bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

1 # Development of a new faecal DNA extraction method (HV-
2 CTAB-PCI) for amplification of mitochondrial and
3 nuclear markers used in genetic analyses of dugongs
4 (*Dugong dugon*)

5 Vicky Ooi[1*], Lee McMichael[2], Margaret E. Hunter[3], Aristide Takoukam Kamla[4,5], Janet M.
6 Lanyon[1]

7

8 [1]School of Biological Sciences, The University of Queensland, St Lucia, Queensland,
9 Australia

10 [2]School of Veterinary Science, The University of Queensland, Gatton, Queensland, Australia

11 [3]U.S. Geological Survey, Sirenia Project, Wetland and Aquatic Research Centre, Gainesville,
12 Florida, USA

13 [4]Aquatic Animal Health Program, College of Veterinary Medicine, University of Florida,
14 Gainesville, Florida, USA

15 [5]African Marine Mammal Conservation Organization, Dizangue, Littoral, Cameroon, Africa

16

17 * Corresponding author

18 E-mail: v.ooi@uqconnect.edu.au

19

20

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

21 # Abstract

22 Non-invasively collected faecal samples are an alternative source of DNA to tissues, that may
23 be used in genetic studies of wildlife when direct sampling of animals is difficult. Although
24 several faecal DNA extraction methods exist, their efficacy varies between species. Previous
25 attempts to amplify mitochondrial DNA (mtDNA) markers from faeces of wild dugongs have
26 met with limited success and nuclear markers (microsatellites) have been unsuccessful. This
27 study aimed to develop a new tool for sampling both mtDNA and nuclear DNA (nDNA)
28 from dugong faeces by modifying approaches used in studies of other large herbivores. First,
29 amplification success of genetic markers from dugong faeces was compared between an
30 established QIAamp and a newly developed DNA extraction method. Faecal DNA extracted
31 using a new 'High Volume-CTAB-PCI' (HV-CTAB-PCI) method was found to achieve
32 comparable amplification results to extraction of dugong skin DNA. As most prevailing
33 practices advocate sampling from the outer surface of a stool to maximise capture of
34 sloughed intestinal cells, this study compared amplification success of mtDNA between the
35 outer and inner layers of faeces, but no difference in amplification was found. Assessment of
36 the impacts of faecal age or degradation on extraction, however, demonstrated that fresher
37 faeces with shorter duration of environmental (seawater) exposure amplified mtDNA and
38 nDNA better than eroded scats. Using the HV-CTAB-PCI method, nDNA was successfully
39 amplified for the first time from dugong faeces. This novel DNA extraction protocol offers a
40 new tool that will facilitate genetic studies of dugongs and other large and cryptic marine
41 herbivores in remote locations.

42

43 # Introduction

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

44 Genetic variation between individuals frequently influences the evolutionary
45 resilience of species [1,2]. Small populations lacking in gene flow have low genetic diversity,
46 which may restrict their ability to adapt to environmental changes, leaving them vulnerable to
47 extinction [3]. As threatened wildlife species dwindle in number, studies addressing genetic
48 variation within and between populations become crucial for continued conservation efforts.
49 To obtain critical genetic information on free-ranging wildlife, most studies rely on direct
50 sampling of body tissues (e.g., blood, skin) as a source of high quality DNA [4]. However,
51 there may be logistical and/or ethical challenges in direct sampling of large, rare, cryptic, or
52 elusive species [5]. In contrast, non-invasive sampling through collection of animal traces,
53 such as faeces, shed hairs, or sloughed skin as a source of DNA, mitigates the need to
54 invasively capture, restrain or in some cases, directly observe the target wildlife [6]. One
55 drawback of sampling biological traces is that quantity and quality of DNA may be
56 compromised [6]. However, if collected and processed appropriately, non-invasive samples
57 may provide comparable genetic information as obtained through direct sampling [e.g., 7].

58 Utilisation of faeces as an alternative source of DNA has its challenges. Apart from a
59 generally low yield of amplifiable target DNA, faeces may contain PCR inhibitors that
60 originate from digestive contents or the environment in which they are voided [8]. DNA from
61 herbivore faeces is often more difficult to amplify compared to that of carnivores, since these
62 may contain plant secondary metabolites that inhibit PCR [9]. The use of faecal DNA for
63 genetic studies of marine species is even more restricted due to degradation and
64 contamination in seawater [10]. Furthermore, the age of faecal samples may impact retrieval
65 of target DNA, i.e., higher yield and quality of target DNA may be extracted from fresh
66 compared to old scats [11].

67 The bulk of mammalian faeces comprises water and soluble molecules such as mucin,
68 and the remainder consists of undigested food and microflora [12]. A small portion of faeces

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

69 comprises epithelial cells sloughed from the host's gastrointestinal tract [13]. Based on the
70 assumption that sloughed cells would adhere to the surface of the egesta as it passes through
71 the gastrointestinal tract, some sampling protocols swab the surface of scats to obtain these
72 cells [14], whilst others resort to either homogenising the faecal sample or scraping, peeling,
73 and/or washing its surface [8]. Surprisingly, the distribution of target cells and thus host DNA
74 within faeces appears to have remained underexplored despite increasing efforts to isolate
75 DNA from scats. Although various extraction methods have been used to retrieve host DNA
76 from faecal samples, there is no 'one-size-fits-all' approach that works optimally for all
77 species. Some widely used extraction techniques include the QIAamp Fast DNA Stool Mini
78 Kit (hereafter referred to as QIAamp) [7,15], phenol-chloroform-isoamyl alcohol (PCI)
79 [11,16] and cetyltrimethyl ammonium bromide (CTAB) [17]. Interestingly, the 2CTAB/PCI
80 protocol has demonstrated success in extracting DNA from herbivore faeces [18].

81 Dugongs (*Dugong dugon*) are vulnerable marine mammals that are declining
82 throughout their range from East Africa to Vanuatu [19]. Dugongs are challenging to study as
83 they often reside in inshore, turbid waters and remote locations where access may be limited
84 [20]. Dugongs spend most of their lives underwater, and their short surface intervals make
85 them difficult to approach and sample directly [21]. Consequently, few genetic data are
86 available for dugongs throughout much of their range. Of all regions supporting dugongs,
87 northern Australia has been best studied [22] due to relative accessibility of some locations
88 by researchers and intermittent recovery of carcasses. Conversely, in many regions outside
89 Australia, dugong densities are low and variable; some populations have been extirpated,
90 some are functionally extinct [23], and most are in decline [19]. These challenges to studying
91 dugongs has rendered a non-invasive DNA sampling approach, i.e., via faeces, increasingly
92 desirable as it may enable population genetic studies in areas where direct sampling is
93 impracticable.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

94 Early genetic studies of dugongs analysed mtDNA from recovered carcasses [21,24].
95 Maternally-inherited mtDNA has high mutation rates, making it an effective marker for
96 distinguishing deep taxonomic relationships and broad population structure [25]. Although
97 mitochondrial studies suggested two dugong maternal lineages in Australia (i.e., north-
98 western, and eastern populations), separated at Torres Strait [21], it was only when nDNA
99 was used that finer scale population structuring was found within southern Queensland [26].
100 A more recent study evaluated population structure of dugongs along the entire eastern
101 Queensland coast, using the same tandemly repeated microsatellites, and found an abrupt
102 genetic break at the Whitsundays region, effectively separating northern and southern clusters
103 [27]. Two less distinct subclusters were found within each of these main clusters, but a lack
104 of tissue samples along the more remote northern coast limited discernment of population
105 structuring along the entire coast. These patterns were also supported by analysis using 464
106 highly discriminatory single nucleotide polymorphisms (SNPs) [27]. A SNP represents
107 variations of a single nucleotide base within the genome of different individuals, where the
108 least abundant allele or nucleotide form occurs in more than 1% of the population [28].
109 Although a greater number of SNPs are required to provide the same resolution as
110 microsatellites due to the biallelic nature of SNPs as opposed to the multiallelic nature of
111 microsatellites, fewer genotyping errors associated with SNPs and their abundance
112 throughout the genome have led to an increase in their use [29]. More importantly, the small
113 target regions for SNP amplifications would allow for higher PCR amplification success in
114 more degraded DNA such as those from faeces [30].

115 In areas where dugong tissues are difficult to obtain, faecal samples have been
116 collected, although their use remains limited as only mtDNA has been successfully amplified,
117 despite efforts to retrieve nDNA [21,27]. Extraction of nDNA is more challenging because
118 target DNA from faeces is often highly degraded, and the amount of available nDNA is less

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

119 than that of mtDNA, since a diploid cell may contain many mitochondria but only possesses
120 two copies of the nuclear genome [31]. However, a recent study has successfully extracted
121 both nDNA and mtDNA from faeces of manatees, phylogenetically related to dugongs, using
122 the QIAamp method with modifications including an additional purification step that
123 removes PCR inhibitors [32]. Manatees are generalist herbivores that consume highly fibrous
124 and abrasive macrophytes [33], whereas dugongs are seagrass specialists that preferentially
125 graze on seagrasses with high nitrogen and low fibre content [34,35]. As dugongs have high
126 apparent digestibility of low fibre seagrass [36], they produce less fibrous faeces compared to
127 manatees [32]. Although these sirenians possess some differences in terms of diet and
128 digestive function [37], it is likely that the extraction methods used for manatees may work in
129 dugongs since their less fibrous diets may reduce the occurrence of PCR inhibitors in their
130 stools.

131 The major aim of this study was to develop and validate a faecal DNA extraction
132 protocol and improve faecal sampling methodologies to enable successful amplification of
133 nuclear and mitochondrial markers used in population genetic studies of dugongs. The main
134 hypothesis was that mtDNA control region and nDNA, represented by a sex marker and ten
135 SNP markers, could be amplified from the outer surface layers of a fresh dugong stool. The
136 first objective was retrieval of sufficient target DNA from dugong faeces, using an optimal
137 DNA extraction method, for consistent amplifications of zinc finger region (ZFX) within the
138 X-chromosome of nDNA and the control region of mtDNA. The second objective
139 investigated whether amplification success of both markers and the amount of target DNA
140 differed between the outer versus inner layers of faeces with different levels of environmental
141 exposure. The third objective sought amplification of ten previously developed SNPs [27],
142 from fresh dugong faeces, using the optimal DNA extraction protocol. A representative SNP

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

143 was also amplified in dugong faeces of different environmental exposure levels (EELs) to
144 explore the impacts of degradation on SNP amplification success.

145

146 Methods

147 Field sampling

148 Two types of dugong faeces were collected: *Ex-ōceanum* faeces were those that were
149 collected when dugongs were held out of water and were thus uncontaminated by seawater,
150 whilst *in-ōceanum* faeces were those that were collected from the ocean and were
151 contaminated by seawater.

152 Matched *ex-ōceanum* faeces (n= 8) and skin tissues (n= 8) were collected from live
153 wild dugongs during the 2018 annual health assessment program in Moreton Bay,
154 Queensland (27.4°S, 153.2°E), Australia [38]. Dorsal skin tissues were sampled using a skin
155 scraper [39], and freshly voided *ex-ōceanum* faeces were collected onto a clean frisbee placed
156 under the anus of each dugong held on deck [38]. Frozen *ex-ōceanum* faecal samples
157 collected in 2016 (n= 8) and 2017 (n= 8) were also used in this study. *Ex-ōceanum* faeces
158 represent the freshest and least degraded of all faecal types and were categorised under
159 Environmental Exposure Level 1 (*ex*-EE-L1). The surfaces of these faeces were textured,
160 ranging from yellow-brown, or light brown to dark brown, and the inner core was usually
161 more dough-like as opposed to fibrous with yellowish pigmentation.

162 *In-ōceanum* faeces of four different EELs or exposure states were collected from the
163 ocean at different time points post-elimination. Fresh *in-ōceanum* faeces with exposure level
164 2 (*in*-EE-L2) (n= 8) were collected from the benthos through free-diving immediately as the
165 feeding herd of dugongs was seen leaving an area (estimated at < 1 h post-elimination).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

166 Floating faeces of exposure level 3 (*in*-EE-L3) (n= 8) were collected from the ocean surface
167 in an area where the feeding herd was spotted nearby (~ 2 – 4 h post-elimination). The *in*-EE-
168 L2 and *in*-EE-L3 faeces have comparable morphologies: cylindrical (large calibre) of
169 variable length. Their inner core usually presents comparable to *ex*-EE-L1; however, fibre
170 proportion can be variable. Their colours are usually dark to light brown. Exposure level 4
171 (*in*-EE-L4) (n= 8) faeces were collected from the ocean surface where individuals were
172 spotted foraging at a distance (~ 5 – 7 h post-elimination). Morphologically, they are usually
173 smaller and/or shorter than *in*-EE-L2 and *in*-EE-L3, though large specimens have been found.
174 The shape of *in*-EE-L4 faeces remains cylindrical, yet jagged edges and lumps emerge in this
175 category. Their colours are darker brown to black on the surface and the inner core are more
176 yellowed with a slightly more fibrous nature. The most eroded faeces, exposure level 5 (*in*-
177 EE-L5) (n= 8), were collected from the ocean surface in an area where no dugongs were
178 spotted (possibly > 7 h post-elimination). The *in*-EE-L5 faeces tend to be in small pieces as
179 they are broken down by ocean currents and/or coprophagous animals since they were
180 exposed to the environment for the longest duration. The eroded faeces commonly appear
181 charred black or dark brown in colour, and the inner core is mostly highly fibrous, though in
182 some cases, they retained less fibre. As they are found in small pieces, the shape can differ
183 from being cylindrical to round-like. (Summary: Table 1, Description: Fig 1).
184
185 **Table 1. Summary on the description and categories of all faecal samples collected and**
186 **used in this study, including the year and place of collection.**

Type of faeces	Environmental exposure level	Abbreviation	Hours post elimination (*approx.*)	Description of faecal collection	Year of collection	Place of collection
Ex-ōceanum	1 (Most fresh)	*Ex*-EE-L1	0	Collected out of water on a boat	2016, 2017, 2018	Moreton Bay, Australia

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

In-ōceanum	2	*In*-EE-L2	~ Less than 1	Collected as feeding herd was seen leaving an area	2022	Fisherman's Gutter, Moreton Bay, Australia
In-ōceanum	3	*In*-EE-L3	~ 2 to 4	Collected where the feeding herd was spotted nearby	2016	Clairview, Australia
In-ōceanum	4	*In*-EE-L4	~ 5 to 7	Collected where feeding herd or individual dugongs were spotted at a distance	2015	Cleveland Bay, Australia
In-ōceanum	5 (Eroded)	*In*-EE-L5	~ More than 7	Collected in an area where no dugongs were spotted	2016	Newry Island, Australia

187

188 **Fig 1. Classification of dugong faeces into five categories as illustrated with an above**
189 **view (top) and transverse cross-section (bottom).**
190 Comparing freshly excreted samples on deck to those most aged in the ocean, a respective
191 transition of surface colour was observed from lighter to darker, yellow-brown, or light
192 brown to black, while sizes trended from large to small, and core composition generally
193 retaining higher fibre proportion over time. (A) *Ex-ōceanum* faeces (*ex*-EE-L1), originating
194 beneath the dugong, were flattened in a shallow frisbee. Thus, the above view displays a
195 wider shape (top) and a flatter profile from the side (bottom). (B) *In-ōceanum* Level 2 and 3
196 faeces. (C) *In-ōceanum* Level 4 faeces. (D) *In-ōceanum* Level 5 faeces.

197

198 Faeces were placed in a ziplock bag upon retrieval, stored in an esky in the field, and
199 later frozen at -20 °C. Skin tissues and a single liver tissue sample from a neonate dugong,
200 were used in this study as positive controls; both mtDNA and nDNA have been previously
201 extracted from these tissue types [26].

