Polar Biology (2018) 41:2275–2287
https://doi.org/10.1007/s00300-018-2364-8

ORIGINAL PAPER

# Conventional and molecular analysis of the diet of gentoo penguins: contributions to assess scats for non-invasive penguin diet monitoring

José C. Xavier[1,2] · Yves Cherel[3] · Renata Medeiros[4] · Nadja Velez[1] · Meagan Dewar[5,6] · Norman Ratcliffe[2] · Ana R. Carreiro[1] · Phil N. Trathan[2]

Received: 15 January 2018 / Revised: 9 May 2018 / Accepted: 14 June 2018 / Published online: 27 June 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

## Abstract

There is a growing search for less invasive methods while studying the diet of Antarctic animals in the wild. Therefore, we compared the diet of gentoo penguins from stomach contents (i.e. through visual identification of prey remains) and scats (i.e. faeces), and further compared prey DNA assay in fresh and old scats. Prey remains identified visually in stomach contents and scats were broadly comparable: the crustaceans and fish were the most important components, with *Themisto gaudichaudii* clearly being the most frequent and numerous prey species in both sampling methods. By mass, differences in species frequency were observed in stomach contents (*Parachaenichthys georgianus*) and scats (*Champsocephalus gunnari*), with the former fish species absent in scats. Differences were detected in the most frequent prey (*T. gaudichaudii* and *Euphausia superba*) and in various fish species, most with bigger sizes in scats. Allometric equations to estimate most crustacean's sizes (i.e. relationships between carapace and mass/total length) are needed. For DNA studies, when comparing DNA from fresh and old scats, both provided similar results that, in general, were also similar to the visual analysis. In order to use penguin scats (along with the use of DNA analyses) for monitoring purposes, allometric equations to estimate mass and size of prey (most crustaceans) and better designed species-specific primers are needed for targeting key prey species (e.g. *Euphausia superba, T. gaudichaudii*). These DNA methodologies can complement other methods (i.e. visual analyses and stomach contents analyses) in monitoring programs of penguins.

**Keywords** *Pygoscelis papua* · Southern Ocean · Prey genetics · Feeding ecology · Conservation

## Introduction

Investigating the diet of an organism is of primary importance for understanding its ecological requirements and its functional role in the ecosystems that it inhabits (Croxall 1987; Pauly et al. 1998; Boyd 2002; Constable 2002). The study of the diet of top predators, such as seabirds, has relied on traditional invasive methods developed over 40 years ago, such as stomach flushing (which involves pumping water through a tube inserted in the oesophagus of a bird and catching the regurgitated contents in, for example, a bag, sieve, or bucket) (Wilson 1984; Duffy and Jackson 1986; Croxall 1993; Barrett et al. 2007; Ratcliffe and Trathan 2011; Karnovsky et al. 2012), because it allowed scientists to determine and quantify both prey species and prey size very well despite their limitations (Clarke and Kerry 1994; Barrett et al. 2007). Understandably, the search for less invasive methods to minimise impact on the populations studied is increasing as animal ethics approval for invasive approaches

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s00300-018-2364-8) contains supplementary material, which is available to authorised users.

✉ José C. Xavier
JCCX@cantab.net

[1] Department of Life Sciences, Institute of Marine Research, University of Coimbra, 3001-401 Coimbra, Portugal

[2] British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK

[3] Centre d´Etudes Biologiques de Chizé, UMR 7372 du CNRS-Université de La Rochelle, 79360 Villiers-En-Bois, France

[4] Cardiff School of Biosciences, Cardiff University, The Sir Martin Evans Building, Museum Av, Cardiff CF10 3AX, UK

[5] School of Applied and Biomedical Sciences, Federation University Australia, Berwick, Australia

[6] School of Life and Environmental Sciences, Deakin University, Burwood, Australia

is increasingly difficult to obtain (Duffy and Jackson 1986; Barrett et al. 2007; Karnovsky et al. 2012). Scats contain remains of prey, are readily available in seabird colonies, can be easily collected and causes minimal disturbance to birds and so may provide an alternative means of quantifying diet. Indeed, hard parts of ingested prey have been recovered from seabird scats (Lumsden and Fladdow 1946; Duffy and Jackson 1986), and used in recent studies (Pompanon et al. 2012; Bowser et al. 2013; McInnes et al. 2016a).

The visual identification of prey remains from scats or regurgitated pellets is less invasive and has the added advantage of providing information on actual prey digested by the adult (as opposed to prey stored to feed the young) but a serious limitation arises from biased recovery of the remains due to differential digestion (particularly for small prey) and the difficulties of identifying well-digested prey (Xavier et al. 2005; Seefelt and Gillingham 2006; Tollit et al. 2007). Stable isotope analyses and fatty acids analyses have, to some degree can be used to compare diet, specifically what is integrated into adult tissue versus what is passed on to chicks (Raclot et al. 1998; Cherel et al. 2008). More recently, molecular analysis of prey DNA in the guts, regurgitations and scats of a wide range of foragers have provided an opportunity to help fill this gap (Symondson 2002; King et al. 2008; Peters et al. 2015; McInnes et al. 2016b; Olmos-Pérez et al. 2017). Prey DNA can be identified from even well-digested remains in dietary samples that have no visible identification features. These techniques are additional complementary dietary tools [and provide broadly comparable and reliable data (e.g. presence/absence) (Deagle et al. 2010)] that have improved the study of multiple trophic links and have been tested in a wide range of organisms, including penguins (Blankenship and Yayanos 2005; Deagle et al. 2010; Murray et al. 2011). However, it also has limitations such as secondary prey, lack of primers to identify prey to species/genus level and inability to accurately quantify prey by number and mass (Barrett et al. 2007; Karnovsky et al. 2012). Nonetheless, molecular techniques remain a valuable tool for examining diets of marine predators (Symondson 2002; Karnovsky et al. 2012; Jarman et al. 2013). In this study, we aim to compare diets inferred from stomach flushing with those inferred from visual identification and DNA detection of prey remains in scats.

Penguins (Spheniscidae) are a key group of the Antarctic marine ecosystem and are important consumers of marine resources in the Southern Ocean (Croxall and Prince 1987; Brooke 2004). Information on the diet and foraging ecology of penguins is vital to parameterise consumption models that are relevant to organisations, such as the Commission for the Conservation of Antarctic Marine Life Resources (CCAMLR). The CCAMLR data collection includes the use of stomach flushing during the Austral summer, the record of wet weight of the whole sample after draining and of each category (i.e. squid, fish and crustaceans (and of the species *Euphausia superba* and *E. crystallorophias* within the crustacean component) (Agnew 1997), with data submitted voluntarily by CCAMLR countries when possible.

Within the penguins, there are species that are known to have high reliance on just a few dietary items during the summer (e.g. Chinstrap penguins *Pygoscelis adeliae*) (Lynnes et al. 2004) and those that have more generalised diets (e.g. Gentoo penguins *Pygoscelis papua*) (Waluda et al. 2017). We selected IUCN Least Concern species of gentoo penguins as our model for several reasons. Gentoo penguins are one of the most important animal species used for ecosystem monitoring in the Southern Ocean (Agnew 1997; Everson 2002) due to their broad geographical range (covering various areas across the Antarctic) and very short foraging ranges from land throughout the year (and consequently sampling prey at a relatively fine scale) (Tanton et al. 2004; IUCN 2010; Wilson 2010); data can be collected during the breeding and non-breeding periods all year round.

The aims of this study were to (1) compare diet composition of gentoo penguins as inferred from visual identification of prey remains in stomach contents and scats, (2) compare diets from scats (identified visually) with those inferred from DNA analysis of old and fresh scats and (3) critically evaluate the potential for the different methods to provide a suitable method for investigating gentoo penguin diet.

