

10.1071/WR19246_AC

© CSIRO 2020

Supplementary Material: *Wildlife Research*, 2020, 47(6), 509-522.

Species in the faeces: DNA metabarcoding as a method to determine the diet of the endangered yellow-eyed penguin

Melanie J. Young^{A,B}, Ludovic Dutoit^A, Fiona Robertson^A, Yolanda van Heezik^A, Philip J. Seddon^A and Bruce C. Robertson^A

^ADepartment of Zoology, University of Otago, PO Box 56, Dunedin 9054, New Zealand.

^BCorresponding author. Email: melanie.young@postgrad.otago.ac.nz

S1. Known prey from previous yellow-eyed penguin stomach flushing studies

Table S1. Chordate and cephalopod taxa (identified to species or genus) in the diet of yellow-eyed penguins (*Megadyptes antipodes*) resident on the Otago coast, New Zealand, identified via stomach flushing by van Heezik (1990a) and Moore and Wakelin (1997), and by DNA metabarcoding in this study. *indicates previously identified taxa might have occurred at a lower resolution than genus in this study (see S4). Taxa classified in the current study were accepted following manual curation to confirm geographical plausibility, as indicated by the curation notes below, after computational filtering. [§]indicates if the geographically plausible congener (i.e. a congener found within the known foraging range of adult yellow-eyed penguins (i.e. waters adjacent Canterbury, Otago, Southland, sub-Antarctic breeding areas), according to Roberts *et al.* 2015) was found in the reference database. Red text indicates geographically implausible taxa, which were excluded from the results.

Prey taxa identified in current and previous yellow-eyed penguin diet studies	Species/genus present in DNA reference database	DNA reference origin	Diagnostic accuracy (this study)	Previously detected by stomach flushing	Curation notes	Geographically plausible congeners found within the known foraging range of adult yellow-eyed penguins
Chordata						
Āhuru <i>Auchenoceros punctatus</i>	Present	This study	Species	Y	a	None
Barracouta/manga <i>Thyrsites atun</i>	Present	GenBank	Species	Y	a	None
Barred grubfish <i>Parapercis</i> sp.	Present	GenBank	Genus	N	d	<i>P. colias</i> [§] <i>P. gilliesii</i>
Blue cod/rāwaru <i>Parapercis colias</i>	Present	This study	Species	Y	c	<i>P. gilliesii</i>
Blue grenadier/hoki <i>Macruronus novaezelandiae</i>	Present	GenBank	Not detected	Y	NA	NA
Blue warehou/warehou kahurangi <i>Seriolella brama</i>	Present	GenBank	Species	Y	c	<i>S. punctata</i> [§] <i>S. caerulea</i> [§]
Bream <i>Tylochromis leonensis</i>	Present	GenBank	Species	N	f	Unknown
Common triplefin/kokopara <i>Forsterygion lapillum</i>	Present	GenBank	Species	Y	c	<i>F. capito</i> <i>F. flavonigrum</i> <i>F. gymnotum</i> <i>F. malcolmi</i> <i>F. nigripenne</i> <i>F. varium</i>
Gudgeon/kuraihana <i>Grahamichthys radiata</i>	Absent	NA	NA	Y	NA	NA
Hake/tikati <i>Merluccius australis</i>	Present	GenBank	Not detected	Y	NA	NA
Hector's lanternfish <i>Lampanyctodes hectoris</i>	Present	GenBank	Not detected	Y	NA	NA
Jackass morwong/tarakihi <i>Nemadactylus macropterus</i>	Present	GenBank	Not detected*	Y	NA	NA
Lamprey/piharau <i>Geotria australis</i>	Absent	NA	NA	Y	NA	NA
Leatherjacket/kokiri <i>Meuschenia scaber</i>	Present	GenBank	Not detected	Y	NA	NA

