Optimised DNA isolation from marine sponges for natural sampler DNA (nsDNA) metabarcoding

Lynsey R. Harper[1,2*], Erika F. Neave[1,4], Graham S. Sellers[3], Alice V. Cunnington[1], María Belén Arias[4], Jamie Craggs[4], Barry MacDonald[5], Ana Riesgo[4,6], Stefano Mariani[1]

[1] School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, L3 3AF, UK
[2] NatureMetrics Ltd, 1 Occam Court, Surrey Research Park, Guildford, GU2 7HJ, UK
[3] Department of Biological and Marine Sciences, University of Hull, Kingston upon Hull, HU6 7RX, UK
[4] Department of Life Sciences, The Natural History Museum, London, SW7 5BD, UK
[5] Fisheries and Oceans Canada, 1 Challenger Drive, Dartmouth, Nova Scotia, B2Y 4A2, Canada
[6] Department of Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales (CSIC), C/José Gutiérrez Abascal 2, 28006, Madrid, Spain

*Corresponding author:
Lynsey R. Harper
lynsey.harper2@gmail.com

Running title: DNA extraction for sponge nsDNA metabarcoding

Abstract

Marine sponges have recently been recognised as natural samplers of environmental DNA (eDNA) due to their effective water filtration and their ubiquitous, sessile and regenerative nature. However, laboratory workflows for metabarcoding of sponge tissue have not been optimised to ensure that these natural samplers achieve their full potential for community survey. We used a phased approach to investigate the influence of DNA extraction procedures on the biodiversity information recovered from sponges. In Phase 1, we compared three treatments of residual ethanol preservative in sponge tissue alongside five DNA extraction protocols. The results of Phase 1 informed which ethanol treatment and DNA extraction protocol should be used in Phase 2, where we assessed the effect of starting tissue mass on extraction success and whether homogenisation of sponge tissue is required. Phase 1 results indicated that ethanol preservative may contain unique and/or additional biodiversity information to that present in sponge tissue, but blotting tissue dry generally recovered more taxa and generated more sequence reads from the wild sponge species. Tissue extraction protocols performed best in terms of DNA concentration, taxon richness and proportional read counts, but the non-commercial tissue protocol was selected for Phase 2 due to cost-efficiency and greater recovery of target taxa. In Phase 2 overall, we found that homogenisation may not be required for sponge tissue and more starting material does not necessarily improve taxon detection. These results combined provide an optimised DNA isolation procedure for sponges to enhance marine biodiversity assessment using natural sampler DNA metabarcoding.

Keywords: biodiversity, DNA extraction, environmental DNA, fish, marine, natural samplers, Porifera, vertebrates

1. Introduction

The emergence of environmental DNA (eDNA) analysis signalled a new age of biodiversity monitoring, facilitating faster, non-invasive, cost-efficient and comprehensive surveys for single species and whole communities than was previously thought possible (Lawson Handley 2015; Deiner et al., 2017; Miya 2021). Such has been the rapid, exponential growth of this field in the last decade (Tsuji et al., 2019; Thalinger et al., 2021) that it is now beginning to branch. In addition to the now relatively standard analyses of aquatic eDNA and soil/sediment eDNA, studies have been published on species detection using eDNA extracted from snow (Franklin et al., 2019; Kinoshita et al., 2019) and air (Clare et al., 2021; Lynggaard et al., 2022; Roger et al., 2022). Innovative research has also demonstrated that water can be used to aggregate eDNA deposited on plants (Valentin et al., 2020) or fishing nets (Russo et al., 2020) for species detection. Less closely intertwined but stemming from the same roots is natural sampler DNA (nsDNA) (Siegenthaler et al., 2019), which refers to the analysis of DNA collected by live organisms. This is not a novel concept as researchers have been analysing the DNA present in faeces and gut contents of consumers (Pompanon et al., 2012; Nørgaard et al., 2021) and hematophagous insects (Schnell et al., 2012; Massey et al., 2021) for years, but more creative applications are now being explored, including flowers and cow dung as samplers of insect assemblages (Thomsen & Sigsgaard, 2019; Sigsgaard et al., 2020; Harper et al., 2021), dung beetles as samplers of mammal communities (Drinkwater et al., 2021), and sponges as samplers of marine biodiversity (Mariani et al., 2019; Turon et al., 2020; Jeunen et al., 2021).

Marine sponges can filter thousands of litres of water in a single day (Gökalp et al., 2020), vastly exceeding the volume that could be collected with artificial devices for aquatic eDNA metabarcoding. As water is filtered, particles are trapped and concentrated in sponge tissue until digestion or excretion, meaning that a biopsy can be taken to provide DNA for metabarcoding analysis (Mariani et al., 2019; Turon et al., 2020; Jeunen et al., 2021). Sponge nsDNA metabarcoding offers several advantages over aquatic eDNA metabarcoding. The high-tech robotics being used to sample large volumes of water for aquatic eDNA (e.g. automated underwater vehicles) are expensive

87 to build and operate, are ineffective for inaccessible and/or complex habitats, and often
88 target a limited set of taxa (Mariani et al., 2019). Although passive filtration has been
89 explored as an alternative to active filtration using peristaltic or vacuum pumps, they
90 require long submergence times and have saturation limits (Kirtane et al., 2020; Bessey
91 et al., 2021b; a). Sponges are present in nearly every aquatic habitat, and sample
92 collection is inexpensive and non-destructive with minimal environmental impact.
93 Provided that biopsies are not destructive and conducted carefully, sponges will
94 regenerate quickly. Therefore, sponges offer a low-tech, affordable and standardised
95 method of surveying marine biodiversity, and could serve as long-term monitoring
96 stations (Mariani et al., 2019). However, there is much work to be done before sponges
97 can be reliably and routinely used in community surveys, specifically sampling and
98 laboratory protocols (Jeunen et al., 2021).

99 Two sponge nsDNA studies to date (Mariani et al., 2019; Turon et al., 2020)
100 utilised samples from existing collections and extracted DNA from tissue with
101 commercial kits. The other study (Jeunen et al., 2021) compared dive surveys and
102 aquatic eDNA to sponge tissue and the ethanol used to preserve sponge tissue, and
103 also extracted DNA using a commercial kit. It remains undetermined whether certain
104 sponge species are better samplers than others and if this is related to phenotype. The
105 spatial extent of biodiversity information recovered by a single sponge is yet to be
106 ascertained along with the optimal mass/volume for biopsies to maximise data
107 generation without being lethal to the sponge sampled (Mariani et al., 2019). The
108 influence of DNA extraction method on biodiversity detection from sponge tissues is
109 unknown (Jeunen et al., 2021), but this has been found to affect detection success with
110 aquatic eDNA (Deiner et al., 2018; Jeunen et al., 2019), bulk invertebrates (Hermans et
111 al., 2018; Majaneva et al., 2018) and faecal samples (Kaunisto et al., 2017).
112 Furthermore, it has not been investigated if samples should be homogenised to ensure
113 DNA is evenly distributed throughout the sample prior to extraction, similar to bulk
114 invertebrate and faecal samples (Gosselin et al., 2017; Pereira-da-Conceicoa et al.,
115 2020). The optimal amount of starting material to maximise species recovery without
116 introducing PCR inhibitors is also unknown, but using larger amounts of soil was found
117 to improve invertebrate detection (Kirse et al., 2021).

118 Here, we focus on the influence of laboratory workflows on the performance of
119 sponge nsDNA metabarcoding, specifically DNA isolation procedures. We used a
120 phased approach with three sponge species to investigate whether residual ethanol
121 preservative, DNA extraction protocol, type of starting material, and amount of starting
122 material influence sponge nsDNA metabarcoding performance using the following
123 extraction success criteria: 1) total DNA yield, 2) target DNA concentration, 3) taxon
124 richness, 4) proportional read counts, and 5) community composition. In Phase 1, we
125 compared three treatments of residual ethanol preservative in sponge tissue (no
126 removal, removal via blotting, removal via centrifugation) alongside five DNA extraction
127 protocols, including commercial kits and non-commercial protocols. Using the optimal
128 ethanol treatment and DNA extraction protocol from Phase 1, we compared different
129 volumes of homogenised sponge tissue to different weights of dried sponge tissue. We
130 synthesise the results from each experimental phase to provide an optimised DNA
131 isolation procedure for sponge tissue to enhance marine biodiversity assessment using
132 sponge nsDNA metabarcoding.
133
134

135 ## 2. Methods

136

137 ### 2.1 Workspace and decontamination procedures

138 A unidirectional workflow was used for sample processing. Sponge tissue was stored
139 and handled in a laboratory dedicated to the processing of environmental samples with
140 low DNA concentrations. Prepared PCR reactions were transported to separate
141 laboratories for: addition of PCR positive controls and PCR amplification; gel
142 electrophoresis, library preparation, and storage; and library quality checks,
143 quantification, and sequencing. Bench space and non-immersible equipment in all
144 laboratories were sterilised before and after use by wiping with fresh paper towel and
145 10% v/v bleach solution (made using Cleanline thin bleach containing 4.53% sodium
146 hypochlorite), followed by 70% v/v ethanol solution. Immersible equipment was
147 sterilised in a 10% v/v bleach bath for at least 10 minutes, then immersed in a 5% v/v
148 Lipsol detergent bath and rinsed with deionised water. Bleach and Lipsol detergent

149 baths were changed weekly or daily during periods of heavy use. Stainless steel 5 mm
150 beads (Qiagen, Germany) were additionally sterilised by heating at 220°C for 3 hours.
151 Plastics and hoods were sterilised with UV light for at least 30 minutes. Full details of
152 decontamination procedures are provided in Supporting Information.
153
154 **2.2 Sponge selection**
155 For this experiment, we used samples from three marine sponges which are the focus
156 of ongoing research employing nsDNA metabarcoding: 1) an invasive sponge present in
157 the aquarium facility of the Horniman Museum and Gardens, London, 2) *Vazella*
158 *pourtalesii*, which is a siliceous sponge that colonises hydroacoustic moorings in Nova
159 Scotia, Canada, and 3) *Phakellia ventilabrum*, which is an axinnellid that forms relatively
160 dense aggregations on rock-sand habitats and is broadly distributed across the North
161 Atlantic. The Horniman invasive sponge provided an opportunity to groundtruth nsDNA
162 metabarcoding in a controlled tank environment with 25 species of known abundance,
163 whereas *V. pourtalesii* and *P. ventilabrum* were collected from natural marine
164 environments with more diverse species assemblages and differed in phenotype as well
165 as filtration efficiency. The Horniman invasive sponge and *V. pourtalesii* were used in
166 Phase 1, whereas *V. pourtalesii* and *P. ventilabrum* were used in Phase 2 due to the
167 greater amounts of starting material required for DNA extraction.
168
169 **2.3 Sponge collection**
170 All sponges were detached from their substrate using a sterile diving knife or sterile
171 disposable surgical scalpel (Swann-Morton No. 21, Fisher Scientific, UK) whilst wearing
172 disposable gloves. Individuals of the Horniman invasive sponge were small thus an ~1
173 cm^3 biopsy was taken. Each biopsy was placed in a 2 ml DNA LoBind tube (Eppendorf,
174 Fisher Scientific, UK) containing 1 ml of molecular grade 100% ethanol (Fisher
175 Scientific, UK) and stored at -20°C for transport to Liverpool John Moores University
176 (LJMU). At LJMU, forceps were used to transfer each biopsy to a new 2 ml DNA LoBind
177 tube containing 1 ml of molecular grade 100% ethanol as water present in tissue may
178 have diluted the original ethanol. Samples were stored at -20°C for three months until
179 DNA extraction. As *V. pourtalesii* colonises hydroacoustic moorings, samples were

180 collected during mooring retrieval undertaken by Ocean Tracking Network and Fisheries
181 and Oceans Canada in July 2021. Whole individuals were removed from the moorings,
182 placed in sterile pots containing 500 ml of molecular grade 100% ethanol, and stored at
183 -20°C once research vessels returned to shore. The ethanol in each pot was changed
184 twice before samples were shipped to LJMU and stored at -20°C for three months until
185 DNA extraction. For *P. ventilabrum,* a ~5 cm^3 biopsy was taken from each individual
186 sampled from Rockall Bank (coordinates: 57.1705, -13.638) at 185-189 m depth in May
187 2018, then placed in a 50 ml falcon tube containing 25 ml of 100% ethanol and stored at
188 4°C. Again, the ethanol was changed twice before samples were shipped to LJMU in
189 January 2021 where they were stored at -20°C for three months until DNA extraction.
190

191 **2.4 DNA extraction**

192 *2.4.1 Phase 1*

193 Three Horniman invasive sponge samples and three *V. pourtalesii* samples were used
194 in Phase 1 (Fig. 1). Each sample from each sponge was subjected to a different ethanol
195 treatment as ethanol is known to interact with chemicals used during DNA extraction
196 and inhibit PCR amplification, but residual ethanol preservative may be present in
197 porous sponge tissue. The ethanol treatments were: 1) no removal of residual ethanol
198 (wet treatment), 2) removal of residual ethanol by blotting sponge tissue against filter
199 paper inside a petri dish (blot treatment), and 3) biopsies were transferred to a new 2 ml
200 DNA LoBind tube and centrifuged for 1 minute at maximum speed, following which the
201 ethanol was removed with a pipette and centrifugation repeated until no more ethanol
202 was observed (spin treatment). After each ethanol treatment, samples were cut into
203 small pieces within a petri dish using dissection scissors. Pieces were added to a weigh
204 boat until a maximum weight of 500 mg was observed. Each subsample was transferred
205 to a high-impact 2 ml screw cap microtube (Starlab, UK) containing a 5 mm stainless
206 steel bead and twice the amount of 1x TE buffer to tissue weight was added to each
207 microtube (e.g. 500 µl to 250 mg). This resulted in 20-30 subsamples per *V. pourtalesii*
208 sample. Each Horniman invasive sponge sample weighed ~500 mg individually thus no
209 subsamples were required. The (sub)samples were homogenised on a Qiagen
210 TissueLyser II for 2 minutes at maximum speed. Using a 1000 µl pipette tip, the

211 homogenate from each subsample was pooled in a 15 ml falcon tube corresponding to
212 each original sample. Twelve 200 µl aliquots of each *V. pourtalesii* sample and twelve
213 50 µl aliquots of the Horniman invasive sponge (Fig. 1) were transferred to individual 2
214 ml DNA LoBind tubes and frozen at -20°C alongside the homogenised samples.

215 The aliquots were used as the starting material for extraction with three
216 commercial and two non-commercial DNA extraction protocols: 1) Qiagen DNeasy
217 PowerSoil Kit (QPS), 2) Qiagen DNeasy Blood and Tissue Kit (QBT), 3) Qiagen
218 DNeasy Blood and Tissue Kit and post-extraction inhibitor removal with the Zymo
219 Research OneStep™ PCR Inhibitor Removal Kit (QBT-ZYMO), 4) Mu-DNA soil protocol
220 (MUS), and 5) Mu-DNA tissue protocol (MUT) with an added inhibitor removal step
221 (Sellers et al*.,* 2018) (Fig. 1). The QPS, QBT, MUS and MUT protocols were used on
222 three aliquots of each sample (*n* = 72) following the manufacturer's instructions or steps
223 outlined in Sellers et al. (2018) with one modification. Elution buffer (100 µl) was added
224 to each spin column, incubated for 5 minutes at room temperature before centrifugation
225 for 1 minute at 10,000 x g, then the eluate was collected and added to the same spin
226 column again for incubation for 5 minutes at room temperature before centrifugation for
227 1 minute at 10,000 x g. An extraction blank (extraction buffers only) was processed
228 alongside samples for each extraction protocol (*n* = 4).