202

203 Experimental Design

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

204 To develop an appropriate protocol for DNA extraction from dugong faeces, two
205 faecal sampling and two processing techniques were compared, and the approach that
206 recovered the highest amount of total (dugong and exogenous) DNA and target (dugong)
207 mtDNA was used in all further extractions. Two DNA extraction methods, i.e., the QIAamp
208 method and a newly developed 'High volume-CTAB-PCI' (HV-CTAB-PCI) method, were
209 tested. Total DNA concentration extracted using each protocol, and the amplification success
210 (including each of PCR and triplicate success) of mtDNA control region and ZFX (sex
211 marker) were compared (S1 Fig A). PCR success represents the percentage of samples
212 amplified, while triplicate success represents the number of replicates amplified out of three
213 technical replicates for each sample.

214 The QIAamp method was used to extract DNA from the outer surface and inner core
215 of dugong faeces of four different EELs (all except *in*-EE-L2 faeces). The amount of DNA
216 used in qPCR reactions was normalised to enable comparisons between faecal samples. Total
217 DNA concentration, relative amount of dugong mtDNA extracted, and amplification success
218 were compared under a two-factorial design (S1 Fig B). The HV-CTAB-PCI method was
219 used to extract DNA from dugong faeces of five different EELs, and both mtDNA control
220 region and ZFX (nDNA) were amplified. Total extracted DNA concentration, and
221 amplification success of each marker were compared between faeces of different EELs (S1
222 Fig B).

223 DNA extracted using the HV-CTAB-PCI method from *in*-EE-L2 faeces was used to
224 trial the amplification of ten SNP primer sets, and amplification success was compared. One
225 SNP marker was selected to be amplified using DNA extracted with the HV-CTAB-PCI
226 method from all *in-ōceanum* faeces, and amplification successes were compared (S1 Fig C).

227

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

228 ## Faecal sampling and processing

229 Using a sterile surgical blade, 220 mg of faecal material was scraped from the outer
230 surface of 16 *ex-ōceanum* faecal samples, each weighed and stored in a 2 mL microcentrifuge
231 tube (Note: For inner core sampling, the stool was broken in half and faecal material from the
232 inner core of faeces was sampled). Each faecal sample was then either transferred into a
233 sterilised mortar and ground into a fine powder with a pestle using liquid nitrogen (N_2)
234 (referred to as 'scrape-and-grind' technique) (n= 8) or left unground (referred to as 'scrape-
235 only' technique) (n= 8) (Fig 2).

236 A further eight *ex-ōceanum* faeces were sampled through swabbing the entire outer
237 surface using a rayon swab (Copan, CA, USA) (referred to as 'swabbing' technique). Once
238 the swab tip was entirely covered in faecal material, it was transferred into a 2 mL tube, and
239 the swab's shaft was trimmed to fit. No further processing was conducted (Fig 2).

240

241 **Fig 2. Faecal sampling and processing techniques used in this study.**

242 Faecal material was either scraped from the outer surface of faeces or swabbed along the
243 entire faeces for DNA extraction. When faecal material was obtained via scraping, it was
244 either further processed by grinding in liquid nitrogen prior to the addition of lysis buffer
245 (Buffer ATL) or added directly with lysis buffer without any processing done. When faecal
246 material was sampled using the swabbing technique, no further processing was performed.

247

248 ## DNA extraction from faeces

249 ### QIAamp protocol

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

250 DNA was extracted using QIAamp Fast DNA Stool Mini Kit (#51604, Qiagen,
251 Germany). When swabbing or scrape-only techniques were used, 1 mL of InhibitEX
252 buffer was added into the 2 mL tube containing the faecal sample [32]. When scrape-
253 and-grind technique was used, 500 μL of InhibitEX buffer was added to the mortar with
254 ground faeces for further grinding prior to the transfer of faecal mixture back into the 2
255 mL tube, and another 500 μL was used to rinse the remaining faeces from the mortar into
256 the same tube. The resultant mixture was vortexed for 1 min (Note: swab was removed
257 after vortexing) and centrifuged for 1 min at 20,000 *g* (14,000 rpm) to pellet the stool
258 particles. The remainder of the extraction was performed following manufacturer's
259 instructions, but by using 800 μL of supernatant from the lysate, and subsequently equal
260 amount of Buffer AL and 95 % ethanol, with a final 100 μL elution volume.

261 **HV-CTAB-PCI protocol**

262 An extraction protocol was developed using the lysis buffers and general concept of the
263 2CTAB/PCI method [18]. Approximately 1 g of faecal material was processed using the
264 scrape-and-grind technique, and 1 mL of Lysis Buffer 1 (LB1: CTAB 2 %, Tris– HCL
265 100 mM, EDTA 20 mM, NaCl 1.4 M, pH 7.5) was added to the mortar to further grind
266 the powdered faeces before the mixture was transferred into a 15 mL tube. This was
267 repeated twice to ensure that any remaining faecal material in the mortar was collected.
268 Another 2 mL of LB1 was added into the same tube, making a total 5 mL LB1 added to
269 the faecal sample. The mixture was vortexed and incubated at 60 °C for 3 h, with
270 occasional mixing for cell lysis. After centrifuging at 3,150 *g* (4,000 rpm) for 12 min, 4
271 mL of supernatant was transferred into a new 15 mL tube and equal volume of phenol:
272 chloroform: isoamyl alcohol (PCI, 21:20:1) was added to the supernatant, and gently
273 mixed. The mixture was centrifuged as above, and 3 mL of the aqueous phase was
274 transferred into a new 15 mL tube. Next, 330 μL of Lysis Buffer 2 (LB2: CTAB 10 %,

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

275 NaCl 0.5 M, pH 5.5) was added to the supernatant, and incubated at 60 °C for 4 h, with
276 occasional mixing, for further lysis. Thereafter, 104 uL of protease (#P5147, Sigma
277 Aldritch, USA) was added to the lysate to digest proteins for 1 h at 60 °C. Then, equal
278 volume (3434 µL) of PCI was added to the mixture, gently mixed, and centrifuged as
279 above. Three mL of the aqueous phase was transferred into a new 15 mL tube and equal
280 volume of isopropanol was added for overnight DNA precipitation at -20 °C. The sample
281 was centrifuged for 20 min at 8000 *g* (5,200 rpm), and all supernatant was decanted. The
282 pellet was washed once with 400 µL of 70 % ethanol. After being vortexed and
283 centrifuged at 3,150 *g* for 12 min, the supernatant was decanted. The pellet was air dried
284 at room temperature for 15 min and resuspended in 250 µL of TE buffer (10 mM Tris–
285 HCl, 1 mM EDTA, pH 8).

286

287 ## DNA extraction from skin

288 DNA was extracted using DNeasy Blood and Tissue kit (#69504, Qiagen, Germany).
289 Approximately 10 mg of dugong skin tissue was ground to fine powder using liquid N_2, and
290 180 µL of Buffer ATL was added into the mortar for further grinding. The mixture was
291 transferred into a 1.5 mL tube, and the addition of another 50 µL of Buffer ATL to the mortar
292 aided recovery of any remaining tissue. The rest of the extraction followed manufacturer's
293 protocol, using a 3 h incubation period, with 100 µL elution in a 1.5 mL tube. A centrifuging
294 step was added following the addition of Buffer AL to remove precipitate, and ethanol was
295 added to the supernatant. Using 2 µL of DNA extract, concentration (ng/µL) and purity
296 (A260/280 and A260/230) of all extracted DNA were measured using NanoDrop™ 1000
297 Spectrophotometer (Thermo Fisher Scientific, USA). DNA isolates were stored at -80 °C
298 until use.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

299

300 # Real-time PCR assays

301 Quantitative polymerase chain reaction (qPCR) was conducted in a CFX96 Touch Real-Time
302 PCR Detection System (Bio-Rad, USA), using 96-well PCR plates, and a 'no template
303 control' was used to detect contamination. DNA extracted from tissue was used as positive
304 control. Melt curve analysis was performed to determine melt temperature of the primers and
305 to detect presence of non-specific amplifications.

306 ## mtDNA and ZFX markers

307 The control region of mtDNA and the ZFX region of nDNA were amplified using
308 specific primers developed by Tol *et al* [40] and McHale *et al* [41], respectively (Table
309 2). These markers were used throughout the entire study except for SNP amplifications.

310 **Table 2. Forward and reverse primer sequences for mtDNA control region and ZFX**
311 **markers used in this study, with their corresponding amplicon sizes and annealing**
312 **temperatures.**

313

Locus	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Amplicon size (bp)	Annealing temp (°C)	Reference
5' end control region	CGCGCGCTATGTACTTCGT	GGGGTAAGTAGTGTAATGCACG	110	65	[40]
ZFX	AAGCATAGTAAAGAGATGCCGT	ATGTCACACTTGTGGGGGTA	230	58	[41]

314

315 PCRs were performed with a 20 µL reaction volume, each consisting of a variable
316 amount of total DNA (Table 3), variable concentration of forward and reverse primers
317 (Table 3), nuclease-free water, and 10 µL of PowerUp™ SYBR™ Green Master Mix
318 (Applied Biosystems, USA).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

319 **Table 3. Summary of the total DNA (ng) and concentrations of forward and reverse**
320 **primers (µM) added to the PCR reaction for the corresponding experiments and DNA**
321 **extraction methods.**

DNA extraction method	Usage	Primer	Total DNA added (ng)	Concentration of forward and reverse primers used (µM)
QIAamp	Compare faecal sampling techniques	mtDNA	50	0.1
QIAamp	Compare DNA extraction method Compare faecal layer sampling on faeces of different quality	mtDNA	100	0.1
QIAamp	Compare DNA extraction method	ZFX	100	0.5
HV-CTAB-PCI	Compare DNA extraction method, Compare faecal quality	mtDNA	Neat DNA extracted/2	0.1
HV-CTAB-PCI	Compare DNA extraction method, Compare faecal quality	ZFX	Neat DNA extracted/2	0.3
HV-CTAB-PCI	SNP amplification	SNP	Neat DNA extracted/2	0.5

322

323 When DNA extracted with QIAamp method was used, the amount of total DNA added to
324 the reactions was standardised and served as a normaliser to enable comparisons between
325 samples. For the comparison of sampling from different layers of faeces of different
326 EELs, a pilot experiment was conducted to determine whether 10 or 100 ng of total DNA
327 should be used in the reactions (S2 Fig). One-hundred ng of total DNA was chosen
328 because eroded faeces appeared to have higher PCR success using this amount, although
329 the difference was not statistically significant. When DNA extracted with HV-CTAB-
330 PCI method was used, a 1:2 dilution factor provided the most consistent amplification of
331 all markers in the dilution experiments and was thus used for all subsequent PCR
332 reactions.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

333 Cycling conditions were 50 °C for 2 min, then 95 °C for 2 min, followed by 45 cycles of
334 95 °C for 15 s, respective annealing temperature (Table 2) for 1 min, and ending with a
335 melt profile from 65 °C to 95 °C, with 0.1 °C increments.

336 SNP markers

337 The 10 SNP primers used were developed by McGowan [27] (Table 4). PCRs were
338 performed as described previously. Total DNA diluted 1:2 was added to each reaction,
339 and 0.5 µM of forward and reverse primers were used (Table 3).

340 **Table 4. Forward and reverse sequences of SNP primers used in this study, with their**
341 **corresponding amplicon sizes and annealing temperatures.**

342

SNP Locus	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Amplicon size (bp)	Annealing temp (°C)	Reference
Dug12	TCAGAAACAGAGCAACCAAAAA	CTTTGGAGTCTGAGCTCTGGA	85	60	
Dug40	TGATGAGCTAAACTAAAACAGAAACC	TGTTTCATGAGGTTAAGGGGA	92	60	
Dug50	GGGCTCCAGCTCTCACTGTA	GGTGATTATCAAACTTCTTTCCAGA	108	60	
Dug51	ACCATGACCAAGTGGGATTC	AAAACAAAATACAAGAACCACATGA	76	60	
Dug53	GATCTGTTGCGTTGCTCTTG	TCTATTTCAGGGCTGGCATG	65	60	[27]
Dug54	GCAGGCTGCACAATTAACTTTTA	GAGCACCCGTAAAGGGAAAC	98	60	
Dug 60	TGCCCAGATTTCTAGCTTGG	TCTCAGTTAGTGGCAACTCCA	71	60	
Dug62	AAGCATCAAAGGAAGATGCAA	GGTAGCATATGATCAGGAACCG	58	60	
Dug63	TTCAAAGTTGACTGATAGTGAGAAA	GCTGTGCTAATTCTGGGGTC	73	60	
Dug64	TGTAGGAAGTCGGGAGTTGG	TTTATTCCCCACTGCGATTC	70	60	

343

344 Cycling conditions were 50 °C for 2 min, 95 °C for 2 min, followed by 45 cycles of 95
345 °C for 15 s, 60 °C for 30 s, 72 °C for 30 s, and 72 °C for 15 min. A melt profile from 65
346 °C to 95 °C, with 0.5 °C increments was performed.

347

348 qPCR intra-assay and inter-assay variability

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

349 PCR reactions were performed in triplicate to account for intra-assay variability, and
350 the results were averaged for final analysis. To account for inter-assay variability, threshold
351 cycle (C_T) was manually set to the same value using the CFX Maestro software (Bio-Rad,
352 USA) when C_T values were compared, and tissue control was used as an inter-plate calibrator
353 in each plate when relative quantity was calculated and compared.

354

355 Primer efficiency and linearity

356 A standard curve was created for both mtDNA and ZFX primers through dilution
357 experiments, starting from 1000 ng of total DNA, with seven 10-fold dilutions for two skin
358 DNA isolates. The linear dynamic range (LDR) and limits of detection (LOD) for each
359 primer sets were determined from dilution curves, and PCR efficiency (*E*) was calculated
360 from the slope of dilution curve as per Ruijter *et al* [42].

361 Five 10-fold serial dilutions and four 10-fold serial dilutions of two faecal DNA
362 extracts, starting from 5 × neat DNA extracted were performed for mtDNA and ZFX primers,
363 respectively. This was used to determine the most appropriate quantity of DNA extracted via
364 the HV-CTAB-PCI method to be added to the PCR reactions.

365

366 Primer specificity

367 Unpurified PCR products were submitted to the Australian Genome Research Facility
368 (AGRF) at The University of Queensland, Australia, for dual-direction Sanger sequencing.
369 Sequences from mtDNA and ZFX amplifications underwent BLAST searches on GenBank,
370 and sequence similarity was used to confirm specific amplification of respective dugong

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

371 DNA. The resultant SNP sequences were compared to amplicon sequences of McCarthy [43]
372 and McGowan [27] to confirm amplification of interested regions, and to determine the SNP
373 allele possessed by the individuals. Specificity of PCR amplicons was additionally verified
374 through a melt curve analysis for each PCR reaction. SNP sequences were aligned and
375 trimmed using MEGA11 software [44].

376

377 Data processing

378 As C_T value is inverse to the amount of target DNA present within a sample, the
379 inverse of C_T value was used for comparative analyses to facilitate visualisation of results.
380 Relative quantity of target mtDNA was calculated in the CFX-Maestro software (Bio-Rad,
381 USA), using a tissue sample as control and *E*. The relative quantity was multiplied by
382 100,000 for easier interpretation of small numbers. For samples that failed to amplify, a C_T
383 value of the number of PCR cycles used + 1 was assigned to enable the determination of
384 lowest possible C_T value difference for comparison.

385

386 Statistical analyses

387 All statistical analyses were performed using *R* (version 4.2.0, [45]). For all
388 parametric models fitted, diagnostic plots were made to check whether the assumptions for
389 homogeneity and normality of residual variance were met. Normality was further confirmed
390 with the Shapiro-Wilk's test. Data were log-transformed when residuals were not normal.
391 When significant difference was found in the initial model, a post-hoc pairwise Tukey test
392 was performed using *lsmeans* package for further pairwise analyses. False discovery rate
393 adjusted p-values were obtained when non-parametric tests were used. Graphs were created

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

394 using *ggplot2* [46], *dplyr*, *hrbrthemes* [47], *Rmisc* [48], *ggpmisc* [49], *reshape*, *ggsignif* [50],
395 and *ggpubr* [51] packages.

