bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1
2
From presence/absence to reliable prey proportions: A field test of dietary DNA for characterizing seabird diets
3
4
Gemma V Clucas[a,b], Andrew Stillman[a,b], Elizabeth C Craig[c].
5
6
7 [a]Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA.
8 [b]Cornell Atkinson Center for Sustainability, Cornell University, Ithaca, NY, USA.
9 [c]Shoals Marine Laboratory, joint program of Cornell University, Ithaca, NY, USA and the
10 University of New Hampshire, Durham, NH, USA.
11
12 Corresponding author: Gemma Clucas, gemma.clucas@cornell.edu.
13
14
15
Abstract
16
17 Climate change and human impacts are causing rapid shifts in species' distributions and
18 abundance, potentially disrupting predator-prey relationships. Monitoring animal diets can
19 elucidate these relationships and potentially provide a mechanistic understanding of population
20 declines, but efficient methods to monitor diets are needed. Dietary DNA can be used to gather
21 data on species' diets with high-taxonomic resolution, but the question of whether it can provide
22 quantitative information on animal diets has so far limited its applications. Here we show that
23 dietary DNA can efficiently provide quantitative information on the relative proportions of fish in
24 the diets of piscivorous seabirds. Over three breeding seasons, we observed common tern chick
25 feeding events and collected fecal samples from chicks. We compared the frequency of
26 occurrence (FOO) and relative read abundance (RRA) of fish prey in the fecal samples to the
27 relative biomass recorded through visual observations and found a high correlation ($R = 0.94$)
28 between RRA and relative biomass at the colony level. We also found that RRA outperformed
29 FOO in capturing interannual changes in relative prey biomass, and that FOO systematically over-
30 estimated the relative importance across all prey categories. The high taxonomic resolution
31 provided by dietary DNA and the high correlation between RRA and relative biomass we found
32 here suggest that dietary DNA can be used as a quantitative metric to monitor seabird diets. This
33 has important applications for studying the impact of changing forage fish availability on seabird
34 populations, but also for harnessing seabirds as sentinels of ecosystem health. Since quantitative
35 diet data can be collected highly efficiently using dietary DNA, this method could provide key
36 information about food web dynamics and forage fish availability as we move towards
37 ecosystem-based fisheries management.
38
39
40
Keywords
41
42 Metabarcoding, diet monitoring, fecal DNA, forage fish.

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

43 ## 1. Introduction
44
45 Species extinctions and population declines continue at alarming rates (Barnosky et al., 2011;
46 Ceballos et al., 2017; Dirzo et al., 2014; Pimm et al., 2014; Rosenberg et al., 2019), and
47 addressing these declines requires a mechanistic understanding of the complex drivers and
48 environmental conditions that impact populations. Monitoring animal diets can provide a critical
49 window into these mechanisms. Diet may be highly influenced by environmental conditions and
50 resource availability, and diet in turn can impact an individual's survival, reproduction, and
51 fitness (Spiller and Dettmers, 2019). Dietary DNA has become an increasingly powerful and
52 popular tool to study these links over the last 15 years due to the ease and taxonomic resolution
53 afforded by dietary DNA over traditional diet monitoring techniques (de Sousa et al., 2019;
54 Pompanon et al., 2012; Valentini et al., 2009). Frequently, fecal samples are used in dietary DNA
55 studies as they can be collected non-invasively, although the method can also be used to identify
56 species from gut contents or regurgitated material. To identify prey species, dietary DNA studies
57 may employ one of three techniques: 1) using targeted PCR to detect prey species of interest
58 through the use of highly specific PCR primers; 2) using a metabarcoding approach, whereby
59 barcoding genes are amplified and sequenced from a broad range of taxa using degenerative or
60 universal primers; and 3) using a metagenomic approach, whereby all DNA molecules contained
61 in the sampled material are sequenced. Fecal metabarcoding is by far the most popular of these
62 methods because of its relative efficiency (vs metagenomic approaches) and its ability to identify
63 multiple species simultaneously (vs targeted PCR), making it a powerful tool to increase our
64 mechanistic understanding of conservation challenges in the Anthropocene.
65
66 A major limitation in the use of dietary DNA in conservation and management is the notion that
67 it cannot produce quantitative estimates of animal diets (Hoenig et al., 2022). That is, the
68 proportion of sequence reads (relative read abundance or RRA) recovered from each prey
69 species does not reflect prey biomass consumed. Indeed, there are many factors that can distort
70 RRA in dietary DNA studies, including prey-specific DNA content and digestibility (Thomas et al.,
71 2014), secondary consumption (Bowser et al., 2013), contamination from the environment or in
72 the lab, PCR stochasticity, preferential amplification and sequencing of certain prey sequences
73 over others, and bioinformatic biases (see Alberdi et al., 2019 for a full review). Captive feeding
74 studies and metabarcoding of tissue mixtures have shown variable results (reviewed in Deagle et
75 al., 2019) with RRA often showing a correlation with biomass but with low accuracy. Therefore,
76 dietary DNA results are often converted to presence/absence data as a conservative alternative
77 to RRA. However, Deagle et al. (2019) showed through simulations that converting sequence
78 counts to presence/absence data can over-estimate the importance of rare diet items because
79 they are counted equally alongside items consumed in greater proportions. Secondary
80 consumption and contamination can also have a greater impact on presence/absence results.
81 Therefore, RRA may give a more accurate representation of population-level diet compared to
82 presence/absence in many circumstances (Deagle et al., 2019). At present, however, few studies
83 have validated this method in wild populations (Verkuil et al., 2020), and more system-specific
84 validation studies are needed before dietary DNA can reach its full potential as an effective tool
85 for conservation.
86

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

87 Seabirds are one of the most threatened groups of vertebrates (Croxall et al., 2012). Monitored
88 populations have declined by 70% between 1950 - 2010 (Paleczny et al., 2015), and threats from
89 interactions with fisheries and climate change continue to raise conservation concerns for these
90 species (Dias et al., 2019; Paleczny et al., 2015; Phillips et al., 2022). Fisheries-induced mortality
91 may occur directly through bycatch, particularly in long-line and trawl fisheries, or indirectly,
92 through competition for prey resources (Anderson et al., 2011; Phillips et al., 2022). Climate-
93 linked seabird declines are often indirect and the result of changing prey resources. For example,
94 climate-induced shifts in prey availability and over-fishing of key prey species has caused
95 declining breeding success of Atlantic puffins in the North Atlantic (Durant et al., 2003; Fayet et
96 al., 2021; Kress et al., 2016; Scopel et al., 2019). Since seabirds tend to forage across large
97 geographic areas, often nest in large colonies, and feed primarily on forage fish and other hard-
98 to-monitor prey species, they have long been recognized as sentinels of marine ecosystem
99 health (Ainley et al., 1996; Cairns, 1987; Furness and Camphuysen, 1997; Piatt et al., 2007), with
100 their diets providing quantitative estimates of the availability of forage fish and squid (Diamond
101 and Devlin, 2003; Montevecchi, 1993; Montevecchi and Myers, 1995; Scopel et al., 2017;
102 Velarde et al., 2019, 1994). However, traditional methods to monitor their diets through
103 stomach content analysis, stable isotopes, or visual observations of chick feeding events
104 frequently lack taxonomic resolution due to the difficulty in identifying prey to species level. This
105 lack of taxonomic specificity contributes to uncertainty in how preyscapes critical to seabird
106 survival will shift under future climate scenarios. Traditional methods can also be invasive or
107 restricted to bill-loading birds. As an alternative, tracking changes in diet through quantitative
108 dietary DNA could greatly increase our ability to monitor real-time impacts on seabird
109 populations and proactively address fisheries and climate-driven threats.
110
111 Our objectives in this study were to (1) determine the extent to which dietary DNA can estimate
112 fish prey proportions in a wild seabird population and (2) investigate the effects of sample size
113 on these estimates. Our focal system was the common tern (*Sterna hirundo*) nesting in the Gulf
114 of Maine, where visual observations have historically been used to monitor prey fed to chicks by
115 adult terns. We compared chick dietary DNA to these visual observations and focused exclusively
116 on fish, excluding the analysis of invertebrate prey, since fish are the primary food source for
117 these birds (Arnold et al., 2020). To validate dietary DNA against visual observations, we
118 calculated four metrics commonly used in diet studies:
119 1) Relative count: the relative proportion of prey taxa based on counts from visual
120 observations.
121 2) Relative biomass: the relative proportion of prey taxa based on estimated prey biomass
122 from visual observations.
123 3) FOO (frequency of occurrence): the proportion of samples in which prey taxa are
124 present in the dietary DNA (i.e. presence/absence of prey species).
125 4) RRA: the relative proportion of DNA sequences of prey taxa in the dietary DNA.
126 Relative prey counts are often used to estimate prey proportions, but relative biomass estimates
127 provide a more informative metric to characterize diets when data on prey size are available.
128 Thus, we focused our inference on assessing whether dietary DNA could capture the relative
129 biomass of prey. We also simulated an annual monitoring scheme to investigate the effects of
130 sample size on the ability of dietary DNA to match estimates from visual observations. If dietary

