Deep-Sea Research I 85 (2014) 110–117

Contents lists available at ScienceDirect

Deep-Sea Research I

journal homepage: www.elsevier.com/locate/dsri

Instruments and Methods

DNA-based identification methods of prey fish from stomach contents of 12 species of eastern North Pacific groundfish

Melanie M. Paquin *, Troy W. Buckley, Richard E. Hibpshman [1], Michael F. Canino

National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA

ARTICLE INFO

Article history:
Received 18 June 2013
Received in revised form
24 November 2013
Accepted 4 December 2013
Available online 18 December 2013

Keywords:
Groundfish
Stomach contents
Prey fish
Restriction digest
Genetics
Alaska

ABSTRACT

Stomach content analysis of eastern North Pacific groundfish has been conducted routinely by researchers interested in understanding trophic interactions between key predator species and their prey. Identification of prey by traditional morphological methods has limitations however, due to the loss of identifiable characters from digestion and morphological similarities between taxa. Furthermore, some forage fish (*e.g.*, osmerids, ammodytids, and juvenile gadids), common prey of Bering Sea and Gulf of Alaska groundfish, are difficult to distinguish because of their slender or fusiform shape, disarticulating easily during digestion. DNA-based identification methods were developed to differentiate among 18 fish species, some that are found at depths greater than 200 m, from four taxonomic families: Ammodytidae and Osmeridae (forage fish), Pleuronectidae (flatfish), and Gadidae (gadid fish). Polymerase chain reaction (PCR) amplification of a 739 basepair section of mitochondrial DNA cytochrome *c* oxidase I and an 862 basepair section of mitochondrial DNA cytochrome *b* was followed by restriction digest assays and resulted in species level resolution for 16 of 18 species of interest. PCR restriction digest assays applied to fish prey from stomach contents of groundfish indicated the presence of several target species, eulachon (*Thaleichthys pacificus*), walleye pollock (*Gadus chalcogrammus*), searcher (*Bathymaster signatus*), rock sole (*Lepidopsetta bilineata*), yellowfin sole (*Limanda aspera*) and either Bering flounder (*Hippoglossoides robustus*) or flathead sole (*H. elassodon*). The PCR restriction digest protocols improved the identification rate of predated fish from stomach contents compared to identification by conventional taxonomic methods alone, and DNA sequence analysis further resolved identification of unknown prey fish samples.

Published by Elsevier Ltd.

1. Introduction

Diet studies of marine fish have traditionally been conducted using prey remains from stomach contents relying on diagnostic characters of prey items for identification. However, highly digested prey remains may preclude accurate identification to the species level. Diet studies conducted by National Marine Fisheries Service scientists at the Alaska Fisheries Science Center (AFSC) have focused on approximately 40 predatory fish species of commercial importance in Alaska fisheries, or species that prey on or compete with harvested North Pacific groundfish and shellfish (Yang et al., 2006). Accurate identification of commercially and trophically important species in stomach contents of groundfish has been the focus of these studies, however, several speciose groups of fishes have similar size and shape with distinctive characteristics mostly limited to the head, caudal and external morphology – anatomical features that are affected during early stages of digestion (Buckley et al., unpub. data). Stomach content analysis conducted at sea during the 2009 and 2010 AFSC groundfish surveys revealed that prey fish that could not be identified to at least the genus level made up 35% of the total prey fish count ($N=2142$) and 9% of the total prey fish weight (57.6 kg) (Buckley et al., in prep.). These percentages represent limitations to accurate estimation of the composition of fish prey in diets of groundfish and accurate trophic dynamics modeling.

In this study, we used a DNA-barcoding approach to increase the identification rate of sampled gut contents. The mitochondrial DNA cytochrome *c* oxidase I (COI) gene, the barcode gene region, has been used extensively for species level discrimination (Hebert et al., 2003; Hajibabaei et al., 2006; Meusnier et al., 2008; Zemlak et al., 2009). DNA-barcode sequences from the COI gene region have been compiled in public databases such as the Barcode of Life Database (BOLD) (Ratnasingham and Hebert, 2007) and GenBank (Benson et al., 2004). In one study, identification of prey fish from stomach contents of an invasive piscivorous species used comparative analyses between unknown prey sequences and reference

* Corresponding author. Tel.: +1 206 526 4794; fax: +1 206 526 6723.
E-mail address: melanie.paquin@noaa.gov (M.M. Paquin).
[1] Present address: Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Box 355672, 3737 Brooklyn Ave NE, Seattle, WA 98105, USA.

0967-0637/$ - see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.dsr.2013.12.002

sequences in BOLD to reveal a diverse number of prey fish species, with a 94% success rate matching query sequences to reference sequences (Valdez-Moreno et al., 2012). BOLD contains sequences for over 5000 fish species (Ward et al., 2009), and due to interest in discriminating between closely related fish species, universal COI fish primers have become available (Ivanova et al., 2007). Another gene, cytochrome *b* (cyt *b*), has also been used extensively for species identification and for many organisms, including fish, has shown relatively high interspecific variation (Lindstrom, 1999; Hyde and Vetter, 2007), thus both cyt *b* and COI are valuable target genes for species identification.

We attempted to develop molecular assays for the accurate identification of 18 marine fish species from four taxonomic families: Gadidae, Pleuronectidae, Osmeridae and Ammodytidae (Table 1). The goal was to easily and accurately distinguish among species within these fish groups and an additional group, Bathymasteridae, even at advanced stages of digestion. Mitochondrial DNA species type was determined using restriction fragment length polymorphisms (RFLP). We tested the polymerase chain reaction (PCR) RFLP assays on prey fish samples from stomach contents, demonstrating that the RFLP analyses facilitated resolution of prey consumption by key predatory groundfish from the eastern North Pacific. Additionally, sequence data from prey items that could not be resolved to species level using morphological characters, and were not classified as target fish groups in this study, further improved identification of fish prey from stomach contents. This approach, using RFLP analysis in combination with standard morphological and meristics methods to classify prey remains, reduced the cost of stomach contents analysis by limiting the number of samples requiring more costly DNA methods such as direct sequencing.