229 The extracts resulting from QBT extraction were halved (Fig. 1), and one half
230 was processed with the Zymo Research OneStep™ PCR Inhibitor Removal Kit
231 (Cambridge Bioscience, UK) for the QBT-ZYMO protocol (*n* = 19). Notably, three preps
232 from the OneStep™ PCR Inhibitor Removal Kit failed, resulting in only 16 extracts for
233 the QBT-ZYMO protocol (Fig. 1). A 25 µl and 5 µl aliquot of each extract (*n* = 92) were
234 transferred to PCR strip tubes (Starlab, UK). All original DNA extracts and 25 µl aliquots
235 were frozen at -20°C until PCR preparation. A Nanodrop was used to measure DNA
236 purity and a Qubit™ 4 Fluorometer (Thermo Scientific, UK) with Qubit™ dsDNA HS
237 Assay Kit (Fisher Scientific, UK) was used to measure total DNA yield from the 5 µl
238 aliquots, which were subsequently discarded.
239
240

241 *2.4.2 Phase 2*

242 Three *V. pourtalesii* samples and three *P. ventilabrum* samples were used in Phase 2
243 (Fig. 1). Based on the findings of Phase 1 (see Results), all samples were blotted
244 against filter paper in a petri dish to remove residual ethanol preservative and extracted
245 with the MUT protocol. Samples from each sponge species were halved (*n* = 12), with
246 each half being used to provide either dry tissue or homogenate as starting material
247 (Fig. 1). Homogenisation of each sample was performed as described in Phase 1.
248 Three dry tissue weights and three homogenate volumes were compared in triplicate for
249 each sample (*n* = 108), and the volume of Lysis Master Mix (1 ml contains 730 µl of
250 Lysis Solution, 250 µl of Tissue Lysis Additive and 20 µl of Proteinase K) for the MUT
251 protocol was adjusted proportionately: 1) 50 mg and 100 µl (≈50 mg) to 300 µl of Lysis
252 Master Mix, 2) 100 mg and 200 µl (≈100 mg) to 600 µl of Lysis Master Mix, and 3) 500
253 mg and 1000 µl (≈500 mg) to 1 ml of Lysis Master Mix. Extraction followed the steps
254 outlined in Sellers et al. (2018) with the same modification for elution as Phase 1.
255 Notably, there was not enough tissue for some samples to obtain triplicate 500 mg
256 weights or 1000 µl volumes, even when additional samples collected from the same
257 location were used as replicates (PHAK4, PHAK5), resulting in 105 DNA extracts. An
258 extraction blank (extraction buffers only) was included for each round of extractions (*n* =
259 4) (Fig. 1). A 25 µl and 5 µl aliquot of each extract (*n* = 109) were transferred to PCR
260 strip tubes (Starlab, UK). All original DNA extracts and 25 µl aliquots were frozen at -
261 20°C until PCR preparation. A Nanodrop was used to measure DNA purity and a
262 Qubit™ 4 Fluorometer was used to measure total DNA yield.

263

264 **2.5 PCR amplification and library preparation**

265 DNA extracts were PCR amplified using the Tele02 primers (F: 5'-
266 AAACTCGTGCCAGCCACC-3', R: 5'-GGGTATCTAATCCCAGTTTG-3') which target a
267 ~167 bp fragment of the mitochondrial 12S ribosomal RNA (rRNA) gene and achieved
268 >98% teleost species detection when tested *in silico* (Taberlet *et al.,* 2018). Although
269 the primers were designed to target teleost fish, mammal and bird DNA may also
270 amplify (Mariani *et al.,* 2021). To ease sample demultiplexing and mitigate cross-

271 contamination and/or tag switching during sequencing, each sample was PCR amplified
272 using the locus primers attached to a unique 8 bp tag shared by the forward and
273 reverse primer. Each tag (*n* = 96) differed by at least three base pairs from other tags
274 and included 2-4 degenerate bases (Ns) at the beginning of the tag sequence to
275 improve clustering during initial sequencing. For Phase 2, samples were processed in
276 two batches to facilitate use of the same tags in different pools for adapter ligation.

277 Each sample was PCR amplified in triplicate using 20 µl reactions consisting of
278 10 µl of MyFi™ Mix (Meridian Bioscience, UK), 1 µl of each forward and reverse primer
279 (10 µM, Eurofins), 0.16 µl of Bovine Serum Albumin (20 mg/ml, Thermo Scientific, UK),
280 5.84 µl of UltraPure™ Distilled Water (Invitrogen, UK), and 2 µl of DNA template. Two
281 PCR positive controls (iridescent catfish, *Pangasianodon hypophthalmus*, tissue DNA at
282 0.05 ng/µl) and two PCR negative controls (molecular grade water) were included in
283 each PCR run (up to 96 reactions). PCR was performed on a T100 Thermal Cycler (Bio-
284 Rad Laboratories Ltd, UK) with the following profile: 95°C for 10 minutes, followed by 40
285 cycles of 95°C for 30 s, 60°C for 45 s, and 72°C for 30 s, and final elongation at 72°C
286 for 5 minutes. PCR products were run on a 2% agarose gel, made with 150 ml of 1x
287 TBE buffer (Invitrogen, UK), 3 g of agarose powder (Fisher Scientific, UK), and 1.5 µl of
288 SYBR™ Safe DNA Gel Stain (Invitrogen, UK), at 100 V for 40 minutes using 1 µl of
289 PCR product to 1 µl of 10x BlueJuiceTM Gel Loading Buffer (Invitrogen, UK). Gels were
290 visualised with a Bio-Rad Gel Doc™ EZ Imaging System.

291 PCR replicates for each sample were pooled prior to purification with Mag-Bind®
292 Total Pure NGS Beads (Omega Bio-tek Inc., GA, USA), following the double size
293 selection protocol established by Bronner et al. (2009). Ratios of 1x and 0.6x magnetic
294 beads to 30 µl of PCR product were used. Eluted DNA (20 µl) was stored at -20°C until
295 quantification using a Qubit™ 4 Fluorometer. Purified samples were normalised and
296 pooled in equimolar concentration. Phase 1 consisted of one pool of 96
297 samples/controls. Phase 2 required two pools - the first consisted of 65
298 samples/controls from Phase 2 and the second consisted of 40 samples/controls from
299 Phase 2 and 25 samples/controls from other projects. Sample pools were purified using
300 the aforementioned bead ratios and elution in 25 µl, then concentrated using a 1x bead
301 ratio and elution in 45 µl.

302 End repair, adapter ligation and PCR were performed using the NEXTFLEX®
303 Rapid DNA-Seq Kit 2.0 for Illumina® Platforms (PerkinElmer, UK) according to the
304 manufacturer's protocol. An Agilent 2200 TapeStation and High Sensitivity D1000
305 ScreenTape (Agilent Technologies, UK) indicated secondary product remained, thus gel
306 extraction was performed on each pool using the GeneJET Gel Extraction Kit (Thermo
307 Scientific, UK) with elution in 20 µl. Each pool was quantified using quantitative PCR
308 (qPCR) on a Rotor-Gene Q (Qiagen, Germany) with the NEBNext® Library Quant Kit for
309 Illumina® (New England Biolabs, UK) and diluted to 1 nM. For Phase 2, 6 µl of each
310 pool were combined into one library. The final libraries and PhiX Control were quantified
311 using qPCR before the Phase 1 library was sequenced at 50 pM with 10% PhiX Control
312 and the Phase 2 library was sequenced at 60 pM with 10% PhiX Control on an Illumina®
313 iSeq™ 100 using iSeq™ 100 i1 Reagent v2 (300-cycle) (Illumina Inc., CA, USA).
314
315 **2.6 Bioinformatics**
316 *2.6.1 Reference databases*
317 Sequence data were processed in two batches to facilitate taxonomic assignment
318 against different reference databases. Sequences for Horniman invasive sponge
319 samples and associated controls were compared against a reference database for the
320 aquarium fish species, the PCR positive control, and common domestic species (human
321 *Homo sapiens*, cow *Bos taurus*, pig *Sus scrofa*, chicken *Gallus gallus*, sheep *Ovis aries*,
322 dog *Canis lupus*, and cat *Felis catus*). Sequences for *V. pourtalesii* and *P. ventilabrum*
323 samples and associated controls were compared against a marine vertebrate reference
324 database, the PCR positive control, and common domestic species. Full details of
325 database creation can be found in the Supporting Information.
326
327 *2.6.2 Data processing and taxonomic assignment*
328 Sequence data were automatically demultiplexed to separate (forward and reverse) fastq files
329 per library using the onboard Illumina MiSeq Reporter software. Library sequence reads were
330 further demultiplexed to sample using a custom Python script. *Tapirs*, a reproducible workflow
331 for the analysis of DNA metabarcoding data (https://github.com/EvoHull/Tapirs), was used for
332 processing and taxonomic assignment of demultiplexed sequence reads. *Tapirs* uses the

333 *Snakemake* workflow manager (Köster & Rahmann, 2012) and a *conda* virtual environment to
334 ensure software compatibility.

335 Raw reads were quality trimmed from the tail with a 5 bp sliding window (qualifying
336 phred score of Q30 and an average window phred score of Q30) using fastp (Chen et al., 2018),
337 allowing no more than 40% of the final trimmed read bases to be below Q30. Poly X/G tail
338 trimming (≥ 5 bp) removed any remaining Illumina iSeq™ 100 sequencing artefacts. Primers
339 were removed by trimming the first 18 bp and 20 bp of the forward and reverse reads
340 respectively. Reads were then tail cropped to a maximum length of 150 bp and reads shorter
341 than 90 bp were discarded. Sequence read pairs were merged into single reads using fastp,
342 provided there was a minimum overlap of 20 bp, no more than 5% mismatches and no more
343 than 5 mismatched bases between pairs. Only forward reads were kept from read pairs that
344 failed to be merged. A final length filter removed any reads longer than 200 bp to ensure
345 sequence lengths approximated the expected fragment size (~167 bp). Redundant sequences
346 were removed by clustering at 100% read identity and length (--derep_fulllength) in *VSEARCH*
347 (Rognes et al., 2016). Clusters represented by less than three sequences were omitted from
348 further processing. Reads were further clustered (--cluster_unoise) to remove redundancies due
349 to sequencing errors (retaining all cluster sizes). Retained sequences were screened for
350 chimeric sequences with *VSEARCH* (--uchime3_denovo).

351 The final clustered, non-redundant query sequences were then compared
352 against the corresponding reference database using *BLAST* (Zhang et al., 2000).
353 Taxonomic identity was assigned using a custom majority lowest common ancestor
354 (MLCA) approach based on the top 2% query *BLAST* hit bit-scores, with at least 90%
355 query coverage. A minimum *BLAST* hit identity filter of 80% was used for the Horniman
356 invasive sponge samples and the aquarium fish reference database, whereas a
357 minimum identity of 98% was used for *V. pourtalesii* and *P. ventilabrum* samples and
358 the marine vertebrate reference database. Of these filtered hits for both batches, 80% of
359 unique taxonomic lineages therein had to agree at descending taxonomic rank (domain,
360 phylum, class, order, family, genus, species) for it to be assigned a taxonomic identity. If
361 a query had a single *BLAST* hit, it was assigned directly to this taxon only if it met all
362 previous MLCA criteria. Read counts assigned to each taxonomic identity were
363 calculated from query cluster sizes. Lowest taxonomic rank was to species and
364 assignments higher than order were classed as unassigned.
365

366 **2.7 Data analysis**

367 *2.7.1 Dataset refinement*

368 Taxonomically assigned data was combined for downstream processing in R v3.6.3 (R
369 Core Team, 2020). All plots were produced using the packages ggplot2 v3.3.1
370 (Wickham, 2016) and ggpubr v0.3.0 (Kassambara, 2020). We split the data by Phase 1
371 (comparison of ethanol treatments and DNA extraction protocols) and Phase 2
372 (comparison of types and amounts of starting material). Assignments were corrected:
373 family and genera containing a single species were reassigned to that species, species
374 were reassigned to domestic subspecies, and misassignments were corrected. Manual
375 reassignment duplicated some assignments thus the read count data for these
376 assignments were merged. Proportional read counts for each taxon were calculated
377 from the total read counts per sample. Taxon-specific sequence thresholds (i.e.
378 maximum sequence frequency for each taxon detected in process controls) were
379 applied to the sponge samples to mitigate cross-contamination and false positives (Fig.
380 S1). A second sequence threshold (0.01%) was applied to the sponge samples to
381 mitigate low noise detections. Higher taxonomic assignments, iridescent catfish (PCR
382 positive control), human, domestic species, and terrestrial species were then removed.
383 Taxonomic assignments remaining in the refined dataset were predominantly of species
384 resolution and considered true positives (Fig. S2).

385

386 *2.7.2 Data summaries*

387 We examined overall fish and mammal taxon richness and read counts across samples
388 processed with each ethanol treatment and DNA extraction protocol in Phase 1 (Fig.
389 S3) and samples of each type and amount of starting material in Phase 2 (Fig. S4) for
390 each sponge species. We also examined detection consistency across extraction
391 replicates, in terms of taxon richness and number of replicates each taxon was detected
392 in, for samples processed with each ethanol treatment and DNA extraction protocol in
393 Phase 1 (Fig. S5) and samples of each type and amount of starting material in Phase 2
394 (Fig. S6) for each sponge species.

395

396 *2.7.2 Statistical analysis*

397 We compared total DNA yield, post-PCR DNA concentration, taxon richness and
398 proportional read counts of individual samples according to ethanol treatment and DNA
399 extraction protocol for Phase 1, and according to type and amount of starting material
400 for Phase 2. Only fish detections were used to compare extraction protocols and
401 ethanol treatments in terms of taxon richness, proportional read counts, and community
402 similarity as the metabarcoding primers used were designed to target fish. The data
403 were not normally distributed and the assumptions of a two-way Analysis of Variance
404 (ANOVA) were violated, thus statistical treatment followed Generalised Linear Models
405 (GLMs) with Quasi-Poisson error family or Gaussian error family (for total DNA yield
406 and post-PCR concentration), Poisson error family and the log-link function or negative
407 binomial error family (for taxon richness), or binomial error family (for proportional read
408 counts). If a significant effect was found using the *drop1* function from the package stats
409 v3.6.3, multiple pairwise comparisons were performed using the *pairs* function from the
410 package emmeans v1.7.2 (Lenth, 2022). Where assumptions of GLMs were violated,
411 the non-parametric Kruskal-Wallis Test from the package stats was used, and multiple
412 pairwise comparisons were performed using the non-parametric Dunn's Test (p-values
413 adjusted with the Benjamini-Hochberg method) from the package FSA v0.8.30 (Ogle et
414 al., 2020). In Phase 1, data were split by ethanol treatment to compare extraction
415 protocols and/or vice versa where significant effects were found using GLMs. In Phase
416 2, data were split by type of starting material to compare amounts of starting material
417 and/or vice versa where significant effects were found using GLMs.