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

131 DNA can match the inferential capabilities of traditional methods such as visual observations,
132 this approach opens doors to track real-time changes in prey availability and more efficiently
133 address the impacts of fisheries management and climate change on prey resources critical to
134 seabird populations.
135
136
137
138 ## 2. Methods
139
140 ### 2.1 Study site and species
141 This study was conducted at a mixed-species tern breeding colony on White and Seavey islands
142 (42°58' N, 70°37' W) located in the Isles of Shoals archipelago, New Hampshire, USA, within the
143 Gulf of Maine. These islands are New Hampshire's primary breeding colony for common, roseate
144 (*S. dougallii*), and Arctic terns (*S. paradisaea*), and the colony is a key conservation area with
145 regional significance. Restoration, management, and research activities have been conducted at
146 this site since 1997 under the auspices of the New Hampshire Fish and Game Department. Data
147 presented here were collected in June and July of 2017, 2018, and 2019 during routine
148 monitoring activities. In the years of this study, the colony annually supported 2000–3000 pairs
149 of common terns, 60–90 pairs of roseate terns, and 1–2 pairs of arctic terns.
150
151 ### 2.2 Visual observations
152 Common terns provision their chicks with whole prey items, typically fish of 0 to 4 years of age
153 and up to 15cm in length (Arnold et al., 2020). Visual observations of chick provisioning were
154 conducted in June and July of each year, either in-person with binoculars from observation
155 blinds or remotely using video footage recorded from an AXIS P5635-E Mk II PTZ Network
156 Camera. Observations occurred between the hours of 05:30 and 21:30. Visual provisioning data
157 were collected in 2017 from video recordings at 13 nests for a total of 160 nest hours, in 2018
158 from video recordings at 14 nests for a total of 167 nest hours, and in 2019 from both video
159 recordings and in-person observations at 22 nests for a total of 310 nest hours.
160
161 For each provisioning event, prey items were visually identified to the highest possible
162 taxonomic resolution. While tern diets included both fish (described below) and invertebrate
163 prey (primarily euphausiids, insects, and squid), invertebrates were not within the scope of this
164 study and excluded from further analysis. Prey fish were grouped into six categories:
165 (1) Hake (likely *Urophycis chuss, U. tennuis,* and *Merluccius bilinearis*).
166 (2) Herring (likely *Clupea harengus* and *Alosa spp*.).
167 (3) Sand lance (*Ammodytes* spp.).
168 (4) Atlantic butterfish (*Peprilus triacanthus*).
169 (5) Other identifiable fish.
170 (6) Unknown fish (including fish that were unidentifiable due to the speed of provisioning,
171 obscured view of identifying features, or lack of confidence on behalf of the observer).
172
173 Prey size was quantified by visually estimating the length of each prey item relative to that of the
174 adult tern's bill to the nearest 0.5 bill lengths, then multiplying by average adult bill length at our

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

175 study site (3.65 cm). The biomass of each fish was calculated by inserting our estimate of fish
176 length into length-weight relationships from the literature (Schneider et al., 2000; Wigley et al.,
177 2003; Winters, 1989). The length-weight equation for Atlantic herring was used for all fish in the
178 herring category, and the equation for white hake was used for all fish in the hake category. For
179 identifiable fish species with no published length-weight relationships, we substituted equations
180 of morphologically similar species for which equations were available. The equation for white
181 hake was also used for fish in the unknown fish category, as white hake has a narrow, fusiform
182 shape, typical of the morphology of many unidentified fish in our study. Provisioning events were
183 further categorized into successful or failed feedings depending on whether the provisioned prey
184 was successfully consumed by the chick. Failed feedings, where prey was either stolen or
185 discarded, were excluded from further analysis for this study.
186

187 ## 2.3 Fecal sample collection

188 We collected fecal samples from common tern chicks (≤ 2 weeks old) if they defecated during
189 handling as part of standard monitoring procedures. Fecal samples were transferred to 2ml
190 tubes pre-filled with 1 ml of DNA/RNA Shield preservation buffer (Zymo Research, Irving, CA)
191 using single-use, individually-wrapped spatulas. Samples were only collected from "clean"
192 surfaces, such as pre-cleaned weighing scales, bird bags, clothing, or hands. To monitor for DNA
193 contamination from these surfaces, we collected negative control samples (hereafter "field
194 blanks") from all surfaces by dipping a spatula into the preservation buffer to dampen it, wiping
195 it on the surface to be sampled two or three times to simulate picking up a sample, and stirring it
196 back into the preservation buffer as if depositing a sample into the tube. We treated field blanks
197 equally to fecal samples for subsequent metabarcoding steps. Samples were stored at ambient
198 temperature, out of direct sunlight, for up to 1 year until DNA extractions took place.
199

200 ## 2.4 DNA extractions and metabarcoding

201 DNA extractions were performed using Zymo Research Quick-DNA Fecal/Soil Miniprep 96 kits.
202 We found that the 96-well bead bashing plates supplied in the kit would leak during beating, so
203 we transferred each sample and its preservation buffer to an individual bead bashing tube to
204 avoid cross-contamination. We bead beat samples on a benchtop vortex set at maximum speed
205 for 20 minutes. The rest of the DNA extraction was performed following the manufacturer's
206 instructions after transferring the centrifuged supernatant from the bead bashing tubes into the
207 kit's 96-well filter plates. We eluted the DNA into a final volume of 60 µl of elution buffer.
208 Extractions were performed alongside other tern fecal sample extractions, with samples mixed
209 among plates to minimize batch effects. We included eight extraction blanks (no template
210 controls) in each plate to monitor for cross-contamination during extractions and performed all
211 DNA extractions and PCR set-ups in a dedicated clean lab.
212

213 Dietary metabarcoding was performed with a 2-step PCR approach. In the first step, the 12S
214 rRNA gene was amplified using MiFish primers (Miya et al., 2015; underlined below) to which we
215 had added TruSeq tails (primer sequences were thus MiFish-U-F-TruSeq: 5'-
216 ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCGGTAAAACTCGTGCCAGC-3'; and MiFish-U-R-
217 Truseq: 5'-TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCATAGTGGGGTATCTAATCCCAGTTTG-3').
218 We included a PCR blank (no template control) with every batch of PCRs to monitor for

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

219 contamination during PCR setup. A mock community (positive control), including DNA extracted
220 directly from fish samples, was also included and added to the plate last. The PCR reaction
221 included 6 µL of KAPA HiFi HotStart ReadyMix 2X (KAPA Biosystems, Wilmington,
222 Massachusetts), 0.7 µL of each of the forward and reverse primers at 5 µM concentration, and
223 4.6 µL of fecal DNA (or molecular grade water for the PCR blanks). For the mock community, we
224 reduced the amount of template DNA to 1 µL and added 3.6 µL of molecular grade water.
225 Thermocycling conditions were: 95 °C for 3 mins, 35 cycles of 98 °C for 30 secs, 65 °C for 30 secs,
226 72 °C for 30 secs, followed by a 5-minute extension at 72 °C. PCR products were visualized using
227 gel electrophoresis to ensure that all blanks were clear. PCRs were performed in duplicate to
228 minimize the effects of PCR stochasticity, with the products from two replicates being combined
229 and diluted depending on the strength of their bands before the second stage PCR.
230

