bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Detection of novel and resident marine species using environmental DNA metabarcoding**
2 **of sediment and water**
3
4 Mr. Luke E. Holman[1]* (ORCID: 0000-0002-8139-3760)
5 Dr. Mark de Bruyn[2,3]
6 Prof. Simon Creer[2]
7 Prof. Gary Carvalho[2]
8 Dr. Julie Robidart[4]
9 Dr. Marc Rius[1,5] (ORCID: 0000-0002-2195-6605)
10
11 [1] School of Ocean and Earth Science, National Oceanography Centre Southampton, University
12 of Southampton, United Kingdom
13 [2] Molecular Ecology and Fisheries Genetics Laboratory, School of Natural Sciences, Bangor
14 University, United Kingdom
15 [3] The University of Sydney, School of Life and Environmental Sciences, Australia
16 [4] Ocean Technology and Engineering Group, National Oceanography Centre Southampton,
17 United Kingdom
18 [5] Centre for Ecological Genomics and Wildlife Conservation, University of Johannesburg, South
19 Africa
20
21
22 *Corresponding Author
23 Email: l.e.holman@soton.ac.uk
24
25
26

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

27 **Abstract**

28 Environmental DNA (eDNA) surveys are an increasingly important tool for biodiversity
29 monitoring, providing unprecedented levels of resolution and sensitivity. Nevertheless, eDNA
30 studies focussing on the detection of novel and resident species remain rare, and little is known
31 about the effects that different environmental sample types have on species detectability. Here
32 we evaluated the ability of eDNA metabarcoding to detect marine species by collecting sediment
33 and water samples and performing eDNA metabarcoding of 18S rRNA and COI genes. We
34 detected vastly different species assemblages between the studied environmental samples; with
35 sediment containing significantly higher number of operational taxonomic units than water. We
36 then compared the obtained datasets with previously published biodiversity surveys and found
37 excellent concordance among different survey techniques, as well as novel species
38 introductions. We conclude that careful consideration of type of environmental sample is needed
39 when conducting biodiversity monitoring using eDNA, especially for studies focusing on
40 community change.

41
42
43
44
45
46
47
48
49
50
51
52

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

53 **Introduction**

54 Anthropogenic activities are causing a global decrease in biodiversity[1,2] that negatively affects
55 ecosystem services and function[3]. Such impacts create an urgent need for tools that rapidly and
56 accurately monitor species diversity. Biodiversity surveys have been criticised for their lack of
57 standardisation and taxonomic resolution[4,5]. One approach that has the potential to overcome
58 these limitations is the use of nucleic acids found in environmental samples, such as water, soil
59 or sediment, to infer presence or absence of living organisms in the local environment[6]. This
60 genetic material, known as environmental DNA (hereafter eDNA), is a poly-disperse mixture of
61 tissue, cells, subcellular fragments and extracellular DNA lost to the environment by
62 organisms[7,8]. Studies using eDNA focus on targeted detection (single species) methods such as
63 qPCR[9-12], or community (multi-species) methods such as metabarcoding[13-15]. Samples can be
64 collected with minimal training and once the methodology is optimised, surveys are highly
65 amenable to automation [16]. Thus, eDNA surveys are highly informative (but see considerations
66 for ensuring validity and replicability[17]) and can complement other biodiversity monitoring
67 methods[18]. Recent work has identified a vast range of viable protocols for the collection,
68 extraction and detection of target nucleic acids from different environmental samples[19-21].
69 Despite this, little is known about the effects that different environmental sample types have on
70 species detectability using eDNA[22].

71 Biodiversity monitoring is normally performed by conducting surveys standardised by time and
72 by reaching a species discovery asymptote[23,24,25]. The results of such surveys often restrict to
73 species that are being targeted at the time, with no ability to retrospectively separate
74 erroneously grouped species in light of new discoveries. This is a critically important aspect for
75 biodiversity monitoring in the sea, mainly because between 9,000-35,000 marine species (2.7%
76 of the total number of estimated marine species) are considered cryptic (i.e., morphologically
77 similar but genetically distinct)[26]. Indeed, many widespread sessile marine species contain

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

78 cryptic lineages as revealed by genetic studies, highlighting the need for an integrated
79 morphological and genetic approach to accurately detect these species[27,28].
80 Non-indigenous species (NIS) are those that have been transported from their native range
81 through human action into a novel geographic location. The impacts of NIS pose a severe threat
82 to agriculture, human health and the conservation of ecosystems and biodiversity[29-32]. Most
83 marine NIS have spread globally via vectors such as transoceanic shipping or the construction
84 of canals interconnecting large water bodies[33,34]. At fine (10s of km) geographical scales, other
85 vectors such as intraregional boating significantly enhance the spread and subsequent impact of
86 NIS[35]. Along coastal areas, studies have highlighted the importance of monitoring marinas and
87 harbours[23], as these are hotspots of NIS and together with other marine infrastructure (e.g.
88 breakwaters, artificial reefs) promote the spread of NIS[36]. In these habitats, NIS often
89 outcompete native species and dominate artificial hard substrata[37,38]. Marinas and harbours
90 have distinct ecological and physico-chemical conditions compared to the surrounding natural
91 environment[39,40]. Consequently, there is a need for specific sampling and surveying protocols to
92 study both native and NIS in these ecologically distinct environments.
93 Here we used eDNA metabarcoding to examine how species detection differs between
94 environmental sample types collected in several distinct marinas. We first documented the
95 differences in both alpha and beta diversity from sediment and water samples. We then
96 compared the eDNA metabarcoding results with previously published biodiversity data to identify
97 key NIS that are both introduced in the study region and / or elsewhere. Subsequently, we
98 identified a number of previously unrecorded NIS in the study region. Finally we discuss the
99 strengths and weaknesses of eDNA metabarcoding for detecting marine NIS, and how this
100 technique can help conservation efforts for both preserving native biodiversity and mitigating the
101 deleterious effects of NIS.
102
103

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

104 **Results**

105 *Raw sequencing results and OTU generation*

106 Sequencing produced a total of 17.8 million paired end reads, with 15.2 million sequences
107 remaining after paired end read merging and quality filtering. The average number of sequences
108 per sample after filtering (excluding control samples) was 200,185 ± 64,019 (s.d). No template
109 control samples contained an average of 811 ± 3,402 (s.d) sequences. One control sample
110 contained ~15,000 sequences that mapped to an operational taxonomic unit (OTU) that had
111 100% identity match to a sequence of a terrestrial fungi (Genbank: FJ804151.1), excluding this
112 OTU gives an average of 51 ± 94 (s.d) sequences per no-template control sample. Denoising
113 produced 8,069 OTUs for COI and 2,433 for 18S with 6,435 and 1,679 remaining respectively
114 after LULU curation. Taxonomic annotation identified 622 OTUs from the 18S rRNA dataset
115 against the SILVA database and 481 OTUs from the COI dataset against the MIDORI database.
116 Taxonomic data from World Register of Marine Species could be retrieved for 200 of the
117 annotated COI OTUs and 190 of the 18S OTUs.
118