Table 1
Prey fish species from four fish groups found in the eastern North Pacific. In this study, PCR restriction digest protocols were developed and tested using tissues from reference specimens for species identification of gadids, flatfish, and forage fish, prey fish commonly found in the diets of eastern North Pacific groundfish. Digested prey fish samples classified to fish groups using morphology were identified using the restriction digest protocols developed in this study and a restriction digest protocol developed by Canino et al. (unpub. data) for the Bathymasterid group.

Fish group/Species	Common name
Gadidae	
Gadus macrocephalus	Pacific cod
G. chalcogrammus	Walleye pollock
Microgadus proximus	Pacific tomcod
Boreogadus saida	Arctic cod
Eleginus gracilis	Saffron cod
Pleuronectidae	
Lepidopsetta polyxystra	Northern rock sole
L. bilineata	Rock sole
Limanda proboscidea	Longhead dab
L. sakhalinensis	Sakhalin sole
L. aspera	Yellowfin sole
Hippoglossoides elassodon	Flathead sole
H. robustus	Bering flounder
Forage fish	
Ammodytes hexapterus	Pacific sand lance
Mallotus villosus	Capelin
Thaleichthys pacificus	Eulachon
Osmerus mordax	Rainbow smelt
Spirinchus starksi	Night smelt
S. thaleichthys	Longfin smelt
Bathymasteridae	
Bathymaster signatus	Searcher
B. leurolepis	Smallmouth ronquil
B. caeruleofasciatus	Alaskan ronquil

2. Materials and methods

2.1. Specimen collection

Muscle or fin tissue was obtained from adult reference specimens from the Gadidae (gadid fish), Pleuronectidae (flatfish), and Ammodytidae and Osmeridae (forage fish) groups, representing 17 of 18 prey fish species targeted for PCR-RFLP assay development (Table 1). These specimens were collected from the eastern Bering Sea, Aleutian Islands, Puget Sound, Washington, and the coast of California during routine trawl surveys, or obtained from the University of Washington Fish Collection. Reference specimens of juvenile *Lepidopsetta polyxystra* were obtained from the eastern Bering Sea and Gulf of Alaska using a beam trawl. Tissue from a minimum of five specimens for each of the 18 species was obtained and preserved either by freezing at −20 °C or in 95% ethanol.

Groundfish were captured during bottom trawl surveys in the Gulf of Alaska in 2009 and the Aleutian Islands in 2010 and fish in various stages of digestion were collected from stomachs of 12 groundfish species (Table 2). Fish displaying signs of regurgitation or net feeding were not selected for this study. Digested prey fish were acquired from between one and 37 stomachs per groundfish species, for a total of 87 stomachs. In the field, prey items from stomachs were separated and identified to the lowest possible hierarchical taxonomic level and frozen for later verification in the laboratory.

2.2. Stomach content analyses

Identification of frozen prey fish items was verified to the extent possible using diagnostic morphological characters, in the laboratory, and a subsample of 89 prey items (from the 87 stomachs, refer to Subsection 2.1) was frozen or preserved in 95% ethanol for later DNA analysis. Of these prey items 54 were classified as gadid, flatfish, forage fish, or bathymasterid. The gadid, flatfish, and forage fish samples were differentiated to species using PCR-RFLP protocols developed in this study. A PCR-RFLP protocol developed by Canino et al. (unpub. data) to distinguish among larvae of three sympatric ronquil species in the bathymasterid group (Table 1) was used to identify bathymasterid samples. The remaining 35 specimens were treated as unidentified teleosts and direct sequencing was used in order to identify them to species.

2.3. DNA analyses

DNA extractions of muscle and finclip samples were performed using a QIAGEN[2] DNeasy kit (QIAGEN Corp., Valencia, CA) using the animal tissue protocol. Extractions and PCR assays of prey items were prepared in a PCR-free laboratory to minimize the risk of contaminating potentially degraded samples with aerosolized PCR product. Additionally, in order to test the repeatability of the genetic results, genomic DNA was re-extracted from 12 prey items and restriction digest and DNA sequence results of the replicate samples were compared.

A universal primer cocktail (C_FishF1t1-C_FishR1t1) containing both forward and reverse primers was used to amplify a 739 base pair (bp) segment of COI for all species following the PCR conditions specified in Ivanova et al. (2007). For the five flatfish species *Lepidopsetta bilineata*, *Limanda sakhalinensis*, *L. aspera*, *Hippoglossoides elassodon*, and *H. robustus* an 862 bp fragment of cyt *b* was amplified for at least three individuals per species. GluDG

[2] Reference to trade names does not indicate endorsement by the National Marine Fisheries Service, NOAA.

Table 2
Prey fish found in the diets of 12 species of groundfish from the Gulf of Alaska and Aleutian Islands. Number of prey fish groups was calculated using the family taxonomic level. Prey fish identifications were determined by one of three methods: morphology, restriction digest assays (limited to gadids, flatfish, forage fish, and bathymasterids) or direct sequencing. A subsample of 89 prey items was retained for species identification using the latter two genetic methods, or for testing restriction digest assays. SACA, spiny dogfish *Squalus acanthias*; BINT, Bering skate *Bathyraja interrupta*; BPAR, Alaska skate *B. parmifera*; RBIN, big skate *Raja binoculata*; RRHI, longnose skate *R. rhina*; GMAC, Pacific cod *Gadus macrocephalus*; GCHA, walleye pollock *G. chalcogrammus*; SVAR, dusky rockfish *Sebastes variabilis*; SALU, Pacific ocean perch *S. alutus*; AFIM, sablefish *Anoplopoma fimbria*; ASTO, arrowtooth flounder *Atheresthes stomias*; HSTE, Pacific halibut *Hippoglossus stenolepis*.