231 The second stage PCR used the diluted product from the first stage as template and added the
232 flow cell binding sites and sequencing primer binding sites. The primer sequences were: forward,
233 5'-AATGATACGGCGACCACCGAGATCTACAXXXXXXXXACACTCTTTCCCTACACGAC-3' and reverse, 5'-
234 CAAGCAGAAGACGGCATACGAGATXXXXXXXXGTGACTGGAGTTCAGACGTGT-3'. The octo-X sites
235 represent the i7 and i5 indexes used to identify samples. We used unique dual indexes, such that
236 any reads that suffered from tag-jumping would be eliminated during de-multiplexing. The PCR
237 reaction used 6 µL of KAPA HiFi HotStart ReadyMix 2X, 0.7 µL of each of the forward and reverse
238 primers at 5 µM concentration, 1 µL of template and 3.6 µl of molecular grade water. The
239 thermal cycling profile was 94 °C for 3 mins, 15 cycles of 94 °C for 20 secs, 72 °C for 15 secs, and
240 a final extension of 7 mins at 72 °C. Sequencing was performed on an Illumina NovaSeq 6000
241 using an SP flow cell with 250bp paired-end reads at the University of New Hampshire's Hubbard
242 Center for Genome Studies. Samples were loaded alongside other projects using a small
243 percentage of a lane.
244

245 ## 2.5 Bioinformatics

246 Bioinformatics were performed using Qiime2 v2021.4 (Bolyen et al., 2019). First we trimmed
247 forward and reverse primers using the cutadapt plugin (Martin, 2011). We then denoised and
248 merged paired-end reads using the DADA2 plugin (Callahan et al., 2016), truncating the forward
249 and reverse reads to 133 and 138 bp, respectively, and specifying a minimum overlap of 50 bp
250 between them. After denoising, we merged all the data across the different sequencing plates
251 and then assigned taxonomy to the sequences using an iterative BLAST method and a custom
252 reference database. To create the reference database, we used the RESCRIPt plugin (Robeson et
253 al., 2021) to download any 12S or mitochondrial genomes from GenBank that originated from
254 fish or birds that were studied in our lab. Downloaded database sequences were cleaned and
255 dereplicated using RESCRIPt default parameters and finally, a human mitochondrial genome was
256 added to the database, as this is a common source of contamination. The iterative BLAST
257 method then took each representative sequence from our samples and blasted it 80 times
258 against the reference database, increasing the percent identity incrementally from 70 – 100 %,
259 thus circumventing the limitation of the BLAST method, which keeps only the first hit that meets
260 the search criteria, rather than the best hit. The script for the iterative BLAST method is available
261 from https://bitbucket.org/dwthomas/qiime2_tools/src/master/mktaxa.py. We chose this
262 method to assign taxonomy rather than training a Naïve Bayes classifier, as this method

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

263 identified all species in our mock community correctly, while a trained classifier mis-identified
264 one species (Atlantic butterfish). Next, we filtered out unassigned reads, reads originating from
265 human contamination, and reads originating from the birds themselves.
266
267 Alpha-rarefaction curves showed that 400 reads were necessary to capture the fish diversity in
268 the fecal samples. We therefore normalized all samples to a depth of 400 reads and dropped any
269 that had fewer. We then manually checked all species assignments by blasting representative
270 sequences against the full GenBank database and checked that the fishes' ranges included the
271 Gulf of Maine using FishBase (www.fishbase.se). Where multiple species occurred in the Gulf of
272 Maine and each had a 98% match or higher with our reference sequence, such as sand lances
273 (*Ammodytes dubius* and *Ammodytes americanus*) and river herrings (*Alosa aestivalis, Alosa*
274 *pseudoharengus,* and *Alosa sapidissima*) we did not attempt to make species assignments and
275 instead used genus-level assignments.
276

277 ## 2.6 Comparisons with visual observations

278 We summarized results from visual observations and dietary DNA using four separate metrics
279 which describe diet composition: relative counts, relative biomass, FOO, and RRA. Using data
280 from visual observations, we calculated relative counts by dividing the number of observations in
281 each prey group by the total number of prey observations. Similarly, we calculated relative
282 biomass by dividing the observed biomass in each group (i.e., number of prey × mass) by the
283 total biomass across observations. Next, we used dietary DNA results to calculate FOO and RRA.
284 Metrics based on dietary DNA had a higher taxonomic resolution than metrics based on visual
285 observations, so we grouped summaries into the following categories to match visual
286 observations: 1) hake, which included red hake, white hake, and silver hake; 2) herring, which
287 included Atlantic herring and river herrings; 3) sand lance; 4) Atlantic butterfish; and 5) other,
288 which included all other species detected in the fecal samples. Unlike the visual observations,
289 there was no "unknown fish" category. We calculated all four metrics separately for the three
290 sampling years and for all years combined. To facilitate comparisons, we omitted "other" and
291 "unknown fish" categories from statistical tests, and we completed analyses using R version 4.2.2
292 (R Core Team, 2022) aided by the package "FSA" (Ogle et al., 2023).
293
294 We used each prey category-year combination as the sampling unit for statistical tests, resulting
295 in a sample size for comparisons of n = 48. First, we calculated the Pearson correlation coefficient
296 for each pair of metrics. Next, we calculated the mean absolute error (MAE) and root mean
297 squared error (RMSE) to measure differences between each pair of metrics while accounting for
298 the magnitude of differences. RMSE is based on the average of squared differences between two
299 metrics, and comparisons based on RMSE tend to give a higher weight to large errors. We
300 compared MAE and RMSE between the four dietary metrics using a Kruskal-Wallis rank sum test
301 followed by a post-hoc Dunn test.
302
303 We also compared the ability of the metrics to track interannual changes in prey consumption. To
304 accomplish this, we calculated the proportional change between years for each metric using the
305 equation
306

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

307 $$\left(\frac{M_{t+1}}{M_t}\right) - 1$$

308

309 Where M_t is the value of a diet metric in an initial year, and M_{t+1} is the value of the same metric
310 in the following year. The results of this equation ranged from-0.90 to 7.80 and can be
311 interpreted as the change between consecutive years, expressed as a proportion of the initial
312 value. We then calculated the Pearson correlation coefficient for interannual changes detected
313 by each pair of metrics (hereafter "interannual correlation").

314

315 Finally, we used a simulation study to test how the relationship between visual observations
316 (specifically relative biomass) and dietary DNA (specifically RRA) changed based on the number
317 of fecal samples collected per year. This approach can (a) investigate whether additional fecal
318 sampling could further improve the correspondence between relative biomass and RRA in this
319 study system, and (b) inform sampling design considerations for future diet monitoring efforts
320 using dietary DNA. To accomplish these objectives, we started by sequentially subsampling the
321 dietary DNA dataset under a simulated sampling design ranging from 10–40 fecal samples/year.
322 We sampled without replacement for 2018 and 2019. For 2017 (n = 30) and simulations of > 30
323 samples/year, we used the full set of 30 samples and then resampled from 2017 to reach the
324 desired number. We repeated this procedure for 500 iterations given each sampling design and
325 calculated RRA for each species category-year at every iteration, resulting in a dataset of 500
326 calculations for each sampling design ranging from 10 to 40 samples/year. Using the relative
327 biomass for each species category-year as our observational "truth", we then calculated MAE,
328 RMSE, and interannual correlation between relative biomass and RRA for each simulated dataset
329 and iteration.

330

331

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

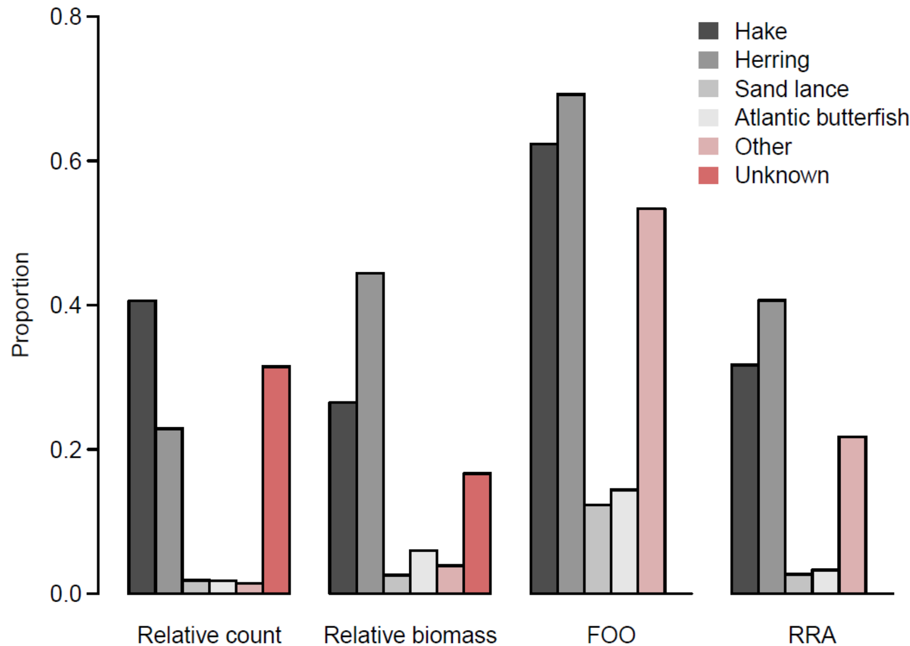
332
3. Results
333
334
3.1 Visual observations