119 *Biodiversity detection*

120 The effect of different water eDNA sample preservation techniques differed between the target
121 amplicons. The 18S rRNA amplicon produced significantly more OTUs in samples preserved by
122 freezing compared to Longmire's preservation method, while the COI amplicon showed no
123 significant difference between preservation treatments (see Supplementary Information 2 for
124 details). As a conservative approach all subsequent analyses used sample data from frozen
125 samples. The minimum number of reads per sample was 137,624 for the COI dataset and
126 117,915 for the 18S dataset and so samples were rarefied to this number of reads. More OTUs
127 in total were detected in the sediment samples compared to the water samples across all sites
128 and both markers as shown in Figure 1b,d. In all cases, both water and sediment samples
129 detected unique OTUs but the mean proportion of unique OTUs detected in water was lower

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

130 (49.2%) in comparison to sediment (73.8%). A 2-way ANOVA testing the effect of eDNA type of
131 environmental sample on number of OTUs generated indicated a significant effect ($p<0.001$) of
132 sample type for both 18S rRNA and COI (See Supplementary Information 3 for full model
133 output). Ordination plots of Bray-Curtis dissimilarity (Fig. 1c,e) showed that OTUs in eDNA found
134 in sediment and water differ in community structure as much as among sites in ordination space.
135 Additionally, the PERMANOVA model indicated highly significant differences ($p<0.001$) between
136 sites and eDNA medium in both the 18S rRNA and COI datasets. Furthermore, eDNA detection
137 medium in the PERMANOVA model explained 23.2% and 32.5% of the variation in the 18S and
138 COI data respectively, while the site explained 34.2% and 30.5% in the COI and 18S rRNA data
139 (See Supplementary Information 4 for full model output). At phylum level (Figure 2), taxonomy
140 does not perfectly predict medium of detection, however a binomial goodness of fit test showed
141 non-random detection proportions in the Nematoda (Bonferroni corrected $p<0.001$), with eDNA
142 detections mostly in sediment.
143
144 *Detection of non-indigenous species*
145 In total 22 NIS to the study region and 29 species documented as NIS in other countries were
146 detected across the four sites (see Supplementary Table 1 for full list). Out of the detected NIS
147 seven were present in the list of 21 NIS previously detected in manual non-native invertebrate
148 surveys at the sites. As shown in Figure 3 the results of the eDNA surveys closely matched
149 those of the manual survey results. Four detections differed from the manual surveys, a single
150 eDNA detection not seen in RA Surveys and three RA detections not seen in eDNA surveys
151 (Figure 3). Remapping of cleaned reads from sites with incongruent detections to respective
152 COI regions (Genbank Accessions: *Austrominius modestus* KY607884; *Bugula neritina*
153 KY235450; *Ficopomatus enigmatus* KX840011) found hits for *Bugula neritina* only (5 reads from
154 a single replicate). These reads were lost during data filtering and so did not feature in the final
155 dataset. A detection of note was 199 reads from the sediment of Site TQ mapping to an OTU

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

156 corresponding with *Arcuatula senhousia* (Asian date mussel), a novel NIS for the UK. Targeted
157 visual surveys on tidal mudflats within two kilometres of Marina TQ confirmed the presence of
158 this species in proximity to the sampling site. Furthermore, we generated COI sequences from
159 these tissue samples (Genbank Accession: MH924820 and MH924821) and matched to known
160 *A. senhousia* sequences confirming the eDNA detection of this species (see Supplementary
161 Information 5 for details of DNA barcoding). Additionally, the nematode *Cephalothrix simula* and
162 the oligochaete *Paranais frici* were also detected using eDNA at site TQ. Both are novel species
163 introductions to the United Kingdom, previously undocumented in academic literature.
164
165 **Discussion**
166 We demonstrated that eDNA from sediment and water samples reveal very different community
167 composition, suggesting that the collection of multiple environmental sample types provides the
168 most comprehensive assessment of community composition. We also found that eDNA
169 metabarcoding shows concordance with published biodiversity surveys for the detection of NIS.
170 Furthermore, we demonstrated that eDNA metabarcoding can detect novel species
171 introductions, suggesting that eDNA surveys are an effective tool to significantly reduce the time
172 between introduction and detection of NIS, and in turn increase the likelihood of successful
173 control and eradication. Our study shows how different environmental sample types can affect
174 our understanding of both whole community composition and particular species of concern (e.g.
175 NIS).
176 The majority of research using eDNA to detect aquatic macrofauna has focused on the
177 collection of water samples, while sediment samples have received comparatively less attention.
178 Moreover, sediment has been shown to harbour 8-1800 times more eDNA compared to water
179 samples in freshwater ecosystems[41]. Here we found dramatic differences in species richness in
180 sediment and water samples, observing a consistently greater number of OTUs detected in
181 sediment compared to water. However, our results indicated the opposite trend when

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

182 considering only fish species; with more fish being detected in seawater samples compared to
183 sediment samples (6 in water, 1 in sediment and water). Similarly, Shaw *et al.*[22] found that
184 sediment 12S rRNA metabarcoding detected fewer fish compared to water in a freshwater lotic
185 environment. More broadly, taxonomy at the level of phylum did not predict if a species was
186 detected in water, sediment or both environmental sample types (except the Nematoda, whose
187 members are predominantly benthic inhabitants). Our study showed that at the level of phyla
188 detection was not significantly different between sediment and water for most taxa. Similarly, we
189 showed that for most NIS both water or sediment samples served as an excellent media for
190 detection. Our study suggests that at lower taxonomic level the species-specific ecology of
191 eDNA (sensu[42]) will result in convergent eDNA occupancy in different environmental sample
192 types, as seen in the case of the fish above. However, further work is needed to clarify how
193 eDNA partitions into adjacent environmental samples across the tree of life.
194 Current eDNA metabarcoding research has identified large variation in the detected biodiversity
195 across small spatial scales in both sediment[43] and water[44]. Additionally fractionation of
196 environmental samples (i.e. sorting samples by particle size class) can produce significant
197 differences in the metabarcoding results between fractions[45,46] indicating significant variation can
198 be found within sites. Here we found similar patterns, with site and environmental sample type
199 containing approximately equivalent OTU biodiversity. Future research should explore how
200 different sample types and eDNA extraction methods affect the detection of biodiversity,
201 especially as eDNA metabarcoding moves from an experimental technique to a routine
202 monitoring tool[47,48].
203 We found that eDNA metabarcoding of water samples accurately detects many NIS species, as
204 seen in previous work [13-15]. In comparing our eDNA data to those collected using existing
205 methods we found close congruence in NIS incidence. The false-negative eDNA detection of *B.*
206 *neritina* was found to be a result of bioinformatic parameters, identifying that choices made
207 during sequence processing can have an effect on the detectability of species in eDNA samples.