	SACA	BINT	BPAR	RBIN	RRHI	GMAC	GCHA	SVAR	SALU	AFIM	ASTO	HSTE
No. of stomachs examined	2	1	1	1	2	17	9	3	1	4	37	9
No. of prey groups	2	1	1	2	2	10	5	2	1	4	5	4
Prey in stomach												
Teleosts												
Bathylagidae (Northern smoothtongue, *Leuroglossus schmidti*)							1					
Osmeridae (Capelin, *Mallotus villosus*)					2	1	8	6	1		49	3
(Eulachon, *Thaleichthys pacificus*)										2	2	1
Gadidae (Pacific cod, *Gadus macrocephalus*)											1	
(Walleye pollock, *G. chalcogrammus*)					1		3			4	7	
(Pacific tomcod, *Microgadus proximus*)						1						
Sebastidae (*Sebastes* spp.)	1											
Cottidae (Armorhead sculpin, *Gymnocanthus galeatus*)											1	
(Ribbed sculpin, *Triglops pingelii*)											1	
(Spatulate sculpin, *Icelus spatula*)											1	
(Thorny sculpin, *I. spiniger*)						1				1		
(*Gymnocanthus*, *Icelus* or *Triglops* spp.)											2	
(Yellow Irish lord, *Hemilepidotus jordani*)												1
Cottoidei (sculpins, poachers, fatheads)						1						
Agonidae (Gray starsnout, *Bathyagonus alascanus*)		1								1		
Agonidae (poachers)						1						
Bathymasteridae (Searcher, *Bathymaster signatus*)						3						
Zoarcidae (Shortfin eelpout, *Lycodes brevipes*)						1						
Stichaeidae (Daubed shanny, *Leptoclinus maculatus*)						7	6	1				
Stichaeidae (pricklebacks)							5					
Cryptacanthodidae (Dwarf wrymouth, *Cryptacanthodes aleutensis*)						2						
Ammodytidae (Pacific sand lance, *Ammodytes hexapterus*)				9		8					35	7
Pleuronectidae (Yellowfin sole, *Limanda aspera*)						1					1	
(Rock sole, *Lepidopsetta bilineata*)				1								
(Northern rock sole, *L. polyxystra*)												1
(Dover sole, *Microstomus pacificus*)						1						
(*Hippoglossoides* spp.)	1		1			1	2					
Unidentified fish prey items	1					1					2	

(Martin and Palumbi, 1993) and one of the following reverse primers, CB3 (Martin and Palumbi, 1993) or CB3R (5′-ATATCATTCTGGCTTAATGTG-3′) (Hyde, unpub.), were used to PCR-amplify a portion of the cyt *b* region and primer pairs were selected based upon the quality of the amplified band, as determined by 2% agarose gel electrophoresis and visualization using ethidium bromide stain. Separate PCR amplifications of COI and cyt *b* were conducted. All PCR reactions for cyt *b* were performed in 25 μL volume reactions containing 2 μL of DNA, 1 × PCR buffer, 1 × BSA, 2.5 mM $MgCl_2$, 0.25 mM deoxynucleotide triphosphates, 0.1 μM of each primer, and 1.25 U *Taq* polymerase (Bioline USA, Inc., Taunton, MA). For the primer pair GluDG and CB3, an initial denaturation step of 30 s at 94 °C was followed by 30 cycles of 30 s at 94 °C, 45 s at 52 °C, and 90 s at 72 °C. For primer pair GluDG and CB3R, an initial denaturation step of 2 min at 90 °C was followed by 36 cycles of 50 s at 94 °C, 2 min at 51 °C, and 1.5 min at 72 °C. A final 5-min extension at 72 °C was added to the end of each thermalcycler profile.

PCR products from fish samples were sequenced using ABI Big Dye chemistry version 3.1 (Applied Biosystems, Inc) and run on an ABI 3730 automated sequencer. M13 forward (5′-TGTAAAACGACGGCCAGT-3′) and reverse (5′-CAGGAAACAGCTATGAC-3′) primers were used to generate COI sequences following Ivanova et al. (2007). Cytochrome *b* sequences were generated using the same primers used in the PCR amplifications. KB BASECALLER quality scores were generated using ABI data collection software version 3.0, and bidirectional sequences were aligned in the computer program SEQUENCHER version 4.9 (Gene Codes Corp.). Both directions were sequenced and aligned for reference specimen samples that showed intraspecific variation in RFLP banding patterns; otherwise, sequence from one direction was analyzed for PCR-RFLP assay development. Likewise, DNA sequence from one direction was analyzed for prey samples not assigned to one of the target fish groups (*i.e.* that could not be identified using the PCR-RFLP protocols) (Table 1), or were classified as unknown teleosts. Replicate DNA extractions of 12 prey items were sequenced in order to verify consistency of sequence data between replicates and to test for contamination from other prey species. Unique COI haplotypes from reference specimens were deposited in BOLD (DNA barcodes of marine fishes from the Northeast Pacific Ocean and Bering Sea, Proj. Code WXYZ http://www.boldsystems.org/). Sequences from samples that could not be identified using the PCR-RFLP protocols, or that were classified as unknown teleosts, were compared to the database of reference sequences in GenBank using the BLAST (Basic Local Alignment Search Tool) query algorithm.

2.4. PCR-RFLP protocol development

Genetic differences among taxa at the COI and cyt *b* genes were used to design PCR-RFLP assays, the most commonly used molecular method for species identification of fish (see review by Teletchea, 2009). Restriction site maps of reference specimen sequences were evaluated in the computer program BIOEDIT (Hall, 1999) and informative site locations for enzymatic cleavage were determined by comparative analyses among species within each fish group (Table 1). The presence of at least one restriction site in the mitochondrial DNA of each species was another criterion for the selection of candidate restriction enzymes so that each species would produce an RFLP banding pattern. This

Table 3
Restriction fragment patterns from single enzyme digests. Digests were performed on both reference specimens and digested prey items. Number of samples tested, observed number of samples with variant banding patterns, number of different variants and the rate of mistaken species identity are reported. The rate of mistaken species identity due to a variant banding pattern was calculated under the assumption that any misidentification was equally likely between the different species within the group being tested with a single restriction digest.