335 We recorded 967 successful feeding events over 637 nest observation hours during the three
336 years of this study (Table 1). Based on visual observations, the most frequently observed prey
337 categories were hake, herring, and unknown fish (Figure 1). A large proportion of visual
338 observations fell into the unknown fish category (Figure 1) as the size, speed, or angle at which
339 fish were provisioned did not allow for a positive species identity to be determined by eye in
340 nearly a third of all provisioning observations (N=297). These unknown fish made up a large
341 proportion by count, but a smaller proportion by biomass, as unidentified fish were smaller on
342 average (5.8 ± 2.0 cm for unidentified fish compared to 6.9 ± 2.4 cm for fish with positive prey
343 IDs). Prey species within the "other identifiable fish" category included, in order of relative
344 frequency, mummichogs (*Fundulus heteroclitus*), cunner (*Tautogolabrus adspersus*), fourbeard
345 rockling (*Enchelyopus cimbrius*), other gadid species (likely haddock [*Melanogrammus*
346 *aeglefinus*], pollock [*Pollachius virens*], and Atlantic cod [*Gadus morhua*]), lumpfish (*Cyclopterus*
347 *lumpus*), Atlantic silverside (*Menidia menidia*), white sucker (*Catostomus commersonii*), and
348 Northern pipefish (*Syngnathus fuscus*). Hake had the highest relative count in the diet for all
349 years combined, but the higher mass of herring (average = 4.2 ± 3.6 g) compared to hake (1.4 ±
350 1.3 g) showed that herring contributed more to the total biomass of nestling diets (Figure 1).
351
352 Table 1. The number of visual observations and dietary DNA samples used in each year to study the diets of
353 common terns.

Year	Number of feeding events observed	Fecal samples collected	Fecal samples included in final analysis
2017	113	40	30
2018	163	55	42
2019	690	80	74

354
355

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



356
357 Figure 1. Diet characterization for wild common tern (*Sterna hirundo*) nestlings observed at a Gulf of Maine breeding
358 colony from 2017-2019. Diet metrics summarized from visual observations included relative count and relative
359 biomass. Diet metrics summarized from dietary DNA include frequency of occurrence (FOO) and relative read
360 abundance (RRA).
361
362

363 ## 3.2 Dietary DNA from fecal samples

364 Overall, we had a high sequencing success rate, allowing us to use 81% of samples in our final
365 analysis after all filtering steps (Table 1). Samples, blanks, and positive controls yielded nearly 63
366 million sequences in total, with a mean of 248,725 reads per sample and 219,574 reads per
367 blank. After adapter removal, denoising, and chimera removal, 73% of reads were retained in the
368 samples and 0.03% were retained in the blanks on average. After removing unassigned, human,
369 and avian sequences and normalizing the depth of all samples to 400 reads, all field-, extraction-,
370 and PCR blanks were dropped because they had fewer than 400 reads remaining. Contamination
371 from sample collection surfaces in the field and from cross-contamination in the lab thus
372 appears to have been negligible. Similar to visual observations, dietary DNA approaches revealed
373 the most prevalent prey categories to be hake and herring across all years.
374

375 ## 3.3 Comparison of visual observations and fecal metabarcoding

376 We recorded 12 fish prey groups (including eight within the “other identifiable fish” category) in
377 our visual observations compared with 19 in our fecal samples. This was largely due to the
378 greater taxonomic resolution afforded by our dietary DNA data. For example, we could
379 differentiate Atlantic herring from river herrings, and gadid species from one another, including
380 hakes, haddock, and Atlantic tomcod, all of which resemble one another as juveniles. Logistical
381 constraints on visual observations resulted in a large proportion of prey items categorized as
382 “unknown”, while there were no unidentifiable prey in the dietary DNA (Figure 1). As a result,

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

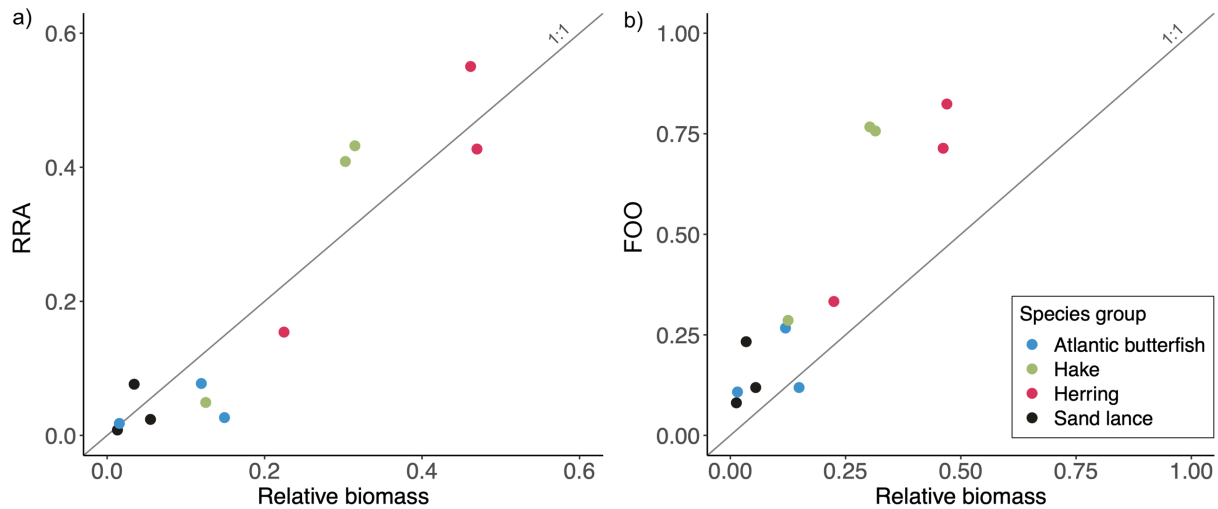
383 dietary DNA summaries estimated a higher proportion of prey categorized as "other" compared
384 to visual observations, reflecting the high taxonomic resolution of DNA approaches.
385
386 Pairwise correlations between each of the dietary metrics examined in this study are
387 summarized in Table 2. Summing across all years of the study, the highest correlation with
388 relative biomass was RRA (r = 0.94; Table 2, Figure 2a). In fact, the strength of this correlation
389 was far higher than the correlation between relative biomass and relative counts and similar to
390 the correlation between RRA and FOO (Table 2). FOO also showed a strong correlation to relative
391 biomass (r = 0.92; Table 2). However, FOO systematically over-estimated the relative importance
392 of nearly all prey categories in all years (Figure 2b) because FOO can range from 0 to 1 in each
393 prey category, while all other metrics (RRA, relative counts, and relative biomass) sum to 1
394 across all prey categories combined.
395
396
397 Table 2. Correlation coefficients and errors among four diet metrics characterizing the diets nestling common terns
398 in the wild. Relative counts and relative biomass come from visual observations, and relative read abundance (RRA)
399 and frequency of occurrence (FOO) come from dietary DNA.