Lemon sole/pātiki <i>Pelotretis flavilatus</i>	Present	GenBank	Not detected	Y	NA	NA
Ling/hokarari <i>Genypterus blacodes</i>	Present	GenBank	Species	Y	g	None
Long-snouted pipefish <i>Stigmatopora macropterygia</i>	Absent	NA	NA	Y	NA	NA
Morid cod <i>Lepidion</i> sp.	Present	GenBank	Genus	N	d	<i>L. microcephalus</i> <i>L. schmidti</i>
Mother-of-Pearl pipefish <i>Vanacampus margaritifer</i>	Present	GenBank	Species	N	f	Unknown
Slender opalfish/kohikohi <i>Hemeroetes artus</i>	Present	This study	Genus	Y	h	<i>H. monopterygius</i> [§] <i>H. morelandi</i> [§] <i>H. pauciradiatus</i> <i>H. macrophthalmus</i>
New Zealand sand diver <i>Tewara cranwellae</i>	Absent	NA	NA	Y	NA	NA
New Zealand banded wrasse <i>Notolabrus fucicola</i>	Present	GenBank	Species	N	c	<i>N. celidotus</i> <i>N. cinctus</i>
New Zealand blueback sprat/kupae <i>Sprattus antipodum</i>	Present	GenBank	Species	Y	b	<i>S. muelleri</i> [§]
New Zealand sole/tarore <i>Peltorhamphus novaezeelandiae</i>	Present	GenBank	Not detected	Y	NA	NA
New Zealand sprat/kupae <i>Sprattus muelleri</i>	Present	GenBank	Species	Y	b	<i>S. antipodum</i> [§]
Opalfish/kohikohi <i>Hemeroetes monopterygius</i>	Present	This study	Genus	Y	h	<i>H. artus</i> [§] <i>H. morelandi</i> [§] <i>H. pauciradiatus</i> <i>H. macrophthalmus</i>
Pigfish/purumoru <i>Congiopodus</i> sp.	Present	This study	Genus	N	h	<i>C. leucopaecilus</i> [§] <i>C. coriaceus</i> [§]
Rattail <i>Coelorinchus aspercephalus</i>	Present	GenBank	Species	Y	c	<i>C. biclinozonalis</i> <i>C. bollonsi</i> [§] <i>C. fasciatus</i> [§] <i>C. innotabilis</i> [§] <i>C. kaiyomaru</i> [§] <i>C. oliverianus</i> [§] <i>C. parvifasciatus</i> [§] <i>C. trachycarus</i>
Red baitfish <i>Emmelichthys nitidus</i>	Present	GenBank	Species	N	a	None
Red cod/hoka <i>Pseudophycis bachus</i>	Present	GenBank	Species	Y	b	<i>P. barbata</i> [§]
Jock Stewart <i>Helicolenus percoides</i>	Present	GenBank	Not detected*	Y	NA	NA
Ribaldo <i>Mora moro</i>	Present	GenBank	Species	N	a	None
Rock fish <i>Sebastes</i> sp.	Present	GenBank	Genus	N	j	Unknown
Rock cod <i>Lotella rhacina</i>	Present	GenBank	Not detected	Y	NA	NA
Salps <i>Thaliacea</i> sp.	Absent	NA	NA	Y	NA	NA
Sandfish <i>Gonorynchus gonorynchus</i>	Present	This study	Not detected	Y	NA	NA
Scarlet wrasse <i>Pseudolabrus miles</i>	Present	GenBank	Species	N	a	None
Scorpionfish/rarai <i>Scorpaena cardinalis</i>	Absent	NA	NA	Y	NA	NA
Seachub <i>Medialuna californiensis</i>	Present	GenBank	Species	N	f	Unknown
Seahorse/kiore moana <i>Hippocampus abdominalis</i>	Present	GenBank	Species	Y	a	None