Restriction enzyme			Group/ Species	No. of samples	No. of samples with variant banding patterns	Number of variant banding patterns	Rate of mistaken species identity (%)	Restriction fragment sizes (bp)
	Units per rxn. (U)	Recognition site						
*Bse*YI	3	C′CAGC	Gadids				*0.8*	
			G. macrocephalus	50	12	1	*0*	391, 348
			G. chalcogrammus	69	1	1	*1.4*	334, 348
			M. proximus	51	11	2	*0*	495, 244
			B. saida	50	1	1	*0*	403, 336
			E. gracilis	42	1	1	*2.4*	495, 144, 100
Serial digests		Flatfish					*0*	
*Bse*RI	4	GAGGAG(N)10′	*L. polyxystra*	51	2	1	*0*	574, 165
			L. proboscidea	5	0	0	*0*	662
*Ear*I[a]	10	CTCTTC(N)1′	*L. bilineata*	8	0	0	*0*	600, 262
			L. sakhalinensis	8	0	0	*0*	600, 178
			L. aspera	5	0	0	*0*	642, 138, 94
			Hippoglossoides spp.	6	0	0	N/A	780, 94
Serial digests		Forage fish					*1.1*	
*Ban*I	2	G′GYRCC	*A. hexapterus*	39	2	2	*2.6*	396, 238, 109
			M. villosus	51	0	0	*0*	396, 343
*Alw*I	5	GGATC(N)4′	*T. pacificus*	49	2	1	*0*	445, 294
			S. starksi	5	0	0	*0*	444, 170, 125
			S. thaleichthys	5	0	0	*0*	597, 142
*Hph*I	5	GGTGA(N)8′	*O. mordax*	47	0	0	*0*	418, 321

[a] *Ear*I digests were performed with cytochrome *b* amplified PCR products.

Table 4
Identification rates are compared for the morphology-only approach and with the addition of genetic analyses, PCR-RFLP protocols and direct sequencing. Rates were calculated as the proportion of fish prey items identified to species or genera from the total number of prey fish collected from stomachs of each groundfish species (for a list of groundfish species abbreviations see Table 2). In some cases identification rates calculated by using both morphology and genetic based methods were $<100\%$ indicating that prey items not identified to species using morphology were also not analyzed using either DNA method, and therefore were not further resolved to species.

	SACA	BINT	BPAR	RBIN	RRHI	GMAC	GCHA	SVAR	SALU	AFIM	ASTO	HSTE
Stomachs, *N* sampled	2	1	1	1	2	17	9	3	1	4	37	9
Fish prey items, *N* sampled	3	1	1	10	3	30	25	7	1	8	102	13
Samples classified to target fish group	1	0	1	10	3	13	9	5	1	4	70	9
Samples resolved to genera or species												
Morphology only	1	0	1	9	3	13	6	5	1	0	67	9
PCR/RFLP method	N/A	N/A	N/A	1	N/A	3	2	N/A	N/A	1	3	N/A
Direct sequencing method	1	1	N/A	N/A	N/A	8	4	2	N/A	3	12	4
Identification rates to genera or species												
Morphology only	0.33	0.00	1.00	0.90	1.00	0.43	0.24	0.71	1.00	0.00	0.66	0.69
Morphology, PCR/RFLP and sequencing	0.67	1.00	N/A	1.00	N/A	0.80	0.48	1.00	N/A	0.50	0.80	1.00

criterion would eliminate mistaken species identity between a species with an anticipated uncleaved PCR product (producing a single band) as the banding pattern and samples that did not cleave at the restriction site due to a mutation or failed restriction digest reaction. Our preliminary analyses of sequences from individuals from across the eastern North Pacific Ocean increased the likelihood of detecting intraspecific variation at restrictions, affecting consistency of RFLP patterns.

Restriction digests were performed in 10 or 20 μL volume reactions with 5 μL of PCR product, 1 × buffer, and 3–10 units (U) of restriction enzyme (New England Biolabs Inc.) (Table 3). Digests were incubated at 37 °C for 1 h and sample fragment patterns were scored visually with 3% agarose gel electrophoresis and ethidium bromide stain. In order to disrupt binding of the restriction enzyme to the DNA substrate after digestion and to ensure a more consistent migration rate of the DNA during electrophoresis (Weber and Osborn, 1969), sodium dodecyl sulfate (SDS) was added for a final sample concentration of 0.1% SDS. Additionally, a PCR-RFLP protocol developed by Canino et al. (unpub. data) was used to discriminate among three sympatric ronquil species (Bathymasteridae) in stomach content samples (Table 1). This digest was performed using 3 U of NEB restriction enzyme *Cac*8I in a 10 μL volume reaction with 5 μL of PCR product and 1 × buffer, with the same incubation profile and visualization techniques as above.

2.5. Prey identification rates

The proportion of prey fish identified to genus or species with and without the genetic methods was calculated for diets of 12 groundfish species in order to demonstrate the improved data

obtained using DNA analysis. The data associated with stomach content samples included predator identity (groundfish species), count and identity of all prey fish items (Table 2). The proportion of prey fish identified to genus or species by morphological methods alone was calculated for each groundfish species by taking the number of prey identified by morphology divided by the number of total prey fish sampled. The proportion of prey fish identified by all methods (morphology and two genetic methods, PCR-RFLP and direct sequencing) was calculated as the sum of samples identified using morphology and both genetic methods divided by the number of prey fish sampled from stomachs of the groundfish species of interest (Table 4).