## Materials and methods

### Comparison between stomach contents and scats (general diet)

#### Stomach contents

Stomach contents were collected throughout the post-breeding period (May–October) in 2009 (stomach samples taken at the middle of each month) from a gentoo penguin colony (Landing beach, Bird Island) at South Georgia. The stomach contents were obtained from non-breeding penguins selected randomly when arriving at the beach to roost (late afternoon, early night). The stomach contents were obtained by stomach flushing, following Xavier et al. (2004) and Xavier et al. (2017); a maximum of three flushes were used, following the CCAMLR Ecosystem Monitoring Program (CEMP) Standard Methods. The British Antarctic Survey (BAS) provided all the support related to the permits for the fieldwork (as fieldwork was carried out from a British Antarctic research base): "The animal procedures used in this study were reviewed and approved by the Joint BAS-Cambridge University Animal Welfare and Ethical Review Committee. Permits to operate were issued by the Government of South Georgia and the South Sandwich Islands". If the first flush provided

only liquid of a green or yellow colour, we assumed their stomachs held no food and so the animal was immediately released. All penguins handled were marked with organic dye to ensure they were only sampled once. The procedure lasted, on average, 15 min. Analyses of the constituent food samples were carried out at the Bird Island research station laboratory within 24 h of collection. Each food sample was analysed by frequency of occurrence (FO; number of stomach samples with a certain taxa present divided by the total number of stomach samples analyses), number (*N*; number of individuals in stomach samples with a certain taxa present divided by the total number of individuals in the stomach samples analyses), estimated mass (*M*; estimated mass of individuals, through available allometric equations, in stomach samples with a certain taxa present divided by the total estimated mass of all individuals in the stomach samples analyses, so results were comparable with those from scats) and size of prey. Identification guides and collections at the British Antarctic Survey (BAS), and allometric equations, were used to identify the crustaceans using their morphology or bony structures (such as carapaces) (Bellan-Santini and Ledoyer 1974; Chekunova and Rynkova 1974; Rakusa-Suszczewski and Stepnik 1980; Kirkwood 1984; Siegel and Mühlenhardt-Siegel 1988; Siegel 1993; Pakhomov and Perissinotto 1996; Boltovskoy 1999), fish (using their otoliths) (Hecht 1987; Williams and McEldowney 1990; Smale et al. 1995; Reid 1996) and cephalopods (using their lower beaks) (Xavier and Cherel 2009). As the otoliths of ?*Gymnoscopelus braueri* were very small and there was a level of uncertainty the authors agreed to mention it in the text as "?*Gymnoscopelus braueri*".

### Scats collection

For comparison with the stomach contents, 15 fresh scat samples were collected from gentoo penguins every 2 weeks from May to October 2009 (i.e. the same time period as the stomach contents collection). These scat samples were collected randomly from ice or rock substrates immediately after defecation. Penguins were not handled and sex was not determined (which may affect the robustness of the results obtained, as gentoo penguins are known to exhibit sexual differences in diets Xavier et al. 2017). The collection of each group of scat samples took generally less than 1 h, either from ice covered ground or rock. Each scat sample was analysed for occurrence of prey types and sizes of those prey, following Xavier et al. (2003).

During October, 3 sets of scats were collected. Set 1 was collected fresh and used for visual identification of prey. Set 2 was collected fresh and used for DNA analyses (see below). Sets 1 and 2 of scats were collected on the same day. Set 3 was collected 1–2 days after defecation (the colony was checked prior to scats collection, to avoid collecting older scats) and also used for DNA analyses (see below). A total of 46 samples were collected—31 fresh samples (16 for visual analysis and 15 for DNA analysis) and 15 old samples for DNA analyses. All these samples for DNA analyses were stored in 70% ethanol for later analyses.

## Molecular analysis

### DNA extraction and amplification

DNA from gentoo penguin scats was extracted using the QIAamp DNA Stool Mini Kit (Qiagen), following the manufacturer's standard protocol. Prior to extraction, the samples were homogenised by mixing for at least 30 min after which a sub-sample (with a volume of approximately 0.6 ml) was used for extraction. Fresh and old samples were extracted separately and two blank extractions were included for each batch of extractions to test for any cross-over contamination.

In order to test for the success of the extraction process, DNA extracts were screened using general primers for Bilateria species (Table 1) since these primers amplify predator DNA that should be present in all the samples. Any samples giving a negative result were tested three times, to confirm that they were indeed a failed extraction. Successful extracts

**Table 1** Molecular analyses of the diet of gentoo penguins

| Target (Taxon, gene) | Primer name | Sequence 5′-3′ | Product size | Annealing temp. | References |
|---|---|---|---|---|---|
| Bilateria | BilSSU1100_F | AGAGGTGAAATTSTTGGAYCG | ~ 245 | 62° | Jarman et al. (2004) |
| Nuclear 18S | BilSSU1300_R | CCTTTAAGTTTCAGCTTTGCA | | | |
| Cephalopoda | Squid28SF | CGCCGAATCCCGTCGCMAGTAAAMGGCTTC | ~ 180 | 60° | Deagle et al. (2005) |
| Nuclear 28S | Squid 28SR | CCAAGCAACCCGACTCTCGGATCGAA | | | |
| Osteichthyes | | CGGTAAAACTCGTGCC | ~ 300 | 56° | Unpublished data |
| 12S | | CCGCCAAGTCCTTTGGG | | | |
| Euphausiidae | EuphMLSUF | TTTATTGGGGCGATAAAAAT | ~ 169 | 54° | Deagle et al. (2007) |
| 16S | EuphMLSUR | TCGAGGTCGYAATCTTTCTTGT | | | |

were screened for Osteichthyes (bony fish), Cephalopoda (cephalopods) and Euphausiidae (krill) specific primers (Table 1) in order to investigate the proportion of bird diets testing positive for these prey types. All the primers had previously been reported in the literature (Table 1), including from previous studies in penguin species, and were the most relevant primers available at the time the lab work was conducted (March 2013). We tested all the primers on blood DNA extracts from gentoo penguins to ensure that predator DNA was not amplified together with the prey. Amphipoda primers (Jarman et al. 2006) were also initially considered but specificity tests showed strong amplification of penguin DNA at the recommended annealing temperature (51 °C) and at temperatures up to 55 °C. At 56 °C, penguin DNA was no longer amplified but the range of amphipods amplified (fresh samples from six different species) also declined. Hence these primers were not considered for analysis.

Amplifications were performed separately for each primer pair, using the Multiplex PCR Kit (Qiagen) in 20 µl reactions containing 1× Multiplex PCR Master Mix, 0.2 µM of each primer and 0.1 mg/ml of BSA (New England Biolabs). The template was 2 µl of the DNA extract. Thermal cycling conditions were as follows: 95 °C for 15 min, 35 cycles (94 °C for 30 s followed by the primer specific annealing temperature for 90 s followed by 72 °C for 90 s), concluding with 72 °C for 10 min. A minimum of three negative controls (the extraction control, plus at least two distilled water blanks) were included in each set of PCR amplifications. Initial PCR reactions were performed using non-modified primers, followed by PCR reactions with modified primers for pyrosequencing (see below). PCR products were separated by electrophoresis in 1.5% agarose gels and visualised by staining with Ethidium bromide and visualised by transillumination with UV light.

## Preparation of DNA libraries for next-generation sequencing (NGS)

Three different NGS libraries were made for each set of faecal samples (old and fresh): (1) a general prey library using Bilateria primers (i.e. general primers that amplify DNA from organisms with bilateral symmetry) for all the successful DNA extractions; (2) a fish library using Osteichthyes primers for the subsets of the samples which tested positive for Osteichthyes DNA; (3) an Antarctic krill library using Euphausiidae primers also on the subsets of sample which tested positive for Euphausiidae DNA. A Cephalopoda library was not prepared due to very small sample size of positives. General Bilateria primers were expected to have a low resolution in terms of prey identification (phylum or class Jarman et al. 2004) and were used mainly to confirm the major groups detected with the other primers and identify the presence of potential prey types (any taxa besides molluscs, crustaceans or bony fish) that were not specifically screened for in this study.

For each primer pair and sample type (old and fresh), the concentration of each individual sample was measured from the gel using a reference ladder and pooled at equimolar concentrations so that the contribution from each individual bird was similar. The DNA concentration of each pool was measured using Qubit and these were also pooled according to their concentration so that each pool contained a weighted contribution from each sample type and primer pair based on the number of individuals in each group. The overall pool sample was sent to Eurofins MWG Operon for amplicon sequencing with the Roche GS-FLX Titanium series chemistry (454). Each primer attached with the 454 fusion sequence (Roche 2012) and labelled with a unique three base pair long tag on the forward and reverse primers (MID tags) so that the two sample types had a unique combination of tags and could run together in the same platform. MID tags were chosen from the list of 454 Standard MID set sequences recommended by Roche (2012).

