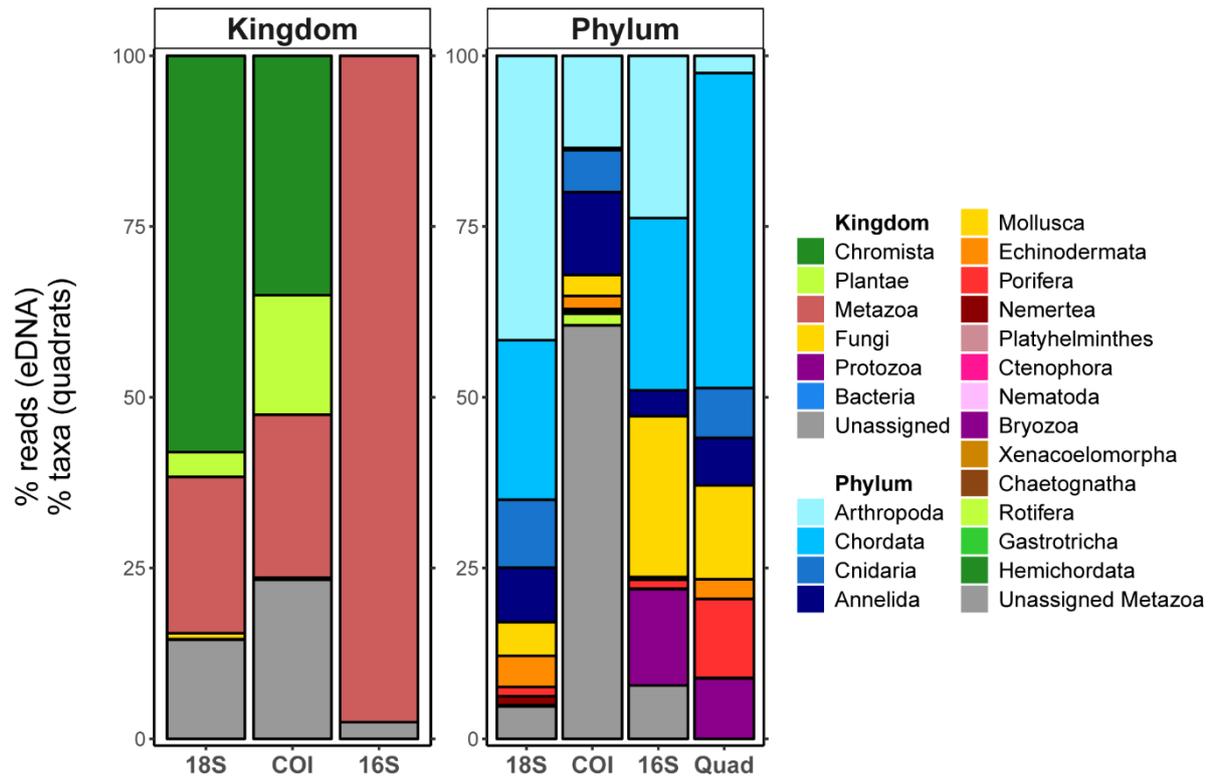
818 **Figures with caption**



819

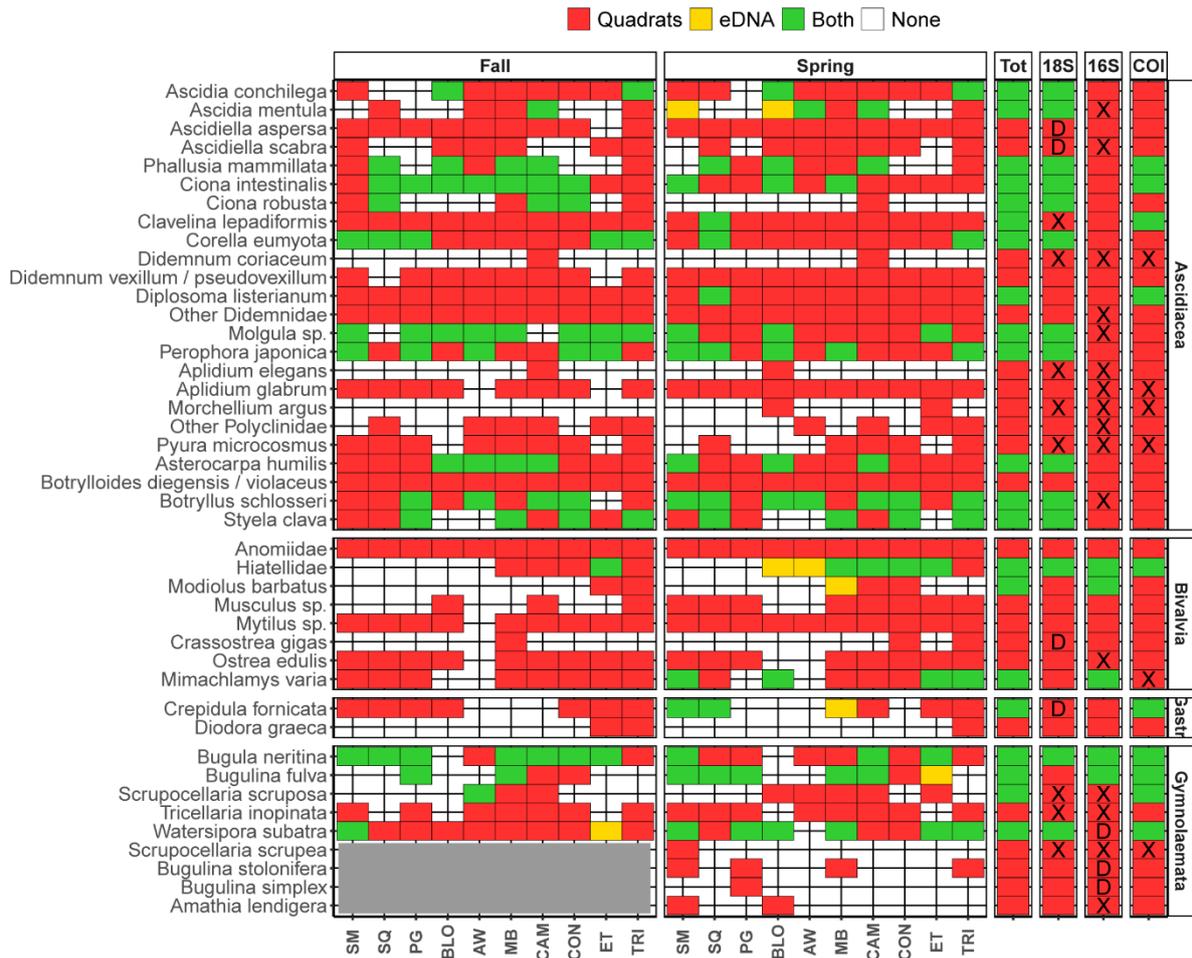
820 **Figure 1 A.** Location of the sampling sites. SM = Saint-Malo, SQ = Saint-Quay-Portrieux,
821 PG = Perros-Guirec, BLO = Bloscon-Roscoff, AW = L'Aber Wrac'h, MB = Moulin Blanc-
822 Brest, CAM = Camaret-sur-Mer, CON = Concarneau, ET = Étel, TRI = La Trinité-sur-Mer.
823 **B.** Picture of one of the study marina (SQ), showing sampled pontoons. **C.** Sampling method
824 (ie, diving under the pontoons) for morphological identification. **D.** Picture showing all the
825 specimens scraped in a quadrat after sorting on site. **E.** Seawater sampling method, with
826 seawater collected in a plastic jar using a sterile 100-mL syringe fixed to a sampling rod.
827 Photo credits: Frédérique Viard (B), Wilfried Thomas (C), Laurent Lévêque (D), and Yann
828 Fontana (E).