3. Results

3.1. PCR-RFLP design

COI sequences were obtained from at least two reference specimen individuals for each target species and restriction

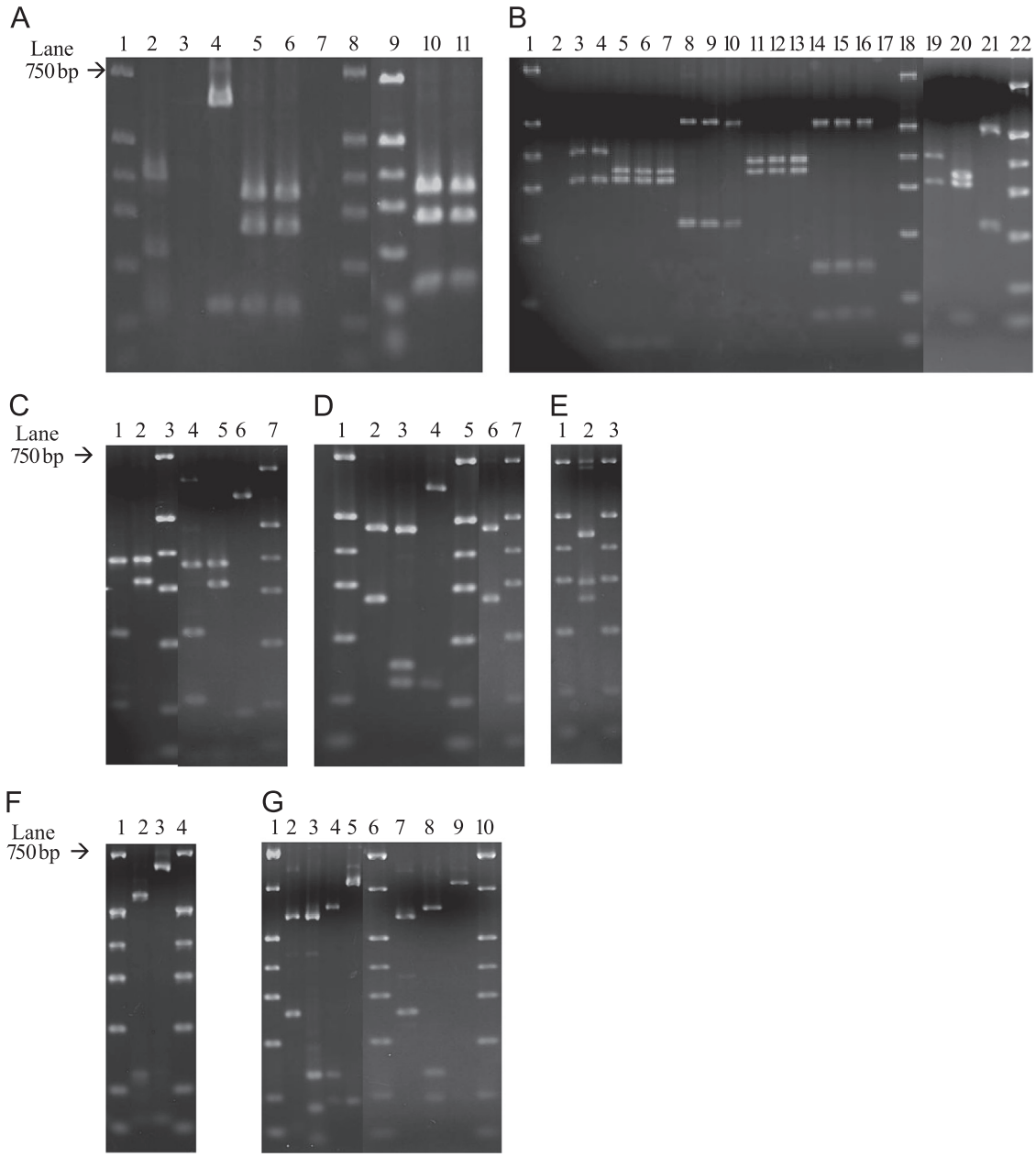

Fig. 1. Restriction fragment patterns of target species. The left half of each image represents fragment patterns from reference specimens separated by Hi-Lo DNA marker(s) (Minnesota Molecular, Inc., Minneapolis, MN) a lane size standard from fragment patterns generated from prey remains. (A) Bathymasterid *Cac*81 restriction fragment patterns. Lane 2, *Bathymaster caeruleofasciatus*; lane 3, empty; lane 4, *B. leurolepis*; lanes 5, 6, 10 and 11, *B. signatus*; lane 7, negative (−) control. (B) Gadid *Bse*YI restriction fragment patterns. Lane 2, empty; lanes 3, 4 and 19, *Gadus macrocephalus*; lanes 5 to 7, and 20, *G. chalcogrammus*; lanes 8 to 10, and 21, *M. proximus*; lanes 11 to 13, *Boreogadus saida*; lanes 14 to 16, *Eleginus gracilis*; lane 17, (−) control. (C–E) Forage fish *Ban*I, *Alw*I, and *Hph*I restriction fragment patterns, respectively. (C) Lanes 1 and 4, *Ammodytes hexapterus*; lanes 2 and 5, *Mallotus villosus*; lane 6, *Thaleichthys pacificus*. (D) Lanes 2 and 6, *T. pacificus*; lane 3, *Spirinchus starksi*; lane 4, *Spirinchus thaleichthys*. (E) Lane 2, *Osmerus mordax*. (F and G) Pleuronectid *Bse*RI and *Ear*I restriction fragment patterns, respectively. (F) Lane 2, *Lepidopsetta polyxystra*; lane 3, *Limanda proboscidea*. (G) Lanes 2 and 7, *Lepidopsetta bilineata*; lane 3, *Limanda sakhalinensis*; lanes 4 and 8, *Limanda aspera*; lanes 5 and 9, *Hippoglossoides robustus*. In some cases, only the results from reference specimens are shown. Arrows indicate 750 bp bands.

enzymes that produced unique banding patterns for each species within a fish group were selected. Species level discrimination was achieved with restriction fragment banding patterns for most of the samples from five gadid species using restriction enzyme *Bse*YI (Table 3). Banding patterns varied within species with higher than expected frequencies for samples of *Gadus macrocephalus* and *Microgadus proximus* and subsequent DNA sequence analysis of these samples revealed nucleotide substitutions at *Bse*YI restriction sites. Banding pattern variants occurred in 24% of *G. macrocephalus* samples and 22% of *M. proximus* samples (Table 3). In cases where a variant banding pattern was indistinguishable between two species, the rate of mistaken species identity was calculated as the frequency of occurrence of the variant banding pattern in the total sample of individuals tested for that species. An overall rate of mistaken identity was also reported and calculated using the frequency of misidentified samples within a target fish group. In contrast, relatively low frequencies of intraspecific banding pattern variants were detected for other species with similar sample sizes, such as forage fish species *Ammodytes hexapterus*, *Mallotus villosus*, *Thaleichthys pacificus*, and *Osmerus mordax*. Distinguishing among six species of forage fish required the sequential use of three separate restriction digests (serial digests) performed with restriction enzymes *Ban*I, *Alw*I, and *Hph*I (Table 3).