418 The package betapart v1.5.1 (Baselga & Orme, 2012) was used to estimate total
419 beta diversity, partitioned by nestedness-resultant (i.e. community dissimilarity due to
420 taxon subsets) and turnover (i.e. community dissimilarity due to taxon replacement),
421 across samples in Phases 1 and 2 with the *beta.multi* function. Total beta diversity and
422 its two components (Jaccard dissimilarity) were estimated for sponge samples in both
423 phases using the *beta.pair* function. For total beta diversity and both partitions, we
424 calculated the average distance of group members to the group centroid in multivariate
425 space using the *betadisper* function in the package vegan v2.5-6 (Oksanen et al.,
426 2019). This was done for each group of samples within ethanol treatments and DNA

427 extraction protocols in Phase 1 and within types and amounts of starting material in
428 Phase 2. The distances of group members to the group centroid were subjected to an
429 ANOVA to test for homogeneity of multivariate dispersions (MVDISP), that is whether
430 the dispersions of one or more groups were different, prior to permutational multivariate
431 analysis of variance (PERMANOVA). Community dissimilarity for total beta diversity and
432 both partitions was visualised with Non-metric Multidimensional Scaling (NMDS) using
433 the *metaMDS* function with 1000 permutations, and tested statistically with
434 PERMANOVA using the function *adonis* in the package vegan. The PERMANOVA
435 included an interaction term between ethanol treatment and DNA extraction protocol for
436 Phase 1 and an interaction term between type and amount of starting material for
437 Phase 2. Pre-defined cut-off values were used for effect size, where PERMANOVA
438 results were interpreted as moderate and strong effects if $R^2 > 0.09$ and $R^2 > 0.25$
439 respectively. These values are broadly equivalent to correlation coefficients of $r = 0.3$
440 and 0.5 which represent moderate and strong effects accordingly (Nakagawa & Cuthill,
441 2007; Macher et al., 2018).
442
443
444 # 3. Results
445
446 ## 3.1 Sequencing performance
447 The Phase 1 library generated 3,757,770 raw sequence reads, of which 3,532,041
448 remained after quality control and merging. Following dereplication, denoising, and
449 chimaera detection, 2,995,356 reads (average read count 31,201 per sample including
450 controls) remained for taxonomic assignment. The Phase 2 library generated 3,005,701
451 raw sequence reads, of which 2,879,596 remained after quality control and merging.
452 Following dereplication, denoising, and chimaera detection, 2,170,282 reads (average
453 read count 16,823 per sample including controls) remained for taxonomic assignment.
454 For the Horniman invasive sponge samples, 85.05% of remaining reads were
455 taxonomically assigned while 15.95% could not be assigned to any taxonomic rank. For
456 *P. ventilabrum* and *V. pourtalesii* samples, 78.87% of remaining reads were
457 taxonomically assigned, while 21.13% could not be assigned to any taxonomic rank.

458 Contamination was restricted to human DNA in one extraction blank and one PCR
459 negative control in Phase 1, but naturally occurring marine species were also detected
460 in extraction blanks and PCR positive controls in Phase 2. Before dataset refinement,
461 29 taxa were identified in Phase 1 and 94 taxa were identified in Phase 2. After dataset
462 refinement, 18 taxa were identified in Phase 1 and 49 taxa were identified in Phase 2
463 (Fig. S2).

464 For simplicity, we only report significant results from the final statistical analyses
465 applied to the data. Full results from all statistical tests, including non-significant results
466 and violations of model assumptions, are detailed in the Supporting Information.
467 Statistical comparisons were not performed for DNA purity due to lack of variation in
468 A260/A280 and A260/A230 ratios between ethanol treatments, extraction protocols,
469 types of starting material, and amounts of starting material.

470

471 **3.2 Phase 1: ethanol treatment and DNA extraction protocol**

472 For the Horniman invasive sponge, total DNA yield significantly differed between
473 extraction protocols for the spin ethanol treatment (Kruskal-Wallis Test: H_4 = 11.314, P
474 = 0.023) and between ethanol treatments for the QBT (Kruskal-Wallis Test: H_2 = 7.200,
475 P = 0.027) and MUT (Kruskal-Wallis Test: H_2 = 7.261, P = 0.027) protocols. For the spin
476 treatment, none of the Dunn's Test adjusted p-values were significant, but the largest
477 differences in effect size were observed between the QBT and MUS protocols (Z = -
478 2.749, P = 0.060), the QBT and MUT protocols (Z = -2.107, P = 0.117), and the QBT
479 and QPS protocols (Z = 2.565, P = 0.052). For both the QBT and MUT protocols, the
480 spin treatment produced lower total DNA yields than the wet treatment (QBT: Z = -
481 2.683, P = 0.022; MUT: Z = -2.695, P = 0.021) (Fig. 2a). With *V. pourtalesii*, total DNA
482 yield significantly differed between extraction protocols for the wet (Kruskal-Wallis Test:
483 H_4 = 13.233, P = 0.010) and blot (Kruskal-Wallis Test: H_4 = 11.833, P = 0.019) ethanol
484 treatments. The QBT protocol produced higher total DNA yields than the QPS protocol
485 with the wet ethanol treatment (Dunn's Test: Z = 3.286, P = 0.010) and the MUT
486 protocol with the blot treatment (Dunn's Test: Z = -2.921, P = 0.035) (Fig. 2a).

487 For the Horniman invasive sponge, ethanol treatment influenced post-PCR DNA
488 concentration with the MUS protocol (Kruskal-Wallis Test: H_2 = 6.006, P = 0.050), where

489 the spin treatment produced lower post-PCR DNA concentrations than the wet
490 treatment (Dunn's Test: Z = -2.395, P = 0.050) (Fig. 2b). For *V. pourtalesii*, post-PCR
491 DNA concentration was positively influenced by the MUT (GLM: 0.669 ± 0.327, t =
492 2.044, P = 0.048) and QBT (GLM: 1.005 ± 0.311, t = 3.229, P = 0.003) protocols, but
493 the MUT protocol (12.13 ± 2.31 ng/µl) produced lower post-PCR DNA concentrations
494 than the QBT protocol (16.97 2.74 ng/µl; pairwise comparison: Z = -3.229, P = 0.011)
495 (Fig. 2b).

496 Only the MUS and QBT extraction protocols produced incidental mammal
497 detections (Fig. 3). Using a GLM, neither extraction protocol or ethanol treatment were
498 found to influence fish taxon richness for the Horniman invasive sponge, but model
499 assumptions were met. For *V. pourtalesii*, no effect of extraction protocol on fish taxon
500 richness was found using Kruskal-Wallis Test (Fig. 4a).

501 Proportional read counts significantly differed between ethanol treatments with
502 the MUS (Kruskal-Wallis Test: H_2 = 6.161, P = 0.046) and QPS (Kruskal-Wallis Test: H_2
503 = 6.006, P = 0.050) protocols for the Horniman invasive sponge. The wet treatment
504 produced higher proportional read counts than the spin treatment with the MUS protocol
505 (Dunn's Test: Z = -2.426, P = 0.046), and than the blot treatment with the QPS protocol
506 (Dunn's Test: Z = -2.395, P = 0.050) (Fig. 4b). For *V. pourtalesii*, no effect of extraction
507 protocol on proportional read counts was found, but proportional reads counts
508 significantly differed between ethanol treatments with the QBT-ZYMO protocol (Kruskal-
509 Wallis Test: H_2 = 7.624, P = 0.022), where the blot treatment produced higher
510 proportional read counts than the spin (Dunn's Test: Z = 2.391, P = 0.025) or wet
511 (Dunn's Test: Z = 2.391, P = 0.050) treatments (Fig. 4b).

512 MVDISP differed between DNA extraction protocols for total beta diversity for *V.*
513 *pourtalesii*, but not either partition for *V. pourtalesii* or the Horniman invasive sponge
514 (Table 1). MVDISP differed between ethanol treatments for total beta diversity and the
515 turnover component for the Horniman invasive sponge, and for the turnover and
516 nestedness-resultant components of beta diversity for *V. pourtalesii*. DNA extraction
517 protocol alone did not influence beta diversity for either sponge species (Table 2; Figs.
518 4ci-iii). Ethanol treatment had a strong positive influence on turnover and total beta
519 diversity of communities detected from the Horniman invasive sponge, but did not

520 influence beta diversity for *V. pourtalesii* (Table 2; Figs. 4ci-iii). DNA extraction protocol
521 combined with ethanol treatment also had a moderate positive effect on turnover for the
522 Horniman invasive sponge (Table 2; Figs. 4ci-iii). Therefore, taxa detected from the
523 Horniman invasive sponge with one ethanol treatment were replaced by different taxa
524 detected with other ethanol treatments, although this finding should be interpreted with
525 caution due to MVDISP heterogeneity.
526
527 **3.3 Phase 2: treatment and amount of sponge tissue**
528 Using GLMs, amount of starting material influenced total DNA yield for both *P.*
529 *ventilabrum* (F_2 = 16.485, P < 0.001) and *V. pourtalesii* (F_2 = 3.795, P = 0.030), but type
530 of starting material influenced total DNA yield for *P. ventilabrum* only (F_1 = 14.641, P <
531 0.001). Model assumptions were met for *P. ventilabrum*. Total DNA yield was higher
532 using homogenised tissue (0.460 ± 0.120, t = 3.840, P < 0.001) and increased as the
533 amount of starting material increased (amount 2: 0.451 ± 0.136, t = 3.321, P = 0.002;
534 amount 3: 0.981 ± 0.165, t = 5.931, P < 0.001). Dried tissue (30.30 ± 2.81 ng/µl)
535 produced lower total DNA yields than homogenate (48.00 ± 3.82 ng/µl; Z = -3.840, P <
536 0.001). Amount 3 (63.10 ± 7.96 ng/µl) produced higher total DNA yields than amount 2
537 (37.10 ± 3.18 ng/µl; Z = -3.474, P = 0.002) or amount 1 (23.70 ± 2.53 ng/µl; Z = -5.931,
538 P < 0.001). Amount 1 also produced lower total DNA yields than amount 2 (Z = -3.321,
539 P = 0.003) (Fig. 5a). Model assumptions were not met for *V. pourtalesii* thus Kruskal-
540 Wallis Test was applied, but amount of starting material was not found to influence total
541 DNA yield.
542 Using GLMs, amount of starting material did not influence post-PCR DNA
543 concentration for either *P. ventilabrum* or *V. pourtalesii*, and type of starting material
544 influenced post-PCR DNA concentration for *P. ventilabrum* only (F_1 = 5.054, P = 0.031).
545 Model assumptions for *P. ventilabrum* were violated thus Kruskal-Wallis Test was
546 applied, but no effect of type of starting material was found (Fig. 5b).
547 *V. pourtalesii* produced more incidental mammal detections than *P. ventilabrum*
548 (Fig. 6). Using GLMs, neither type or amount of starting material were found to influence
549 fish taxon richness for *P. ventilabrum*, and type of starting material influenced fish taxon
550 richness for *V. pourtalesii* only (χ_1 = 6.451, P = 0.011). Type of starting material had a

551 positive influence on taxon richness (0.476 ± 0.190, t = 2.514, P = 0.012), where taxon
552 richness was higher using homogenate (2.780 ± 0.334) than dried tissue (1.720 ±
553 0.253; Z = -2.514, P = 0.012) (Fig. 6a).

554 Using GLMs, neither type or amount of starting material were found to influence
555 fish proportional read counts for *P. ventilabrum*, and only amount of starting material
556 influenced fish proportional read counts for *V. pourtalesii* (F_2 = 3.266, P = 0.047). Model
557 assumptions for *V. pourtalesii* were violated thus Kruskal-Wallis Test was applied, but
558 no effect of amount of starting material was found (Fig. 6b).

559 MVDISP did not differ between types of starting material for total beta diversity or
560 its partitions for either sponge. MVDISP did not differ between amounts of starting
561 material for total beta diversity or its partitions for *V. pourtalesii*, but did differ between
562 amounts for total beta diversity and turnover for *P. ventilabrum* (Table 3). For *P.*
563 *ventilabrum*, amount of starting material influenced turnover and total beta diversity, but
564 type of starting material did not influence beta diversity (Table 4; Figs. 7ci-iii). For *V.*
565 *pourtalesii*, type of starting material influenced turnover and total beta diversity, but
566 amount of starting material did not influence beta diversity (Table 4; Figs. 7ci-iii).
567 Therefore, taxa detected with one type or amount of starting material were replaced by
568 different taxa detected with the other type (*V. pourtalesii*) or amounts (*P. ventilabrum*) of
569 starting material, although the latter should be interpreted with caution due to MVDISP
570 heterogeneity.

571

572

573 ## 4. Discussion

574

575 We have illustrated how the performance of sponge nsDNA metabarcoding can be
576 affected by DNA isolation procedures, specifically treatment to remove residual ethanol
577 prior to DNA extraction, DNA extraction protocol, type of starting material, and amount
578 of starting material. These effects can also be sponge-specific. Our experiments have
579 provided an optimised DNA isolation procedure that will inform the development and
580 application of sponge nsDNA metabarcoding for aquatic biodiversity assessment.

581

582 **4.1 Phase 1: ethanol treatment and DNA extraction protocol**

583 The QBT and MUT extraction protocols outperformed others for different ethanol
584 treatments of sponge tissue in terms of total DNA yield and post-PCR DNA
585 concentration. Lower DNA purity was observed with the MUT and QBT-ZYMO protocols
586 for the wet ethanol treatments of *V. pourtalesii* tissue and Horniman invasive sponge
587 tissue respectively, but otherwise extraction protocols produced consistent purity ratios.
588 Extraction protocol was not found to influence fish taxon richness, proportional read
589 counts or community composition (arguably the most important extraction success
590 criteria) for either sponge species.

591 The spin ethanol treatment produced lower total DNA yields and post-PCR DNA
592 concentrations with certain extraction protocols for the Horniman invasive sponge, but
593 no differences were observed between ethanol treatments for *V. pourtalesii*. Ethanol
594 treatment did not influence fish taxon richness for either sponge species, but did
595 influence fish proportional read counts with certain extraction protocols, where the wet
596 or blot treatments produced higher fish proportional read counts than the spin treatment.
597 Ethanol treatment influenced community composition of Horniman invasive sponge
598 tissue but not *V. pourtalesii* tissue, where different taxa were detected in Horniman
599 invasive sponge tissue with different ethanol treatments.

600 DNA extraction protocols have been found to influence the performance of
601 metabarcoding across sample types and habitats (Deiner et al., 2018; Jeunen et al.,
602 2019; Johnson et al., 2019; Geraldi et al., 2019; Sanches & Schreier, 2020), thus it was
603 not surprising that this also turned out to be the case for sponge tissue. All extraction
604 protocols produced some biodiversity information, but some protocols were more
605 effective than others, which has implications for detection of rare and/or low density
606 species in diverse and complex marine systems using sponge nsDNA metabarcoding.
607 Our findings echo those of Sellers et al. (2018), where the Mu-DNA protocols were
608 comparable to or outperformed commercial kits, and other studies which found the QBT
609 protocol produced higher DNA yield than other commercial kits and unbranded
610 extraction protocols (Hinlo et al., 2017; Deiner et al., 2018; Jeunen et al., 2019). Mu-
611 DNA is advantageous due to adaptability and cost-efficiency. Although an inhibitor
612 removal step is not included in the tissue protocol published in Sellers et al. (2018), it

613 can be added from the soil or water protocols, which is desirable for obtaining high DNA
614 yields whilst removing inhibitors (Jeunen et al., 2019). A QBT extraction costs four times
615 as much as a MUT extraction (Sellers et al., 2018), and requires a commercial post-
616 extraction inhibitor removal kit (e.g. Zymo Research) to remove potential PCR inhibitors.
617 Reduced cost means that the Mu-DNA protocols are more scalable for large projects
618 and buffer volumes can be adjusted proportionately to the amount of starting material
619 without quickly depleting budgets.