## Data analysis

Stomach contents and scats were compared through Chi-square tests (comparing frequencies) and Mann–Whitney *U* tests (comparing number, estimated mass and sizes, after checking for normality). For DNA studies, the proportion of faecal samples from which DNA was successfully extracted and the proportion of those containing each of the prey types tested (Osteichthyes, Euphausiidae and Cephalopoda) were compared between old and fresh scat samples using a Chi-square contingency table test or a Fisher's exact test.

For the NGS data, the cutadapt package (Martin 2011) was used to de-multiplex the pooled pyrosequencing sequences based on the forward and reverse primers and MID tags and to remove all the adapters (including primers and MID tags). Sequences missing any of the adapters were discarded. Reads from 12S, 18S and 16S amplicons were filtered to a minimum length of 240 bp, 180 bp and 100 bp, respectively, and merged into a master file for each target group. The maximum sequencing length found for each gene was 269 bp for 12S, 201 bp for 18S and 128 bp for 16S; these represent normal variations in fragment sizes for rDNA genes so there was no need to filter sequences for maximum length.

Reads were dereplicated using -fastx_uniques in USEARCH v10.0.240 (Edgar 2010) and singletons were removed. To cluster reads into OTUs (Operational Taxonomic Units), the UPARSE pipeline was used for 12S amplicons analysis with a 97% clustering (Edgar 2013) while 18S and 16S were analysed in the UNOISE algorithm (Edgar 2016) with a 99% clustering, as suggested in previous works for this target groups (Bachy et al. 2013;

Edgar and Flyvbjerg 2015). During this process, three chimeras were detected and discarded for 12S and 10 chimeras were detected and discarded for 18S (no chimeras were detected for 16S). The total number of sequences retrieved, sequences lost, uniques and singletons for each gene can be found in Supplementary Material Appendix 1.

All commands and associated python scripts in the analysis are provided in the GitHub repository: https://github.com/AnaCarreiro/Xavier_et_al_GentooPenguins. NCBI database (Zhang et al. 2000; Morgulis et al. 2008) was used to taxonomically classify OTUs through MegaBLAST, and only results with 100% query cover were considered as matches. Each target group result was analysed using different analysis parameters: for 12S gene a minimum 90% Identity and *E*-value of $1^{-100}$; for 18S gene a minimum of 99% identity and *E*-value of $1^{-94}$; and for 16S a minimum of 99% identity and *E*-value of $1^{-57}$. These thresholds were defined considering each fragment size, but also based in previous works using these genes (e.g. Bachy et al. 2013) since a lower sequence similarity threshold would result in a mixing of different taxa with no ecological sense. For each OTU, all the reads matching the thresholds defined were considered and analysed together in order to classify each group to the lowest taxonomic level possible. Taxon (e.g. species, genus, family) was assigned if the query sequences clustered monophyleticaly at that level, producing an identical match in BLAST, higher to any other taxa. Species or genus levels identifications were not assigned if the identity match was below 99%, regardless of meeting the previous criteria; such cases were discussed as probable genus.

## Results

### Stomach contents versus scats (general diet)

A total of 43 samples with food (12 individuals were handled but no stomach contents were obtained) from stomach contents of gentoo penguins from May to October (13–15 individuals handled per month) were analysed visually, following Xavier et al. (2017). The most important components were crustaceans [86.0% of frequency of occurrence (FO); 98.4% of number (*N*); 59.1% by estimated mass (*M*)] and fish (88.4% FO; 1.5% *N*; 39.5% *M*), with the species *Themisto gaudichaudii* (74.4% FO; 76.9% *N*; 30.0% *M*), by frequency and by number, and *Parachaenichthys georgianus* (4.7% FO; 0.1% *N*, 28.1% *M*), by mass. *Euphausia superba* was the second most important species by FO and *N* (65.1% FO, 6.2% *N*, 7.1% M) (Table 2; Supplementary Material Appendix 2). In terms of biodiversity, 38 taxa were identifiable (24 crustaceans, 11 fish, two cephalopods and one other taxa) in the stomach contents.

A total of 168 scats were analysed. Crustaceans (95.8% FO; 86.0% *N*; 62.8% *M*) and fish (78.0% FO; 13.9% *N*; 37.0% *M*) were also the most frequent components. In terms of species, *T. gaudichaudii* (78.0% FO; 70.5% *N*; 25.2% *M*) was the most important by FO and by *N*, and *Champsocephalus gunnari* by mass (20.2% FO; 1.7% *N*; 30.7% *M*). Antarctic krill *E. superba* (26.8% FO; 6.5% *N*) was the second most important prey by FO and by *N* (Table 2; Supplementary Material Appendix 2). In terms of biodiversity, 30 taxa (16 crustaceans, 11 fish and three cephalopods) were identifiable in scats.

**Table 2** Frequency of occurrence (*F*), number (*N*) and estimated mass (*M*) of prey collected from stomach contents ($n=43$) and scats ($n=168$) from gentoo penguins (*Pygoscelis papua*) collected in 2009, through the post-breeding period, from May to October (of prey with > 5% mass; detailed information for all prey in Supplementary Material Appendix)

| Taxa | Stomach contents | | | Scats | | |
|---|---|---|---|---|---|---|
| | *F* (%) | *N* (%) | *M* (%) | *F* (%) | *N* (%) | *M* (%) |
| Crustaceans | 86.0 | 98.4 | 59.1 ± 41.3 | 95.8 | 86.0 | 62.8 ± 37.3 |
| Amphipoda | | | | | | |
| *Themisto gaudichaudii* | 74.4 | 76.9 | 30.0 | 78.0 | 70.5 | 25.2 |
| Euphausiacea | | | | | | |
| *Euphausia superba* | 65.1 | 6.2 | 7.1 | 26.8 | 6.5 | 5.7 |
| Fish | 88.4 | 1.5 | 39.5 ± 41.6 | 78.0 | 13.9 | 37.0 ± 37.2 |
| Bathydraconidae | | | | | | |
| *Parachaenichthys georgianus* | 4.7 | 0.1 | 28.1 | – | – | – |
| Channichthyidae | | | | | | |
| *Champsocephalus gunnari* | 37.2 | 0.2 | 14.8 | 20.2 | 1.7 | 30.7 |
| *Pseudochaenichthys georgianus* | 4.7 | < 0.1 | 0.9 | 1.2 | 0.1 | 9.9 |
| Nototheniidae | | | | | | |
| *Gobiotothen gibberifrons* | 4.7 | < 0.1 | 5.2 | – | – | – |
| *Lepidonotothen larseni* | 30.2 | 0.3 | 7.5 | 11.9 | 1.8 | 18.3 |
| Cephalopods | 2.4 | < 0.1 | 1.4 ± 9.2 | 2.4 | 0.1 | 0.2 ± 2.4 |
| Others | 11.6 | < 0.1 | < 0.1 ± 0.1 | – | – | – |

When comparing results from stomach contents and scats collected in the same time period, crustaceans was the most important component (in FO, *N* and *M*) in both methods, with no significant differences in FO (i.e. FO of crustaceans and fish in stomach contents and scats were not significant ($\chi^2_1 = 0.81$, $P = 0.483$)). However, there was a higher biodiversity of crustaceans in stomach contents ($n = 24$ species) than in scats ($n = 16$ species; Table 2) and the *N* in crustaceans was higher than fish in stomach contents (98.4 and 1.5%, respectively) in comparison with scats (86% by *N* in crustaceans and 13.9% in scats; Table 2). In term of species, *T. gaudichaudii* was clearly the most important species by FO and *N* in both sampling methods (see above; Table 2). By mass, different fish species were the most important prey (i.e. *P. georgianus* in stomach contents and *C. gunnari* in scats), with *P. georgianus* absent in scats. *E. superba* were the second most important prey by FO and *N* in both stomach contents and scats.