829

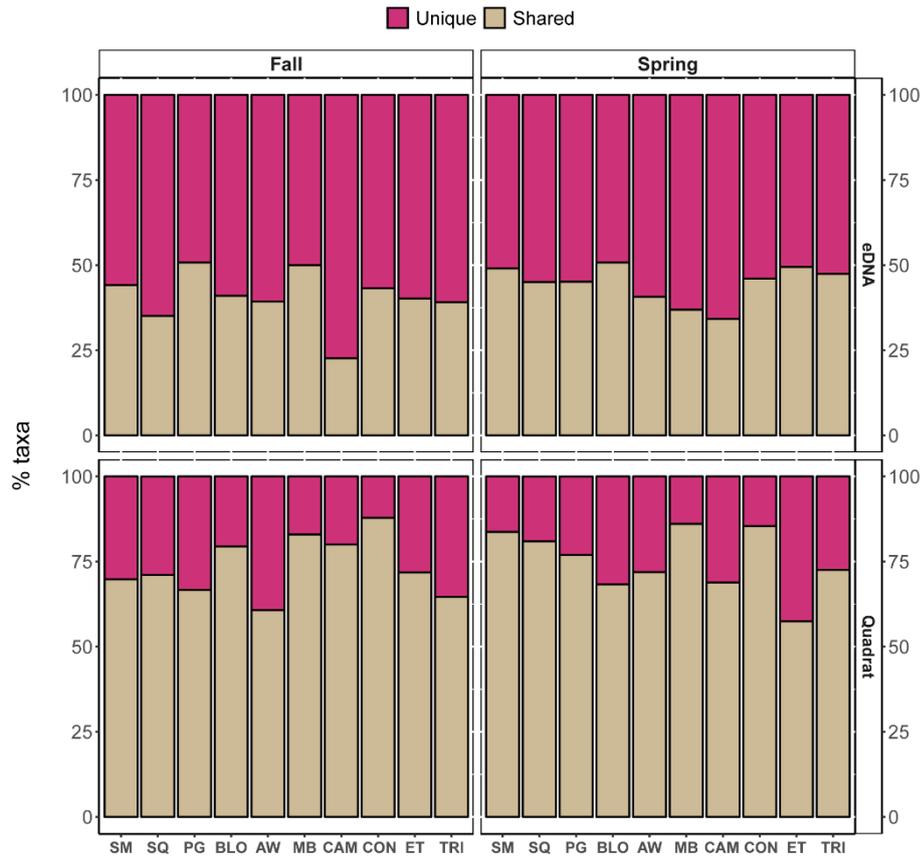


830 **Figure 2** Percentage of reads assigned to different kingdoms (left) or metazoan phyla (right),
 831 over all samples, for each of the three markers, as well as for quadrats data (Quad).