	Correlation	MAE	RMSE
Relative biomass vs RRA	0.94	0.062	0.074
Relative biomass vs FOO	0.92	0.199	0.245
Relative biomass vs relative counts	0.66	0.104	0.133
RRA vs relative counts	0.92	0.087	0.139
RRA vs FOO	0.96	0.196	0.223
FOO vs relative counts	0.81	0.229	0.284

400
401

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

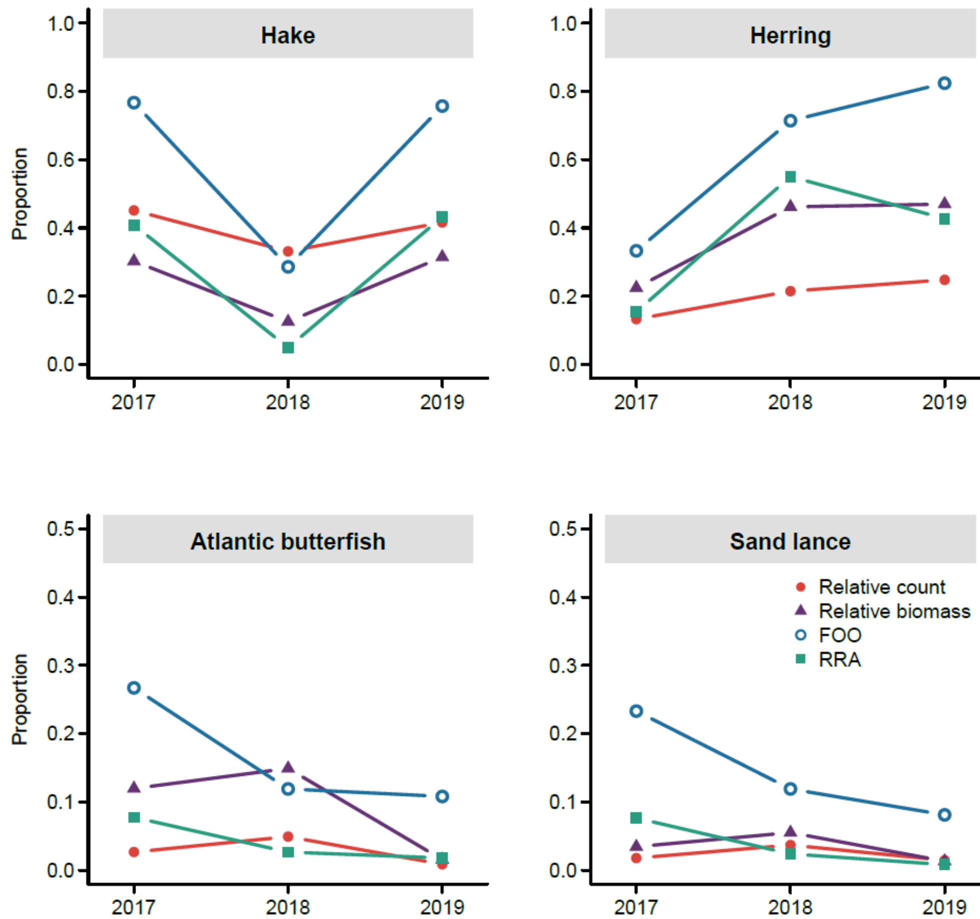
402



403
404
405 Figure 2. a) The high correlation between RRA and relative biomass (r = 0.94) falls close to a 1:1 relationship. b) FOO
406 also shows a high, positive correlation (r = 0.92), but tends to overestimate the relative importance of most prey
407 groups relative to their biomass.
408
409

410 MAE and RMSE measurements for each pair of metrics showed strong differences between
411 comparisons (Kruskal-Wallis χ^2 = 14.5, p = 0.002). MAE and RMSE error rates were lowest when
412 comparing RRA to relative biomass and highest when comparing FOO to relative counts and
413 relative biomass (Table 2). Notably, error rates from the RRA vs. relative biomass comparison
414 were even lower than comparisons based on the same data collection techniques, such as
415 relative biomass vs. relative counts (both summarized from visual diet observations) and RRA vs.
416 FOO. Post-hoc tests showed strong evidence that RRA provided a better match to relative
417 biomass (Dunn test Z = 2.5, p = 0.04) and relative counts (Dunn test Z = 2.9, p = 0.021) compared
418 to FOO. When examining year-to-year changes in diet metrics across four focal prey categories
419 (Figure 3), the interannual correlation between dietary DNA metrics and relative biomass was
420 strong for both RRA (r = 0.77) and FOO (r = 0.74). Interannual correlations were substantially
421 weaker when comparing relative counts to RRA (r = 0.14) and FOO (r = 0.11), highlighting the
422 usefulness of biomass-based measurements over relative counts for detecting and tracking shifts
423 in diets over time.
424
425

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

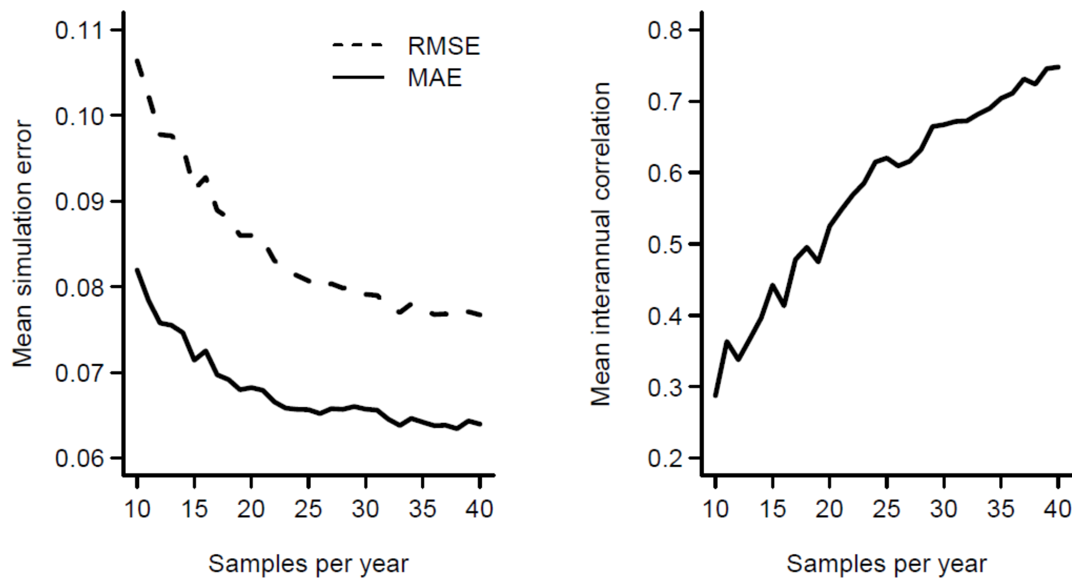


426
427 Figure 3. Interannual changes in diet metrics based on visual observations (relative counts, relative biomass) and
428 dietary DNA (RRA, FOO) show variation in temporal patterns of diet composition based on the method used to
429 characterize diet.
430
431

432 ### 3.4 The effect of sample size on proportional estimates of diet

433 By simulating fecal datasets with sampling designs ranging from 10–40 successfully sequenced
434 samples per year, we found that mean simulation error between RRA and relative biomass
435 (measured by both RMSE and MAE) approached a lower asymptote at 35–40 samples/year
436 (Figure 4). At a sample size of 40 samples/year, RRA and relative biomass had a mean RMSE of
437 0.08 (95% CI = 0.07, 0.09), and a mean MAE of 0.06 (95% CI = 0.05, 0.07). The interannual
438 correlation between simulated RRA and relative biomass continued to increase between 10–40
439 samples/year ranging from a low of 0.28 (95% CI = -0.43, 0.80) at 10 samples/year to a maximum
440 of 0.75 (95% CI = 0.72, 0.80) at 40 samples/year. This suggests that future studies should aim to
441 collect at least 40 samples per year to accurately estimate diet proportions in this study system.