Comparing prey sizes (i.e. total length, Table 3; Supplementary Material Appendix 3) of the most important prey (with > 10% FO) found in stomach contents and in scats, the species *T. gaudichaudii* (Mann-Whitney *U* test, $U = 1231353$, $P < 0.0001$; bigger in scats, Fig. 1), *E. superba* (Mann–Whitney *U* test, $U = 5698$, $P < 0.0001$; bigger in stomach contents, Fig. 1), *G. georgiana* (Mann–Whitney *U* test, $U = 771$, $P < 0.0001$; bigger in scats), ?*Gymnoscopelus braueri* (Mann–Whitney *U* test, $U = 27358$, $P < 0.0001$; bigger in scats), *M. microps* (Mann–Whitney *U* test, $U = 408$, $P < 0.0001$; bigger in scats), *C. gunnari* (Mann–Whitney *U* test, $U = 3077$, $P < 0.00011$; bigger in scats) showed significant differences, whereas for *B. securiger* (Mann–Whitney *U* test, $U = 76$, $P = 0.913$) and *L. Larseni* (Mann–Whitney *U* test, $U = 2351$, $P = 0.267$) there were no differences in sizes.

## Comparison between visual identification and DNA based scat analyses

DNA was successfully amplified from all 30 scat samples (15 old and 15 fresh) using general Bilateria primers. Visual analysis of fresh scats (collected in October only; $n = 16$) showed that *E. superba* occurred in 100% of the samples which was higher than the occurrence of Euphausiidae through DNA amplification (80% for fresh scat samples and 93.3% for old scat samples) but a higher diversity of Euphausiidae taxa was identified using DNA analysis (Tables 4, 5). Fish occurred in 81.3% of the samples analysed visually (see Table 4 for the list of fishes identified visually, via their otoliths), while DNA detected them in 73.3 and 93.3% of fresh and old scats, respectively. Cephalopoda were not found in any samples analysed visually; DNA for this group was found only in fresh samples and only in two out of the 15 samples. The differences in prey detection between fresh and old scat samples are not statistically significant (all $P > 0.483$) suggesting that 1- to 2-day-old scats are as good as fresh samples in terms of successful amplifications of prey groups. Amplification success was lower for the Osteichthyes and Euphausiidae primers when the primers had the fusion sequence and MID tags attached (by 17 and 3%, respectively).

Regarding the results of the NGS analysis, UPARSE detected 10 OTUs for the 12S fish primers and UNOISE detected 14 OTUs for the 18S Bilateria primers (of which five were either human contamination, predator or parasites)

**Table 3** Sizes of prey identified from stomach contents ($n = 43$) and scats ($n = 168$) from gentoo penguins (*Pygoscelis papua*) collected in 2009, through the post-breeding period, from May to October (of species with > 5% Mass; detailed information for all prey in Supplementary Material Appendix)

| Taxa | Stomach contents | | | | Scats | | | |
|---|---|---|---|---|---|---|---|---|
| | Total length (mm) | | | | Total length (mm) | | | |
| | Mean | Min. | Max. | SD | Mean | Min. | Max. | SD |
| *Crustaceans* | | | | | | | | |
| Amphipoda | | | | | | | | |
| *Themisto gaudichaudii* | 22.08 | 15.0 | 35.0 | 1.8 | 22.8 | 15.0 | 35.0 | 1.8 |
| Euphausiacea | | | | | | | | |
| *Euphausia superba* | 44.6 | 27.0 | 58.0 | 4.5 | 38.3 | 31.7 | 47.2 | 2.2 |
| *Fish* | | | | | | | | |
| Bathydraconidae | | | | | | | | |
| *Parachaenichthys georgianus* | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a |
| Channichthyidae | | | | | | | | |
| *Champsocephalus gunnari* | 148.3 | 89.4 | 401.2 | 51.8 | 107.0 | 49.7 | 323.8 | 44.4 |
| *Pseudochaenichthys georgianus* | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a |
| Nototheniidae | | | | | | | | |
| *Gobiotothen gibberifrons* | n/a | n/a | n/a | n/a | | | | |
| *Lepidonotothen larseni* | 89.1 | 38.7 | 181.2 | 38.8 | 89.7 | 41.5 | 171.7 | 34.0 |

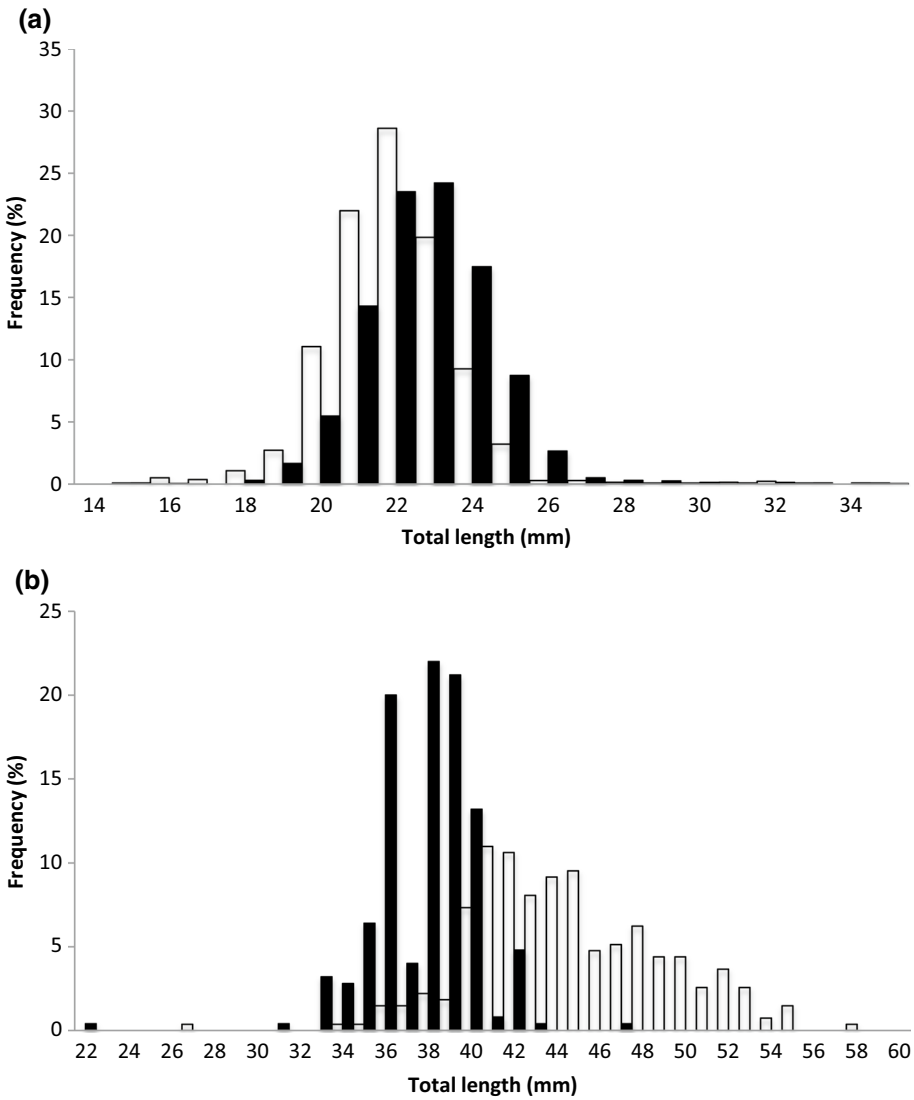


**Fig. 1** **a** *Themisto gaudichaudii* sizes from stomach contents (white bars; $n = 1402$ individuals measured from 43 penguins handled) and scats (black bars; $n = 2416$ individuals measured from 168 scats collected) from gentoo penguins. **b** *Euphausia superba* sizes from stomach contents (white bars; $n = 273$ individuals measured from 43 penguins handled) and scats (black bars; $n = 250$ individuals measured from 168 scats collected) from gentoo penguins

and 8 OTUs for 16S krill primers (of which one was discarded due to low quality query match). For the Bilateria primers, 68.4% of sequences from fresh scats were predator as opposed to 91.6% from the old scats. A further 3.7% of sequences from fresh scats and 4.1% from the old scats were identified as parasitic flatworms (Cestoda), possibly of the genus *Tetrabothrius*. Results of prey identified from the NGS analysis are shown in Table 5. As expected, prey identification from general Bilateria primers is hard to achieve at low taxonomic levels but it was possible to identify major prey sub-classes and some orders. A small proportion of sequences of copepods and cephalopods were identified in the fresh samples but not the old. This is in accordance with amplification results from the Cephalopoda primers but, regarding the copepods, it can be due to the reduced number of prey sequences obtained for old samples because of the higher preference of predator amplification. Fish, including Perciformes, and crustaceans, including krill species, were identified in both fresh and old samples. In both cases, krill comprised over 50% of the total prey sequences retrieved. The Perciformes together with another two unknown teleost fish orders comprised most of the remaining sequences for the old samples, while for the fresh samples this was split between the unknown teleost fish orders and one unknown crustacean group (also identified in the old samples but representing a much smaller proportion of the sequences).