Short-finned eel/tuna <i>Anguilla australis</i>	Present	GenBank	Not detected	Y	NA	NA
Short-snouted pipefish <i>Lissocampus filum</i>	Absent	NA	NA	Y	NA	NA
Silver conger <i>Gnathopis habenatus</i>	Absent	NA	NA	Y	NA	NA
Silver warehou/warehou hiriwa <i>Serirolella punctata</i>	Present	GenBank	Genus	Y	NA	<i>S. brama</i> [§] <i>S. caerulea</i> [§]
Silverside <i>Argentina elongata</i>	Present	This study	Species	Y	a	None
Smelt/paraki <i>Retropinna retropinna</i>	Present	GenBank	Not detected	Y	NA	NA
Southern bastard cod <i>Pseudophycis barbata</i>	Present	GenBank	Species	N	b	<i>P. bachus</i> [§]
Southern blue whiting <i>Micromesistius australis</i>	Present	GenBank	Species	N	e	None
Snipefish <i>Macroramphosus scolopax</i>	Present	GenBank	Species	N	a	None
Spiny gurnard <i>Lepidotrigla sp.</i>	Present	GenBank	Genus	N	d	<i>L. brachyoptera</i> <i>L. robinsi</i>
Stargazer/pūwhara <i>Kathetostoma giganteum</i>	Present	GenBank	Species	Y	c	<i>K. binigrasella</i>
Violet cod <i>Antimora rostrata</i>	Present	GenBank	Species	N	a	None
Violet warehou <i>Schedophilus velaini</i>	Present	GenBank	Species	N	j	Unknown
Cephalopoda						
Arrow squid/wheketere <i>Nototodarus sloanii</i>	Present	GenBank	Genus	Y	c	<i>N. gouldi</i>
Greater hooked squid/ngū <i>Moroteuthopsis ingens</i>	Present	GenBank	Species	Y	a	None
Knobby argonaut/ pūpū tarakihi <i>Argonauta nodosa</i>	Present	GenBank	Species	N	c	<i>A. argo</i>
Octopus/wheke <i>Macroctopus maorum</i>	Present	GenBank	Genus	Y	a	None
Pacific bobtail squid <i>Sepioloidea pacifica</i>	Absent	NA	NA	Y	NA	NA
Yellow octopus/wheke <i>Enteroctopus sp.</i>	Present	GenBank	Genus	N	a	<i>E. zealandicus</i>
Arthropoda						
Squillid shrimp <i>Heterosquilla sp.</i>	Absent	NA	NA	Y	NA	NA
Krill <i>Nyctiphanes australis</i>	Absent	NA	NA	Y	NA	NA

Curation notes for Table S1

- a.** Geographically plausible species selected by the naïve-Bayes classifier and monotypic in their genus in New Zealand's exclusive economic zone (EEZ). *Listed by species.*
- b.** All geographically plausible species and congeners found in New Zealand's EEZ were selected by the naïve-Bayes classifier. *Listed by species.*
- c.** Geographically plausible species selected by the naïve-Bayes classifier, but not all geographically plausible New Zealand congeners found within the known foraging range of adult yellow-eyed penguins were accounted for in the reference database. *Listed by species.*
- d.** Geographically implausible species selected by the naïve-Bayes classifier and not all geographically plausible New Zealand congeners found within the known foraging range of adult yellow-eyed penguins were accounted for in the reference database. *Listed by genus.*
- e.** Geographically implausible species selected by the naïve-Bayes classifier, but the genus is geographically plausible, and contains only one species in New Zealand's EEZ. *Listed by species.*
- f.** Geographically implausible species selected by the naïve-Bayes classifier with no known New Zealand congeners. *Removed from results. Listed by species, in red.*
- g.** Geographically plausible genus selected by the naïve-Bayes classifier, and the genus contains only one species in New Zealand's EEZ. *Listed by species.*
- h.** Geographically plausible genus manually classified from validated reference material. *Listed by genus.*
- i.** Geographically plausible genus selected by the naïve-Bayes classifier but grouped separately from one or more geographically plausible New Zealand congeners identified by the classifier. *Listed by genus.*
- j.** Geographically plausible genus selected by the naïve-Bayes classifier, but recent change to or questionable taxonomy. *Listed by genus.*

