THE UNIVERSITY OF AUCKLAND
FACULTY OF SCIENCE

uniservices

c/- The University of Auckland
Private Bag 92019, AMC
Auckland 1142, New Zealand
www.uniservices.co.nz

Auckland UniServices Limited
Level 10, UniServices House,
70 Symonds Street, Auckland
+64 9 373 7522

Determining the diet of New Zealand king shag using DNA metabarcoding

New Zealand King Shag, (*Leucocarbo carunculatus*) on Blumine Island, Marlborough Sounds, New Zealand in 2016 (Wikipedia commons).

Aimee van der Reis & Andrew Jeffs

Report Prepared For: Department of Conservation, Conservation Services Programme, Project BCBC2019-05. DOC Marine Science Advisors Graeme Taylor and Dr Karen Middlemiss.

November 2020

INTRODUCTION

The New Zealand king shag (*Leucocarbo carunculatus*) is an endemic seabird that is classed as nationally endangered (Miskelly et al., 2008). The population is confined to a small number of colonies located around the coastal margins of the outer Marlborough Sounds (South Island, New Zealand); with surveys suggesting the population is currently stable (~800 individuals surveyed in 2020; Aquaculture New Zealand, 2020; Schuckard et al., 2015). Monitoring the colonies has become a priority and research is being conducted to better understand their population dynamics and basic ecology to improve the management of the population, particularly in relation to human activities such as fishing, aquaculture and land use (Fisher & Boren, 2012).

The diet of the New Zealand king shag is strongly linked to the waters surrounding their colonies and it has been suggested that anthropogenic activities, such as marine farm structures, may displace foraging habitat that could affect the population of New Zealand king shag (Fisher & Boren, 2012). However, GPS tracking of king shag movements has shown that these birds roost and forage around mussel farms (Aquaculture New Zealand, 2020).

The objective of the present work was to examine the diet of the New Zealand king shag using a next-generation sequencing method (NGS), DNA metabarcoding. This molecular method uses 'DNA barcoding' in conjunction with NGS to reliably identify a variety of species from a single sample (e.g., from regurgitated king shag pellets containing mostly indigestible components of the diet).

The DNA is extracted from the sample and the species 'barcode' region is amplified by universal primers that are gene specific and have the ability to amplify short sequences from a broad taxonomic range (Deagle et al., 2014; Kress et al., 2015).

The genes that are targeted by such primers are usually housekeeping genes (e.g., mitochondrial cytochrome *c* oxidase one - COI) that have been conserved in most species over time, but contain enough DNA sequence variation to allow discrimination between closely related species, thus providing each species with a unique 'barcode'. This unique barcode makes it possible to identify organisms to species level when sequences are matched to a DNA repository, after having undergone quality control and filtering using bioinformatics software such as Qiime 2 (Bolyen et al., 2019).

DNA metabarcoding has provided a solution to the limitations of the highly labour intensive traditional method of using microscopy of gut contents. These traditional methods also suffer from difficulties in reliably identifying degraded remnants of food items. The major benefits of DNA metabarcoding include the ability to identify food species regardless of the state of the gut content, thus enabling the identification of food species in regurgitated pellets and faecal matter, while traditional methods require the food items to be largely intact or include recognisable hard parts that survive digestive processes. Furthermore, this molecular method does not require taxonomic expertise to obtain species level identification of food items and is relatively rapid, cost-effective and accurate. In contrast, microscopic analyses of gut contents can provide counts of individual food items where persistent diagnostic elements remain in the gut contents, such as squid beaks, whereas currently DNA metabarcoding can only reliably provide presence/absence data for food species in gut contents.

A preliminary king shag DNA metabarcoding diet study (7 regurgitated pellets and 5 faecal samples) was conducted in late 2019. This allowed for molecular protocols to be optimized, which included strict sterilisation of working areas to avoid contamination and optimization of PCR protocols to avoid erroneous over amplification. A list was also compiled of likely dietary species and whether or not they were present in GenBank (Benson et al., 2013), if they were not efforts were made to obtain species of interest so they could be added as a reference in the DNA databases. After some work in developing the DNA recovery methods, the preliminary results proved successful for both pellet and faecal samples, allowing for the study to be expanded. The main study was subsequently focused on regurgitated pellets rather than faecal samples, because the pellets represented discrete feeding periods for the birds and were easy to collect intact, minimising contamination. However, faecal samples are of value and provide a possible route for dietary analysis in future studies, as the DNA extraction methods were sufficiently robust to allow the main dietary items to be detected in both regurgitated pellets and faecal matter.

The objectives of this expanded study (185 pellets) was to; 1) determine the diet of the king shag, 2) identify the sex of the king shag that had regurgitated the pellet, and 3) investigate whether differences in the diet may occur due to area (i.e., differences among colonies) or between sexes.

MATERIALS AND METHODS

King shag collection and DNA amplification

Regurgitated pellets from New Zealand king shag were collected between March 2019 and March 2020 from seven colonies in the Marlborough Sounds (South Island, New Zealand) by Mike Bell, Wildlife Management International Limited, whilst visiting colonies for other research activities. A total of 28, 43, 47, 42, 24, 10 and 23 pellets were collected from Blumine, Duffers Reef, North Trio, Tawhitinui, The Twins, The Haystack and White Rocks, respectively (Figure 1; Table 1). The samples were placed in 90% ethanol immediately upon collection to best preserve the DNA for identifying taxa in gut content using molecular methods.

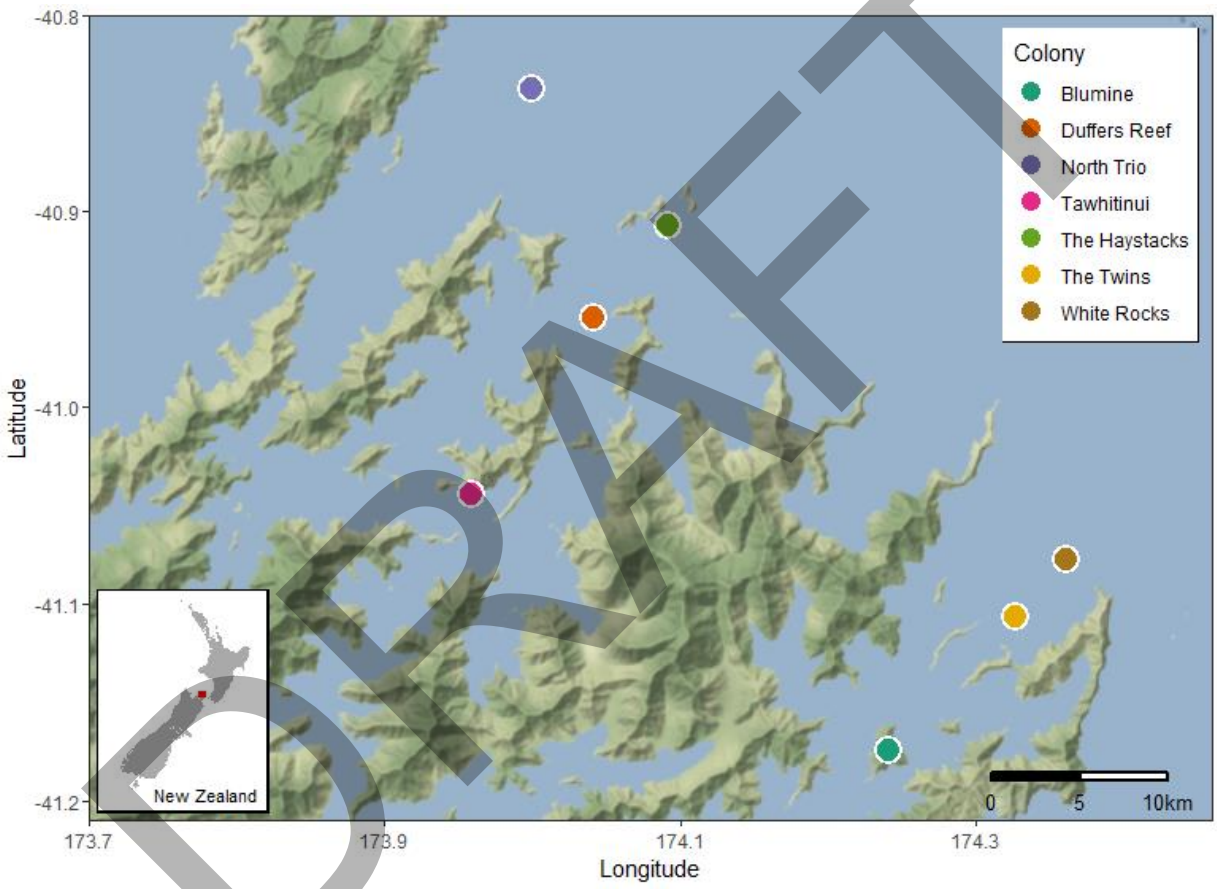


Figure 1. Regurgitated pellets collected from seven New Zealand king shag colonies in the Marlborough Sounds. ***The location of the Marlborough Sounds, is identified by the red polygon on the inset map, and are located at the top of the South Island, New Zealand.***

Table 1. The number of regurgitated pellets collected from seven New Zealand king shag colonies in the Marlborough Sounds and the number used for DNA metabarcoding.

Colony	Mar 19	Nov 19	Mar 20	Total (collected)	Total (metabarcoding)
Blumine			28	28	28
Duffers Reef	3	40		43	33
North Trio	6	41		47	34
Tawhitinui	14	28		42	33
The Twins			24	24	24
The Haystack	10			10	10
White Rocks	8		15	23	23

The pellets were dissected at the University of Auckland and the tissue material was separated from the hard items (e.g., bones and shells) allowing for a homogenized subset of gut digesta tissue to be used for DNA extractions. E.Z.N.A.® Mollusc DNA Kit (Omega Bio-Tek Inc, Georgia, USA) was used for the extractions, following the manufacturer's protocol, to remove mucopolysaccharides associated with marine invertebrates which are known to interfere with DNA extraction and amplification (Palmer, 2008). DNA extractions were done in sets according to king shag colony and the polymerase chain reactions (PCR) were done using MyTaq Red Mix (Bioline, London, UK) master mix; 7 µl MyTaq Red Mix, 0.5 µl of each primer, 5 µl UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen - Thermo Fisher Scientific, Massachusetts, USA), 1 µl DNA and 2 µl BSA (1%) when necessary for optimal DNA amplification per reaction. Negative controls were included in every set of DNA extractions (extraction blank - no tissue added) and every PCR run (PCR blank - no DNA added) to check for possible contamination.

The sex of the king shag that had regurgitated the pellet was identified using a non-ratite universal primer pair (2550 and 2718; Fridolfsson & Ellegren, 1999) which produces a single band for males (600 bp) and a double band for females (450 bp and 600 bp). PCR protocol: 94 ºC - 4 min, 8×[touchdown: 94 ºC - 30 s, 57 ºC - 45 s, 72 ºC – 45 s], 30×[standard: 94 ºC - 30 s, 51 ºC – 45 s, 72 ºC - 45 s], 74 ºC - 5 min. The PCR products were run on a 3% agarose gel, and visualised using Gel Red (Biotium, Fremont, California, USA), in a Gel DocTM XR+ (Bio-Rad Laboratories Inc., California, USA).