396 A linear mixed effects model (LMEM) from *nlme* package [52] was used to compare
397 total DNA concentration and inverse C_T between different faecal sampling techniques, where
398 random effects of individuals within PCR plates were accounted for. It was also used to
399 compare log total DNA concentration recovered between different DNA extraction methods,
400 accounting for the random effects of individuals between PCR plates. For all primers, PCR
401 success (binary outcome: successful—if one or more amplified versus unsuccessful—none
402 amplified) between DNA extraction methods was compared using Pearson's Chi-square test
403 (χ^2), with a Monte Carlo simulation utilising 10,000 replicates, followed by post-hoc pairwise
404 Chi-square tests. The number of amplifications within a triplicate (i.e., Triplicate success: 0,
405 1, 2, or 3) for all primers, was compared between DNA extraction methods using Kruskal-
406 Wallis test and post-hoc pairwise Wilcoxon rank sum (W) test.

407 To determine the effects of environmental exposure and layers of faeces sampled on
408 the amplification of dugong DNA, a full interactive generalised linear model (GLM) with all
409 variables (sampling layer, faecal EELs, year of collection, individuals) was fitted to compare
410 the total DNA concentration and relative quantity of dugong mtDNA. The Akaike
411 Information Criterion (AIC) or stepAIC function from *MASS* package [53] was used to
412 determine significant variables through a dual-direction stepwise model selection. A final
413 additive LMEM was fitted to compare total DNA concentration and log relative quantity of
414 dugong mtDNA between faecal sampling layers and faecal EELs, accounting for random
415 effects of different individuals. The relationship between PCR success and log total DNA was
416 tested using a quadratic GLM with binomial responses.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

417 A likelihood ratio Chi-square test (LR-Chi-square test) was used to compare between
418 interactive (sampling layer*faecal EELs) and additive (sampling layer + faecal EELs) GLMs,
419 and a final additive GLM with biased reduction was fitted, using *brglm2* package [54], to
420 compare PCR success between the two explanatory variables. A LR-Chi-square test was also
421 used to compare between four multinomial log-linear models (from *nnet* package; [53]) fitted
422 with interactive, additive, and single explanatory effects on the triplicate success of samples.
423 A post-hoc pairwise W test was then used to compare triplicate success between sampling
424 layer and faecal EELs.

425 When HV-CTAB-PCI method was used, total DNA concentration was compared
426 between faecal EELs using a one-way analysis of variance (ANOVA), whilst PCR and
427 triplicate successes were compared between different faecal EELs using the χ^2 and Kruskal-
428 Wallis tests as described earlier. Similarly, two non-parametric tests were also used to
429 compare the PCR and triplicate successes between different primers and faecal types in SNP
430 amplifications.

431

432 Ethics Statement

433 Dugong samples were obtained under The University of Queensland Animal Ethics Permit
434 SBS/181/18, Scientific Purposes Permit WISP14654414, Moreton Bay Marine Parks Permit
435 MPP18-001119, and Great Barrier Reef Marine Park Permit G14/36987.1.

436

437 Results

438 PCR efficiencies, LDR and LOD

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

439 PCR efficiency of tissue DNA was 96.3 % and 86.2 % for mtDNA and ZFX primers,
440 respectively (Fig 3A). All amplifications were within LDR (1-1000ng of total DNA) and
441 LOD (0.001 ng of total DNA) for both primers. The *E* of faecal DNA indicated slight
442 inhibition (mtDNA= 106.3 %, ZFX= 104.7 %), which was addressed through dilutions (Fig
443 3B).

444

445 **Fig 3. PCR dilution curves showing the efficiency (*E*) calculated in percentage for each**
446 **primer sets.**

447 (A) The dilution curve of tissue DNA isolates extracted using the Qiagen DNeasy method.
448 (B) The dilution curve of faecal DNA isolates extracted using the HV-CTAB-PCI method.
449 The *E* showed slight inhibition for both primers, using faecal DNA isolates. The initial total
450 DNA amount (ng) used for PCR reactions was determined from these dilutions.

451

452 DNA extraction method comparison

453 The scrape-and-grind technique yielded significantly higher concentrations of total
454 DNA (70.7 ± 5.3 ng/µL; mean ± standard error) compared to the scrape-only (34.9 ± 4.8
455 ng/µL, Tukey-HSD: $T_{14} = 8.361$, $p < 0.001$) and swabbing (20.4 ± 2.8 ng/µL, Tukey-HSD:
456 $T_{14} = 11.762$, $p < 0.001$) techniques (Fig 4A). Total DNA concentrations, yielded using the
457 scrape-only technique, were significantly higher than that of swabbing (Tukey-HSD: $T_{14} =$
458 3.401, $p = 0.011$). The scrape-and-grind technique also had higher inverse C_T values (0.040 ±
459 0.001) compared to scrape-only (0.033 ± 0.001, Tukey-HSD: $T_{14} = 6.027$, $p < 0.001$) and
460 swabbing (0.035 ± 0.001, Tukey-HSD: $T_{14} = 4.413$, $p = 0.002$) techniques (Fig 4B).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

461 However, the inverse C_T values were not significantly different between scrape only and
462 swabbing techniques (Tukey-HSD: $T_{14} = 1.613$, $p = 0.273$).

463

464 **Fig 4. Comparison between faecal sampling and processing techniques to determine the**
465 **best sampling protocol for DNA extraction.**

466 Violin plot shows the distribution of the data, whilst the box plot shows the first (Q1), second
467 (Q2) and third (Q3) quartiles, minimum, maximum values, and outliers. All DNA was
468 extracted using the QIAamp protocol. (A) Comparison of total DNA concentration extracted
469 using each sampling and processing technique. (B) Comparison of the inverse C_T values from
470 mtDNA amplification using faecal DNA extracted with each sampling and processing
471 technique. Similar letters denote no statistical significance, while different letters denote
472 statistical significance with an alpha value of 0.05.

473

474 Total DNA concentration extracted from faeces using HV-CTAB-PCI method (691.9
475 ± 51.9 ng/µL) was significantly higher than that extracted using QIAamp method (158.5 ±
476 15.8 ng/µL, Tukey-HSD: $T_{14} = 11.701$, $p < 0.001$, Fig 5A). The former method extracted
477 more total DNA than control (145.4 ± 12.6 ng/µL, Tukey-HSD: $T_{14} = 12.196$, $p < 0.001$), but
478 total DNA extracted using QIAamp method was not significantly different from control
479 (Tukey-HSD: $T_{14} = 0.496$, $p = 0.875$). On average, the purity of faecal DNA extract was
480 higher using QIAamp rather than HV-CTAB-PCI method (Table 5).

481

482 **Fig 5. Comparisons between two DNA extraction methods and a tissue control.**

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

483 (A) A violin plot with box plot showing the total DNA concentration of each DNA extraction
484 method. (B) A stacked bar graph showing the percent amplification that was successful (PCR
485 success) and unsuccessful between all DNA extraction methods for each primer. (C) A bar
486 graph showing the mean ± standard error of the number of replicates that amplified in
487 triplicate for each DNA extraction method. Similar letters (with same colour and letter case)
488 denote no statistical significance, while different letters denote statistical significance (α=
489 0.05). Different statistical comparisons made were indicated by different colour and letter
490 case.

491 **Table 5. Purity (A260/280 and A260/230) values of DNA extracted for each sample type**
492 **using the Qiagen kits and a newly developed HV-CTAB-PCI method.**

DNA Extraction Method	Sample Type	A260/280 (mean ± std error)	A260/230 (mean ± std error)
Qiagen DNeasy	Tissue	2.12 ± 0.01	1.82 ± 0.04
QIAamp	Faeces	1.95 ± 0.01	1.27 ± 0.03
HV-CTAB-PCI	Tissue	2.06 ± 0.01	2.14 ± 0.01
HV-CTAB-PCI	Faeces	1.67 ± 0.01	1.06 ± 0.03

493

494 PCR success and triplicate success for mtDNA amplification were similar between all
495 DNA extraction methods, as all replicates of all samples amplified successfully (Fig 5B, C).
496 For ZFX amplification, PCR success of HV-CTAB-PCI method was the same as control, and
497 PCR success of those two were significantly higher than that of the QIAamp method (χ^2=
498 7.273, p = 0.026). All samples amplified in control and HV-CTAB-PCI method, but only
499 37.5% of faecal samples amplified in the QIAamp method (Fig 5B). The control had
500 significantly higher triplicate success for ZFX amplification (Mean, $\bar{x}$ amplification= 3/3)
501 compared to HV-CTAB-PCI ($\bar{x}$ amplification= 2.1/3, *W* test: Z= 1.86, p = 0.032) and
502 QIAamp ($\bar{x}$ amplification= 0.8/3, *W* test: Z= 2.69, p = 0.004) methods (Fig 5C). However,

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

503 triplicate success of HV-CTAB-PCI method was higher than that of QIAamp method (*W*-test:
504 Z= 1.86, p = 0.032).

505

506 Effects of EELs and faecal layers sampled

507 Effects of sampling from different faecal layers did not depend on faecal EELs as all
508 interactive models had higher AIC values compared to additive models (Table 6). Same
509 results were obtained for LR-Chi-Square test (Deviance= 0.022 – 0.832, p > 0.842).

510 **Table 6. The AIC values obtained for interactive and additive models fitted in *R* for**
511 **different variables compared using a dual-direction stepwise model selection.**

Response variable	Type of model	AIC values
Total DNA concentration	Interactive	349.98
	Additive	346.74
Log relative amount of mtDNA	Interactive	930.97
	Additive	929.56

512

513 *Impacts of faecal layer sampled*

514 For faeces of all EELs, total DNA concentration extracted from the outer surface
515 of faeces was significantly higher than that extracted from their inner core (Tukey-HSD:
516 T= 5.601, p < 0.001 for all faeces; Fig 6A). The outer surface of *ex*-EE-L1 faeces had at
517 least 2.87× more mtDNA than their inner core, while the outer surface of *in*-EE-L3, L4,
518 and L5 faeces had at least 0.14×, 0.65×, and 0.95× more mtDNA than their inner core,
519 respectively (Fig 6B). On average, the outer surface of all stools had at least 0.16× more
520 dugong mtDNA than the inner core of faeces. However, none of the differences were

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

521 significant (Tukey-HSD: T= 1.222, p = 0.919 for all faeces). PCR success (Biased-
522 Reduction-GLM: Z= 0.373, p = 0.709; Fig 6C) and triplicate success (Likelihood-ratio=
523 0.832, p = 0.842; Fig 6D) of mtDNA amplification were the same regardless of whether
524 outer surface or inner core of faeces were sampled. As the layer of faeces sampled had no
525 significant effect on the amplification success, only the outer surface of faeces was used
526 for further comparisons.

527

528 **Fig 6. Comparison of the variables measured between outer surface and inner core of**
529 **faeces across different environmental exposure levels (EELs) representing varying**
530 **quality.**

531 (A) Total DNA concentration extracted from different layers of various faeces. No statistical
532 significance was present between the faecal types, but the outer surface yielded a higher
533 concentration of total DNA compared to the inner core for all faeces. (B) Log relative
534 quantity of dugong mtDNA amplified from different layers of various faeces. (C) The percent
535 mtDNA amplification that was successful and unsuccessful between different layers of
536 various faeces. PCR success was statistically similar between all faecal types. (D) Triplicate
537 success between different layers of various faeces. Similar letters denote non-significance,
538 while different letters denote statistical significance (α= 0.05).

539 ***Impacts of environmental exposure***

540 When the QIAamp method was used, the total DNA concentration extracted from
541 *ex*-EE-L1 (87.0 ± 12.8 ng/µL), *in*-EE-L3 (55.2 ± 5.1 ng/µL), *in*-EE-L4 (86.5 ± 10.0
542 ng/µL), and *in*-EE-L5 (59.2 ± 6.06 ng/µL) faeces were not significantly different from
543 each other regardless of the faecal layer sampled (Tukey-HSD: $-2.877 < T_{28} < 0.970$, p >
544 0.123; Fig 6A). The amount of dugong mtDNA extracted from *ex*-EE-L1 faeces was at

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

545 least 46× less than that extracted from *in*-EE-L3 faeces (Tukey-HSD: $T_{30} = -4.548$, $p < 0.001$), but was 247,253× higher than that isolated from *in*-EE-L5 faeces (Tukey-HSD: $T_{30} = 5.105$, $p < 0.001$) (Fig 6B). Amount of dugong mtDNA was not significantly different between *ex*-EE-L1 and *in*-EE-L4 faeces (Tukey-HSD: $T_{29} = 1.910$, p= 0.246). However, amount of dugong mtDNA in *in*-EE-L3 faeces was at least 38,322× and 11,697,397× higher than that of the *in*-EE-L4 (Tukey-HSD: $T_{29} = 6.604$, p= < 0.001) and *in*-EE-L5 (Tukey-HSD: $T_{30} = 14.405$, p= < 0.001) faeces, respectively. The amount of dugong mtDNA was the same in *in*-EE-L4 and L5 faeces (Tukey-HSD: $T_{29} = 3.049$, p= 0.081).

554 PCR success for mtDNA amplification was similar between faeces of all EELs (Tukey-HSD: $-1.583 < Z < 1.775$, $p > 0.285$; Fig 6C). Despite insignificant results, all *in*-EE-L3 and L4 faeces amplified, but only 70% of *ex*-EE-L1 and 68.8% of *in*-EE-L5 faeces amplified successfully. However, *in*-EE-L3 faeces had higher triplicate success ($\bar{x}$ amplification= 3/3) compared to *ex*-EE-L1 ($\bar{x}$ amplification= 2.3/3, *W*-test: Z= 1.680, p= 0.046), *in*-EE-L4 ($\bar{x}$ amplification= 2.4/3, *W*-test: Z= 2.530, p= 0.006) and *in*-EE-L5 faeces ($\bar{x}$ amplification= 1.1/3, *W*-test: Z= 4.68, $p < 0.001$) (Fig 6D). Triplicate success of *ex*-EE-L1 faeces was not significantly different compared to that of *in*-EE-L4 faeces (*W*-test: Z= 0.444, p= 0.672) but was significantly higher than that of *in*-EE-L5 faeces (*W*-test: Z= 2.530, p= 0.006). The *in*-EE-L4 faeces also had a higher triplicate success than *in*-EE-L5 faeces (*W*-test: Z= 3.011, p= 0.001).

565 When the HV-CTAB-PCI method was used, the total DNA concentration extracted from *ex*-EE-L1 faeces (742.3 ± 70.3 ng/µL; mean ± standard error) was not significantly different to that extracted from *in*-EE-L2 faeces (559.4 ± 95.4 ng/µL, ANOVA: $T_{36} = 1.934$, p= 0.318), but was significantly higher than that extracted from *in*-EE-L3 (324.1 ± 19.8 ng/µL, ANOVA: $T_{36} = 4.372$, p= 0.001), *in*-EE-L4 (134.8 ± 14.4

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

570 ng/µL, ANOVA: $T_{36} = 9.253$, $p < 0.001$) and *in*-EE-L5 (198.3 ± 34.8 ng/µL, ANOVA:
571 $T_{36} = 7.597$, $p < 0.001$) faeces (Fig 7A). Total DNA concentration of *in*-EE-L2 faeces was
572 not significantly different compared to that of *in*-EE-L3 faeces (ANOVA: $T_{36} = 2.369$, p =
573 0.147), but was significantly higher that than of *in*-EE-L4 (ANOVA: $T_{36} = 7.113$, p <
574 0.001) and *in*-EE-L5 (ANOVA: $T_{36} = 5.503$, $p < 0.001$) faeces. The *in*-EE-L3 faeces also
575 had higher concentration of total DNA compared to *in*-EE-L4 (ANOVA: $T_{36} = 4.744$, p <
576 0.001) and *in*-EE-L5 (ANOVA: $T_{36} = 3.134$, $p = 0.027$) faeces, but total DNA
577 concentration between *in*-EE-L4 and L5 faeces was not significantly different (ANOVA:
578 $T_{36} = 1.610$, $p = 0.501$).