620 Residual ethanol in tissue can retain unique DNA that may leech from tissue into
621 ethanol (Zizka et al., 2018; Marquina et al., 2019; Jeunen et al., 2019; Derycke et al.,
622 2021; Persaud et al., 2021; Wang et al., 2021), but also react with buffers used for DNA
623 extraction and/or inhibit PCR amplification downstream (Schrader et al., 2012). Our
624 DNA concentration results seem to indicate that residual ethanol contained DNA unique
625 and/or additional to that present in sponge tissue (at least for the Horniman invasive
626 sponge) as higher total DNA yields and post-PCR DNA concentrations were observed
627 with the wet ethanol treatment. However, this did not translate to higher fish taxon
628 richness, and conflicting effects were found for fish proportional read counts from each
629 sponge species, where the wet ethanol treatment generated higher proportional read
630 counts for the Horniman invasive sponge and the blot ethanol treatment produced
631 higher proportional read counts for *V. pourtalesii*. Beta diversity analyses also indicated
632 that different taxa were detected using different ethanol treatments for the Horniman
633 invasive sponge. This reaffirms findings of studies on bulk invertebrate samples (Zizka
634 et al., 2018; Derycke et al., 2021; Persaud et al., 2021; Wang et al., 2021) and marine
635 sponges (Jeunen et al., 2019) where eDNA from ethanol preservative detected different
636 taxa to DNA from tissue, although eDNA from ethanol preservative generally detected
637 more taxa than bulk invertebrate tissue but less taxa than marine sponge tissue.

638 Sponge species, extraction protocol and ethanol treatments appear to have
639 specific interactions and effects on total DNA yield, DNA purity, target DNA
640 concentration, taxon recovery, and sequence generation. This has been observed for
641 aquatic eDNA in terms of filter material and extraction protocol (Deiner et al., 2017),
642 although other studies have suggested that these aspects of the eDNA workflow do not
643 interact with each other (Sanches & Schreier, 2020). A more elaborate experiment

644 incorporating numerous individuals of different sponge species would be required to test
645 this hypothesis. Based on our results, we recommend that these future studies focus on
646 comparing tissue-based extraction protocols and DNA extraction from ethanol alone
647 versus sponge tissue alone versus ethanol and sponge tissue combined. For the
648 purposes of experimental Phase 2 in the present study, blotting sponge tissue dry after
649 ethanol preservation followed by DNA extraction with the MUT protocol appeared to be
650 the best combination for extracting DNA from tissue of naturally occurring marine
651 sponges like *V. pourtalesii*.
652
653 **4.2 Phase 2: treatment and amount of sponge tissue**
654 Homogenised tissue produced higher total DNA yields for *P. ventilabrum* tissue but not
655 *V. pourtalesii* tissue. DNA from homogenised tissue from both sponge species was of
656 lower purity. Type of starting material did not influence post-PCR DNA concentration for
657 either sponge species. Homogenised tissue resulted in higher fish taxon richness for *V.*
658 *pourtalesii*, although not *P. ventilabrum*, and type of starting material did not influence
659 fish proportional read counts for either sponge species. Different communities were
660 observed using different types of starting material for *V. pourtalesii*, but community
661 composition from homogenate and dried tissue was highly consistent for *P. ventilabrum*.
662 Larger amounts of starting material produced higher total DNA yields for *P.*
663 *ventilabrum* tissue, but not *V. pourtalesii* tissue. DNA purity was consistent across
664 different amounts of starting material for both sponge species. Amount of starting
665 material did not influence post-PCR DNA concentration, fish taxon richness or fish
666 proportional read counts for either sponge species. However, different communities
667 were observed using different amounts of starting material for *P. ventilabrum*, but not for
668 *V. pourtalesii*.
669 Homogenisation is often used with complex sample types to account for uneven
670 distribution of organisms/DNA throughout the sample by facilitating the release of DNA
671 via bead beating, blending or vortexing, but introduces contamination risk via more
672 handling of samples and reusable equipment (Gosselin et al., 2017; Marquina et al.,
673 2019; Pereira-da-Conceicoa et al., 2020; Hestetun et al., 2021; Zizka et al., 2022). Until
674 now, the distribution of eDNA in marine sponge tissue and whether homogenisation was

675 required was unknown. Our results suggest that eDNA has a relatively even distribution
676 in sponge tissue. Type of starting material did not influence fish taxon richness for *P.*
677 *ventilabrum* or fish proportional read counts for either sponge species. Although higher
678 taxon richness was observed with homogenised *V. pourtalesii* tissue, the gains were
679 marginal compared to dried *V. pourtalesii* tissue. A homogenous distribution of eDNA in
680 sponge tissue offers additional advantages to using these natural samplers for marine
681 biodiversity assessment over aquatic eDNA which can exhibit substantial spatial
682 heterogeneity (Bruce et al., 2021). Dried tissue appears to be sufficient for taxon
683 recovery and would avoid the contamination risk associated with homogenisation.
684 However, sponge nsDNA accumulation and degradation rates still require investigation
685 to quantify temporal variation in eDNA signals from these natural samplers (Jeunen et
686 al., 2021) compared to aquatic eDNA accumulation and degradation rates which differ
687 by target taxa and habitat (Harrison et al., 2019). This has been done in an artificial
688 capacity (Cai et al., 2022), but must also be investigated in natural settings.
689 Total DNA yield increased as volume of homogenate and weight of dry tissue
690 increased. This relationship was expected (bar the risk of introducing inhibitors by using
691 too much tissue for DNA extraction) as using more starting material was found to
692 improve detection success for metabarcoding of bulk invertebrate (Zizka et al., 2022)
693 and soil (Kirse et al., 2021) samples. However, more starting material did not result in
694 higher fish taxon recovery or fish proportional read counts, which were consistent
695 across volumes of homogenate and weights of dried tissue. Different amounts of
696 starting material also produced similar communities for V. pourtalesii, albeit different
697 communities were observed with different amounts of P. ventilabrum tissue. Therefore,
698 there seems to be little benefit to using more starting material in sponge nsDNA
699 metabarcoding as more complex extraction procedures are required to handle larger
700 quantities of tissue. Instead, it may be more advantageous to invest effort in performing
701 more extraction replicates with smaller amounts of starting material (Lanzén et al.,
702 2017; Hestetun et al., 2021).
703

704 **4.3 Future directions**

705 Our findings have advanced the field of sponge nsDNA metabarcoding and emphasise
706 the importance of sequencing samples used in comparisons of DNA extraction
707 procedures rather than relying only on DNA concentration or purity values and gel
708 images. Nevertheless, a number of unanswered questions remain. The sample size
709 used in our study for each sponge, ethanol treatment, and DNA extraction protocol as
710 well as type and amount of starting material was small due to the various permutations
711 on DNA extraction methods being investigated and the available budget. Thousands of
712 sponges occupy marine systems at every depth and latitude across the globe (Morganti
713 et al., 2021). They can vary in shape, size, and hardness giving rise to variable
714 characteristics, including height, diameter, depth, growth form (free-standing,
715 encrusting, tubular, massive), substrate (hard rock, soft sediment, animal, floating
716 debris), spicules (spongin fibre, calcium carbonate, silica), colour, symmetry (radial
717 symmetry or asymmetrical), body plan (Asconoid, Syconoid, Leuconoid), and number
718 and cross-sectional area of oscula. They also differ in terms of reproduction (sexual or
719 asexual), feeding (carnivorous or herbivorous), defence (toxicity or boring activity),
720 disease (fungi, viruses, cyanobacteria), and symbionts (fish, shrimps, plants, bacteria,
721 algae) (Kahn et al., 2015; Kandler et al., 2019; Morganti et al., 2021). Only by pooling
722 empirical nsDNA-inferred taxon inventories from a variety of wild sponges in natural
723 settings will it be possible to gauge a more comprehensive understanding of the
724 robustness and reliability of this approach across habitats and geographical regions.

725 Different sponges may not have the same pumping and filtration efficiency or
726 even have viable tissue for DNA extraction due to microbial loads and PCR inhibitors
727 (Cai et al., 2022). Therefore, future research must identify the best natural samplers in
728 both benthic and pelagic habitats from those sponges that are accessible and can
729 withstand repeat sampling as well as understand how different sponge characteristics
730 influence eDNA recovery (Mariani et al., 2019; Jeunen et al., 2021). The spatial and
731 temporal resolution of sponge nsDNA under different environmental conditions and
732 across seasons must be investigated as we do not know the spatial extent or recency of
733 eDNA trapped in sponge tissue (Turon et al., 2020) and whether eDNA in sponge tissue
734 is affected by the same factors that hasten degradation of aquatic eDNA (e.g.

735 temperature, pH, salinity, microbial activity), but these factors are known to influence the
736 pumping and filtration action of sponges (Gökalp et al., 2020; Morganti et al., 2021). It is
737 also undetermined whether sponges exhibit selectivity for eDNA particles of different
738 sizes akin to filter membranes used for aquatic eDNA (Turner et al., 2014) and the
739 sampling effort (whether replicates from the same individual or replicates from different
740 individuals) required for sponges to achieve set species detection probabilities (Turon et
741 al., 2020).

742 In addition to these biological considerations, there are further technical
743 considerations for sponge nsDNA metabarcoding. Many species were only detected in
744 one extraction replicate performed on tissue from wild *V. pourtalesii* specimens,
745 emphasising the potential stochasticity of sponge tissue and need for extraction
746 replicates, similar to sediment (Lanzén et al., 2017; Hestetun et al., 2021) and bulk
747 invertebrate samples (Elbrecht & Steinke, 2018). Marker and primer choice is crucial to
748 metabarcoding success. Existing sponge nsDNA studies have focused on 12S and 16S
749 rRNA primers for the detection of fish (Mariani et al., 2019; Turon et al., 2020; Jeunen et
750 al., 2021), but as future studies diverge and other markers are explored, primers that
751 minimise non-target amplification (Collins et al., 2019) and blocking primers
752 (Boessenkool et al., 2012) to prevent amplification of DNA from the sponge itself could
753 become essential (Jeunen et al., 2021).

754 Finally, sequencing depth may play an important role in detection of low
755 concentration eDNA signals in sponge tissue, whether these represent species that are
756 rare, small or release less DNA into the environment. Very few studies have compared
757 sequencing platforms for metabarcoding, but comparisons have been made for the
758 popular Illumina MiSeq with the larger Illumina NovaSeq (Singer et al., 2019) and with
759 the smaller Illumina iSeq 100 (Nakao et al., 2021). The Illumina NovaSeq detected more
760 species from seawater than the Illumina MiSeq, and higher sequencing depth was also
761 found to improve species detection from sediment (Lanzén et al., 2017) and faecal
762 (Alberdi et al., 2018) samples with metabarcoding. However, species detection rates
763 and community composition were highly similar between the Illumina MiSeq and
764 Illumina iSeq 100 (Nakao et al., 2021). Given these disparities, future studies should
765 carefully consider sequencing depth for nsDNA metabarcoding when testing the

766 performance of new natural samplers and addressing the outstanding questions
767 identified here.
768 These biological and technical considerations are well-exemplified by the
769 Horniman invasive sponge in our study which detected 4 of 25 (after dataset
770 refinement) or 11 of 25 (before dataset refinement) species present in the aquarium
771 tank. The sponge individuals were relatively young, having regrown over a few weeks
772 after attempts to remove the sponge from the aquarium tank. This should have been
773 enough time for eDNA to accumulate in the tissues of these sponges (Cai et al., 2022),
774 but they may not have reached a large enough size to adequately pump and filter eDNA
775 from the volume of water held in the tank. The BLAST identity for taxonomic assignment
776 also had to be reduced to 80% in order to recover more species known to be present in
777 the aquarium tank. This suggests that the primers used may have insufficient taxonomic
778 resolution to discriminate between tropical fish species, namely members of the families
779 Acanthuridae, Chaetodontidae, Labridae, Pomacanthidae, Serranidae and Siganidae.
780
781
782 ## 5. Conclusions
783
784 Our study represents another step forward in the development of sponge nsDNA
785 metabarcoding as a means to assess aquatic biodiversity, and highlights how
786 methodological choices (specifically DNA isolation procedures) can influence the taxa
787 and subsequent communities recovered using this approach. Based on our results, we
788 recommend blotting sponge tissue dry to remove ethanol preservative prior to extraction
789 with a tissue-based protocol, whether commercial or non-commercial. However, we
790 note that more extensive comparisons of eDNA in ethanol preservative only, sponge
791 tissue only, and ethanol and sponge tissue combined extracted with tissue-based
792 protocols would be worthwhile. Homogenisation does not seem to be necessary for
793 sponges, eliminating a time-consuming step that presents a contamination risk. Amount
794 of starting material did not substantially influence taxon recovery or community
795 composition, but further comparisons of multiple extraction replicates with small
796 amounts of sponge tissue versus a single extraction with a large amount of tissue are

needed to determine whether time is better invested in simpler but more numerous extractions or more complex but fewer extractions. With these additional comparative studies, we will have a comprehensive understanding of the optimal DNA isolation procedure for sponge nsDNA metabarcoding which future investigations, whether of a biological or technical nature, can draw upon to advance this growing field.

Acknowledgements

This study was funded by the UK Natural Environment Research Council grant NE/T007028/1. A.R. was also supported by the Spanish Ministry of Science and Innovation grant (RYC2018-024247-I) and CSIC intramural grant (202030E006). M.B.A. was supported by the National Agency for Research and Development (ANID), fellowship program Postdoctorado en el extranjero/2019 - 74200143.

Author Contributions

S.M. and L.R.H. designed the study; J.C., A.R., and B.M. facilitated sample collection; L.R.H, E.F.N. and A.C. performed the laboratory work with guidance from G.S., M.B.A. and A.R.; G.S. performed the bioinformatic analyses; S.M. and A.R. obtained the funding; L.R.H. analysed the data and wrote the manuscript, which all authors revised.