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

208 Indeed, this has previously been shown in metabarcoding of bulk tissue samples[49] and work is
209 urgently needed to determine the effects of bioinformatic parameters, variable primer binding
210 sites and the role of DNA barcodes in reference databases for the detection of NIS in eDNA
211 samples. The remaining incongruent detections may be as a result of community turnover
212 among the survey dates or seasonal phenology. Indeed marine coastal communities have been
213 shown to shift in community composition across seasons and reproductive cycles[50,51]. It is
214 therefore recommended to combine eDNA metabarcoding with existing survey methods where
215 possible, as both approaches provide reciprocal validation data. For example, important NIS
216 may be missed in surveys based solely on eDNA[e.g. 52], and eDNA metabarcoding may detect
217 rare species that are often missed using other type of surveys [53].
218 In this study we identified several NIS currently unrecorded in the United Kingdom and
219 confirmed the eDNA detection with targeted local surveys. The case of *A. senhousia* is
220 particularly relevant as this species is spreading globally[54] and has the potential to dramatically
221 alter benthic biodiversity when invasive[55,56]. Coincidentally, field surveys along the south coast
222 of the United Kingdom have recently confirmed the presence of *A. senhousia*[57]. The use of
223 routine eDNA surveys has the potential to accurately assess the magnitude of the spread of *A.*
224 *senhousia* along this coast, providing key information for coastal biodiversity managers. Moving
225 forwards, the detection of NIS could be facilitated through autonomous sampling and eDNA
226 surveys[16] to provide live species introduction data in introduction hotspots. Additionally combing
227 these techniques with eDNA biobanking[58] could provide an eDNA reference database for
228 specific geographical regions that have high biosecurity risk, providing an invaluable resource
229 for biodiversity managers and researchers to examine the process of invasion through time.
230 Taken together, our study shows how effective eDNA metabarcoding is as monitoring tool of
231 novel and resident marine species and how it allows for an unprecedented sampling replicability
232 and accuracy of different environmental sample types.
233

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

234 **Methods**

235 *Study sites*

236 Four marinas were selected from around the United Kingdom to represent variation in modelled
237 invasion potential[59], known NIS present[60] and surrounding benthic habitat type[61]. Importantly, all
238 chosen marinas have been surveyed previously for NIS and so there is a good understanding of
239 the expected NIS in these areas[62,63]. Marina access was contingent on anonymity and so marina
240 names and exact locations are not presented, Fig. 1a shows approximate locations. Marina *TQ*
241 is an open marina subject to tides and varying salinity located in Southampton Water on the
242 north coast of the English Channel. Marina *PQ* is a loch marina open during high tide to the
243 Bristol Channel and the Celtic Sea. Marina *TB* is located at the mouth of the River Blackwater
244 open to the North Sea. Marina *HH* is located on the Isle of Anglesey and is open to the Celtic
245 Sea.

246

247 *Environmental DNA sampling*

248 Surveys were conducted during May 2017. A total of 24 sampling points were randomly selected
249 within each site. At each point 50ml of water was collected from 10cm below the surface using a
250 sterile 60ml Luer lock syringe and filtered through a 0.22mm polyethersulfone Sterivex filter
251 (Merck Millipore, Massachusetts USA). After collecting seawater from eight locations (400ml
252 total volume) the filter was changed, resulting in a total of three filters per site. To test the effect
253 of different sample preservation methods sampling was performed in duplicate. One set of three
254 filters had ~1.5ml sterile Longmire's solution (100mM Tris,10mM EDTA, 10mM NaCl, 0.5%
255 SDS) applied in the inlet valve [64]. The second set of three filters were kept on ice for no longer
256 than eight hours before being frozen at -20°C. During the surveys, a sediment sample was
257 collected at the first water sampling site and then after every 3rd water sample, for a total of 9
258 per site. We used a UWITEC Corer (UWITEC, Mondsee, Austria) to collect a sediment core
259 (600mm tall x 60mm diameter). Using a sterile disposable spatula, a subsample of 10-20g of

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

260 sediment was taken from the top 2cm of the core, taking care to avoid sampling the sides of the
261 core. The subsample was stored in a sterile plastic zip container and kept on ice for no longer
262 than eight hours before begin frozen at -80°C. Due to equipment malfunction no sediment
263 sample could be taken for Site HH. Disposable gloves were changed after collection of each
264 sample. All reused equipment was washed thoroughly and soaked in 10% bleach between sites,
265 before rinsing in DNAse-free sterile water.
266
267 *eDNA extraction*
268 DNA extractions were performed in a PCR–free cleanroom, separate from main laboratory
269 facilities. No high copy templates, cultures or amplicons were permitted in this sterile laboratory.
270 DNA extractions from water samples followed Spens *et al.*[20] using the $SX_{CAPSULE}$ method. Briefly,
271 preservative solution was removed from the outlet and filters were dried at room temperature for
272 two hours, 720µl Qiagen buffer ATL (Qiagen, Hilden, Germany) and 80µl Proteinase K was
273 added to the filter and all samples were digested overnight at 56ºC. After digestion, samples
274 were processed using the Qiagen DNeasy Blood and Tissue Kit as per manufacturer
275 instructions. The final elution was 200µl PCR grade water.
276 Sediment extractions were performed using the Qiagen DNeasy Powermax Soil Kit using the
277 manufacturer recommended protocol. For each site the nine samples were mixed to form three
278 pooled samples; for each extraction, 10g of pooled sample was processed. A total of ten
279 samples were processed, three each per site with a single no blank, as per manufacturer's
280 instructions.
281
282 *Inhibition testing*
283 To ensure extracted DNA was free of PCR inhibitors a Primer Design Real-Time PCR Internal
284 Control Kit (PrimerDesign, Southampton, UK) was used. qPCR reactions were performed
285 following the manufacturer's protocol. Inhibition due to co-purified compounds from DNA

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

286 extraction protocols would produce an increase in cycle threshold number in comparison to no
287 template controls. All samples were successfully processed and no samples showed indication
288 of PCR inhibition.
289
290 *Primer selection and library preparation*
291 Two sets of primers were chosen for metabarcoding the environmental samples: a 313bp
292 section of the standard DNA barcoding region of the cytochrome c oxidase subunit I gene (COI)
293 using primers described in Leray *et al.*[65]; and a variable length target of the hypervariable v4
294 region of the nuclear small subunit ribosomal DNA (18S) using primers from Zhan *et al.*[66].
295 Sequencing libraries were prepared using a 2-step PCR approach as detailed in Bista *et al.*[67],
296 this method amplifies the target region in PCR 1 annealing universal adapters onto which
297 sample specific indices and sequencing primers are annealed in PCR 2. In contrast to Bista *et*
298 *al.*[67] we used unique dual-matched indexes for PCR 2 to avoid index crosstalk associated with
299 combinatorial indexing[68]. PCR 1 was prepared in a PCR–free room, separate from main
300 laboratory facilities. Briefly, PCR 1 reactions were conducted in 20ml volumes containing 10µl
301 Amplitaq GOLD 360 2X Mastermix (Applied Biosystems, California, USA), 0.8µl (5 nmol ml^{-1}) of
302 each forward and reverse primer and 2µl of undiluted environmental DNA extract. The reaction
303 conditions for PCR were an initial denaturation step at 95°C for 10 minutes followed by 20
304 cycles of 95°C for 0:30, variable annealing temp (46°C for COI 50°C for 18S) for 0:30, and
305 extension at 72°C for 1:00. A final extension at 72°C was performed for 10:00 minutes. The
306 PCR product was cleaned using AMPure XP beads (Beckman Coulter, California, USA) at a 0.8
307 beads:sample ratio as manufacturer's instructions. PCR 2 reactions were conducted in 20µl
308 volumes containing 10µl Amplitaq GOLD 360 2X Mastermix, 0.5µl (10 nmol ml^{-1}) of both forward
309 and reverse primers and 5µl of undiluted cleaned PCR1 product. PCR conditions were an initial
310 denaturation step at 95°C for 10 minutes followed by 15 cycles of 95°C for 0:30, annealing at
311 55°C for 0:30, and extension at 72°C for 1:00. A final extension at 72°C was performed for 10:00