S2. Occurrence and within-sample filtering of taxa

As we had used a conservative filtering approach to limit false positives and to include only those samples with >2000 reads, we opted not to impose a third filter as an arbitrary within-sample read threshold for individual taxa. Any taxa that were left after the complete filtering process were considered as being present if they were represented by at least one read. This decision was made to ensure comparability to previous stomach flushing studies, wherein a single left or right otolith might have indicated the presence of a species in yellow-eyed penguin diet, provided the otolith was not eroded (van Heezik 1990, Moore and Wakelin 1997). We note that DNA metabarcoding has a similar temporal window to stomach flushing (1-4d; Deagle *et al.* 2010), therefore we expected that some taxa will have less reads than others within a sample, due to prey size, digestion efficiency, and differential digestion. In addition, we presented diet information at the highest taxonomic rank possible, therefore imposing a within-sample filter based on a percentage of the number of reads within each sample would result in a significant reduction in %FOO because of the larger number of taxa differentiated, compared to studies where a lower taxonomic rank, and thus fewer taxonomic groups were determined (e.g. identification to Class, Cavallo *et al.* 2018).

To give an indication of the sensitivity of our data to within-sample filtering, we present a range of filter options to arbitrarily limit %FOO. In Table S2, we indicate the differences in %FOO of the top ten ray-finned fishes, using a variety of within-sample filters: (a) excluding observations <10 reads per taxon within each sample was applied; (b) excluding observations using a threshold of <0.1% of all reads per sample, and; (c) excluding observations using a threshold of <1.0% of all reads per sample.

Table S2. Frequency of occurrence (%FOO) of the top ten ray-finned fish taxa detected in the diet of yellow-eyed penguins using DNA metabarcoding in this study. Subsequent within-sample filtering (excluding <10 reads per taxon within each sample, or excluding a taxon from a sample if the number of reads <0.1% - <1.0% of the sample's total) was not applied to the Results to ensure comparability with previous stomach flushing studies.

Ray-finned fish taxa	Common name	%FOO (this study)	%FOO excluding <10 reads	%FOO excluding <0.1% reads	%FOO excluding <1.0% reads
<i>Parapercis colias</i>	Blue cod/rāwaru	100.0	98.4	97.9	93.3
<i>Hemerocoetes sp.</i>	Opalfish/kohikohi	98.4	96.4	95.9	80.3
<i>Argentina elongata</i>	Silverside	90.7	83.9	82.9	57.5
<i>Congiopodus sp.</i>	Pigfish/purumorua	76.2	71.5	67.4	40.4
<i>Seriolella brama</i>	Blue warehou/warehou kahurangi	62.2	54.9	54.9	29.5
<i>Emmelichthys nitidus</i>	Red baitfish	62.2	56.0	52.8	21.2
<i>Pseudophycis bachus</i>	Red cod/hoka	61.1	56.0	53.9	30.1
<i>Sprattus antipodum</i>	Blueback sprat/kupae	47.7	45.6	42.0	21.2
<i>Auchenoceros punctatus</i>	Āhuru	21.2	17.1	17.1	7.8
<i>Sprattus muelleri</i>	New Zealand sprat/kupae	18.1	16.6	15.0	5.7

S3. Sequence analysis according to sample type and age class

Chordates

Following PCR using chordate primers, 218 processed samples had visible amplicons (n_{visamp} , Table S3), indicating that amplified DNA was present in 69.6% of all samples collected. Computational filtering to optimise species richness based on alpha rarefaction curves and exclusion of spurious species resulted in 193 processed samples that contained >2000 target ray-finned fish DNA sequences ($n_{>A2000}$), representing 61.7% of all samples collected and 88.5% of processed samples that contained target prey DNA sequences following PCR (Table S4.1). We detected a difference in the proportions of samples that were retained ($\chi^2 = 2.94$, $df = 3$, $\alpha = 0.05$, $p = 0.4$), with just 71.1 % of fresh samples from adults being retained after computational filtering.