579

580 **Fig 7. Comparison of variables measured in faeces with different environmental**
581 **exposure levels (EELs) representing varying quality.**

582 (A) Total DNA concentration extracted from each faecal type. (B) Graph shows a positive
583 correlation between PCR success and total DNA concentration. Lines represent mean ±
584 standard error. (C) Percent amplification that was successful and unsuccessful for mtDNA
585 and ZFX primers between different faecal types. (D) Triplicate success between the faecal
586 types for both primers. Similar letters (with same colour and letter case) denote no statistical
587 significance, while different letters denote statistical significance ($\alpha= 0.05$). Different
588 statistical comparisons made were indicated by different colour and letter case.

589

590 PCR success and triplicate success (Kruskal-Wallis: $\chi^2= 4.125$, $p= 0.389$) for
591 mtDNA amplification were the same across faeces of all EELs, as all samples amplified
592 successfully (Fig 7C, D). Mean triplicate success for faeces of all EELs was 3, except for
593 *in*-EE-L4 faeces which was 2.9. For ZFX amplification, faeces with higher total DNA

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

594 concentration tended to have higher PCR success (GLM: Z= 2.558, p=0.011; Fig 7B).
595 PCR success of *ex*-EE-L1 faeces was similar to that of *in*-EE-L2 faeces (χ^2= 1.195, p=
596 0.471), but was significantly higher than that of *in*-EE-L3 (χ^2= 5.885, P= 0.032), *in*-EE-
597 L4 (χ^2= 13.39, p< 0.001), and *in*-EE-L5 (χ^2= 11.46, p= 0.003) faeces (Fig 7C). All *ex*-EE-
598 L1 faeces amplified, but only 87.5% and 50.0% of *in*-EE-L2 and L3 faeces amplified,
599 respectively. Only 12.5% of both *in*-EE-L4 and L5 faeces amplified. PCR success of *in*-
600 EE-L2 faeces was not significantly different compared to that of *in*-EE-L3 faeces (χ^2=
601 2.618, p= 0.283), but was higher than that of *in*-EE-L4 (χ^2= 9.00, P= 0.012) and *in*-EE-L5
602 faeces (χ^2= 7.244, p= 0.015). PCR success of *in*-EE-L3, L4, and L5 faeces was not
603 significantly different from one another ($0.275 < \chi^2 < 2.618$, p> 0.284). Triplicate success
604 of *ex*-EE-L1 faeces ($\bar{x}$ amplification= 2.1/3) was similar to that of *in*-EE-L2 faeces ($\bar{x}$
605 amplification= 1.9/3, *W-t*est: Z= 0.711, p= 0.761), but higher than that of *in*-EE-L3 ($\bar{x}$
606 amplification= 0.6/3, *W*-test: Z= 2.197, p= 0.014), *in*-EE-L4, and L5 (both $\bar{x}$
607 amplification= 0.1/3, both *W*-tests: Z= 2.759, p= 0.003) faeces (Fig 7D). The *in*-EE-L2
608 faeces also had a higher triplicate success compared to *in*-EE-L3 (*W*-test: Z= 1.794, P=
609 0.036), *in*-EE-L4, and L5 (both *W*-tests: Z= 2.473, p= 0.007) faeces, but triplicate success
610 of *in*-EE-L3, L4 and L5 faeces were not significantly different from each other (*W*-tests:
611 Z < 1.029, p> 0.152).

612 SNP amplifications in *in-ōceanum* faeces

613 All *in*-EE-L2 faecal samples amplified for all dugong SNP primers, except for
614 primers Dug12 and Dug63, in which 50% and 75% of all samples amplified, respectively
615 (Fig 8A). The SNP regions were successfully located and aligned, showing the form of allele
616 possessed (S3 Fig). However, PCR success was not significantly different between
617 amplification of different primer sets ($\chi^2 < 5.33$, p> 0.079; Fig 8A). Triplicate success of

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

618 primer Dug12 ($\bar{x}$ amplification= 1.3/3) was significantly lower than that of primers Dug50,
619 Dug53, and Dug62 (All $\bar{x}$ amplifications= 3/3, *W*-tests: Z= 1.694, p= 0.045) (Fig 8B).
620 Triplicate success of primer Dug63 ($\bar{x}$ amplification= 1.3/3) was equal to that of Dug12, but
621 significantly lower than that of all other primers (All $\bar{x}$ amplifications > 2.8/3, *W*-tests: Z >
622 2.115, p < 0.017).

623

624 **Fig 8. PCR success and triplicate success between SNP primers and faeces with different**
625 **environmental exposure levels (EELs), representing its quality.**

626 (A) PCR success of different SNP primers amplified. No statistical significance exists
627 between primers used. (B) Triplicate success of different SNP primers amplified. (C) PCR
628 success between faeces of different EELs. (D) Triplicate success between faeces of different
629 EELs. Similar letters denote no statistical significance, while different letters denote
630 statistical significance (α= 0.05).

631

632 For amplification of SNP primer Dug54 in all *in-ōceanum* faeces, the PCR success of
633 *in*-EE-L2 faeces was significantly higher than that of *in*-EE-L5 faeces (χ^2= 9.600, p= 0.008),
634 but similar to that of *in*-EE-L3 and L4 faeces (χ^2< 5.333, p > 0.077) (Fig 8C). All *in*-EE-L2
635 faeces amplified, and 87.5%, 50.0%, and 25.0% of *in*-EE-L3, L4, and L5 faeces amplified,
636 respectively. The *in*-EE-L3 faeces also had a higher PCR success compared to *in*-EE-L5
637 faeces (χ^2= 6.349, p= 0.042), but it had a similar PCR success as *in*-EE-L4 faeces (χ^2= 2.618,
638 p= 0.283). The PCR success of *in*-EE-L4 and L5 faeces was not significantly different (χ^2=
639 1.067, p= 0.603). The triplicate success of *in*-EE-L2 faeces ($\bar{x}$ amplification= 2.9/3) was
640 significantly higher than that of the *in*-EE-L3 ($\bar{x}$ amplification= 1.6/3, *W*-test: Z= 2.382, p=
641 0.009), *in*-EE-L4, and L5 faeces (both $\bar{x}$ amplification= 0.5/3, both *W*-tests: Z= 2.929, p=

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

642 0.002) (Fig 8D). The *in*-EE-L3 faeces also had higher triplicate success compared to *in*-EE-
643 L4 (*W*-test: Z= 1.937, p= 0.026) and *in*-EE-L5 faeces (*W*-test: Z= 1.706, p= 0.044). However,
644 triplicate success of *in*-EE-L4 and L5 faeces was the same (*W*-test: Z= 0.438, p= 0.669).

645

646 # Discussion

647 This study demonstrated that mitochondrial and nuclear markers can be amplified
648 from dugong faeces, using a novel 'High volume-CTAB-PCI' (HV-CTAB-PCI) DNA
649 extraction method. Numerous modifications of the 2CTAB/PCI method by Vallet *et al* [18]
650 resulted in this efficacious method for DNA isolation from a large quantity of faecal material.
651 The amplification success of mtDNA was found to be similar regardless of whether faecal
652 material was sampled from the outer surface or inner core of a stool. However, the period to
653 which faeces were exposed to the environment had significant impacts on amplification
654 success of nDNA markers. Although the amplification success of mtDNA was adequate with
655 the QIAamp method, the use of HV-CTAB-PCI resulted in improvements, especially for
656 eroded faeces (*in*-EE-L5). Regardless of the method used, nuclear markers amplified better in
657 fresher faeces than eroded ones. Therefore, the hypothesis suggesting that QIAamp method
658 would enable both mtDNA and nDNA amplifications when sampling from a stool's outer
659 surface was rejected; the hypothesis suggesting increased efficacy of amplification in fresher
660 faeces was supported by these results.

661

662 ## Development of HV-CTAB-PCI method

663 The HV-CTAB-PCI method developed in this study was cost-effective, user-friendly,
664 and highly efficacious at isolating sufficient quantity and quality of dugong DNA from scats.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

665 Although mtDNA amplification was highly successful using DNA extracted following the
666 QIAamp protocol, ZFX amplification was inconsistent, even for fresh faecal samples. This
667 contrasts with Takoukam Kamla's [32] results showing that the QIAamp protocol enabled
668 amplifications of both microsatellites and mtDNA from manatee scats. In his study, pre-
669 amplification enhanced the amount of manatee DNA that was originally extracted, thus
670 improving the subsequent PCRs for nDNA amplification. However, pre-amplification in pilot
671 experiments (including touchdown protocol) conducted in this study yielded no ZFX (nDNA)
672 amplifications in the final PCRs. This may suggest that there is only a scant amount of DNA
673 in dugong faeces. If the amount of dugong DNA originally extracted was of a negligible
674 proportion compared to the robust quantities of exogenous DNA, then such a dilute presence
675 of dugong DNA would render an infinitesimal probability that dugong DNA is pipetted into
676 each well of the PCR plates [55]. If no dugong DNA was pipetted into the PCR reactions,
677 then pre-amplification would not increase the quantity of target DNA, thus resulting in non-
678 amplification of target DNA in subsequent PCRs. Another notable contributor to Takoukam
679 Kamla's success was likely the addition of a purification step post-DNA isolation, removing
680 PCR inhibitors that may have been co-extracted with the DNA [32]. However, the consistent
681 amplification of dugong mtDNA and results from dilution experiments performed in this
682 study showed only low levels of PCR inhibition. Consequently, an additional purification
683 step was not utilised in this study, in an attempt to minimise DNA loss [56]. Furthermore, this
684 study showed a positive correlation between concentration of total DNA extracted and PCR
685 success. This again hinted at a problem of scarcity of DNA in dugong faeces rather than
686 issues of PCR inhibitors. As dugongs feed on a more restricted diet [35] compared to
687 manatees [33], it is possible that there are fewer and less diverse plant secondary metabolites
688 (e.g., tannins) to inhibit PCR in dugong faeces. Additionally, unlike the open ocean where
689 dugongs are found, the estuarine and freshwater water bodies that manatees inhabit often

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

690 contain high concentrations of tannins due to the decomposition of plant material from the
691 surrounding forests and mangroves. Furthermore, the surface of dugong faeces is generally
692 smoother and less fibrous than that of manatees. This may affect the relative quantity of
693 epithelial cells found within dugong faeces, as the rougher, harder scats of manatees may
694 stimulate more mucus secretion that entraps exfoliated cells [13].

695 A simple way to increase total DNA extracted from faeces is to increase the total
696 amount of faecal material used for DNA extraction. This study utilised 220 mg of dugong
697 faeces for DNA extraction with the QIAamp protocol as this is the maximum amount
698 recommended by the manufacturer. Despite the retention of a higher volume of supernatant
699 that contains DNA and the increase in subsequent volumes of extraction buffers used, this
700 method did not isolate sufficient target DNA for consistent ZFX amplification in this study.
701 Takoukam Kamla [32] managed to use a slightly larger amount (300 – 1260 mg) of manatee
702 faeces without increasing the volumes of any reagents used. However, that might have
703 contributed to the lower purity of those DNA extracts since there may have been insufficient
704 reagents to efficiently clean and protect the extracted DNA, making an extra purification step
705 essential. Since target DNA insufficiency was a problem in this study, the HV-CTAB-PCI
706 method was developed to enable a more economical higher volume extraction of faeces. This
707 method was modified from Vallet *et al*'s [18] 2CTAB/PCI protocol that was successful in
708 extracting DNA from faeces of herbivorous primates. As the reagents used were self-
709 prepared and the function of each chemical was known, this new method provided the
710 flexibility to modify usage and concentrations of any reagents to suit DNA extraction from
711 dugong faeces.

712 This novel HV-CTAB-PCI method is markedly different to the 2CTAB/PCI
713 approach. The minimum effective sample-to-reagent ratio used in this method (1:5) was first
714 determined through a 1 mL-by-1 mL addition of Lysis Buffer 1 to 1 g of faecal sample. This

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

715 differs from the ratios used in the 2CTAB/PCI approach [18], which would require large
716 volumes of reagents, when scaled up, to extract DNA from 1 g of faeces. Instead of an
717 overnight incubation, cell lysis duration was reduced to 3 h to shorten the DNA extraction
718 procedure. To separate DNA from proteins and other impurities, an effective equal PCI-to-
719 supernatant ratio (1:1) was utilised [57]. Proteins were digested using protease instead of the
720 widely used Proteinase K in the HV-CTAB-PCI method as protease was found to be effective
721 at removing proteins, and costs less than the latter enzyme. As the presence of RNA did not
722 present problems in PCR amplifications, RNAse was not used in this newly developed
723 protocol. To minimise DNA loss, the DNA pellet retrieved at the end of extraction was
724 washed once only [56]. Despite a reduced washing frequency, PCR was not significantly
725 inhibited and the amplification successes of mtDNA and ZFX were higher using this new
726 protocol compared to the QIAamp method. This again showed that PCR inhibitors were
727 unlikely to be a major issue within dugong faeces.

728

729 Sampling from different layers of faeces of varying quality

730 Amplification success and relative quantity of mtDNA did not differ between DNA
731 extracted from the outer surface and inner core of dugong stools. This result was surprising as
732 faecal DNA extractions in most molecular scatology studies have sampled the outer surface
733 of a scat only, e.g., [8,14]. This study concurs with Stenglein *et al* [58] which found no
734 difference in PCR success and error rates of DNA extracted from outer and inner layers of
735 brown bear scats. Although lower allelic dropout and higher genotyping success rates were
736 achieved when DNA from the outer surface of stools was used [58], it would be improper to
737 conclude that most of the target animal cells are located on the outer surface as those
738 measures only reflect the quality of DNA extracted. As the total DNA yield tended to be

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

739 higher when outer surface was sampled, it is possible that the small sample size (total n= 32)
740 in this study limited the detection of significance between outer and inner layers of faecal
741 stools. However, it is likely that most of the DNA extracted was from organisms that were
742 not of interest (e.g., bacteria, plants) since the relative amount of dugong mtDNA was the
743 same regardless of layers sampled. Instead, we hypothesise that the target animal cells and
744 therefore their DNA would be heterogeneously distributed throughout the faecal mass. Since
745 faeces are unformed in the upper regions of the colon, exfoliated epithelial cells from those
746 regions would likely be incorporated within the faecal mass through peristaltic contractions
747 of the colon [12,59].

748 The degree to which voided dugong faeces were exposed to the environment was
749 found to significantly influence amplification success and relative amount of recovered DNA.
750 Comparing mtDNA amplifications, highly eroded stools were found to amplify poorly
751 compared to fresher ones. Eroded faeces were anticipated to amplify poorly as DNA degrades
752 over time [60]. Epithelial cells undergo apoptosis as they detach from the basement
753 membrane [61], where enzymes that degrade DNA break down the phosphodiester bonds that
754 form the backbone of DNA [62]. This breaks the DNA into progressively shorter fragments
755 and eventually into nucleotides [62]. The new HV-CTAB-PCI method, however, greatly
756 improved mtDNA amplification, allowing consistent amplification even in eroded faecal
757 samples. As faecal DNA is often highly degraded, the extraction of DNA from a larger
758 amount of faecal material, used in the HV-CTAB-PCI method, can increase the chance for
759 more target DNA of sufficient length or quality to be extracted, thus allowing for a higher
760 amplification success. As such, this newly developed DNA extraction method also enabled
761 the comparison of ZFX amplifications in faeces of different environmental exposure levels
762 (EELs). The amplification of ZFX marker was found to be highest in freshest faeces (*ex*-EE-
763 L1) compared to older ones. Interestingly, the amplification success and relative quantity of

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

764 mtDNA were found to be higher in fresh *in-ōceanum* faeces (*in*-EE-L3) compared to the *ex-*
765 *ōceanum* faeces (*ex*-EE-L1), in contrast to what one might expect. The amount of amplifiable
766 target DNA is influenced by the physical loss and chemical degradation (e.g., fragmentation)
767 of DNA. Although the loss of target DNA could be higher when faeces is transported in the
768 ocean, sea salts such as Ca^{2+} and Mg^{2+},which DNA has a higher affinity for compared to Na^{+}
769 [63,64], may help to stabilise or preserve the integrity of DNA [64,65,66]. Therefore, the
770 higher relative quantity of mtDNA in *in*-EE-L3 faeces may be due to preservative effects of
771 salt. However, the higher DNA loss in *in*-EE-L4 and L5 faeces may have outweighed any
772 preservative effects on DNA, leading to the observed results. Although it remains unknown
773 why the ZFX amplifications failed to reproduce the same trend, it is possible that the
774 structural differences (circular versus linear) or the cellular location of mtDNA versus nDNA
775 may have influenced their degradative vulnerability [67].