For DNA metabarcoding of recovered dietary DNA, a mitochondrial DNA universal primer pair was used, Illumina Nextera™ library adapters (NexAd) were added (Illumina, San Diego, CA, USA), which targeted a portion of the COI region (313 bp; mlCOIintF; Leray et al., 2013 and jgHCO2198; Geller et al., 2013). PCR protocol: 94 ºC – 4 min, 30×[standard: 94 ºC - 30s, 45 ºC - 30s, 72 ºC - 1 min], 72 ºC - 5 min. The PCR products were run on a 1.6% agarose gel and visualised as above.

DNA metabarcoding

Selection of samples for metabarcoding

Overall, 185 pellet samples were selected for DNA metabarcoding: 28 from Blumine, 33 from Duffers Reef, 34 from North Trio, 33 from Tawhitinui, 24 from The Twins, 10 from The Haystack and 23 from White Rocks (Table 1). In cases where samples need to be selected from a larger collection, selection was based on the pellet's appearance (e.g., fresh or dried out), quantity of hard parts and DNA amplification success. Seven negative controls (one control per DNA extraction set/colony) were included in the metabarcoding run to monitor possible atmospheric contamination. PCRs were performed in triplicate for all samples and for DNA extraction negative controls. The PCR triplicates were pooled together before DNA clean-up proceeded.

DNA clean-up and pooling

Agencourt AMPure XP (Beckman Coulter, Brea, CA, USA) was used following the Illumina (Illumina, San Diego, CA, USA) protocol for PCR clean-up (Illumina, 2013). The concentration of the purified PCR products was determined using Qubit™ dsDNA HS Assay Kit (Invitrogen™, Thermo Fisher Scientific Inc., Waltham, MA, USA) following the manufacturer's instructions. The PCR products were brought to equal molarity, 2 ng μl^{-1} where possible. Sequencing was done through Auckland Genomics (Auckland, New Zealand) where indexing, using the Nextera™ DNA library Prep Kit and the second round of PCR clean-up occurred before sequencing on an Illumina MiSeq™ System (2×300 pair-end; single lane was used).

Metabarcoding protocol

Raw demultiplexed sequencing data was received from Auckland Genomics and Cutadapt v2.3 (Martin, 2011) was used to trim the forward and reverse primers from the sequences. Primers were removed if an exact sequence match could be found, and no indels or errors were allowed in the primer sequence and all untrimmed sequences were discarded. Qiime 2 (version 2020.8; Bolyen et al., 2019) was used to visualize the initial sequence quality. DADA2 (within Qiime 2; Callahan et al., 2016) was used for sequence filtering based on quality scores, denoising, merging and chimera formation to ensure only high quality paired-end sequences were retained (Table 2). The Naïve Bayes Classifier was used for assigning taxonomy at minimum confidence threshold of 70% confidence from the curated database, Midori (COI database - GenBankRelease239 as of August 2020; Machida et al., 2017). The Genbank database (2020-07-v5; Benson et al., 2013) was also used for assignment, using the megablast option BLASTn (Morgulis et al., 2008) with an e-value threshold of 0.001 and a percentage identity of 80%.

The resulting sequences were further filtered and analysed with R Studio® (version 3.6.1; R Core Team R Core Team, 2018). Sequences were retained if they were ≥250 bp and ≤350 bp. The results from the databases were consolidated to create one amplicon sequence variant (ASV; Callahan et al., 2017) taxonomic assignment list. ASVs assigned to class "Aves" (classified as host DNA) were filtered out from the dataset and ASVs identified as contamination were removed (Table 2).

Gut content analyses

Two samples (NT28 and TT18) were removed during sequence quality filtering and the seven DNA negatives were removed, resulting in 183 samples for gut content analyses. The dataset was transformed into a more reliable and conserved binary dataset (i.e., presence-absence). All results (i.e., figures) were produced using ggplot2 version 3.2.1 (R package; Wickham, 2017).

The binary dataset was used to investigate the frequency of fish species, as fish are known to be the main dietary target of New Zealand king shags (Lalas & Brown 1998). An analysis of variance (ANOVA) was used to assess diversity of fish species among colonies and Tukey's HSD was used to find means of fish diversity that were significantly different among colonies. A non-metric multidimensional scaling (NMDS) and a permutational multivariate analysis of variance (PERMANOVA) using distance matrices were performed using Vegan (version 2.5-6; Oksanen et al., 2019) to visualize and test if differences occurred among colonies and between sexes. As significant differences were detected, a generalized linear model (GLM) was run. The resulting data was then explored to identify which fish species were different, using emmeans (version 1.4.3; Lenth, 2019).

RESULTS

Sex ID

Host DNA was sufficient in the majority of extracted pellet content DNA samples to determine the sex of the king shag that had regurgitated the pellet. Only 7 pellets could not be sexed. In total 68 pellets were regurgitated by females and 108 from males.

Sequence filtering and taxonomic assignment

A total of 884 ASVs passed filtering (650,034 reads; Table 2) and were assigned to 30 Phyla (Table 3). The taxonomic composition of each sample varied (Figure 2), but overall seven phyla were identified to be in >10% of the pellets: Annelida (13%; 22 ASVs; segmented worms), Arthropoda (58%; 239 ASVs), Ascomycota (16%; 73 ASVs; fungi), Bacteroidetes (42%; 107 ASVs; bacteria), Chordata (99%; 240 ASVs), Nematoda (40%; 74 ASVs; roundworms) and Platyhelminthes (12%; 19 ASVs; flatworms). Overall worms (Annelida, Acanthocephala, Nematoda, Nemertea and Platyhelminthes) occurred in 95 samples (52%).

Table 2. Stepwise filtering of reads using Qiime 2 and R.

Filtering Steps	Reads after filtering
Initial (no filtering)	9,470,138
Cutadapt	8,686,874
Quality score	5,192,759
Denoised	5,173,616
Merged	5,148,334
Non-chimeric	5,142,653
Length (250-350 bp)	4,994,133
Host (Class: Aves)	650,482
DNA negative assessment	650,169
Sequence confidence assignment	650,034

Table 3. Phyla identified from king shag regurgitated pellets. ***Phyla that occurred in >10% of the pellets are in bold. Amplicon sequence variants – ASVs.***

Phyla	ASVs	Frequency of occurrence (n = 183)	Frequency of occurrence (%)
Acanthocephala	7	4	2.19
Amoebozoa	8	4	2.19
Annelida	**22**	**24**	**13.11**
Apicomplexa	2	4	2.19
Arthropoda	**239**	**107**	**58.47**
Ascomycota	**73**	**29**	**15.85**
Bacteroidetes	**107**	**77**	**42.08**
Basidiomycota	8	6	3.28
Blastocladiomycota	1	7	3.83
Bryozoa	6	6	3.28
Chlorophyta	1	1	0.55
Chordata	**240**	**181**	**98.91**
Cnidaria	11	14	7.65
Echinodermata	4	4	2.19
Euglenozoa	1	1	0.55
Magnoliophyta	1	1	0.55
Mollusca	5	12	6.56
Myzozoa	2	2	1.09
Nematoda	**74**	**73**	**39.89**
Nemertea	2	2	1.09
Ochrophyta	10	11	6.01
Oomycota	3	3	1.64
Platyhelminthes	**19**	**22**	**12.02**
Porifera	7	7	3.83
Proteobacteria	17	13	7.10
Protozoa	1	1	0.55
Rhodophyta	4	5	2.73
Rotifera	4	1	0.55
Stramenopiles	3	3	1.64
Zygomycota	2	3	1.64

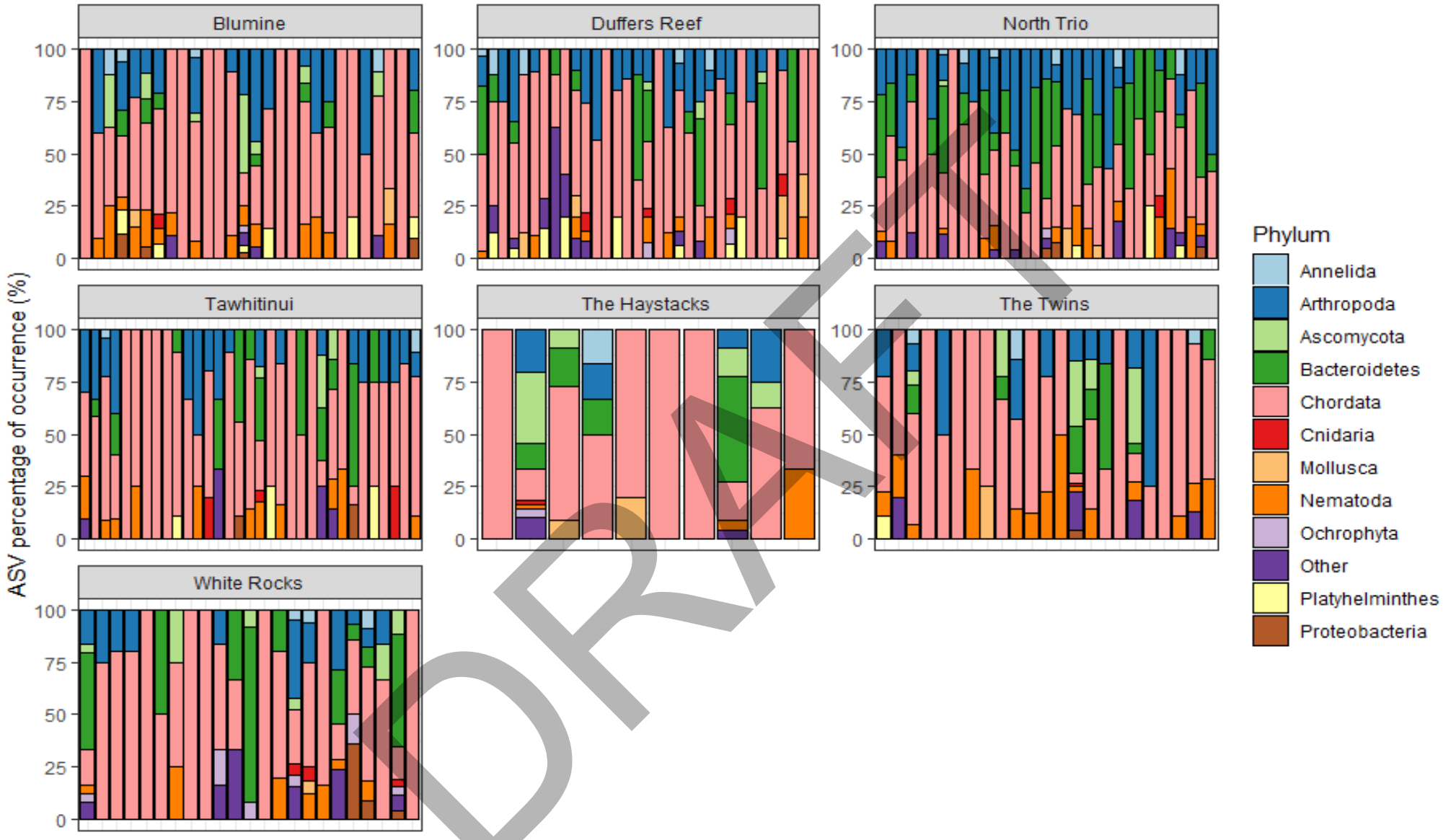


Figure 2. Taxonomic composition at phyla level for each New Zealand king shag regurgitated pellet, group by colony of collection. ***'Other' represents phyla that made up less than 5% of the gut content of a pellet. Percentage occurrence is the proportion of total amplicon sequence variants (ASVs) per pellet.***

Dietary items

More than 300 species were identified from the pellets (Table S1), of which 32 were fish (Table 4; Figure 3). Fish (class: Actinopteri and Actinopterygii), considered to be the main New Zealand king shag dietary item, occurred in 181 samples (99%). The overall diversity of fish species found in king shag pellet samples was significantly different (p-value <0.05) among colonies, however, only Duffers Reef (median: 4 species per pellet) vs. Blumine (median: 2 species per pellet) was significantly different out of the pairwise combinations (Figure 4). The NMDS and PERMANOVAS revealed weak yet significant clustering for colony (Blumine: 28, Duffers Reef: 33, North Trio: 33, Tawhitinui: 33, The Twins: 23, The Haystack: 10 and White Rocks: 23) and sex (F: 68 and M: 108) (p-value < 0.05; Appendix 1).