832



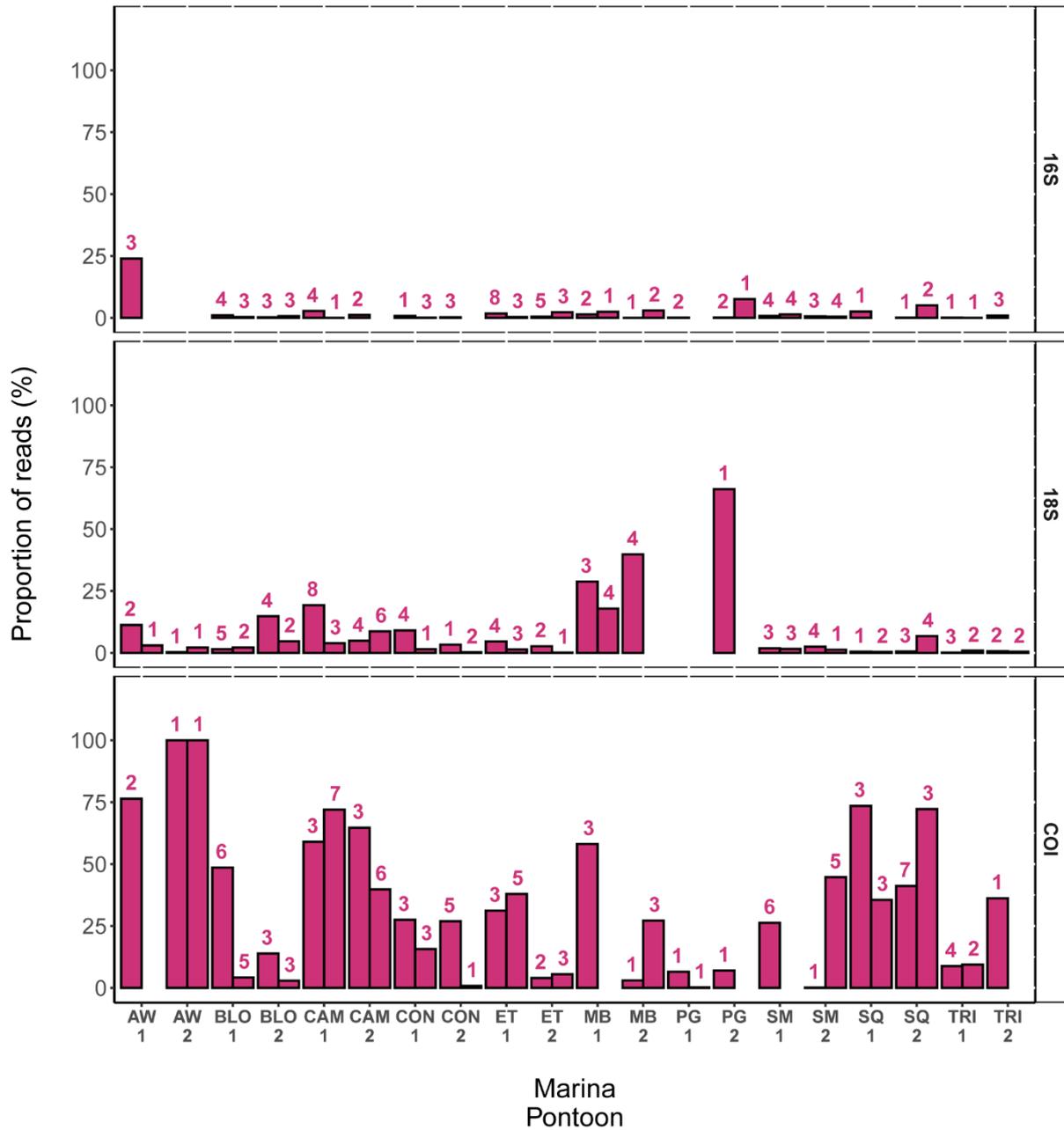
834 **Figure 3** Comparison between water-eDNA-metabarcoding and morphology-based detection.
 835 Only the taxa observed within quadrats and belonging to the four targeted classes are
 836 considered. Taxa found with both methods within the same marina are displayed in green,
 837 those found only in quadrats are in red, and those only found in water-eDNA are in yellow.
 838 The first two panels correspond to samples from fall 2017 and spring 2018, respectively.
 839 Metabarcoding detection is recorded across all markers. The column Tot represents the
 840 presence of each taxon in the whole dataset. The last three columns represent the presence of
 841 a taxon in the whole dataset (all marinas and dates combined) for each marker separately.
 842 Black crosses indicate the lack of reference sequences for a particular taxon and D indicates
 843 that this marker was not able to discriminate a particular taxon. The grey area represents
 844 species that were not isolated during morphological identification in fall 2017. They were all
 845 pooled in a unique category called ‘erected bryozoans’. See Fig. 1 for location codes.



847

848 **Figure 4** Proportion of taxa recovered with water-eDNA-metabarcoding across all markers or
 849 within quadrats, in samples from the two pontoons of a marina for a given season (shared) or
 850 in samples from only one pontoon (unique). See Fig. 1 for location codes.

851



852

853 **Figure 5** Proportion of reads associated to taxa found in only one of the two 2-L replicates
 854 (unique taxa) for each pontoon of each marina. This proportion is calculated by dividing the
 855 sum of reads for unique taxa by the sum of all reads for the four target groups in each
 856 replicate. The results for each replicate are shown separately (one bar per replicate). The
 857 number of unique taxa is indicated at the top of each bar. See Fig. 1 for location codes.
 858 Results are presented separately for each marker. Detailed and additional data (values per taxa
 859 unique to one replicate and for taxa shared between pontoons) are provided in Fig. S6.