Cephalopods

Following PCR using cephalopod primers, 106 samples containing visible amplicons remained after computational filtering ($n_{>T2000}$), indicating that amplified DNA was present in 33.9% of all samples collected. As cephalopods were filtered differently, based on the *a priori* understanding of the smaller contribution cephalopods make to yellow-eyed penguin diet compared to ray-finned fishes (van Heezik 1990a; Moore and Wakelin 1997), the number of processed samples that contained cephalopod DNA sequences after computational filtering was reduced to 48/106, representing 15.3% of all samples collected (Table S3). We detected no difference in the proportions of samples containing cephalopod DNA that were retained ($\chi^2 = 33.2$, $df = 3$, $\alpha = 0.05$, $p < 0.01$). The proportion of processed samples that were retained after computational filtering were lowest for fresh samples compared to latrines.

Table S3. Yellow-eyed penguin faecal samples collected ($n_{\text{collected}}$) from the Otago coast during the 2016/17 season, relative to age class and sample quality ($n = 313$). Computational filtering ($n_{>A2000}$ for ray-finned fishes, and $n_{>T2000}$ for cephalopods) was based on alpha rarefaction curves pooled across processed faecal samples that contained visible amplicons following PCR (n_{visamp}). Species and genus-level assignment was conducted on computationally filtered samples ($n_{>A2000}$ and $n_{>T2000}$) and presented in the Results.

Primer	Sample quality	Age class	$n_{\text{collected}}$	n_{visamp}	$n_{>A2000}$	%filtered
Chordates	Fresh	Adult	62	38	27	71.1 %
		Chick	24	13	13	100 %
	Latrine	Adult	158	103	101	98.1 %
		Chick	69	64	52	81.3 %
Total - Chordates			313	218	193	88.5 %
Primer	Sample quality	Age class	$n_{\text{collected}}$	n_{visamp}	$n_{>T2000}$	%filtered
Cephalopods	Fresh	Adult	62	22	7	31.8 %
		Chick	24	7	1	14.3 %
	Latrine	Adult	158	41	20	48.8 %
		Chick	69	36	20	55.6 %
Total - Cephalopods			313	106	48	45.3 %

S4. Removal of low-resolution taxa detected using DNA metabarcoding of yellow-eyed penguin faeces

Table S4. Raw sample occurrence (n_r), frequency of occurrence (% FOO_r), total counts of DNA sequence reads (I_r), and relative read abundance (% RRA_r) of lower resolution (class/order/family) taxa detected in yellow-eyed penguin faeces using DNA metabarcoding, prior to removal of these lower resolution taxa from the analyses (after computational filtering). The highest taxonomic rank identified is given in bold font.

Phylum	Class	Order	Family	n_r	% FOO_r	I_r	% RRA_r	Geographically plausible congeners
Chordates								
Chordata	Actinopterygii	Centrarchiformes	Cheilodactylidae (Moki and Tarakihi)	91	47.2	42838	1.30	<i>Nemadactylus macropterus</i>
Chordata	Actinopterygii	NA	Emmelichthyidae (Bonnetmouths)	1	0.5	215	<0.01	<i>Emmelichthys nitidus</i> <i>Plagiogeneion rubiginosum</i>
Chordata	Actinopterygii	Perciformes	Sebastidae (Sea perches)	10	0.5	2080	<0.01	<i>Helicolenus percoides</i>
Chordata	Actinopterygii	Scombriformes	Centrolophidae (Warehou)	21	0.5	1900	<0.01	<i>Seriolella brama</i> <i>S. caerulea</i> <i>S. punctata</i>
Chordata	Actinopterygii	Perciformes (Perch-like fishes)	-	1	10.9	193	0.05	-
Chordata	Actinopterygii	-	-	1	5.2	378	0.11	-
Total: All ray finned fish identified below genus-level (discarded)				108	56.0	47604	1.49	
Total: All ray finned fish identified above genus-level (used in Results)				193	100	2736711	98.51	