The fish primers identified one OTU of lanternfish (Myctophidae) and nine OTUs within the Perciformes, including two species level (*Champsocephalus gunnari* and *Parachaenichthys charcoti*), six non-identified Nototheniidae (two of which are likely to be *Dissostichus* sp., one likely to be *Pagothenia sp.* and one likely to be *Notothenia* sp.) and one non-identified Channichthyidae. With the exception of one of the possible *Dissostichus* sp., which comprised a small proportion of sequences of the fresh samples only, all the fish taxa were found in both fresh and old samples.

**Table 4** Taxa identified/detected through convention (visual) identification in scats collected during October 2009 from Gentoo penguins (*Pygoscelis papua*) (*F* = frequency of occurrence)

| Taxa | Visual identification ($n = 16$) | |
|---|---|---|
| | *F* | *F* (%) |
| Crustaceans | 16 | 100.0 |
| Amphipoda | | |
| *Themisto gaudichaudii* | 4 | 25.0 |
| Decapoda | | |
| *Chorismus antarcticus* | 2 | 12.5 |
| Decapoda spp. | 1 | 6.3 |
| Euphausiacea | | |
| *Euphausia superba* | 16 | 100.0 |
| Fish | 13 | 81.3 |
| Channichthyidae | | |
| *Champsocephalus gunnari* | 4 | 25.0 |
| Myctophidae | | |
| *Gymnoscopelus braueri* | 7 | 43.8 |
| Mureanolepididae | | |
| *Muraenolepis microps* | 5 | 31.3 |
| Nototheniidae | | |
| *Lepidonotothen larseni* | 2 | 12.5 |
| Osteichthyes | 1 | 6.3 |
| Unidentifiable fish | 2 | 12.5 |

The proportion of sequences comprising each OTU varies between the two types of sample but in both cases the most represented prey are *P. charcoti* and the non-identified nototheniids, and the least represented is the non-identified Channichthyidae.

Six OTUs were identified with the krill primers, including four species of *Euphausia* and two of *Thysanoessa*. All OTUs were present in both fresh and old samples with equivalent proportions of sequences of each prey type in both groups with *E. superba* comprising around 50% of all the krill sequences in either case.

## Discussion

### Comparison between stomach contents and scats using conventional analyses (i.e. visually)

Gentoo penguins are amongst the species of top predators in the Southern Ocean used for Antarctic ecosystem monitoring (Agnew 1997; Everson 2002). The diet of gentoo penguins may provide insight into the prey stocks such as *E. superba* that can be related to environmental conditions locally (Croxall et al. 1999; Reid et al. 2005; Ratcliffe and Trathan 2011; Xavier et al. 2017). As gentoo penguins are resident, they can provide information all year round. This study shows that, using stomach contents and scats, gentoo penguins at South Georgia feed primarily on crustaceans and fish, relying very little on cephalopods, during the austral winter. This is in accordance with results from the various studies available for this region using dissections or stomach flushing on gentoo penguins during Austral summer and Austral winter (Croxall and Prince 1980; Croxall et al. 1988; Williams 1990; Kato et al. 1991; Croxall et al. 1997, 1999; Berrow et al. 1999; Xavier et al. 2017). In terms of biodiversity, the number of crustacean taxa differed between stomach contents and scats (higher in stomach contents) but similar for fish and cephalopod taxa. The main prey species by frequency of occurrence and by number (i.e. *Themisto gaudichaudii*) were the same in both stomach contents and scats (see results; Table 2). *T. gaudichaudii* proved to be an easily identifiable prey species, both in stomach contents and in scats, showing that it can resist digestion and go through the digestive tract, as other crustaceans do (e.g. *E. superba*). By mass, different prey species were the most important prey (*P. georgianus* for stomach contents and *C. gunnari* in scats). Such result suggests that otoliths of *P. georgianus* may get easily eroded, as they were not encountered in scats, which have implications on the fish component estimations using this method.

Sizes of some of the most important prey (e.g. *Themisto gaudichaudii* and *Euphausia superba*) differed between methods, but caution is necessary because *Euphausia superba* total length was measured directly from the individuals in stomach contents, whereas carapace lengths were measured in scats (with allometric equations used to estimate total length), which can add a level of error (Reid and Measures 1998; Goebel et al. 2007). Allometric equations for numerous crustaceans (e.g. from carapace size to mass/total length) are still urgently needed (see Supplementary Material Appendix 2) to have a proper evaluation of mass contribution by prey. Also, small carapaces and otoliths may tend to get more digested to be measured in scats which may lead to differences between methods (e.g. absence of *P. georgianus* otoliths in scats). Future studies should measure carapaces from both methods to maintain comparability between methods. Overall, stomach contents provided similar information to scats (but note the differences mentioned above, particularly on the fish component), and scats could be used as an alternative/complementary method for monitoring crustaceans, but attention should be paid to potential biases towards prey that need allometric equations to estimate their original size and mass. Stable isotopic analyses applied to scats can also allow the estimation of the diets (through mixed models) of crustaceans and fish (Horswill et al. 2016). Future research should also focus on the identification of prey that can be easily mistaken by other prey. For example, *E. frigida* was only found in stomach contents

**Table 5** Taxa identified from next-generation sequencing of fresh ($n=15$) and old (1–2 days, $n=15$) scats from gentoo penguins (*Pygoscelis papua*), collected in October 2009, using DNA fragments from three different genes

| Target Gene | Class | Subclass | Order | Family | Genus/Species | Fresh (%) | Old (%) |
|---|---|---|---|---|---|---|---|
| 18S | Cephalopoda | Unknown Subclass | | | | 0.24 | 0 |
| | Maxillopoda | Copepoda | Calanoida | | | 0.62 | 0 |
| | | | Unknown Order | | | 0.38 | 0 |
| | Actinopterygii | Neopterygii | Perciformes | Unknown Family | | 4.29 | 14.89 |
| | | | Unknown Teleost Order[a] | | | 10.19 | 27.15 |
| | | Unknown Subclass | | | | 1.54 | 0.29 |
| | Malacostrata | Eumalacostrata | Euphausiacea | Unknown Family | | 63.55 | 55.77 |
| | | Unknown Subclass | | | | 19.2 | 1.9 |
| 12S | Actinopterygii | Neopterygii | Myctophiformes | Myctophidae | Unknown | 18.04 | 9.37 |
| | | | Perciformes | Nototheniidae | Unknown Genus[b] | 22.03 | 50.78 |
| | | | | Channichthyidae | *Champsocephalus gunnari* | 11.71 | 13.03 |
| | | | | | Unknown Genus | 2.93 | 1.5 |
| | | | | Bathydraconidae | *Parachaenichthys charcoti* | 45.29 | 25.31 |
| 16S | Malacostrata | Eumalacostrata | Euphausiacea | Euphausiidae | *Euphausia superba* | 45.78 | 55.57 |
| | | | | | *Euphausia tricantha* | 0.03 | 0.23 |
| | | | | | *Euphausia frigida* | 0.08 | 0.05 |
| | | | | | Other *Euphausia* sp. | 0.47 | 0.32 |
| | | | | | *Thysanoessa macrura* | 50.90 | 43.35 |
| | | | | | Other *Thysanoessa* sp.[a] | 2.76 | 0.48 |

Percentages refer to proportion of sequences that comprise each prey type

[a]Classification contains 2 OTU's

[b]Classification contains 6 OTU's

that could be digestible and harder to identify in scats; the carapaces of *E. frigida* can be confused with *E. superba* if not familiar with euphausiid taxonomy. However, DNA techniques can aid to identify such less conspicuous species (see below), but are unable to provide information on size classes of prey.