Data Accessibility

Raw sequence reads have been archived on the NCBI Sequence Read Archive (BioProject: PRJNA854174; BioSamples: SAMN29421799 - SAMN29421894 [Phase 1] and SAMN29444657 - SAMN29444784 [Phase 2]; SRA accessions: SRR19906974 - SRR19907069 [Phase 1] and SRR19912485 - SRR19912613 [Phase 2]). Scripts and

827 corresponding data have been deposited in a dedicated Zenodo repository
828 (https://doi.org/10.5281/zenodo.6815267).
829
830
831 ## References
832

833 Alberdi, A., Aizpurua, O., Gilbert, M.T.P., & Bohmann, K (2018) Scrutinizing key steps
834 for reliable metabarcoding of environmental samples. *Methods in Ecology and*
835 *Evolution*, **9**, 134–147. https://doi.org/10.1111/2041-210X.12849
836 Baselga, A., & Orme, C.D.L. (2012) betapart: an R package for the study of beta
837 diversity: Betapart package. *Methods in Ecology and Evolution*, **3**, 808–812.
838 https://doi.org/10.1111/j.2041-210X.2012.00224.x
839 Bessey, C., Gao, Y., Truong, Y.B., Miller, H., Jarman, S.N., & Berry, O. (2021)
840 Comparison of materials for rapid passive collection of environmental DNA. *bioRxiv*,
841 2021.11.08.467842. https://doi.org/10.1111/1755-0998.13640
842 Bessey, C., Jarman, S.N., Simpson, T., Miller, H., Stewart, T., Keesing, J.K., & Berry,
843 O. (2021) Passive eDNA collection enhances aquatic biodiversity analysis.
844 *Communications Biology*, **4**, 236. https://doi.org/10.1038/s42003-021-01760-8
845 Boessenkool, S., Epp, L.S., Haile, J., Bellemain, E., Edwards, M., Coissac, E.,
846 Willerslev, E., & Brochmann, C. (2012) Blocking human contaminant DNA during
847 PCR allows amplification of rare mammal species from sedimentary ancient DNA.
848 *Molecular Ecology*, **21**, 1806–1815. https://doi.org/10.1111/j.1365-
849 294X.2011.05306.x
850 Bronner, I.F., Quail, M.A., Turner, D.J., & Swerdlow, H. (2009) Improved Protocols for
851 Illumina Sequencing. *Current Protocols in Human Genetics*, **18**, 18.2.1–18.2.42.
852 https://doi.org/10.1002/0471142905.hg1802s62
853 Bruce, K., Blackman, R., Bourlat, S.J., Hellström, A.M., Bakker, J., Bista, I., Bohmann,
854 K., Bouchez, A., Brys, R., Clark, K., Elbrecht, V., Fazi, S., Fonseca, V., Hänfling, B.,
855 Leese, F., Mächler, E., Mahon, A.R., Meissner, K., Panksep, K., Pawlowski, J.,
856 Schmidt Yáñez, P., Seymour, M., Thalinger, B., Valentini, A., Woodcock, P.,
857 Traugott, M., Vasselon, V., & Deiner, K. (2021) A practical guide to DNA-based
858 methods for biodiversity assessment. *Advanced Books*.
859 https://doi.org/10.3897/ab.e68634
860 Cai, W., Harper, L.R., Neave, E.F., Shum, P., Craggs, J., Arias, M.B., Riesgo, A., &
861 Mariani, S. (2022) Environmental DNA persistence and fish detection in captive
862 sponges. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13677
863 Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018) fastp: an ultra-fast all-in-one FASTQ
864 preprocessor. *Bioinformatics*, **34**, i884–i890.
865 https://doi.org/10.1093/bioinformatics/bty560
866 Clare, E.L., Economou, C.K., Faulkes, C.G., Gilbert, J.D., Bennett, F., Drinkwater, R., &
867 Littlefair, J.E. (2021) eDNAir: proof of concept that animal DNA can be collected from
868 air sampling. *PeerJ*, **9**, e11030. https://doi.org/10.7717/peerj.11030
869 Collins, R.A., Bakker, J., Wangensteen, O.S., Soto, A.Z., Corrigan, L., Sims, D.W.,
870 Genner, M.J., & Mariani, S. (2019) Non-specific amplification compromises

environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution*, **10**, 1985–2001. https://doi.org/10.1111/2041-210X.13276

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N., Pfrender, M.E., & Bernatchez, L. (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, **26**, 5872–5895. https://doi.org/10.1111/mec.14350

Deiner, K., Lopez, J., Bourne, S., Holman, L., Seymour, M., Grey, E.K., Lacoursière, A., Li, Y., Renshaw, M.A., Pfrender, M.E., Rius, M., Bernatchez, L., & Lodge, D.M. (2018) Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. *Metabarcoding and Metagenomics*, **2**, e28963. https://doi.org/10.3897/mbmg.2.28963

Derycke, S., Maes, S., Van den Bulcke, L., Vanhollebeke, J., Wittoeck, J., Hillewaert, H., Ampe, B., Haegeman, A., Hostens, K., & De Backer, A. (2021) Detection of Macrobenthos Species With Metabarcoding Is Consistent in Bulk DNA but Dependent on Body Size and Sclerotization in eDNA From the Ethanol Preservative. *Frontiers in Marine Science*, **8**, 637858. https://doi.org/10.3389/fmars.2021.637858

Drinkwater, R., Williamson, J., Clare, E.L., Chung, A.Y.C., Rossiter, S.J., & Slade, E. (2021) Dung beetles as samplers of mammals in Malaysian Borneo—a test of high throughput metabarcoding of iDNA. *PeerJ*, **9**, e11897. https://doi.org/10.7717/peerj.11897

Elbrecht V., & Steinke, D. (2018) Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology*, **64**, 380–387. https://doi.org/10.1111/fwb.13220

Franklin, T.W., McKelvey, K.S., Golding, J.D., Mason, D.H., Dysthe, J.C., Pilgrim, K.L., Squires, J.R., Aubry, K.B., Long, R.A., Greaves, S.E., Raley, C.M., Jackson, S., MacKay, P., Lisbon, J., Sauder, J.D., Pruss, M.T., Heffington, D., & Schwartz, M.K. (2019) Using environmental DNA methods to improve winter surveys for rare carnivores: DNA from snow and improved noninvasive techniques. *Biological Conservation*, **229**, 50–58. https://doi.org/10.1016/j.biocon.2018.11.006

Geraldi, N.R., Díaz-Rúa, R., Shea, L.A., & Duarte, C.M. (2019) Performance of extraction methods for extracellular DNA from sediments across marine habitats. *Environmental DNA*, **4**, e6372. https://doi.org/10.1002/edn3.48

Gökalp, M., Kooistra, T., Rocha, M.S., Silva, T.H., Osinga, R., Murk, A.J., & Wijgerde, T. (2020) The Effect of Depth on the Morphology, Bacterial Clearance, and Respiration of the Mediterranean Sponge *Chondrosia reniformis* (Nardo, 1847). *Marine Drugs*, **18**. https://doi.org/10.3390/md18070358

Gosselin, E.N., Lonsinger, R.C., & Waits, L.P. (2017) Comparing morphological and molecular diet analyses and fecal DNA sampling protocols for a terrestrial carnivore: Noninvasive Diet Analysis Methodologies. *Wildlife Society Bulletin*, **41**, 362–369. https://doi.org/10.1002/wsb.749

Harper, L.R., Niemiller, M.L., Benito, J., Paddock, L.E., Knittle, E., Molano-Flores, B., & Davis, M.A. (2021) BeeDNA: microfluidic environmental DNA metabarcoding as a tool for connecting plant and pollinator communities. *bioRxiv*, 2021.11.11.468290. https://doi.org/10.1101/2021.11.11.468290

915 Harrison, J.B., Sunday, J.M., & Rogers, S.M. (2019) Predicting the fate of eDNA in the
916 environment and implications for studying biodiversity. *Proceedings of The Royal*
917 *Society B*, **286**, 20191409. https://doi.org/10.1098/rspb.2019.1409
918 Hermans, S.M., Buckley, H.L., & Lear, G. (2018) Optimal extraction methods for the
919 simultaneous analysis of DNA from diverse organisms and sample types. *Molecular*
920 *Ecology Resources*, **18**, 557–569. https://doi.org/10.1111/1755-0998.12762
921 Hestetun, J.T., Lanzén, A., Skaar, K.S., & Dahlgren, T.G. (2021) The impact of DNA
922 extract homogenization and replication on marine sediment metabarcoding diversity
923 and heterogeneity. *Environmental DNA*, **3**, 997–1006.
924 https://doi.org/10.1002/edn3.223
925 Hinlo, R., Gleeson, D., Lintermans, M., & Furlan, E. (2017) Methods to maximise
926 recovery of environmental DNA from water samples. *PLoS ONE*, **12**, e0179251.
927 https://doi.org/10.1371/journal.pone.0179251
928 Jeunen, G.-J., Cane, J.S., Ferreira, S., Strano, F., von Ammon, U., Cross, H., Day, R.,
929 Hesseltine, S., Ellis, K., Urban, L., Pearson, N., Olmedo-Rojas, P., Kardailsky, A.,
930 Gemmell, N.J., & Lamare, M. (2021) Assessing the utility of marine filter feeders for
931 environmental DNA (eDNA) biodiversity monitoring. *bioRxiv*, 2021.12.21.473722.
932 https://doi.org/10.1101/2021.12.21.473722
933 Jeunen, G.-J., Knapp, M., Spencer, H.G., Taylor, H.R., Lamare, M.D., Stat, M., Bunce,
934 M., & Gemmell, N.J. (2019) Species-level biodiversity assessment using marine
935 environmental DNA metabarcoding requires protocol optimization and
936 standardization. *Ecology and Evolution*, **9**, 1323–1335.
937 https://doi.org/10.1002/ece3.4843
938 Johnson, M.D., Cox, R.D., & Barnes, M.A. (2019) Analyzing airborne environmental
939 DNA: A comparison of extraction methods, primer type, and trap type on the ability to
940 detect airborne eDNA from terrestrial plant communities. *Environmental DNA*, **1**,
941 176–185. https://doi.org/10.1002/edn3.19
942 Kahn, A.S., Yahel, G., Chu, J.W.F., Tunnicliffe, V., & Leys, S.P. (2015) Benthic grazing
943 and carbon sequestration by deep-water glass sponge reefs. *Limnology and*
944 *Oceanography*, **60**, 78–88. https://doi.org/10.1002/lno.10002
945 Kandler, N.M., Wooster, M.K., Leray, M., Knowlton, N., de Voogd, N.J., Paulay, G., &
946 Berumen, M.L. (2019) Hyperdiverse Macrofauna Communities Associated with a
947 Common Sponge, Stylissa carteri, Shift across Ecological Gradients in the Central
948 Red Sea. *Diversity*, **11**, 18. https://doi.org/10.3390/d11020018
949 Kassambara A (2020) ggpubr: 'ggplot2' Based Publication Ready Plots. R package
950 version 0.4.0. https://CRAN.R-project.org/package=ggpubr
951 Kaunisto, K.M., Roslin, T., Sääksjärvi, I.E., & Vesterinen, E.J. (2017) Pellets of proof:
952 First glimpse of the dietary composition of adult odonates as revealed by
953 metabarcoding of feces. *Ecology and Evolution*, **7**, 8588–8598.
954 https://doi.org/10.1002/ece3.3404
955 Kinoshita, G., Yonezawa, S., Murakami, S., & Isagi, Y. (2019) Environmental DNA
956 Collected from Snow Tracks is Useful for Identification of Mammalian Species.
957 *Zoological Science*, **36**, 198–207. https://doi.org/10.2108/zs180172
958 Kirse, A., Bourlat, S.J., Langen, K., & Fonseca, V.G. (2021) Unearthing the potential of
959 soil eDNA metabarcoding—towards best practice advice for invertebrate biodiversity

assessment. *Frontiers in Ecology and Evolution*, **9**, 630560. https://doi.org/10.3389/fevo.2021.630560

Kirtane, A., Atkinson, J.D., & Sassoubre, L.M. (2020) Design and Validation of Passive Environmental DNA Samplers (PEDS) using Granular Activated Carbon (GAC) and Montmorillonite Clay (MC). *Environmental Science & Technology*, **54**, 11961–11970. https://doi.org/10.1021/acs.est.0c01863

Köster, J., & Rahmann, S. (2012) Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics*, **28**, 2520–2522. https://doi.org/10.1093/bioinformatics/bts480

Lanzén, A., Lekang, K., Jonassen, I., Thompson, E.M., & Troedsson, C. (2017) DNA extraction replicates improve diversity and compositional dissimilarity in metabarcoding of eukaryotes in marine sediments. *PLoS ONE*, **12**, e0179443. https://doi.org/10.1371/journal.pone.0179443

Lawson Handley, L. (2015) How will the "molecular revolution" contribute to biological recording? *Biological Journal of the Linnean Society*, **115**, 750–766. https://doi.org/10.1111/bij.12516

Lenth, R.V. (2022). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.2. https://CRAN.R-project.org/package=emmeans

Lynggaard, C., Bertelsen, M.F., Jensen, C.V., Johnson, M.S., Frøslev, T.G., Olsen, M.T., & Bohmann, K. (2022) Airborne environmental DNA for terrestrial vertebrate community monitoring. *Current Biology*, **32**, 701–707.e5. https://doi.org/10.1016/j.cub.2021.12.014

Macher, J.-N., Vivancos, A., Piggott, J.J., Centeno, F.C., Matthaei, C.D., & Leese, F. (2018) Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers. *Molecular Ecology Resources*, **18**, 1456–1468. https://doi.org/10.1111/1755-0998.12940

Majaneva, M., Diserud, O.H., Eagle, S.H.C., Hajibabaei, M., & Ekrem, T. (2018) Choice of DNA extraction method affects DNA metabarcoding of unsorted invertebrate bulk samples. *Metabarcoding and Metagenomics*, **2**, e26664. https://doi.org/10.3897/mbmg.2.26664

Mariani, S., Baillie, C., Colosimo, G., & Riesgo, A. (2019) Sponges as natural environmental DNA samplers. *Current Biology*, **29**, PR401–PR402. https://doi.org/10.1016/j.cub.2019.04.031

Mariani, S., Harper, L.R., Collins, R.A., Baillie, C., Wangensteen, O.S., McDevitt, A.D., Heddell-Cowie, M., & Genner, M.J. (2021) Estuarine molecular bycatch as a landscape-wide biomonitoring tool. *Biological Conservation*, **261**, 109287. https://doi.org/10.1016/j.biocon.2021.109287

Marquina, D., Esparza-Salas, R., Roslin, T., & Ronquist, F. (2019) Establishing arthropod community composition using metabarcoding: Surprising inconsistencies between soil samples and preservative ethanol and homogenate from Malaise trap catches. *Molecular Ecology Resources*, **19**, 1516–1530. https://doi.org/10.1111/1755-0998.13071

Massey, A.L., Bronzoni, R.V. de M., da Silva, D.J.F., Allen, J., de Lazari, P.R., dos Santos Filho, M., Canale, G.R., Bernardo, C.S.S., Peres, C.A., & Levi, T. (2021) Invertebrates for vertebrate biodiversity monitoring: comparisons using three insect taxa as iDNA samplers. *Molecular Ecology Resources*, **22**, 962–977. https://doi.org/10.1111/1755-0998.13525

Miya, M. (2021) Environmental DNA Metabarcoding: A Novel Method for Biodiversity Monitoring of Marine Fish Communities. *Annual Review of Marine Science*, **14**, 161–185. https://doi.org/10.1146/annurev-marine-041421-082251

Morganti, T.M., Ribes, M., Moskovich, R., Weisz, J.B., Yahel, G., & Coma, R. (2021) In situ Pumping Rate of 20 Marine Demosponges Is a Function of Osculum Area. *Frontiers in Marine Science*, **8**, 583188. https://doi.org/10.3389/fmars.2021.583188

Nakagawa, S., & Cuthill, I.C. (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, **82**, 591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