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



442
443 Figure 4. The match between diet metrics summarized from dietary DNA (RRA) and visual observations (relative
444 biomass) increases with increasing fecal sample size. Error rates (mean absolute error [MAE] and root mean squared
445 error [RMSE]) reach a lower asymptote and interannual correlations remain high at 35–40 samples per year. Error
446 rates and correlations were generated from an iterative, simulated dietary DNA sampling scheme ranging from 10 to
447 40 samples per year compared against observed relative biomass.
448
449
450
451 ## 4. Discussion
452
453 We used dietary DNA and visual observations of feeding events in a wild seabird population to
454 determine to what extent dietary DNA could be used to estimate prey proportions. We found
455 that the relative read abundance (RRA) of fish taxa recovered from fecal samples accurately
456 captured the relative biomass consumed at the population-level over the three years of the
457 study, with the relationship between RRA and relative biomass falling close to 1:1. The frequency
458 of occurrence (FOO) of prey taxa, a commonly-used metric in dietary DNA studies, also showed a
459 high correlation with relative biomass but over-estimated the importance of most fish taxa. In
460 addition, we found that RRA slightly outperformed FOO in its ability to capture interannual
461 changes in diet, and RRA produced similar estimates to relative biomass from high-effort visual
462 observations. By simulating an annual fecal sampling program, we found that 35–40 successfully
463 analyzed samples per year was necessary to produce RRA estimates that matched biomass
464 estimates from visual observations. This range of sample effort provides a helpful benchmark for
465 future dietary DNA studies aiming to characterize seabird diets and track climate- or human-
466 induced change over time, although we caution that seabirds with more diverse diets (e.g., in the
467 tropics) may require larger sample sizes.
468
469 Our study adds to the body of evidence that supports the use of RRA in dietary DNA studies
470 (Deagle et al., 2019), particularly in avian systems (Verkuil et al., 2022). Deagle et al. (2019)
471 suggested that RRA is a semi-quantitative metric, but encouraged its use over presence/absence

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

472 metrics, such as FOO, because occurrence-based metrics can over-estimate the importance of
473 prey consumed in small amounts (and low-level contamination). In our study, the very high
474 correlation between RRA and relative biomass suggests that RRA is more than a semi-
475 quantitative metric; it can give quantitative estimates of prey proportions in piscivorous seabirds.
476 However, we caution that the low taxonomic resolution of visual observations in this study
477 limited the number of prey categories that we could compare against dietary DNA metrics. In
478 systems where the preyscape of seabirds is more diverse (or where prey can be compared at
479 higher taxonomic resolution than was possible here), we may expect to see a larger benefit to
480 using RRA over FOO due to the inflated impact of rare taxa on occurrence-based metrics. In
481 other systems with especially high diet diversity, such as insectivorous birds or herbivorous
482 mammals, validation studies and the use of mock communities are still advised.
483
484 Compared to traditional diet monitoring techniques, dietary DNA has some important
485 advantages, particularly if we can use it to estimate prey proportions reliably as exhibited in this
486 field test. Firstly, dietary DNA can provide greater taxonomic resolution than techniques that rely
487 on the morphological identification of prey through either visual observations or the
488 identification of hard parts, for example, from prey remains in fecal matter, regurgitates, or
489 stomach contents (Hoenig et al., 2022). Expertise is often required to morphologically
490 differentiate prey, and even then, species-level identifications can remain impossible.
491 Observational studies can also suffer from observer bias and errors (Farmer et al., 2012), which
492 are largely avoided with dietary DNA. In addition, a high number (almost one third) of our visual
493 observations could not be identified at all due to the size, speed, or angle at which the fish were
494 provisioned to chicks, whereas dietary DNA does not generally suffer from this limitation
495 provided reference databases are complete. Dietary DNA can also provide greater taxonomic
496 resolution than stable isotope-based methods, although, unlike stable isotope analyses, dietary
497 DNA only provides a snapshot of the diet, while stable isotopes can integrate over longer time
498 scales depending on the tissues analyzed (reviewed in Wiley et al., 2017). Dietary DNA also
499 provides a means to collect diet information from a broader range of individuals with limited
500 disturbance, for example from both adults and nestlings (Bowser et al., 2013; Fayet et al., 2021).
501 This ability to rapidly 'scale up' population-level diet monitoring opens doors to more effectively
502 track changes in prey availability over time and the potential impacts of these changes on
503 sensitive populations (Barrett et al., 2007).
504
505 Although dietary DNA can provide complementary information on diet composition, it cannot
506 estimate feeding rates, prey sizes, or prey quality. These metrics can have a significant effect on
507 breeding success in seabirds (Davoren and Montevecchi, 2003; Fayet et al., 2021; Lamb et al.,
508 2017; Schrimpf et al., 2012; Wanless et al., 2005) and often provide important information to
509 guide seabird conservation efforts. Therefore, where existing diet monitoring programs exist,
510 dietary DNA should be considered as a tool to increase the taxonomic resolution of the data
511 collected and provide efficient estimates of diet composition as a supplement to field-based
512 efforts. However, dietary DNA may prove especially useful when diet monitoring is constrained
513 due to logistical challenges, such as an inability to observe feeding events or support observers in
514 the field over extended periods of time. Since many seabird colonies are located on remote

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

515 islands far from field stations, the ability to collect dietary DNA data during relatively short visits
516 to colonies is a major advantage.
517
518 Overall, dietary DNA, when applied to seabirds, opens doors to more efficient, cost-effective
519 tracking of climate- and fisheries-induced impacts on marine ecosystems. Dietary DNA has
520 already been used to identify albatross populations with a heavy reliance on fishery discards,
521 putting them at greater risk of bycatch (McInnes et al., 2017), and to elucidate the diets of
522 endangered species such as the yellow-eyed penguin (Young et al., 2020). With the majority of
523 fisheries resources already over-fished (Pauly et al., 1998) and climate-induced shifts in forage
524 fish distributions (Kleisner et al., 2017; Pinsky et al., 2013), it is likely that seabird diets are
525 shifting at unprecedented rates, yet we have not, until recently, had the tools to monitor this
526 change in more than a few species. In some systems, estimates of forage fish availability may be
527 more readily obtained from seabird diet data than through traditional net-based surveys (e.g.
528 Scopel et al., 2017) and could prove especially valuable for numerous species and regions where
529 data are limited or nonexistent (Cairns, 1988). Increasing the availability and accuracy of seabird
530 diet data is timely, as fisheries management moves towards ecosystem approaches that aim to
531 increase sustainability. By providing high-resolution estimates of prey proportions in seabird
532 diets, we may provide the data necessary to parameterize food web models for ecosystem-
533 based fisheries management (Hall and Mainprize, 2004; Pikitch et al., 2004) and track changes in
534 forage fish availability on an annual basis using seabirds as indicators (Scopel et al., 2017; Velarde
535 et al., 2013). Moving forward, the ability to monitor real-time changes in seabird diets through
536 dietary DNA, combined with efforts to scale-up diet monitoring across populations, represents a
537 key pathway to advance our mechanistic understanding of seabird declines and effectively
538 address the challenges posed by fisheries management and global change.
539

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

540 ## Acknowledgements
541
542 We would like to thank all the field and lab technicians that helped collect data for this project,
543 including Elizabeth Ford, Taylor Ouelette, Caitlin Bowman, Amber Litterer, Aliya Caldwell, Olivia
544 Smith, and Casey Coupe. Special thanks to Bronwyn Butcher, Irby Lovette, Jennifer Seavey,
545 Adrienne Kovach, Brian Trevelline, and Jeffrey Hall for support and advice throughout. This study
546 was supported by a seed grant from the University of New Hampshire School of Marine Science
547 and Ocean Engineering. GC and AS were supported by postdoctoral fellowships from the Cornell
548 Atkinson Center for Sustainability and Cornell Lab of Ornithology Rose Postdoctoral Program. The
549 Isles of Shoals Tern Conservation Project is supported by the New Hampshire Fish and Game
550 Department's Nongame and Endangered Wildlife Program, the United States Fish and Wildlife
551 Service (New Hampshire State Wildlife Grants), and the Shoals Marine Laboratory. We thank
552 Michael Marchand and the New Hampshire Fish and Game Department, Nongame and
553 Endangered Wildlife Program for leadership in New Hampshire tern conservation efforts. The
554 handling of live birds was permitted under protocol #160403 approved by University of New
555 Hampshire's Institutional Animal Care and Use Committee and permit 06543 under the U.S.
556 Geological Survey's Bird Banding Laboratory. This paper is contribution 201 to the Shoals Marine
557 Laboratory.
558