Lefteyed flounders (Bothidae) were found to be the dominant species occurring in the pellets. Witch (*Arnoglossus scapha*) had the highest frequency of occurrence (122 samples), followed by crested flounder (*Lophonectes gallus*[1]; 79 samples) (Table 4; Figure 3). Crested flounder was found more commonly in pellets sampled from North Trio (found to have a significant difference when compared to Tawhitinui, The Haystacks, The Twins and White Rocks), while witch was found to be more common in pellets from Blumine (found to have a significant difference when compared to Duffers Reef, Tawhitinui and The Haystacks) (Table 5; Figure 4 and 5). A significant difference for the smooth leather jacket (*Meuschenia scaber*) was detected between colony pairwise comparisons Blumine vs. Duffers Reef and Blumine vs. The Haystacks, with <5% being detected in pellets collected from Blumine (Table 5).

Malacostraca (class) were identified in 81 pellets and some potential dietary species were identified including shrimps (e.g., *Pterygosquilla schizodontia* - 1 sample) and crabs (e.g., *Munida gregaria* - 12 samples, *Nectocarcinus bennetti* - 5 samples and *Halicarcinus* spp. - 4 samples). *Octopus huttoni* (class: Cephalopoda) is another likely dietary item and was identified in 9 samples.

Witch, scaly gurnard (*Lepidotrigla brachyoptera*), smooth leather jacket and speckled sole (*Peltorhamphus latus*) were found to be significantly different in the diet when comparing females and males (Table 5). Only smooth leather jacket was identified to have a higher frequency of occurrence in males. The other three fish species had a higher occurrence in females (Table 5).

Non-Dietary Items[2]

Many non-dietary items were detected which included more prevalent species found in the pellets such as bacteria (e.g., *Flavobacterium* spp. - 58 pellets) and parasitic worms (e.g., *Contracaecum osculatum* – 20 pellets), but also infrequent species likely consumed as secondary dietary items or by chance, such as a brittle star (*Amphiura* magellanica - 1 pellet) and brown seaweed (*Carpomitra costata* – 3 pellets).

[1] The crested flounder sequence matched the Australian species, *L. gallus*, but it is likely that the sequence will also match the New Zealand crested flounder, *L. mongonuiensis*, when added to the DNA reference databases (McMillan et al., 2019).

[2] Please note that non-dietary items were only briefly analysed and discussed in this report.

Table 4. Species of fish detected in the king shag regurgitated pellets. *Species in bold are likely not the correct match*.

Family	Genus	Species	Common name	Frequency of occurrence (n = 183)	Frequency of occurrence (%)
Bothidae	*Arnoglossus*	*Arnoglossus scapha*	Witch	122	66.67
Bothidae	***Lophonectes***	***Lophonectes gallus***	**Crested flounder**	**79**	**43.17**
Rhombosoleidae	*Pelotretis*	*Pelotretis flavilatus*	Southern lemon sole	66	36.07
Monacanthidae	*Meuschenia*	*Meuschenia scaber*	Smooth leatherjacket	53	28.96
Rhombosoleidae	*Peltorhamphus*	*Peltorhamphus latus*	Speckled sole	47	25.68
Percophidae	*Hemerocoetes*	*Hemerocoetes morelandi*	Duckbill fish	42	22.95
Sebastidae	*Helicolenus*	*Helicolenus percoides*	Red gurnard perch	38	20.77
Triglidae	*Lepidotrigla*	*Lepidotrigla brachyoptera*	Scaly gurnard	32	17.49
Labridae	***Notolabrus***	***Notolabrus tetricus***	**Blue-throated wrasse**	**18**	**9.84**
Tripterygiidae	*Matanui*	*Matanui profundum*	Deepwater triplefin	13	7.10
Scorpaenidae	*Scorpaena*	*Scorpaena papillosa*	Red scorpianfish	10	5.46
Pinguipedidae	*Parapercis*	*Parapercis colias*	Blue cod	9	4.92
Rhombosoleidae	*Peltorhamphus*	*Peltorhamphus novaezeelandiae*	New Zealand sole	7	3.83
Rhombosoleidae	*Rhombosolea*	*Rhombosolea tapirina*	Greenback flounder	6	3.28
Percidae	***Etheostoma***	***Etheostoma radiosum***	**Freshwater species**	**5**	2.73
Syngnathidae	*Hippocampus*	*Hippocampus abdominalis*	Big-belly seahorse	5	2.73
Moridae	*Pseudophycis*	*Pseudophycis bachus*	Red codling	4	2.19
Labridae	*Pseudolabrus*	*Pseudolabrus miles*	Scarlet wrasse	3	1.64
Moridae	*Pseudophycis*	*Pseudophycis breviuscula*	Northern bastard codling	3	1.64
Serranidae	*Caesioperca*	*Caesioperca lepidoptera*	Butterfly perch	3	1.64
Zeidae	*Zeus*	*Zeus faber*	John Dory	3	1.64
Carangidae	*Trachurus*	*Trachurus novaezelandiae*	Yellowtail horse mackerel	2	1.09
Pomacentridae	***Pomachromis***	***Pomachromis fuscidorsalis***	**Tropical species**	**2**	1.09
Centrolophidae	*Seriolella*	*Seriolella brama*	Blue warehou	1	0.55
Congiopodidae	*Congiopodus*	*Congiopodus coriaceus*	Deepsea pigfish	1	0.55
Merlucciidae	*Macruronus*	*Macruronus novaezelandiae*	Blue grenadier	1	0.55
Percophidae	*Hemerocoetes*	*Hemerocoetes artus*	Duckbill fish	1	0.55
Serranidae	***Plectranthias***	***Plectranthias winniensis***	**Tropical species**	**1**	0.55
Serranidae	***Plectropomus***	***Plectropomus leopardus***	**Tropical species**	**1**	0.55
Tripterygiidae	*Blennodon*	*Blennodon dorsalis*	Giant triplefin	1	0.55
Tripterygiidae	*Forsterygion*	*Forsterygion flavonigrum*	Yellow-and-black triplefin	1	0.55
Uranoscopidae	*Kathetostoma*	*Kathetostoma giganteum*	Giant stargazer	1	0.55

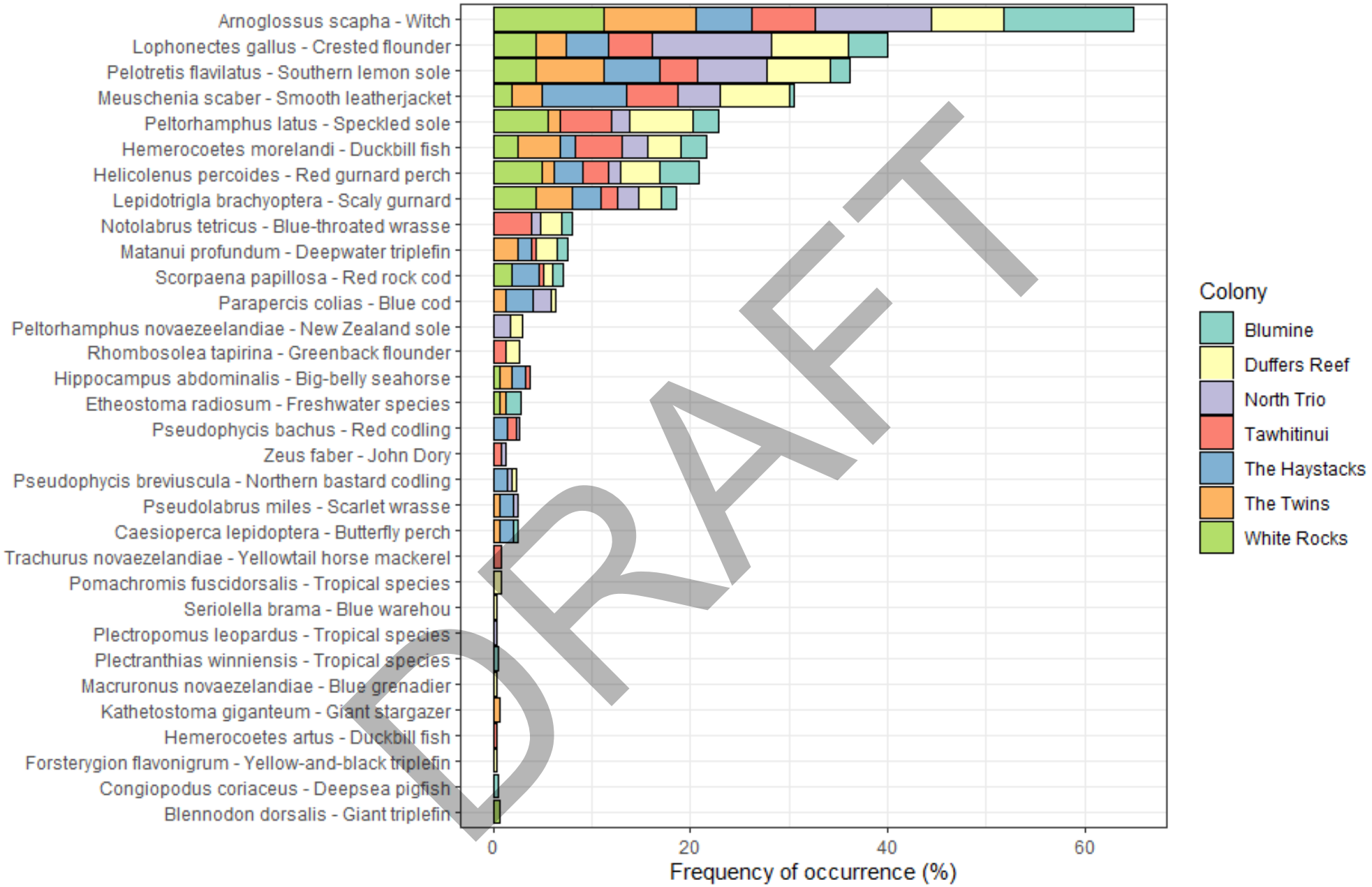


Figure 3. Frequency of occurrence (%) (presence-absence) of fish species detected among colonies in the regurgitated pellets from king shags.