Phylum	Class	Order	Family	n_r	% FOO_r	I_r	% RRA_r	Geographically plausible congeners
Cephalopods								
Mollusca	Cephalopoda	Teuthida	Octopoteuthidae (Squid)	3	2.8	93	2.46	<i>Taningia danae</i> <i>Octopoteuthis megaptera</i>
Mollusca	Cephalopoda	Teuthida (Squid)	-	1	0.9	187	0.98	-
Total: All cephalopods identified below genus-level (discarded)				4	3.8	280	3.44	
Total: All cephalopods identified above genus-level (used in Results)				51	48.1	10514	96.56	

S5. Non-target taxa amplified using 16S Cephalopod primers

Table S5. Non-target taxa amplified using 16S Cephalopod primers, after clustering. No computational filtering has been applied. 64.8% of the non-target sequences that were amplified are attributed to ray-finned fishes, and 23.7% were unassigned. Asterisks indicate potential field or laboratory contamination.

Phylum	Class	Total number of reads
(Unassigned)	-	248842
Annelida	Clitellata	97
Annelida	Polychaeta	24
Arthropoda	-	22617
Arthropoda	Arachnida	22
Arthropoda	Collembola	9
Arthropoda	Hexanauplia	10
Arthropoda	Insecta	21082
Arthropoda	Malacostraca	32
Chordata	-	2206
Chordata	Actinopterygii	679727
Chordata	Amphibia*	766
Chordata	Aves*	18503
Chordata	Mammalia*	51642
Cnidaria	Scyphozoa	82
Echinodermata	-	10
Echinodermata	Holothuroidea	34
Mollusca	Gastropoda	17
Nemertea	Enopla	5
Nemertea	Pilidiophora	93
Total non-target taxa DNA sequence reads		1049227

S6. DNA sequence alignment of the yellow-eyed penguin blocking primer and the 16S_SHORT_Chord primers



Figure S1. DNA sequence alignment of the yellow-eyed penguin *Megadyptes antipodes* blocking primer used in this study against yellow-eyed penguin 16S sequence (Accession No: DQ137158), the 16S_SHORT_Chord_F and R primers, and the little penguin *Eudyptula minor* blocking primer used by Deagle *et al.* (2010).

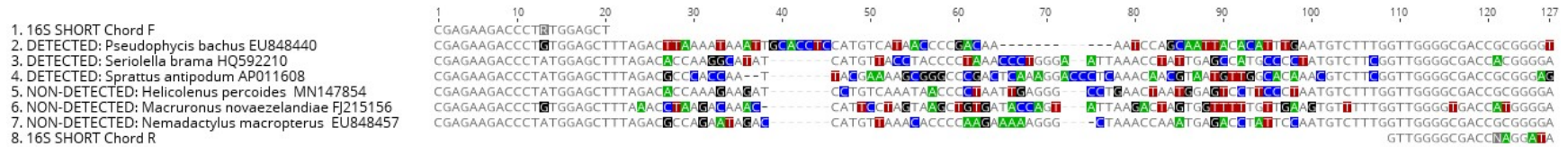


Figure S2. DNA sequence alignment of the 16S_SHORT_Chord_F and 16S_SHORT_Chord_R primers with 16S sequence (including Accessions Numbers) for three ray-finned fish species that were detected in this study (red cod *Pseudophycis bachus*, blue warehou *Seriolaella brama*, and New Zealand blueback sprat *Sprattus antipodum*). Three species previously identified in yellow-eyed penguin diet that were not detected in this study (Jock Stewart *Helicolenus percoides*, hoki *Macruronus novaezelandiae*, and tarakihi *Nemadactylus macropterus*) are also presented, to indicate primer sequence similarity with detected species.