Restriction site variation at both the COI and cyt *b* gene regions was insufficient to discriminate among all seven flatfish species using a single restriction enzyme for either gene region. Restriction digest banding patterns of *L. polyxystra* and *L. proboscidea* distinguished these two flatfish species when samples of COI-amplified PCR product were digested with restriction enzyme *Bse*RI (Table 3). Identification of three of the remaining five flatfish species was accomplished using cyt *b*-amplified PCR product with restriction enzyme *Ear*I (Table 3). Identification of the two *Hippoglossoides* species (sister taxa) could only be resolved to the genus level using this same digest reaction.

3.2. Prey samples identification

RFLP protocols were used to identify 54 prey fish collected from stomach contents that had been classified using morphological methods as forage fish, flatfish, gadids, or bathymasterids. Included were 12 samples that had been DNA-extracted twice. Results showed consistent banding patterns between the duplicate samples and sequence data corroborated identifications made using RFLP analyses. The 54 samples analyzed with PCR-RFLP protocols identified one ronquil species *Bathymaster signatus* (Fig. 1A), three gadid species *G. macrocephalus*, *G. chalcogrammus*, and *M. proximus* (Fig. 1B), three forage fish species *A. hexapterus* and *M. villosus* (Fig. 1C), and *T. pacificus* (Fig. 1D), two flatfish species *L. bilineata* and *L. aspera* (Fig. 1G) and the flatfish genus *Hippoglossoides* (*H. elassodon* or *H. robustus*) (Fig. 1G). RFLP protocols differentiated three species of forage fish (*A. hexapterus*, *M. villosus*, and *T. pacificus*) from prey item samples classified morphologically as forage fish (Fig. 2) even though fragment banding patterns of two *M. villosus* samples contained additional faint bands, potentially indicating DNA contamination by other prey species. The dominant bands accurately identified the species.

3.3. Identification rates

Ten of twelve samples that had been morphologically classified as unidentified gadid, flatfish, forage fish, or bathymasterid (the same twelve samples that were DNA extracted twice) were

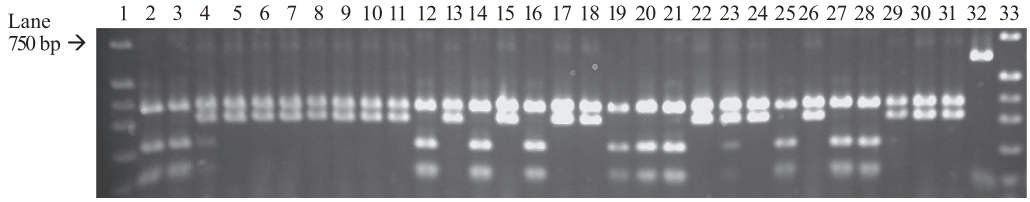

Fig. 2. *Ban*I digest results of stomach contents ($n=31$). The samples were identified as *Ammodytes hexapterus* (lanes 2, 3, 12, 14, 16, 19 to 21, 25, 27 and 28), *Mallotus villosus* (lanes 4 to 11, 13, 15, 17, 18, 22 to 24, 26, and 29 to 31), and *Thaleichthys pacificus* (lane 32) using restriction fragment patterns. Lanes 1 and 33, Hi-Lo DNA marker (Minnesota Molecular, Inc., Minneapolis, MN). The arrow indicates 750 bp bands.

Table 5
Sequence similarity scores (% maximum identity) for fish prey items (*N*). Prey species body-type descriptors are relative body types compared to other species identified from the gut content sample. Predator species sampled for stomach contents is indicated.

Prey fish	Family	Body type	*N* prey fish	% Identity	Predator
Ammodytes hexapterus	Ammodytidae	Slender-bodied	2	99	*Hippoglossus stenolepis*, *Gadus macrocephalus*
Bathyagonus alascanus	Agonidae	Moderately-elongate	2	100	*Anoplopoma fimbria*, *Bathyraja interrupta*
Cryptacanthodes aleutensis	Cryptacanthodidae	Moderately-elongate	2	100	*Gadus macrocephalus*
Gadus chalcogrammus	Gadidae	Fusiform	2	100	*Atheresthes stomias*
Hemilepidotus jordani	Cottidae	Subcircular	1	100	*Hippoglossus stenolepis*
Icelus spatula	Cottidae	Subcircular	1	99	*Atheresthes stomias*
Icelus spiniger	Cottidae	Subcircular	2	100	*Anoplopoma fimbria*, *Gadus macrocephalus*
Lepidopsetta polyxystra	Pleuronectidae	Ovate, flattened	1	99	*Hippoglossus stenolepis*
Leptoclinus maculatus	Stichaeidae	Moderately-elongate	4	99	*Sebastes variabilis*, *Gadus chalcogrammus*, *Gadus macrocephalus*
Leuroglossus schmidti	Bathylagidae	Moderately-elongate	1	99	*Gadus chalcogrammus*
Mallotus villosus	Osmeridae	slender-bodied	11	99	*Gadus macrocephalus*, *Atheresthes stomias*, *Gadus chalcogrammus*, *Sebastes variabilis*
Microstomus pacificus	Pleuronectidae	Ovate, flattened	1	99	*Gadus macrocephalus*
Sebastes spp.[a]	Sebastidae	Fusiform	1	99	*Squalus acanthias*
Thaleichthys pacificus	Osmeridae	Slender-bodied	2	99	*Anoplopoma fimbria*, *Hippoglossus stenolepis*
Triglops pingelii	Cottidae	Moderately-elongate	3	93, 98, 99	*Atheresthes stomias*