776 Although this study achieved a low amplification success in more eroded faeces, the
777 chance of success may be enhanced by scaling up the amount of starting faecal material used
778 for DNA extraction. If that fails, separation of host DNA from bacterial DNA could be
779 trialled using the host-DNA enrichment method by Chiou and Bergey [68] as this may
780 improve the purity of DNA extracted. However, whole genome amplification may need to be
781 performed prior to enrichment if the amount of target DNA extracted is insufficient for an
782 efficient separation. Alternatively, host epithelial cells could be separated from a large
783 proportion or the entire faecal mass using the methodology developed by Matsushita et al.
784 [69] and DNA extraction could then be directly performed on those cells.

785

786 SNP amplifications using faecal DNA

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

787 This study showed a proof-of-concept for the potential utilisation of faecal DNA in
788 population genetic studies of dugongs, as all ten SNPs were successfully amplified using the
789 dugong faecal DNA extracted with the new HV-CTAB-PCI method. The SNP primers did
790 not significantly influence amplification success, but interestingly, primers Dug12 and Dug63
791 amplified less than other SNP primers. One plausible explanation is that those target sites
792 may be more sensitive to DNA degradation. For example, Johnston *et al* [70] showed that
793 regions where SNP markers showed polyploidy due to historical genome duplication in
794 Atlantic salmon had higher sensitivity to DNA degradation and thus had lower genotyping
795 success. Alternatively, the primers used in this study that were designed for those target sites
796 may not have been optimal. The impact of DNA degradation was also reflected in SNP
797 amplifications as fresh faeces amplified more successfully than eroded ones. Since SNP
798 amplicons are generally shorter than microsatellites [30], the difficulty for SNP amplification
799 in eroded faeces may indicate that the DNA extracted from these faeces was too degraded or
800 fragmented for primers to bind and amplify at the target site. That said, only one SNP marker
801 (Dug54) was used to compare the amplification success between faeces of different EELs
802 which limits the extent of inferences that can be made.

803

804
Conclusion

805 This research has successfully pioneered a novel DNA extraction protocol (HV-
806 CTAB-PCI) for an herbivorous marine mammal, the dugong, and has demonstrated the most
807 efficient sampling methodologies to maximise retrieval of target DNA from faecal samples,
808 to enable amplification of both mitochondrial and nuclear markers used in genetic population
809 studies. This study also provided a preliminary idea on the distribution of sloughed epithelial
810 cells and thus target DNA within faecal stools, which may help researchers increase the

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

811 efficiency and success for target DNA recovery. The main limitation of this research was the
812 inability to isolate DNA of sufficient quality from the oldest, most eroded stools for
813 amplification of target nDNA. Despite that, multiple suggestions have been made to increase
814 the chances for DNA retrieval. As the first study to demonstrate consistent amplification of
815 nDNA from dugong faeces, the development of the HV-CTAB-PCI method will facilitate
816 non-invasive genetic studies in areas where direct sampling is unfeasible. For instance, faecal
817 samples have been collected from some remote and turbid water regions in northern
818 Queensland where dugongs could not be sampled directly. These samples may now be
819 analysed to gain a fuller understanding on the population structure of dugongs along the
820 entire eastern Queensland coast. Future application of this approach could include broad-
821 scale population genetic studies on dugongs throughout the Indo-Pacific region where
822 resources for direct sampling are typically unavailable but obtaining samples for genetic
823 analysis is critical.

824

825 # Acknowledgements

826 Thanks to the UQ Dugong Team for assistance in sample collection, especially Helen
827 Peereboom. Morgan McCarthy provided SNP amplicon sequences, Jennifer Seddon advised
828 on SNP analysis, Simone Blomberg advised on statistics, while Seth and Matthew Kazmer
829 assisted with hand-drawn illustrations. Any use of trade, firm, or product names is for
830 descriptive purposes only and does not imply endorsement by the U.S. Government.

831

832 # References

833 1. García-Dorado A, Caballero A. Neutral genetic diversity as a useful tool for

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

834 conservation biology. Conserv Genet. 2021;22: 541–5. https://doi.org/10.1007/s10592-
835 021-01384-9
836 2. Kardos M, Armstrong EE, Fitzpatrick SW, Hauser S, Hedrick PW, Miller JM, et al.
837 The crucial role of genome-wide genetic variation in conservation. Proc Natl Acad Sci
838 U S A. 2021;118(48): e2104642118. https://doi.org/10.1073/pnas.2104642118.
839 3. Frankham R. Genetics and extinction. Biol Conserv. 2005;126(2): 131–40.
840 https://doi.org/10.1016/j.biocon.2005.05.002.
841 4. Wong P, Wiley EO, Johnson WE, Ryder OA, O'brien SJ, Haussler D, et al. Tissue
842 sampling methods and standards for vertebrate genomics. Gigascience. 2012;1: 2047-
843 217X. https://doi.org/10.1186/2047-217X-1-8.
844 5. Fernando P, Vidya TNC, Rajapakse C, Dangolla A, Melnick DJ. Reliable noninvasive
845 genotyping: Fantasy or reality? J Hered. 2003;94(2): 115–23.
846 https://doi.org/10.1093/jhered/esg022.
847 6. Taberlet P, Luikart G, Waits LP. Noninvasive genetic sampling: Look before you leap.
848 Trends Ecol Evol. 1999;14(8): 323–7. https://doi.org/10.1016/S0169-5347(99)01637-
849 7.
850 7. Boston ESM, Puechmaille SJ, Scott DD, Buckley DJ, Lundy MG, Montgomery IW, et
851 al. Empirical assessment of non-invasive population genetics in bats: Comparison of
852 DNA quality from faecal and tissue samples. Acta Chiropt. 2012;14(1): 45–52.
853 https://doi.org/10.3161/150811012X654259.
854 8. Ramón-Laca A, Soriano L, Gleeson D, Godoy JA. A simple and effective method for
855 obtaining mammal DNA from faeces. Wildlife Biol. 2015;21(4): 195–203.
856 https://doi.org/10.2981/wlb.00096.
857 9. Schrader C, Schielke A, Ellerbroek L, Johne R. PCR inhibitors - occurrence, properties
858 and removal. J Appl Microbiol. 2012;113(5): 1014–26. https://doi.org/10.1111/j.1365-

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

859 2672.2012.05384.x.

860 10. Gillett RM, White BN, Rolland RM. Quantification and genetic profiling of DNA
861 isolated from free-floating feces of the North Atlantic right whale (*Eubalaena*
862 *glacialis*). Mar Mammal Sci. 2008;24(2): 341–55. https://doi.org/10.1111/j.1748-
863 7692.2008.00192.x.

864 11. Zhang M, Wei M, Dong Z, Duan H, Mao S, Feng S, et al. Fecal DNA isolation and
865 degradation in clam *Cyclina sinensis*: Noninvasive DNA isolation for conservation and
866 genetic assessment. BMC Biotechnol. 2019;19(1): 4–11.
867 https://doi.org/10.1186/s12896-019-0595-6.

868 12. Ryan L, Wong YT, Dwyer KM, Clarke D, Kyprian L, Craig JM. Coprocytobiology: a
869 technical review of cytological colorectal cancer screening in fecal samples. SLAS
870 Technol. 2021;26(6): 591–604. https://doi.org/10.1177/24726303211024562.

871 13. White V, Scarpini C, Barbosa-Morais NL, Ikelle E, Carter S, Laskey RA, et al.
872 Isolation of stool-derived mucus provides a high yield of colonocytes suitable for early
873 detection of colorectal carcinoma. Cancer Epidemiol Biomarkers Prev. 2009;18(7):
874 2006–13. https://doi.org/10.1158/1055-9965.EPI-08-1145.

875 14. Rutledge LY, Holloway JJ, Patterson BR, White BN. An improved field method to
876 obtain DNA for individual identification from wolf scat. J Wildl Manage. 2009;73(8):
877 1430–5. https://doi.org/10.2193/2008-492.

878 15. Tende T, Hansson B, Ottosson U, Bensch S. Evaluating preservation medium for the
879 storage of DNA in African lion *Panthera leo* faecal samples. Curr Zool. 2014;60(3):
880 351–8. https://doi.org/10.1093/czoolo/60.3.351.

881 16. Parsons KM. Reliable microsatellite genotyping of dolphin DNA from faeces. Mol
882 Ecol Notes. 2001;1(4): 341–4. https://doi.org/10.1046/j.1471-8278.2001.00098.x.

883 17. Green ML, Herzing DL, Baldwin JD. Noninvasive methodology for the sampling and

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

884 extraction of DNA from free-ranging Atlantic spotted dolphins (*Stenella frontalis*).
885 Mol Ecol Notes. 2007;7(6): 1287–92. https://doi.org/10.1111/j.1471-
886 8286.2007.01858.x.
887 18. Vallet D, Petit EJ, Gatti S, Levréro F, Ménard N. A new 2CTAB/PCI method
888 improves DNA amplification success from faeces of Mediterranean (*Barbary*
889 *macaques*) and tropical (lowland gorillas) primates. Conserv Genet. 2008;9(3): 677–
890 80. https://doi.org/10.1007/s10592-007-9361-8.
891 19. Marsh H, Sobtzick S. *Dugong dugon* (amended version of 2015 assessment). IUCN
892 Red List Threat Species. 2019;8235: eT6909A160756767.
893 https://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T6909A160756767.en
894 20. Plön S, Thakur V, Parr L, Lavery SD. Phylogeography of the dugong (*Dugong dugon*)
895 based on historical samples identifies vulnerable Indian Ocean populations. PLoS One.
896 2019;14(9): 1–19. https://doi.org/10.1371/journal.pone.0219350.
897 21. Tikel D. Using a genetic approach to optimise dugong (*Dugong dugon*) conservation
898 management. PhD Thesis, James Cook Universsity. 1997. Available from:
899 http://researchonline.jcu.edu.au/28125/
900 22. Marsh H, OShea TJ, Reynolds III JE. Ecology and conservation of the Sirenia. Vol.
901 18. Cambridge: Cambridge University Press; 2011. pp. 327–393. Available from:
902 http://ebooks.cambridge.org/ref/id/CBO9781139013277
903 23. Kayanne H, Hara T, Arai N, Yamano H, Matsuda H. Trajectory to local extinction of
904 an isolated dugong population near Okinawa Island, Japan. Sci Reports. 2022
905 Apr;12(1): 1–8. Available from: https://www.nature.com/articles/s41598-022-09992-2
906 24. McDonald B. Population genetics of dugongs around Australia: Implications of gene
907 flow and migration. PhD Thesis, The University of Queensland. 2005.
908 25. Harrison RG. Animal mitochondrial DNA as a genetic marker in population and

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

909 evolutionary biology. Trends Ecol Evol. 1989;4(1): 6–11.
910 https://doi.org/10.1016/0169-5347(89)90006-2.
911 26. Seddon JM, Ovenden JR, Sneath HL, Broderick D, Dudgeon CL, Lanyon JM. Fine
912 scale population structure of dugongs (*Dugong dugon*) implies low gene flow along
913 the southern Queensland coastline. Conserv Genet. 2014;15(6): 1381–92.
914 https://doi.org/10.1007/s10592-014-0624-x.
915 27. McGowan A. Ecological genetics of dugongs (*Dugong dugon*) in Queensland. PhD
916 Thesis, The University of Queensland. 2019.
917 28. Brookes AJ. The essence of SNPs. Gene. 1999;234(2): 177–86.
918 https://doi.org/10.1016/S0378-1119(99)00219-X.
919 29. Morin PA, Luikart G, Wayne RK. SNPs in ecology, evolution and conservation.
920 Trends Ecol Evol. 2004;19(4): 208–16. https://doi.org/10.1016/j.tree.2004.01.009.
921 30. Freire-Aradas A, Fondevila M, Kriegel AK, Phillips C, Gill P, Prieto L, et al. A new
922 SNP assay for identification of highly degraded human DNA. Forensic Sci Int Genet.
923 2012;6(3): 341–9. http://dx.doi.org/10.1016/j.fsigen.2011.07.010
924 31. Andréasson H, Nilsson M, Budowle B, Lundberg H, Allen M. Nuclear and
925 mitochondrial DNA quantification of various forensic materials. Forensic Sci Int.
926 2006;164(1): 56–64. https://doi.org/10.1016/j.forsciint.2005.11.024.
927 32. Takoukam Kamla A. Genetic diversity, diet, and habitat quality of the African manatee
928 (*Trichechus senegalensis*) in the downstream of the Sanaga River watershed,
929 Cameroon. PhD Thesis, University Of Florida. 2019.
930 33. Castelblanco-Martínez DN, Morales-Vela B, Hernández-Arana HA, Padilla-Saldivar J.
931 Diet of the manatees (*Trichechus manatus manatus*) in Chetumal Bay, Mexico. Lat
932 Am J Aquat Mamm. 2009;7(1–2): 39-46. https://doi.org/10.5597/lajam00132.
933 34. Lanyon JM. The nutritional ecology of the dugong in tropical north Queensland. PhD

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

934 Thesis, Monash and James Cook University. 1991.
935 35. Lanyon JM, Sanson GD. Mechanical disruption of seagrass in the digestive tract of the
936 dugong. J Zool. 2006;270(2): 277–89. https://doi.org/10.1111/j.1469-
937 7998.2006.00135.x.
938 36. Murray RM, Marsh H, Heinsohn GE, Spain AV. The role of the midgut caecum and
939 large intestine in the digestion of sea grasses by the dugong (Mammalia: Sirenia).
940 Comp Biochem Physiol -- Part A Physiol. 1977;56(1): 7–10.
941 https://doi.org/10.1016/0300-9629(77)90432-7.
942 37. Best RC. Foods and feeding habits of wild and captive Sirenia. Mamm Rev.
943 1981;11(1): 3–29. https://doi.org/10.1111/j.1365-2907.1981.tb00243.x.
944 38. Lanyon JM, Sneath HL, Long T, Bonde RK. Physiological response of wild dugongs
945 (*Dugong dugon*) to out-of-water sampling for health assessment. Aquat Mamm.
946 2010;36(1): 46–58. https://doi.org/10.1578/AM.36.1.2010.46.
947 39. Lanyon JM, Sneath HL, Kirkwood JM, Slade RW. Establishing a mark-recapture
948 program for dugongs in Moreton Bay, south-east Queensland. Aust Mammal.
949 2002;24(1): 51–6. https://doi.org/10.1071/AM02051.
950 40. Tol SJ, Harrison M, Groom R, Gilbert J, Blair D, Coles R, et al. Using DNA to
951 distinguish between faeces of *Dugong dugon* and *Chelonia mydas*: non-invasive
952 sampling for IUCN-listed marine megafauna. Conserv Genet Resour. 2021;13(2):
953 115–7. https://doi.org/10.1007/s12686-020-01187-z
954 41. McHale M, Broderick D, Ovenden JR, Lanyon JM. A PCR assay for gender
955 assignment in dugong (*Dugong dugon*) and West Indian manatee (*Trichechus*
956 *manatus*). Mol Ecol Resour. 2008;8(3): 669–70. https://doi.org/10.1111/j.1471-
957 8286.2007.02041.x.
958 42. Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den hoff MJB, et