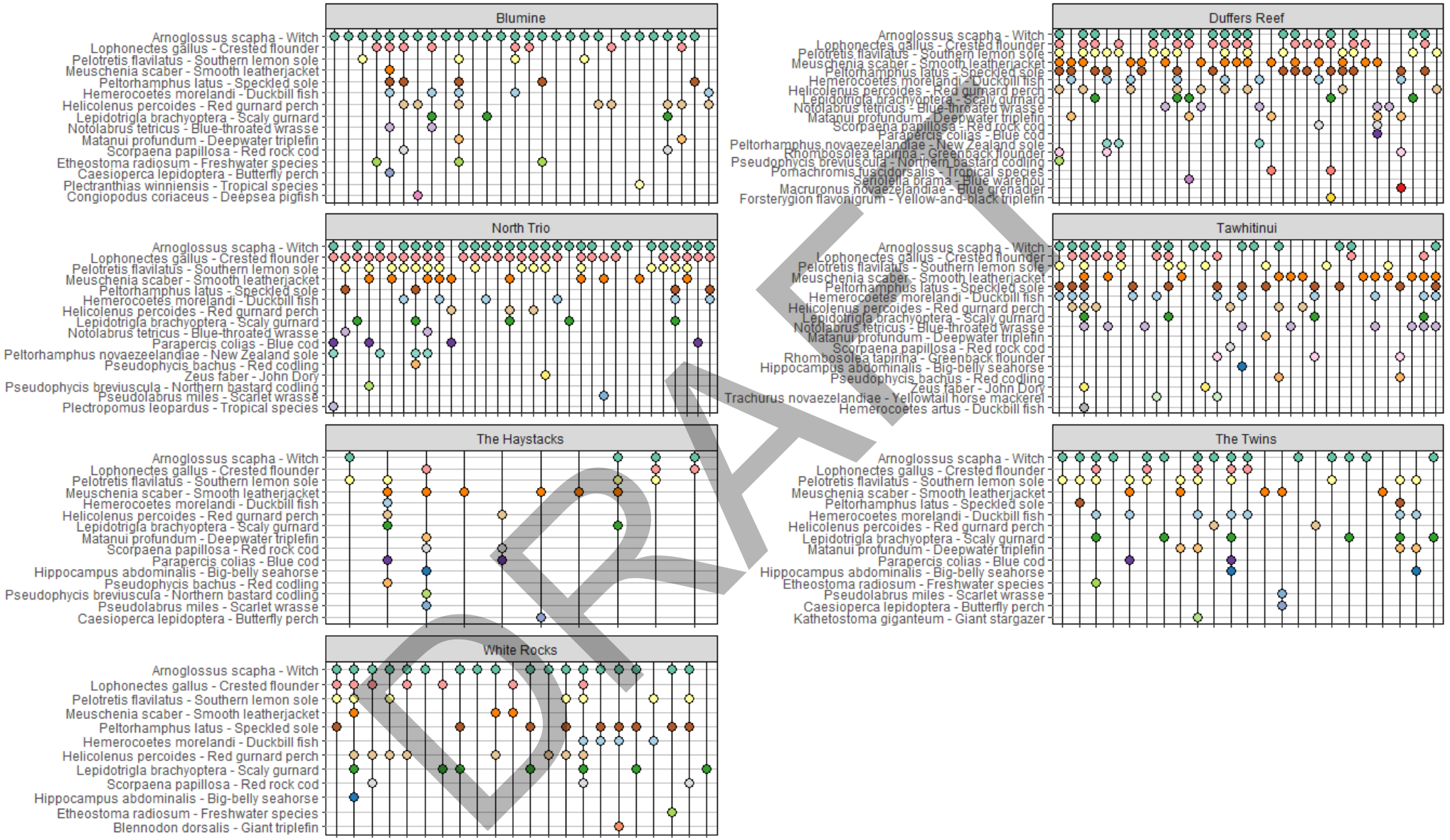


Figure 4. Presence-absence plot of fish species found in every king shag pellet, grouped by colony. ***Every tick along the x axis represents a pellet and every point along the y axis represents the presence of a species of fish. The fish species are colour coded.***

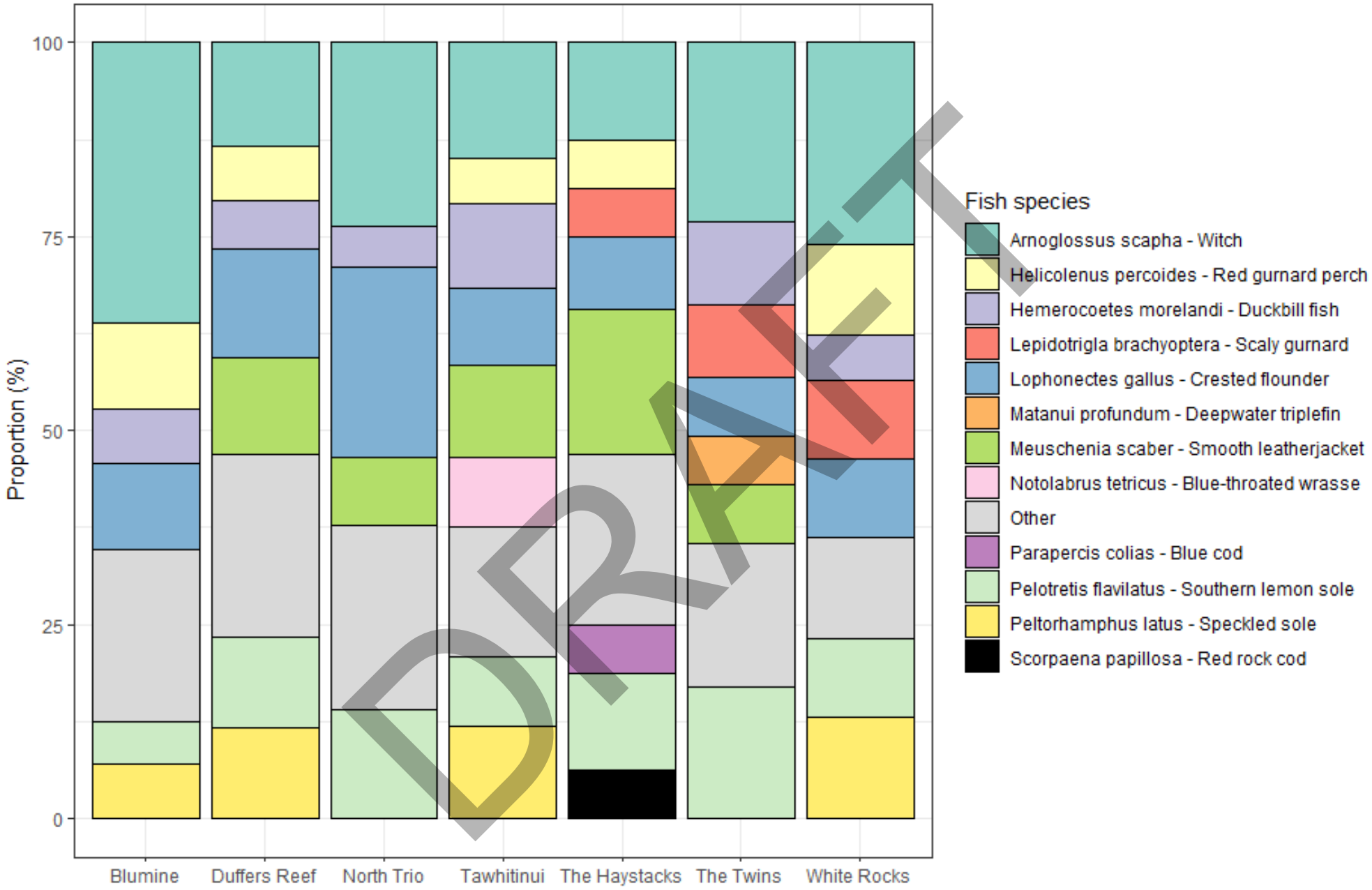


Figure 5. The proportion of dietary items detected in each king shag colony. ***If a species occurred in less than 5% of pellets it was classed as 'other'.***

Table 5. Fish species identified from king shag pellets to be significantly different (p-value <0.05) among colonies and between sex. ***The item within each pairwise comparison that has the greater frequency of occurrence (%) is indicated with an asterisk.***

Variable	Species – common name	Pairwise comparison	p-value
Colony	*Arnoglossus scapha* - witch	Blumine* vs. Duffers Reef	0.034
		Blumine* vs. Tawhitinui	0.017
		Blumine* vs. The Haystacks	0.038
	Lophonectes gallus - Crested flounder	Blumine vs. North Trio*	<0.001
		North Trio* vs. Tawhitinui	<0.001
		North Trio* vs. The Haystacks	0.038
		North Trio* vs. The Twins	<0.001
		North Trio* vs. White Rocks	0.004
	Meuschenia scaber - Smooth leatherjacket	Blumine vs. Duffers Reef*	0.042
		Blumine vs. The Haystacks*	0.035
Sex	*Arnoglossus scapha* - witch	F* vs. M	0.010
	Lepidotrigla brachyoptera - Scaly gurnard	F* vs. M	0.042
	Meuschenia scaber - Smooth leatherjacket	F vs. M*	<0.001
	Peltorhamphus latus - Speckled sole	F* vs. M	<0.001

Discussion

This is the first study that has used DNA metabarcoding to determine the diet of the New Zealand king shag and is one of very few DNA metabarcoding bird diet studies undertaken in New Zealand. The frequency of occurrence of fish species detected is comparable to the microscopic study, using otoliths and exoskeletal remains, run in parallel on the same samples (Lalas and Schuckard 2020 in pers. comm.). This comparison provides further proof that host DNA does not impede the diversity of species detected and thus is a very effective method to determine diet with ease (Devloo-Delva et al., 2018; Piñol et al., 2014). The host DNA was effectively utilised to sex the majority of the birds that regurgitated the pellet in this study, as New Zealand king shags do not prey on other birds and thus no other bird DNA would have been present in the pellets.

The diet of New Zealand king shags is of interest due to it being nationally endangered. It has been suggested that marine farm sites may potentially affect their prey species through changes to the benthic environment. The significant differences identified in the fish present in the regurgitated pellets among the colonies and between sexes are likely due to differences in the geographic location of the colonies and their associated feeding habitats. Although GPS tracking of king shags indicates that individuals will forage over extensive areas of coast.