## Comparison of DNA analyses between old and freshly collected scats (to assess levels of degradation) with those analysed visually

Our study compares for the first time fresh and old faecal samples from gentoo penguins, using molecular techniques, that successfully provided valuable dietary information, despite possible methodological issues (some primers were not used, i.e. Amphipoda primers, see Materials and methods) and difficulties detecting some prey types from old samples if these are competing with predator DNA (Table 5). This is relevant for future monitoring programs as collecting scats can be time restricted [i.e. programs have only a limited time at a given colony, and having the possibility to collect old scats for DNA analyses allows faster sample collection, without noticeable DNA damage (but dependent on location, time of the year and environmental conditions such as temperature or weather)] (Symondson 2002; Valentini et al. 2009; Pompanon et al. 2012).

Furthermore, this is useful since sometimes it is harder to obtain fresh samples, or to clearly differentiate fresh from old samples, making it possible to define a simpler protocol to collect scats for monitoring in the future. Also, our DNA techniques confirm the high amplification success of prey DNA from scats, supporting other studies that have used similar techniques with penguins (Deagle et al. 2005, 2007, 2010; Jarman et al. 2013). McInnes et al. (2017) found differences in amplification success between fresh and recent (wet) scats and significant differences in prey composition between fresh and dry scats of shy albatross *Thalassarche cauta* but these were collected from dirt or rock in warmer weather, suggesting that ice as a substrate is effective at preserving DNA integrity.

Although with the molecular method it was not always possible to identify prey to species or genus level, even when using group specific primers for fish or krill, this approach can clearly provide valuable dietary information, sometimes impossible to obtain through any other means. Primer choice is key for the detection of a wide range of prey and to assure that the DNA of the target prey species has been previously sequenced for the particular gene/

sequence of choice. However, primer choice is a trade off at various levels including fragment size, specificity and availability: the larger the fragment, the better the identification resolution but the less likely it is to be successfully amplified from scats; primers that amplify the widest possible range of prey species while avoiding amplifying the predator or environmental noise (e.g. bacteria) are hard to design. Similarly, from otoliths, beaks or certain euphausiids (e.g. carapaces of *E. superba*), it is possible to identify the item to species level visually in many cases but some items remained unidentified and others can be missed altogether (e.g. in the sample of scats used for comparing both methods, copepods were detected using molecular analysis but not by visual analysis). Molecular analysis can, therefore, be used in combination with visual analysis to aid in the identification of remains (Alonso et al. 2014; Lu et al. 2016), with all these methods complementing each other.

Although great caution must be used to draw any quantitative conclusions from the NGS analysis (as the number of sequences does not necessarily reflect the quantity consumed or any errors if amplification does not establish an association), dietary shifts in terms of prey composition have been identified with changes between months/years using penguins (Jarman et al. 2013). Furthermore, the low incidence of cephalopods and high incidence of krill, particularly *E. superba*, nototheniids, and species such as the Mackerel icefish *C. gunnari* suggested in terms of proportion of sequences retrieved from these groups, are in agreement with the visual analysis and previous studies (Croxall et al. 1999; Ratcliffe and Trathan 2011). It was possible to identify a greater number of krill species using the molecular approach when compared to the visual identification, even considering prey identified visually from the larger number of samples collected through the post-breeding period. Two species of fish detected through molecular methods were not detected visually: *P. charcoti* (known to be distributed further south (Gon and Heemstra 1990), being from the same family as *P. georgianus*, which was an important species in stomach contents; see results) and a toothfish (*Dissostichus* sp.). For the latter, two OTUs matched equally to the two species of toothfish (although *D. eleginoides* is the only *Dissostichus* species to occur around South Georgia Gon and Heemstra 1990), meaning that both species could have been consumed or a case of OTU over-splitting, with two OTUs representing the same species. Although the overall diversity of fish prey found using molecular analysis was greater (ten different taxa), only two taxa could be identified to species level, compared to four using visual analysis. The lack of reference sequences for some of these species on GenBank hinders their identification to species level. An extensive reference collection of DNA sequences of identified potential prey would greatly aid these analyses in the future. Indeed, data from this study can be revisited as more reference sequences become available for comparison.

Current standards of practice regarding using NGS to investigate animal's diets would include performing three different PCRs with the same primers in each sample, in order to minimise PCR bias in the NGS run and including negative primer combinations (i.e. primer combinations used on PCR negatives—without DNA) in the NGS run for control during the bioinformatics process (McInnes et al. 2017). These standards were not considered at the time the laboratory work for this study was performed and could have implications in the results obtained, namely lower prey diversity or potential comparisons bias between old and fresh scats due to minor contamination or primer tag jumping (Schnell et al. 2015)—this would be particularly worrying if many prey types occurred in a very small proportion in one of the groups compared to the other but this is not the case for most prey types (Table 5) and those that do should not change the nature of our conclusions. Nevertheless, despite the technical limitations of the molecular approach, considering the relatively small sample sizes for comparisons between the molecular methods and visual analysis, the consistency in prey composition between the two methods is very reassuring.

*Themisto gaudichaudii* was shown to be an important prey from the stomach analysis but could not be detected in the molecular analysis since none of the primers used in the NGS amplified amphipods (when our study was conducted). Designing a species-specific primer to target this particular prey would be a useful development for future studies. More recently published Metazoa COI primers (Leray et al. 2013) might also allow the identification of this species in future studies and consequently be incorporated in future monitoring programs. Other primers, also developed recently, that can be useful for studying penguin diet include a primer pair for 16S rDNA targeting Decapoda, marine Isopoda and Euphausiacea (Waap 2015).

## Critically evaluate the potential use of scats as a suitable alternative, non-invasive method for investigating gentoo penguin diet

Our study showed that scats provide a valuable technique to identify the most important component (crustaceans) and species by frequency of occurrence and number (i.e. *Themisto gaudichaudii*) in the diet of gentoo penguins, similar to the information obtained from stomach contents. However, by mass considerable differences occurred (e.g. the most important fish prey in stomach contents was absent in scats) and differences in sizes. As this study focused on a single sampling period (during Winter), a more complete test relevant for monitoring programs would be assessing the diets using these methods between years (ideally with

different prey). Also, it would be useful to develop a monitoring program that could reflect the diet but also inferences in the feeding and foraging ecology of predators, such as penguins.

In terms of methodologies applied, through DNA analyses, information could be obtained from all fresh and old scats and was similar to that obtained from visual identification; lower diversity in DNA analyses in certain taxa (e.g. overall crustaceans) was due to the lack of adequate primers, while lower diversity of krill species in visual analysis is probably due to difficulty of detecting small degraded prey consumed probably in small quantities. DNA techniques have the ability to analyse large amounts of samples, at different times and across geographical scales (Jarman et al. 2013). Despite the many advantages of molecular methods to study animal diets (Clare 2014), these techniques (NGS in particular) are usually costly when compared to more conventional methods. Typically, a minimum of £1500 GBP is required for lab costs (excluding any staff payments) even for a small number of samples (whereas for visual identification, an undergraduate/MSc student project under supervision of an expert (and a collection) could be cheaper that the value presented but more labour intensive). Such costs are steadily decreasing as the new generation of sequencing technology evolves and its use becomes more widespread, but it will always be a more expensive option than visual analysis. A less costly alternative is to design specific primers to target key prey species to routinely investigate the proportion of gentoo penguin's faecal samples testing positive for these prey types, similar to Thalinger et al. (2016). This is a simple way of investigating yearly dietary shifts potentially caused by changes in prey abundance without the need for a detailed dietary study (using NGS), which could be done less frequently. Either way, access to a well-equipped lab is essential (e.g. general lab consumables, vortex, centrifuges, heat block, PCR machines, gel electrophoreses system, DNA concentration calculator) or established collaborations with institutes that have such facilities. For monitoring purposes, considering having such a consortium of laboratories endorsed by contributing nations with samples (as this work cannot be done in the field), with data going to an international database accessible free (e.g. SCAR Diet and energetics database; https://www.scar.org/data-products/southern-ocean-diet-energetics/), could be a way forward.