[a] This sample aligned with reference sequences from the *Sebastes* rockfish group *S. zacentrus*, *emphaeus*, *variagatus*, and *wilsoni* in GenBank.

identified to species using the PCR-RFLP protocols (Table 4). Two unknown flatfish samples from these 12 samples were determined to be non-target fish species based on their RFLP banding patterns and subsequently identified using COI sequences. Prey fish ($N=37$) morphologically classified as unidentified teleosts (including the two unknown flatfish samples) from 34 stomachs taken from 8 groundfish species (*Bathyraja parmifera*, *Raja binoculata*, *R. rhina*, and *Sebastes alutus* were excluded) were identified by comparison of COI nucleotide sequence data to the database of sequences in GenBank and BOLD (Table 5). The identification of 10 prey items analyzed using RFLPs and an additional 36 samples identified using direct sequencing illustrates the improved identification accuracy of stomach contents compared to the traditional method (Table 4).

4. Discussion

The RFLP protocols developed in this study allowed for accurate identification of 16 marine fish species, the majority of which are commonly encountered in stomach content surveys of eastern North Pacific groundfish. Because of the relative certainty with which these common prey species are assigned to taxonomic groups in the laboratory, the PCR-RFLP analyses increased the accuracy of species identification of digested fish by overcoming the limitations of traditional morphology-based methods (Table 4). Except for the additional PCR amplification of the cyt *b* gene region for a subset of the flatfish species, the protocols can be performed relatively quickly. Although the cost savings were relatively small in this study - analyzing 10 of 47 samples with RFLPs instead of with direct sequencing - the savings in laboratory costs could be substantial for a larger sample size of prey fish items preserved for DNA analysis. The cost per sample for restriction digests, if using for example an average of 3 U of restriction enzyme per sample, is approximately $0.20. This works out to approximately five reactions per $1 of restriction enzyme based on the cost of a typical enzyme ($65 for 1000 U). In comparison, the cost to sequence samples (unpurified PCR product) at the University of Washington's High-Throughput Genomics Unit in Seattle, Washington, is $2.73 per sample (based on university affiliation pricing). Additional costs per sample for DNA extraction and PCR are the same for both DNA methods of species identification and are therefore excluded from the calculation of cost. Also excluded from the calculation is the cost of preparation time required for running and scoring the visualization gel since preparation time required for sequencing analysis (quantification and dilution of samples) is similar. Recent developments in DNA analysis methods of species identification such as mini-barcodes (Hajibabaei et al., 2006) and express barcodes (Ivanova et al., 2009) still require sequencing analysis whereas the cost of supplies needed to run an agarose gel for visualization and scoring of restriction fragment patterns is negligible. The only additional costs incurred from the RFLP method that would need to be considered are initial costs associated with PCR-RFLP assay development. The success with the RFLP analyses to identify fish species in stomach contents collected from 12 predatory groundfish species, over half of which are key predator species (Lang and Livingston, 1996; Yang et al., 2006), suggests that these techniques could be expanded for analyzing gut contents from a greater diversity of groundfish and their prey. Although this study focused on fish species in gut contents, these methods could also be expanded to include invertebrate prey taxa.

In addition, the PCR-RFLP methods developed in this study allow for species identification of juvenile *L. polyxystra* and *L. bilineata*, which can have ambiguous morphometric characteristics. Studies that attempt to determine the ontogenetic variability in geographic distributions of fishes are dependent upon the correct species identification of juveniles (Matarese et al., 2003). However, with the limited data presented here for *L. bilineata*, additional samples would help verify restriction fragment sizes for this species. As a cautionary note, when confirming species-specific fragment patterns from PCR-RFLP analyses, restriction fragment pattern variants were typically detected in species with a sample size >40 individuals (Table 3). Most minor intraspecific RFLP haplotypes, usually resulting from a gain or loss of a restriction site, occurred in relatively low frequencies, although two gadid species, *G. macrocephalus* and *M. proximus*, exhibited substantial frequencies of alternative haplotypes (Table 3). In these species, alternative haplotypes usually produced unique RFLP banding patterns and thus did not result in misidentifications. For samples analyzed using restriction digests that resulted in intraspecific restriction fragment pattern variation, or that failed to cleave, sequence data was analyzed for nucleotide substitutions at restriction sites.

Direct sequencing of an additional 37 prey item samples classified as unknown teleosts, or that could not be classified to genus after visual examination of the specimens, further improved the identification accuracy rate of gut contents. This study focused on analyzing prey fish remains and empty stomachs containing chime were not considered. Sequence similarity scores of prey item samples ranged from 93–100%, and only one sample failed (89% species similarity) due to poor quality sequence (Table 5). Slender-bodied and moderately-elongate species made up the majority (75%) of the sampled prey. In a few cases, unknown prey item samples could only be identified to genus or had a relatively low sequence similarity score indicating either a limited representation of sequences in the database for a particular species from the species entire geographic range (*e.g.* cottids), or the close genetic relationship between species within a taxonomic group (*e.g.* cottids or sebastids). Species with very similar COI nucleotide sequence data (sequence similarity scores $\geq 98\%$), such as rockfish, would likely require sequence data from multiple gene regions (see Hyde and Vetter, 2007) in order to discriminate among the approximately 37 species inhabiting Alaskan waters (Butler et al., 2012).