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

959 al. Amplification efficiency: Linking baseline and bias in the analysis of quantitative
960 PCR data. Nucleic Acids Res. 2009;37(6): e45. https://doi.org/10.1093/nar/gkp045.
961 43. McCarthy M. Reading the bones - how archived skulls can inform temporal genetic
962 variation in central Queensland's dugong population. PhD Thesis, The University of
963 Queensland. 2018.
964 44. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis
965 Version 11. Mol Biol Evol. 2021;38(7): 3022–3027.
966 https://doi.org/10.1093/molbev/msab120.
967 45. R Core Team. R: A language and environment for statistical computing. R Foundation
968 for Statistical Computing. 2020. https://www.R-project.org/
969 46. Wickham H. ggplot2: Elegant graphics for data analysis. In: Springer-Verlag. New
970 York. 2016.
971 47. Rudis B. hrbrthemes: Additional themes, theme components and utilities for 'ggplot2'.
972 R package version 0.8.0. 2020. https://CRAN.R-project.org/package=hrbrthemes.
973 48. Hope RM. Rmisc: Ryan miscellaneous. R package version 1.5.1. 2022.
974 https://CRAN.R-project.org/package=Rmisc.
975 49. Aphalo P. ggpmisc: Miscellaneous extensions to 'ggplot2'. R package version 0.4.7.
976 2022. https://CRAN.R-project.org/package=ggpmisc.
977 50. Ahlmann-Eltze C, Patil I. ggsignif: R package for displaying significance brackets for
978 'ggplot2'. PsyArxiv. 2021. doi:10.31234/osf.io/7awm6
979 51. Kassambara A. ggpubr: 'ggplot2' based publication ready plots. R package version
980 0.4.0. 2020. https://CRAN.R-project.org/package=ggpubr.
981 52. Pinheiro J, Bates D, R Core Team. nlme: Linear and nonlinear mixed effects models. R
982 package version 3. 2022; 1-157. https://CRAN.R-project.org/package=nlme.
983 53. Venables WN, Ripley BD. Modern applied statistics with S. 4th ed. New York:

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

984 Springer; 2002.
985 54. Kosmidis I. brglm2: Bias reduction in generalized linear models. R package version
986 0.8.2. 2021. https://CRAN.R-project.org/package=brglm2.
987 55. Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, et al.
988 Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic
989 Acids Res. 1996;24(16): 3189–94. https://doi.org/10.1093/nar/24.16.3189.
990 56. Guo W, Jiang L, Bhasin S, Khan SM, Swerdlow RH. DNA extraction procedures
991 meaningfully influence qPCR-based mtDNA copy number determination.
992 Mitochondrion. 2009;9(4): 261–5. https://doi.org/10.1016/j.mito.2009.03.003.
993 57. Green MR, Sambrook J. Isolation of high-molecular-weight DNA using organic
994 solvents. Cold Spring Harb Protoc. 2017;4: 356–9.
995 https://doi.org/10.1101/pdb.prot093450.
996 58. Stenglein JL, De Barba M, Ausband DE, Waits LP. Impacts of sampling location
997 within a faeces on DNA quality in two carnivore species. Mol Ecol Resour.
998 2010;10(1): 109–14. https://doi.org/10.1111/j.1755-0998.2009.02670.x.
999 59. Ahlquist DA, Harrington JJ, Burgart LJ, Roche PC. Morphometric analysis of the
1000 “mucocellular layer” overlying colorectal cancer and normal mucosa: relevance to
1001 exfoliation and stool screening. Hum Pathol. 2000;31(1): 51–7.
1002 https://doi.org/10.1016/S0046-8177(00)80198-7.
1003 60. Piggott MP. Effect of sample age and season of collection on the reliability of
1004 microsatellite genotyping of faecal DNA. Wildl Res. 2004;31(5): 485–93.
1005 https://doi.org/10.1071/WR03096.
1006 61. Strater J, Wedding U, Barth TFE, Koretz K, Elsing C, Moller P. Rapid onset of
1007 apoptosis in vitro follows disruption of β1- integrin/matrix interactions in human
1008 colonic crypt cells. Gastroenterology. 1996;110(6): 1776–84.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

1009 https://doi.org/10.1053/gast.1996.v110.pm8964403.

1010 62. Kawane K, Motani K, Nagata S. DNA degradation and its defects. Cold Spring Harb
1011 Perspect Biol. 2014;6(6): 1–14. https://doi.org/10.1101/cshperspect.a016394.

1012 63. Korolev N, Lyubartsev AP, Rupprecht A, Nordenskiöld L. Competitive binding of
1013 Mg^{2+}, Ca^{2+}, Na^{+}, and K^{+} ions to DNA in oriented DNA fibers: Experimental and
1014 Monte Carlo simulation results. Biophys J. 1999;77(5): 2736–49.
1015 https://doi.org/10.1016/S0006-3495(99)77107-9.

1016 64. Sigel RKO, Sigel H. A stability concept for metal ion coordination to single-stranded
1017 nucleic acids and affinities of individual sites. Acc Chem Res. 2010;43(7): 974–84.
1018 https://doi.org/10.1021/ar900197y.

1019 65. Shved N, Haas C, Papageorgopoulou C, Akguel G, Paulsen K, Bouwman A, et al. Post
1020 mortem DNA degradation of human tissue experimentally mummified in salt. PLoS
1021 One. 2014;9(10): e110753. https://doi.org/10.1371/journal.pone.0110753.

1022 66. Cruz-León S, Vanderlinden W, Müller P, Forster T, Staudt G, Lin YY, et al. Twisting
1023 DNA by salt. Nucleic Acids Res. 2022;50(10): 5726–38.
1024 https://doi.org/10.1093/nar/gkac445.

1025 67. Foran DR. Relative degradation of nuclear and mitochondrial DNA: An experimental
1026 approach. J Forensic Sci. 2006;51(4): 766–70. https://doi.org/10.1111/j.1556-
1027 4029.2006.00176.x.

1028 68. Chiou KL, Bergey CM. Methylation-based enrichment facilitates low-cost,
1029 noninvasive genomic scale sequencing of populations from feces. Sci Rep.
1030 2018;1975(8): 1–10. https://doi.org/10.1038/s41598-018-20427-9.

1031 69. Matsushita H, Matsumura Y, Moriya Y, Akasu T, Fujita S, Yamamoto S, et al. A new
1032 method for isolating colonocytes from naturally evacuated feces and its clinical
1033 application to colorectal cancer diagnosis. Gastroenterology. 2005;129(6): 1918–27.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

1034 https://doi.org/10.1053/j.gastro.2005.10.007.

1035 70. Johnston SE, Lindqvist M, Niemelä E, Orell P, Erkinaro J, Kent MP, et al. Fish scales
1036 and SNP chips: SNP genotyping and allele frequency estimation in individual and
1037 pooled DNA from historical samples of Atlantic salmon (*Salmo salar*). BMC
1038 Genomics. 2013;14(1): 1-13. https://doi.org/10.1186/1471-2164-14-439.

1039

1040 Supporting information

1041 **S1 Fig. Flowcharts summarising the experimental designs and approaches taken in this**
1042 **study.**

1043 (A) The first stage of this study targeted the development of a DNA extraction protocol for
1044 dugong faeces. (B) The second stage of this study assessed the influence of sampling between
1045 outer surface and inner core of faeces, in addition to impacts of faecal environmental
1046 exposure, on the amplification success of mtDNA and ZFX markers. (C) The final stage of
1047 this study sought to amplify ten SNP markers to show a proof-of-concept that population
1048 genetic markers used for genetic studies of dugongs can be amplified using their faecal DNA.
1049 The impacts on amplification were compared between different *In-ōceanum* faeces.

1050

1051 **S2 Fig. Results from a pilot study conducted to determine initial amount of total DNA**
1052 **(10 and 100 ng) to be added into qPCR assay for comparison of mtDNA amplification**
1053 **between outer surface and inner core of faeces with different EELs.**

1054 (A) Stacked bar charts of PCR success for the different initial total DNA used. No difference
1055 was found between the initial DNA amount used (χ^2= 0.381, p= 1.000). (B) Violin plots

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

1056 incorporating box plots of triplicate success for the different initial total DNA. No difference
1057 was found between the total DNA used (Kruskal-Wallis χ^2= 0.804, df= 1, p= 0.370).

1058

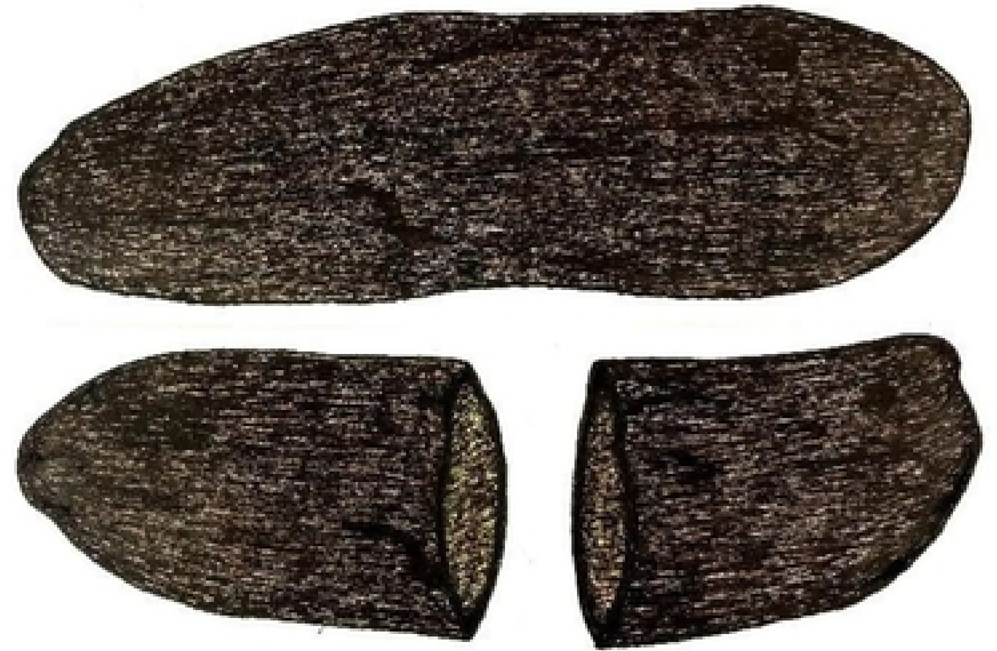
1059 **S3 Fig. Alignment of SNP sequences amplified using ten dugong SNP primers with two**
1060 **forms of alleles known.**

1061 Three random samples were chosen to be shown in the plot for each primer set. (A) Primer
1062 Dug12. (B) Primer Dug40. (C) Primer Dug50. (D) Primer Dug51. (E) Primer Dug53. (F)
1063 Primer Dug54. (G) Primer Dug60. (H) Primer Dug62. (I) Primer Dug63. (J) Primer Dug64.

preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright
under 17 USC 105 and is also made available for use under a CC0 license.

A. *Ex-ōceanum*— EE-Level 1

B. *In-ōceanum*— EE-Level 2 and 3



C. *In-ōceanum*— EE-Level 4

D. *In-ōceanum*— EE- Level 5

Fig1

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

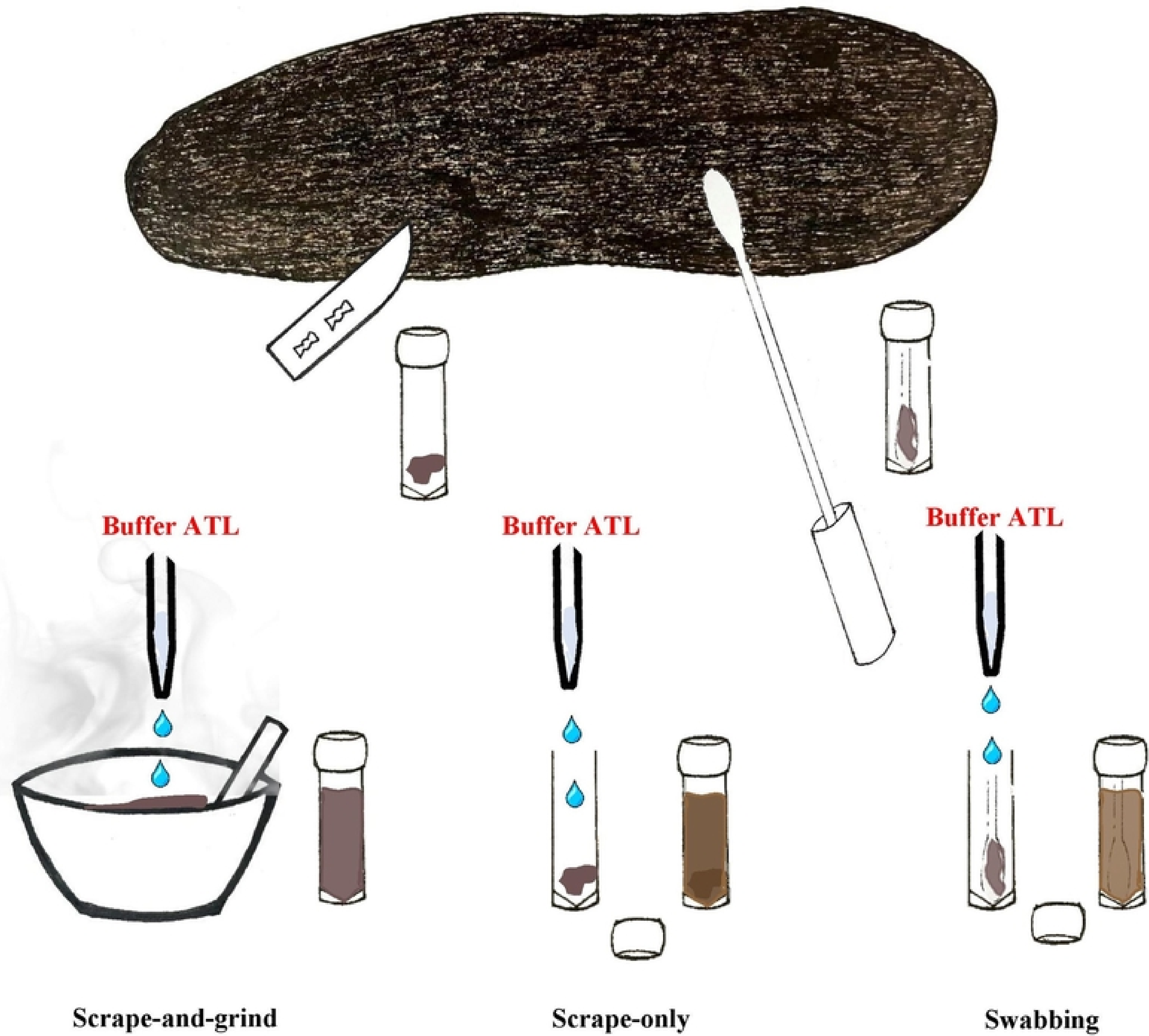


Fig2

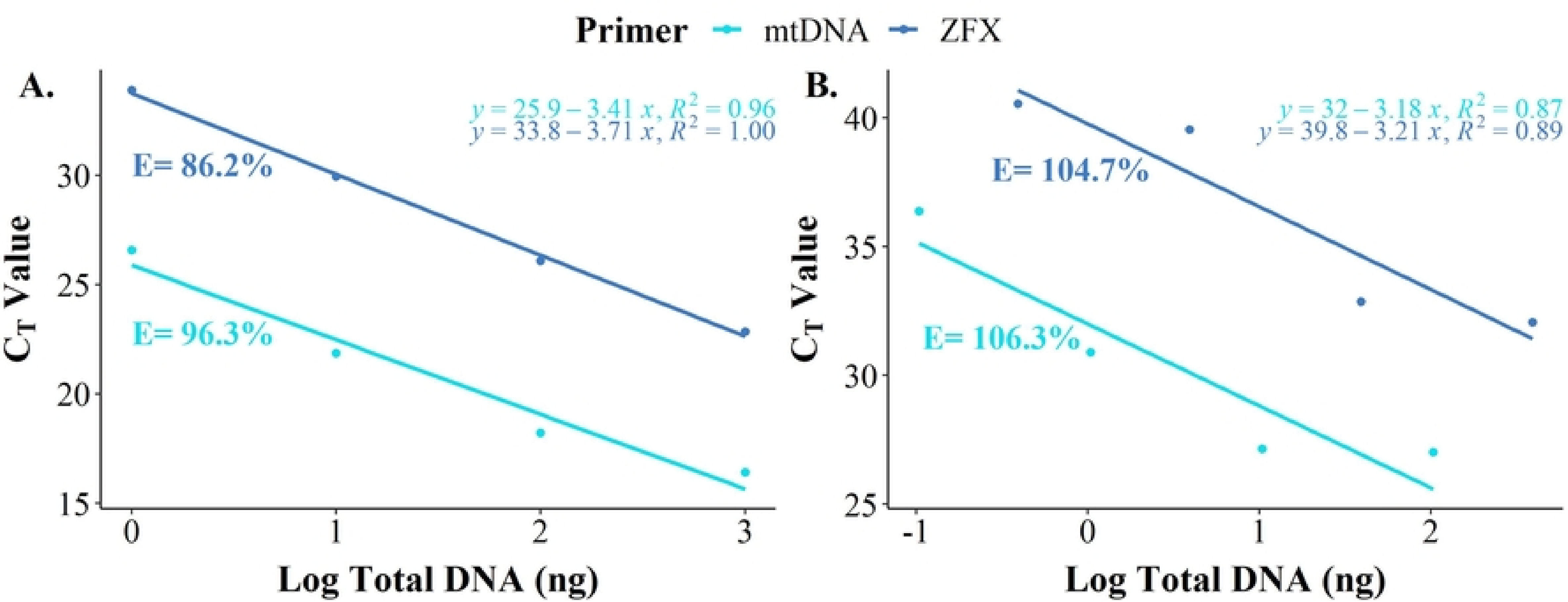


Fig3

A.