Nakao, R., Inui, R., Akamatsu, Y., Goto, M., Doi, H., & Matsuoka, S. (2021) Illumina iSeq 100 and MiSeq exhibit similar performance in freshwater fish environmental DNA metabarcoding. *Scientific Reports*, **11**, 15763. https://doi.org/10.1038/s41598-021-95360-5

Nørgaard, L., Olesen, C.R., Trøjelsgaard, K., Pertoldi, C., Nielsen, J.L., Taberlet, P., Ruiz-González, A., De Barba, M., & Iacolina, L. (2021) eDNA metabarcoding for biodiversity assessment, generalist predators as sampling assistants. *Scientific Reports*, **11**, 6820. https://doi.org/10.1038/s41598-021-85488-9

Ogle, D.H., Doll, J.C., Wheeler, P., & Dinno, A. (2021) FSA: Fisheries Stock Analysis. R package version 0.9.1. https://github.com/droglenc/FSA

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs E., & Wagner, H. (2020). vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Pereira-da-Conceicoa, L., Elbrecht, V., Hall, A., Briscoe, A., Barber-James, H., & Price, B. (2020) Metabarcoding unsorted kick-samples facilitates macroinvertebrate based biomonitoring with increased taxonomic resolution, while outperforming environmental DNA. *Environmental DNA*, **3,** 353–371. https://doi.org/10.1002/edn3.116

Persaud, S.F., Cottenie, K., & Gleason, J.E. (2021) Ethanol eDNA Reveals Unique Community Composition of Aquatic Macroinvertebrates Compared to Bulk Tissue Metabarcoding in a Biomonitoring Sampling Scheme. *Diversity*, **13**, 34. https://doi.org/10.3390/d13010034

Pompanon, F., Deagle, B.E., Symondson, W.O.C., Brown, D.S., Jarman, S.N., & Taberlet, P. (2012) Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology*, **21**, 1931–1950. https://doi.org/10.1111/j.1365-294X.2011.05403.x

Roger, F., Ghanavi, H.R., Danielsson, N., Wahlberg, N., Löndahl, J., Pettersson, L.B., Andersson, G.K.S., Boke Olén, N., & Clough, Y. (2022) Airborne environmental DNA metabarcoding for the monitoring of terrestrial insects—A proof of concept from the field. *Environmental DNA*. https://doi.org/10.1002/edn3.290

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, **4**, e2584. https://doi.org/10.7717/peerj.2584

Russo, T., Maiello, G., Talarico, L., Baillie, C., Colosimo, G., D'Andrea, L., Di Maio, F., Fiorentino, F., Franceschini, S., Garofalo, G., Scannella, D., Cataudella, S., & Mariani, S. (2020) All is fish that comes to the net: metabarcoding for rapid fisheries

catch assessment. *Ecological Applications,* **31**, e02273. https://doi.org/10.1002/eap.2273

Sanches, T.M., & Schreier, A.D. (2020) Optimizing an eDNA protocol for estuarine environments: Balancing sensitivity, cost and time. *PLoS ONE*, **15**, e0233522. https://doi.org/10.1371/journal.pone.0233522

Schnell, I.B., Thomsen, P.F., Wilkinson, N., Rasmussen, M., Jensen, L.R.D., Willerslev, E., Bertelsen, M.F., & Gilbert, M.T.P. (2012) Screening mammal biodiversity using DNA from leeches. *Current Biology*, **22**, R262–R263. https://doi.org/10.1016/j.cub.2012.02.058

Schrader, C., Schielke, A., Ellerbroek, L. & Johne, R. (2012) PCR inhibitors - occurrence, properties and removal. *Journal of Applied Microbiology*, **113**, 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Sellers, G.S., Di Muri, C., Gómez, A., & Hänfling, B. (2018) Mu-DNA: a modular universal DNA extraction method adaptable for a wide range of sample types. *Metabarcoding and Metagenomics*, **2**, e24556. https://doi.org/10.3897/mbmg.2.24556

Siegenthaler, A., Wangensteen, O.S., Soto, A.Z., Benvenuto, C., Corrigan, L., & Mariani, S. (2019) Metabarcoding of shrimp stomach content: Harnessing a natural sampler for fish biodiversity monitoring. *Molecular Ecology Resources*, **19**, 206–220. https://doi.org/10.1111/1755-0998.12956

Sigsgaard, E.E., Olsen, K., Hansen, M.D.D., Hansen, O.L.P., Høye, T.T., Svenning, J.-C., & Thomsen, P.F. (2020) Environmental DNA metabarcoding of cow dung reveals taxonomic and functional diversity of invertebrate assemblages. *Molecular Ecology*, **30**, 3374-3389. https://doi.org/10.1111/mec.15734

Singer, G.A.C., Fahner, N.A., Barnes, J.G., McCarthy, A., & Hajibabaei, M. (2019) Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. *Scientific Reports*, **9**, 5991. https://doi.org/10.1038/s41598-019-42455-9

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018) *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press.

Thalinger, B., Deiner, K., Harper, L.R., Rees, H.C., Blackman, R.C., Sint, D., Traugott, M., Goldberg, C.S., & Bruce, K. (2021) A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. *Environmental DNA*, **3**, 823–836. https://doi.org/10.1002/edn3.189

Thomsen, P.F., & Sigsgaard, E.E. (2019) Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. *Ecology and Evolution*, **9**, 1665–1679. https://doi.org/10.1002/ece3.4809

Tsuji, S., Takahara, T., Doi, H., Shibata, N., & Yamanaka, H. (2019) The detection of aquatic macroorganisms using environmental DNA analysis—A review of methods for collection, extraction, and detection. *Environmental DNA*, **1**, 99–108. https://doi.org/10.1002/edn3.21

Turner, C.R., Barnes, M.A., Xu, C.C.Y., Jones, S.E., Jerde, C.L., & Lodge, D.M. (2014) Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods in Ecology and Evolution*, **5**, 676–684. https://doi.org/10.1111/2041-210X.12206

Turon, M., Angulo-Preckler, C., Antich, A., Præbel, K., & Wangensteen, O.S. (2020) More Than Expected From Old Sponge Samples: A Natural Sampler DNA

Metabarcoding Assessment of Marine Fish Diversity in Nha Trang Bay (Vietnam). *Frontiers in Marine Science*, **7**, 605148. https://doi.org/10.3389/fmars.2020.605148

Valentin, R.E., Fonseca, D.M., Gable, S., Kyle, K.E., Hamilton, G.C., Nielsen, A.L., & Lockwood, J.L. (2020) Moving eDNA surveys onto land: Strategies for active eDNA aggregation to detect invasive forest insects. *Molecular Ecology Resources*, **20**, 746–755. https://doi.org/10.1111/1755-0998.13151

Wang, Y., Chen, K., Gao, J., Wang, M., Dong, J., Xie, Y., Giesy, J.P., Jin, X. & Wang, B. (2021) Environmental DNA of preservative ethanol performed better than water samples in detecting macroinvertebrate diversity using metabarcoding. *Diversity & Distributions*, **27**, 1989–2002. https://doi.org/10.1111/ddi.13284

Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000) A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, **7**, 203–214. https://doi.org/10.1089/10665270050081478

Zizka, V.M.A., Geiger, M.F., Hörren, T., Kirse, A., Noll, N.W., Schäffler, L., Scherges, A.M., & Sorg, M. (2022) Recommendations for tissue homogenisation and extraction in DNA metabarcoding of Malaise trap samples. *bioRxiv*, 2022.01.25.477667. https://doi.org/10.1101/2022.01.25.477667

Zizka, V.M.A., Leese, F., Peinert, B., & Geiger, M.F. (2018) DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method. *Genome*, **62**, 122–136. https://doi.org/10.1139/gen-2018-0048

1121 **Table 1.** Summary of analyses statistically comparing homogeneity of multivariate
1122 dispersions (MVDISP) between the communities (ANOVA) produced by different
1123 extraction protocols and ethanol treatments. Relative contributions of taxon turnover
1124 and nestedness-resultant to total beta diversity (Jaccard dissimilarity) for each protocol
1125 and treatment are given in brackets.
1126

	Mean distance to centroid ± SE	df	F	*P*
Horniman invasive sponge				
DNA extraction protocol				
Turnover		4	0.059	0.993
MUT (100.00%)	0.375 ± 0.268			
MUS (100.00%)	0.333 ± 0.267			
QBT (96.92%)	0.446 ± 0.276			
QBT-ZYMO (88.89%)	0.413 ± 0.054			
QPS (96.55%)	0.429 ± 0.286			
Nestedness-resultant		4	2.013	0.118
MUT (0.00%)	$4.939 \times 10^{-9} \pm 4.647 \times 10^{-17}$			
MUS (0.00%)	$9.697 \times 10^{-9} \pm 2.341 \times 10^{-16}$			
QBT (3.08%)	0.091 ± 0.015			
QBT-ZYMO (11.11%)	0.110 ± 0.046			
QPS (3.45%)	$8.700 \times 10^{-9} \pm 1.361 \times 10^{-16}$			
Total beta diversity		4	0.132	0.969
MUT (100.00%)	0.375 ± 0.268			
MUS (100.00%)	0.333 ± 0.267			
QBT (100.00%)	0.500 ± 0.199			
QBT-ZYMO (100.00%)	0.413 ± 0.053			
QPS (100.00%)	0.429 ± 0.286			
Ethanol treatment				
Turnover		2	12.343	**<0.001**
Blot (96.73%)	0.500 ± 0.286			
Spin (0.00%	$2.363 \times 10^{-9} \pm 7.259 \times 10^{-17}$			
Wet (0.00%)	$1.129 \times 10^{-9} \pm 1.784 \times 10^{-17}$			
Nestedness-resultant		2	0.294	0.747
Blot (3.27%)	$3.833 \times 10^{-6} \pm 2.959 \times 10^{-11}$			
Spin (100.00%)	0.039 ± 0.019			
Wet (100.00%)	0.036 ± 0.018			
Total beta diversity		2	8.593	**0.001**
Blot (100.00%)	0.500 ± 0.286			
Spin (100.00%)	0.039 ± 0.019			
Wet (100.00%)	0.036 ± 0.018			
Vazella pourtalesii				

DNA extraction protocol				
Turnover		4	1.902	0.204
MUT (98.52%)	0.572 ± 0.077			
MUS (100.00%)	0.500 ± 2.465 x 10^{-32}			
QBT (100.00%)	0.500 ± 0.000			
QBT-ZYMO (100.00%)	0.575 ± 2.115 x 10^{-5}			
QPS	0.000			
Nestedness-resultant		4	1.620	0.260
MUT (1.48%)	0.078 ± 0.006			
MUS (0.00%)	0.000 ± 0.000			
QBT (0.00%)	1.273 x 10^{-8} ± 1.819 x 10^{-18}			
QBT-ZYMO (0.00%)	6.332 x 10^{-9} ± 1.203 x 10^{-16}			
QPS	0.000			
Total beta diversity		4	5.107	**0.024**
MUT (100.00%)	0.593 ± 0.030			
MUS (100.00%)	0.500 ± 0.000			
QBT (100.00%)	0.500 ± 1.233 x 10^{-32}			
QBT-ZYMO (100.00%)	0.577 ± 3.513 x 10^{-9}			
QPS	0.000			
Ethanol treatment				
Turnover		2	10.920	**0.003**
Blot (99.36%)	0.651 ± 0.009			
Spin (98.08%	0.521 ± 0.079			
Wet (0.00%)	1.343 x 10^{-8} ± 0.000			
Nestedness-resultant		2	11.748	**0.002**
Blot (0.64%)	2.915 x 10^{-9} ± 5.948 x 10^{-17}			
Spin (1.92%)	0.137 ± 0.026			
Wet (100.00%)	0.333 ± 0.000			
Total beta diversity		2	4.055	0.051
Blot (100.00%)	0.635 ± 0.009			
Spin (100.00%)	0.559 ± 0.041			
Wet (100.00%)	0.333 ± 0.000			

1129 **Table 2.** Results of PERMANOVA used to compare variation in community composition
1130 produced by different extraction protocols and ethanol treatments, individually and
1131 combined.
1132

	df	F	R^2	*P*	Stress
Horniman invasive sponge					0
Turnover					
DNA extraction protocol	4	2.782	0.060	0.057	
Ethanol treatment	2	67.309	0.720	**0.001**	
DNA extraction protocol : Ethanol treatment	7	2.883	0.108	**0.025**	
Nestedness-resultant					0
DNA extraction protocol	4	0.670	0.122	0.569	
Ethanol treatment	2	-2.124	-0.193	0.947	
DNA extraction protocol : Ethanol treatment	7	0.361	0.115	0.704	
Total beta diversity					0
DNA extraction protocol	4	2.152	0.063	0.087	
Ethanol treatment	2	46.617	0.677	**0.001**	
DNA extraction protocol : Ethanol treatment	7	2.132	0.108	0.054	
Vazella pourtalesii					
Turnover					0
DNA extraction protocol	4	0.677	0.263	0.950	
Ethanol treatment	2	1.117	0.217	0.365	
DNA extraction protocol : Ethanol treatment	2	0.683	0.133	0.830	
Nestedness-resultant					5.113×10^{-5}
DNA extraction protocol	4	1.984×10^{-15}	0.895	0.112	
Ethanol treatment	2	-2.317×10^{-15}	-0.523	0.877	
DNA extraction protocol : Ethanol treatment	2	2.783×10^{-15}	0.628	0.098	
Total beta diversity					0
DNA extraction protocol	4	0.772	0.288	0.886	
Ethanol treatment	2	1.007	0.188	0.427	
DNA extraction protocol : Ethanol treatment	2	0.816	0.152	0.723	

1133
1134

1135 **Table 3.** Results of ANOVA used to compare homogeneity of multivariate dispersions
1136 (MVDISP) between the communities produced by different types and amounts of
1137 starting material. Relative contributions of taxon turnover and nestedness-resultant to
1138 total beta diversity (Jaccard dissimilarity) for each type and amount of starting material
1139 are given in brackets.
1140

	Mean distance to centroid ± SE	df	F	*P*
Phakellia ventilabrum				
Type of starting material				
Turnover		1	0.172	0.681
Dried tissue (85.02%)	0.115 ± 0.021			
Homogenate (88.16%)	0.131 ± 0.010			
Nestedness-resultant		1	1.729	0.196
Dried tissue (14.98%)	0.074 ± 0.004			
Homogenate (11.84%)	0.051 ± 0.002			
Total beta diversity		1	0.059	0.809
Dried tissue (100.00%)	0.165 ± 0.026			
Homogenate (100.00%)	0.175 ± 0.011			
Amount of starting material				
Turnover		2	6.141	**0.005**
1 (79.35%)	0.106 ± 0.005			
2 (79.46%)	0.096 ± 0.004			
3 (92.15%)	0.247 ± 0.038			
Nestedness-resultant		2	0.425	0.657
1 (20.65%)	0.073 ± 0.005			
2 (20.54%)	0.059 ± 0.002			
3 (7.85%)	0.049 ± 0.002			
Total beta diversity		2	4.898	**0.013**
1 (100.00%)	0.153 ± 0.009			
2 (100.00%)	0.137 ± 0.005			
3 (100.00%)	0.287 ± 0.030			
Vazella pourtalesii				
Type of starting material				
Turnover		1	0.327	0.570
Dried tissue (97.03%)	0.489 ± 0.059			
Homogenate (96.63%)	0.451 ± 0.046			