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

312 minutes. PCR 2 products were cleaned using AMpure XP beads as above and normalised
313 according to their fluorescence using the Quibit HS Assay Kit (Thermofisher Scientific,
314 Massachusetts, USA). These normalised samples were pooled into the two PCR targets and the
315 pools were then quantified as per manufacturer's instructions using the NEBNext Library Quant
316 qPCR kit (New England Biolabs, Massachusetts, USA).
317 Blank filters, DNA extraction kits and positive controls where collected, extracted and sequenced
318 as the experimental treatments (detailed in Supplementary Information 1). Samples were pooled
319 at an equimolar ratio and sequenced using the Illumina MiSeq instrument (Illumina, San Diego,
320 USA) with a V3 2 x 300bp kit.
321
322 *Bioinformatic analyses*
323 Samples were demultiplexed using the Illumina MiSeq control software (v 2.6.2.1). The
324 demultiplexed data was analysed using a custom pipeline written in the R programming
325 language[69] hosted at https://github.com/leholman/metabarTOAD the steps are as follows.
326 Forward and reverse paired end reads were merged using the -fastq_mergepairs option of
327 USEARCH v.10.0.240[70] with maximum difference of 15, percent identity of 80% and quality filter
328 set at maximum expected errors of 1. Both the forward and reverse primer sequences were
329 matched using Cutadapt v.1.16[71] and only sequences containing both primer regions were
330 retained. Sequences were discarded if they were outside of a defined length boundary (303-
331 323bp for COI, 375-450bp for 18S) using Cutadapt. Sequences were then pooled, singletons
332 were discarded and sequences were quality filtered with a maximum expected error of 1 using
333 the -fastq_filter option of vsearch v.2.4.3[72]. Sequences were then denoised and chimeras filtered
334 using the unoise3 algorithm implemented in USEARCH. The resultant operational taxonomic
335 units (OTUs) were curated using the LULU package v.0.1.0 in R[73]. An OTU by sample table was
336 produced by mapping the merged and trimmed reads against the curated OTUs using
337 USEARCH, with the raw query read assigned to the OTU with the best match (highest e value)

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

338 within 97% identity. The OTU x sample table was filtered in R as follows. To minimise the
339 chance of spurious OTUs being included in the final dataset any record with less than 3 raw
340 reads were changed to zero and any OTU that did not appear in more than one sample was
341 removed from the analysis. OTUs found in negative controls were removed from the analysis.
342
343 *Taxonomic assignment*
344 Assigning correct taxonomy to an unknown set of marine sequences can be challenging as
345 large databases require vast computational resources for query matching; many databases
346 contain errors and the taxonomy of some marine groups is uncertain. With such limitations in
347 mind, we assigned taxonomy using a BLAST v.2.6.0+ search[74] returning the single best hit
348 (largest e value) from databases within 97% of the query using a custom R script to parse the
349 raw blast results. The MIDORI database (UNIQUE_20180221)[75] was used for the COI data and
350 the SILVA database (SSU r132)[76] was used for the 18S rRNA data. The match taxa tool from
351 the World Register of Marine Species[77] was used to filter the data for marine species and check
352 the classification. Remaining annotations were checked against the World Register of
353 Introduced Marine Species[78] to determine non-indigenous status.
354
355 *Statistical analyses*
356 All statistical analyses were conducted in R v3.5.0. The Vegan R package v.2.5.2[79] was used to
357 rarefy samples to the minimum sample read depth for each amplicon. The number of OTUs per
358 site/condition was calculated as the number of OTUs with a non-zero number of normalized
359 reads after summing the reads across all three site level replicates. To test if there was a
360 significant difference between the number of OTUs generated by sediment and water eDNA,
361 individual non-summed replicate sample data was used to build a two-way ANOVA model with
362 the formula *number_of_OTUs~sedimentorwater*site* implemented in R using the function aov().
363 Non-metric multidimensional scaling ordination plots were generated from Bray-Curtis

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

364 dissimilarity values derived using Vegan. A Permutation Analysis of Variance (PERMANOVA)[80]
365 was performed using the Bray Curtis dissimilarity following the model
366 *dissimilarity_matrix~sedimentorwater*site* implemented in R using the function *adonis* from the
367 *vegan* package. OTUs with taxonomic assignment were separated into those found in sediment,
368 water or both media and the OTUs were then collapsed at the Phylum level to explore
369 taxonomic patterns of detection in water or sediment. Phyla with less than eight OTUs were
370 combined. To test for non-random counts of species detection between water and sediment
371 within taxa an exact binomial test was performed between counts of species detected in water
372 and sediment. Half the number of counts for species detected in both water and sediment were
373 added to water and sediment, with non-integer values conservatively rounded down to the
374 nearest whole number. A Bonferroni correction for multiple comparisons was applied across the
375 p values from the exact binomial tests. Records from manual surveys previously conducted for
376 non-native invertebrates at the sample sites[60,62,81,82] were compared with the detected species
377 from metabarcoding data.
378
379
380
381
382
383
384
385
386
387
388
389

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

390 **References**

391 1 Sala, E. & Knowlton, N. Global Marine Biodiversity Trends. *Annu. Rev. Environ. Resour.*
392 **31**, 93-122 (2006).
393 2 Butchart, S. H. M. *et al.* Global Biodiversity: Indicators of Recent Declines. *Science* **328**,
394 1164-1168 (2010).
395 3 Worm, B. *et al.* Impacts of biodiversity loss on ocean ecosystem services. *Science* **314**,
396 787-790 (2006).
397 4 Oliver, I. & Beattie, A. J. A Possible Method for the Rapid Assessment of Biodiversity.
398 *Conser. Biol.* **7**, 562-568 (1993).
399 5 Fitzpatrick, M. C., Preisser, E. L., Ellison, A. M. & Elkinton, J. S. Observer bias and the
400 detection of low-density populations. *Ecol. Appl.* **19**, 1673-1679 (2009).
401 6 Thomsen, P. F. & Willerslev, E. Environmental DNA - An emerging tool in conservation
402 for monitoring past and present biodiversity. *Biol. Cons.* **183**, 4-18 (2015).
403 7 Turner, C. R. *et al.* Particle size distribution and optimal capture of aqueous macrobial
404 eDNA. *Methods Ecol. Evol.* **5**, 676-684 (2014).
405 8 Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A. & Boehm, A. B.
406 Quantification of Environmental DNA (eDNA) Shedding and Decay Rates for Three
407 Marine Fish. *Environ. Sci. Technol.* **50**, 10456-10464 (2016).
408 9 Dougherty, M. M. *et al.* Environmental DNA (eDNA) detects the invasive rusty crayfish
409 *Orconectes rusticus* at low abundances. *J. Appl. Ecol.* **53**, 722-732 (2016).
410 10 Simpson, T. J. S., Dias, P. J., Snow, M., Muñoz, J. & Berry, T. Real-time PCR detection
411 of *Didemnum perlucidum* (Monniot, 1983) and *Didemnum vexillum* (Kott, 2002) in an
412 applied routine marine biosecurity context. *Mol. Ecol. Resour.* **17**, 443-453 (2017).
413 11 Wood, S. A., Zaiko, A., Richter, I., Inglis, G. J. & Pochon, X. Development of a real-time
414 polymerase chain reaction assay for the detection of the invasive Mediterranean