Total DNA Concentration (ng/µL)

75
50
25

a
b
c

Swabbing
Scrape-Only
Scrape-and-Grind

Faecal Sampling Technique

B.

$1/C_T$ Value

0.045
0.040
0.035

a
a
b

Swabbing
Scrape-Only
Scrape-and-Grind

Faecal Sampling Technique

Fig4

A.

Total DNA Concentration (ng/μL)

750
500
250

b
a
a

Tissue Control-DNeasy
Faeces-QIAamp
Faeces-HV-CTAB-PCI

DNA Extraction Method

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517804; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

B.

PCR Outcome: Not successful, Successful

Tissue Control DNeasy
Faeces QIAamp
Faeces HV-CTAB-PCI

a A a B a A

Percent Amplification

100%
75%
50%
25%
0%

mtDNA ZFX mtDNA ZFX mtDNA ZFX

Primer

C.

Primer: mtDNA, ZFX

a A a a
B
C

Triplicate Success

3
2
1
0

Tissue Control DNeasy
Faeces QIAamp
Faeces HV-CTAB-PCI

DNA Extraction Method

Fig5

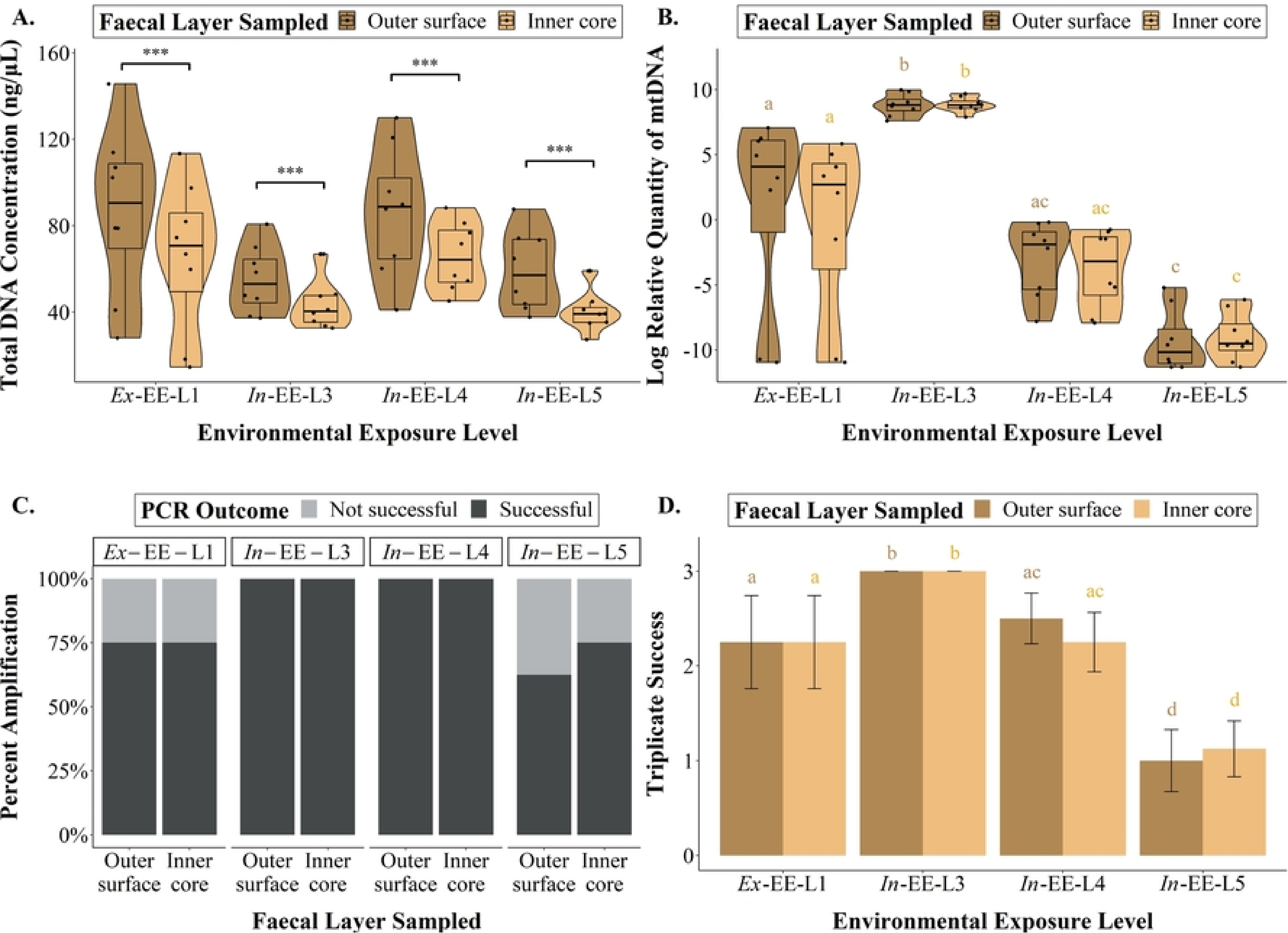
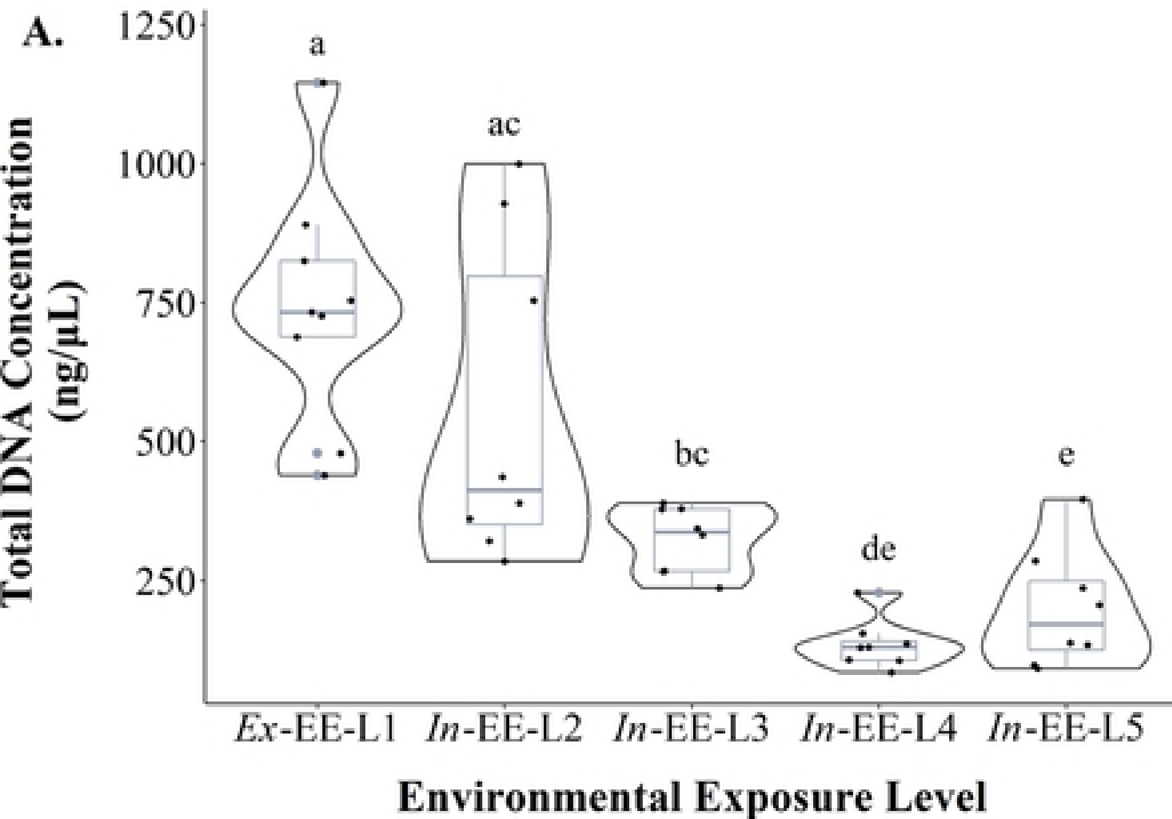
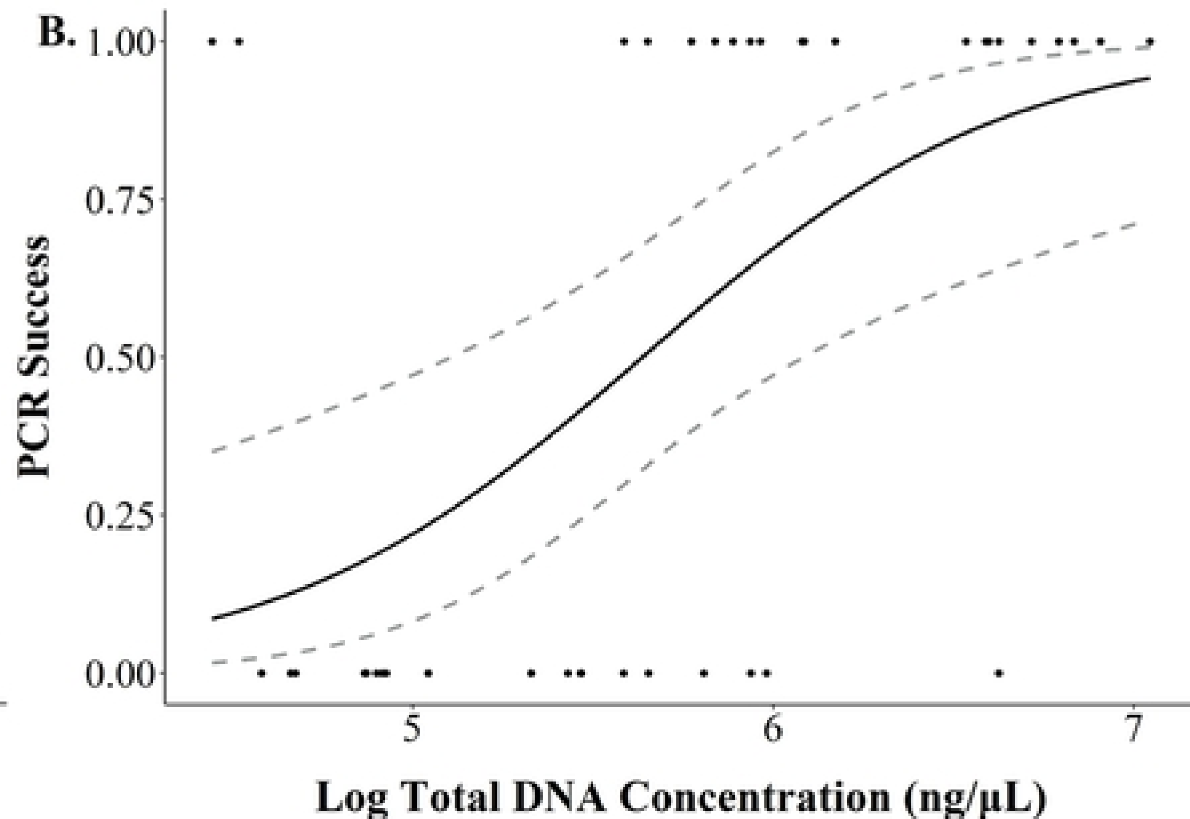
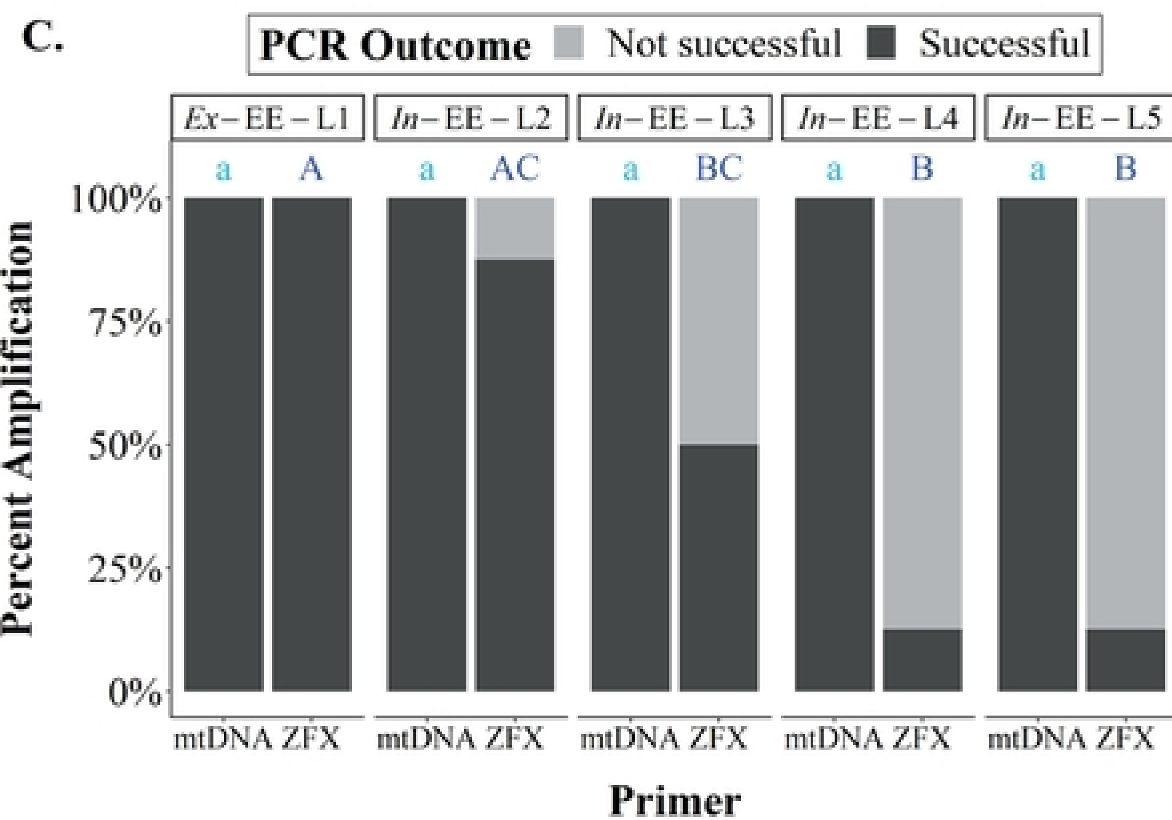