Nestedness-resultant		1	0.696	0.409
Dried tissue (2.97%)	0.146 ± 0.019			
Homogenate (3.37%)	0.186 ± 0.035			
Total beta diversity		1	0.143	0.707
Dried tissue (100.00%)	0.569 ± 0.024			
Homogenate (100.00%)	0.554 ± 0.014			
Amount of starting material				
Turnover		2	0.165	0.849
1 (96.29%)	0.513 ± 0.025			
2 (97.05%)	0.477 ± 0.055			
3 (95.02%)	0.480 ± 0.044			
Nestedness-resultant		2	0.645	0.530
1 (3.71%)	0.132 ± 0.010			
2 (2.95%)	0.177 ± 0.030			
3 (4.98%)	0.115 ± 0.037			
Total beta diversity		2	0.022	0.978
1 (100.00%)	0.569 ± 0.015			
2 (100.00%)	0.563 ± 0.021			
3 (100.00%)	0.559 ± 0.012			

1143 **Table 4.** Results of PERMANOVA used to compare variation in community composition
1144 according to type and amount of starting material, individually and combined.
1145

	df	F	R^2	*P*	Stress
Phakellia ventilabrum					
Turnover					0.065
Type of starting material	1	1.917	0.034	0.194	
Amount of starting material	2	7.513	0.269	**0.004**	
Type of starting material : Amount of starting material	2	1.981	0.071	0.141	
Nestedness-resultant					0.045
Type of starting material	1	1.941	0.054	0.194	
Amount of starting material	2	0.385	0.021	0.659	
Type of starting material : Amount of starting material	2	-0.756	-0.042	0.994	
Total beta diversity					0.077
Type of starting material	1	1.323	0.026	0.265	
Amount of starting material	2	5.571	0.220	**0.001**	
Type of starting material : Amount of starting material	2	1.638	0.065	0.135	
Vazella pourtalesii					
Turnover					0.119
Type of starting material	1	3.058	0.067	**0.012**	
Amount of starting material	2	0.664	0.030	0.757	
Type of starting material : Amount of starting material	2	0.507	0.022	0.863	
Nestedness-resultant					0.190
Type of starting material	1	-6.207	-0.145	1.000	
Amount of starting material	2	2.410	0.113	0.196	
Type of starting material : Amount of starting material	2	2.105	0.098	0.249	
Total beta diversity					0.113
Type of starting material	1	2.055	0.045	**0.019**	
Amount of starting material	2	1.026	0.045	0.396	
Type of starting material : Amount of starting material	2	0.751	0.033	0.849	

1146
1147

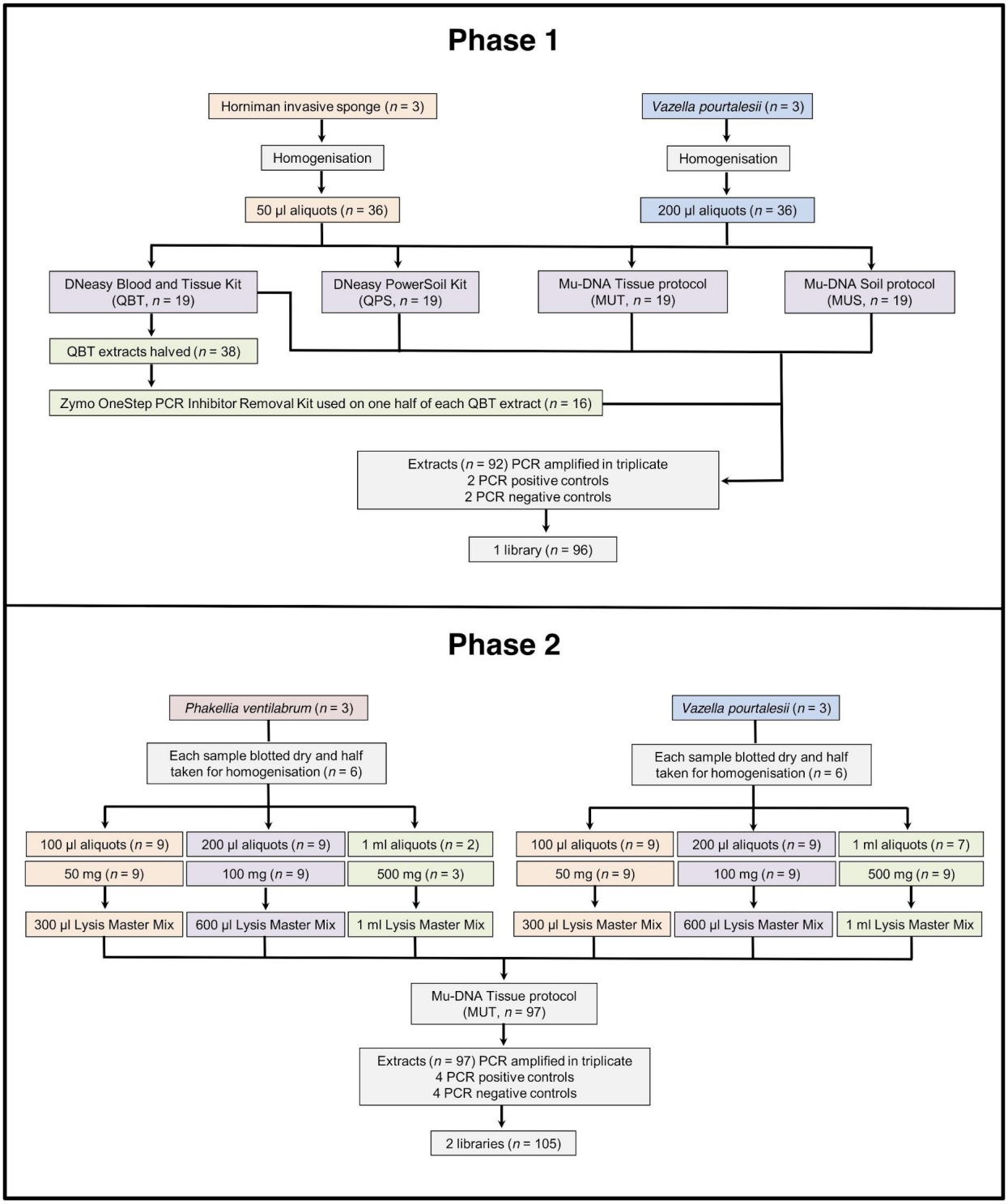

Figure 1. Flow diagram illustrating the laboratory workflow for the phased experiment.

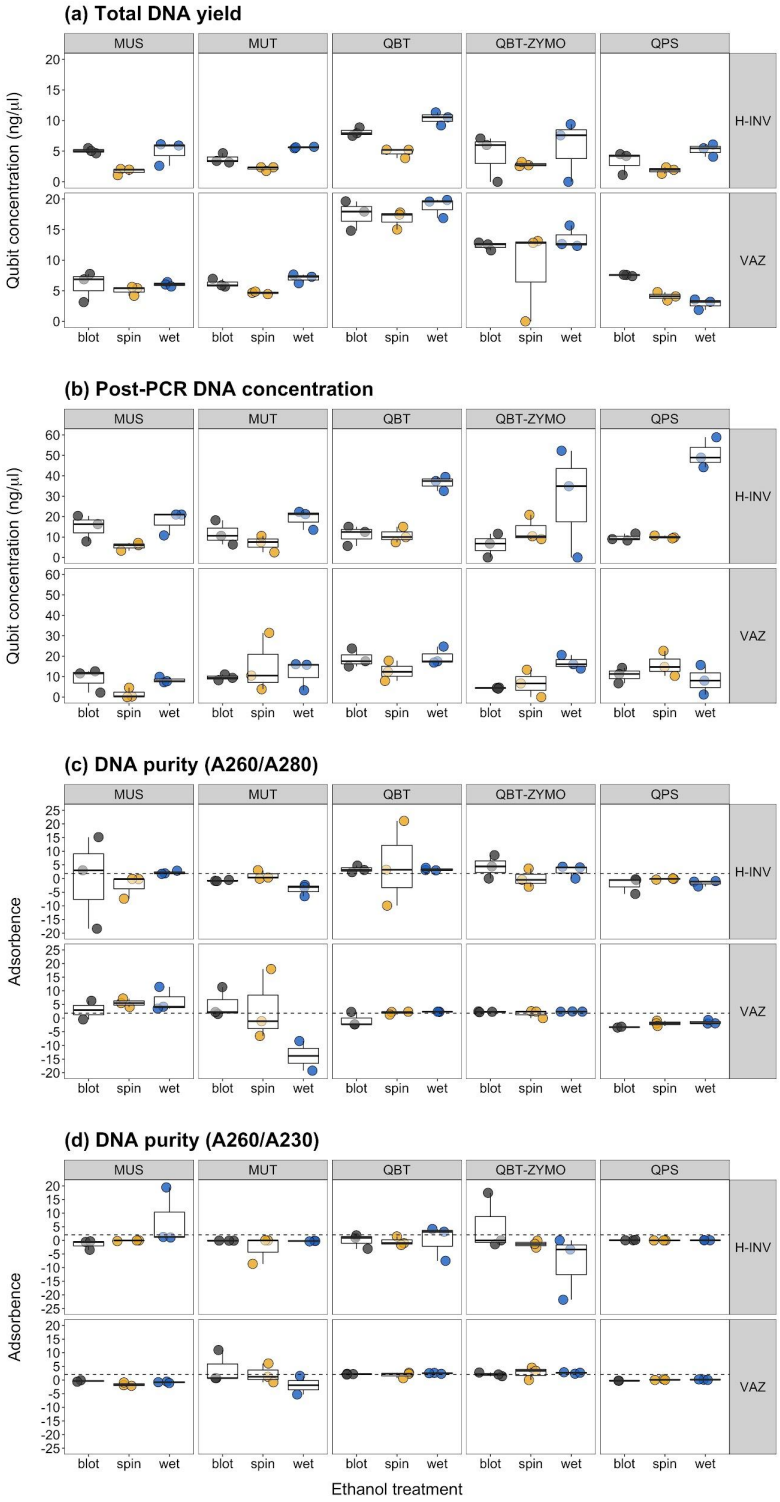

Figure 2. Box plots summarising **(a)** total DNA yield, **(b)** post-PCR DNA concentration, **(c)** A260/280 ratios, and **(d)** A260/A230 ratios for each sample from each sponge species (Horniman invasive sponge (H-INV) and *Vazella pourtalesii* (VAZ)) processed with each ethanol treatment and each DNA extraction protocol.

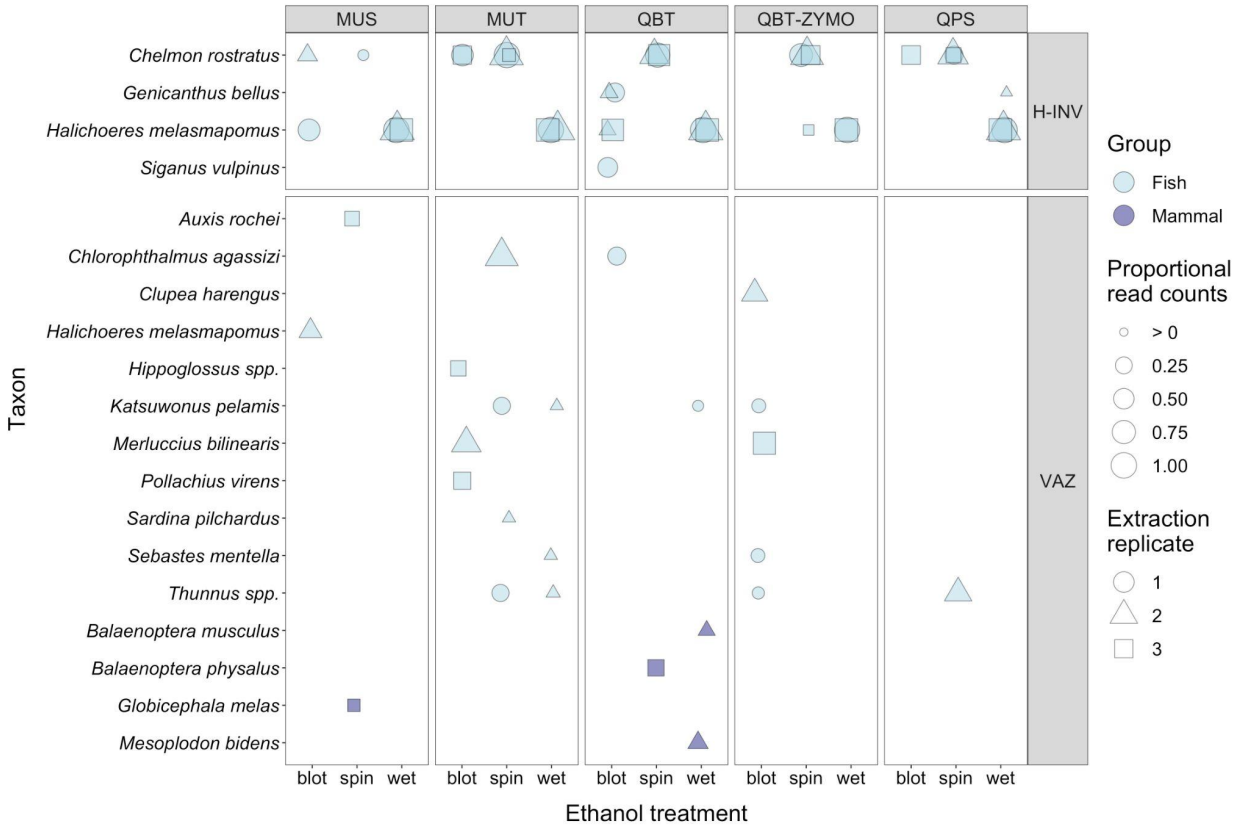

Figure 3. Bubble plot summarising proportional read counts for each taxon and detection consistency across extraction replicates for samples from each sponge species (Horniman invasive sponge (H-INV) and *Vazella pourtalesii* (VAZ)) that were processed with each ethanol treatment (blot, spin, wet) and DNA extraction protocol (MUS, MUT, QBT, QBT-ZYMO and QPS).