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

415 fanworm, *Sabella spallanzanii*, in environmental samples. *Environ. Sci. Pollut. R.* **24**,
416 17373-17382 (2017).
417 12 Kim, P., Kim, D., Yoon, T. J. & Shin, S. Early detection of marine invasive species,
418 *Bugula neritina* (Bryozoa: Cheilostomatida), using species-specific primers and
419 environmental DNA analysis in Korea. *Mar. Environ. Res.* **139**, 1-10 (2018).
420 13 Borrell, Y. J., Miralles, L., Do Huu, H., Mohammed-Geba, K. & Garcia-Vazquez, E. DNA
421 in a bottle-Rapid metabarcoding survey for early alerts of invasive species in ports. *PLoS*
422 *One* **12**, e0183347 (2017).
423 14 Grey, E. K. *et al.* Effects of sampling effort on biodiversity patterns estimated from
424 environmental DNA metabarcoding surveys. *Sci. Rep.* **8**, 8843 (2018).
425 15 Lacoursiere-Roussel, A. *et al.* eDNA metabarcoding as a new surveillance approach for
426 coastal Arctic biodiversity. *Ecol. Evol.* **8**, 7763-7777 (2018).
427 16 McQuillan, J. S. & Robidart, J. C. Molecular-biological sensing in aquatic environments:
428 recent developments and emerging capabilities. *Curr. Opin. Biotechnol.***45**, 43-50 (2017).
429 17 Goldberg, C. S. *et al.* Critical considerations for the application of environmental DNA
430 methods to detect aquatic species. *Methods Ecol. Evol.* **7**, 1299-1307 (2016).
431 18 Deiner, K. *et al.* Environmental DNA metabarcoding: transforming how we survey animal
432 and plant communities. *Mol. Ecol.* **26**, 5872-5895 (2017).
433 19 Deiner, K. *et al.* Optimising the detection of marine taxonomic richness using
434 environmental DNA metabarcoding: the effects of filter material, pore size and extraction
435 method. *Metabarcod. Metagenom.* **2**, e28963 (2018).
436 20 Spens, J. *et al.* Comparison of capture and storage methods for aqueous macrobial
437 eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods*
438 *Ecol. Evol.* **8**, 635-645 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

439 21 Sellers, G. S., Di Muri, C., Gómez, A. & Hänfling, B. Mu-DNA: a modular universal DNA
440 extraction method adaptable for a wide range of sample types. *Metabarcod. Metagenom.*
441 **2**, e24556 (2018).
442 22 Shaw, J. L. A. *et al.* Comparison of environmental DNA metabarcoding and conventional
443 fish survey methods in a river system. *Biol. Cons.* **197**, 131-138 (2016).
444 23 Ashton, G., Books, K., Shucksmith, R. & Cook, E. Rapid assessment of the distribution
445 of marine non-native species in marinas in Scotland. *Aquat. Invasions* **1**, 209-213 (2006).
446 24 Campbell, M. L., Gould, B. & Hewitt, C. L. Survey evaluations to assess marine
447 bioinvasions. *Mar. Pollut. Bull.* **55**, 360-378 (2007).
448 25 Bishop, J. D., Wood, C. A., Leveque, L., Yunnie, A. L. & Viard, F. Repeated rapid
449 assessment surveys reveal contrasting trends in occupancy of marinas by non-
450 indigenous species on opposite sides of the western English Channel. *Mar. Pollut. Bull.*
451 **95**, 699-706 (2015).
452 26 Appeltans, W. *et al.* The magnitude of global marine species diversity. *Curr. Biol.* **22**,
453 2189-2202 (2012).
454 27 Pérez-Portela, R., Arranz, V., Rius, M. & Turon, X. Cryptic speciation or global spread?
455 The case of a cosmopolitan marine invertebrate with limited dispersal capabilities. *Sci.*
456 *Rep.* **3**, 3197 (2013).
457 28 Rius, M. *et al.* Ecological dominance along rocky shores, with a focus on intertidal
458 ascidians. *Oceanogr. Mar. Biol. Ann. Rev.* **55**, 55-86 (2017).
459 29 Bax, N., Williamson, A., Aguero, M., Gonzalez, E. & Geeves, W. Marine invasive alien
460 species: a threat to global biodiversity. *Mar. Policy* **27**, 313-323 (2003).
461 30 Lovell, S., Stone, S. & Fernandez, L. The Economic Impacts of Aquatic Invasive
462 Species: A Review of the Literature. *Agri. Resour. Econ. Rev.* **35**, 195-208 (2006).

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

463 31 Ricciardi, A., Hoopes, M. F., Marchetti, M. P. & Lockwood, J. L. Progress toward
464 understanding the ecological impacts of nonnative species. *Ecol. Monogr.* **83**, 263-282
465 (2013).
466 32 Mazza, G., Tricarico, E., Genovesi, P. & Gherardi, F. Biological invaders are threats to
467 human health: an overview. *Ethol. Ecol. Evol.* **26**, 112-129 (2014).
468 33 Molnar, J. L., Gamboa, R. L., Revenga, C. & Spalding, M. D. Assessing the global threat
469 of invasive species to marine biodiversity. *Front. Ecol. Environ.* **6**, 485-492 (2008).
470 34 Nunes, A. L., Katsanevakis, S., Zenetos, A. & Cardoso, A. C. Gateways to alien
471 invasions in the European seas. *Aquat. Invasions* **9**, 133-144 (2014).
472 35 Murray, C. C., Pakhomov, E. A. & Therriault, T. W. Recreational boating: a large
473 unregulated vector transporting marine invasive species. *Divers. Distrib.* **17**, 1161-1172
474 (2011).
475 36 Airoldi, L., Turon, X., Perkol-Finkel, S. & Rius, M. Corridors for aliens but not for natives:
476 effects of marine urban sprawl at a regional scale. *Divers. Distrib.* **21**, 755-768 (2015).
477 37 Glasby, T. M., Connell, S. D., Holloway, M. G. & Hewitt, C. L. Nonindigenous biota on
478 artificial structures: could habitat creation facilitate biological invasions? *Mar. Biol.* **151**,
479 887-895 (2007).
480 38 Dafforn, K. A., Johnston, E. L. & Glasby, T. M. Shallow moving structures promote
481 marine invader dominance. *Biofouling* **25**, 277-287 (2009).
482 39 Rivero, N. K., Dafforn, K. A., Coleman, M. A. & Johnston, E. L. Environmental and
483 ecological changes associated with a marina. *Biofouling* **29**, 803-815 (2013).
484 40 Foster, V., Giesler, R. J., Wilson, A. M. W., Nall, C. R. & Cook, E. J. Identifying the
485 physical features of marina infrastructure associated with the presence of non-native
486 species in the UK. *Mar. Biol.* **163**, 163-173 (2016).
487 41 Turner, C. R., Uy, K. L. & Everhart, R. C. Fish environmental DNA is more concentrated
488 in aquatic sediments than surface water. *Biol. Cons.* **183**, 93-102 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