Overall, the results presented here show that these methods (i.e. diet characterised from stomach contents and from scats, identified visually and by DNA techniques), possibly with other techniques (e.g. stable isotopic and fatty acid analyses that provided valuable information over space and time) (Bearshop et al. 2006; Polito et al. 2011; Karnovsky et al. 2012; Gorman et al. 2014), could contribute to ecosystem monitoring programs, such as CEMP. The methodology used here, along with similar DNA techniques (Deagle et al. 2005, 2010; Jarman et al. 2013), is applicable to monitoring the diet of any bird or mammal and could be used to address a variety of ecological questions. Indeed, such methodology could be applied to the study of the diet of less generalist species (e.g. Adélie and Chinstrap penguins) that consume a high percentage of crustaceans (with small variety of species) with known primers. DNA techniques applied to scats are useful to complement the identification of prey visually (i.e. when it is too digested to be identified visually) from samples from stomach contents (this method is still considered very useful to get the size of prey and species identification (when prey individuals are complete) but there are still important limitations to using next-generation sequencing techniques as a part of a long-term regular monitoring programs (e.g. need of good primers, costs and rapid advances of the field with quick changes of standards of practice) (see Clare 2014 for a detailed review of the advantages and limitations of molecular methods).

**Acknowledgements** We thank *Antarctic science* Bursary for the Antarctic Science scholarship awarded to this project and the editor Dieter Piepenburg, Tom Hart and two other referees for their comments. JX is supported by the Investigator FCT program (IF/00616/2013) and is part of the SCAR Ant-ERA, SCAR EGBAMM, PROPOLAR and ICED programs. This study benefited from the strategic program of MARE, financed by FCT (MARE- UID/MAR/04292/2013). We thank Naomi Treble and Alexandra McCubbin for their support in the lab.

## Compliance with ethical standards

**Conflict of interests** The authors have no conflict of interests.

## References

Agnew DJ (1997) Review—the CCAMLR ecosystem monitoring programme. Antarct Sci 9:235–242

Alonso H, Granadeiro JP, Waap S, Xavier J, Symondson WOC, Ramos JA, Catry P (2014) An holistic ecological analysis of the diet of Cory's shearwaters using prey morphological characters and DNA barcoding. Mol Ecol 23:3719–3733

Bachy C, Dolan JR, López-García P, Deschamps P, Moreira D (2013) Accuracy of protist diversity assessments: morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. ISME J 7:244–255

Barrett RT et al (2007) Diet studies of seabirds: a review and recommendations. ICES J Mar Sci 64:1675–1691

Bearshop S, Phillips RA, McGill R, Cherel Y, Dawson DA, Croxall JP (2006) Stable isotopes indicate sex-specific and long-term individual foraging specialization in diving seabirds. Mar Ecol Prog Ser 311:157–164

Bellan-Santini D, Ledoyer M (1974) Gammariens (Crustacea, Amphipoda) des Iles Kerguelen et Crozet. Tethys 5:635–708

Berrow SD, Taylor RI, Murray AWA (1999) Influence of sampling protocol on diet determination of gentoo penguins *Pygoscelis papua* and Antarctic fur seals *Arctocephalus gazella*. Polar Biol 22:156–163

Blankenship LE, Yayanos AA (2005) Universal primers and PCR of gut contents to study marine invertebrate diets. Mol Ecol 14:891–899
Boltovskoy D (1999) South Atlantic zooplankton. Backhuys Publishers, Leiden
Bowser AK, Diamond AW, Addison JA (2013) From puffins to plankton: a DNA-based analysis of a seabird food chain in the northern Gulf of Maine. PLoS ONE 8:e83152
Boyd IL (2002) Estimating food consumption of marine predators: Antarctic fur seals and macaroni penguins. J Appl Ecol 39:103–119
Brooke ML (2004) The food consumption of the world´s seabirds. Proc R Soc Lond B 271:S246–S248
Chekunova VI, Rynkova TI (1974) Energy requirements of the Antarctic crustacean *Euphausia superba* DANA. Oceanology 14:434–440
Cherel Y, Le Corre M, Jaquemet S, Ménard F, Richard P, Weimerskirch H (2008) Resource partitioning within a tropical seabird community: new information from stable isotopes. Mar Ecol Prog Ser 366:281–291
Clare E (2014) Molecular detection of trophic interactions: emerging trends, distinct advantages, significant considerations and conservation applications. Evol Appl 7:1144–1157
Clarke J, Kerry K (1994) The effects of monitoring procedures on Adélie penguins. CCAMLR Sci 1:155–164
Constable AJ (2002) CCAMLR ecosystem monitoring and management: future work. CCAMLR Sci 9:233–253
Croxall JP (1987) Seabirds: feeding ecology and role in marine ecosystems. Cambridge University Press, Cambridge
Croxall JP (1993) Diet. In: Laws RM (ed) Antarctic seals: research methods and techniques. Cambridge University Press, Cambridge, pp 268–290
Croxall JP, Prince PA (1980) The food of gentoo penguins *Pygoscelis papua* and macaroni penguins *Eudyptes chrysolophus* at South Georgia. Ibis 122:245–253
Croxall JP, Prince PA (1987) Seabirds as predators on marine resources, especially krill, at South Georgia. In: Croxall JP (ed) Seabirds: feeding ecology and role in marine ecosystems. Cambridge University Press, Cambridge, pp 347–368
Croxall JP, Davis RW, Oconnell MJ (1988) Diving patterns in relation to diet of gentoo and macaroni penguins at South Georgia. Condor 90:157–167
Croxall JP, Prince PA, Reid K (1997) Dietary segregation in South Georgia seabirds. J Zool 242:531–556
Croxall JP, Reid K, Prince PA (1999) Diet, provisioning and productivity responses of marine predators to differences in availability of Antarctic krill. Mar Ecol Prog Ser 177:115–131
Deagle BE, Tollit DJ, Jarman SN, Hindell MA, Trites AW, Gales NJ (2005) Molecular scatology as a tool to study diet: analyses of prey DNA in scats from captive Steller sea lions. Mol Ecol 14:1831–1842
Deagle BE, Gales NJ, Evans K, Jarman SN, Robinson S, Trebilco R, Hindell MA (2007) Studying seabird diet through genetic analysis of faeces: a case study on macaroni penguins (*Eudyptes chrysolophus*). PLoS ONE 2:e831
Deagle BE, Chiaradia A, McInnes I, Jarman SN (2010) Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? Conserv Genet 11:2039–2048
Duffy DC, Jackson S (1986) Diet studies of seabirds: a review of methods. Colonial Waterbirds 9:1–17
Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461
Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10:996–998
Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. bioRxiv:081257
Edgar RC, Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 31:3476–3482
Everson I (2002) Consideration of major issues in ecosystem monitoring and management. CCAMLR Sci 9:213–232
Goebel ME, Lipsky JD, Reiss CS, Loeb VJ (2007) Using carapace measurements to determine the sex of Antarctic krill, *Euphausia superba*. Polar Biol 30:307–315
Gon O, Heemstra PC (1990) Fishes of the Southern Ocean. JLB Smith Institute of Ichthyology, Grahamstown
Gorman KB, Williams TD, Fraser WR (2014) Ecological sexual dimorphism and environmental variability within a community of Antarctic Penguins (Genus *Pygoscelis*). PLoS ONE 9:e90081
Hecht T (1987) A guide to the otoliths of Southern Ocean fishes. S Afr J Antarctic Res 17:2–87
Horswill C et al (2016) Drivers of intrapopulation variation in resource use in a generalist predator, the macaroni penguin. Mar Ecol Prog Ser 548:233–247
IUCN (2010) IUCN Red List of Threatened Species. Version 20104. www.iucnredlist.org
Jarman SN, Deagle BE, Gales NJ (2004) Group-specific polymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples. Mol Ecol 13:1313–1322
Jarman S, Redd K, Gales N (2006) Group-specific primers for amplifying DNA sequences that identify Amphipoda, Cephalopoda, Echinodermata, Gastropoda, Isopoda, Ostracoda and Thoracica. Mol Ecol 6:268–271
Jarman SN et al (2013) Adélie penguin population diet monitoring by analysis of food DNA in scats. PLoS ONE 8:e82227
Karnovsky NJ, Hobson KA, Iverson SJ (2012) From lavage to lipids: estimating diets of seabirds. Mar Ecol Prog Ser 451:263–284
Kato A, Williams TD, Barton TR, Rodwell S (1991) Short-term variation in the winter diet of gentoo penguins *Pygoscelis papua* at South Georgia during July 1989. Mar Ornithol 19:31–38
King RA, Read DS, Traugott M, Symondson WOC (2008) Molecular analysis of predation: a review of best practice for DNA-based approaches. Mol Ecol 17:947–963
Kirkwood JM (1984) A guide to the Decapoda of the Southern Ocean. ANARE Res Notes 11:1–47
Leray M et al (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Front Zool 10:34
Lu Z, Xu S, Song N, Gao T, Tian J, Han J (2016) Analysis of the diet of finless porpoise (*Neophocaena asiaeorientalis sunameri*) based on prey morphological characters and DNA barcoding. Conserv Genet Resour 8:523–531
Lumsden WHR, Fladdow AJ (1946) The food of the Shag *P. aristotelis* in the Clyde Sea area. Anim Ecol 15:35–42
Lynnes A, Reid K, Croxall J (2004) Diet and reproductive success of Adélie and chinstrap penguins: linking response of predators to prey population dynamics. Polar Biol 27:544–554
Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12
McInnes JC, Emmerson L, Southwell C, Faux C, Jarman SN (2016a) Simultaneous DNA-based diet analysis of breeding, non-breeding and chick Adélie penguins. R Soc Open Sci 3:150443
McInnes JC, Raymond B, Phillips RA, Jarman SN, Lea M-A, Alderman R (2016b) A review of methods used to analyse albatross diets—assessing priorities across their range. ICES J Mar Sci 73:2125–2137
McInnes JC, Alderman R, Deagle BE, Lea M-A, Raymond B, Jarman SN (2017) Optimised scat collection protocols for dietary DNA metabarcoding in vertebrates. Methods Ecol Evol 8:192–202

Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA (2008) Database indexing for production MegaBLAST searches. Bioinformatics 24:1757–1764

Murray DC et al (2011) DNA-based faecal dietary analysis: A comparison of qPCR and high throughput sequencing approaches. PLoS ONE 6:e25776

Olmos-Pérez L, Roura Á, Pierce GJ, Boyer S, González ÁF (2017) Diet composition and variability of wild *Octopus vulgaris* and *Alloteuthis media* (Cephalopoda) paralarvae: A Metagenomic Approach. Front Physiol 8:321

Pakhomov E, Perissinotto R (1996) Trophodynamics of the hyperiid amphipod Themisto gaudichaudi in the South Georgia region during late austral summer. Mar Ecol Prog Ser 134:91–100

Pauly D, Christensen V, Dalsgaard J, Froese R, Jr Torres F (1998) Fishing down marine food webs. Science 279:860–863

Peters KJ, Ophelkeller K, Bott NJ, Deagle BE, Jarman SN, Goldsworthy SD (2015) Fine-scale diet of the Australian sea lion (*Neophoca cinerea*) using DNA-based analysis of faeces. Mar Ecol 36:347–367

Polito MJ, Trivelpiece WZ, Karnovsky NJ, Ng E, Patterson WP, Emslie SD (2011) Integrating stomach content and stable isotope analyses to quantify the diets of Pygoscelid penguins. PLoS ONE 6:e26642

Pompanon F, Deagle BE, Symondson WO, Brown DS, Jarman SN, Taberlet P (2012) Who is eating what: diet assessment using next generation sequencing. Mol Ecol 21:1931–1950

Raclot T, Groscolas R, Cherel Y (1998) Fatty acid evidence for the importance of myctophid fishes in the diet of king penguins, *Aptenodytes patagonicus*. Mar Biol 132:523–533

Rakusa-Suszczewski S, Stepnik R (1980) Three species of krill from Admiralty Bay (King George, South Shetlands), in summer 1978/79. Pol Arch Hydrobiol 27:273–284

Ratcliffe N, Trathan PN (2011) A review of the diet and at sea-distribution of penguins breeding within the CCAMLR convention area. CCAMLR Sci 18:75–114

Reid K (1996) A guide to the use of otoliths in the study of predators at South Georgia. British Antarctic Survey, Cambridge

Reid K, Measures J (1998) Determining the sex of Antarctic krill *Euphausia superba* using carapace measurements. Polar Biol 19:145–147

Reid K, Croxall JP, Briggs DR, Murphy EJ (2005) Antarctic ecosystem monitoring: quantifying the response of ecosystem indicators to variability in Antarctic krill. ICES J Mar Sci 62:366–373

Roche (2012) 454 Sequencing system guidelines for amplicon experimental design. 454 Life Sciences Corp. Branford, USA

Schnell IB, Bohmann K, Thomas M, Gilbert P (2015) Tag jumps illuminated—reducing sequence-to-sample misidentifications in metabarcoding studies. Mol Ecol Resour 15:1289–1303

Seefelt NE, Gillingham JC (2006) A comparison of three methods to investigate the diet of breeding double-crested cormorants (*Phalacrocorax auritus*) in the Beaver Archipelago, northern Lake Michigan. Hydrobiol 567:57–67

Siegel V (1993) Review of length-weight relationships for Antarctic krill. SC-CAMLR Sel Sci Pap 9:145–155

Siegel V, Mühlenhardt-Siegel U (1988) On the occurrence and biology of some Antarctic Mysidacea (Crustacea). Polar Biol 8:181–190

Smale MJ, Watson G, Hecht T (1995) Otolith atlas of Southern African marine fishes, vol 1. Ichthyological Monographs of the JLB Smith Institute of Ichthyology, Grahamstown

Symondson WOC (2002) Molecular identification of prey in predator diets. Mol Ecol 11:627–641

Tanton JL, Reid K, Croxall JP, Trathan PN (2004) Winter distribution and behaviour of gentoo penguins *Pygoscelis papua* at South Georgia. Polar Biol 27:299–303

Thalinger B, Oehm J, Mayr H, Obwexer A, Zeisler C, Traugott M (2016) Molecular prey identification in Central European piscivores. Mol Ecol Resour 16:123–137

Tollit DJ, Heaslip SG, Barrick RL, Trites AW (2007) Impact of diet-index selection and the digestion of prey hard remains on determining the diet of the Steller sea lion (*Eumetopias jubatus*). Can J Zool 85:1–15

Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. Trends Ecol Evol 24:110–117

Waap S (2015) Trophic relationships among pelagic predators of the deep seas of the Madeira Islands. Doctoral dissertation, Cardiff University (United Kingdom)

Waluda CM, Hill SL, Peat HJ, Trathan PN (2017) Long-term variability in the diet and reproductive performance of penguins at Bird Island, South Georgia. Mar Biol 164:39

Williams TD (1990) Foraging ecology and diet of gentoo penguins (*Pygoscelis papua*) at South Georgia during winter and an assessment of their winter prey consumption (WG-CEMP-90/16). Working Group for the CCAMLR Ecosystem Monitoring Program, Stockholm

Williams R, McEldowney A (1990) A guide to the fish otoliths from waters off the Australian Antarctic Territory, Heard and Macquarie Island vol 75. ANARE Research Notes

Wilson RP (1984) An improved pump for penguins and other seabirds. J Field Ornithol 55:109–112

Wilson RP (2010) Resource partitioning and niche hyper-volume overlap in free-living Pygoscelid penguins. Funct Ecol 2010:649–657

Xavier JC, Cherel Y (2009) Cephalopod beak guide for the Southern Ocean. British Antarctic Survey, Cambridge

Xavier JC, Croxall JP, Reid K (2003) Inter-annual variation in the diet of two albatross species breeding at South Georgia: implications for breeding performance. Ibis 145:593–610

Xavier JC, Trathan PN, Croxall JP, Wood AG, Podestá GP, Rodhouse PG (2004) Foraging ecology and interactions with fisheries of wandering albatrosses at South Georgia. Fish Oceanogr 13:324–344

Xavier JC, Croxall JP, Cresswell KA (2005) Boluses: an effective method to assess the proportions of cephalopods in the diet of albatrosses. Auk 122:1182–1190

Xavier JC et al (2017) Sexual and individual foraging segregation in Gentoo penguins *Pygoscelis papua* from the Southern Ocean during an abnormal winter. PLoS ONE 12:e0174850

Zhang Z, Schwartz S, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7:203–214