Although the number of groundfish stomachs collected varied greatly between species, the composition of diets of the 12 species of eastern North Pacific groundfish indicates the relatively high occurrence of the osmerids, ammodytids, and gadids in stomach contents. Evident even with a limited sample of stomach contents is the diversity of taxonomic fish groups found in the diets of *G. macrocephalus*, *G. chalcogrammus*, *Atheresthes stomias*, and to a lesser extent *Anoplopoma fimbria* and *Hippoglossus stenolepis* (Table 2). DNA analysis improved prey fish identification in nine of the groundfish species and the number of species revealed from the prey remains varied greatly (Table 5). For *A. stomias* 47% of the prey remains identified by DNA analysis were *Mallotus villosus*, one of the fish species targeted for RFLP analysis development. In contrast, DNA analysis of prey remains from *G. macrocephalus* stomachs revealed the presence of nine species. Of particular interest was the presence of commercially important groundfish species, *G. chalcogrammus* and *G. macrocephalus*, in the prey fish samples identified by DNA analysis. Three of the 12 groundfish species, *G. chalcogrammus*, *A. fimbria*, and *A. stomias*, contained at least one sample of *G. chalcogrammus*, raising the question: how much of the diet that remains as unidentified teleost or gadid is composed of this species?

In conclusion the use of both genetic methods, PCR-RFLP protocols developed to identify abundant prey fish species and direct sequencing of unidentifiable prey fish remains, could facilitate more cost-effective species identification in gut content studies.

Acknowledgments

We thank the following agencies and people for the collection of specimens: Alaska Fisheries Science Center's RACE Groundfish

and Shellfish Assessment Programs and crews of the fishing vessels, the Northwest Fisheries Science Center biologists on their West Coast Groundfish Slope-Shelf Bottom Trawl Survey, and WA Palsson of the Alaska Fisheries Science Center. We are grateful to KP Maslenikov and TW Pietsch of the University of Washington Fish Collection for curation and collections support and to the technicians at the University of Washington's High-Throughput Genomics Unit for generating sequence data. We also thank SE Syrjala for helpful discussions concerning data analysis. We thank AC Matarese, IB Spies, and three anonymous reviewers for their reviews. A North Pacific Research Board Grant #924 to MFC and TWB provided funding. The findings and conclusions in the paper are those of the authors and do not necessarily represent the views of the National Marine Fisheries Service.

References

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Wheeler, D.L., 2004. GenBank: update. Nucleic Acids Res. 32, D23–D26. (Database issue).

Butler, J.L., Love, M.S., Laidig, T.E., 2012. A Guide to the Rockfishes, Thornyheads, and Scorpionfishes of the Northeast Pacific. University of California Press, California.

Hajibabaei, M., Smith, M.A., Janzen, D.H., Rodriguez, J.J., Whitfield, J.B., Hebert, P.D.N., 2006. A minimalist barcode can identify a specimen whose DNA is degraded. Mol. Ecol. Notes 6, 959–964.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Hebert, P.D., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B Biol. Sci. 270, 313–321.

Hyde, J.R., Vetter, R.D., 2007. The origin, evolution, and diversification of rockfishes of the genus *Sebastes* (Cuvier). Mol. Phylogenet. Evol. 44, 790–811.

Ivanova, N.V., Borisenko, A.V., Hebert, P.D.N., 2009. Express barcodes: racing from specimen to identification. Mol. Ecol. Resour. 9, 35–41.

Ivanova, N.V., Zemlak, T.S., Hanner, R.H., Hebert, P.D.N., 2007. Universal primer cocktails for fish DNA barcoding. Mol. Ecol. Notes 7, 544–548.

Lang, M., Livingston, P.A., 1996. Food habits of key groundfish species in the eastern Bering Sea slope region. US Department of Commerce, National Oceanic and Atmospheric Administration Tech. Memo. NMFS-AFSC-67, pp. 111.

Lindstrom, D.P., 1999. Molecular species identification of newly hatched Hawaiian amphidromous gobioid larvae. Mar. Biotechnol. 1, 167–174.

Martin, A.P., Palumbi, S.R., 1993. Protein evolution in different cellular environments: cytochrome b in sharks and mammals. Mol. Biol. Evol. 10, 873–891.

Matarese, A.C., Blood, D.M., Picquelle, S.J., Benson, J.L., 2003. Atlas of abundance and distribution patterns of ichthyoplankton from the Northeast Pacific Ocean and Bering Sea ecosystems: based on research conducted by the Alaska Fisheries Science Center (1972–1996). NOAA Prof. paper NMFS, vol. 1, pp. 281.

Meusnier, I., Singer, G.A.C., Landry, J.F., Hickey, D.A., Hebert, P.D.N., Hajibabaei, M., 2008. A universal DNA mini-barcode for biodiversity analysis. BMC Genomics 9.

Ratnasingham, S., Hebert, P.D.N., 2007. BOLD: The Barcode of Life Data System. Mol. Ecol. Notes 7, 355–364 ⟨http://www.barcodinglife.org⟩.

Teletchea, F., 2009. Molecular identification methods of fish species: reassessment and possible applications. Rev. Fish Biol. Fish. 19, 265–293.

Valdez-Moreno, M., Quintal-Lizama, C., Gomez-Lozano, R., Garcia-Rivas, M.D.C., 2012. Monitoring an alien invasion: DNA barcoding and the identification of lionfish and their prey on coral reefs of the Mexican Caribbean. PLoS ONE 7, e36636.

Ward, R.D., Hanner, R., Hebert, P.D., 2009. The campaign to DNA barcode all fishes, FISH-BOL. J. Fish Biol. 74, 329–356.

Weber, K., Osborn, M., 1969. The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. J. Biol. Chem. 244, 4406–4412.

Yang, M.-S., Dodd, K., Hibpshman, R., Whitehouse, A., 2006. Food habits of groundfishes in the Gulf of Alaska in 1999 and 2001. US Department of Commerce, National Oceanic and Atmospheric Administration Tech. Memo., 199. (NMFS-AFSC-164).

Zemlak, T.S., Ward, R.D., Connell, A.D., Holmes, B.H., Hebert, P.D., 2009. DNA barcoding reveals overlooked marine fishes. Mol. Ecol. Resources, vol. 9, pp. 237–242.