New Zealand king shags primarily rely on Bothidae, witch (*A. scapha*) and crested flounder, as their main dietary items making up 35% of the total fish detected in pellets. However, they also consume a range of other fish species accounting of the remaining 65% of fish detected in the pellets. In addition, a number of mobile non-fish species were detected and may be minor prey items or they may be secondary dietary items (i.e., the gut contents of consumed fish), such as octopus and crab.

The crested flounder has not been previously mentioned as king shag prey, possibly due to its otoliths being mistaken for that of small witch. The DNA sequences for these two species vary by ~15% and the difference in identity was confirmed by another curated database, BOLD (https://www.boldsystems.org/). The DNA results provide strong evidence that witch and crested flounder were both consumed. Sequences from the pellet contents match crested flounder with 99.7% confidence, but when matched to witch the confidence dropped by ~10% and vice versa. Thus, these two species of fish are reliably distinguished using DNA.

The majority of dietary taxa matched previous records from regurgitated pellets and published reports based on incidental observations (e.g., *M. gregaria* - lobster krill and *Scorpaena papillosa* - red scorpionfish) (Lalas & Brown, 1998). Other species that have not been previously mentioned as being possible dietary items but were found to be in this study include John Dory (*Zeus faber*) and the big-belly seahorse (*Hippocampus abdominalis*). Many sequences were assigned to crabs, but the majority of species identified were not of high confidence due to New Zealand crab species being poorly referenced in DNA databases (Dowle et al., 2016; van der Reis et al., 2018; Zaiko et al., 2016). Overall, the results suggest king shags are opportunistic generalists having a wide diversity of fish species in their diet, but benthic fish species are more commonly targeted.

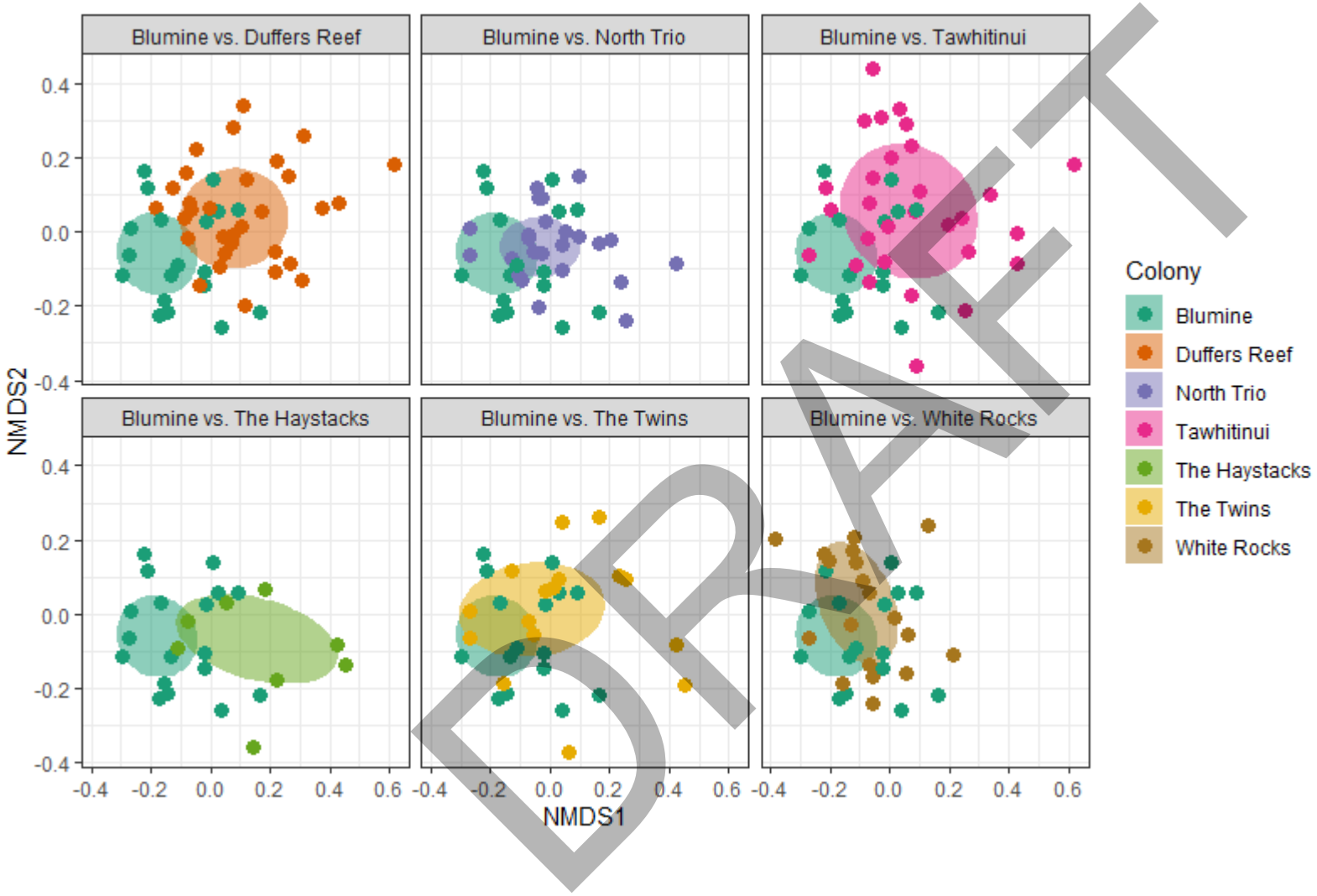
The two most dominant species in the king shag pellets were witch (present in 67% of pellets) and crested flounder (43%). Neither fish species are targeted by commercial or recreational fishers

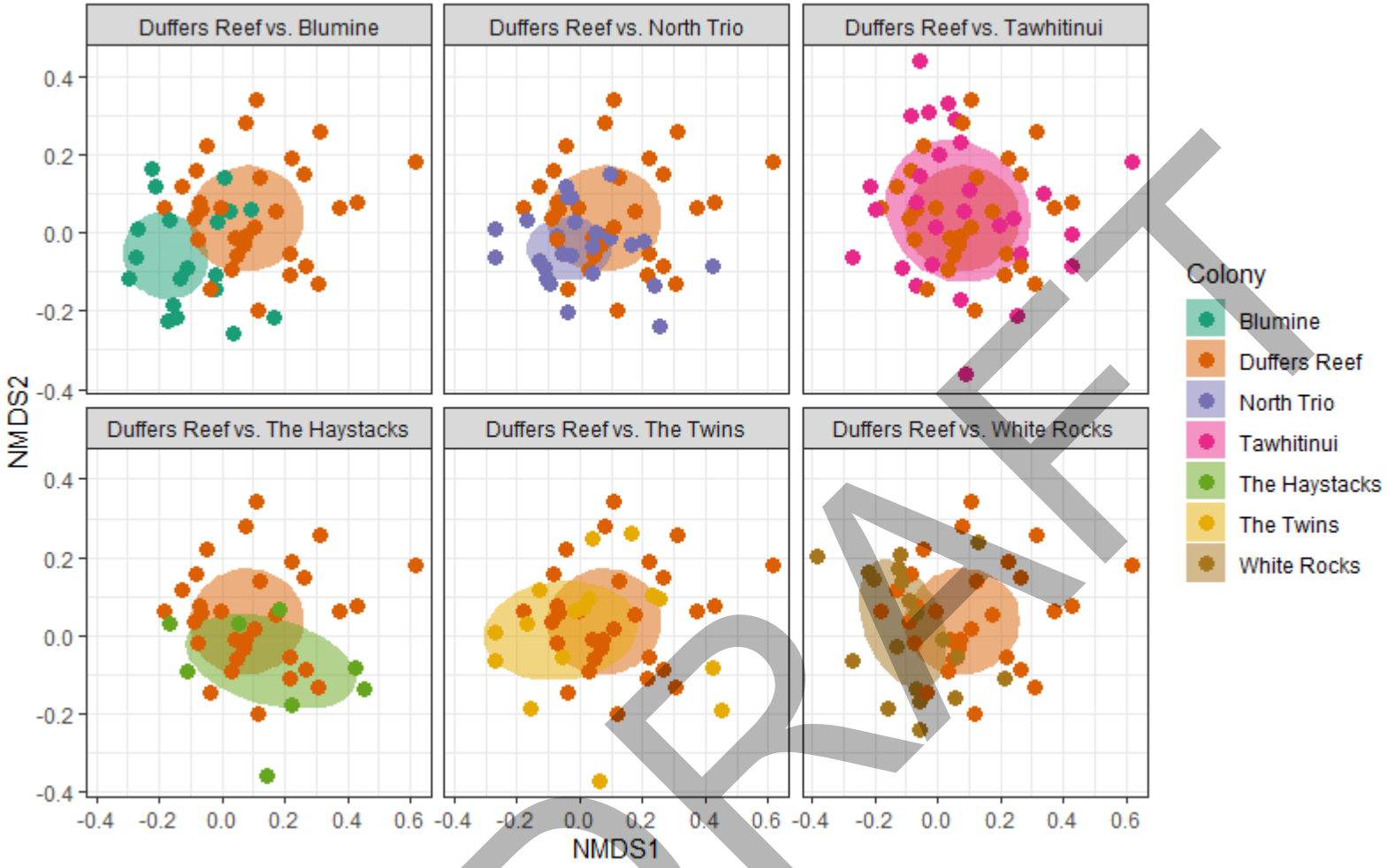
(Manikiam, 1969). In contrast, the southern lemon sole (*Pelotretis flavilatus*) is of commercial and recreational interest and was less frequent in the pellets (36%).