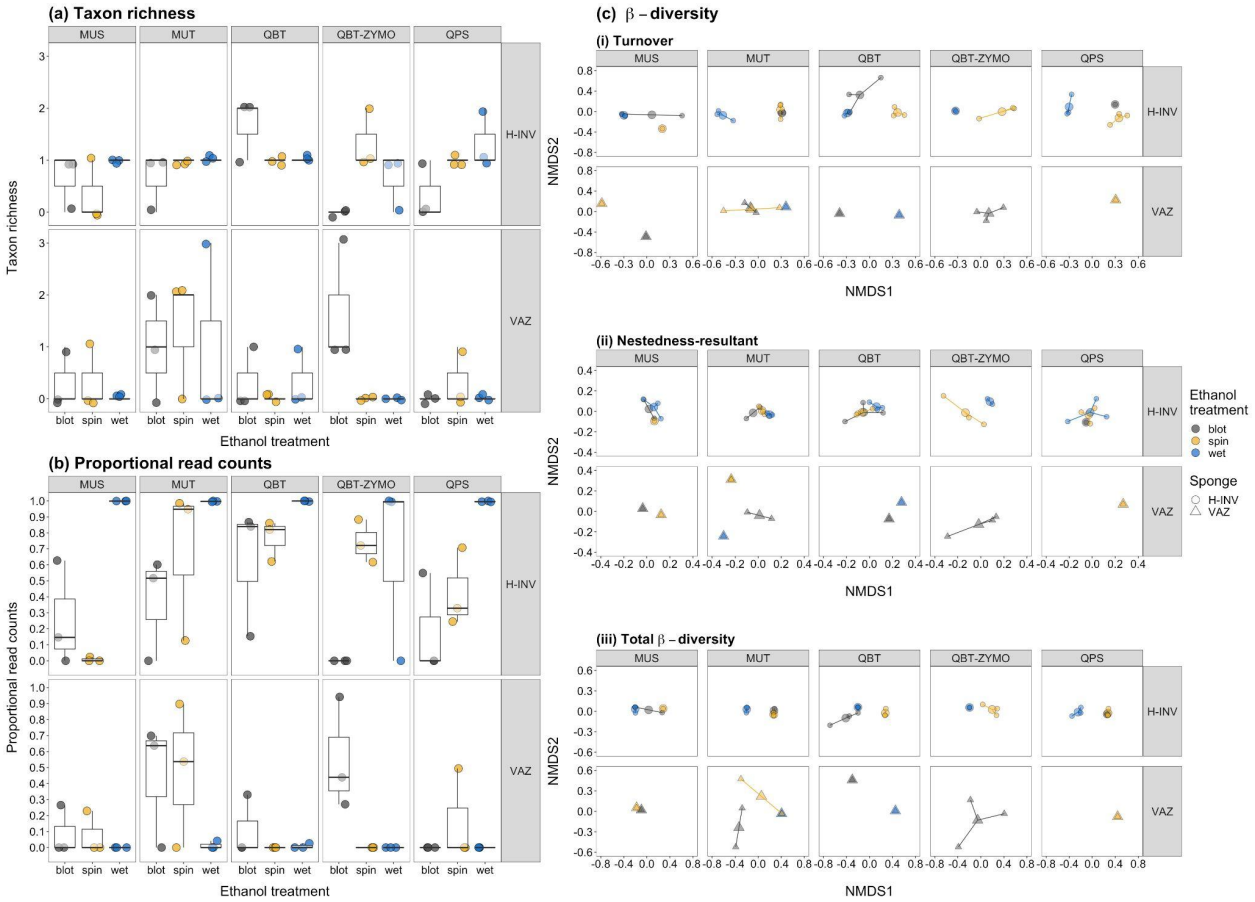

Figure 4. Differences in fish taxon richness **(a)**, proportional read counts **(b)** and community similarity **(c)**, including turnover **(i)** and nestedness-resultant **(ii)** components of total beta diversity **(iii)**, for samples processed with each ethanol treatment (blot, spin, wet) and each DNA extraction protocol (MUT, QBT and QBT-ZYMO). Abbreviations: Horniman invasive sponge (H-INV), *Vazella pourtalesii* (VAZ).

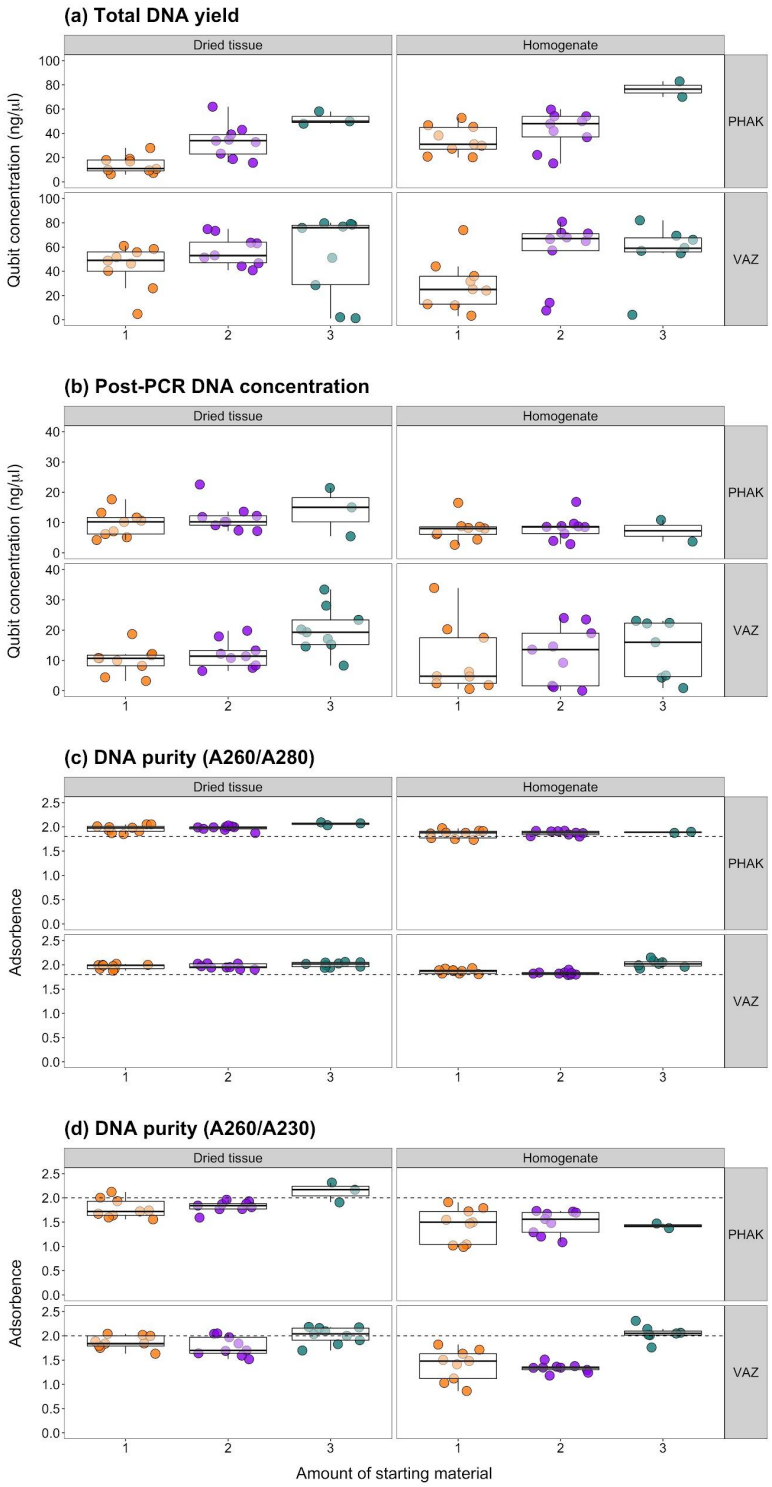

Figure 5. Box plots summarising **(a)** total DNA yield, **(b)** post-PCR DNA concentration, **(c)** A260/280 ratios, and **(d)** A260/A230 ratios for each sample from each sponge species (*Phakellia ventilabrum* (PHAK) *Vazella pourtalesii* (VAZ)) processed using different types and amounts of starting material.

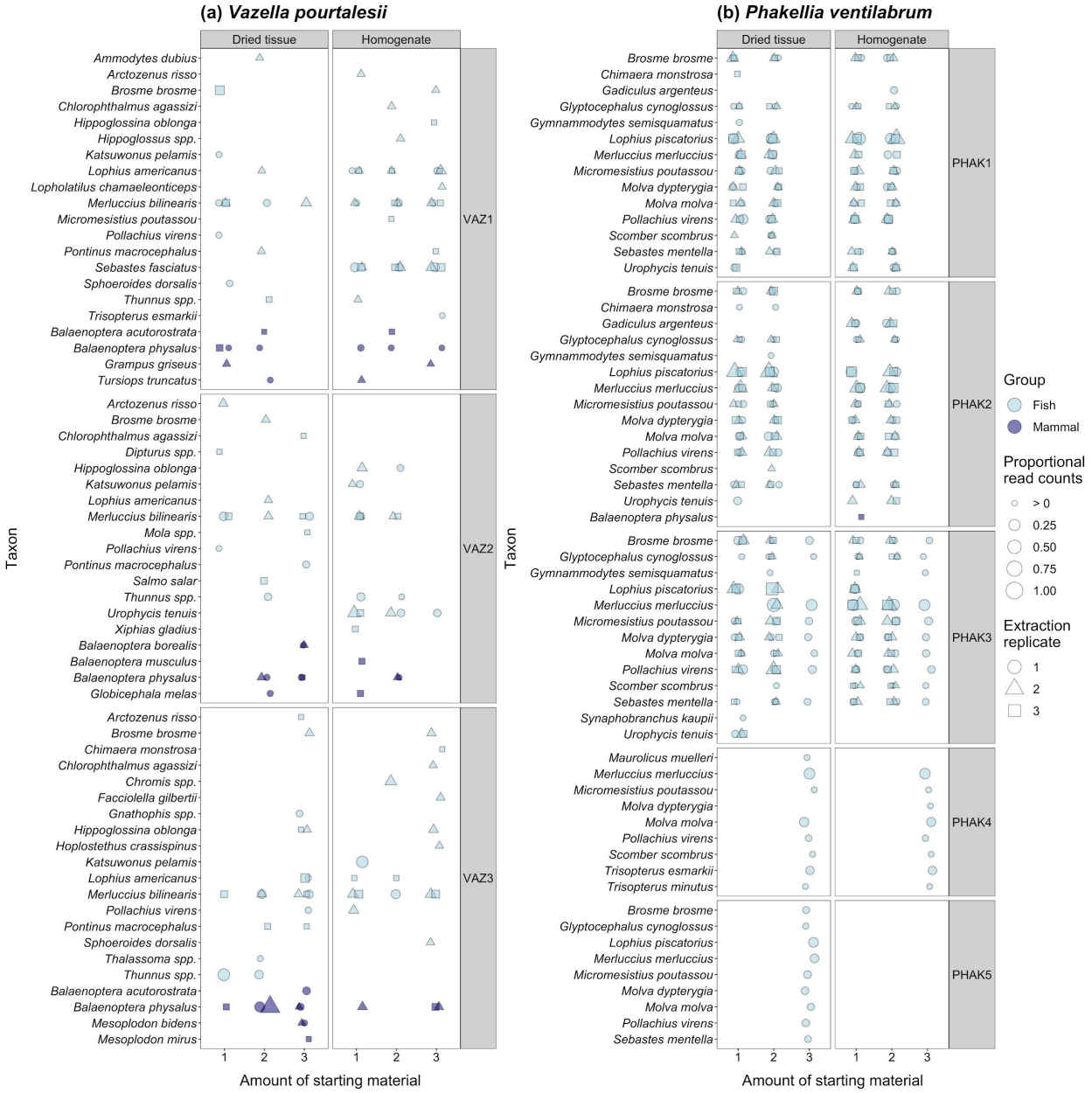

Figure 6. Bubble plots summarising proportional read counts for each taxon and detection consistency across extraction replicates for different types and amounts of starting material from **(a)** *Vazella pourtalesii* (VAZ) and **(b)** *Phakellia ventilabrum* (PHAK).

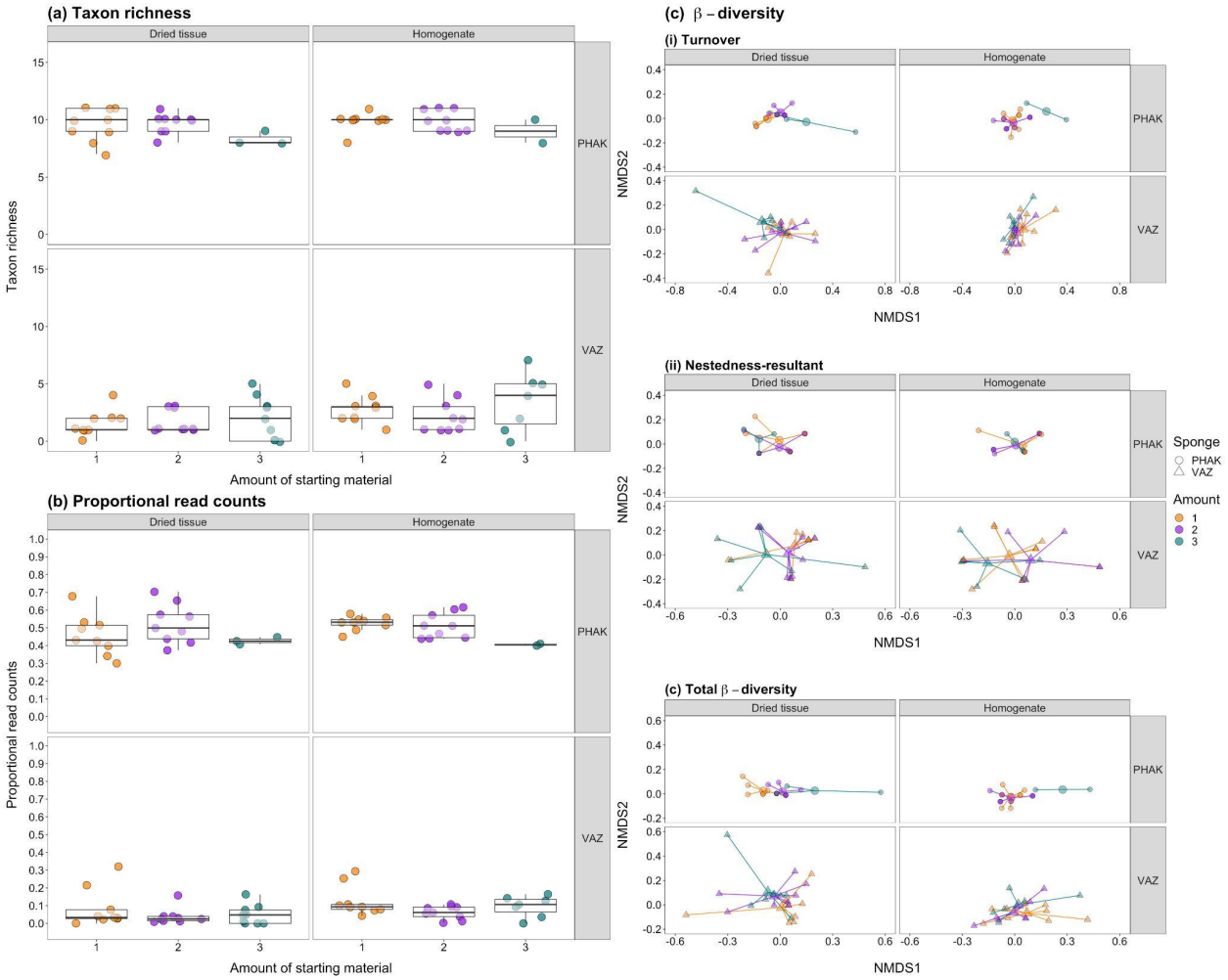

Figure 7. Differences in taxon richness **(a)**, proportional read counts **(b)** and community similarity **(c)**, including turnover **(i)** and nestedness-resultant **(ii)** components of total beta diversity **(iii)**, for fish detected in samples of each type and amount of starting material from each sponge species (*Phakellia ventilabrum* (PHAK), *Vazella pourtalesii* (VAZ)).