489 42 Barnes, M. A. & Turner, C. R. The ecology of environmental DNA and implications for
490 Conserv. Genet.. *Conserv. Genet.* **17**, 1-17 (2016).
491 43 Nascimento, F. J. A., Lallias, D., Bik, H. M. & Creer, S. Sample size effects on the
492 assessment of eukaryotic diversity and community structure in aquatic sediments using
493 high-throughput sequencing. *Sci. Rep.* **8**, 11737 (2018).
494 44 O'Donnell, J. L. *et al.* Spatial distribution of environmental DNA in a nearshore marine
495 habitat. *PeerJ* **5**, e3044 (2017).
496 45 Wangensteen, O. S., Cebrian, E., Palacín, C. & Turon, X. Under the canopy:
497 Community-wide effects of invasive algae in Marine Protected Areas revealed by
498 metabarcoding. *Mar. Pollut. Bull.* **127**, 54-66 (2018).
499 46 Wangensteen, O. S., Palacin, C., Guardiola, M. & Turon, X. DNA metabarcoding of
500 littoral hard-bottom communities: high diversity and database gaps revealed by two
501 molecular markers. *PeerJ* **6**, e4705 (2018).
502 47 Pawlowski, J. *et al.* The future of biotic indices in the ecogenomic era: Integrating
503 (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Sci. Total*
504 *Environ.* **637-638**, 1295-1310 (2018).
505 48 Aylagas, E., Borja, A., Muxika, I. & Rodríguez-Ezpeleta, N. Adapting metabarcoding-
506 based benthic biomonitoring into routine marine ecological status assessment networks.
507 *Ecol. Indic.* **95**, 194-202 (2018).
508 49 Scott, R. *et al.* Optimization and performance testing of a sequence processing pipeline
509 applied to detection of nonindigenous species. *Evol. Appl.* **11**, 891-905 (2018).
510 50 Stachowicz, J. J. & Byrnes, J. E. Species diversity, invasion success, and ecosystem
511 functioning: disentangling the influence of resource competition, facilitation, and extrinsic
512 factors. *Mar. Ecol. Prog. Ser.* **311**, 251-262 (2006).
513 51 Sutherland, J. P. & Karlson, R. H. Development and Stability of the Fouling Community
514 at Beaufort, North Carolina. *Ecol. Monogr.* **47**, 425-446 (1977).

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

515 52 Wood, S. A. *et al.* Considerations for incorporating real-time PCR assays into routine
516 marine biosecurity surveillance programmes: a case study targeting the Mediterranean
517 fanworm (*Sabella spallanzanii*) and club tunicate (*Styela clava*). *Genome* **0**, 1-10 (2018).
518 53 Blackman, R. C. *et al.* Detection of a new non-native freshwater species by DNA
519 metabarcoding of environmental samples – first record of *Gammarus fossarum* in the
520 UK. *Aquat. Invasions* **12**, 177-189 (2017).
521 54 Bachelet, G. *et al.* A round-the-world tour almost completed: first records of the invasive
522 mussel *Musculista senhousia* in the north-east Atlantic (southern Bay of Biscay). *Mar.*
523 *Biodivers. Rec.* **2**, 2002–2005 (2009).
524 55 Crooks, J. A. Assessing invader roles within changing ecosystems: historical and
525 experimental perspectives on an exotic mussel in an urbanized lagoon. *Biol. Invasions* **3**,
526 23-36 (2001).
527 56 Mistri, M. The non-indigenous mussel *Musculista senhousia* in an Adriatic lagoon:
528 Effects on benthic community over a ten year period. *J. Mar. Biol. Assoc. U. K.* **83**, 1277-
529 1278 (2003).
530 57 Barfield, P., Holmes, A., Watson, G. & Rowe, G. First Evidence of *Arcuatula senhousia*
531 (Benson, 1842), The Asian Date Mussel in UK waters. *J. Conchol.* **43**, 217-222 (2018).
532 58 Jarman, S. N., Berry, O. & Bunce, M. The value of environmental DNA biobanking for
533 long-term biomonitoring. *Nat. Ecol. Evol.* **2**, 1192-1193 (2018).
534 59 Pearce, F., Peeler, E. & Stebbing, P. Modelling the Risk of the Introduction and Spread
535 of Non-Indigenous Species in the UK and Ireland. *Project Report for E5405W. CEFAS*
536 (2012).
537 60 Bishop, J., Wood, C., Yunnie, A. & Griffiths, C. Unheralded arrivals: non-native sessile
538 invertebrates in marinas on the English coast. *Aquat. Invasions* **10**, 249-264 (2015).
539 61 Calewaert, J.-B., Weaver, P., Gunn, V., Gorringe, P. & Novellino, A. in *Quantitative*
540 *Monitoring of the Underwater Environment: Results of the International Marine Science*

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

541 *and Technology Event MOQESM´14 in Brest, France* (eds B. Zerr *et al.*) 31-46
542 (Springer International Publishing, Cham, 2016).
543 62 Wood, C. A., Bishop, J. D. D., Rennocks, L. & Crundwell, R. Non-Native Species Rapid
544 Assessment Surveys in English Marinas (E Anglia & W coast). (2016).
545 63 Wood. RAS 2015: Non-Native Species Rapid Assessment Surveys in English Marinas (E
546 Anglia & W coast). (2016).
547 64 Renshaw, M. A., Olds, B. P., Jerde, C. L., McVeigh, M. M. & Lodge, D. M. The room
548 temperature preservation of filtered environmental DNA samples and assimilation into a
549 phenol-chloroform-isoamyl alcohol DNA extraction. *Mol. Ecol. Resour.* **15**, 168-176
550 (2015).
551 65 Leray, M. *et al.* A new versatile primer set targeting a short fragment of the mitochondrial
552 COI region for metabarcoding metazoan diversity: application for characterizing coral
553 reef fish gut contents. *Front. Zool.* **10**, 34 (2013).
554 66 Zhan, A. *et al.* High sensitivity of 454 pyrosequencing for detection of rare species in
555 aquatic communities. *Methods Ecol. Evol.* **4**, 558-565 (2013).
556 67 Bista, I. *et al.* Annual time-series analysis of aqueous eDNA reveals ecologically relevant
557 dynamics of lake ecosystem biodiversity. *Nat. Commun.* **8**, 14087 (2017).
558 68 MacConaill, L. E. *et al.* Unique, dual-indexed sequencing adapters with UMIs effectively
559 eliminate index cross-talk and significantly improve sensitivity of massively parallel
560 sequencing. *BMC Genomics* **19**, 30 (2018).
561 69 R_Core_Team. R: A Language and Environment for Statistical Computing. *ISBN 3-*
562 *900051-07-0* (2018).
563 70 Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads.
564 *Nat. Methods* **10**, 996-998 (2013).
565 71 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing
566 reads. *EMBnet J.* **17**, 10-12 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