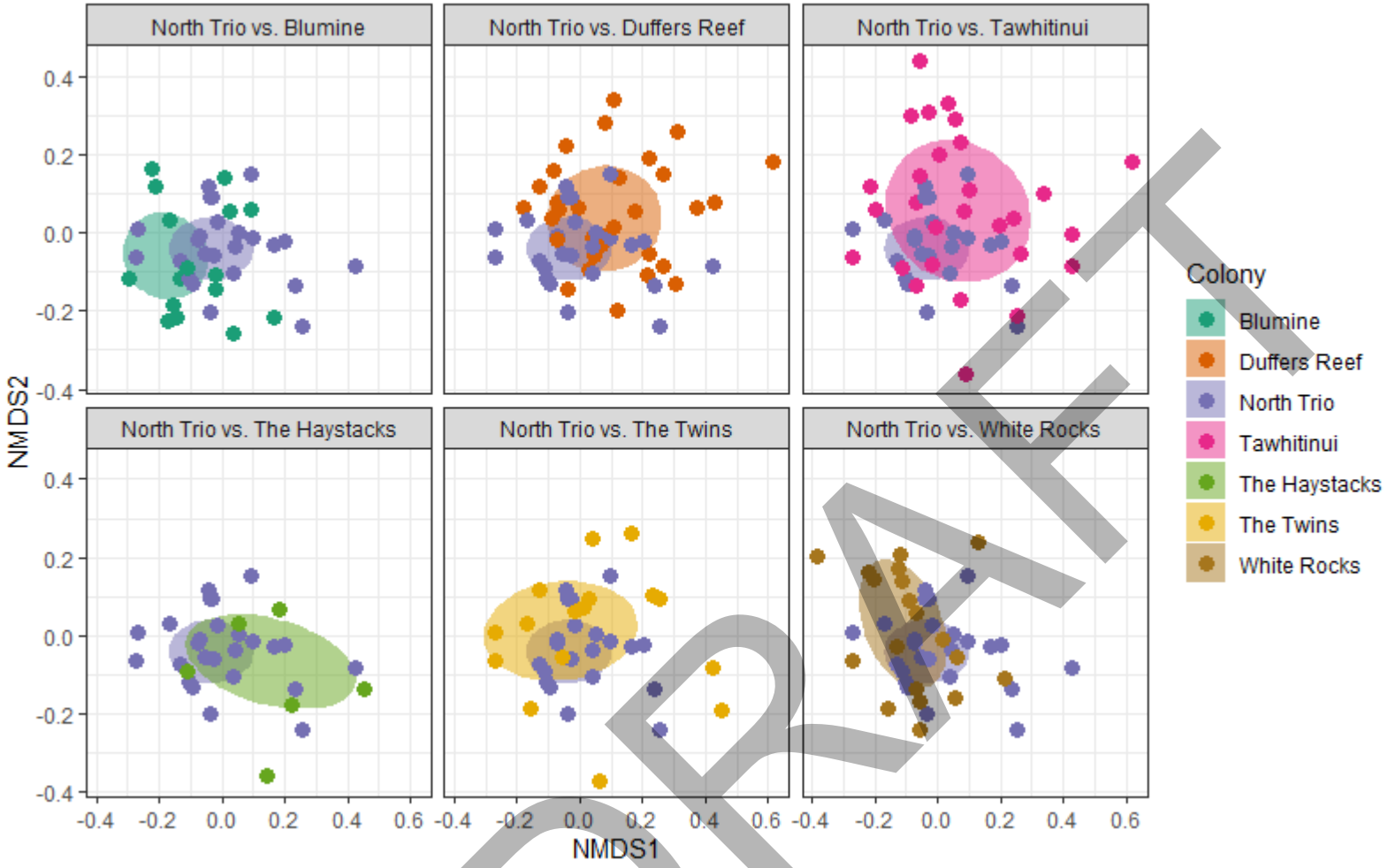
As DNA metabarcoding identifies all species present in the pellets, many non-dietary items were also detected, which can also potentially provide further insight into the king shags themselves. Microbiota have a large impact on host health and can provide greater insight to the diet, but is poorly studied in birds (Grond et al., 2018). Bacteroidetes, a phylum containing multiple bacterial species, was identified in 42% of pellets and some of these species are known to be pathogens to birds, but some may also have been transmitted through the food web via infected fish that were consumed (e.g., *Flavobacterium psychrophilum* - detected in 10 pellets; a Flavobacteriaceae species) (Thomas et al., 2011). It was noticeable that pellets from North Trio had a higher percentage of occurrence for Bacteroidetes (26 pellets; Figure 2). This may be due to widespread transmission in the colony itself or due to infected fish being more abundant at their specific foraging sites. Using DNA metabarcoding it was also possible to detect other species of interest in the pellets including feather mites (e.g., *Proctophyllodes* sp. - present in 5 pellets) and parasitic worms known to species of Phalacrocoracidae (e.g., *Contracaecum osculatum* - Garbin et al., 2011; *Andracantha sigma* - Presswell et al., 2018).

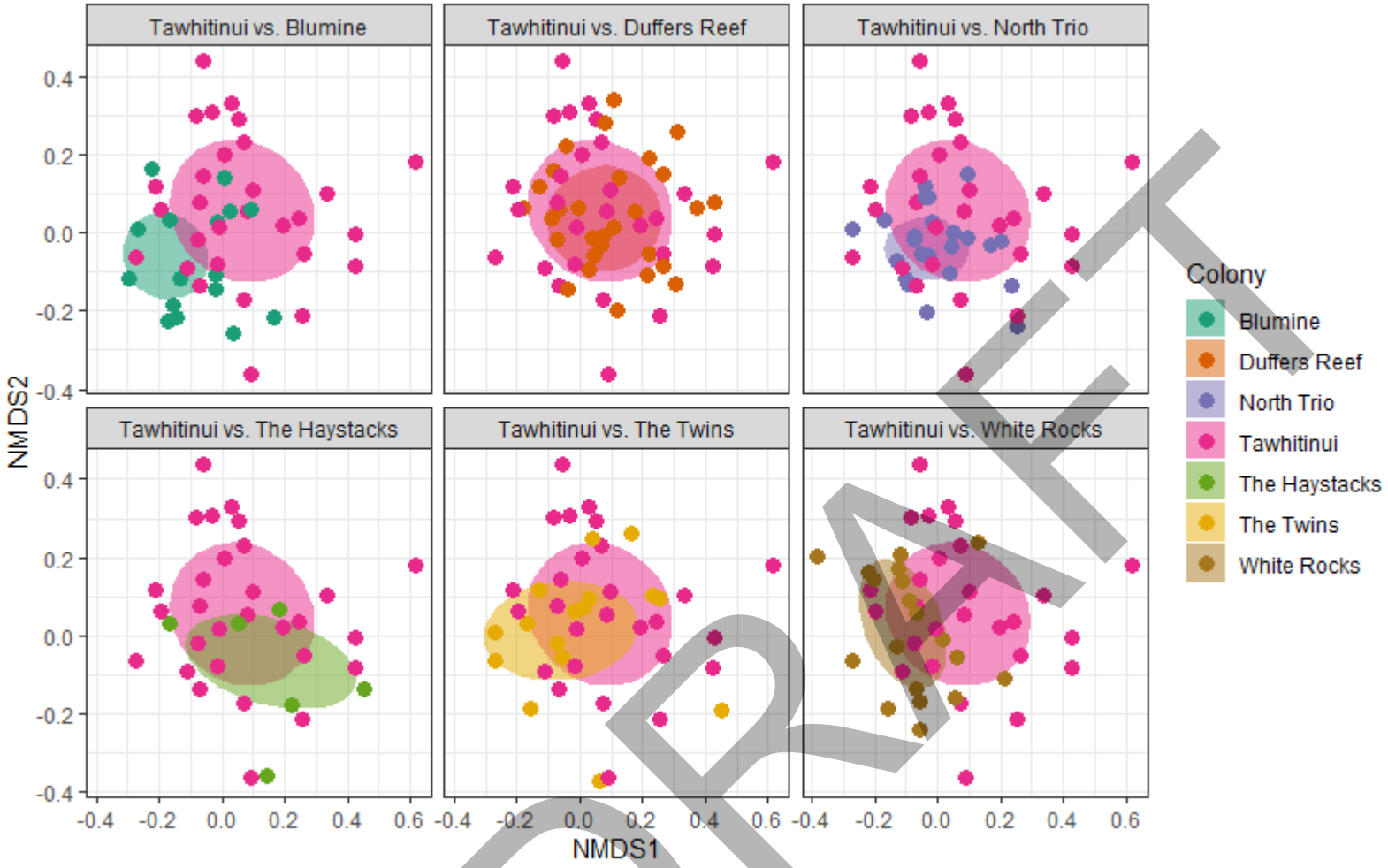
Appendix 1:

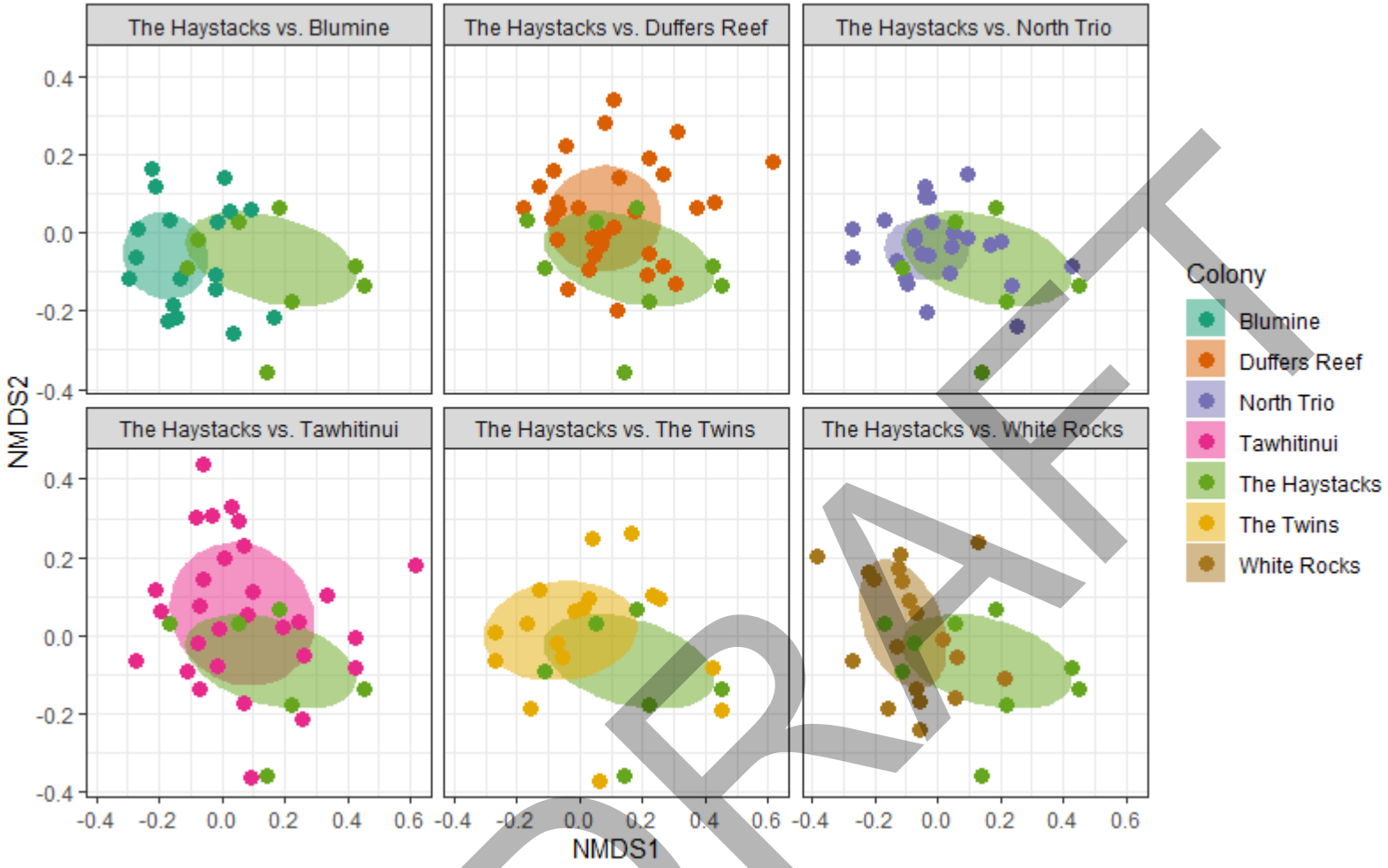
Figure A1: Nonmetric multidimensional scaling (NMDS) was used to visualize the differences in king shag regurgitated pellets at species level (fish only) among colonies and between sexes. ***A PERMANOVA revealed that the clustering was significant (p <0.05).***