S7. Alpha rarefaction curves used to filter samples

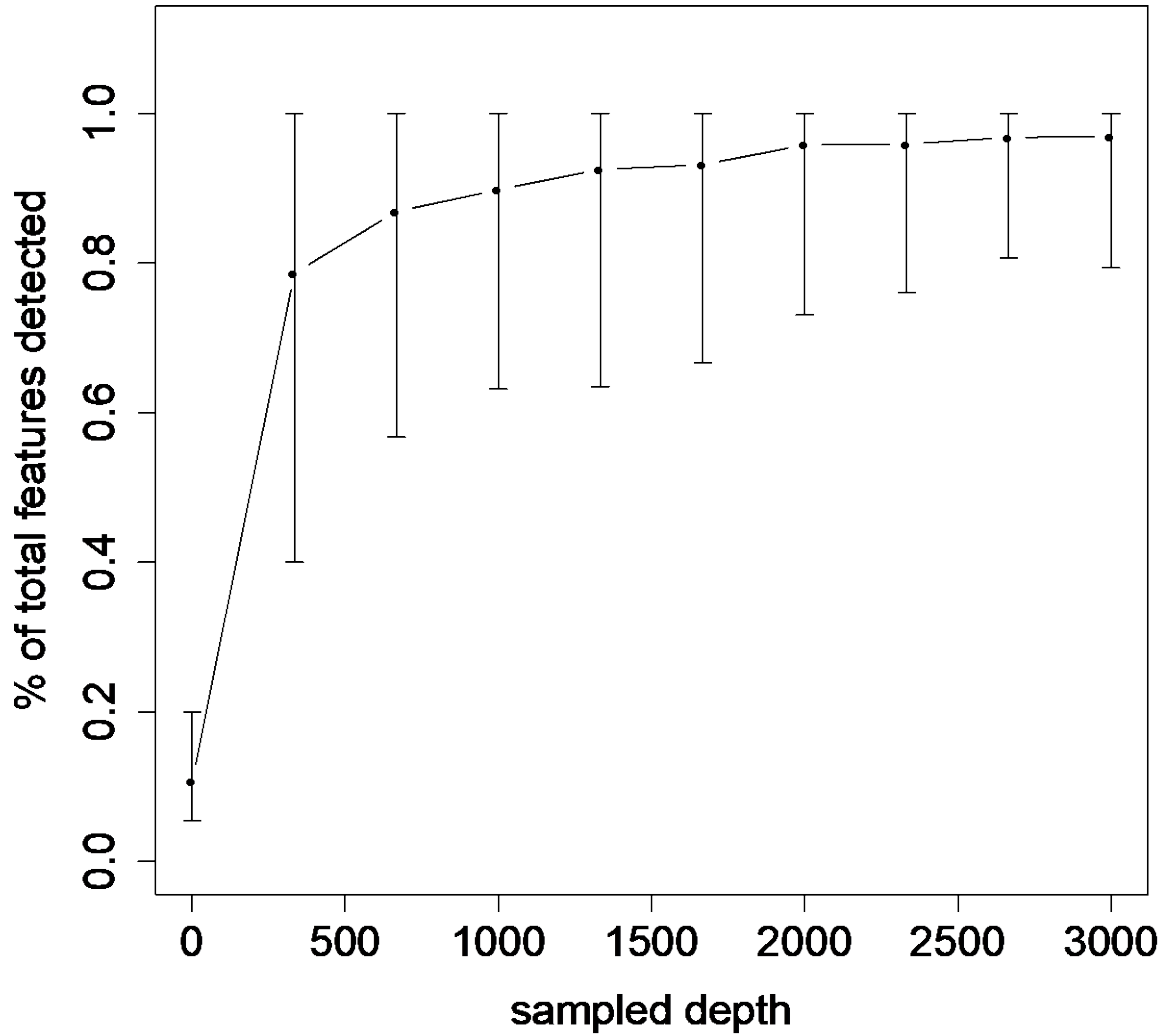


Figure S3. Alpha rarefaction curve for ray-finned fish DNA sequences amplified in yellow-eyed penguin faeces based on re-sampling 184 samples with more than 3000 reads once at 10 different sampling depths. The y-axis represents the mean percentage of the total features (i.e. independent types of sequences) within each sample detected at each sequence count depth (x-axis). The error bars delimit the 95% confidence intervals.