567 72 Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahe, F. VSEARCH: a versatile open
568 source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
569 73 Frøslev, T. G. *et al.* Algorithm for post-clustering curation of DNA amplicon data yields
570 reliable biodiversity estimates. *Nat. Commun.* **8**, 1188 (2017).
571 74 Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinform.* **10**, 421
572 (2009).
573 75 Machida, R. J., Leray, M., Ho, S. L. & Knowlton, N. Metazoan mitochondrial gene
574 sequence reference datasets for taxonomic assignment of environmental samples. *Sci.*
575 *Data* **4**, 170027 (2017).
576 76 Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data
577 processing and web-based tools. *Nucleic Acids Res.* **41**, D590-596 (2013).
578 77 WoRMS_Editorial_Board. *World Register of Marine Species*,
579 <http://www.marinespecies.org> (2018).
580 78 Ahyong, S. *et al. World Register of Introduced Marine Species (WRiMS)*,
581 <http://www.marinespecies.org/introduced> (2018).
582 79 Oksanen, J. *et al.* Vegan: community ecology package. *R package* **1**, 17 (2011).
583 80 Anderson, M. J. in *Wiley Stats Ref: Statistics Reference Online* (eds N. Balakrishnan *et*
584 *al.*) 1-15 (John Wiley & Sons, Ltd, 2014).
585 81 Wood, C. A., Bishop, J. D. D., Nall, C. R. & Rennocks, L. RAS 2016. Non-Native Species
586 Rapid Assessment Surveys in English Marinas (NE & SW coasts). 42 (2017).
587 82 Wood, C. A., Bishop, J. D. D. & Yunnie, A. L. E. Comprehensive Reassessment of NNS
588 in Welsh Marinas. (2015).
589
590
591
592

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

593 **Acknowledgements**

594 We are grateful to John Bishop and Chris Wood from the Marine Biological Association of the
595 United Kingdom for sharing information on marinas and their excellent NIS survey data. We
596 thank the staff of the Environmental Sequencing Facility from the National Oceanography
597 Centre Southampton for advice and assistance during library preparation. We acknowledge the
598 Department of Geography and Environment from the University of Southampton for access to
599 coring equipment and lab space. LH was supported by the Natural Environmental Research
600 Council (grant number NE/L002531/1).

601

602 **Competing Interests**

603 The authors declare no competing interests.

604

605 **Author Contributions**

606 L.E.H. and M.R. designed the experiment, L.E.H. collected samples, generated and analysed
607 the data, designed all figures and wrote the first draft of the paper. All authors contributed
608 critically to the drafts and gave final approval for publication.

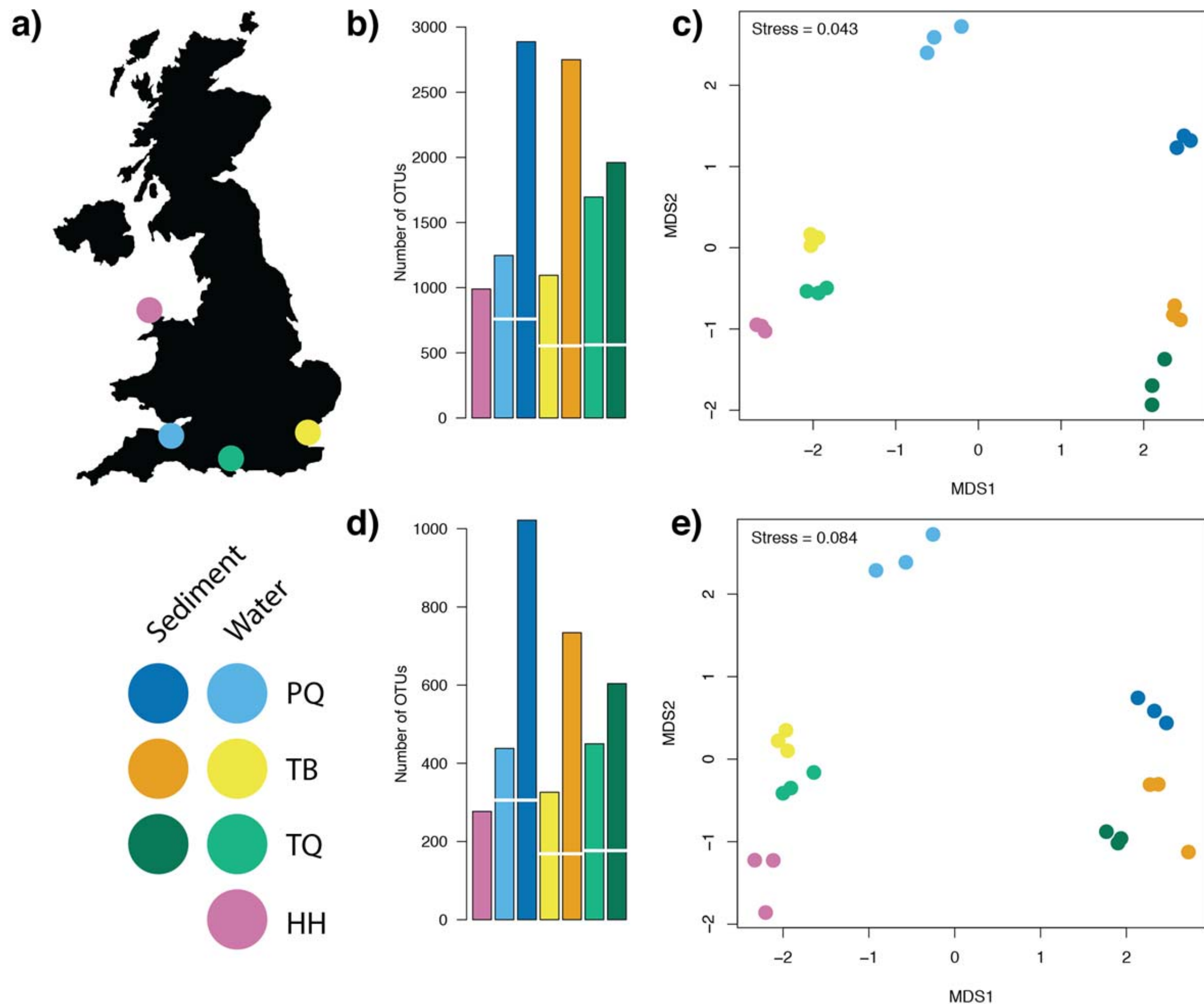
609

610 **Data accessibility**

611 Raw Illumina sequencing data is available from the European Nucleotide Archive under
612 accession number (data will be uploaded upon acceptance, available to reviewers upon
613 request).
614 Associated metadata, script and intermediate files can be found on GitHub with the following
615 DOI: 10.5281/zenodo.1453959.