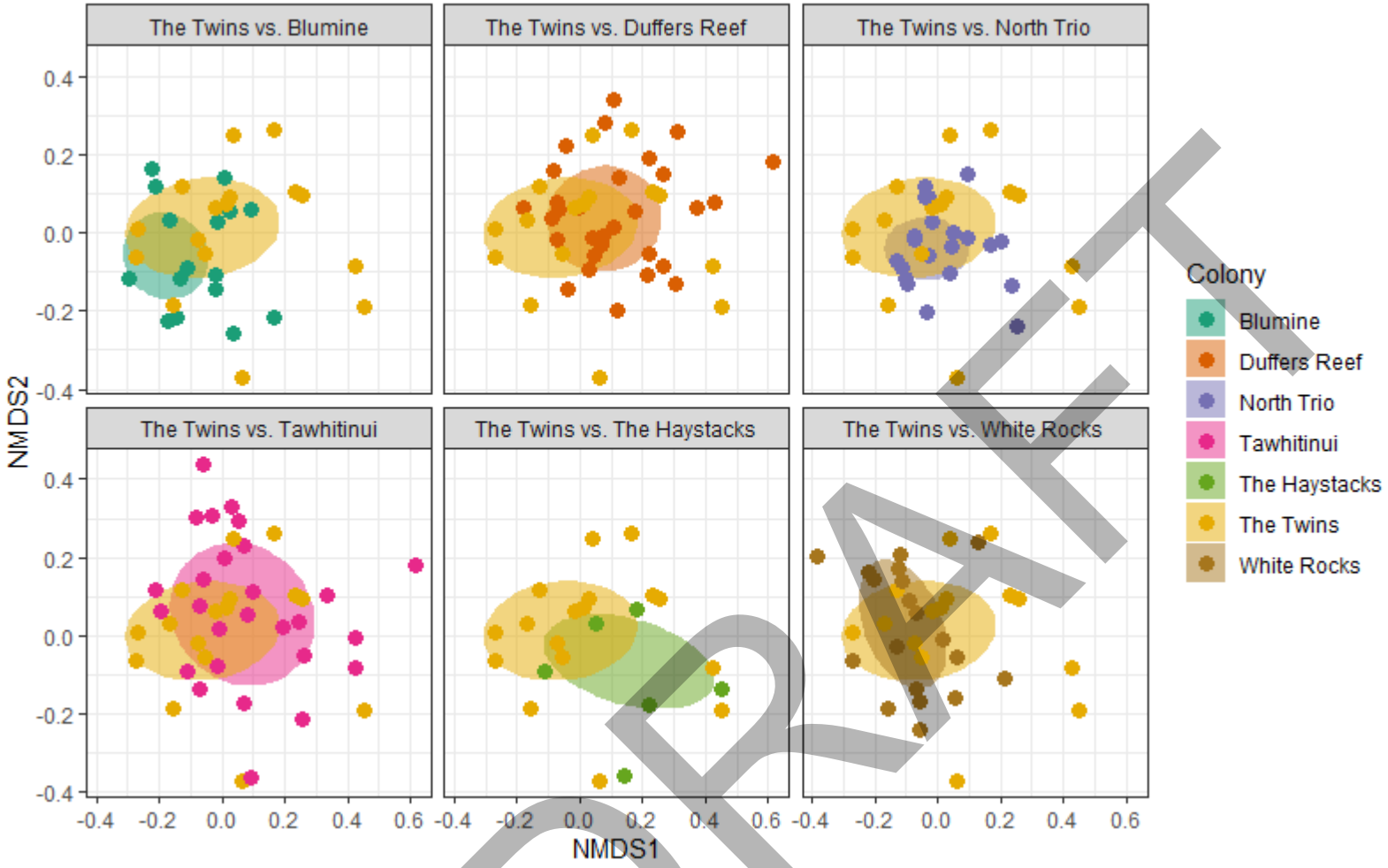


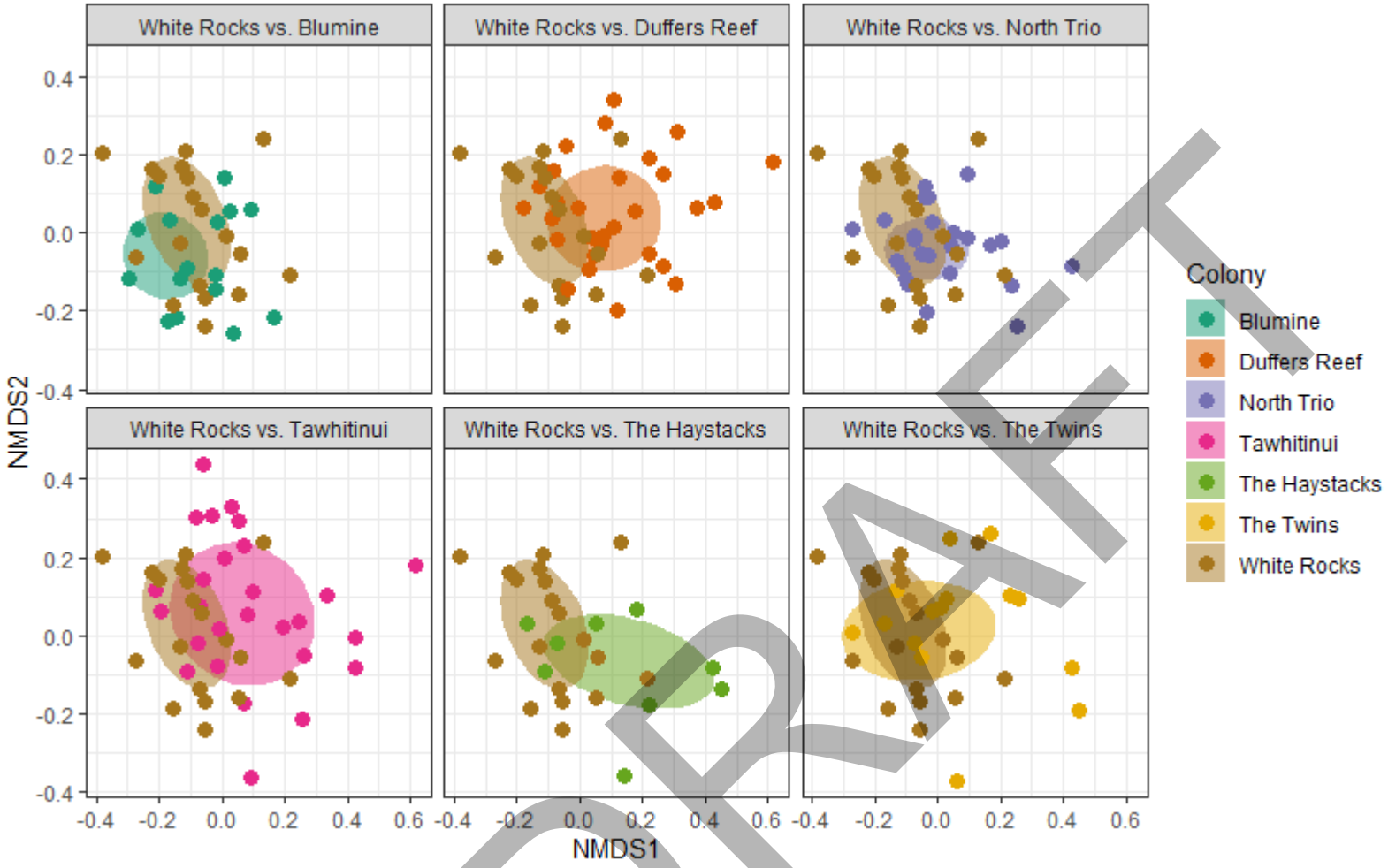

Duffers Reef vs. Blumine
Duffers Reef vs. North Trio
Duffers Reef vs. Tawhitinui
Duffers Reef vs. The Haystacks
Duffers Reef vs. The Twins
Duffers Reef vs. White Rocks
NMDS2
NMDS1
0.4
0.2
0.0
-0.2
-0.4
-0.4
-0.2
0.0
0.2
0.4
0.6
Colony
Blumine
Duffers Reef
North Trio
Tawhitinui
The Haystacks
The Twins
White Rocks

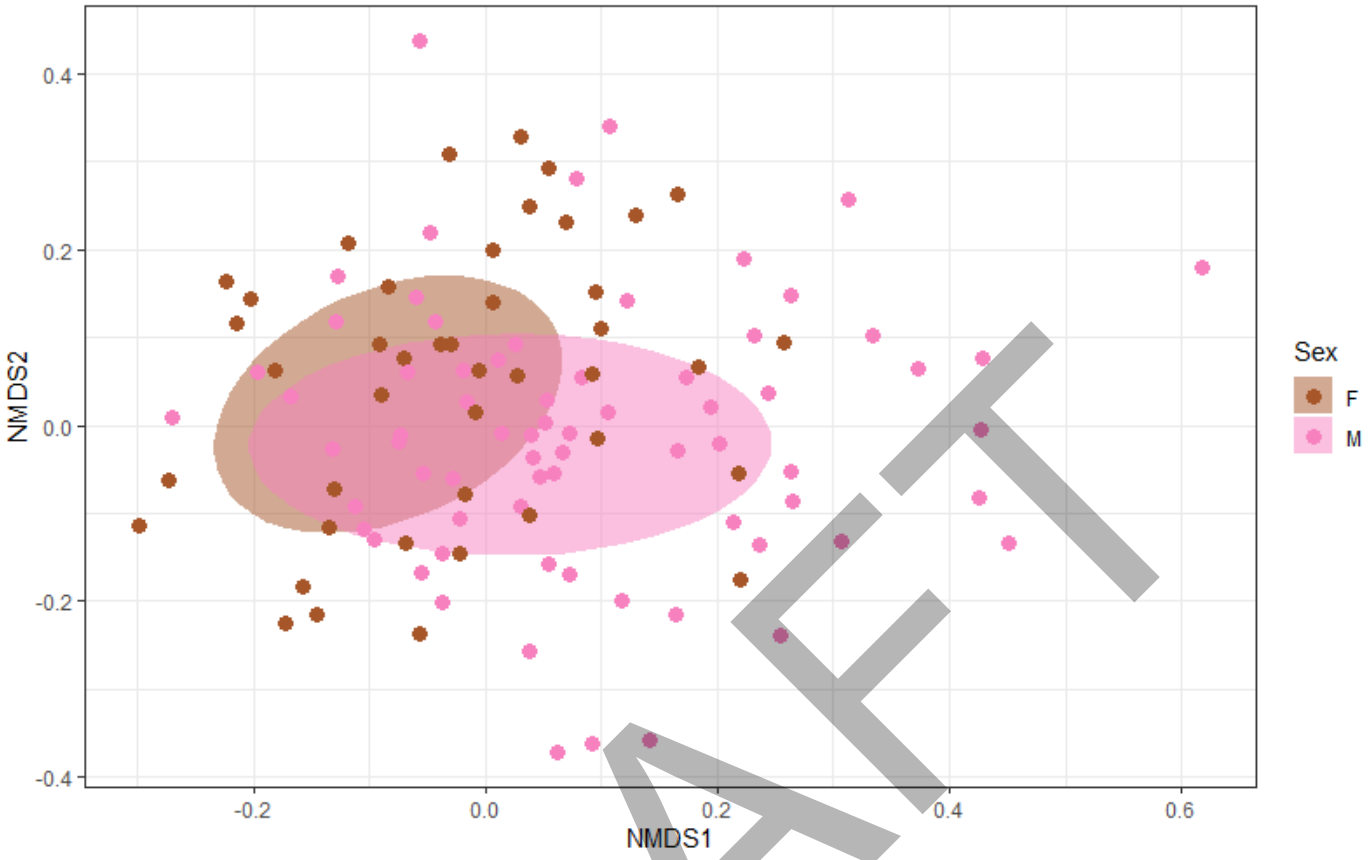

North Trio vs. Blumine
North Trio vs. Duffers Reef
North Trio vs. Tawhitinui
North Trio vs. The Haystacks
North Trio vs. The Twins
North Trio vs. White Rocks
NMDS2
NMDS1
0.4
0.2
0.0
-0.2
-0.4
-0.4
-0.2
0.0
0.2
0.4
0.6
Colony
Blumine
Duffers Reef
North Trio
Tawhitinui
The Haystacks
The Twins
White Rocks