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

559 ## 5. Literature Cited
560

561 Ainley, D.G., Spear, L.B., Allen, S.G., 1996. Variation in the diet of Cassin's auklet reveals spatial,
562 seasonal, and decadal occurrence patterns of euphausiids off California, USA. Mar. Ecol.
563 Prog. Ser. 137, 1–10. https://doi.org/10.3354/meps137001
564 Alberdi, A., Aizpurua, O., Bohmann, K., Gopalakrishnan, S., Lynggaard, C., Nielsen, M., Gilbert,
565 M.T.P., 2019. Promises and pitfalls of using high-throughput sequencing for diet analysis.
566 Mol. Ecol. Resour. 19, 327–348. https://doi.org/10.1111/1755-0998.12960
567 Anderson, O.R.J., Small, C.J., Croxall, J.P., Dunn, E.K., Sullivan, B.J., Yates, O., Black, A., 2011.
568 Global seabird bycatch in longline fisheries. Endanger. Species Res. 14, 91–106.
569 https://doi.org/10.3354/esr00347
570 Arnold, J.M., Oswald, S.A., Nisbet, I.C.T., Pyle, P., Patten, M.A., 2020. Common Tern (Sterna
571 hirundo), version 1.0., in: Billerman, S.M. (Ed.), Birds of the World. Ithaca, NY, USA.
572 https://doi.org/https://doi.org/10.2173/bow.comter.01
573 Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C.,
574 McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B., Ferrer, E.A., 2011. Has the Earth's
575 sixth mass extinction already arrived? Nature 471, 51–57.
576 https://doi.org/10.1038/nature09678
577 Barrett, R.T., Camphuysen, K., Anker-Nilssen, T., Chardine, J.W., Furness, R.W., Garthe, S.,
578 Hüppop, O., Leopold, M.F., Montevecchi, W.A., Veit, R.R., 2007. Diet studies of seabirds: A
579 review and recommendations. ICES J. Mar. Sci. 64, 1675–1691.
580 https://doi.org/10.1093/icesjms/fsm152
581 Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H.,
582 Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn,
583 C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R.,
584 Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst,
585 M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick,
586 K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S.,
587 Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K. Bin, Keefe, C.R., Keim, P., Kelley, S.T.,
588 Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.X.,
589 Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J.,
590 Melnik, A. V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A.,
591 Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E.,
592 Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A.,
593 Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P.,
594 Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y.,
595 Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C.,
596 Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R.,
597 Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data
598 science using QIIME 2. Nat. Biotechnol. 37, 852–857. https://doi.org/10.1038/s41587-019-
599 0209-9
600 Bowser, A.K., Diamond, A.W., Addison, J.A., 2013. From puffins to plankton: A DNA-based
601 analysis of a seabird food chain in the Northern Gulf of Maine. PLoS One 8, 1–16.
602 https://doi.org/10.1371/journal.pone.0083152

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

603 Cairns, D.K., 1988. Seabirds as Indicators of Marine Food Supplies. Biol. Oceanogr. 5, 261–271.
604 https://doi.org/10.1080/01965581.1987.10749517
605 Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016.
606 DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods 13,
607 581–583. https://doi.org/10.1038/nmeth.3869
608 Ceballos, G., Ehrlich, P.R., Dirzo, R., 2017. Biological annihilation via the ongoing sixth mass
609 extinction signaled by vertebrate population losses and declines. Proc. Natl. Acad. Sci. U. S.
610 A. 114, E6089–E6096. https://doi.org/10.1073/pnas.1704949114
611 Croxall, J.P., Butchart, S.H.M., Lascelles, B., Stattersfield, A.J., Sullivan, B., Symes, A., Taylor, P.,
612 2012. Seabird conservation status, threats and priority actions: A global assessment. Bird
613 Conserv. Int. 22, 1–34. https://doi.org/10.1017/S0959270912000020
614 Davoren, G., Montevecchi, W., 2003. Signals from seabirds indicate changing biology of capelin
615 stocks. Mar. Ecol. Prog. Ser. 258, 253–261. https://doi.org/10.3354/meps258253
616 de Sousa, L.L., Silva, S.M., Xavier, R., 2019. DNA metabarcoding in diet studies: Unveiling
617 ecological aspects in aquatic and terrestrial ecosystems. Environ. DNA 1, 199–214.
618 https://doi.org/10.1002/edn3.27
619 Deagle, B.E., Thomas, A.C., McInnes, J.C., Clarke, L.J., Vesterinen, E.J., Clare, E.L., Kartzinel, T.R.,
620 Eveson, J.P., 2018. Counting with DNA in metabarcoding studies: How should we convert
621 sequence reads to dietary data? Mol. Ecol. https://doi.org/10.1111/mec.14734
622 Diamond, A.W., Devlin, C.M., 2003. Seabirds as indicators of changes in marine ecosystems:
623 ecological monitoring on Machias Seal Island. Environ. Monit. Assess. 88, 153–175.
624 Dias, M.P., Martin, R., Pearmain, E.J., Burfield, I.J., Small, C., Phillips, R.A., Yates, O., Lascelles, B.,
625 Borboroglu, P.G., Croxall, J.P., 2019. Threats to seabirds: A global assessment. Biol. Conserv.
626 237, 525–537. https://doi.org/10.1016/j.biocon.2019.06.033
627 Dirzo, R., Young, H.S., Galetti, M., Ceballos, G., Isaac, N.J.B., Collen, B., 2014. Defaunation in the
628 Anthropocene. Science (80-.). 345, 401–406. https://doi.org/10.1126/science.1251817
629 Durant, J.M., Anker-Nilssen, T., Stenseth, N.C., 2003. Trophic interactions under climate
630 fluctuations: The Atlantic puffin as an example. Proc. R. Soc. B Biol. Sci. 270, 1461–1466.
631 https://doi.org/10.1098/rspb.2003.2397
632 Farmer, R.G., Leonard, M.L., Horn, A.G., 2012. Observer effects and avian-call-count survey
633 quality: Rare-species biases and overconfidence. Auk 129, 76–86.
634 https://doi.org/10.1525/auk.2012.11129
635 Fayet, A.L., Clucas, G. V., Anker-Nilssen, T., Syposz, M., Hansen, E.S., 2021. Local prey shortages
636 drive foraging costs and breeding success in a declining seabird, the Atlantic puffin. J. Anim.
637 Ecol. 90, 1152–1164. https://doi.org/10.1111/1365-2656.13442
638 Furness, R.W., Camphuysen, K., 1997. Seabirds as monitors of the marine environment, in: ICES
639 Journal of Marine Science. pp. 726–737. https://doi.org/10.1006/jmsc.1997.0243
640 Hall, S.J., Mainprize, B., 2004. Towards ecosystem-based fisheries management. Fish Fish. 5, 1–
641 20.
642 Hoenig, B.D., Snider, A.M., Forsman, A.M., Hobson, K.A., Latta, S.C., Miller, E.T., Polito, M.J.,
643 Powell, L.L., Rogers, S.L., Sherry, T.W., Toews, D.P.L., Welch, A.J., Taylor, S.S., Porter, B.A.,
644 2022. Current methods and future directions in avian diet analysis. Ornithology 139, 1–28.
645 https://doi.org/10.1093/ornithology/ukab077
646 Kleisner, K.M., Fogarty, M.J., McGee, S., Hare, J.A., Moret, S., Perretti, C.T., Saba, V.S., 2017.

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

647 Marine species distribution shifts on the U.S. Northeast Continental Shelf under continued
648 ocean warming. Prog. Oceanogr. 153, 24–36. https://doi.org/10.1016/j.pocean.2017.04.001
649 Kress, S.W., Shannon, P., O'neal, C., 2016. Recent changes in the diet and survival of Atlantic
650 puffin chicks in the face of climate change and commercial fishing in midcoast Maine, USA.
651 https://doi.org/10.1139/facets-2015-0009
652 Lamb, J.S., Satgé, Y.G., Jodice, P.G.R., 2017. Diet composition and provisioning rates of nestlings
653 determine reproductive success in a subtropical seabird. Mar. Ecol. Prog. Ser. 581, 149–164.
654 https://doi.org/10.3354/meps12301
655 Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing
656 reads. EMBnet. J. 17, 10–12.
657 McInnes, J.C., Jarman, S.N., Lea, M.-A., Raymond, B., Deagle, B.E., Phillips, R.A., Catry, P.,
658 Stanworth, A., Weimerskirch, H., Kusch, A., Gras, M., Cherel, Y., Maschette, D., Alderman,
659 R., 2017. DNA Metabarcoding as a Marine Conservation and Management Tool: A
660 Circumpolar Examination of Fishery Discards in the Diet of Threatened Albatrosses. Front.
661 Mar. Sci. 4, 277. https://doi.org/10.3389/fmars.2017.00277
662 Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J.Y., Sato, K., Minamoto, T., Yamamoto, S.,
663 Yamanaka, H., Araki, H., Kondoh, M., Iwasaki, W., 2015. MiFish, a set of universal PCR
664 primers for metabarcoding environmental DNA from fishes: Detection of more than 230
665 subtropical marine species. R. Soc. Open Sci. 2, 150088.
666 https://doi.org/10.1098/rsos.150088
667 Montevecchi, W.A., 1993. Birds as indicators of change in marine prey stocks, in: Furness, R.W.,
668 Greenwood, J.J. (Eds.), Birds as Monitors of Environmental Change. Chapman & Hall,
669 London, pp. 217–266. https://doi.org/10.1007/978-94-015-1322-7_6
670 Montevecchi, W.A., Myers, R.A., 1995. Prey harvests oí seabirds reílect pelagic fish and squid
671 abundance on multiple spatial and temporal scales. Mar. Ecol. Prog. Ser., 117, 1–9.
672 https://doi.org/10.2307/44634813
673 Ogle, D.H., Doll, J.C., Wheeler, A.P., Dinno, A., 2023. FSA: Simple Fisheries Stock Assessment
674 Methods. R Packag. version 0.9.4.
675 Paleczny, M., Hammill, E., Karpouzi, V., Pauly, D., 2015. Population trend of the world's
676 monitored seabirds, 1950-2010. PLoS One 10, 1–11.
677 https://doi.org/10.1371/journal.pone.0129342
678 Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., Torres, F., 1998. Fishing Down Marine Food
679 Webs. Science (80-.). 279, 860–863.
680 Phillips, R.A., Fort, J., Dias, M.P., 2022. Conservation status and overview of threats to seabirds,
681 Conservation of Marine Birds. Elsevier Inc. https://doi.org/10.1016/B978-0-323-88539-
682 3.00015-7
683 Piatt, J.F., Sydeman, W.J., Wiese, F., 2007. Introduction: A modern role for seabirds as indicators.
684 Mar. Ecol. Prog. Ser. 352, 199–204. https://doi.org/10.3354/meps07070
685 Pikitch, E.K., Santora, C., Babcock, E.A., Bakun, A., Bonfil, R., Conover, D.O., Dayton, P., Doukakis,
686 P., Fluharty, D., Heneman, B., Houde, E.D., Link, J., Livingston, P.A., Mangel, M., McAllister,
687 M.K., Pope, J., Saisnbury, K.J., 2004. Ecosystem-based fishery management. Science (80-.).
688 305, 346–347. https://doi.org/10.1126/science.294.5548.1920
689 Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts,
690 C.M., Sexton, J.O., 2014. The biodiversity of species and their rates of extinction,