616

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



617

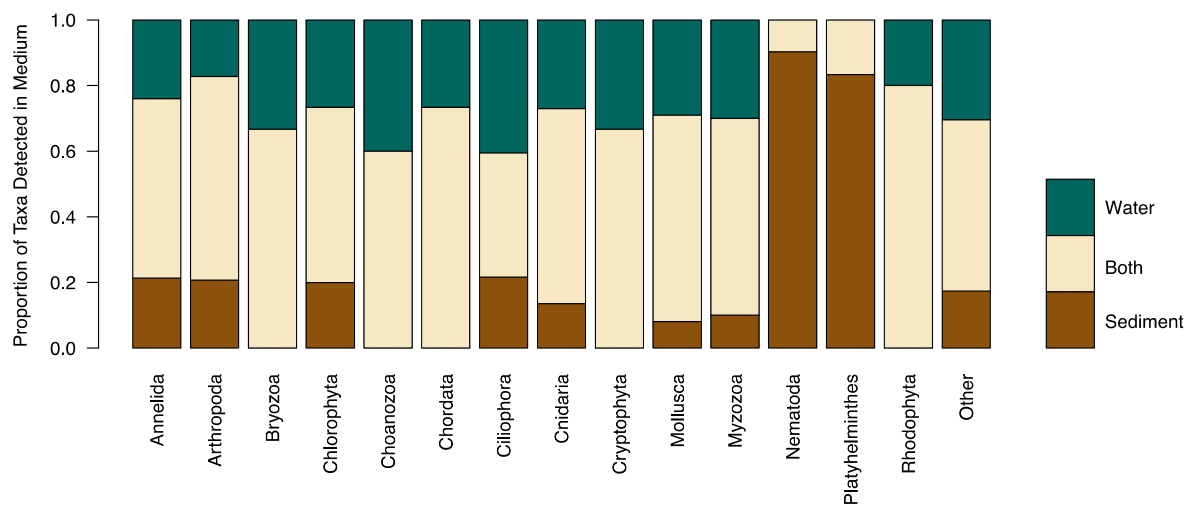
618

619 ***Fig. 1.*** *a) Map of United Kingdom indicating the geographic position of the sampled sites, a legend is*
620 *provided below indicating the four sites (PQ, TB, TQ and HH) and colours for water and sediment eDNA*
621 *samples for each site. Barplots detailing number of OTUs detected across sampling sites and eDNA*
622 *sample type for COI b) and 18S rRNA d) metabarcoding of UK marinas, the break in bars indicates the*
623 *number of shared OTUs between sediment and water eDNA samples. Non-metric multidimensional*
624 *scaling ordination plots based on Bray-Curtis dissimilarities between: c) COI and e) 18S rRNA*
625 *metabarcoding of marina sediment and water eDNA samples.*

626

627

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



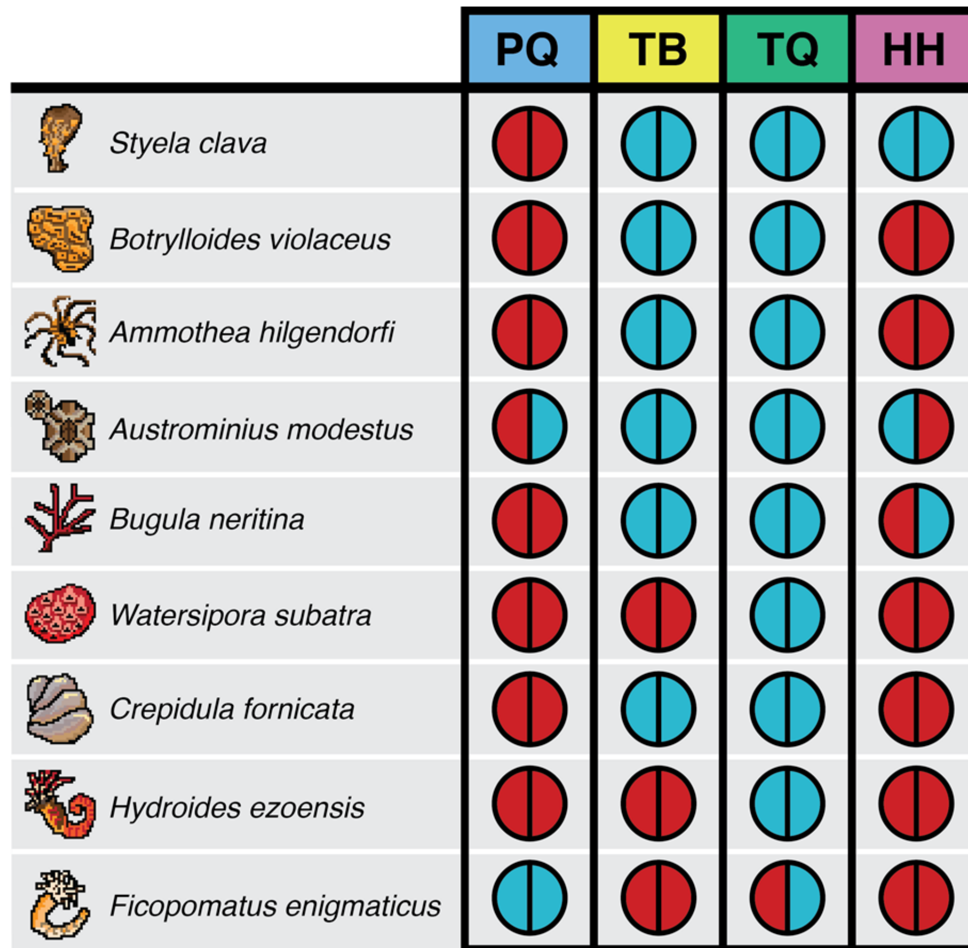
628

629 ***Fig. 2.*** *Horizontal stacked barchart detailing proportion of OTUs detected in eDNA from sediment, water*
630 *or both sediment and water across the top 14 phyla for pooled 18S rRNA and COI metabarcoding of UK*
631 *marinas.*

632

633

bioRxiv preprint doi: https://doi.org/10.1101/440768; this version posted December 9, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



634

635

636 ***Fig. 3.*** *Presence absence diagram for seven non-indigenous species across four sampling sites. For each*

637 *species-location the left semi-circle indicates the detection using eDNA metabarcoding surveys of 18S*

638 *rRNA and COI fragments, and the right semi-circle indicates the detection from Rapid Assessment*

639 *surveys. Blue indicates a positive detection for that species-location and red indicates no detection.*

640

641

642

643

644