Tawhitinui vs. Blumine
Tawhitinui vs. Duffers Reef
Tawhitinui vs. North Trio
Tawhitinui vs. The Haystacks
Tawhitinui vs. The Twins
Tawhitinui vs. White Rocks
NMDS2
NMDS1
0.4
0.2
0.0
-0.2
-0.4
-0.4
-0.2
0.0
0.2
0.4
0.6
Colony
Blumine
Duffers Reef
North Trio
Tawhitinui
The Haystacks
The Twins
White Rocks


The Haystacks vs. Blumine
The Haystacks vs. Duffers Reef
The Haystacks vs. North Trio
The Haystacks vs. Tawhitinui
The Haystacks vs. The Twins
The Haystacks vs. White Rocks
NMDS2
NMDS1
0.4
0.2
0.0
-0.2
-0.4
0.6
Colony
Blumine
Duffers Reef
North Trio
Tawhitinui
The Haystacks
The Twins
White Rocks


The Twins vs. Blumine
The Twins vs. Duffers Reef
The Twins vs. North Trio
The Twins vs. Tawhitinui
The Twins vs. The Haystacks
The Twins vs. White Rocks
NMDS2
NMDS1
0.4
0.2
0.0
-0.2
-0.4
-0.4
-0.2
0.0
0.2
0.4
0.6
Colony
Blumine
Duffers Reef
North Trio
Tawhitinui
The Haystacks
The Twins
White Rocks


White Rocks vs. Blumine
White Rocks vs. Duffers Reef
White Rocks vs. North Trio
White Rocks vs. Tawhitinui
White Rocks vs. The Haystacks
White Rocks vs. The Twins
NMDS2
NMDS1
0.4
0.2
0.0
-0.2
-0.4
-0.4
-0.2
0.0
0.2
0.4
0.6
Colony
Blumine
Duffers Reef
North Trio
Tawhitinui
The Haystacks
The Twins
White Rocks



References

Aquaculture New Zealand. (2020). *King shag numbers steady*. https://www.aquaculture.org.nz/2020/06/09/king-shags-steady/

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013). GenBank. *Nucleic Acids Research, 41*(D1), D36-D42. https://doi.org/10.1093/nar/gks1195

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., ... (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology, 37*, 852–857. https://doi.org/10.1038/s41587-019-0209-9

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The Isme Journal, 11*, 2639. https://doi.org/10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods, 13*(7), 581-583. https://doi.org/10.1038/nmeth.3869

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters, 10*(9), 20140562. https://doi.org/10.1098/rsbl.2014.0562

Devloo-Delva, F., Huerlimann, R., Chua, G., Matley, J., Heupel, M., Simpfendorfer, C., & Maes, G. (2018). How does marker choice affect your diet analysis: Comparing genetic markers and digestion levels for diet metabarcoding of tropical-reef piscivores. *Marine and Freshwater Research*. https://doi.org/10.1071/MF17209

Dowle, E. J., Pochon, X., C. Banks, J., Shearer, K., & Wood, S. A. (2016). Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: A case study using freshwater macroinvertebrates. *Molecular Ecology Resources, 16*(5), 1240-1254. https://doi.org/10.1111/1755-0998.12488

Fisher, P., & Boren, L. (2012). New Zealand king shag (*Leucocarbo carunculatus*) foraging distribution and use of mussel farms in Admiralty Bay, Marlborough Sounds. *Notornis, 59*(3&4), 105-115. http://notornis.osnz.org.nz/new-zealand-king-shag-leucocarbo-carunculatus-foraging-distribution-and-use-mussel-farms-admiralty-b

Fridolfsson, A.-K., & Ellegren, H. (1999). A Simple and Universal Method for Molecular Sexing of Non-Ratite Birds. *Journal of Avian Biology, 30*(1), 116-121. https://doi.org/10.2307/3677252

Garbin, L., Mattiucci, S., Paoletti, M., González-Acuña, D., & Nascetti, G. (2011). Genetic and morphological evidences for the existence of a new species of *Contracaecum* (Nematoda: Anisakidae) parasite of *Phalacrocorax brasilianus* (Gmelin) from Chile and its genetic relationships with congeners from fish-eating birds. *Journal of Parasitology, 97*(3), 476-492. https://doi.org/10.1645/ge-2450.1

Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources, 13*(5), 851-861. https://doi.org/10.1111/1755-0998.12138

Grond, K., Sandercock, B. K., Jumpponen, A., & Zeglin, L. H. (2018). The avian gut microbiota: community, physiology and function in wild birds. *Journal of Avian Biology, 49*(11), e01788. https://doi.org/https://doi.org/10.1111/jav.01788

Illumina. (2013). *16S metagenomic sequencing library preparation: Preparing 16S ribosomal RNA gene amplicons for the Illumina MiSeq system* (Report no. 15044223 Rev. B). Retrieved from

https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf

Kress, W. J., García-Robledo, C., Uriarte, M., & Erickson, D. L. (2015). DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology and Evolution, 30*(1), 25-35. https://doi.org/10.1016/j.tree.2014.10.008

Lalas, C., & Brown, D. (1998). The diet of New Zealand king shags (Leucocarbo carunculatus) in Pelorus Sound [Article]. *Notornis, 45*(2), 129-139. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0032414708&partnerID=40&md5=6ae822f54284de305b450cb414faae3b

Lenth, R. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version 1.3.4*. https://doi.org/https://CRAN.R-project.org/package=emmeans

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (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. *Frontiers in Zoology, 10*, 34. https://doi.org/10.1186/1742-9994-10-34

Machida, R. J., Leray, M., Ho, S.-L., & Knowlton, N. (2017). Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples [Data Descriptor]. *Scientific Data, 4*, 170027. https://doi.org/10.1038/sdata.2017.27

Manikiam, J. (1969). A Guide to The Flatfishes (Order Heterosomata) of New Zealand. *Tuatara, 17*(3). http://nzetc.victoria.ac.nz/tm/scholarly/tei-Bio17Tuat03-t1-body-d3.html

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads [next generation sequencing; small RNA; microRNA; adapter removal]. *2011, 17*(1), 3. https://doi.org/10.14806/ej.17.1.200

McMillan, P.J., Francis, M.P., James, G.D., Paul, L.J., Marriott, P., Mackay, E., Wood, B.A., Stevens, D.W., Griggs, L.H., Baird, S.J., Roberts, C.D., Stewart, A.L., Struthers, C.D., Robbins, J.E. (2019). New Zealand fishes. A field guide to common species caught by bottom, midwater, and surface fishing. New Zealand Aquatic Environment and Biodiversity Report No. 208. https://docs.niwa.co.nz/library/public/NZAEBR-208.pdf

Miskelly, C., Dowding, J., Elliott, G., Hitchmough, R., Powlesland, R., Robertson, H., Sagar, P., Scofield, P., & Taylor, G. (2008). Conservation status of New Zealand birds, 2008. *Notornis, 55*(3), 117-135.

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., & Schaffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics, 24*(16), 1757-1764. https://doi.org/10.1093/bioinformatics/btn322

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package* (Version 2.5-6). Comprehensive R Archive Network (CRAN). Retrieved from https://CRAN.R-project.org/package=vegan

Palmer, A. D. N. (2008). DNA Isolation and Amplification from Formaldehyde-Fixed Animal Tissues Rich in Mucopolysaccharides, Pigments, and Chitin. *Preparative Biochemistry & Biotechnology, 39*(1), 72-80. https://doi.org/10.1080/10826060802589635

Piñol, J., Andrés, V. S., Clare, E. L., Mir, G., & Symondson, W. O. C. (2014). A pragmatic approach to the analysis of diets of generalist predators: The use of next-generation sequencing with no blocking probes. *Molecular Ecology Resources, 14*(1), 18-26. https://doi.org/10.1111/1755-0998.12156

Presswell, B., García-Varela, M., & Smales, L. R. (2018). Morphological and molecular characterization of two new species of Andracantha (Acanthocephala: Polymorphidae) from New Zealand shags (Phalacrocoracidae) and penguins (Spheniscidae) with a key to the species. *Journal of Helminthology, 92*(6), 740-751. https://doi.org/10.1017/S0022149X17001067

R Core Team. (2018). *R: A language and environment for statistical computing* (Version 1.0.143). R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Schuckard, R., Melville, D., & Taylor, G. (2015). Population and breeding census of New Zealand king shag (Leucocarbo carunculatus) in 2015. *62*(4), 209-218.

Thomas, F., Hehemann, J.-H., Rebuffet, E., Czjzek, M., & Michel, G. (2011). Environmental and Gut Bacteroidetes: The Food Connection [Review]. *Frontiers in Microbiology, 2*(93). https://doi.org/10.3389/fmicb.2011.00093

van der Reis, A. L., Laroche, O., Jeffs, A. G., & Lavery, S. D. (2018). Preliminary analysis of New Zealand scampi (*Metanephrops challengeri*) diet using metabarcoding. *PeerJ, 6*, e5641. https://doi.org/10.7717/peerj.5641

Wickham, H. (2017). *Tidyverse: Easily install and load 'Tidyverse' packages* (Version 1.1.1). Comprehensive R Archive Network (CRAN). Retrieved from https://CRAN.R-project.org/package=tidyverse

Zaiko, A., Schimanski, K., Pochon, X., Hopkins, G. A., Goldstien, S., Floerl, O., & Wood, S. A. (2016). Metabarcoding improves detection of eukaryotes from early biofouling communities: Implications for pest monitoring and pathway management. *Biofouling, 32*(6), 671-684. https://doi.org/10.1080/08927014.2016.1186165