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

691 distribution, and protection. Science (80-.). 344. https://doi.org/10.1126/science.1246752
692 Pinsky, M.L., Worm, B., Fogarty, M.J., Sarmiento, J.L., Levin, S.A., 2013. Marine taxa track local
693 climate velocities. Science (80-.). 341, 1239–1242.
694 Pompanon, F., Deagle, B.E., Symondson, W.O.C., Brown, D.S., Jarman, S.N., Taberlet, P., 2012.
695 Who is eating what: Diet assessment using next generation sequencing. Mol. Ecol. 21,
696 1931–1950. https://doi.org/10.1111/j.1365-294X.2011.05403.x
697 R Core Team, 2022. R: A language and environment for statistical computing.
698 Robeson, M.S., O'Rourke, D.R., Kaehler, B.D., Ziemski, M., Dillon, M.R., Foster, J.T., Bokulich, N.A.,
699 2021. RESCRIPt: Reproducible sequence taxonomy reference database management, PLoS
700 Computational Biology. https://doi.org/10.1371/journal.pcbi.1009581
701 Rosenberg, K. V., Dokter, A.M., Blancher, P.J., Sauer, J.R., Smith, A.C., Smith, P.A., Stanton, J.C.,
702 Panjabi, A., Helft, L., Parr, M., Marra, P.P., 2019. Decline of the North American avifauna.
703 Science (80-.). eaaw1313. https://doi.org/10.1126/SCIENCE.AAW1313
704 Schneider, J.C., Alexander, G.R., Merna, J., 2000. Modules for Lake and Stream Surveys, in:
705 Schneider, J.C. (Ed.), Manual of Fisheries Survey Methods II: With Periodic Updates.
706 Michigan Department of Natural Resources, Fisheries Special Report 25, Ann Arbor.
707 Schrimpf, M.B., Parrish, J.K., Pearson, S.F., 2012. Trade-offs in prey quality and quantity revealed
708 through the behavioral compensation of breeding seabirds. Mar. Ecol. Prog. Ser. 460, 247–
709 259. https://doi.org/10.3354/meps09750
710 Scopel, L., Diamond, A., Kress, S., Shannon, P., 2019. Varied breeding responses of seabirds to a
711 regime shift in prey base in the Gulf of Maine. Mar. Ecol. Prog. Ser. 626, 177–196.
712 https://doi.org/10.3354/meps13048
713 Scopel, L.C., Diamond, A.W., Kress, S.W., Hards, A.R., Shannon, P., 2017. Seabird diets as
714 bioindicators of Atlantic herring recruitment and stock size: a new tool for ecosystem-based
715 fisheries management. Can. J. Fish. Aquat. Sci. 15, 1–15. https://doi.org/10.1139/cjfas-
716 2017-0140
717 Spiller, K.J., Dettmers, R., 2019. Evidence for multiple drivers of aerial insectivore declines in
718 North America. Condor 121, 1–13. https://doi.org/10.1093/condor/duz010
719 Thomas, A.C., Jarman, S.N., Haman, K.H., Trites, A.W., Deagle, B.E., 2014. Improving accuracy of
720 DNA diet estimates using food tissue control materials and an evaluation of proxies for
721 digestion bias, in: Molecular Ecology. pp. 3706–3718. https://doi.org/10.1111/mec.12523
722 Valentini, A., Pompanon, F., Taberlet, P., 2009. DNA barcoding for ecologists. Trends Ecol. Evol.
723 https://doi.org/10.1016/j.tree.2008.09.011
724 Velarde, E., Anderson, D.W., Ezcurra, E., 2019. Seabird clues to ecosystem health. Science (80-.).
725 365, 116–117.
726 Velarde, E., Ezcurra, E., Anderson, D.W., 2013. Seabird diets provide early warning of sardine
727 fishery declines in the Gulf of California. Sci. Rep. 3, 1–6. https://doi.org/10.1038/srep01332
728 Velarde, E., Tordesillas, M.S., Vieyra, L., Esquivel, R., 1994. Seabirds as indicators of important
729 fish populations in the Gulf of California. Calif. Coop. Ocean. Fish. Investig. Prog. Rep. 35,
730 137–143.
731 Verkuil, Y.I., Nicolaus, M., Ubels, R., Dietz, M.M., Annabet, S., Kim, G., Peter, K., Christiaan, D.K.,
732 2020. DNA metabarcoding successfully quantifies relative abundances of arthropod taxa in
733 songbird diets : a validation study using camera-recorded diets 1–27.
734 Verkuil, Y.I., Nicolaus, M., Ubels, R., Dietz, M.W., Samplonius, J.M., Galema, A., Kiekebos, K., de

bioRxiv preprint doi: https://doi.org/10.1101/2024.03.22.586275; this version posted March 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

735 Knijff, P., Both, C., 2022. DNA metabarcoding quantifies the relative biomass of arthropod
736 taxa in songbird diets: Validation with camera-recorded diets. Ecol. Evol. 12, 1–17.
737 https://doi.org/10.1002/ece3.8881
738 Wanless, S., Harris, M.P., Redman, P., Speakman, J.R., 2005. Low energy values of fish as a
739 probable cause of a major seabird breeding failure in the North Sea. Mar. Ecol. Prog. Ser.
740 294, 1–8. https://doi.org/10.3354/meps294001
741 Wigley, S.E., McBride, H.M., McHugh, N.J., 2003. Length-weight relationships for 74 fish species
742 collected during NEFSC research vessel bottom trawl surveys, 1992-1999. Natl. Mar. Fish.
743 Serv. Woods Hole, MA, USA.
744 Wiley, A.E., James, H.F., Ostrom, P.H., 2017. Emerging techniques for isotope studies of avian
745 ecology., in: Webster, M.S. (Ed.), The Extended Specimen: Emerging Frontiers in Collections-
746 Based Ornithological Research. CRC Press: Boca Raton, FL, pp. 89–109.
747 Winters, G.H., 1989. Life history parameters of sand lances (Ammodytes spp.) from the coastal
748 waters of eastern Newfoundland. J. Northwest Atl. Fish. Sci. 9, 5–11.
749 Young, M.J., Dutoit, L., Robertson, F., Van Heezik, Y.M., Seddon, P.J., Robertson, B.C., 2020.
750 Species in the faeces: DNA metabarcoding as a method to determine the diet of the
751 endangered yellow-eyed penguin. Wildl. Res. 47, 509–522.
752