Fig6







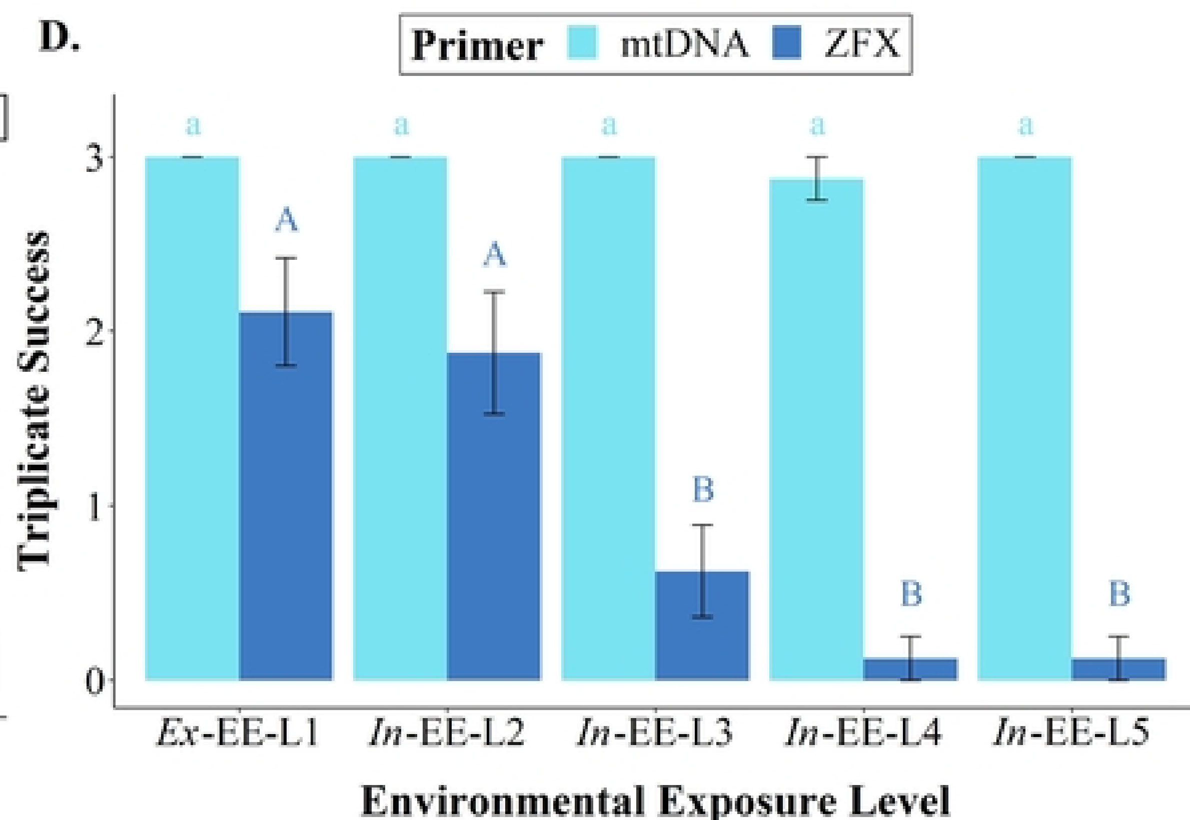
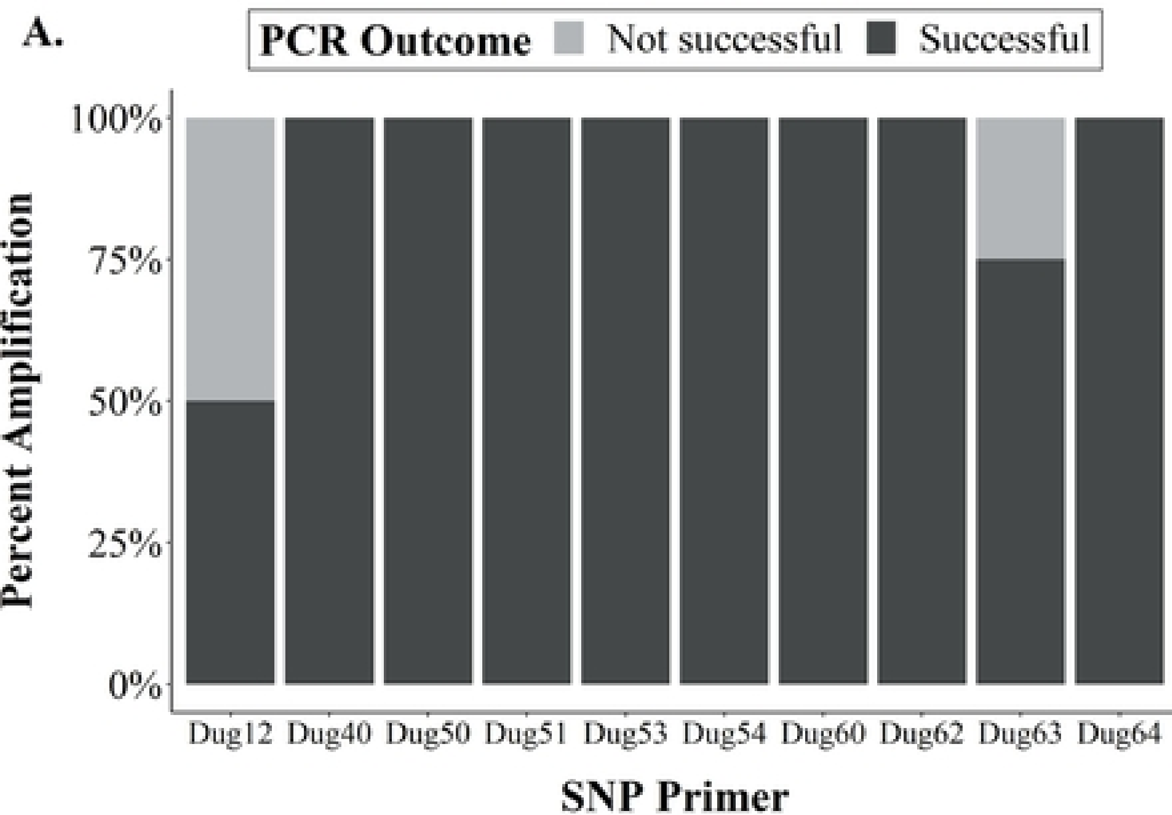
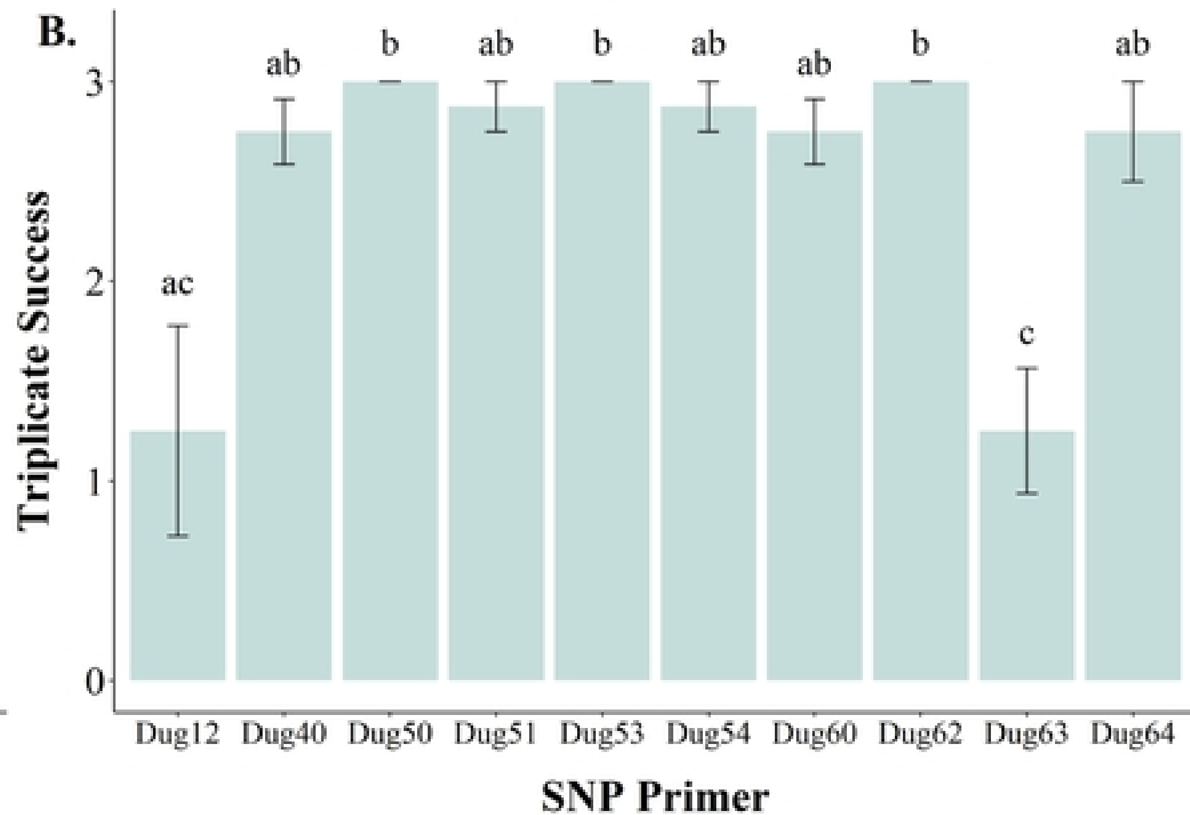
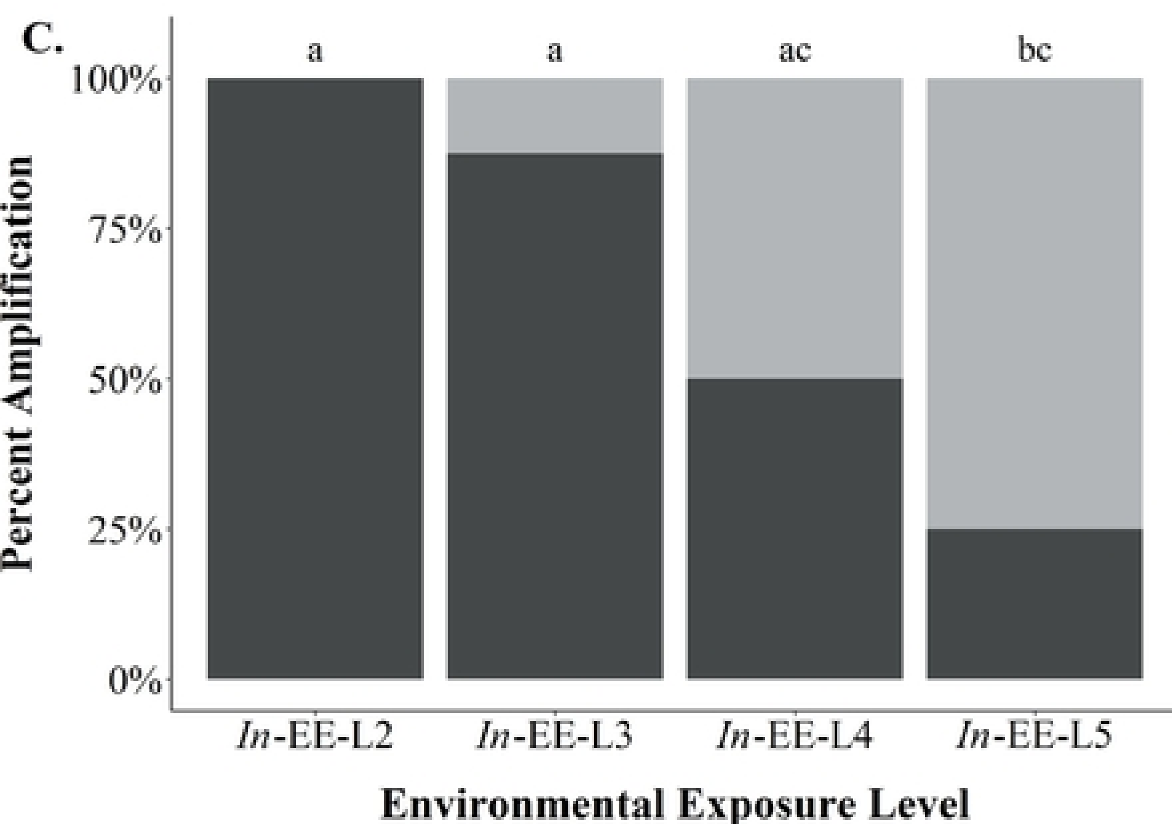


Fig7







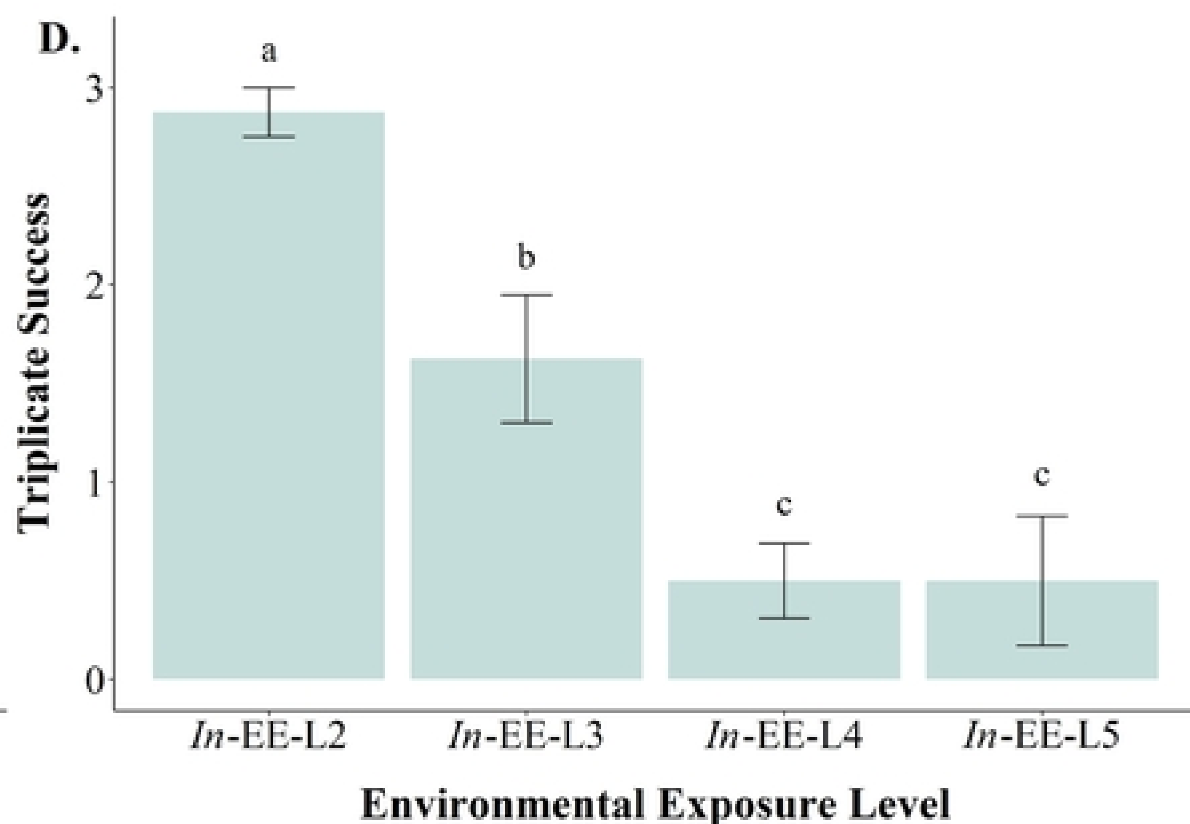


Fig8